# Comparison of Mono- and Bifunctional Forms of *Haemophilus influenzae* TyrA Protein

Hyun Young Kim

A Thesis

In the Department

of

Chemistry and Biochemistry

Presented in Partial Fullfilment of the Requirements

For the Degree of Master of Science at

Concordia University

Montréal, Quebec, Canada

April 2013

**CONCORDIA UNIVERSITY**

**School of Graduate Studies**

This is to certify that the thesis prepared

By: Hyun Young Kim

Entitled: Comparison of Mono- and Bifunctional Forms of *Haemophilus influenzae* TyrA Protein

And submitted in partial fulfillment of the requirements for the degree of

**Master of Science (Chemistry)**

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Dr. Ann English Chair

Dr. Paul Joyce Examiner

Dr. Peter Pawelek Examiner

Dr. Joanne Turnbull Supervisor

Approved by _____

Chair of Department or Graduate Program Director

_____ _____

Date Dean of Faculty

# Abstract

**Comparison of Mono- and Bifunctional Forms of a *Haemophilus influenzae* TyrA Protein**

**Hyun Young Kim**

The TyrA protein family of dehydrogenases includes prephenate dehydrogenase (PD) which is dedicated to the synthesis of L-tyrosine (L-Tyr). PD is a monofunctional enzyme in many bacteria, but in *Haemophilus influenzae*, PD is bifunctional and is fused to chorismate mutase (CM). The present work is focused on understanding the importance of the CM domain in PD structure and stability in the presence of guanidine hydrochloride (Gdn-HCl) and variable temperatures (VT). Towards this goal, both full-length (FL) protein and an engineered monofunctional variant (Δ80 CM-PD) form of TyrA from *H. influenzae* were overexpressed and purified using nickel affinity chromatography. Measurements by far-UV circular dichroism (CD) and steady-state fluorescence spectroscopies suggested that both proteins were completely unfolded by 4.2 M Gdn-HCl, although in the presence of 1.5 M to 2.4 M Gdn-HCl, the FL protein was more prone than Δ80 CM-PD to time-dependent protein precipitation. Thermal denaturation studies using far-UV CD spectroscopy and intrinsic fluorescence measurements revealed similar melting temperature of ~52$^{\circ}$C for both proteins. VT-FTIR experiments showed that the temperature-mediated unfolding pathways of the two enzymes were very similar but not identical. Additionally the active site ligands, $NAD^+$ and L-Tyr, stabilized intramolecular β-sheets of both proteins by ~5$^{\circ}$C. VT experiments also support the idea that *H. influenzae* TyrA proteins retain their dimeric nature during unfolding and highlight the importance of the dimer interface in TyrA structure and function.

**Acknowledgments**

First and foremost, I thank the almighty God for blessing me to complete the graduate studies.

I wish to express my sincere and deepest gratitude to my supervisor, Dr. Joanne Turnbull for her constant patience as a teacher. Throughout the research, she provided encouragement, sound advice, good teaching, and good ideas. I would have been lost without her.

I would like to thank my committee members, Dr. Paul Joyce and Dr. Peter Pawelek as well as Dr. Ashraf Ismail who patiently guided me through the research. Your kind assistance with instruments, wise advice, and ideas on various applications helped me greatly.

I am indebted to many colleagues and faculty members in chemistry and biochemistry department for providing a stimulating and fun environment in which to learn and grow. I am especially grateful to Natalie Khor, Nare Ko and the lab members, Peter Quashie, Nathalie Rachael, Natascha Hotz and Rina Shlaifer for their friendship, and advice.

I would like to thank my family for providing a loving environment for me. My mom, Min Ju Lee, and my brother, Hans Young Kim, you are the best supporters. Lastly and most importantly, I would like to give my special thanks to my husband, Chris, whose consistent support and love enables me to complete this work.

## TABLE OF CONTENTS

## LIST OF FIGURES

## LIST OF TABLES

## LIST OF ABBREVIATIONS

| | |
|---|---|
| 3CB | three component buffer |
| AD | arogenate dehydrogenase |
| ADT | arogenate dehydratase |
| Ala (A) | alanine |
| Amp | ampicillin |
| Arg (R) | Arginine |
| Asp (N) | Asparagine |
| AT | aminotransterase |
| BSA | bovine serum albumin |
| CD | circular dichrosim |
| CM | chorismate mutase |
| $D_2O$ | deuterium oxide |
| Da | Dalton |
| DAHP | deoxy-D-arabino-heptulosonate 7-phosphate |
| DNA | deoxyribonucleic acid |
| dNTPs | deoxynucleotide triphosphates |
| DTT | dithiothreitol |
| EDTA | ethylenediamine tetra-acetic acid |
| ESI-Tof | electron spray ionization Quadrupole Time-of-Flight |
| FL CM-PD | *H. influenzae* full-length CM-PD |
| FSD | Fourier self-deconvolution |
| FTIR | Fourier transform infrared |
| Gdn-HCl | guanidine hydrochloride |
| Glu (Q) | glutamic acid |
| Gly (G) | Glycine |
| Gβ1-TyrA | β1-domain of protein G fused to N-terminus of PD domain |
| H-bond | hydrogen bond |
| His (H) | Histidine |
| HPP | p-hydroxyphenylpyruvate |
| IPTG | isopropyl-β-D-thiogalactopyranoside |
| $k_{cat}$ | turnover number |
| $K_m$ | Michaelis constant |
| LB | Luria-Bertani broth |
| Lys (K) | Lysine |
| MES | N-morpholino ethane sulphonic acid |
| MS | mass spectrometry |
| MW | molecular weight |
| $NAD(P)^+$ | oxidized form of nicotinamide adenine dinucleotide (phosphate) |

| | |
|---|---|
| NATA | free tryptophan analog N-acetyl tryptophanamide |
| NAYA | free tyrosine analog N-acetyl tyrosinamide |
| NEM | 4-ethylmorpholine |
| Ni-NTA | nickel-nitrilotriacetic acid |
| PAT | prephenate aminotransferase |
| PD | prephenate dehydrogenase |
| PDT | prephenate dehydratase |
| Phe (F) | L-phenylalanine |
| PMMA | Poly(methyl methacrylate), transparent thermoplastic |
| PMSF | phenyl-methyl-sulfonyl fluoride |
| PP | phenylpyruvate |
| $\theta_M$ | molar ellipticity |
| $\theta_{MRW}$ | mean residue weight ellipticity |
| Ser (S) | Serine |
| SDS-PAGE | sodium dodecylsulfate polyacrylamide gel electrophoresis |
| Tris | tris(hydroxymethyl)aminomethane |
| Trp (W) | L-tryptophan |
| TS | transition state |
| TSA | transition state analog |
| Tyr (Y) | L-tyrosine |
| UV | ultra violet |
| $V_{max}$ | maximum velocity |
| VT | variable temperature |
| WT | wild-type |
| Δ80 CM-PD | 80 amino acids deleted at the N-terminus of *H. influenzae* full-length CM-PD |

# Chapter 1: INTRODUCTION

## 1.1 Biosynthesis of Aromatic Amino Acids via the Shikimate and Common Pathways

Aromatic amino acids are essential components in the biosynthesis of living organisms as products and precursors in metabolic pathways. Among the 20 amino acids, three possess aromatic side chains and include L-tryptophan (Trp), L-phenylalanine (Phe) and L-tyrosine (Tyr). The biosynthesis of aromatic amino acids uses both the shikimate pathway, named for the intermediate compound, shikimate (1) and the common pathway (1).

The shikimate pathway starts with the condensation of phosphoenolpyruvate (PEP) and erythrose 4-phosphate (E4P) to yield deoxy-D-arabino-heptulosonate 7-phosphate (DAHP). DAHP is then cyclized to form shikimate which after three more steps forms chorismate, a precursor of all three aromatic amino acids (2) as well as aromatic vitamins, quinones, folates and siderophores (2) (Fig 1). Beyond chorismate, the pathway divides into the terminal pathway called the common pathway, specific for each aromatic amino acid.

Because the shikimate and common pathways occur only in microorganisms, fungi and plants, but not in animals (3), the enzymes of these pathways are potential targets for the design of inhibitors which can serve as antimicrobials, fungicides, herbicides and anti-parasitic agents. As an example, glyphosate (trade name: Roundup®) is a well-known herbicide which kills plants by interfering with the synthesis of essential aromatic amino acids. Glyphosate is a structural analog of PEP and it inhibits 5-

enolpyruvylshikimate-3-phosphate synthase (EPSPS) which catalyzes the reaction of PEP and shikimate-3-phosphate to form 5-enolpyruvyl shikimate -3-phosphate (ESP) (4). By interrupting the formation of ESP, which is subsequently dephosphorylated to form chorismate, it disables the biosynthesis of the aromatic amino acids. Further studies elucidating the enzymes in the pathways will help to identify other valuable inhibitor targets and to provide new insights into microbial common pathway.

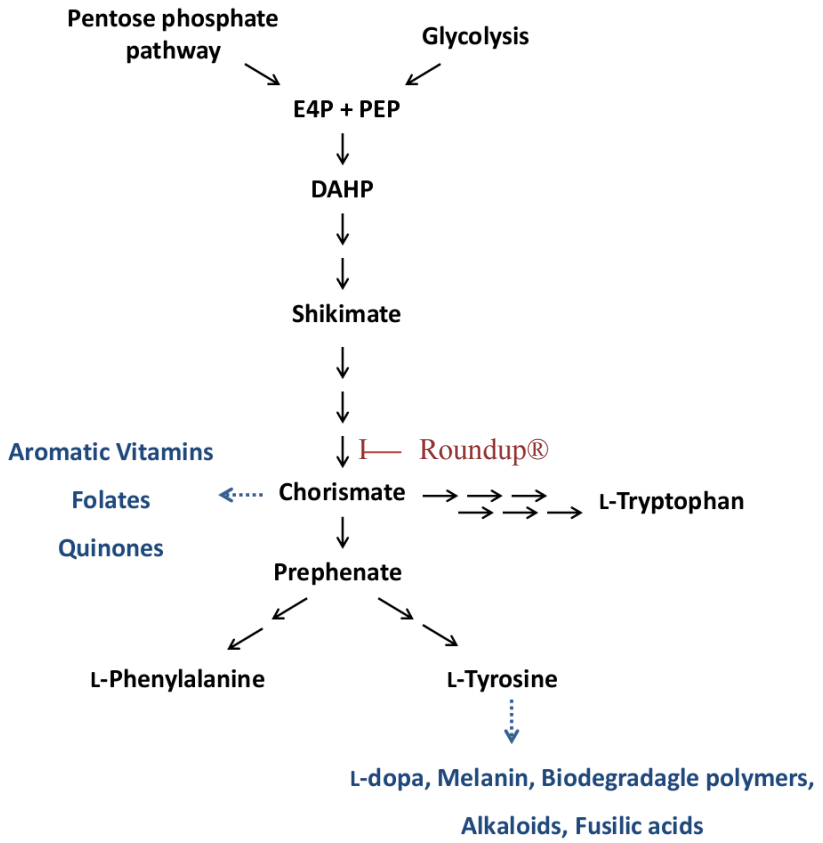



**Figure 1: Aromatic amino acid biosynthesis pathway.** Modified from (1, 5). The precursors, phosphoenolpyruvate (PEP) and erythrose 4-phosphate (E4P) are provided by glycolysis and the pentose phosphate pathway (1). The arrows indicate distinct enzymatic steps.

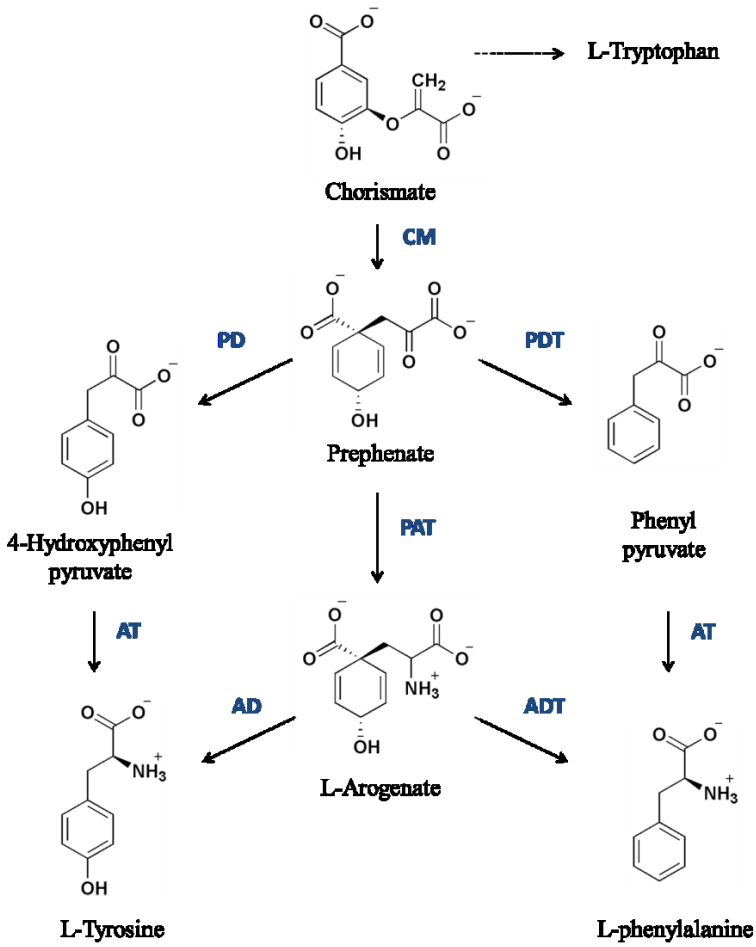



**Figure 2: Biosynthesis of L-phenyalanine and L-tyrosine via the common and the L-arogenate pathways.** The enzymes involved in the pathways are CM (chorismate mutase), PD (prephenate dehydrogenase), PDT (prephenate dehydratase), AT (aminotransferase), PAT (prephenate aminotransferase), ADT (arogenate dehydratase), AD (arogenate dehydrogenase).

Fig 2 shows in more detail the chemical transformations in the common pathway. For L-Phe and L-Tyr, chorismate undergoes a Claisen rearrangement to form prephenate in a reaction catalyzed by chorismate mutase (CM). Prephenate is then oxidatively decarboxylated in the presence of $NAD^+$ by prephenate dehydrogenase (PD) to yield p-hydroxyphenylpyruvate (HPP) and carbon dioxide, or dehydrated followed by decarboxylation by prephenate dehydratase (PDT) to yield phenylpyruvate (PP) (1). These compounds are then transaminated to either L-Tyr or L-Phe which can serve as

feedback regulators of the respective synthetic enzymes. The tyrosine branch of the pathway is of growing interest because L-Tyr is a precursor for valuable compounds such as L-dopa, an anti-Parkinson's drug (6, 7, 8), melanin (9), and biodegradable polymers (10). Therefore, studies elucidating the structures, catalytic mechanisms, and modes of regulation of enzymes in this pathway could not only aid in the design of inhibitors with therapeutic properties, but also lead to the production of bacterial strains containing tyrosine-insensitive enzymes which would be of interest for metabolic engineering purposes (11, 12).

**1.2 Arogenate Pathway to L-Tyrosine and L-Phenylalanine**

As shown in Fig 2, L-Tyr and L-Phe are synthesized from HPP and PP, respectively. However, there is an alternate pathway to these amino acids which uses L-arogenate as an intermediate. L-arogenate was discovered in 1974 by Jensen *et al* (13) from *Agmenellum quadruplicatum* and named pretyrosine to reflect its position in the common pathway. L-arogenate is formed from prephenate by transamination. Next, L-arogenate undergoes dehydration catalyzed by arogenate dehydratase (ADT) or oxidative decarboxylation by arogenate dehydrogenase (AD) to yield L-Phe and L-Tyr, respectively. L-arogenate is structurally similar to prephenate except that the alanyl group is replaced with a pyruvyl side chain at C-1 (14).

The combination of these two routes, the HPP/PP and the arogenate routes, contributes to the diversity in the biosynthesis of L-Tyr and L-Phe. These amino acids are synthesized using the arogenate pathway in most plants and some cyanobacteria, whereas this pathway is absent in proteobacteria such as *E. coli* and yeast (10, 15). In

some cyanobacteria and several other micoorganisms such as *Brevibacterium flavum*, the arogenate pathway is the only means to produce L-Tyr while L-Phe is synthesized from PP (15). HPP and arogenate pathways coexist in other bacterial species such as *Zymomonas mobilis* and *Pseudomonas aeruginosa* (15, 16).

### 1.3 The TyrA Protein Family

TyrA protein family members are comprised of dehydrogenases dedicated to L-Tyr biosynthesis. They can be categorized by their specificities for the cyclohexadienyl substrate, prephenate or L-arogenate (17). Prephenate dehydrogenase ($TyrA_p$) and arogenate dehydrogenase ($TyrA_a$) use prephenate and L-arogenate respectively, while cyclohexadienyl dehydrogenases ($TyrA_c$) can use either substrate in the decarboxylation reaction. TyrA proteins also use $NAD^+$, or $NADP^+$ or both as cofactors depending on the organism (17). $NAD^+$ is preferred by $TyrA_p$ whereas $NADP^+$ is used by mainly $TyrA_a$ proteins. Jensen and coworkers have used extensive phylogenetic analysis combined with amino acid sequence alignments to help predict the identity of residues that might dictate substrate specificity within the various TyrA enzymes (17). For example, specificity for the pyridine nucleotide co-substrate has been determined by studying the region of the enzyme which is in contact with the ribose ring since the only difference between $NAD^+$ and $NADP^+$ is the esterified phosphate group at the adenosine 2'-ribose (15, 17). When the glycine-rich regions of the ADP-binding βαβ fold (Rossmann βαβ) of TyrA proteins sequences are aligned (18), a negatively charged aspartate is found at the consensus position 36 (*E. coli* TyrA numbering) that is critical for hydrogen bonding with the ribose diol group in $NAD^+$-specific TyrA enzymes (*e.g. E.*

*coli* and *Z. mobilis*) (15, 17). In contrast, many $NADP^+$ specific enzymes (*e.g.* *Acinetobacter sp.* and *Synechocystis sp.*) retain an Arg at position 37 to interact with the cofactor's phosphate ribose oxygen (15). Residues that help specify prephenate or L-arogenate specificity are currently not well defined.

Most TyrA proteins exist as mono- or bifunctional dehydrogenases, that is, PD activity is either contained within a single polypeptide or is linked at its N or C terminus to an additional activity. Among the complete genomic sequences which are available only three trifunctional PDs have been identified although only TyrA from *Archaeoglobus fulgidus* has been recombinantly expressed and preliminarily characterized. Studies on *A. fulgidus* TyrA have shown that this enzyme encodes three enzyme activities: PD, CM, PDT thus synthesizing PP and HPP from the same polypeptide chain (19). Examples of these TyrA proteins are shown in Table 1. The focus of this thesis is on the *Haemophilus influenzae* bifunctional CM-PD and its engineered monofunctional PD form.

### 1.4 *E. coli* TyrA: the Structural Relationship between the Two Activities

The most well-studied TyrA protein is the bifunctional *E. coli* CM-PD. *E. coli* CM-PD is a homodimer composed of two ~42 kDa monomers (35). Sequence alignment of CM-PD with CM-PDT (the bifunctional enzyme that synthesizes L-Phe) suggests that of the 373 residues in the CM-PD polypeptide chain, the N-terminal 100 residues encode the mutase domain while the remaining 273 residues are responsible for the dehydrogenase domain (36, 37). There has been much interest in the spatial geometry of

the active site(s) of this bifunctional TyrA, since the product of the CM reaction, prephenate, is the substrate for the consecutive reaction catalyzed by PD.

**Table 1. Examples of mono-, bi- and trifunctional TyrA proteins.**

<table>
<tr><th colspan="3">TyrA</th></tr>
<tr><th>Monofunctional</th><th>Bifunctional</th><th>Trifunctional</th></tr>
<tr><td>PD</td><td>CM fused at N-terminus</td><td rowspan="6">Nanoachaem equitans[19]<br>Archaeoglobus fulgidus[19]<br>♦Ectocarpus siliculosus</td></tr>
<tr><td>*Aquifex aeolicus[20, 21]<br>*Streptococcus mutans[22]<br>Bacillus subtilis[23]<br>Mycobacterium tuberculosis[24,25]<br>*Streptococcus thermophilus<br>*Corynebacterium glutamicum</td><td rowspan="3">Escherichia coli[30]<br>Erwinia herbicola[31]<br>*Haemophilus influenzae[32]<br>Aerobacter aerogenes[38,39]</td></tr>
<tr><td>AD</td></tr>
<tr><td>*Synechocystis[26]<br>◊Nicotiana silvestris[27]</td></tr>
<tr><td>CDH</td><td>EPSP fused at C-terminus[24]</td></tr>
<tr><td>Zymomonas mobilis[28]<br>Pseudomonas stutzeri[29]</td><td>Pseudomonas fluorescens[33]<br>Acinetobacter calcoaceticus[34]</td></tr>
</table>

PD: prephenate dehydrogenase, AD: arogenate dehydrogenase, CDH: cyclohexadienyl dehydrogenase, CM: chorismate mutase, EPSP: phosphoshikimate carboxyvinyltransferase. * Crystal structures (monofunctional forms) are available in Protein Data Bank. ◊ Herbaceous plant, ♦ Brown alga.

The idea that both CM and PD reactions were catalyzed at a single active site came from studies on CM-PD from *Aerobacter aerogenes* which showed the parallel loss of both activities in the presence of heat and cysteine modifying reagents (38, 39). Further support came from kinetic studies of the *E. coli* protein which showed that the Michaelis constant of prephenate in the PD reaction was similar to its dissociation constant as an inhibitor in the CM reaction (40). Evidence for two distinct active sites stemmed from the results of site-directed mutagenesis studies of *E. coli* CM-PD from the Turnbull laboratory (40, 41) which showed that a single amino acid substitution in either

the CM domain (K37Q) or the PD domain (H197N) completely eliminated the respective activities. In addition, compounds mimicking the proposed transition-state of each reaction resulted in preferential inhibition of each of the activities (42, 43). The idea of overlapping active sites stemmed from inhibition studies of *E. coli* CM-PD by Christopherson *et al* (44); they showed that the dicarboxylic acid malonate could combine with the enzyme when chorismate was bound, while the larger diethyl derivative of malonate prevented the binding of both chorismate and prephenate. Computer simulation of reaction rates and radioactive substrate channeling experiments by Heyde showed that some prephenate synthesized from chorismate by CM-PD could be converted directly to HPP indicating that the two active sites are structurally linked (39). Further evidence supporting the close spatial proximity of both active sites came from site-directed mutagenesis studies by Turnbull and coworkers (45) who showed that amino acid replacements in the PD domain (K178R, H189N, R286A) impaired both CM and PD activities.

The first report of an engineered monofunctional PD derived from a bifunctional CM-PD came from studies on *Erwinia herbicola* TyrA (31). Jensen and coworkers based their conclusion on the ability of a deletion construct to complement the growth of a *pheA- tyrA-* strain in the presence of prephenate and L-Phe. Since this *tyrA* encoded over 60% of the mutase domain, it was not surprising that sufficient interactions to support a monofunctional PD were retained (31). Ganem and coworkers extended this study by attempting to express and purify several N- and C-terminal deletion variants yielding monofunctional CM and PD enzymes from *E. coli* CM-PD (47). The proteins, however, were either mostly inactive or highly unstable indicating the structural interdependence of

the CM and PD domains. In contrast, bifunctional *E. coli* CM-PDT involved in the biosynthesis of L-Phe has been reported to possess non-interacting sites and a separate regulatory domain that binds the end product (48). Interestingly, a monofunctional PD has been generated by deleting the N-terminal 80 residues of *H. influenzae* CM-PD (The full sequence of *H. influenzae* TyrA is in Appendix 1). This variant is both fully folded and crystallizable (32). Additionally, initial studies by Quashie showed that both the bifunctional enzyme and the engineered monofunctional PD were active (49). These two *H. influenzae* TyrA enzymes are the focus of my thesis work.

**1.5 Chorismate Mutase Reaction and Structure**

Chorismate mutase catalyzes the conversion of chorismate to prephenate by a Claisen rearrangement. Although this rearrangement is possible without the enzyme, the process is accelerated over $10^6$ fold in the presence of CM (50). Both the uncatalyzed (51, 52) and the enzyme-catalyzed reaction (53) are believed to proceed via a chair-like transition state (TS) formed from a chorismate conformation resembling its less stable diaxial form (53) (Fig 3). Ganem and coworkers solved the crystal structure of the independently expressed CM domain of *E. coli* CM-PDT (the "mini-mutase") in complex with a transition state analog, endo-oxabicyclic diacid, which binds over 300 times more tightly to *E. coli* CM-PD than chorismate (52). The structure (Fig 4) reveals the importance of Lys 39' and Gln 88 in stabilizing the ether oxygen as well as Arg 11' and Arg 28 in maintaining the important diaxial conformation found in the TS of the reaction. Extensive site-directed mutagenesis by Turnbull (54) and others (59, 60, 61) has shown the functional importance of active site residues in the CM-catalyzed reaction.

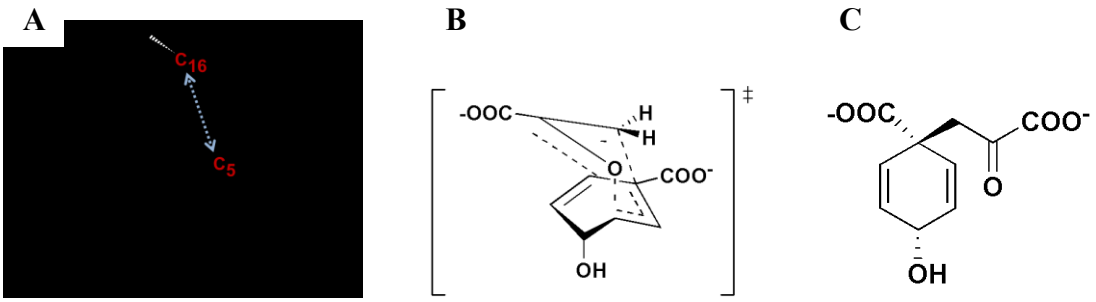



**Figure 3: Rearrangement of chorismate through a transition-state complex.** (a) Near attack conformation of chorismate. (b) Transition state. (c) Prephenate. Zhang *et al* (53) proposed that the chorismate (diequitorial form) undergoes a rearrangement and rotation of the C3-O13 which places C16 π-orbital directed towards the C5 atom: a near-attack conformation (NAC). TS stabilization energy from NAC to TS in the presence and absence of CM supported the idea that the efficiency of the enzymatic reaction is largely due to the stabilization of NAC species (53). Adapted from Zhang *et al* (53).

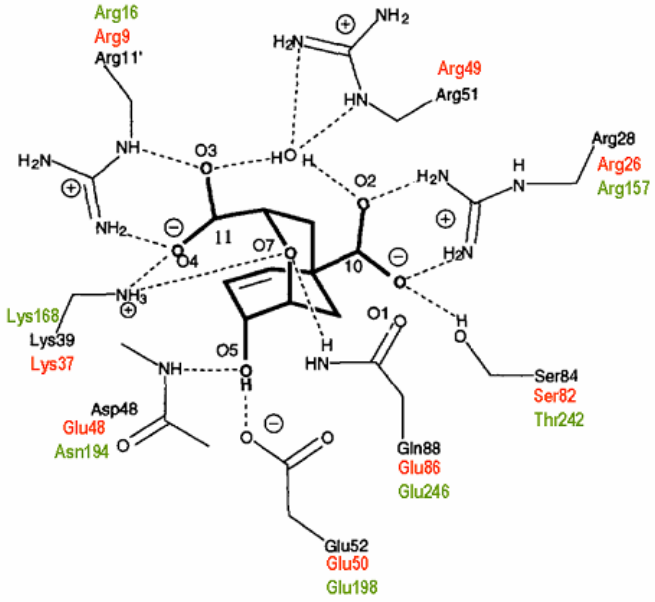



**Figure 4: Schematic diagram of the crystal structure of the active site of *E. coli* "mini-mutase" complexed with endo-oxabicyclic diacid**. Residues from *E. coli* "mini-mutase" (black) and *S. cerevisiae* (green) were determined by X-ray crystallography while those from *E. coli* CM-PD (red) were elucidated through sequence alignment. Note that the active site residues originate from both monomers (one monomer has primed residues). This picture was adapted from Lee and coworkers (52).

With regard to their structure, CMs belong to two classes (Fig 5): the less abundant aroH, composed of trimeric α/β barrels, and the more ubiquitous AroQ, composed of α-helical dimers and of which the *E. coli* mini-mutase is an example. The

mutase domain of other bifunctional CM-PDs (*e.g. E. coli*, *H. influenzae*) likely adopts an AroQ fold. Additionally, CMs from both families generally have a shared active site (one per monomer), with residues from both monomers contributing to each catalytic center.

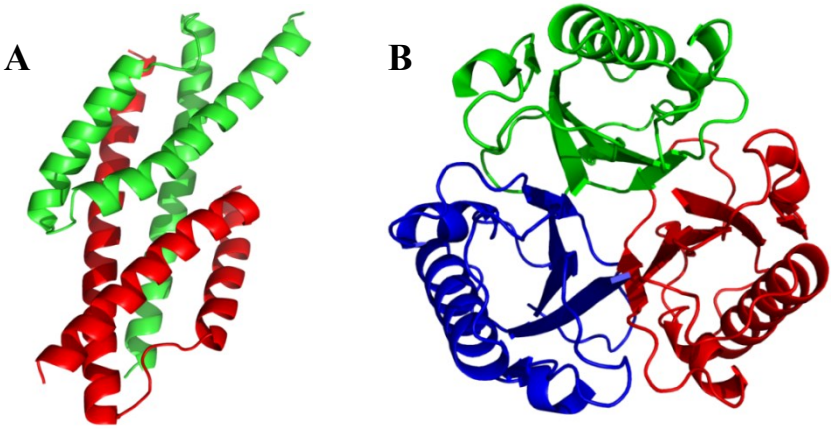



**Figure 5: Chorismate mutase from (A) AroQ and (B) AroH classes.** (A) *E. coli* CM from CM-PDT belongs to AroQ class. Each monomer is colored in green and red. (B) *B. subtilis* CM belongs to AroH class. Each monomer is colored in green, red and blue. This picture was created using PYMOL (PDB accession number: 1ECM and 2CHS, respectively) (30, 55).

### 1.6 Prephenate Dehydrogenase-Catalyzed Reaction

Prephenate is oxidatively decarboxylated by prephenate dehydrogenase in the presence of $NAD^+$ to yield HPP and $CO_2$. A chemical mechanism for the PD-catalyzed reaction of *E. coli* CM-PD was proposed earlier based on the results of isotope effect studies (59), peptide mapping (42), pH profiles (42, 59), and site-directed mutagenesis (45, 54). The results of the analysis of initial velocity patterns in the presence of products and dead-end inhibition allowed elucidation of the kinetic mechanism, indicating that *E. coli* PD conforms to a rapid equilibrium random mechanism with catalysis as the rate limiting step (35, 42). Both $^{13}C$ and deuterium kinetic isotope effect studies showed that the hydride transfer and the decarboxylation occurred in the same chemical step (59). Fig

6 shows the proposed catalytic mechanism of the *E. coli* PD reaction highlighting key active site residues involved in the reaction. The H-bond acceptor in *E. coli* CM-PD is reported to be His197 while Arg294 is reported to have a role in the binding of prephenate (42, 45, 54).

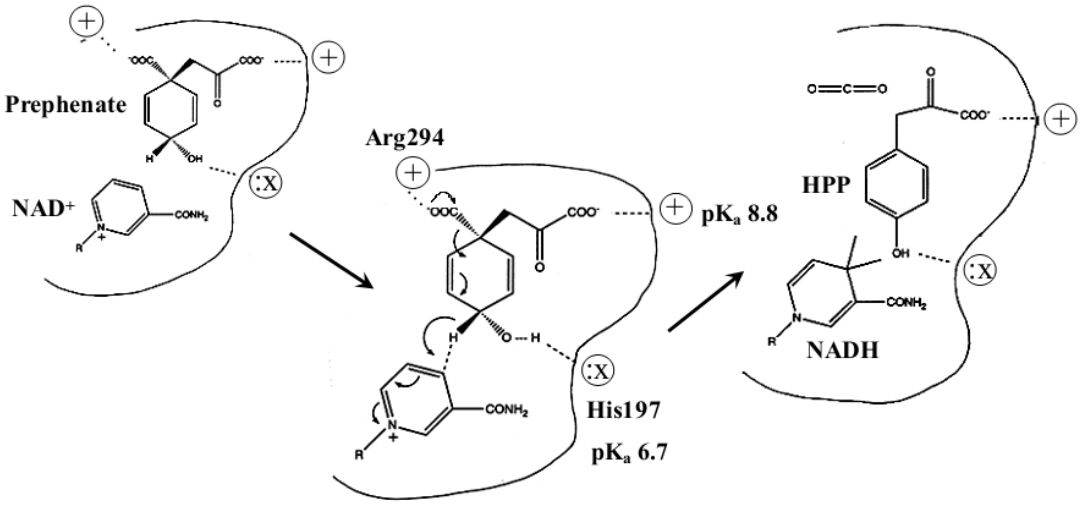



**Figure 6: Proposed mechanism for the reaction catalyzed by prephenate dehydrogenase of *E. coli* CM-PD.** Prephenate and $NAD^+$ bind to different cavities in the active site of the PD domain. Deprotonated His197 ($pK_a$ ~6.7) is believed to assist hydride transfer from the 4–hydroxyl group of prephenate to $NAD^+$ and concomitant decarboxylation by polarizing prephenate's 4-hydroxyl group. The formation of carbon dioxide and the aromatic HPP provides the diving force to lower the activation energy barrier. Arg294 is proposed to interact with the ring carboxylate of prephenate bound to the enzyme-$NAD^+$ complex. Another protonated group (unknown identity) with a $pK_a$ of 8.8 is believed to interact with the pyruvyl side chain of prephenate to lock prephenate in the active site. This mechanism was proposed by Christendat and Turnbull (45).

L-Tyr, the end product of the pathway, is a feedback inhibitor of bifunctional CM-PD in the presence of $NAD^+$ (41, 60). The affinity of the enzyme for L-Tyr is markedly increased in the presence of $NAD^+$ and vice-versa (41). Interestingly, CM is reported to be insensitive to L-Tyr's effects in the absence of $NAD^+$ (40). Hudson and colleagues (60) showed through ultracentrifugation studies that dimeric *E. coli* TyrA formed a tetramer in the presence of L-Tyr and $NAD^+$ and proposed that this tetramer

was inactive. Double reciprocal plots generated by varying prephenate as a function of L-Tyr concentrations were concave upwards and the inhibition appeared competitive with respect to prephenate although there is considerable debate as to whether or not L-Tyr binds at a separate allosteric site (41, 60).

It is worth noting that the detailed mechanistic studies on *E. coli* CM-PD have been conducted without the aid of 3D structural information. Diffraction quality crystals have yet to be obtained for this 84 KDa bifunctional protein in spite of the efforts of many groups over the past 25 years.

## 1.7 Crystallographic Studies of TyrA Proteins from *H. influenzae* and Other Organisms

In 2006, the first three dimensional structure of a TyrA protein from any organism along with its biophysical characterization were reported by the laboratories of Christendat and Turnbull (20, 21). An N-terminal deletion variant of a monofunctional PD (Δ19PD) from the hyperthermophilic bacterium, *Aquifex aeolius* was recombinantly expressed in *E. coli* and its structure was solved to 1.9Å resolution in the presence of $NAD^+$. In 2009, more structures were obtained in the presence of $NAD^+$, HPP and other product analogs, including L-Tyr (63). Also in 2006, Legrande and coworkers reported the first crystal structure of a monofunctional arogenate dehydrogenase from the cyanobacterium, *Synechocystis sp* (strain PCC 6803/Kazusa). The structure was resolved to 1.6 Å with $NADP^+$ bound to the protein (26). Since then, there have been a number of other TyrA protein structures reported: a monofunctional PD from *Streptococcus mutans, Corynebacterium glutamicum, Streptococcus thermophilus* and the PD domain of

bifunctional PD from *Haemophilus influenzae*. Of importance to this thesis, the Joint Centre for Structural Genomics deposited the coordinates in the PDB (2008) and later published (2010) the crystal structure of the independently expressed PD domain of bifunctional *H. influenzae* CM-PD bound with L-Tyr and $NAD^+$ determined to 2.0 Å resolution (32). The PD domain was obtained by expressing residues 81-377 of the bifunctional enzyme thereby deleting ~80% of the mutase domain. The first solution studies were initiated by Quashie in the Turnbull laboratory on both the mono- and bifunctional forms of this TyrA (49).

*H. influenzae* CM-PD is a homodimer with monomer molecular weight of ~43 kDa. The crystallizable PD domain is also dimeric in solution and its global structure shows a remarkable similarity with *A. aeolicus*, *Synechocystis sp* and the other crystallized TyrA proteins. Each monomer of *H. influenzae* PD consists of a nucleotide-binding domain (residue 92-243) at the N-terminal region of the polypeptide chain and an α-helix dimerization domain (residues 244-371) at the C-terminus. The protein nucleotide-binding domain contains a modified Rossmann fold with seven parallel stranded β-sheet and five α-helices. The dimerization domain has seven helices that form the dimer interface which is intertwined into a helical bundle (32) (Fig 7). The active site (one per monomer) is at the interface of the two domains. Residues in the PD domain of *H. influenzae* and *E. coli* CM-PD (residues ~ 100-373) share about 57% amino acid sequence identity whereas there is only ~18% sequence identity with the PD from monofunctional *A. aeolicus* TyrA. Nonetheless, the overall geometry of TyrA active site appears well conserved in all three organisms, including many of the residues that are involved in substrate binding and catalysis in *E. coli* TyrA (Fig 7 inset) (49).

### 1.8 The Structure and Stability of TyrA Protein

There have been very few studies examining the relationship between the structure of TyrA proteins and their stability. However, all studies point to the importance of the dimerization interface in stabilizing TyrA. Previous studies by Christendat and Turnbull (45) examining the elution profile of *E. coli* CM-PD by size exclusion chromatography in the presence of the chemical denaturant guanidine hydrochloride (Gdn-HCl) suggested that the unfolding of this TyrA protein proceeded via an unordered dimer rather than first dissociating to folded monomers. Similarly size exlusion studies by Bonvin *et al* (20) from the Turnbull lab revealed that in the presence of Gdn-HCl, *A. aeolicus* TyrA likely proceeded via a pathway involving partially unfolded dimers which then associated into a higher molecular weight species before collapsing to unfolded monomers. Both of these studies from the Turnbull lab also included the examination of the Gdn-HCl-induced loss of secondary structure of *E. coli* and *A. aeolicus* TyrA as followed by circular dichrosim spectroscopy. Very recent studies by Osuna *et al* (65) further highlighted the importance of the dimerization interface to the stability of TyrA. These authors replaced the CM domain of bifunctional *E. coli* CM-PD by the β1-domain of protein G (Gβ1) and fused it to the N-terminus of PD domain (Gβ1-TyrA). The resulting protein was fully active and retained its dimeric conformation.

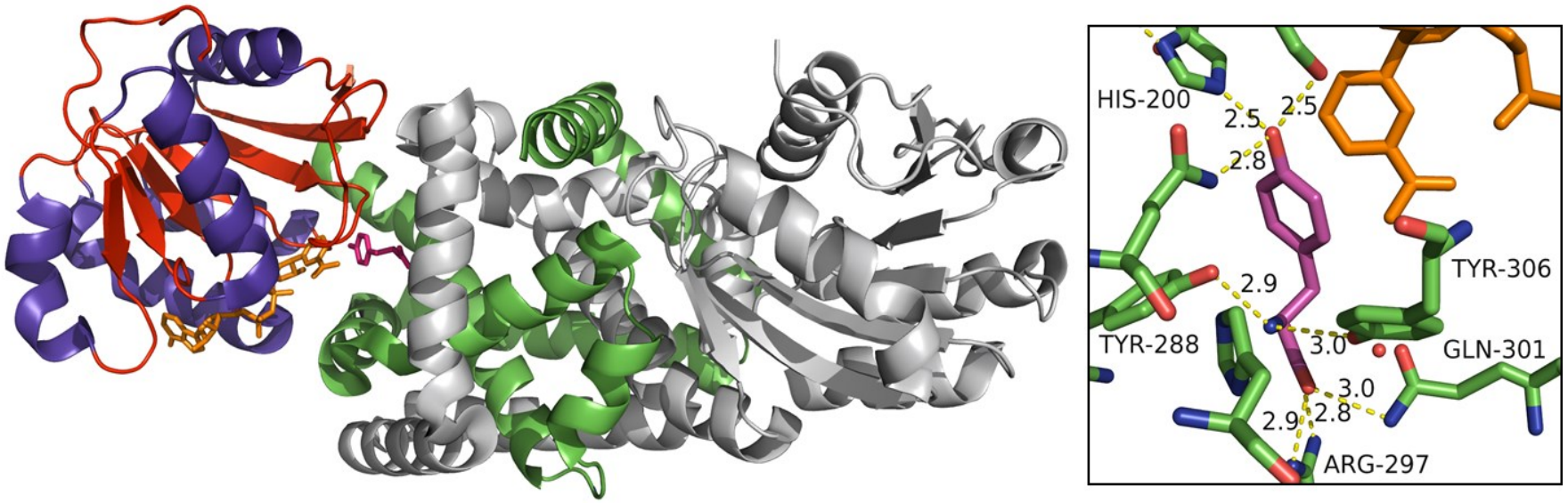



**Figure 7: Crystal structure of *H. influenzae* dimeric PD with bound L-Tyr and $NAD^+$.** The monofunctional PD was generated by deleting 80 amino acids at the N-terminal of the protein thus eliminating 80% of the CM domain. One monomer is shown in light grey and the other in green (C-terminal dimerization domain), purple (α-helix) and red (β-sheet) at the N-terminal, a modified Rossman fold. $NAD^+$ is shown in orange and L-Tyr in pink. This picture was generated using Pymol (64) using the PDB coordinates 2PV7. The inset shows a close-up view of the active site containing L-Tyr and $NAD^+$.

However, when site-directed mutagenesis in Gβ1 interrupted the dimeric conformation of the N-terminal domain, Gβ1-TyrA could not be expressed, supporting the importance of dimeric conformation of the two domains in stabilizing TyrA structure. As previously mentioned, evidence suggests that there is likely a striking interdependence of the CM and PD domains of *E. coli* CM-PD. However, more comprehensive studies are required in order to understand the influence imparted by each domain on the global structure and stability of the bifunctional protein.

The availability of both mono- and bifunctional forms of *H. influenzae* TyrA has allowed us to further examine the stability and solution structure of TyrA. The main focus of this thesis is to examine the effect that the mutase domain may have on PD stability in the presence of temperature, or the chemical denaturant, Gdn-HCl. We used circular dichroism and FTIR spectroscopies as probes of secondary structure, and intrinsic fluorescence emission as a probe of tertiary structure.

The following sections will introduce how the approaches mentioned above can be used to provide valuable information regarding protein structure in solution.

### 1.9 The Chaotropic Agent, Guanidine Hydrochloride

A biologically active protein exists in a compact, highly ordered conformation referred to as the native state. The native state is a result of a balance between the force favoring the unfolded state (conformational entropy) and the one favoring the folded state, which comprises mainly non-covalent interactions within the polypeptide chain and between the chain and its surrounding solvent. Under physiological conditions, the

native (folded) and denatured (unfolded) states of protein are in equilibrium: Native (folded, N) ↔ denatured (unfolded, D). The conformational stability of a protein can be deduced from the free energy change, ΔG, of this equilibrium reaction (66). In the presence of a chemical chaotropic agent, such as guanidine hydrochloride, the folding/ unfolding equilibrium is shifted so that an equilibrium constant can be measured. This can be extrapolated to determine the conformational stability of a protein in physiological conditions (66). Guanidine hydrochloride in a solution at pH 7 is protonated to guanidinium (Fig 8) which interacts with the hydrophobic regions of a protein through its flat, nonpolar surface, while exposing its polar/charged edges to the solvent (67). This denaturant also thermodynamically stabilizes the denatured chain which facilitates solvation of the nonpolar surface (including the backbone) of the denatured protein (67). In this thesis, the stability of *H. influenzae* TyrA proteins, both FL and truncated form (Δ80) was studied in the presence of guanidine hydrochloride as a denaturant in order to investigate the stabilizing effect of the CM domain on PD structure.

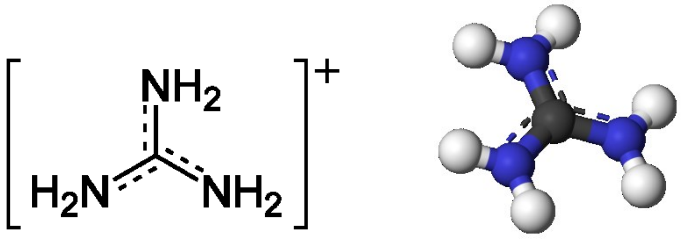



**Figure 8: The structure of guanidinium.** The ball and stick model of guanidinium (on the right) has nitrogen in blue, carbon in black and hydrogen in white.

### 1.10 Application of CD Spectroscopy to Assessing Protein Secondary Structure

Circular dichroism refers to the difference in absorption between left and right-handed circularly polarized light in an optically active or chiral molecule (68). CD

spectropolarimeters measure this difference in circularly polarized light, known as elliptical polarization, as a function of wavelength. Elliptical polarization in the far-UV portion of the light spectrum (170-250 nm) is absorbed by amide and carbonyl groups in the protein peptide backbone. The energy absorption of the electronic transitions depends on the configuration of the backbone and thus the secondary structure of the protein. In far-UV CD, the α–helix yields the most intense and characteristic signal at 208 nm and 222 nm (68, 69). Summarized in the table below are the most common secondary structural features in proteins/peptides with their characteristic CD signatures in the far-UV region.

**Table 2. The characteristic far-UV CD signatures of common secondary structures (69)**

| Secondary structure | $\lambda_{min}$ (nm) | $\lambda_{max}$ (nm) | intensity |
|---|---|---|---|
| α-helix | 208, 222 | 190 | Strongest bands |
| β-turn | No single characteristic signal | | |
| β-sheet | 217 | 191-197 | Less intense than α-helix, simpler |
| Random structure | 192-201 | 208-212 | Large $\lambda_{min}$ and small $\lambda_{max}$ |

The relative amounts of α-helices, β-sheets, turns or unordered structures in a protein can be estimated from the CD spectrum using several commercially available web-based algorithms (68, 69, 70, 71). Algorithms are based on sophisticated curve-fitting procedures involving sets of spectra of well-characterized proteins of known secondary structure. One of the most popular packages, CONTIN, devised by Provencher and Glockner, analyzes the far-UV ellipticities from 240-190 nm as a linear combination of the CD spectra of 16 proteins whose structures have been determined by X-ray crystallography (73).

## 1.11 Application of Fluorescence Emission Spectroscopy in Assessing the Protein Tertiary Structure

Fluorescence is a molecular phenomenon whereby light energy absorbed by π electron rich molecules (Tyr and Trp in proteins) is emitted at a lower energy or longer wavelength. Changes in tertiary structure of a protein can be assessed by fluorescence emission spectroscopy by monitoring alterations in the local environment surrounding Trp and Tyr residues. Trp fluorescence emission is highly sensitive to its local environment. Trp emission spectra can be perturbed in response to conformational transitions and protein unfolding (73). Trp is particularly sensitive to quenching due to its ability to donate electrons from the excited indole ring (73). Hence, a solvent exposed Trp shows a $\lambda_{max}$ emission at about 355 nm with a low intensity (quenched by solvent) whereas a completely buried Trp can exhibit a $\lambda_{max}$ emission of 330 nm or less with higher intensity. Tyr generally does not show a dramatic $\lambda_{max}$ emission shift, but its quantum yield is easily influenced by environment because Tyr transfers emission energy to nearby Trp residues.

*H. influenzae* FL CM-PD has 13 Tyr per monomer (12 of which are in the PD domain), and 4 Trp per monomer (all in the PD domain). Δ80 CM-PD possesses one less Tyr but has the same number of Trp per monomer as FL CM-PD. Chemically induced unfolding of *H. influenzae* TyrA monitored by fluorescence emission and circular dichroism spectroscopies can be used to correlate the changes in tertiary structure in relation to the secondary structure.

## 1.12 Application of FTIR Spectroscopy in Assessing the Protein Secondary Structure

Fourier transform infrared (FTIR) spectroscopy is an effective solution technique for detecting changes in H-bonding patterns within proteins and is often used as a complementary technique to CD spectroscopy in the analysis of protein secondary structure.

The study of protein conformation using infrared spectroscopy commenced in the early 1950's by Elliot and Ambrose (74). They showed the correlation between the secondary structure of the polypeptide backbone and the mid-infrared region spectra (400 to 4000 $cm^{-1}$) of proteins. The characteristic infrared amide bands of proteins are shown in Table 3. Amide I and II bands are most frequently used for assessing the secondary structures of the polypeptide chain. In order to avoid the strong infrared absorption of water molecules between 1700 and 1600 $cm^{-1}$ (amide I region of the protein) due to the H-O-H bending vibration, proteins are reconstituted in $D_2O$ instead of water thus shifting these D-O-D bending vibrations to a distinct spectroscopic window at 1300-1200 $cm^{-1}$. These spectra are referred to as amide I' and II'. The amide I (I') band originates primarily from the C=O stretching vibration of the protein backbone. Changes in conformation cause perturbations in the H-bonding and the polarity of the C=O bond which results in a difference in absorption. These absorptions are very characteristic and can be assigned as components of secondary structure, such as the α-helix, parallel and anti-parallel β-sheets, turns and unordered structures (Table 4).

For better resolution of these bands, computational band-narrowing, Fourier self-deconvolution (FSD) or second derivatives can be used. These methods mathematically

decrease the width of infrared bands leading to an increased separation between individual components. In contrast to CD measurements, bands from β-sheet structures are easily distinguishable from those of the α-helix (Table 4) making FTIR an excellent spectroscopic tool for monitoring changes in the secondary structural elements by physiochemical perturbations. The amide II' band monitors the compactness of the protein. At ~1555-1520 $cm^{-1}$, the rate of H/D exchange of N-H groups in the protein can be recorded as a function of time, temperature, pressure or pH change. A rapid drop in the amide II' band indicates a faster H/D exchange implicating the easily solvent accessible tertiary structure of a protein.

**Table 3. Characteristic infrared amide bands of proteins** (75).

| **Designation** | **Wavenumber ($cm^{-1}$)** | **Assignment** |
|---|---|---|
| A | 3300 | N-H stretching |
| B | 3110 | N-H stretching |
| I | 1653 | 80% C=O stretching; 10% C-N stretching; 10% N-H bending |
| II | 1567 | 60% N-H bending; 40% C-N stretching |
| III | 1299 | 30% C-N stretching; 30% N-H bending; 10% C=O stretching; 10% O=C-N bending; 20% other |
| IV | 627 | 40% O=C-N bending; 60% other |
| V | 725 | N-H bending |
| VI | 600 | C=O bending |
| VII | 200 | C-N torsion |

**Table 4. Characteristic amide I band assignment of protein secondary structure** (75).

| Wavenumber ($cm^{-1}$) | Secondary structure assignment |
|---|---|
| 1695-1670 | Intermolecular β-structure |
| 1690-1680 | Intramolecular β-structure |
| 1666-1659 | '3-turn' helix |
| 1657-1648 | α-helix |
| 1645-1640 | Random coil |
| 1640-1630 | Intramolecular β-structure |
| 1625-1610 | Intermolecular β-structure |

### 1.13 Two-Dimensional IR Correlation Spectroscopy

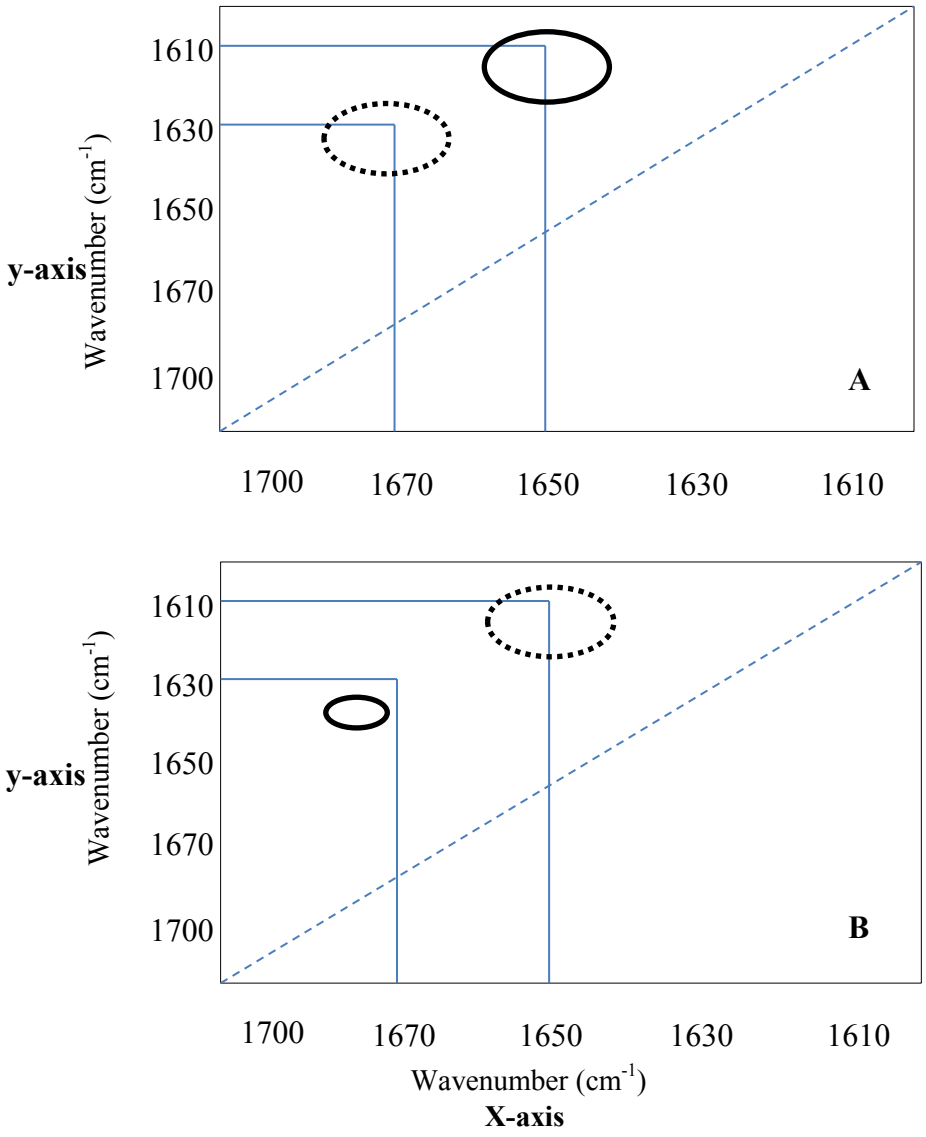
The position and the relative intensity of IR bands of a sample are sensitive to subtle changes in the sample environment (75). These subtle changes may not be apparent in 1D spectra, thus 2D IR correlation spectroscopy was developed (76). Identification of molecular interactions and the sequence of events (*e.g.* unfolding of protein structure) can be deduced using correlation bands. The algorithm for 2D IR correlation spectroscopy detects the dynamic changes occurring in the spectra; the resulting equation from the algorithm generates two maps. The synchronous map represents the similarity of the time-dependent spectral intensity variations at two different wavenumbers while the asynchronous map represents the difference in the dynamic behavior. These maps are analyzed using a set of rules (76). The analysis of these two maps can be used, for example, to determine the sequence of temperature-induced unfolding of protein structure in a variable temperature experiment.

In order to assign the sequence of unfolding, two signs must be determined: the first sign originates from the synchronous map and the second sign from the

asynchronous map. Only the top halves of the synchronous and asynchronous maps are used to determine the signs. To define the first sign, one needs to observe how two peaks move relative to each other: a positive sign is given if two peaks move together (increase or decrease), a negative sign is given if two peaks move oppositely. For the second sign, it is necessary to examine the cross-point of two different peaks whose wavenumbers are assigned on the x-axis and y-axis. For example, to determine how an α-helix component (1650 $cm^{-1}$) is different in dynamic behavior from intramolecular β-sheet (1630 $cm^{-1}$) as a function of temperature, one must assign the wavenumber of the α-helix band in the x-axis and the intramolecular β-sheet in the y-axis and confirm where the cross-point falls. The result will be a positive sign if the cross-point falls within or on the solid line, whereas a negative sign will be assigned if the cross-point lands within or on the dotted line. When there is no cross peak, it means one of two things: the intensity is too weak to assign positive or negative values, or the changes in intensities are directly coupled producing a null effect (denoted as 0). When two signs from the synchronous and asynchronous maps are different, the peak intensities assigned on the y-axis occur before those on the x-axis on asynchronous map, whereas it is the opposite when the two signs are the same.

Fig 9 shows a simple example of how to interpret 2D correlation maps. One examines the synchronous map (from Fig 9A) first to determine the first sign. At 1650 $cm^{-1}$ (x-axis) and 1610 $cm^{-1}$ (y-axis), the cross-point is within the solid line, therefore; the sign will be positive. The cross-point at 1670 $cm^{-1}$ (x-axis) and 1630 $cm^{-1}$ (y-axis) falls within dotted line, thus this sign will be negative. These signs are noted on the table in Fig 9. For the second signs, the cross-point from asynchronous map (Fig 9B) is used.

Using the same method, the cross peak at 1650 $cm^{-1}$ (x-axis) and 1610 $cm^{-1}$ (y-axis) is negative while the intersection of 1670 $cm^{-1}$ (x-axis) and 1630 $cm^{-1}$ (y-axis) is zero. From the table, the order of appearance or disappearance of peaks can be determined: when the two signs are different, those on the y-axis occur before the corresponding x-axis. As noted in the table in Fig 9, analysis of the peaks at 1650 $cm^{-1}$ (x-axis) and 1610 $cm^{-1}$ (y-axis), indicates that the event at 1610 $cm^{-1}$ occurs before 1650 $cm^{-1}$. However, for intensities at 1670 $cm^{-1}$ (x-axis) and 1630 $cm^{-1}$ (y-axis), the two peaks appear simultaneously since the second sign is zero.




| | | x-axis ($cm^{-1}$) | |
|---|---|---|---|
| y-axis ($cm^{-1}$) | | **1670 (α-helix)** | **1650 (α-helix)** |
| | **1610 (intermolecular β-sheet)** | **0 , 0** | **+ , - (→)** |
| | **1630 (intramolecular β-sheet)** | **- , 0** | **0 , 0** |

**Figure 9: An example of (A) synchronous and (B) asynchronous 2D correlation maps.** The first sign in the table is assigned from the cross-point of a synchronous map (blue) and the second sign is from that of an asynchronous map (red). The solid and dotted ovals are a correlation of a change in a frequency on the x-axis with that on the y-axis. The wave numbers can be assigned to specific secondary structures.

## 1.14 Research Objectives

The main goal of this project was to compare the chemical and thermal stabilities of *H. influenzae* CM-PD and its engineered monofunctional PD, Δ80 CM-PD, as efforts towards understanding the importance of the mutase domain on PD structure. Both full-length and Δ80 CM-PD were expressed and purified using Ni-NTA affinity chromatography. Denaturation of the proteins by temperature and by the chemical chaotrope guanidine hydrochloride (Gdn-HCl) were followed using circular dichroism (CD) and intrinsic fluorescence emission as probes of secondary and tertiary structural changes, respectively. The results of the chemical denaturation experiments performed on *H. influenzae* TyrA in two different buffer systems were compared with that of *E. coli* CM-PD. The thermal denaturation of both *H influenzae* proteins was further monitored by using 1D and 2D correlation FTIR spectroscopy in an effort to delineate the temperature-mediated unfolding pathway of TyrA proteins in the presence and absence of the mutase domain. The thermal stabilization afforded by the binding of L-Tyr and $NAD^+$ to *H. influenzae* TyrA proteins was also examined. In order to maintain a consistent thermal denaturation environment measured by both FTIR and CD spectroscopies, the protein samples were prepared using the same conditions (high concentration of protein, similar buffer components) allowing an improved comparison of the two complementary methods. To our knowledge this is the first reported study of its kind. Our results were interpreted in light of the crystal structure for *H. influenzae* Δ80 CM-PD in complex with $NAD^+$ and L-Tyr. This work extends the solution studies of *H. influenzae* CM-PD of Quashie (49) monitoring the relationship between CM and PD structure and activity.

# Chapter 2: MATERIALS AND METHODS

## 2.1 Chemicals, Reagents and Enzymes

The free acid form of chorismic acid was prepared from *Klebsiella pneumonia* 62-1 as described by Rieger and Turnbull (77), while prephenate (sodium salt) was obtained from enzymatic conversion of chorismic acid as described by Dudzinski and Morrison (78). $NAD^+$ (grade I, free acid) was purchased from Roche. Substrates were prepared in appropriate buffers (pH 7.5) and stored at -20ºC or -80ºC. Concentrations of stock solutions of chorismic acid and $NAD^+$ were determined spectrophotometrically using published extinction coefficients (79) while prephenate concentrations were determined from enzyme-mediated end-point analysis (80). L-Tyr was purchased from ICN Biochemicals Inc. Ampicillin (sodium salt), dithiothreitol (DTT) and guanidine hydrochloride (Gdn-HCl, Biotech grade >99.5%) were all from BioShop. A 1 M stock of DTT was made in distilled water and stored in -20ºC in 1 mL aliquots. Roche supplied the EDTA-free protease inhibitor cocktail, benzonase nuclease and deoxynucleotide triphosphates (dNTPs). The dNTPs mixture was prepared to a concentration of 5 mM, and stored in 100 µL aliquots at -20ºC. Phenyl-methyl-sulfonyl fluoride (PMSF) was from EMD Biosciences Inc. and it was prepared as a 0.5 M stock in methanol and stored at -20ºC in 200 µL aliquots. Methanol, acetonitrile (ACN), formic acid (FA) and $C_4$ Zip tips were from Fisher Scientific. $D_2O$ was purchased from Sigma-Aldrich as 10 g sealed vial and stored in 200 µL aliquots in eppendorf tubes at room temperature to avoid contamination with $H_2O$. Protein assay dye agent and molecular weight standards were from Bio-Rad, while NAP™-5 size exclusion buffer exchange columns prepacked with DNA grade Sephadex™ G-25 were from Amersham Biosciences. Nickel-nitrilotriacetic

acid (Ni-NTA) SuperFlow chromatography resin was purchased from Qiagen and regenerated before protein purification as per manufacturer's instructions. Dialysis membrane (12-14 kDa molecular weight cutoff) was from Spectrapor. Thrombin protease was obtained from GE Healthcare Life Science. Free tyrosine analog, N-acetyl tyrosinamide (NAYA), and tryptophan analog, N-acetyl tryptophanamide (NATA), were supplied from Sigma-Aldrich. All other chemical reagents and solvents were purchased commercially.

*Haemophilus influenzae tyrA* was amplified from genomic DNA and DNA encoding both the full-length (FL) and the Δ80 variant CM-PD and cloned into *E. coli* expression vector pET-15b (Novagen) by Quashie (49) to yield the recombinant plasmids pPKQ1 and pPKQ2, respectively. Recombinant wild-type *E. coli* CM-PD plasmid (Clone G) was constructed in the Turnbull lab previously by cloning *E. coli TyrA* into the same expression vector (36). Accordingly, recombinant proteins contained a cleavable N-terminal hexahistidine (hexa-His) tag to facilitate protein purification. BL21 (DE3) Gold cells $F^-$, *omp*T, *hsd*$S_B$($r_B$-, $m_B$-), *dcm*, *gal*, λ(DE3) (Stratagene) were used for protein expression and were rendered competent using calcium chloride (79). Bacterial growth was sustained in Luria-Bertani (LB) broth containing 1% (w/v) Bio-tryptone, 0.5 % (w/v) yeast extract and 1 % (w/v) NaCl in double distilled water and supplemented with 100 μg/mL ampicillin (Amp).

**2.2 Bacterial Transformation**

Protein expression was performed by amplifying plasmid in *E. coli* BL21-Gold (DE3) competent cells as described by Sambrook and Russell (82) with some

modifications. Plasmid DNA (2-5 µL, 50-100 ng) was mixed with 50 µL of competent cells. The mixture was incubated on ice for 30 min followed by heat-shock (42°C for 45 s), then placed on ice for 2 min. This entire mixture was then plated on LB/agar supplemented with 100 µg/mL Amp and grown for about 16 h at 37°C.

### 2.3 Expression and Purification of Full-length and Δ80 Variant of *H. influenzae* CM-PD

Proteins were purified based on the procedure of Sun *et al.* (63) with some modifications. Following transformation of pPKQ plasmids into BL21 (DE3) Gold cells, a single colony was selected and grown in 50 mL of LB/Amp by incubating at 37°C for 16 h with agitation at 250 rpm. Ten mL of this culture was used to inoculate 2 × 1 L of the same medium which was then incubated at 37°C with shaking until bacterial growth reached logarithmic phase ($OD_{600}$ 0.5-0.9). It was observed that cultures induced at early log phase ($OD_{600}$ 0.5-0.6) yielded higher protein per gram of wet cell pellet than those induced in late log phase ($OD_{600}$ 0.9-1.0). Protein expression was induced by the addition of IPTG to a final concentration of 0.5 mM. The induced culture was incubated with shaking at 18°C for 16 h. The cells were collected by centrifugation at 7000 × g for 15 min at 4°C. When samples were processed immediately, cell pellets were weighed and frozen at -80°C for 20 min, otherwise the pellets could be stored at -20°C for up to two months. Frozen pelleted cells were thawed and resuspended in 5 mL/g (cell pellet) ice cold cell lysis buffer (50 mM potassium phosphate, 500 mM NaCl, 5mM imidazole, 10% glycerol, EDTA-free protease inhibitor cocktail, 1 µL/mL benzonase nuclease, 5 mM PMSF and 10 mM β-mercaptoethanol, pH 7.5). After ten up-and-down strokes of a

Dounce homogenizer, the sample was processed through two passes (20 mL per pass) of 18 000 psi in a pre-chilled Thermo Spectronic pressure cell. The cell lysate was separated into pellet and supernatant (cell-free extract) by centrifugation at 45 000 × g for 45 min at 4°C. All purification steps were performed in the cold room at 4°C with ice cold buffer. A column (1.2 cm × 30 cm) was packed with Ni-NTA SuperFlow resin (~13 mL) and equilibrated with 5 × column volumes of binding buffer (50 mM potassium phosphate, 500 mM NaCl, 5 mM imidazole, 10% glycerol, 10 mM β-mercaptoethanol, pH 7.5). To facilitate its flow rate during all chromatography steps, a Pharmacia P1 peristaltic pump was used. The supernatant of the cell lysate was applied at a flow rate of 1 mL/min and the immediate flow-through (FT) was collected. Next, ~500 mL of wash buffer (binding buffer but with 30 mM imidazole) was applied at a flow-rate of 1 mL/min for the first 10 mL, 3 mL/min for next 10 mL, and then the reminder at 4 mL/min. The eluant from the washing step was collected in 50 mL fractions and 1 mL aliquots were checked spectrophotometically ($OD_{280}$) for protein content. Washing was continued until the $OD_{280}$ of the eluant fractions was less than 0.01. His-tagged CM-PD was eluted from Ni-NTA resin using elution buffer (binding buffer but with 400 mM imidazole) at a flow-rate of 1 mL/min. One mL fractions were collected in microfuge tubes containing 1 µL of 1 M DTT and 10 µL of 0.1 M EDTA. Fractions containing the highest PD specific activity were pooled. Thrombin (1.4 units/ mg protein) was added to the pooled protein which was then dialyzed overnight at 4ºC in 4 L of 50 mM potassium phosphate, 500 mM NaCl, 10% glycerol, 1 mM EDTA, 1 mM DTT, 21 mM calcium citrate and 2.5 mM $CaCl_2$, pH 7.5. SDS-PAGE and ESI-MS were used to confirm thrombin cleavage of the His tag. Glycerol (15% v/v) and DTT (10 mM) were added to the sample which was

stored in 400 µL aliquots (4 – 6 mg/mL) at -80°C. The yield of homogeneous Δ80 CM-PD (80-100 mg per 2 L culture) was generally ~15% higher than that of FL CM-PD. Both TyrA proteins retained excellent PD activity and behavior in cold storage (-80°C) up to a year.

**2.4 Protein Concentration Determination**

The Bio-Rad protein assay kit (Bio-Rad Laboratory) and bovine serum albumin (BSA, Sigma) as a standard were used to determine the protein concentration. The stock of BSA standard was prepared in 10 mM Tris-HCl, pH 7.4 and filtered with 0.2 µm syringe filters. BSA standard concentration was determined by reading $OD_{280}$ and using an extinction coefficient of 0.667 mL/mg/cm (83). Protein content also was determined from $OD_{280}$ measurements using the extinction coefficients predicted from the amino acid sequences of FL and Δ80 CM-PD (41370 and 39880 $M^{-1}cm^{-1}$, respectively) using PeptideMass Tool available on the ExPaSy Server (84).

**2.5 ESI-MS of CM-PD**

As a routine analysis, *H. influenzae* Tyr A protein masses from each preparation were confirmed by ESI-ToF-MS. Each protein sample was prepared as described in Millipore's Zip-Tip manual with a minor modification outlined by Quashie (49). The bed resin of the $C_4$ Zip-Tip was rinsed three times using 100% acetonitrile (ACN), then three times with 5% ACN + 0.1% Formic Acid (FA). This procedure involved pipetting up-and-down with the pipette volume set to 10 µL. Before loading on a $C_4$ Zip-Tip, the

protein sample (20 - 40 μL) was acidified in order to increase binding to the resin by adding 50% FA to be 0.1% of total volume. If precipitation was observed the FA volume was decreased. Binding of the pre-acidified protein to the washed Zip-Tip was achieved by pipetting up-and-down 10 times through the Zip-Tip. Final washing of the Zip-Tip was performed by rinsing with 5 % ACN + 0.1% FA three times. Protein was then eluted from the Zip-Tip by rinsing ten times with 40 μL of 50% ACN + 50% MilliQ water + 0.1 % FA. Binding, washing and elution steps were repeated two times to increase the concentration of protein for better spectral data.

Protein samples were loaded into a Waters Micromass Q-ToF-II mass spectrometer by direct injection with a 50 μL syringe at a constant flow rate of 1.0 μL/min. Samples were analyzed in positive ion mode and in a mass range (*m/z*) of 400 to 3000 using the following instrument's parameters: capillary voltage of 3500 V, cone voltage 45 V, multiplier 550 V, MCP 1800 V, resolution of 8000. The instrument was calibrated using myoglobin. MassLynx 4.0 software (Waters Micromass) was used for data analysis and the deconvolution of charge envelopes.

**2.6 Polyacrylamide Gel Electrophoresis**

Protein purity was evaluated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), using a 4% stacking gel (pH 6.8) and a 12% resolving gel (pH 8.3) as described by Sambrook and Russell (82). Acrylamide (29% w/v) and N, N'-methylene-bis-acrylamide (1% w/v) were mixed to prepare a 30% acrylamide stock solution. Tris buffers, 1.5 M at pH 8.3 and 0.5 M at pH 6.8 (pH adjusted with HCl), were prepared as stock solutions. The acrylamide and Tris buffer solutions were kept at 4 $^{\circ}$C.

Ammonium persulfate (10% w/v) was stored at -20°C in 200 µL aliquots while SDS (10 % w/v) was stored at room temperature. Electrophoresis buffer (25 mM Tris, 250 mM glycine and 0.1% SDS, pH 8.3) was used as an electrolyte. Protein samples were prepared by diluting with 1:1 (v/v) 2 × SDS gel-loading buffer (1.5 M Tris-HCl, 4% SDS, 20% glycerol, 0.2 M DTT and a small amount of bromophenol blue), and then boiled for 10 min along with unstained protein molecular weight marker (Fermentas) before loading on the gel. Electrophoresis was performed at 60 V as the protein migrated through the stacking gel and increased to 150V using a Mini-PROTEAN® tetra Electrophoresis System from Bio-Rad. Proteins were visualized by incubating the gel in a staining solution of 0.25% (w/v) Brilliant blue R (Sigma), 40% (v/v) methanol and 7% (v/v) acetic acid followed by incubating overnight in a destaining solution of 10:9:1 (v/v/v) methanol: distilled water: glacial acetic acid.

**2.7 Determination of Enzyme Activities and Kinetic Parameters**

The conversion of chorismate to prephenate catalyzed by chorismate mutase (CM) was monitored by following a decrease in absorbance at 273 nm. The oxidative decarboxylation of prephenate in the presence of $NAD^+$ catalyzed by prephenate dehydrogenase was monitored by the increase in absorbance at 340 nm as described by Turnbull and Morrison (40). All assays were conducted at 30°C in a Varian Cary 50 spectrophotometer equipped with a thermostat water-jacket cell holder using a 1 cm path-length quartz cuvette (for CM activity) or a PMMA semi-micro 1.5 mL Plastibrand® cuvette (for PD activity) in a total reaction volume of 1 mL. Changes in the absorbance were recorded continuously for 3-5 min after the addition of enzyme. The initial reaction

rates were calculated using Cary WinUV kinetics software from the linear portions of the progress curves. For the determinations of the kinetic parameters of the reactions, the assays were conducted in a three-component buffer (3CB) consisting of 0.1 M MES, 0.1 M 4-ethylmorpholine, 51 mM diethanolamine, pH 7.2 supplemented with 1 mM EDTA and 1 mM DTT.

To monitor CM and PD activities during protein preparation, 10 µL of protein-containing fractions were added to 990 µL of CM or PD reaction mix. CM reaction mix consisted of 100 mM Tris-HCl (pH 7.5), 1 mM EDTA, 1 mM DTT and 0.25 mM chorismate, while PD mix contained 100 mM Tris-HCl, (pH 7.5), 1 mM EDTA, 1 mM DTT, 2 mM $NAD^+$ and 0.25 mM prephenate. In all assays, the reaction mix or buffer was incubated at 30°C for 1 min in the cuvette, and then substrate was added (if required) and the reaction initiated with enzyme.

The values of $K_m$ and $k_{cat}$ of full-length and Δ80 CM-PD were calculated by fitting initial velocity data from substrate saturation curves to the Michaelis-Menten equation ($V_0 = (V_{max}[S])/([S] + K_m)$ using GraFit v 5.0 (Erithacus Software). A unit of activity was defined as the amount of enzyme required to produce 1 µmol product/ min at 30°C. Calculations for units of activity, specific activity and $k_{cat}$ are indicated:

Units of CM activity (µmol/min/mL):

$$\text{Units (µmol/min/mL)} = \frac{\Delta OD_{273}/\text{min}}{1\text{cm x } 2630\ M^{-1}\ cm^{-1}} \text{ x } \frac{10^6\ \mu\text{mol}}{\text{mol}} \text{ x } \frac{1\text{L}}{10^3\ \text{mL}} \text{ x dilution factor}$$

Units of PD Activity (µmol/min/mL):

$$\text{Units (µmol/min/mL)} = \frac{\Delta OD_{340}/\text{min}}{1\text{cm x } 6400\ M^{-1}\ cm^{-1}} \text{ x } \frac{10^6\ \mu\text{mol}}{\text{mol}} \text{ x } \frac{1\text{L}}{10^3\ \text{mL}} \text{ x dilution factor}$$

Specific Activity (μmol/min/mg) = $\frac{\text{Units}}{\text{mg/mL protein}}$

$k_{cat}$ ($s^{-1}$) = $V_{max}$ (μmol/sec/mg) x $\frac{\text{mg}*}{\text{mmoL active site}}$ x $\frac{1\text{mmol}}{10^3\ \mu\text{moL}}$

*Amount of protein/mol active site (one active site/monomer) varied with the primary sequence of the protein and the 3 residues (G,S,H) added from protein expression in pET-15b following thrombin cleavage. Exact masses in g/mol were 43303.7 (FL CM-PD) and 34088.0 (Δ80 CM-PD) as determined by ESI-MS.

## 2.8 Far-UV Circular Dichroism Spectroscopy

Far-UV circular dichroism (CD) studies were performed using a Jasco-815 spectropolarimeter connected to a Pelletier heating/cooling temperature control system. Proteins were dialyzed or buffer exchanged (NAP-5 columns) into phosphate buffered saline (PBS: 137 mM NaCl, 27 mM KCl, 10 mM $Na_2HPO_4$, 2 mM $KH_2PO_4$, pH 7.4) supplemented with 10% glycerol and diluted to a final concentration of 3 – 5 μM monomer. Spectra were recorded at 25°C by averaging the accumulations of five scans from 240 nm to 200 nm (bandwidth of 1 nm) using a scan rate of 20 nm/min with 0.2 cm path-length quartz cuvette. Other parameters were 0.2 nm step resolution, 0.5 sec response time and a sensitivity of 100 mdeg. The spectra were corrected for the contribution of the buffer to the absorbance.

## 2.9 Chemical Denaturation Probed by Far-UV CD Spectroscopy

Chemical denaturation of TyrA proteins with guanidine hydrochloride (Gdn-HCl) was performed by taking spectral scans from 200 nm to 240 nm at 25°C, and the average of five scans (scan rate of 50 nm/min and other instrument parameters were reported in section 2.8). Protein samples were prepared as described as in section 2.8 in

PBS/glycerol or in 50 mM $K_2HPO_4$ with the appropriate volume of stock Gdn-HCl (7.2 M) to have increasing 0.3 M increments of denaturant in each sample. Gdn-HCl stock was made in either PBS/glycerol or in 50 mM $K_2HPO_4$ and its molar concentration was determined using the measured refractive index and the relationship below (85, 86):

$$57.147(\Delta N)+38.68(\Delta N)^2-91.60(\Delta N)^3$$

where $\Delta N$ was the difference in refractive index between the denaturant solution and buffer at the sodium D line. A drop of denaturant solution or buffer is placed on a refractometer (Concordia undergraduate teaching lab) and the compensating drum is rotated until the borderline is sharp and achromatic. The refractive index of denaturant solution or buffer is read from the display window. A total of 17 samples (600 µL) containing *H. influenzae* or *E. coli* CM-PD proteins (3.5 µM to 5 µM), Gdn-HCl (final concentration from 0 M to 6 M) and PBS/10% glycerol or 50 mM $K_2HPO_4$ were prepared. Samples in PBS/glycerol were incubated for 4 h at room temperature before measurements were taken while those in 50 mM $K_2HPO_4$ were incubated overnight at room temperature. All the spectra were corrected for the absorbance of buffer and Gdn-HCl. The resulting data were plotted and manipulated using Microsoft Excel.

**2.10 Thermal Denaturation Probed by Far-UV CD Spectroscopy**

For variable temperature (VT) experiments performed with a 0.2 cm path-length cell, changes in ellipticity at 222 nm were recorded from 25°C to 95°C by using a Pelletier temperature controlling system interfaced with the Jasco 815 spectropolarimeter's software temperature ramping program. Protein samples were prepared in PBS/glycerol

at a concentration of 3.5-5 μM. Parameters used were: ΔT of 20°C/h, a 0.2°C step resolution, and a 0.5 sec response time. At the beginning and at the end of the VT experiment (25°C and 95°C, respectively), spectral scans from 200 - 240 nm were recorded. Melting temperature ($T_m$) was determined by using the first derivative processing function supplied by the instrument's software.

For VT experiments performed with 6 or 12 μm pathlength cells and at a high (800-1500 μM) protein concentration in $D_2O$, an FTIR cell interfaced with its thermocouple temperature controlling system was used. Accordingly, the CD spectropolarimeter was adapted specially to accommodate the FTIR heating unit (Fig 10). The average of five spectral scans from 180 nm to 245 nm was recorded for every 2°C increase from 25°C to 95°C. Protein was prepared as described in section 2.12. The temperature was increased manually giving a 4 min equilibration time between readings. Other parameters were the same as outlined in section 2.8 except a scan rate of 50 nm/min was used. All the spectra were corrected for the contribution to the absorbance by the buffer. The content of secondary structure was predicted from the CD spectra (185 - 240 nm) using the CONTIN analysis program with a reference data set of three (88). This program was accessed through Dichroweb (70, 71).

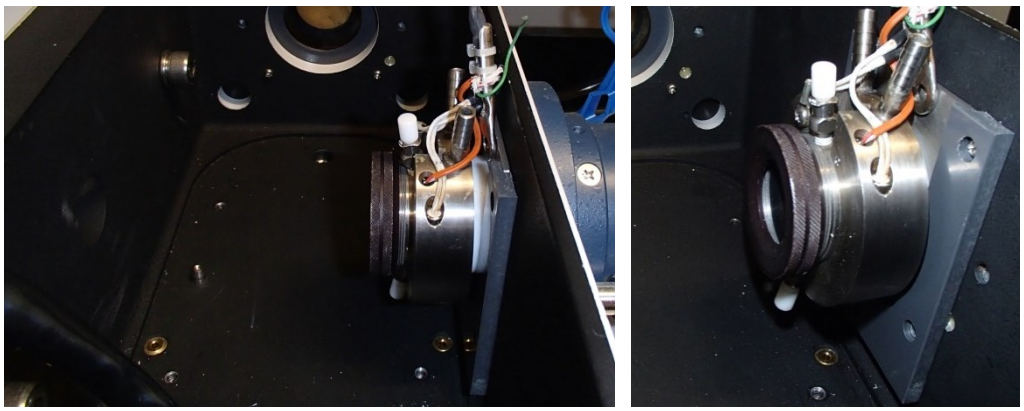


**Figure 10: Jasco-815 spectropolarimeter equipped with an FTIR cell.** A special casing was permanently mounted in the spectropolarimeter to accommodate the movable FTIR cell. The cell's gasket was interfaced to an externally supplied thermocouple temperature controlling system. The wires of the heating unit were fed through a port in the rear of the instrument at right-angles to the cell.

## 2.11 Fluorescence Spectroscopy

Steady-state fluorescence emission spectra of Tyr A proteins were recorded at 25°C using a Varian quartz cuvette (0.5 mL) with 1 cm pathlength in a Varian Cary Eclipse spectrofluorometer. For these experiments, protein samples were prepared as described in section 2.8, whereas the samples for chemical and thermal denaturation experiments (with 3.5-5 μM monomer) were prepared as described in sections 2.9 and 2.10. The excitation wavelength was set at 280 nm or 295 nm and the intrinsic fluorescence emission was recorded from 300 - 400 nm. Other parameters were: fast scan rate, sensitivity 600 V, 5 nm excitation and emission slit width, CAT mode to average 15 unsmoothed scans. All the spectra were corrected for the contribution of buffer and were exported into Microsoft Excel. Fluorescence emission spectra of the Tyr and Trp analogs, N-acetyl tyrosinamide (NAYA) and N-acetyl-L-tryptophanamide (NATA) were also measured to compare the changes in protein fluorescence emission (the $\lambda_{max}$ shift and intensity changes) under high temperature (95°C) and in the presence

of the denaturant, Gdn-HCl. Stock solutions of these analogs were prepared by multiplying the concentration of protein monomers by the number of Trp and Tyr in each protein (4 Trp and 13 Tyr in FL CM-PD and 4 Trp and 12 Tyr Δ80 CM-PD) in the same buffer as protein samples (PBS/glycerol or 50 mM $K_2HPO_4$). To observe the effect of high temperature on intensity of fluorescence emission intensity, NATA/NAYA samples were heated to 95°C prior to recording the spectra.

**2.12 Variable Temperature Fourier Transform Infra-Red Spectroscopy (FTIR)**

Fourier transform infrared (FTIR) spectroscopy was used to confirm/compare the protein secondary structure measured by CD. *H. influenzae* CM-PD proteins (3 - 5 mg, 1 - 1.5 mL) were dialyzed overnight in 2 L of 10 mM $K_2HPO_4$ buffer (pH 7.5) supplemented with 1 mM DTT. These samples were then placed in glass test tubes and lyophilized to dryness using a ThermoSavant Freeze Dryer. The resulting pellets were stored at room temperature (or -20°C for long term storage) to prevent condensation. This dry sample was weighed (0.5 mg to 1.1 mg) and reconstituted with 15 µL deuterium oxide ($D_2O$, Sigma-Aldrich) to yield a final protein concentration of ~30 to 70 mg/mL. The sample was gently vortexed to facilitate the dissolution of protein. The actual protein concentration was less than that calculated due to the contribution of buffer components (15% $K_2HPO_4$ and 1% DTT by weight per 1 mg of lyophilized sample). The protein concentration was estimated by reconstituting the same amount (in mg) of BSA powder (100% protein and no buffer) in $D_2O$ and measuring the ellipticity in the far-UV CD. For the same weight of protein, about 15-20% higher ellipticity at 222 nm was detected for BSA compared to the CM-PD sample. An 8 µL aliquot of the protein

sample was placed between $CaF_2$ windows fitted with a 50 μm spacer in a temperature controlled IR cell (Omega CN8500). Nicolet 550 IR spectrometer was interfaced with a DGS detector, connected to a Whatman air-purging system. VT experiments were performed in one cycle of heating and cooling from 25°C to 95°C with 2°C increments and 4 min equilibration time between each temperature. Experimental parameters were 128 scans with 4 $cm^{-1}$ resolution. Data were corrected for water and buffer contributions. The analysis of spectra was performed with OMNIC 7.3 software with the assistance of Dr. Ashraf Ismail from Department of Food Science, McGill University, Montréal. Fourier Self Deconvolution (FSD) was done at bandwidth of 19 – 28 and enhancement at 2.8. Different protein secondary structures were assigned from the amide I'components (87).

To ensure complete H-D exchange (monitored at the amide II' band of 1640 $cm^{-1}$), prior to VT experiments, samples were incubated at 25°C overnight (~20 to 24 h). The FTIR spectra of proteins were recorded every 4 min at 25°C to monitor the changes in amide II' band during the 24 h incubation.

**2.13 Two Dimensional (2D) Analysis of FTIR Data**

For further investigations of the temperature-mediated protein unfolding pathway, 2D correlation analysis was performed. 2D analysis (74) involves the interpretation of synchronous and asynchronous maps which were generated from 1D FTIR spectra using GRAMS Al spectroscopy software by Dr. Ismail. The synchronous map represents coupled intensity changes while an asynchronous map represents sequential or

unsynchronized changes in spectral intensities. The sequence of unfolding of a protein can be deduced by analyzing both maps using a set of rules (76). The rules to analyze the 2D correlation maps were described in section 1.13.

# CHAPTER 3: RESULTS

## 3.1 Purification of TyrA Proteins

Full-length (FL) and Δ80 CM-PD previously cloned into pET-15b were recombinantly expressed in *E. coli* BL 21 DE3 cells. After the induction of protein expression by IPTG and centrifugation of the culture, the cells were disrupted using high pressure treatment. TyrA was then purified using SuperFlow$^{TM}$ nickel NTA resin. The purified proteins contained a plasmid-encoded N-terminal hexa-His tag which was cleaved by thrombin. The purification was monitored by denaturing gel electrophoresis (Fig 11) and by PD activity assays.

Fig 11 shows that FL and Δ80 CM-PD proteins were well-expressed as judged by Coomassie Blue staining. Intense bands were observed at ~43000 Da (lane 2) and ~34000 Da (lane 3) corresponding to the migration of untagged monomers of FL and Δ80 CM-PD, respectively. Proteins appeared to be purified to near homogeneity with yields of ~40-50 mg/L of cell culture.

ESI-Q-ToF MS analysis was used to verify the masses of the purified TyrA proteins and showed values of 43305.5 and 34088.6 a.m.u for FL and Δ80 CM-PD, respectively (spectra not shown). These values were obtained using Masslynx software and were in agreement with the theoretical masses (FL CM-PD: 43303.7 a.m.u., Δ80 CM-PD: 34088.0 a.m.u.) calculated by PeptideMass Tool from ExPaSy server (82).

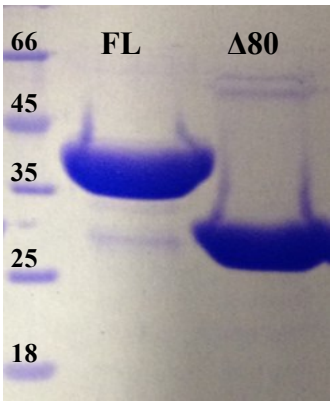



**Figure 11: SDS-PAGE analysis of purified *H. influenzae* TyrA proteins.** Lane 1: molecular weight (kDa) marker; Lane 2: FL CM-PD; Lane 3: Δ80 CM-PD

### 3.2 Determination of Steady-State Kinetics Parameters

The kinetic parameters of the mutase and dehydrogenase reactions catalyzed by *H. influenzae* TyrA were reported in order to determine the extent to which the deletion of the mutase domain affected the PD-catalyzed reaction. The steady-state kinetic parameters $k_{cat}$, $K_m$ and $k_{cat}/K_m$ were derived from the analysis of substrate saturation curves and the values are shown in Table 5. For comparison, WT *E. coli* CM-PD was also reported with the expectation that bifunctional TyrA proteins from *H. influenzae* and *E. coli* would yield similar results based on the high sequence conservation of proposed active site residues between the two enzymes (48).

Reaction rates were followed for the mutase-catalyzed reactions with chorismate as the variable substrate. As predicted, FL CM-PD from *H. influenzae* was an effective catalyst possessing values of kinetic parameters that were comparable to those of the *E. coli* enzyme (49). The Δ80 CM-PD possessed no mutase activity, consistent with the deletion of the enzyme's N-terminal domain. Steady-state kinetic parameters for the

dehydrogenase-catalyzed reactions were obtained by varing one substrate (prephenate or $NAD^+$) while keeping the other at fixed and saturating (or near saturating) concentrations.

Surprisingly, deletion of the mutase domain of *H. influenzae* CM-PD did not markedly alter the apparent binding of prephenate and $NAD^+$ or the enzyme's turnover number in the dehydrogenase reaction. In agreement with the results obtained by Quashie (49), values for the kinetic parameters for the reactions catalyzed by the two TyrA forms differed by less than a factor of two. Additionally, there was good agreement between the kinetic parameters of the reactions catalyzed by *H. influenzae* and *E. coli* CM-PDs.

## 3.3 Native Structure of *H. influenzae* TyrA

### 3.3.1 Analysis of Secondary Structure of TyrA by Far-UV CD Spectroscopy

Far-UV CD spectroscopy measurements were recorded to determine if deletion of the mutase domain had altered the global secondary structure of *H. influenzae* CM-PD. Low concentrations of protein were used (0.1 mg/mL or ~3.5 μM monomers) to allow comparison of the results obtained by far-UV CD and intrinsic fluorescence emission which monitors tertiary structural changes (presented in section 3.3.2).

Fig 12A compares the molar ellipticities of *H. influenzae* FL and Δ80 CM-PDs recorded from 200-240 nm at 25°C. In units of molar ellipticity ($\theta_M$), the CD signal is independent of protein concentration but dependent on protein mass. In the far-UV CD, the α–helix yields the most intense and characteristic signal (68, 69).

**Table 5. Summary of kinetic parameters for the reactions catalyzed by TyrA proteins from *H. influenzae* and *E. coli*.**

| CM-PD form | Mutase | | | Dehydrogenase | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Chorismate | | | Prephenate | | | $NAD^+$ | | |
| | $K_m$ (μM) | $k_{cat}$ ($s^{-1}$) | $k_{cat}/K_m$ ($M^{-1}s^{-1}$) | $K_m$ (μM) | $k_{cat}$ ($s^{-1}$) | $k_{cat}/K_m$ ($M^{-1}s^{-1}$) | $K_m$ (μM) | $k_{cat}$ ($s^{-1}$) | $k_{cat}/K_m$ ($M^{-1}s^{-1}$) |
| *E.coli*[49] | 44.0 ± 11.3 | 10.7 ± 0.6 | $2.4 \times 10^5$ | 43.5 ± 9.2 | 30.4 ± 1.4 | $7.0 \times 10^5$ | 195 ± 38 | 30.4 ± 1.4 | $1.6 \times 10^5$ |
| FL | 68.5 ± 8.7 | 31.0 ± 3.2 | $3.6 \times 10^5$ | 79.8 ± 20.6 | 53.7 ± 3.2 | $6.7 \times 10^5$ | 105.7 ± 21.8 | 34.2 ± 1.8 | $3.2 \times 10^5$ |
| Δ80 | ND | ND | ND | 98.0 ± 39.1 | 20.0 ± 1.5 | $2.3 \times 10^5$ | 233.8 ± 42.9 | 26.1 ± 2.3 | $1.1 \times 10^5$ |

Initial rates for the mutase-catalyzed reactions were recorded using chorismate concentrations between 20–700 μM. Kinetic parameters for PD were determined from substrate saturation curves, varying prephenate from 20-700 μM at 2 mM $NAD^+$, and varying $NAD^+$ from 20-2000 μM keeping prephenate constant at 1.5 mM. Kinetic parameters for both reactions were obtained by fitting data from substrate saturation via non-linear regression analysis to the equation for a rectangular hyperbola using GraFit 4.0. Kinetic parameters for *E. coli* CM-PD were from Quashie (49). ND denotes activity not detected. All enzymatic assays were carried out at 30°C and pH 7.2 as described in section 2.7, using ~3 μg protein per reaction.

The spectra revealed a double minimum at 208 nm and 222 nm indicating that both proteins possessed a high degree of α-helical structure. As expected, the more intense signal of FL CM-PD reflects the contribution of the highly helical 80-residue mutase domain (section 1.7) to the global secondary structure of the bifunctional protein. However, when the data were reported as molar ellipticity per amide bond or mean residue weight ellipticity (Fig 12B), the spectra of two proteins were almost superimposable. Thus, per residue, FL and Δ80 CM-PD have similar α-helical content. The slightly stronger signal at 222 nm observed for the FL protein (Fig 12B) is consistent with the influence of the helical mutase domain averaged globally per residue.

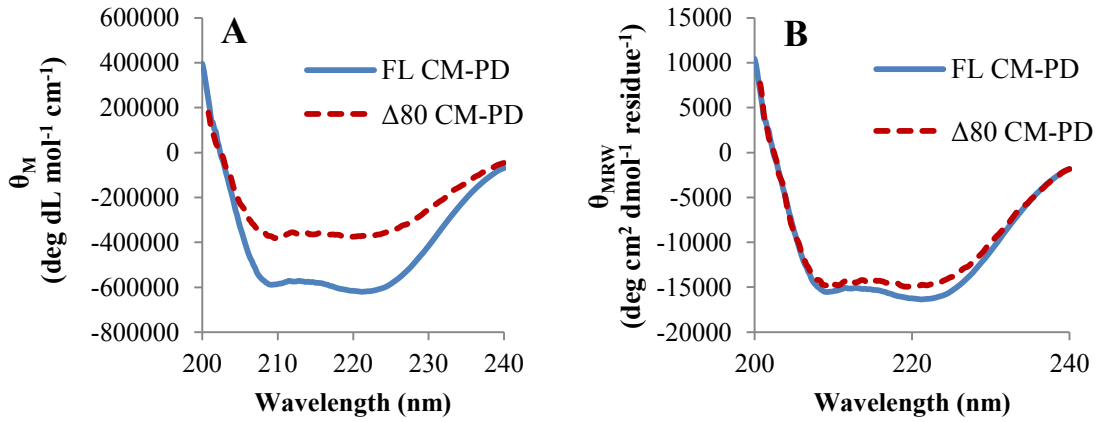



**Figure 12: Far-UV CD spectra of *H. influenzae* FL and Δ80 CM-PD at 25°C.** Data were expressed as (A) $\theta_M$, molar ellipticity or (B) $\theta_{MRW}$, mean residue weight ellipticity. Proteins (0.1 mg/mL) were reconstituted in PBS (pH 7.5) containing 10% glycerol and measurement were recorded on a Jasco 815 Spectropolarimeter at 25°C using a 0.2 cm pathlength rectangular cell. The curves are the result of the average of 5 scans performed at a rate of 20 nm/min. Ellipticity in milideg, was converted into molar ellipticity ($\theta_M$) using the relationship: $\theta_M = 100 \cdot \theta / m \cdot l$ where $\theta$ was the observed ellipticity in degree, $m$ was the molar protein concentration, and $l$ was the pathlength of the cell in cm (68). Mean residue weight ellipticity ($\theta_{MRW}$) was calculated using the relationships: $\theta_{MRW} = \theta \cdot MRW / 10 \cdot c \cdot l$ where $\theta$ was the observed ellipticity in degree, $MRW$ was the mean residue weight, $c$ was the concentration in g/L and $l$ was the pathlength of the cell in cm. $MRW$ can be calculated from the equation: MRW = $M/(N-1)$ where $M$ was the molar mass of the protein in Daltons and $N$ was the number of residues in the polypeptide chain of the protein (68).

### 3.3.2 Analysis of Secondary Structure of TyrA by FTIR Spectroscopy

Further analysis of the secondary structure of *H. influenzae* FL and Δ80 CM-PDs were conducted using Fourier transform infrared spectroscopy. Fig 13 shows the FTIR spectra of FL and Δ80 CM-PD at 25°C immediately after resuspension of the dried protein samples in deuterated buffer. As anticipated, the FL protein showed a slightly higher absorbance at the α-helix frequencies (b, 1650cm$^{-1}$) compared to Δ80 CM-PD, while absorbances for intramolecular β-sheets (c, 1630cm$^{-1}$) for both proteins were identical. These results reflect the additional helical structure contributed by the CM domain in the FL protein which also is observed by far-UV CD (Fig 13), whereas both FL and Δ80 CM-PD contain the β-sheet rich PD domain. The low absorbance for the amide II' band (e, 1540cm$^{-1}$) immediately after the dissolution of proteins in $D_2O$ indicated a fast exchange of N-H to N-D and suggested that neither protein was densely compact; a significant number of polypeptide amide groups were solvent accessible.

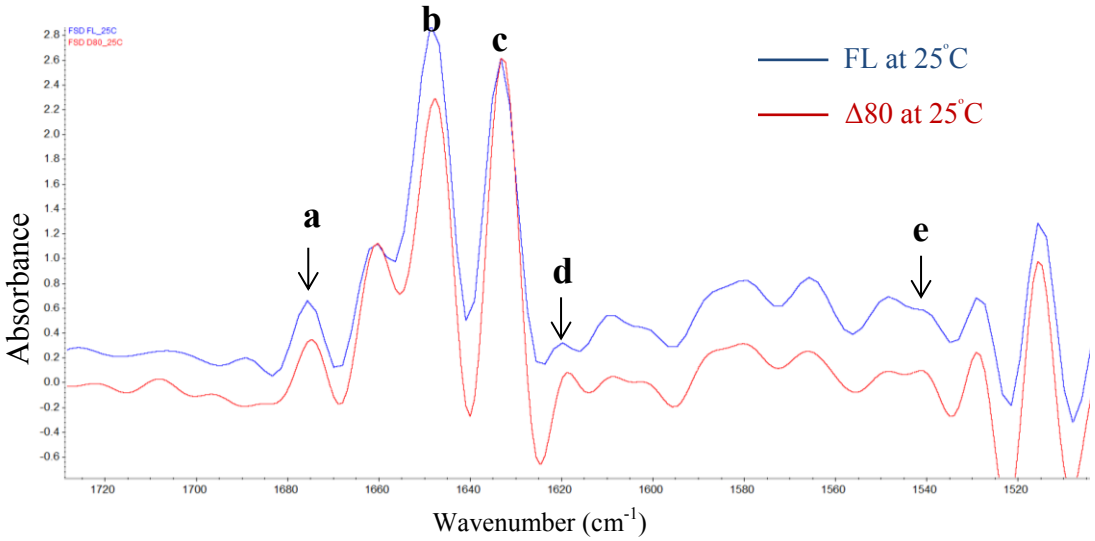



**Figure 13: FTIR spectral analysis of FL and Δ80 CM-PD at 25°C.** Proteins were dialyzed in 10 mM $K_2HPO_4$ buffer supplemented with 1 mM DTT overnight. About 0.5 - 1 mg of lyophilized dried powder (0.5 mg and 1.1 mg for FL and Δ80 CM-PD, respectively) was resuspended in 15 µL $D_2O$. Protein concentrations were ~30 mg/mL (800 µM) and 70 mg/mL (1500 µM) for FL and Δ80 CM-PD, respectively. Measurements were carried out on a Nicolet FTIR Spectrometer scanning from 4000 - 400 $cm^{-1}$ as described in section 2.12. Each spectrum represents the average of 128 scans. After baseline correction, Fourier self-deconvolution was performed on each spectra using OMNIC software. Peak annotations are as follows; (a) high frequency intermolecular β-sheets (~1680 $cm^{-1}$), (b) α-helices (~1650 $cm^{-1}$), (c) intramolecular β-sheets (~1630 $cm^{-1}$), (d) low frequency β-sheets (~1618 $cm^{-1}$), (e) amide II' band (~1540 $cm^{-1}$).

### 3.3.3 Analysis of Tertiary Structure of TyrA by Intrinsic Fluorescence Emission

The environment surrounding intrinsic fluorophores in proteins (Tyr, Trp) were used as a probe of the tertiary structure of *H. influenzae* TyrA proteins. The fluorescence emission spectra of FL, Δ80 CM-PD and the Trp and Tyr analogs NATA and NAYA are shown in Fig 14. After excitation at 280 and 295 nm, both TyrA proteins showed a $\lambda_{max}$ emission of about 332 nm indicating that Trp residues in the folded protein were partially buried. In contrast, the fully solvent exposed fluorophores in the Trp and Tyr analogs NATA and NAYA displayed $\lambda_{max}$ emission of about 355 nm. At $\lambda_{ex}$ 280 nm, the emission intensity of FL protein appeared slightly elevated compared to Δ80 CM-PD

(panel A), likely due to an additional Tyr per monomer of the bifunctional protein; at $\lambda_{ex}$ 280 nm, both Trp and Try were excited. The emission spectrum at $\lambda_{ex}$ 295 nm was recorded to focus on the contribution to the emission by Trp only (panel B). The emission spectra for both FL and Δ80 CM-PD proteins appeared to be identical.

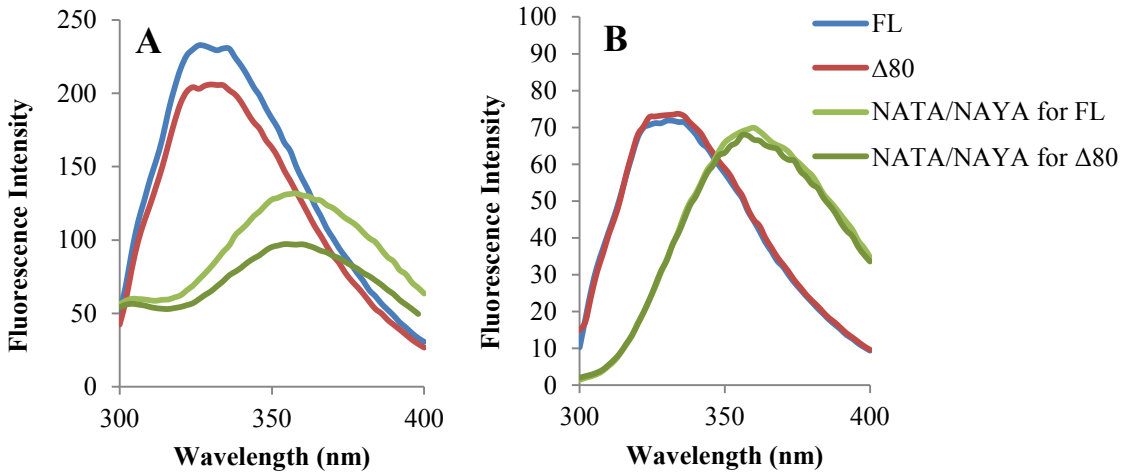



**Figure 14: Intrinsic fluorescence emission spectra of FL and Δ80 CM-PD.** Excitation wavelength at (A) 280 nm and (B) 295 nm. Proteins (0.1 mg/mL, 3.5 µM monomer) were reconstituted in PBS, pH 7.5 supplemented with 10% glycerol. Trp and Tyr analogs (NATA and NAYA) were prepared as same concentration as Trp and Tyr in protein (14 µM NATA/ 45.5 µM NAYA for FL CM-PD and 14 µM NATA/ 42 µM NAYA for Δ80 CM-PD) in the same buffer as TyrA proteins. Emission spectra were recorded at 25°C using excitation and emission slit width of 5 nm and were result of 15 scans at fast scan speed. Contribution to the signal from buffer was subtracted.

## 3.4 Chemical Stability Studies of *H. influenzae* TyrA

### 3.4.1 Analysis of Guanidine-Induced Unfolding of TyrA by Far-UV CD Spectroscopy

Since there was no significant difference in the global secondary and tertiary structures of FL and Δ80 CM-PDs, chemical denaturation experiments were conducted to

probe any differences in the stability of the proteins. Guanidine hydrochloride (Gdn-HCl) was chosen as a denaturant for this study, since it was previously reported that *E. coli* CM-PD showed incomplete unfolding in the presence of 8M urea (45). The protein concentrations of 0.1 mg/mL (3.5 μM) monomers were used in order to draw comparisons between the results obtained by far-UV CD and fluorescence emission.

Preliminary studies were conducted to determine an appropriate incubation time with the Gdn-HCl to ensure the unfolding reaction has reached equilibrium. Protein samples were incubated for 0 h, 1 h, 4 h and 24 h at room temperature in 0 M, 3 M, 4 M, and 5.3 M Gdn-HCl in PBS/glycerol buffer. Far-UV CD scans from 200-245 nm showed that ellipticity measurements at 222 nm were identical at all time points after 1 h (data not shown). For convenience, an incubation time of 4 h was selected for further studies.

Fig 15 shows that after 4 h of incubation in PBS/glycerol with increasing Gdn-HCl concentration, both FL (panel A) and Δ80 CM-PD (panel B) lose intensity at 208 nm and 222 nm, ellipticities which are characteristic of α-helices. While the loss of helical signal was generally inversely proportional to increasing amounts of Gdn-HCl, surprisingly for FL CM-PD (panel A), the ellipticity at 1.8 M Gdn-HCl was almost superimposable with that at 3.0 M Gdn-HCl. This unexpected result at 1.8 M Gdn-HCl was not observed for Δ80 CM-PD (panel B) or for a comparator protein bifunctional *E. coli* CM-PD (panel C).

Previous chemical denaturation studies of *E. coli* CM-PD by Christendat and Turnbull (45) had been conducted after incubating the enzyme for 18 h in 50 mM $K_2HPO_4$ buffer containing increasing Gdn-HCl concentrations. Again, for comparative

purposes the unfolding of the TyrA proteins were monitored under these same conditions and the results are shown in Fig 15 D and E for *H. influenzae* FL and Δ80 CM-PD, respectively (*E. coli* CM-PD data not shown). The far-UV CD spectra for FL and Δ80 CM-PD showed decreases in ellipticity as Gdn-HCl concentrations increased. However, for FL protein (panel D) at 1.5, 1.8 and 2.4 M Gdn-HCl, the measurements deviated from the predictable trend and were nearly superimposable with the spectrum at 6.0 M Gdn-HCl. At these lower concentrations of Gdn-HCl, the sample for FL protein was clearly turbid. In contrast, the spectra for Δ80 CM-PD (panel E) did not show such irregular changes in ellipticity with the increasing concentrations of Gdn-HCl and no visible precipitation was noted in these samples.

The spectra from Fig 15 were re-plotted in Fig 16 to show only ellipticity changes at 222 nm.

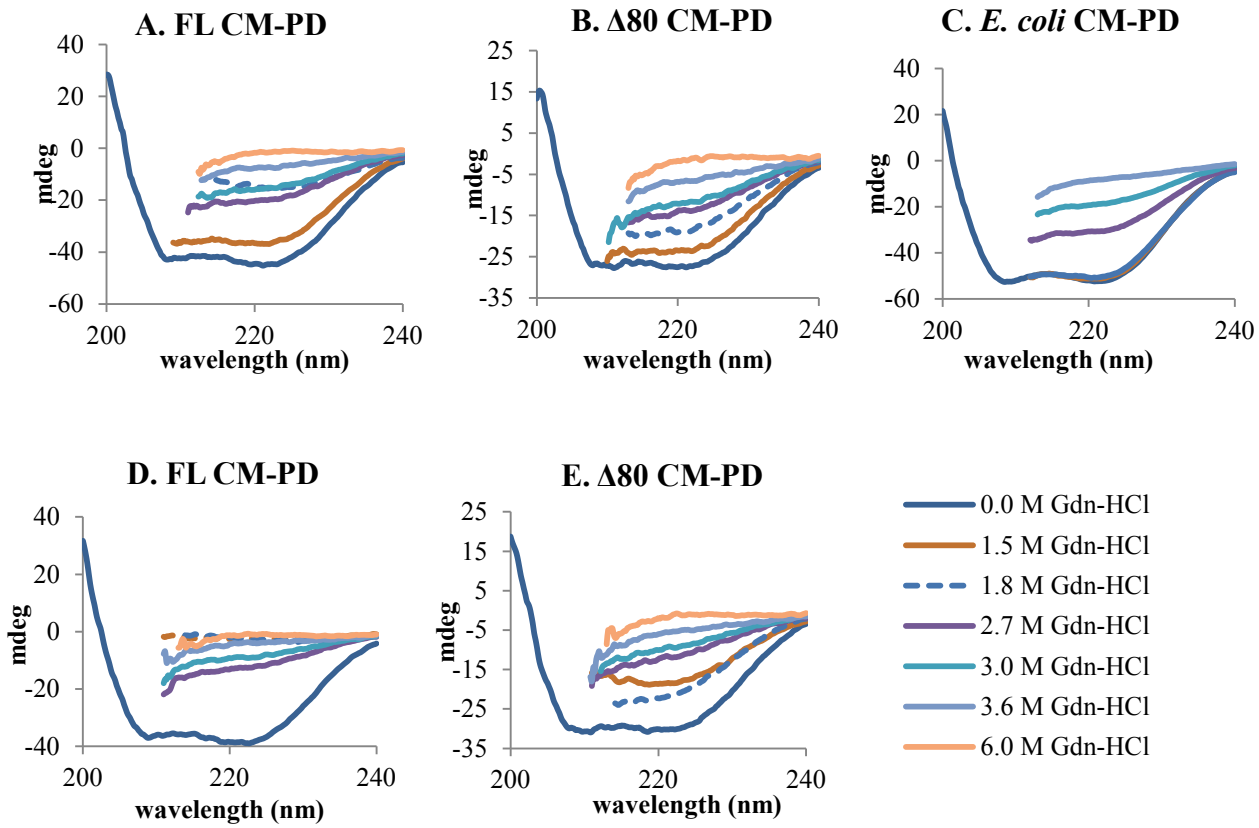



**Figure 15: Far-UV CD spectra of *H. influenzae* TyrA in the presence of increasing concentrations of Gdn-HCl**. Proteins (0.1 mg/mL) were reconstituted in a buffer of PBS/glycerol (pH 7.5) containing Gdn-HCl and incubated for 4 h (A, B, and C). For D and E, samples were reconstituted in 50 mM $K_2HPO_4$ containing Gdn-HCl and incubated for 18 h. Experimental parameters are identical as in Fig 12. Spectra were scanned at a rate of 50 nm /min. A wide range of concentrations of Gdn-HCl were recorded (18 samples) but only representative curves at selected Gdn-HCl concentrations are shown.

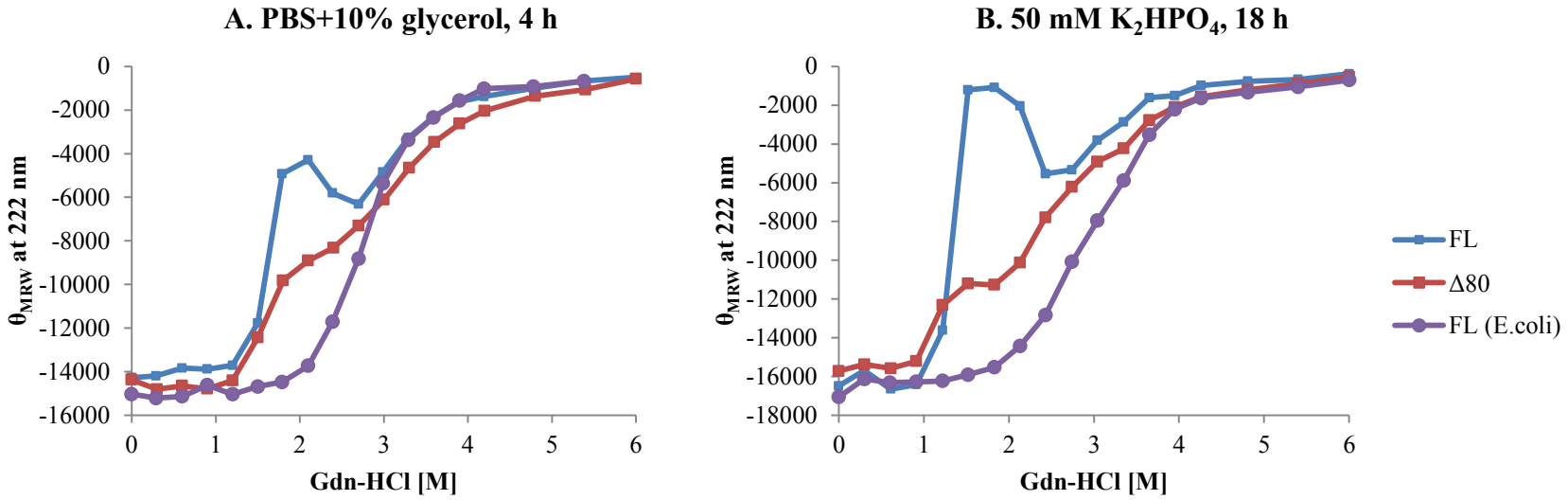



**Figure 16: Gdn-HCl denaturation of *H. influenzae* and *E. coli* TyrA proteins monitored by far-UV CD at 222 nm as a function of Gdn-HCl concentration**. The spectra of TyrA proteins from Fig 15 were re-plotted to show changes in MRW ellipticies at 222 nm. (A) *H. influenzae* and *E. coli* TyrA proteins were reconstituted in a buffer of PBS/glycerol containing Gdn-HCl and incubated for 4 h. (B) *H. influenzae* and *E. coli* TyrA proteins were reconstituted in 50 mM $K_2HPO_4$ containing Gdn-HCl and incubated for 18 h. The dramatic decrease in MRW 222 nm for FL CM-PD at 1.5 M, 1.8 M, 2.1 M and 2.4 M Gdn-HCl in 50 mM $K_2HPO_4$ after incubation for 18 h, were attributed to visible protein precipitation. No visible precipitation was observed at the higher concentrations of Gdn-HCl.

This replot (Fig 16) more clearly shows that the irregular decreases in mean residue weight ellipticity at 222 nm are most prominent for *H. influenzae* FL CM-PD with 1.5 M, 1.8 M, 2.1 M and 2.4 M Gdn-HCl in PBS/glycerol and $K_2HPO_4$ buffers. In both buffers, the spectra for Δ80 CM-PD showed a plateau from 1.8 M to 2.1 M Gdn-HCl indicative of an intermediate structure in the unfolding. In contrast, *E. coli* CM-PD (panel A) showed a sigmoidal curve indicating a cooperative denaturation after 4 h incubation in PBS /glycerol containing Gdn-HCl. This curve is in good agreement with that reported by Christendat and Turnbull (45) in 50 mM $K_2HPO_4$ after 18 h incubation and in the present study (panel B).

We further explored the possibility that any intermediate states in the unfolding of *H. influenzae* TyrA might be due to precipitated/aggregated proteins. Accordingly, UV-visible absorbance scans from 200-500 nm were recorded for protein samples incubated at different concentrations of Gdn-HCl (Fig 17). The elevated absorbance readings for FL CM-PD at 1.5, 1.8, 2.1 and 2.4 M Gdn-HCl (panel A), and Δ80 CM-PD at 1.8 M Gdn-HCl (panel B) are indicative of light scattering caused by protein precipitation.

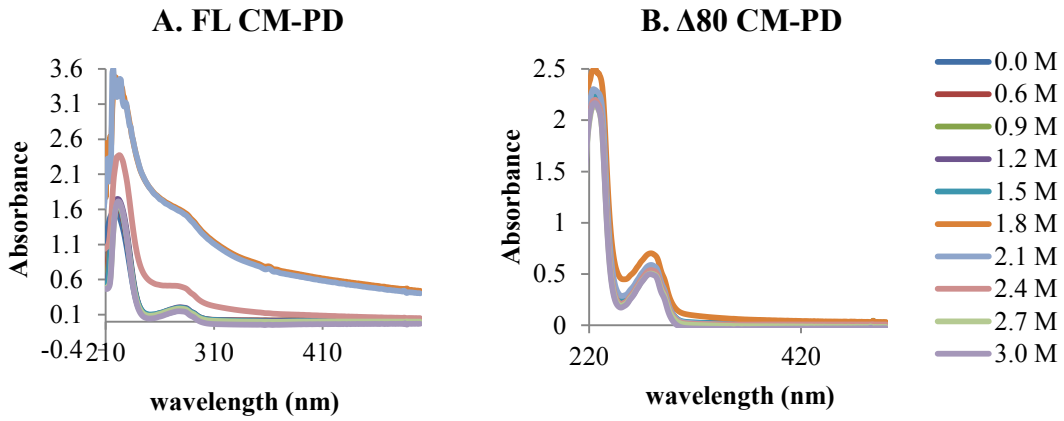



**Figure 17: UV-Vis spectral scans of *H. influenzae* TyrA in the presence of Gdn-HCl.** Proteins were incubated for 4 h in PBS/glycerol containing various concentration of Gdn-HCl (listed above). UV-Vis absorbance spectra from 200 to 500 nm were recorded using an Agilent 8453 spectrophotometer, with an average of 150 scans at a fast scan rate at room temperature using a 1 cm pathlength cuvette. The contribution by buffer was subtracted.

### 3.4.2 Time-Dependent Denaturation of FL CM-PD by Gdn-HCl

To further examine the time-dependent denaturation of *H. influenzae* CM-PD, protein was incubated from 0-24 h in PBS/glycerol containing 0, 0.6, 1.5, 2.5, 3.6, and 5.4 M of Gdn-HCl at 25°C.

Changes in ellipticity at 222 nm were monitored by far-UV CD spectroscopy (Fig 18). At 0.6, 3.6, and 5.4 M Gdn-HCl, there were little or no change in the measurements as a function of time, confirming that under these conditions the equilibrium of unfolding was rapidly achieved. At 1.5 and 2.5 M Gdn-HCl, however, the ellipticity at 222 nm gradually decreased as the incubation time increased from 0 to 14 h after which time the ellipticity changes remained constant. The data suggested that at these Gdn-HCl concentrations the unfolding reaction had not reached equilibrium until 14 h. However, it

was noted that these samples became more turbid at the longer incubation times. We concluded that Gdn-HCl at 1.5 and 2.4 M could catalyze the formation of an insoluble aggregate of TyrA in a time-dependent fashion.

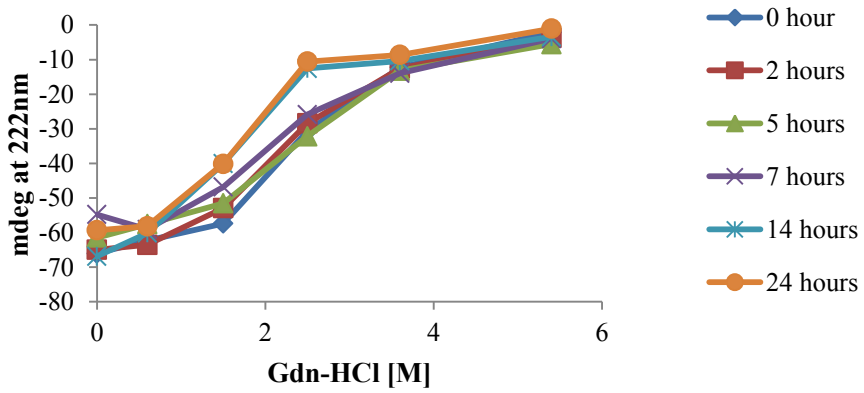



**Figure 18: Time-dependent Gdn-HCl denaturation studies of FL CM-PD monitored by far-UV CD spectroscopy.** Protein was prepared as described in Fig 12 in PBS/glycerol in a reaction volume of 1 mL at 25°C. A single sample at specific concentration was used to measure the ellipticities at each time points. Incubation time after the addition of Gdn-HCl is indicated. Spectra were recorded as described in Fig 12. Figure from Amiri (89).

### 3.4.3 Refolding Studies of Gdn-HCl Denatured FL CM-PD

The reversibility of the Gdn-HCl-induced unfolding reaction of FL CM-PD was monitored by far-UV CD spectroscopy with special attention to the effects at 1.5 and 2.5 M Gdn-HCl. Protein samples were incubated for 24 h in PBS/ glycerol buffer containing 0.6, 1.5, 2.5, 3.6, and 5.4 M Gdn-HCl and were then dialyzed overnight to remove the Gdn-HCl (see Fig 19 legend). Far UV-CD spectra were recorded before and after dialysis and the changes in ellipticity were followed at 222 nm (Fig 19).

As expected, protein incubated at 0.6, 3.6, and 5.4 M Gdn-HCl showed the predictable decreases in ellipticity at 222 nm reflecting protein unfolding. Following dialysis to remove Gdn-HCl (blue), FL CM-PD regained considerable but not full α-helical structure. In contrast, incubation with 1.5 and 2.5 M Gdn-HCl for 24 h resulted in a dramatic reduction in ellipticity (red) and a visible turbidity of the sample which was not reversed by removing the denaturant by extensive dialysis. The results in Figs 18 and 19 suggested that protein aggregation was kinetically controlled; FL CM-PD could be refolded by dialysis to remove high concentrations of Gdn-HCl but only if the length of time the sample remained at 1.5 M and 2.5 M was relatively short.

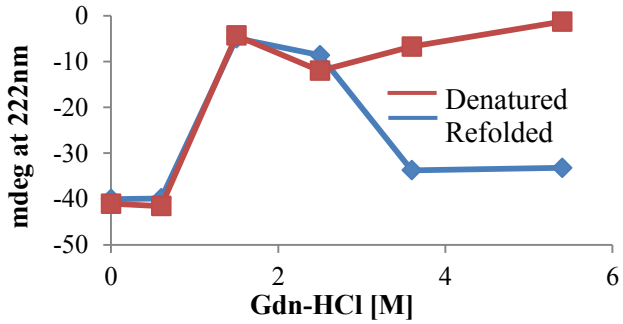



**Figure 19: Refolding of FL CM-PD by extensive dialysis after Gdn-HCl denaturation**. Protein was prepared as described in Fig 12 in PBS/glycerol at 0 M, 0.6 M, 1.5 M, 2.5 M, 3.6 M, and 5.4 M Gdn-HCl. After the addition of Gdn-HCl, samples were incubated for 24 h at room temperature then dialyzed at 4°C in the same buffer (2 L without Gdn-HCl) to dilute the Gdn-HCl to 0 M. Spectra were recorded for each sample as described in Fig 12 before and after the dialysis. The ellipticity at 222 nm was plotted as a function of Gdn-HCl. Figure from Amiri (89).

### 3.4.4 Analysis of Guanidine-Induced Unfolding of TyrA by Intrinsic Fluorescence Spectroscopy

Fig 20 shows the Gdn-HCl-induced changes in tertiary structure as monitored by intrinsic fluorescence emission. Protein samples used in this study were the same as those analyzed by far-UV CD spectroscopy (section 3.4.1). Additionally, the solutions of NAYA and NATA were analyzed to provide a comparison of the intrinsic fluorescence emission of completely exposed intrinsic fluorophores (90, 91).

Fig 20 A and B show the fluorescence emission spectra of FL and Δ80 CM-PD at $\lambda_{ex}$ 280 nm recorded after 4 h of incubation in the presence of PBS/glycerol buffer containing Gdn-HCl. As Gdn-HCl concentrations increased, $\lambda_{max}$ emission of both TyrA proteins shifted to higher wavelengths concomitant with a decrease in intensity. By 4.2 M Gdn-HCl the proteins were completely unfolded and the fluorophores completely solvent exposed as ascertained by a shift in $\lambda_{max}$ to ~ 355 nm, identical to the curve at 6 M Gdn-HCl and that of NATA/NAYA (dotted green curve). The decrease in fluorescence intensity at high concentrations of Gdn-HCl is the result of quenching by solvent of exposed Tyr residues in the unfolded protein (73). In accordance with the results from far-UV CD studies, FL CM-PD showed an irregularly low fluorescence intensities at 1.8 and 2.1 M Gdn-HCl (panel A) which coincided with visible turbidity of the protein. In contrast, Δ80 CM-PD showed a predictable Gdn-HCl-mediated decrease in fluorescence intensity without visible protein precipitation (panel B). The fluorescence changes in emission intensity at 320 nm are shown in panel C and clearly highlight the Gdn-HCl-induced intermediate in the unfolding of the two TyrA proteins.

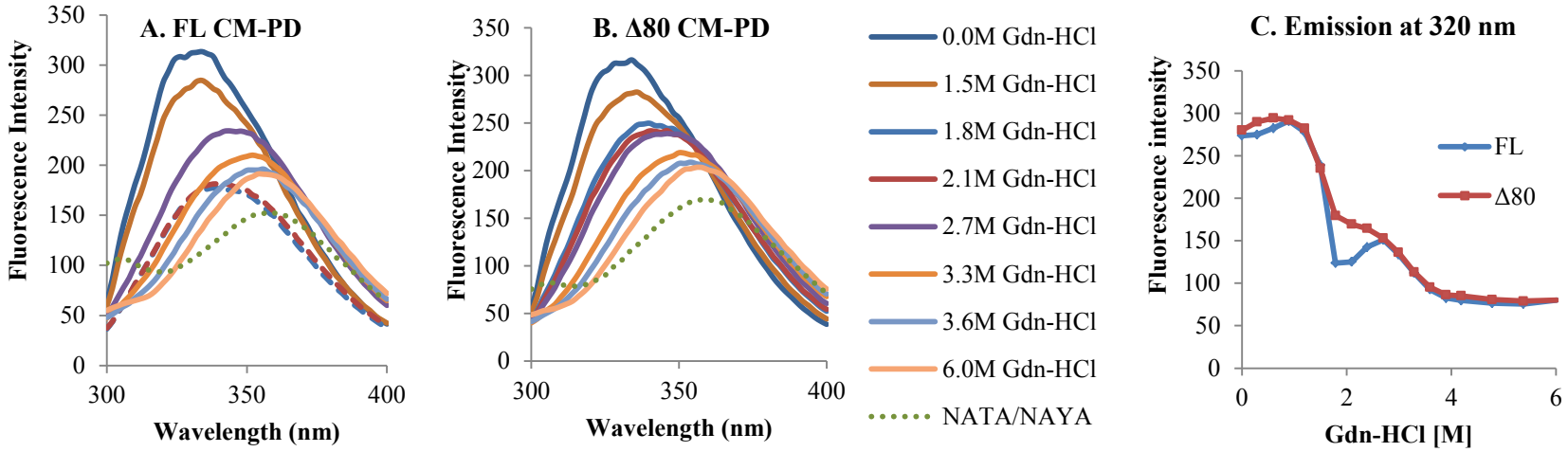



**Figure 20: Fluorescence emission spectra of *H. influenzae* TyrA after 4 h incubation with increasing Gdn-HCl.** Proteins were prepared as described in Fig 15 in PBS/glycerol, pH 7.5. The protein samples were incubated for 4 h after the addition of Gdn-HCl. NATA/NAYA standards were prepared at the same concentration of Tyr and Trp found in the protein samples (section 2.11) in PBS/glycerol, also incubated for 4 h. Emission spectra recorded at 25°C after excitation at 280 nm using the slit width of 5 nm and were the result of 15 scans at fast scan speed. Contributions to the signal from buffer and Gdn-HCl were subtracted. (C) Emission intensity at 320 nm was replotted as a function of Gdn-HCl for comparison between FL and Δ80 CM-PD. At 320 nm, the intensity difference of TyrA in the presence of 0 M and 6 M Gdn-HCl was the greatest. The emission spectra recorded after excitation at 295 nm showed comparable trends to those reported above (data not shown).

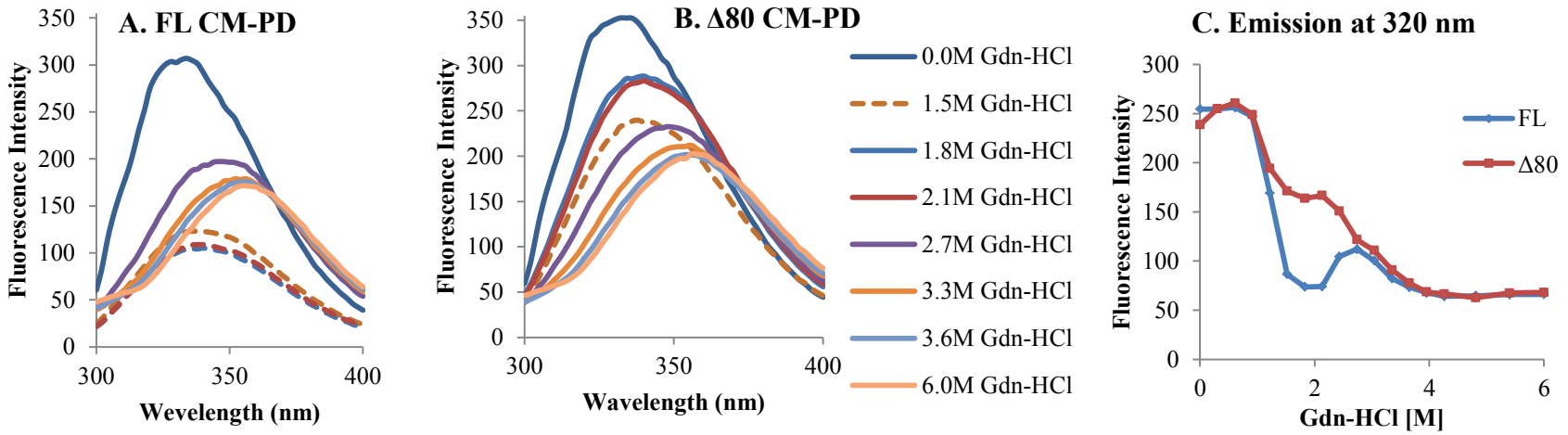



**Figure 21: Fluorescence emission spectra of *H. influenzae* TyrA after 18 h incubation with increasing Gdn-HCl.** Proteins were prepared as described in Fig 15 in 50mM $K_2HPO_4$ at pH 7.5. The protein samples were incubated overnight (~18 h) after the addition of Gdn-HCl. Emission spectra are recorded as listed in Fig 20. Contributions to the signal from buffer and Gdn-HCl were subtracted. (C) Emission intensity at 320 nm was replotted as a function of Gdn-HCl for comparison between FL and Δ80 CM-PD. The emission spectra recorded after excitation at 295 nm showed comparable trends to those reported above (data not shown).

When FL and Δ80 CM-PD proteins were incubated for 18 h in 50 mM $K_2HPO_4$ buffer (Fig 21 A and B), the spectra showed the expected red shift in $\lambda_{max}$ with increasing concentrations of Gdn-HCl. The unusual decrease in fluorescence emission intensity was prominent in the spectra for FL CM-PD at Gdn-HCl concentration of 1.5, 1.8 and 2.1 M (panel A) and at 1.5 M for Δ80 CM-PD (panel B). We attributed this change in fluorescence emission intensity (Fig 21, panel C) to visible protein precipitation which was much more pronounced than when the proteins were incubated for 4 h in PBS/glycerol buffer (Fig 20, panel C).

When the data from Fig 20 and 21 were replotted as $\lambda_{max}$ emission versus increasing concentrations of Gdn-HCl, the resulting curves (Fig 22) were superimposable, suggesting that emission maxima were not influenced by protein precipitation. Probed by this measurement, the loss of Gdn-HCl-induced tertiary structure is similar for both TyrA proteins regardless of the presence or absence of the CM domain.

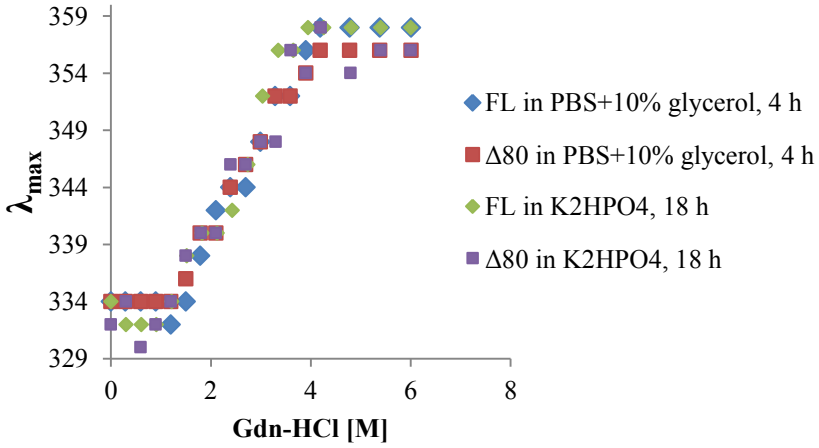



**Figure 22: $\lambda_{max}$ from fluorescence emission spectra of FL and Δ80 CM-PD as a function of Gdn-HCl**. The buffer and incubation times are noted above. Data were derived from Fig 20 and Fig 21.

## 3.5 Thermal Stability Studies of *H. influenzae* TyrA

### 3.5.1 Analysis of Temperature-Induced Unfolding of TyrA by Far-UV CD Spectroscopy

The thermal stability of *H. influenzae* TyrA proteins was probed by analyzing the changes in ellipticity at 222 nm as a function of increasing temperature (25–95°C) using far-UV CD spectroscopy. Variable temperature (VT) experiments were conducted in PBS/10% glycerol as this buffer had been previously reported to help stabilize *E. coli* CM-PD under different physiochemical conditions (92). The thermal denaturation curves are shown in Fig 23 and revealed that FL and Δ80 CM-PD responded differently to increasing temperature. While both proteins show a cooperative transition between 50 and 56 °C, yielding similar $T_m$ values (52°C and 50°C for FL and Δ80 CM-PD, respectively), the FL protein continued to lose secondary structure after the transition, concomitant with precipitation, while the ellipticity of the Δ80 variant remained constant. To determine the loss of protein by thermally-induced precipitation, samples recorded at the end of the VT experiment were centrifuged and the amount of soluble protein was quantified by a Bradford assay. Only 10% of the total FL protein remained soluble while over 60% of Δ80 CM-PD was retained in solution.

The far-UV CD spectra were recorded from 200-240 nm before and after the VT experiment (Fig 24). At 95°C, values of MRW ellipticity for FL CM-PD were close to zero over this wavelength range indicating that there was little soluble protein remaining and/ or that the protein had lost considerable secondary structure. By contrast, Δ80 CM-PD showed a pronounced minimum at 218 nm, suggesting that the truncated protein had

formed a structure dominated by β-sheets when heated to 95°C. Thermal denaturation was not reversible (49).

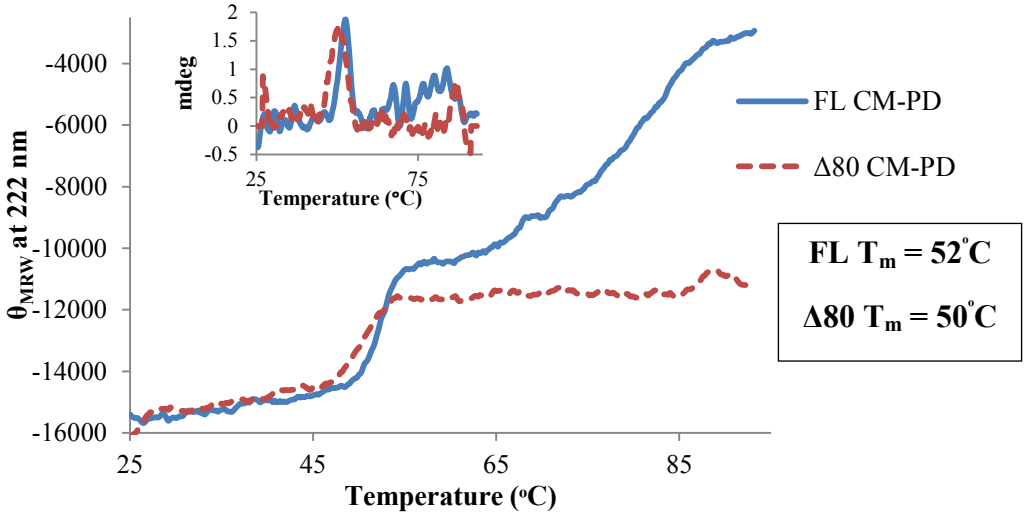



**Figure 23: Changes in mean residue ellipticity at 222 nm as a function of temperature**. Proteins (0.1 mg/mL) were reconstituted in PBS/glycerol, pH 7.5. Data were recorded at 222 nm using a 0.2 cm pathlength rectangular cell and a ramping speed of 20°C/h. Ellipticity in mdeg was converted to MRW ellipticities as described in Fig 12. Inset shows the 1st derivative of the changes at 222 nm; the inflection point was used to estimate the melting temperature ($T_m$).

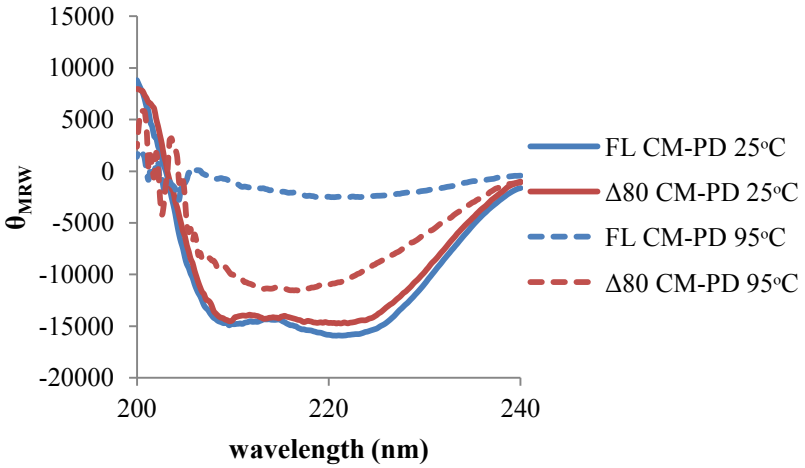



**Figure 24: Changes in mean residue ellipticity before and after thermal denaturation**. Proteins (0.1 mg/mL) were reconstituted in PBS/glycerol, pH 7.5. Data were recorded from 200 nm to 240 nm using 0.2 cm pathlength rectangular cell. Other instrument parameters were listed in Fig 12. CD units in mdeg were converted to MRW ellipticities as described in Fig 12.

### 3.5.2 Analysis of Temperature-Induced Denaturation of TyrA by Intrinsic Fluorescence Emission

In order to probe the changes in tertiary structure of *H. influenzae* TyrA proteins as a function of temperature, intrinsic fluorescence emission scans were recorded before (25°C) and after (95°C) thermal denaturation and the results are shown in Fig 25. Fluorescence intensity at the higher temperature decreases dramatically due to a shortened emission lifetime of the fluorophores (93). The Tyr and Trp analogs, NATA and NAYA, were used to normalize the intensity changes in protein fluorescence emission due to high temperature. At 25°C, emission scans of both proteins showed $\lambda_{max}$ value of 332 nm, indicative of partially buried Trp residues in the folded protein. After heating the proteins to 95°C the $\lambda_{max}$ emission shifted to ~346 nm. By comparison NATA/NAYA emission scans revealed $\lambda_{max}$ values of 358 nm, suggesting that the TyrA proteins were not fully unfolded after thermal denaturation and their intrinsic fluorophores remained partially buried.

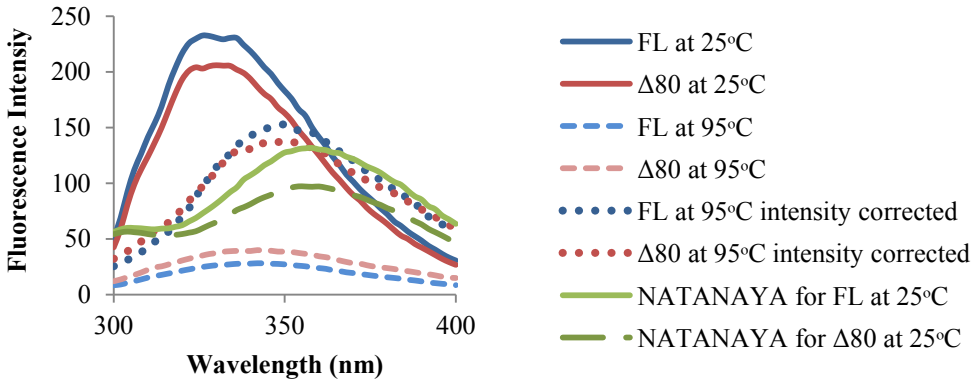



**Figure 25: $\lambda_{max}$ from fluorescence emission spectra of TyrA proteins before and after thermal denaturation**. Samples of protein and NATA/NAYA standards were prepared as described in Fig 20. The protein samples were incubated before (25°C) and after (95°C) the VT experiments for 3 min at the appropriate temperature. NATA/NAYA standards prepared at room temperature were then incubated at 25°C and 95°C for 3 min. Emission spectra were recorded after excitation at 280 nm. Other instrument parameters were same as previously described (Fig 14). Decrease in fluorescence intensities of protein's at 95°C were normalized by adding the difference between fluorescence emission intensity of NATA/NAYA at 25°C and 95°C at all wavelengths (modified from 91, 94). Curves for NATA/NAYA at 95°C not shown.

### 3.5.3 Analysis of Temperature–Induced Unfolding of TyrA by FTIR Spectroscopy

In order to further analyze structural changes that occur during thermal denaturation of TyrA, variable temperature (VT) FTIR spectroscopy was used. Amide I' and amide II' bands are excellent probes of protein secondary and tertiary structure, respectively.

Preliminary VT-FTIR studies were conducted by collecting spectra immediately after proteins were dissolved in $D_2O$. Consequently, the intensity changes of the amide I' bands following the temperature-mediated unfolding of secondary structure could be influenced by the band broadening of H-D exchange in the amide band region. The rate of H-D exchange is dependent on the solvent accessibility of protein structures. For comparison, both proteins were incubated in $D_2O$ for 24 h to ensure complete H-D exchange prior to VT experiments.

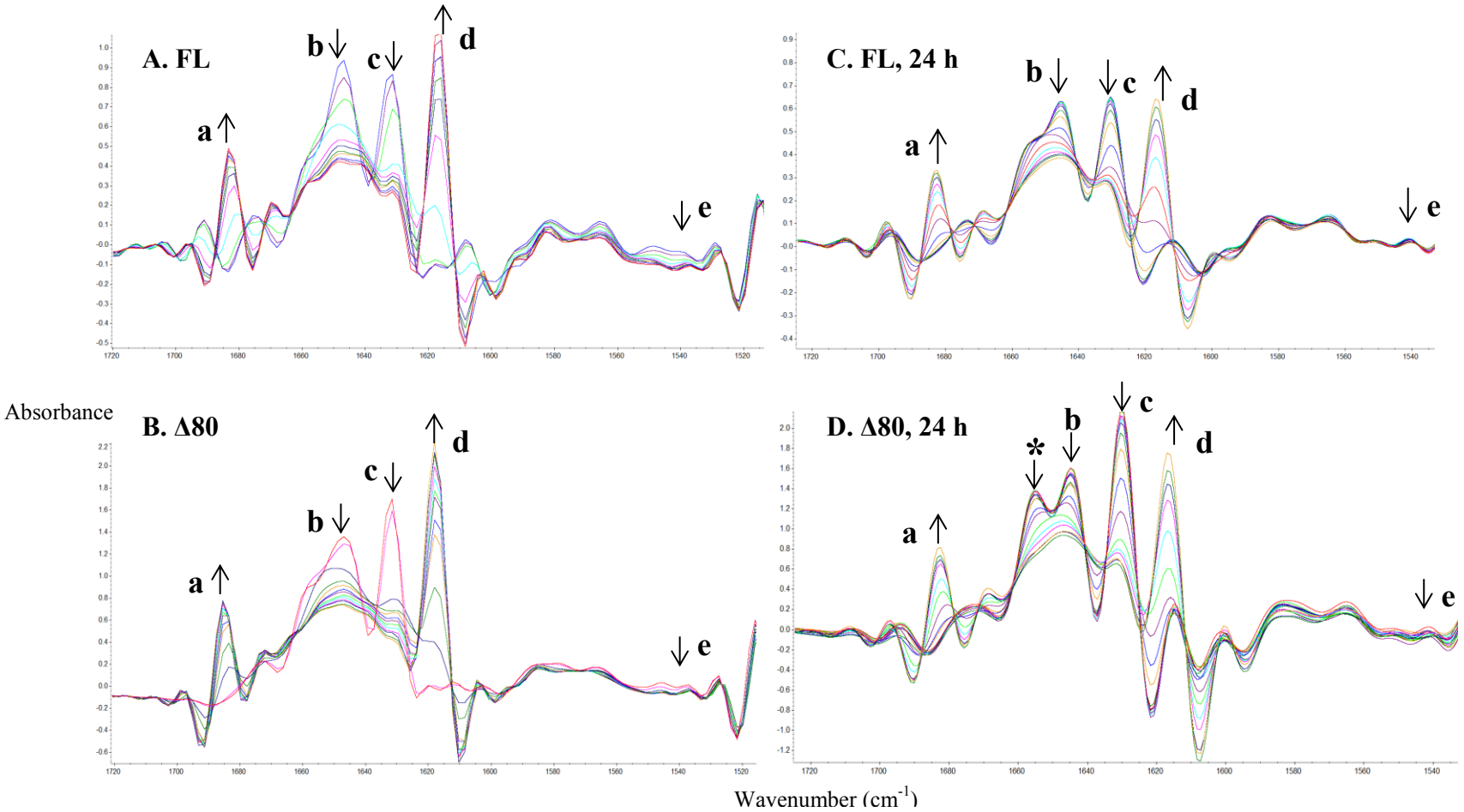



**Figure 26: VT-FTIR spectral analysis of *H. influenzae* TyrA.** Lyophilized protein samples were prepared as described in Fig 13. Thermal denaturation began immediately after dissolution of proteins (A and B) in $D_2O$ or after 24 h of incubation (C and D) in $D_2O$. Spectra were recorded as listed in Fig 13. For C and D, denaturation was carried out from 25°C to 95°C in 2°C increments and 4 min equilibration time, while for A and B, from 35°C-95°C with 5°C increment and 4 min equilibration time. Peak annotations are: (a) high frequency intermolecular β-sheets (~1680 $cm^{-1}$), (b) α-helices (~1650 $cm^{-1}$), (c) intramolecular β-sheets (~1630 $cm^{-1}$), (d) low frequency β-sheets (~1618 $cm^{-1}$), (e) amide II' band (~1540 $cm^{-1}$). (*) assigned to α-helices in the PD domain (1657 $cm^{-1}$). Arrows indicate the increase (↑) or decrease (↓) of the peak intensities as a function of temperature.

Fig 26 shows a comparison of the VT-FTIR spectra of *H. influenzae* TyrA proteins immediately after dissolution of proteins in $D_2O$ (A and B) or after 24 h of incubation of proteins in $D_2O$ (C and D). The amide I' and II' bands were assigned (95) after resolution enhancement by application of Fourier self-deconvolution (FSD).

The amide II' band intensity at 1540 $cm^{-1}$ (denoted e) showed a significant difference between complete and incomplete H-D exchanged protein samples. When VT experiments were performed on protein samples immediately after resuspension in $D_2O$ (panels A and B), the intensity of amide II' band gradually decreased indicating the continuous H-D exchange of amide groups of protein as a function of temperature. By contrast, samples that had been incubated for 24 h in $D_2O$ (panels C and D) to ensure complete H-D exchange showed negligible changes in the amide II' band ($\leq 0.001$) with temperature.

The intensity changes of the amide I' bands of both FL and Δ80 CM-PD as a function of temperature appeared to follow the same general trends whether or not H-D exchange was complete before VT experiments were initiated. The absorbance for α-helices (denoted b; 1649 $cm^{-1}$) and intramolecular β-sheets (denoted c; 1631 $cm^{-1}$) decreased rapidly with increasing temperature while the absorbance for intermolecular β-sheets (aggregation bands, denoted a and d, respectively; 1680 $cm^{-1}$and 1618 $cm^{-1}$) increased. The intensity of intermolecular β-sheets for both TyrA proteins showed very little change after 65°C indicating that protein aggregation/ denaturation was complete by this temperature. The aggregation band remained constant upon cooling from 95°C to 25°C (data not shown) demonstrating the irreversibility of the thermal denaturation process. Interestingly, after incubation of Δ80 CM-PD in $D_2O$ for 24 h, the spectra of

this sample revealed a pronounced peak at ~1657 $cm^{-1}$ (panel D) which we attributed to an additional α-helix band. The preliminary Peak Resolution analysis (Appendix 4) confirmed the presence of this α-helix band in FL and Δ80 CM-PD although it is less prominent in the spectra of FL CM-PD. The α-helix band at ~1657 $cm^{-1}$ may be attributed to α-helices within the PD domain.

For better visual representation of the absorbance changes in the amide I' region, intensities at specific bands at each temperature from Fig 26 were re-plotted in Fig 27. Table 6 also shows the temperature at which a specific band commenced unfolding (in black) and its overall $T_m$ value (in red).

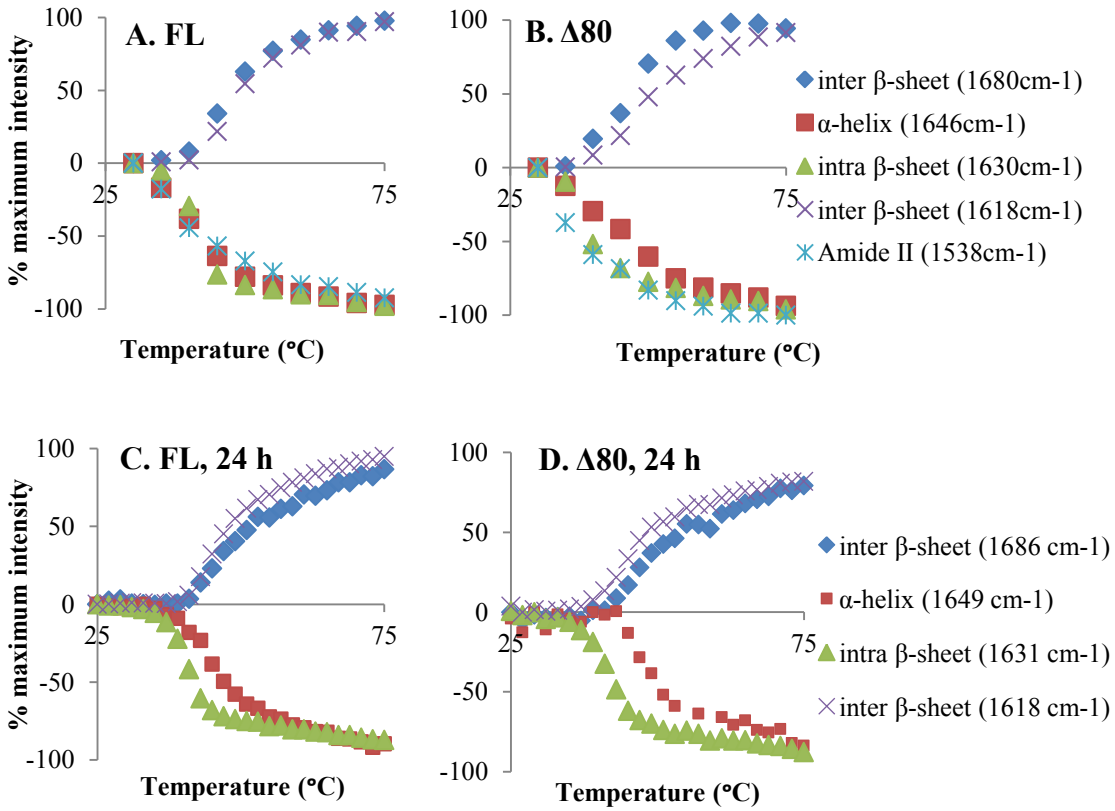



**Figure 27: Analysis of band-specific changes in absorbance as a function of temperature for *H. influenzae* TyrA**. The spectra from Fig 26 is re-plotted to show changes in intensities according to the relationship: % maximum intensity = $[(I_{tp} - I_{25}) / (I_{95} - I_{25})] \times (\pm)\ 100\%$, intensities at a given temperature ($I_{tp}$), at 25°C ($I_{25}$) and at 95°C ($I_{95}$). (+ or -) 100% is multiplied depending on the increasing (+) or decreasing (-) band intensity as a function of temperature. Bands at 1686 $cm^{-1}$ and 1618 $cm^{-1}$ increased together and were assigned to the formation of intermolecular β-sheet as a result of protein aggregation. The $T_m$ of both FL and Δ80 CM-PD in all conditions were about 47 °C as calculated from the midpoint of the curves for β-aggregation. Amide II' band for samples incubated for 24 h in $D_2O$ were not plotted because the loss of intensity at 1540 $cm^{-1}$ as a function of temperature was negligible (≤ 0.001 absorbance change).

**Table 6. Analysis of Fig 27**

| | | No incubation* | | 24 h incubation* | |
|---|---|---|---|---|---|
| | | A. FL | B. Δ80 | C. FL | D. Δ80 |
| Loss | intramolecular β-sheet (1631 $cm^{-1}$) Amide II band (1544 $cm^{-1}$) | 35/ 40 | 35/ 40 | 35/ 40 | 35/ 42 |
| | **α-helices (1649 $cm^{-1}$)** | **35/ 40** | **35/ 43** | **39/ 45** | **45/ 47** |
| Gain | intermolecular β-sheet (aggregation bands) (1686 $cm^{-1}$ and 1618 $cm^{-1}$) | 40/ 47 | 40/ 47 | 41/ 47 | 41/ 47 |

This table indicates a temperature at which specific secondary structures started to unfold (black) and the observed $T_m$ (red). *VT experiments were performed on protein samples after immediate dissolution in $D_2O$ or after samples were incubated for 24 h in $D_2O$.

The intramolecular β-sheets of both FL and Δ80 CM-PD start to denature at 35°C coincident with changes in the amide II' band, followed by the formation of the aggregation bands at 1680 $cm^{-1}$and 1618 $cm^{-1}$ (starting at ~40°C) regardless of the sample's incubation time in $D_2O$. Fig 27 clearly shows a difference however, in the unfolding of the α-helix in both TyrA proteins before (panel A and B) or after 24 h of incubation (panel C and D) in $D_2O$. Intensity changes of the α-helix band in FL CM-PD (panel A) were completely overlaid with those of an intramolecular β-sheet suggesting the parallel loss of these two secondary structures when VT experiments commenced immediately after sample preparation. After 24 h of incubation in $D_2O$ prior to the VT experiments, however, the α-helical structure started to unfold much later than the intramolecular β-sheets for both TyrA proteins. Intensity change of the α-helix band in FL CM-PD (panel C) commenced at 39°C, ~4°C higher than intramolecular β-sheet. Similarly, unfolding of α-helices in Δ80 CM-PD (panel D) commenced at 45°C, which is 10°C higher than the unfolding of intramolecular β-sheets. $T_m$ values calculated from the unfolding curves of α-helices in both TyrA proteins after samples were 24 h incubation in $D_2O$ were higher by ~5°C compared to those samples without incubation (Table 6). The $T_m$ values of α-helices in Δ80 CM-PD were also higher by 2°C than the $T_m$ of α-helices of FL CM-PD.

In summary, the data indicated that the changes in the α-helical intensity commenced immediately in VT experiments due to N-D exchange in the amide I' region unless samples were incubated in $D_2O$ for 24 h. Therefore, with adequate H-D exchange prior to VT experiments, the changes observed for α-helices were unaffected by changes from amide I' region. The α-helices of both proteins were more thermally stable than

intramolecular β-sheets. Additionally the α-helices of Δ80 CM-PD appeared more stable than those in FL CM-PD.

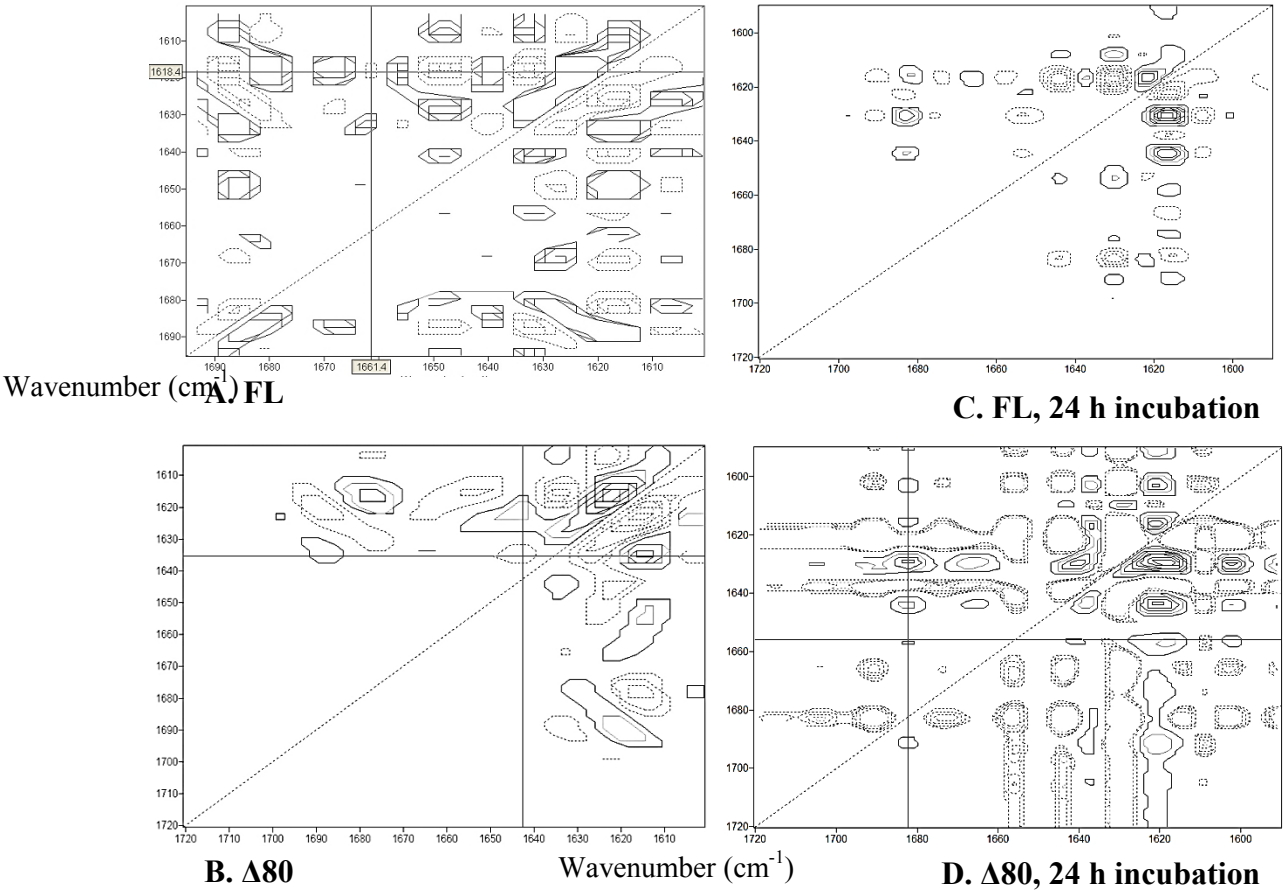



**Figure 28: Asynchronous 2D spectra of *H. influenzae* TyrA proteins**. 2D spectra were generated from 1D spectra from 25°C to 55°C because the $T_m$ was 47 °C for both proteins. The band assignments were the same as in Fig 27: high frequency intermolecular β-sheets, ~1680 $cm^{-1}$; α-helices, ~1650 $cm^{-1}$; intramolecular β-sheets, ~1630 $cm^{-1}$; low frequency β-sheets, ~1618 $cm^{-1}$; amide II' band, ~1540 $cm^{-1}$.

The sequence of unfolding of FL and Δ80 CM-PD was verified by 2D correlation analysis of the FTIR spectra. 2D analysis is a useful tool in studying the sequence of events by detecting the subtle changes in carbonyl stretching due to protein conformational changes. Synchronous and asynchronous maps are involved in 2D analysis. Fig 28 shows the asynchronous maps for FL and Δ80 CM-PD from 25°C to

55°C with or without 24 h incubation in $D_2O$ prior to VT experiments. The results are tabulated in Tables 7 and 8 (refer to Appendix 2 A and B for the synchronous maps).

**Table 7. Sequence of unfolding of FL and Δ80 CM-PD from 25°C to 55°C**

| | | x-axis on asynchronous map | | |
|---|---|---|---|---|
| | **Wavenumber** ($cm^{-1}$) | **1684 ↑** intermolecular β-sheet | **1650 ↓** α-helix | **1631 ↓** intramolecular β-sheet |
| **y-axis on asynchronous map** | **1618 ↑** intermolecular β-sheet | +, 0 | -, - → | -, - ← |
| | **1631 ↓** intramolecular β-sheet | -, + → | +, 0 | |
| | **1650 ↓** α-helix | -, + → | | |
| Sequence of events | | | | |
| ⇩ | Unfolding of α-helix (1650 $cm^{-1}$) and intramolecular β-sheet (1631 $cm^{-1}$)<br>Formation of intermolecular β-sheet (1618 and 1684 $cm^{-1}$) | | | |

The up and down arrows indicate increase and decrease of peaks as a function of temperature.

**Table 8. Sequence of unfolding of FL and Δ80 CM-PD from 25°C to 55°C after 24 h incubation in $D_2O$.**

| | | x-axis on asynchronous map | | |
|---|---|---|---|---|
| | **Wavenumber** ($cm^{-1}$) | **1684 ↑** intermolecular β-sheet | **1650 ↓** α-helix | **1631 ↓** intramolecular β-sheet |
| **y-axis on asynchronous map** | **1618 ↑** intermolecular β-sheet | +, + ← | -, 0 | -, - ← |
| | **1631 ↓** intramolecular β-sheet | -, + → | +, - → | |
| | **1650 ↓** α-helix | -, 0 | | |
| Sequence of events | | | | |
| ⇩ | Unfolding of intramolecular β-sheet (1631 $cm^{-1}$)<br>Unfolding of α-helix (1650 $cm^{-1}$)<br>Formation of intermolecular β-sheet(1618 and 1684 $cm^{-1}$) | | | |

The sequence of unfolding of the secondary structural components of both TyrA proteins between 25°C and 55°C is very similar. When proteins were allowed to undergo complete

H-D exchange (Table 8) prior to the VT experiments (after 24 h incubation in $D_2O$), intramolecular β-sheet (1631 $cm^{-1}$) unfolded first, and followed by α-helix (1650 $cm^{-1}$) which was followed by the formation of intermolecular β-sheets (aggregation at 1684 and 1618 $cm^{-1}$). This pattern contrasted the results obtained when H-D exchange was incomplete prior to VT experiments (Table 7) and showed α-helix (1650 $cm^{-1}$) and intramolecular β-sheet (1631 $cm^{-1}$) unfolded concurrently which was then followed by β-aggregation. Despite the differences in the complexity of the asynchronous maps, the sequence of unfolding of FL and Δ80 CM-PD's assigned major secondary structures appeared to be identical.

### 3.5.4 Analysis of Temperature-Induced Unfolding of TyrA in the Presence of Ligands Using FTIR Spectroscopy

To further probe the thermal stability of secondary structure components of TyrA proteins, VT-FTIR measurements were conducted in the presence of the active site ligands, L-Tyr and $NAD^+$. For complete H-D exchange, protein samples were incubated at 25°C overnight (over 18 h) in $D_2O$ prior to VT experiments.

The spectra of FL and Δ80 CM-PD were shown in Fig 29. As expected, the intensity of α-helices (b) and intramolecular β-sheet (c) decrease while the intensities of intermolecular β-sheets (a and d) increased as a function of temperature. The temperature mediated-intensity changes in the amide I' bands were replotted in Fig 30. A $T_m$ value of 51°C derived from the midpoint of the β-aggregation band was calculated for both FL and Δ80 CM-PD in complex with ligands; the binding of ligands stabilized the secondary

structure of TyrA proteins by 4°C compared to TyrA proteins in the absence of ligands (47°C in Table 6).

The most striking observation in Fig 30, however, was that the unfolding of α-helices and intramolecular β-sheets were coincident for both proteins (panel A and B) in contrast to the results observed in the absence of ligands (see Fig 27 C and D). Surprisingly though, this occurs by a destabilization of the α-helices. While $T_m$ values for the unfolding of α-helices of both proteins (~45°C) remained unchanged in the presence and absence of ligands (Table 10), in the presence of ligands the α-helices commenced unfolding at 35°C, temperatures which were 4-10°C lower than for unliganded FL (39°C) and Δ80 CM-PD (45°C), respectively (Tables 10). Accordingly, the temperature-mediated unfolding of the helices appeared less cooperative for the liganded proteins. In contrast, the unfolding of intramolecular β-sheets in the TyrA proteins commenced at the same temperature (35°C) regardless of whether or not ligands were present (Table 10). However, the $T_m$ of the intramolecular β-sheets of the ligand-bound proteins at 45°C was about 3-5°C higher than the TyrA proteins without ligands (Table 10).

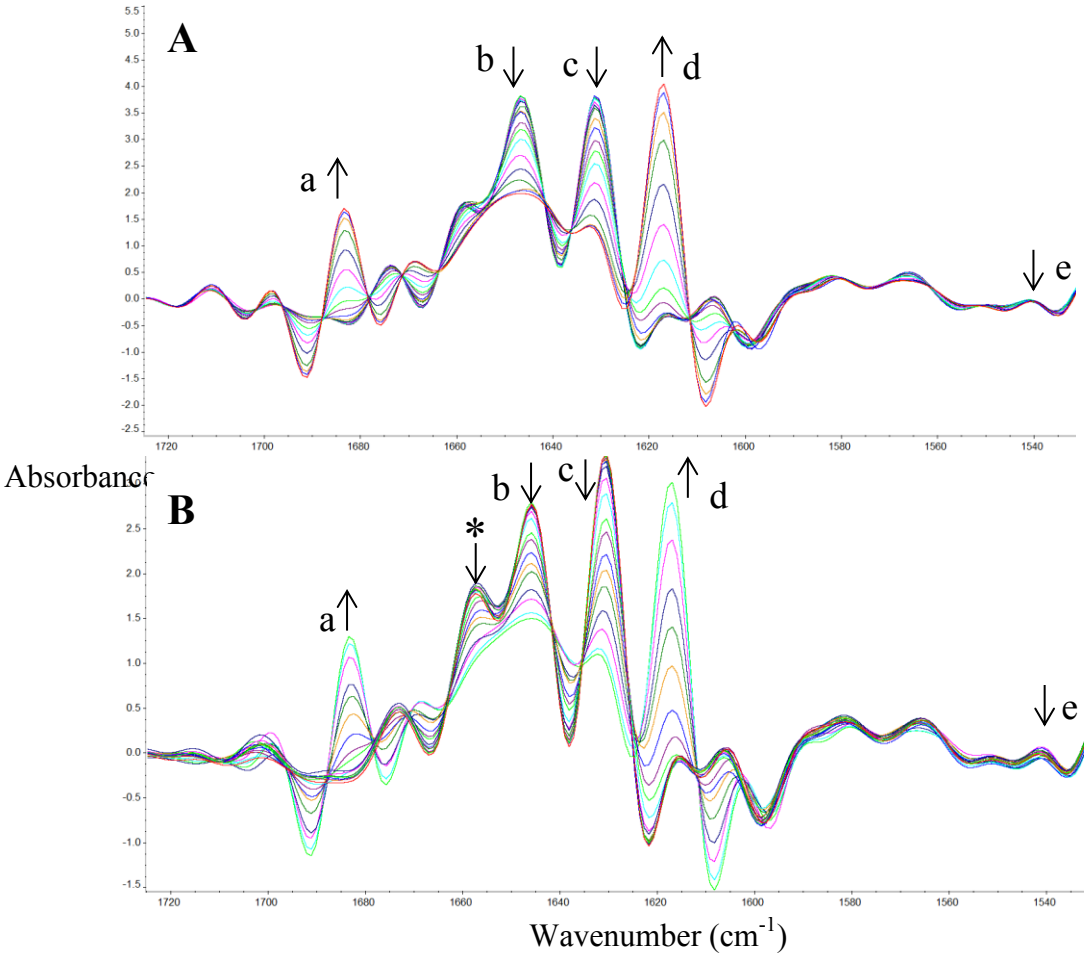



**Figure 29: VT- FTIR spectral analysis of (A) FL and (B) Δ80 CM-PD with ligands after complete H-D exchange**. Proteins were dialyzed overnight (refer to section 2.12) in 10 mM $K_2HPO_4$ and 1 mM DTT with ligands, and lyophilized overnight. The dried powder was weighed to 1.2 mg and reconstituted in 15 µL $D_2O$. The concentration of ligands of L-Tyr and $NAD^+$ after reconstitution in $D_2O$ was 2 mM each whereas the TyrA proteins were about 2 - 2.3 mM. The molar ratio of ligands to protein was about 1:1. The concentration of L-Tyr could not be increased due to the low solubility of L-Tyr (~2 mM at pH 7.0) (96). $NAD^+$ was added to facilitate the binding of L-Tyr (21). Instrument parameters were as listed in Fig 26. Peak annotations are; (a) high frequency intermolecular β-sheets (~1680 $cm^{-1}$), (b) α-helices (~1650 $cm^{-1}$), (c) intramolecular β-sheets (~1630 $cm^{-1}$), (d) low frequency β-sheets (~1618 $cm^{-1}$), (e) amide II' band (~1540 $cm^{-1}$). Arrows indicate the increase (↑) or decrease (↓) of the peak intensities as a function of temperature. (*) assigned to α-helices in the PD domain (1657 $cm^{-1}$).

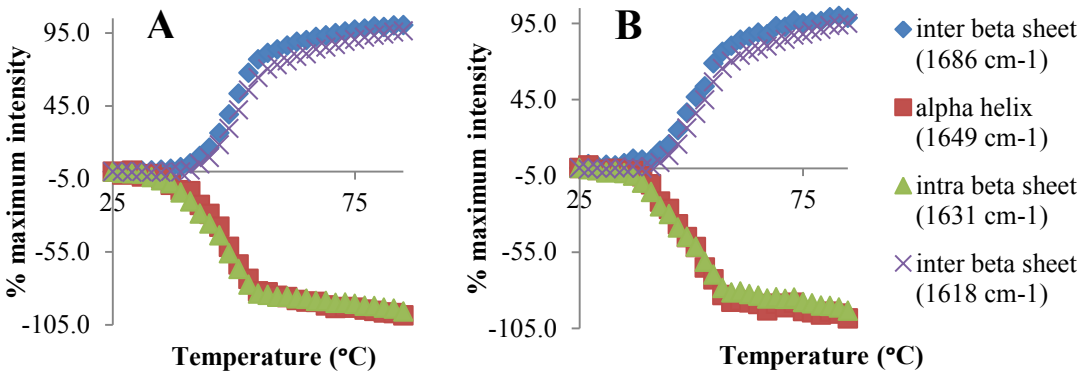



**Figure 30: Analysis of band-specific changes in absorbance as a function of temperature for (A) FL and (B) Δ80 CM-PD with ligands.** The 1686 $cm^{-1}$ and 1618 $cm^{-1}$ bands move together and were assigned to the increase in intermolecular β-sheet as a result of protein aggregation. The $T_m$ of ~51°C was calculated using the curve for intermolecular β-sheet. The band assigned to α-helices (1649 $cm^{-1}$) was lost at the same rate as intramolecular β-sheet band (1631 $cm^{-1}$) as temperature increases.

**Table 9. Analysis of Fig 30.**

| | | A. FL and B. Δ80 |
|---|---|---|
| Loss | intramolecular β-sheet (1631 $cm^{-1}$)<br>Amide II band (1544 $cm^{-1}$) | 35/ 45 |
| | **α-helices (1649 $cm^{-1}$)** | **35/ 45** |
| Gain | intermolecular β-sheet<br>(aggregation bands)<br>(1686 $cm^{-1}$ and 1618 $cm^{-1}$) | 40/ 51 |

This table indicates a temperature which specific secondary structure started to unfold (black) and the observed $T_m$ (red).

**Table 10. Analysis of changes in $T_m$ and temperatures of the start of unfolding of secondary structure of unliganded and liganded TyrA proteins**

| | | TyrA no ligands (Fig 27 C & D) | | TyrA ligands bound (Fig 30) | TyrA ligands bound – TyrA no ligands* | |
|---|---|---|---|---|---|---|
| | | FL | Δ80 | FL & Δ80 | FL | Δ80 |
| Loss | intra β-sheet (1631 $cm^{-1}$) | 35/ 40 | 35/ 42 | 35/ 45 | 0/5 | 0/3 |
| | **α-helices (1649 $cm^{-1}$)** | **39/ 45** | **45/ 47** | **35/ 45** | **-4/0** | **-10/-2** |
| | α-helices (1657 $cm^{-1}$) | 37/ 43 | 37/ 44 | 37/ 45 | 0/2 | 0/1 |
| Gain | β-aggregation (1686 $cm^{-1}$ , 1618 $cm^{-1}$) | 41/ 47 | 41/ 47 | 40/ 51 | -1/4 | -1/4 |

This table indicates a temperature which specific secondary structure started to unfold (black) and the observed $T_m$ (red). * $T_m$ and temperature at which secondary structure started to unfold of unliganded TyrA protein subtracted from liganded TyrA protein. Positive value means the stabilization of secondary structure by bound ligands and negative value means the destabilization of secondary structure by bound ligands.

## 3.6 Analysis of Temperature-Induced Unfolding of TyrA by Far-UV CD Spectroscopy Under FTIR Experimental Conditions

For a more accurate comparison of thermal stability of TyrA measured by FTIR and far-UV CD spectroscopies, VT CD measurements were recorded using protein concentrations and buffer conditions identical to those used in the FTIR experiments. Accordingly, the Jasco 815 spectropolarimeter was modified so that the instrument could accommodate the small pathlength FTIR cell and heating unit. To our knowledge, these are the first VT experiments ever reported to allow the comparison of these two techniques. Proteins were prepared as described in Fig 13 and were incubated for 24 h in $D_2O$ prior to CD measurements. Accordingly, protein concentrations were ~200-400 fold higher than the 3.5 μM reported in the initial CD experiments (Fig 23). Additionally, using the CONTIN algorithm accessed through Dichroweb (70), the relative

contributions of the major secondary structural components of TyrA were calculated from the experimental data and then used to construct the theoretical curves.

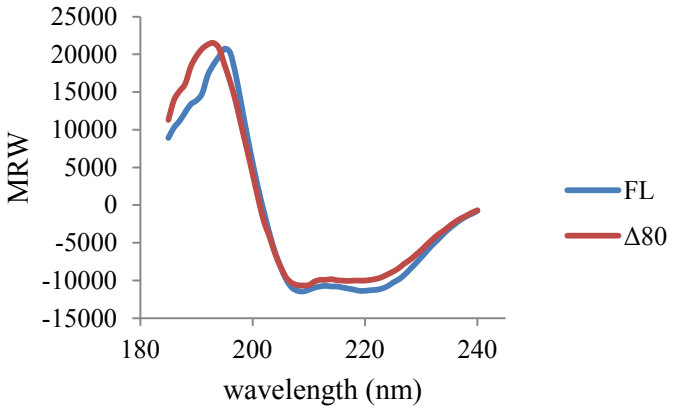



**Figure 31: Far-UV CD spectra of *H. influenzae* TyrA at high concentrations of protein**. Data were expressed as MRW, mean residue weight ellipticity. Protein samples are prepared as described in Fig 13 at concentrations of 800 µM and 1500 µM monomer for FL and Δ80 CM-PD, respectively. Measurements were recorded on a Jasco 815 Spectropolarimeter at 25°C using a spacer pathlength of 12 µm (for FL CM-PD) and 6 µm (for Δ80 CM-PD). The curves were the result of the average of 5 scans from 180 to 240 nm performed at a rate of 50 nm/min. Milideg was converted to molar residue weight ellipticity, MRW ($\theta_{MRW}$) using the relationship described in Fig 12.

Experimentally derived spectra (Fig 32 A, C) were then fitted (Fig 32 B, D) using Dichroweb (CONTIN) (70, 71). Fig 31 shows the far-UV spectra of TyrA proteins recorded at 25°C. Measurable absorbances of TyrA proteins from 180-240 nm were obtained by using a much smaller pathlength cell (6 or 12 µm spacers) to compensate for the high protein concentrations. Both TyrA proteins clearly showed spectra characteristic of α-helical structures (68) - a double minimum at 208 and 222 nm, and a maximum at about 193 nm. At the higher concentration of protein, FL CM-PD also showed a slightly stronger intensity at 222 nm, (a feature also reported when using lower concentrations of protein Fig 12), which was attributed to the helical CM domain averaged globally per

residue. The data indicated that at 25°C there were no secondary structure changes to the TyrA proteins caused by the higher protein concentrations. Fig 32 shows the VT far-UV spectra derived experimentally for both TyrA proteins (panel A and C) and the theoretical curves (panel B and D) reconstructed using CONTIN. All of the reconstructed curves showed good fits to the experimental far-UV CD scans as reflected in the normalized root mean square deviation (NRMSD up to 0.1) for both proteins. The spectral overlay of the experimental data and the CONTIN reconstructed curves are shown in Appendices 3A and 3B.

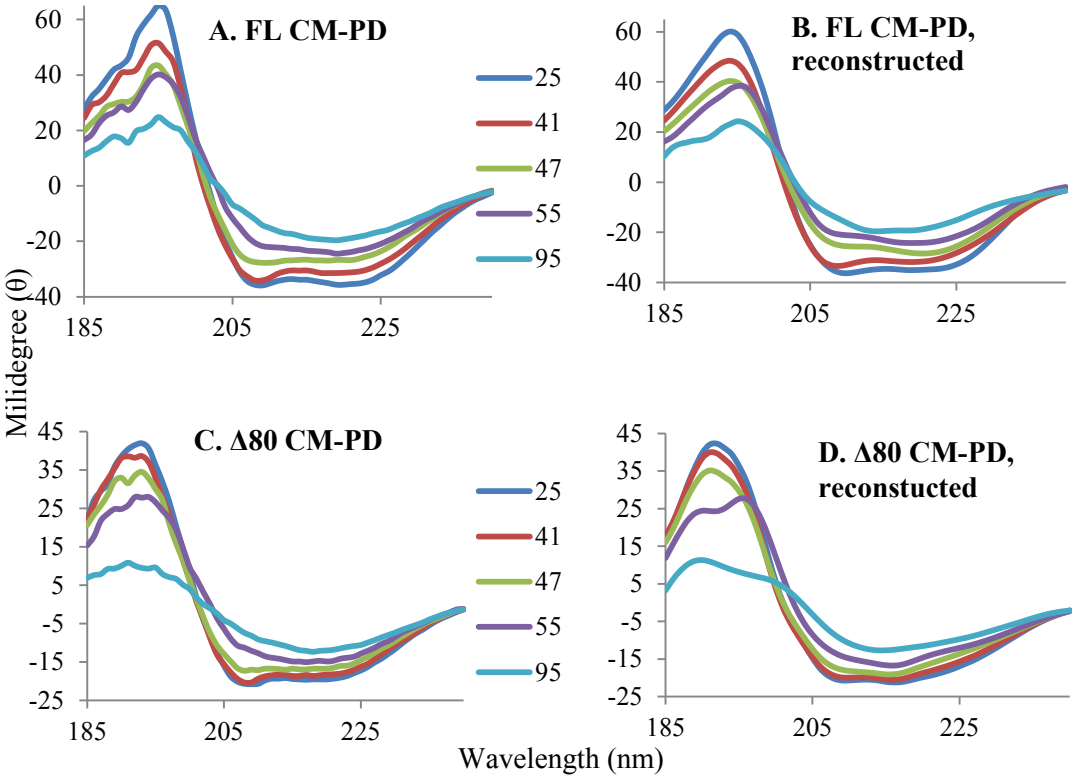



**Figure 32: Far UV-CD measurements of TyrA proteins derived experimentally (A and C) and by curve fitting (B and D).** FL (30 mg/mL monomer) and Δ80 (70 mg/mL monomer) CM-PD samples were prepared as described in Fig 13. Measurements were recorded on a Jasco 815 Spectropolarimeter from 25°C to 95°C with 2°C increment (manually increased), 4 min equilibrium time. Scans were recorded at 35 different temperatures, but only representative curves at selected temperature are shown. Spacers with pathlengths of 12 and 6 µm were used for FL and Δ80 CM-PD, respectively. The curves were the results of the average of 5 scans performed at a rate of 50 nm/min.

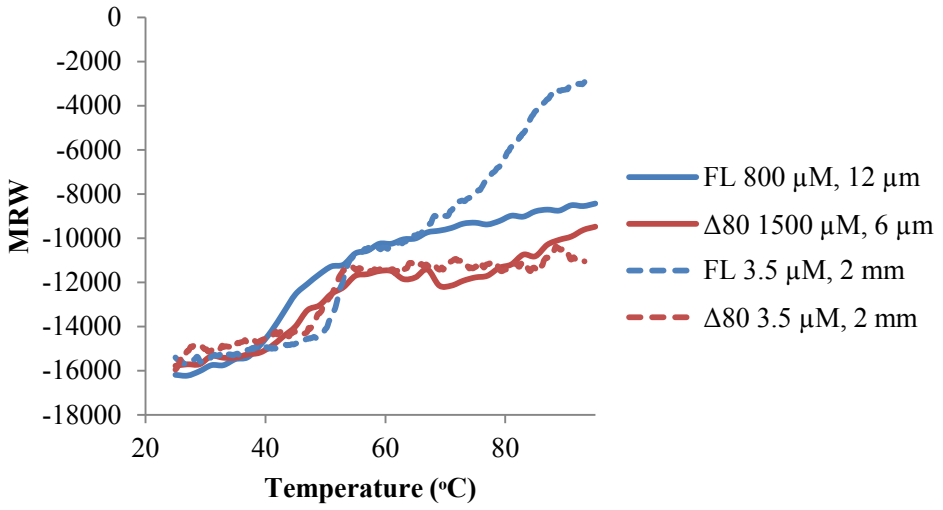



**Figure 33: Changes of mean residue ellipticity at 222 nm as a function of temperature**. TyrA proteins at various concentrations were measured with different pathlength cells. Intensities at 222 nm from Fig 31 were replotted as a function of temperature for 800 µM FL (monomer) and 1500 µM Δ80 CM-PD (monomer) prepared in $K_2HPO_4$ buffer. The 3.5 µM monomers of each protein prepared in PBS/glycerol were recorded at 222 nm at a ramping speed of 20°C/h (from Fig 23). CD units in milideg were converted to MRW ellipticities as described in Fig 12. The melting temperature ($T_m$) derived from the curves for concentrated TyrA proteins were ~47°C while the ones at lower concentration were about 52°C.

Fig 33 shows the change in ellipticity at 222 nm of TyrA proteins (at high and low protein concentrations) as a function of temperature. The data show that while the temperature-mediated unfolding curves of Δ80 CM-PD produced at the two protein concentrations were similar, the curves for FL CM-PD under these conditions were strikingly different. At a high protein concentration of FL CM-PD ellipticity values remained constant at the high temperatures mirroring the results for Δ80 CM-PD. Additionally, $T_m$ values for FL and Δ 80 CM-PD obtained from these far-UV CD measurements at high and low protein concentration ones were ~47°C and ~52°C, respectively. Thus the value obtained at the high protein concentration (~47°C) using far-UV CD measurements was in good agreement with the value (47°C) obtained for both proteins using FTIR spectroscopy.

The changes in the percent secondary structure component of TyrA proteins calculated from Dichroweb's CONTIN algorithm were plotted in Fig 34 as a further confirmation of the results by VT-FTIR experiments (Fig 27). The data showed that α-helix content decreased as temperature increases for both TyrA proteins. Interestingly, the α-helix in Δ 80 CM-PD commenced unfolding at a higher temperature (45°C) than in FL CM-PD. This is also consistent with the results from VT-FTIR spectroscopy (Fig 27 C and D). Unfortunately, the β-sheet structure could not be further separated into its intermolecular (protein aggregation) and intramolecular β-strand components. The major difference between the curves for FL and Δ 80 CM-PD was the marked reduction in unordered structure at 60°C.

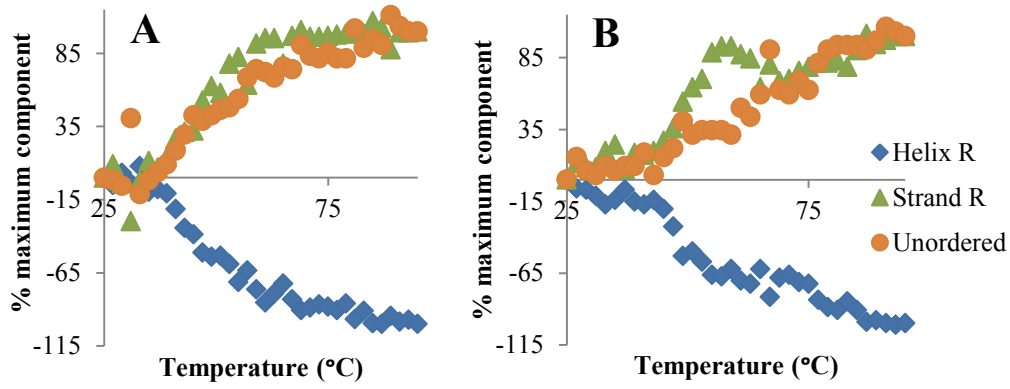



**Figure 34: The calculated changes in secondary structure component as a function of temperature for (A) FL and (B) Δ80 CM-PD**. The changes of percent secondary structural components were calculated using the relationship in Fig 27. $T_m$ was ~47 °C as measured by FTIR in identical experimental conditions. Helix R: α-Helix regular, Strand R: β- strand regular and, unordered: other structure that could not be defined as secondary structure.

# CHAPTER 4: DISCUSSION

The present study reports the first indepth unfolding studies performed on mono- and bifunctional forms of *H. influenzae* PD in the presence of temperature and the chemical denaturant, guanidine hydrochloride. The results were correlated with the crystal structure of monofunctional *H. influenzae PD* (Δ80 CM-PD, PDB:2PV7) and *E. coli* CM (PDB:1ECM) when applicable. Additionally, a comparison of the activity and stability of mono- and bifunctional forms of PD from the same organism has provided clues regarding the importance of the presence of the CM domain in TyrA structure. In an attempt to identify the contribution of the CM domain in the stabilization of the protein structure, thermal stability studies were extended to *H. influenzae* TyrA proteins with the active site ligands, $NAD^+$ and L-Tyr.

## Mono- and bifunctional *H. influenzae* PDs were stable and active dimers

Monofunctional *H. influenzae* PD (Δ80 CM-PD) was previously expressed and purified by Quashie (49) in order to facilitate comparison of the results with the crystal structure of this enzyme (PDB: 2PV7) (32). Our results are in agreement with Quashie (49) who showed that the presence or absence of the mutase domain does not markedly affect the catalytic function of PD. The kinetic parameters $k_{cat}$ and $K_m$ for the dehydrogenase reaction catalyzed by FL protein were within a factor of two of those determined for Δ80 CM-PD (Table 5). Quashie also confirmed, by size exclusion chromatography, that both FL and Δ80 CM-PD exist as dimers in solution (49) indicating that the absence of the mutase domain does not interfere with the dimerization of the PD domain. The crystal structure of *H. influenzae* Δ80 CM-PD also confirmed the dimeric nature of TyrA. The importance of the dimerization of each domain was probed by

studies on *E. coli* Gβ1-TyrA by Osuna *et al* (65). In their study, bifunctional *E. coli* TyrA (CM-PD) was engineered in order to replace the CM domain with the dimeric β1-domain of protein G (Gβ1-TyrA) (65). Site-directed mutagenesis on the Gβ1 domain of this bifunctional protein appeared to yield a stable and catalytically active PD enzyme provided that the quaternary structure of the β1-domain of protein G was retained allowing the dimerization of PD (65). It is important to note that all monofunctional TyrA proteins studied to date (*e.g. A. aeolicus* (20), *Synechocystis sp*. (26), *S. mutans* (22), and *M. tuberculosis* (24)) revealed an extensive C-terminal dimerization domain.

Our results build on the findings of others who have probed the importance of the polypeptide extension at PD's N-terminus to TyrA function and stability. Ganem and his co-workers engineered monofunctional *E. coli* PD fragments from bifunctional CM-PD encompassing residue 93-373 and 96-373 (47). These monofunctional PD variants, which contained only four to seven residues of the CM domain, showed reduced activity and poor stability in cold storage (47). In contrast, we have shown that *H. influenzae* Δ80 CM-PD, which possesses a longer extension (20 residues), possessed excellent PD activity (Table 5) which was retained over term cold storage (49). The diversity in activity/stability between the *E. coli* and *H. influenzae* engineered variants could be due to the differences in the primary sequence of these enzymes (they share about 60% amino acid sequence identity) and/or that the dimer interface was compromised in the *E. coli* constructs.

**The presence of the mutase domain makes predictable contributions to the global secondary and tertiary structures of TyrA**

Far-UV CD spectroscopy (Fig 12) showed that *H. influenzae* FL and Δ80 CM-PD proteins are highly α-helical. This is in agreement with the crystal structures of the independently expressed homologous CM domain of *E. coli* CM-PDT (Fig 5) and the *H. influenzae* Δ80 CM-PD (Fig 7). When the ellipticities were corrected for the differences in polypeptide length and concentration (MRW) (Fig 12), FL and Δ80 CM-PD showed a similar α-helical content. The small increases in ellipticity at 222 nm for FL CM-PD compared to the truncated variant likely reflect the addition of 80 residues from the highly α-helical mutase domain. FTIR spectra of *H. influenzae* FL protein also showed a slightly higher absorbance at the α-helix frequencies of the amide I' band compared to Δ80 CM-PD at 25°C due to the presence of the CM domain, while the absorbances at intramolecular β-sheet frequencies from the PD domain were comparable for the two proteins (Fig 13).

Our results using steady-state fluorescence spectroscopy showed ~10% higher emission for FL protein compared to Δ80 CM-PD upon excitation at 280 nm (Fig 14A) consistent with an additional Tyr residue, Tyr 47 (out of 13 Tyr residues per bifunctional monomer) in the CM domain. In contrast, the emission spectra for both TyrA proteins after excitation at 295 nm were almost identical (Fig 14B) as both proteins contain the same number of Trp residues (four per monomer). The emission maxima for both FL and Δ80 CM-PD after excitation at 280 nm and 295 nm are identical at ~332 nm because of the fluorescence energy transfer of Tyr to Trp. The value also indicates that globally, the Trp residues must be partially solvent exposed when compared with the $\lambda_{max}$ of 355 nm of the free Trp and Tyr analogs, NATA and NAYA (Fig 14). These results are supported by the crystal structure of Δ80 CM-PD which shows only two Trp and six Tyr per

monomer are solvent exposed (Fig 35). Importantly, the comparable $\lambda_{max}$ of the FL and Δ80 CM-PD proteins from fluorescence emission spectra suggests that the presence of CM domain does not alter the tertiary structure of the protein.

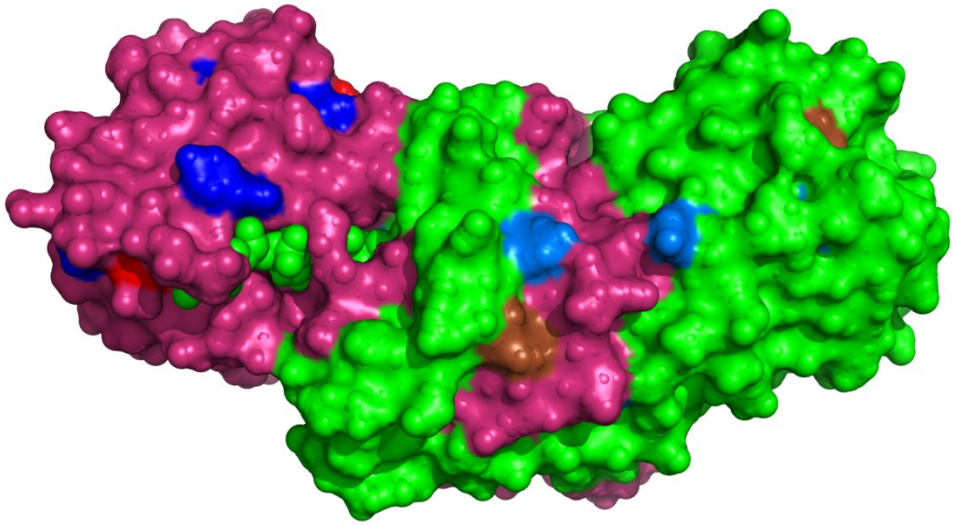


**Figure 35: Surface accessibility of Trp and Tyr in Δ80 CM-PD.** Surface exposed Trp (2 per monomer) are shown in red. Surface exposed Tyr (6 per monomer) residues are shown in blue. There are 4 Trp and 12 Tyr per monomer. One monomer of the dimer is in green and the other is in pink. Fluorophores in the other monomers are in different shade of red and blue. This was generated using Pymol, PDB: 2PV7.

## Stability and unfolding of TyrA

There have been very few detailed studies reported on the stability of TyrA proteins to temperature and/or chaotropic agents. Those studies to date have centered on *A. aeolicus* monofunctional PD and *E. coli* CM-PD performed in the Turnbull laboratory (20, 45, 98). Since the structure and activity of *H. influenzae* TyrA was not markedly perturbed under ambient conditions, as previously outlined, we set out to determine if VT and/or the presence of the chaotrope Gdn-HCl could reveal any differences in the stability of FL CM-PD and the Δ80 deletion variant. Chemical and thermal denaturations

of these proteins were monitored by far-UV CD, intrinsic fluorescence emission and FTIR spectroscopies.

**The unfolding of *H. influenza*e TyrA occurs via a metastable intermediate**

Our results show that both FL and Δ80 CM-PDs have similar global stabilities in the presence of Gdn-HCl. Both proteins commenced unfolding at the same concentration of Gdn-HCl (1.5 M) and were completely denatured by 4.2 M Gdn-HCl (Fig 16). Additionally, an identical $D_{1/2}$ value of 2.6 M Gdn-HCl was calculated for both proteins based on Gdn-HCl-mediated changes in fluorescence $\lambda_{max}$ emission (Fig 22). The same spectroscopic measurements clearly showed, however, that the unfolding pathways of *H. influenzae* FL and Δ80 CM-PDs are complex and are not identical for the two proteins. Both proteins showed a multi-state complex unfolding pattern that was punctuated by an insoluble intermediate in the presence of Gdn-HCl (Fig 16). FL CM-PD, however, showed much more pronounced visible turbidity at 1.5 M-2.4 M Gdn-HCl in two different buffer systems (PBS/glycerol and $K_2HPO_4$). Protein precipitation/ aggregation at specific concentrations of Gdn-HCl was also verified by UV-Visible spectroscopy which is sensitive to light scattering by particles (Fig 17). The formation of the insoluble aggregates of FL CM-PD also appeared to be time-dependent. Over most concentrations of Gdn-HCl the unfolding reactions reached equilibrium in less than 2 h but between 1.5 M and 2.5 M Gdn-HCl precipitation increased over time (Fig 18). Additionally, the aggregation appeared irreversible once established. This is illustrated in Fig 19 which showed that protein incubated at 1.5 M and 2.5M Gdn-HCl for 24 h was not refolded during dialysis, while protein samples at other concentrations of denaturant refolded after

dialysis. Together these results reinforce the importance of kinetics effects in protein aggregation.

Aggregation occurs when the balance between the net conformational stability of a protein (mainly from hydrophobic interaction, hydrogen bonding and van der Waals' force) and a protein's conformational entropy is disrupted (100). The thermodynamic stability of the native protein conformation is generally not much larger (5-20 kcal/mole) than the unfolded conformation, thus, small perturbations in the environment, *i.e.* by Gdn-HCl or increased temperature, can disturb the stability of the protein and promote its unfolding (101). Studies on *A. aeolicus* PD and its truncated variant, Δ19PD (20, 98), tryptophan synthase from *E. coli* (102), and methionine adenosyltransferase from *Methanococcus jannaschii* (103) have shown irreversible aggregation and/or multistate, complex Gdn-HCl-induced unfolding pathways. In order to analyze the aggregation pathways, the framework of Lumry-Eyring is often cited (104) which is described in the scheme in Fig 36 and in the caption in Fig 37.

$$\text{(Step 1)} \quad N \leftrightarrow TS^* \rightarrow A_I$$

$$\text{(Step 2)} \quad A_I + A_m \rightarrow A_{m+I}$$

**Figure 36: Lumry-Eyring framework of protein aggregation**. N represents the native protein, TS* is a transition state preceding the formation of an aggregation intermediate $A_I$, $A_m$ is an aggregate while $A_{m+I}$ is the aggregate with m+I protein molecules.

The model in Fig 36 describes the rate of formation of aggregates as a combination of thermodynamic and kinetic factors. Free energy is required to first form the TS* which will lead to $A_I$ that can only propagate by association with other $A_m$ to form $A_{m+I}$. The

rate-limiting step, therefore, could be either the formation of TS* (unimolecular, step 1) or the collision frequency (bimolecular, step 2).

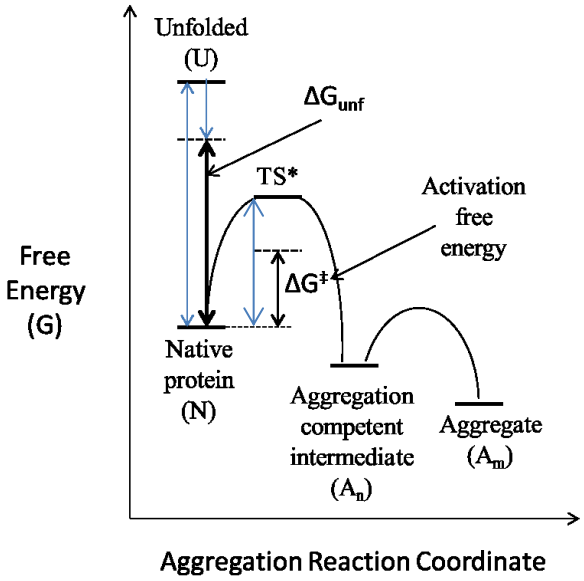



**Figure 37: Schematic reaction coordinate diagram of protein aggregation in the presence of Gdn-HCl**. Curved lines represent the kinetic energy barriers. Blue arrows are the energy required in the absence of Gdn-HCl. Note that the activation free energy required to form TS* and to unfold the protein is significantly decreased in the presence of Gdn-HCl. Adapted from Chi *et al* (97).

Fig 37 illustrates the reduced activation energy required to form the TS* from N in the presence of Gdn-HCl (black) which eventually leads to facile aggregation. According to Wyman linkage theory, binding of a ligand in a two-state equilibrium shifts the equilibrium toward the state with the greater binding (105). Therefore, the binding of Gdn-HCl shifts the thermodynamic stability between native and unfolded proteins in favor of the unfolded state. Gdn-HCl prefers to interact with protein surfaces thus the higher surface accessible area of the unfolded protein is favored (101).

Ma *et al* (106) reported that two of the parameters determining unfolding mechanisms are the size and amino acid composition of the protein. Thus, the more pronounced Gdn-HCl-induced precipitation observed for *H. influenzae* FL CM-PD may be due to its N-terminal extension (CM domain, 80 residues). Accordingly, the surface area of the unfolded state will be increased as well as the frequency of the TS* form of the unfolded protein which could interact and lead to the formation of aggregates. *Amino acid composition must also play an important role, however*. It is of interest to note that the homologous *E. coli* CM-PD which contains a CM domain and two PD subdomains ($NAD^+$ binding and dimerization) undergoes an apparent cooperative two-state unfolding pathway as monitored by far-UV CD spectroscopy (Fig 16) (45). Although *H. influenzae* and *E. coli* TyrA proteins are evolutionarily related they share only ~60% amino acid sequence identity. Additionally, the monofunctional *A. aeolicus* PD (311 amino acids), of similar size to *H. influenzae* PD (293 amino acids) but not homologous, was shown by far-UV CD and fluorescence experiments to proceed through several intermediates prior to fully unfolded (20, 98).

It is difficult to have a complete understanding of the complex unfolding pathways of *H. influenzae* TyrA proteins. However, there are reports where the loss of activity was tested to understand the nature of the intermediates. Gel filtration chromatography, sedimentation velocity experiments, and site-directed mutagenesis were also used to monitor the changes in quaternary structure in the course of intermediate/ aggregation formation (20, 102, 103). For example, size exclusion chromatography studies in combination with activity measurements of Δ19PD from *A. aeolicus* showed at a low concentration of Gdn-HCl the enzyme retained its dimeric structure and was active.

However, at a high concentration of Gdn-HCl the protein formed an oligomer with no enzymatic activity (20). *E. coli* CM-PD also remained dimeric at 3.5 M Gdn-HCl even after >90% of its secondary structure was lost (45). The analysis of the changes in structure of *H. influenzae* FL and Δ80 CM-PD and their enzymatic activities in the presence of denaturant will provide further insight into the relationship between the structure and activity of these proteins.

**Temperature-mediated unfolding pathways of *H. influenzae* TyrA proteins are similar but not identical and illustrate the strength of the dimerization domain**

Analysis of the thermal denaturation of FL and Δ80 CM-PDs by far-UV CD and FTIR spectroscopies showed curves consistent with irreversible, moderately cooperative temperature-induced unfolding. VT-CD measurements were performed at low protein concentrations (3.5 μM monomers in PBS/glycerol) and at high protein concentrations under conditions used for VT-FTIR experiments (~1 mM monomers, $K_2HPO_4$). These represent the first studies of its kind on a TyrA protein. $T_m$ values derived from these curves suggested that the midpoints of thermal unfolding of both proteins were similar; ~52°C from CD experiments using low protein concentration (Fig 23) and 47°C from far-UV CD (Fig 33) and FTIR measurements (Fig 27) using high protein concentrations. High temperature increases the thermal kinetic energy of the transition state of an unfolded protein which facilitates aggregation; *i.e.*, the probability a collision with other partially unfolded (or native) protein which triggers further unfolding (step 2, Fig 37) increases proportionally with temperature and protein concentration. As expected the $T_m$ values of the TyrA proteins were lower (47°C) at the higher protein concentrations.

At low protein concentrations, however, the changes in ellipticity monitored by far-UV CD indicated that FL and Δ80 CM-PD might undergo a different temperature-mediated unfolding process. After the transition, the ellipticity at 222 nm continued to decrease as the FL protein formed more insoluble aggregates (Fig 23). Δ80 CM-PD at 95°C retained a distinct minimum at 218 nm (Fig 24) indicating that the aggregate was composed of predominantly intermolecular β-sheets.

These differences were much less apparent at high protein concentrations. FTIR and CD experiments performed at high protein concentrations showed that both proteins formed β-aggregates at high temperatures (Fig 26). This also highlights the utility of conducting spectroscopic measurements under the same physiochemical conditions in order to draw more accurate conclusions. Also worth noting fluorescence emission confirmed that at 95°C both proteins retained compact aggregated structures (Fig 25). The emission maximum at 346 nm suggested that some Tyr and Trp residues were still partially buried.

FTIR is an exquisite technique for monitoring α-helices, intramolecular β-sheets, and intermolecular β- sheets through the amide I (I') bands, as well as the compactness of protein structure through H-D exchange associated with the amide II (II') bands. Accordingly, the unfolding of the main secondary structures of TyrA proteins and their aggregation as intermolecular β-sheets after heating was confirmed by VT-FTIR (Fig 26 and 27). Our results also indicated that both proteins are not highly compact structures as revealed by the low absorbance of the amide II' band (Fig 26, peak e). Nevertheless H-D exchange of the protein's amide groups is not complete unless the proteins are incubated for ~18 h in $D_2O$ prior to VT experiments. Additionally, the incomplete H/D exchange

within both TyrA proteins (Fig 26 and 27 A and B) appeared to influence the absorbance of amide I' bands of the α-helices which resulted in an unfolding sequence of: concomitant loss of intramolecular β-sheet and α-helix followed by the appearance of intermolecular β-sheet (aggregation) (Table 7). In contrast, the 2D correlation analysis from VT-FTIR spectra of both FL and Δ80 CM-PD after 24 h of incubation in $D_2O$ showed a different unfolding sequence: loss of intramolecular β-sheets followed by loss of α-helices which is accompanied by the appearance of intermolecular β-sheets (Table 8). From the crystal structure of *H. influenzae* Δ80 CM-PD, we can infer that TyrA unfolding commences from the N-terminal $NAD^+$ binding domain which is rich in intramolecular β-sheet structure (Fig 7). The helix bundle (14 intertwined helices) involved in dimerization within TyrA's C-terminal domain was more resistant to heat-induced unfolding implying that *H. influenzae* TyrA retains its dimeric structure while unfolding.

The presence of the mutase domain in the FL protein does not appear to offer protection toward thermal unfolding of the $NAD^+$ binding domain even though both domains are contiguous in the primary sequence of bifunctional TyrA. This conclusion stems from several observations: 1) ***both*** proteins' intramolecular β-sheets unfold before α-helices; 2) close inspection of the FTIR data show that the sequence of unfolding of the two proteins are similar but not identical and point to a ***destabilizing effect*** of the CM domain. The α-helices within the FL protein commenced unfolding ~6°C *prior* to those associated with Δ80 CM-PD (39°C compared to 45°C), while the intramolecular β-sheets of both proteins started to unfold at the same temperature (35°C) (Table 10). The α-helices in FL CM-PD appeared more thermally unstable compared to the αhelices in the

Δ80 CM-PD and the temperature-mediated transition less cooperative. Our results can be compared to trends observed in other enzymes in the common pathway. The engineered *E. coli* mutase domain (mini-mutase) from CM-PDT showed a $T_m$ of 76°C by VT far-UV CD measurements which is significantly higher than that reported for bifunctional *E. coli* CM-PD (57°C) (20). Although these $T_m$ experiments were conducted under slightly different conditions (PPS buffer, pH 7.5 for CM-PD and PBS buffer, pH 6.5 for the "mini-mutase") the results suggest that the addition of a CM domain to PD may not necessarily stabilize the structure of TyrA.

The binding of the active site ligands $NAD^+$ and L-Tyr appeared to alter the patterns of unfolding of α-helices and intramolecular β-sheets, reflecting a concomitant loss of both types of secondary structures. This was observed for both FL and Δ80 CM-PD (Fig 30). For the liganded TyrA proteins, the unfolding of α-helices commenced 4 and 10°C lower than for the unliganded FL and Δ80 CM-PD proteins, respectively, yet they showed an identical $T_m$ of 45°C; the $T_m$ values of the intramolecular sheets were ~3°C and 5°C higher than for the unliganded FL and Δ80 proteins, respectively. Therefore, binding of the ligands not only caused α-helices to unfold less cooperatively and at a lower temperature but also stabilized the intramolecular β-sheets. The presence of the CM domain does not appear to stabilize the secondary structure of FL CM-PD which might be attributed to the fact that $NAD^+$ and L-Tyr bind to the PD portion of the polypeptide chain (Fig 7). This is in keeping with the crystal structure which shows $NAD^+$ binding to the β-sheet rich Rossmann fold motif. Globally both TyrA proteins are more thermally stable in the presence of $NAD^+$ and L-Tyr as reflected in the 4°C increase in $T_m$ values of the intermolecular β-sheets (Fig 30). Direct binding studies using

isothermal titration microcalorimetry could help correlate the changes in $T_m$ with the strength of interaction of these active site ligands and the enzyme.

## CHAPTER 5: PERSPECTIVES AND FUTURE DIRECTION

In summary, we have investigated the structural stability of engineered mono- and bifunctional *H. influenzae* TyrA proteins in the presence of biochemical and biophysical denaturants. The work presented in this thesis represents some of the first in-solution studies on TyrA proteins from *H. influenzae* probing the stability of the two forms of the enzyme. Variable temperature experiment monitored by far-UV circular dichroism using a high concentration of protein with μm pathlength cells was the first attempt to align the experimental parameters from CD with those from FTIR spectroscopy.

Surprisingly, the presence of the CM domain does not appear to impart additional global stabilization to PD. In fact, the presence of the CM domain appears to help promote the formation of insoluble intermediate(s) in the Gdn-HCl-unfolding pathway. The sequence of events of the temperature-mediated unfolding pathway is also somewhat different for the two proteins. Probed by 1D and 2D FTIR analyzes and by far-UV CD under comparable conditions together suggested that the intramolecular β-sheets of the two proteins commenced unfolding first and at the same temperature, while the α-helices unfolded next concomitant with β-aggregation, although the helices of Δ80 CM-PD appeared much more resistant to thermal denaturation than those found within the FL protein. It may be that Δ80 CM-PD adopts a more compact global structure; analytical ultracentrifugation experiments may provide further insight.

The unfolding of α-helices (one of the terminal steps of the unfolding process) also highlights the strength of the dimerization interface contributed by the extensive helices at C-terminal domain within each monomer. Further analysis of the FTIR spectra will be required to probe the specific secondary structures that are involved in

dimerization in the CM and PD domains. Preliminary analysis by enhanced peak resolution has allowed the assignment of a new band for α-helices (1657 $cm^{-1}$, Appendix 4) which might be specific for the PD domain. Further analysis by enhanced peak resolution will be required to determine if we can assign an α-helical component contributed solely by the CM domain

In order to investigate the intermediates in the course of the chemically-induced unfolding, urea could be used as a denaturant; urea, unlike the guanidinium ion is uncharged and could interact differently with TyrA. To further investigate intermediate structures in the presence of Gdn-HCl or urea, 8-anilinonaphthalene-1-sultonate (ANS) could be used for the quantitative measurement of partially unfolded but soluble proteins. Lastly, determination of kinetic parameters in variable temperature as well as in the presence of chemical denaturants will help correlate the structures of the CM and PD domains with the function of the *H. influenzae* TyrA proteins.

## REFERENCES


1. Sprenger, GA, (2006). Aromatic Amino Acids. In Wendisch VF (Ed.), *Amino Acids Biosynthesis- Pathways, Regulation and Metabolic Engineering* (93). Heidelberg, Germany: Springer.
2. Barlett PA, Nakagawa Y, Johnson CR, Reich SH and Luis A, (1988) Chorismate Mutase Inhibitors: Synthesis and evaluation of some potential transition state analogs. *J Org Chem* 53:3195-3210.
3. Herrmann KM, and Weaver LM (1999) The shikimate pathway. *Annu Rev Plant Physiol Plant Mol Biol* 50:473–503.
4. Schonbrunn E, Eschenburg S, Shuttleworth WA, Schloss JV, Amrhein N, Evans JNS, and Kabsch W, (2001) Interaction of the herbicide glyphosate with its target enzyme 5-enolpyruvylshikimate 3-phosphate synthase in atomic detail. *Proc Nat Acad Sci* 98: 1376-1380.
5. Coggins JR, Abell C, Evans LB, Frederickson M, Robinsin DA, Roszak AW and Lapthron AP, (2003) Experiences with the shikimate pathway enzymes as targets for rational drug design. *Biochem Soc Trans* 31: 548-552.
6. Seetharam G, and Saville B, (2002) L-Dopa production from tyrosinase immobilized on zeolite. *Enz Microbial Technol* 31: 747-753.
7. Ikram UH, and Ali A, (2002) Microbiological transformation of L-tyrosine to 3,4-dihydroxyphenyl L-alanine (L-Dopa) by mutant strain of *Aspergillus oryzae* UV-7. *Curr Microbiol* 45:88-93.
8. Krama M, Bongaerts J, Bovenberg R, Kremer S, Muller U, Orf S, Wubbolts M, and Raeven L, (2003) Metabolic engineering for microbial production of shikimic acid. *Metab Eng* 5:277-283.
9. Cabrera VN, Martinez A, Pinero S, Lagunas VH, Tinoco R, de Anda R, Vazquez DR, Bolivar F, and Gosset G, (2006) Expression of the melA gene from Rhizobium etli CFN42 in *Escherichia coli* and characterization of the encoded tyrosine. *Enz Microbial Technol* 38:772-779.
10. Bourke SL, and Kohn J, (2003) Polymers derived from the amino acids L-tyrosine: polycarbonates, polyarylates and copolymers with poly (ethylene glycol). *Adv Drug Delivery Rev* 55:447-466.

11. Chavez-Bejar MI, Laura AR, Lopez H, Hernandez-Chavez G, Martinez A, Ramirez OT, Bolivar F, and Gosset G, (2008) Metabolic engineering of *Escherichia coli* for L-tyrosine production by expression of genes coding for the chorismate mutase domain of the native chorismate mutase-prephenate dehydratase and a cyclohexadienyl dehydrogenase from *Zymomonas mobilis*. *Appl Environ Microbiol* 74: 3284-3290.
12. Chavez-Bejar MI, Baez-Viveros JL, Martinez A, Bolivar F, and Gosset G, (2012) Biotechnological production of L-tyrosine and derived compounds. *Process Biochem* 47:1017-1026.
13. Sternmark SL, Pierson DL, and Jensen RA, (1974) Blue-Green bacteria synthesize L-Tyrosine by the pretyrosine pathway. *Nature* 247: 290-292.
14. Zamir LO, Jensen RA, Arison B, Douglas A, Albers-Schonberg G, and Bowen JR, (1980) Structure of arogenate (pretyrosine), an amino acid intermediate of aromatic biosynthesis. *J Am Chem Soc* 102: 4499-4505.
15. Bonner CA, Disz T, Hwang K, Song J, Vonstein V, Overbeek R, and Jensen RA, (2008) Cohesion group approach for evolutionary analysis of TyrA, a protein family with wide-ranging substrate specificities. *Microbiol Mol Biol Rev* 72:13-53.
16. Xie G, Keyhani NO, Bonner CA, and Jensen RA, (2003) Ancient origin of the tryptophan operon and the dynamics of evolutionary change. *Microbiol Mol Biol Rev* 67:303-342.
17. Song J, Bonner CA, Wolinsky M, and Jensen RA, (2005) The TyrA family of aromatic-pathway dehydrogenases in phylogenetic context. *BMC Biology* 3:13-43.
18. Wierenga RK, Terpstra P, and Hol WGJ, (1986) Prediction of the occurrence of the ADP-binding β α β-fold in proteins, using an amino-acid sequence fingerprint. *J Mol Biol* 187:101-107.
19. Lim S, Springstead JR, Yu M, Bartkowski W, Schroder I, and Monbouquette HG, (2009) Characterization of a key trifunctional enzyme for aromatic amino acid biosynthesis in *Archaeoglobus fulgidus*. *Extremophiles* 13:191-198.
20. Bonvin J, Aponte RA, Marcantonio M, Singh S, Christendat D and Turnbull JL, (2006) Biochemical characterization of prephenate dehydrogenase from the hyperthermophilic bacterium *Aquifex aeolicus*. *Protein Sci* 15:1417-1432.

21. Sun W, Singh S, Zhang R, Turnbull JL, and Christendat D, (2006) Crystal structure of prephenate dehydrogenase from *Aquifex aeolicus*. Insights into the catalytic mechanism. *J Biol Chem* 281:12919-12928.
22. Ku H, Park S, Yang I and Kim S (2010) Expression and functional characterization of prephenate dehydrogenase from *streptococcus mutans*. *Process Biochem* 45:607-612.
23. Champney WS, and RA Jensen, (1970) The enzymology of prephenate dehydrogenase in *Bacilus subtilis*. *J Biol Chem* 245:3763-3770.
24. Yang Q, Yu K, Yan L, Li Y, Chen C and Li X, (2011) Structural view of the regulatory subunit of aspartate kinase from *Mycobacterium tuberculosis*. *Protein cell* 2: 745-754.
25. Xu S, Yang Y, Jin R, Zhang M, and Wang H (2006) Purification and characterization of a functionally active Mycobacterium tuberculosis prephenate dehydrogenase. *Protein Expr Purif* 49:151-158.
26. Legrand P, Dumas R, Seux M, Rippert P, Ravelli R, Ferrer JL and Matringe M, (2006) Biochemical characterization and crystal structure of *Synechocystis* arogenate dehydrogenase provide insights into catalytic reaction. *Structure* 14:767-776.
27. Gaines CG, Byng GS, Whitaker RJ and Jensen RA, (1982) The effect of wastewater discharge on biomass production and nutrient content of *Cyperus papyrus* and *Miscanthidium violaceum* in the Nakivubo wetland, Kampala, Uganda. *Plantae* 156:233-240.
28. Zhao G, Xia T, Ingram LO, Jensen RA, (1993) An allosterically insensitive class of cyclohexadienyl dehydrogenase from *Zymomonas mobilis*. *Eur J Biochem* 212:157-165.
29. Fischer RS, Zhao G, and Jensen RA, (1991) Cloning, sequencing and expression of the P-protein gene (pheA) of *Pseudomonas stutzero* in *Escherichia coli*: implications for evolutionary relationships in phenylalanine biosynthesis. *J Gen Microbiol* 137:1293-1301.
30. Lee AY, Karplus PA, Ganem B, and Clardy J, (1995) Atomic-structure of the buried catalytic pocket of *Escherichia coli* chorismate mutase. *J Am Chem Soc* 117: 3627-3628.

31. Xia T, Zhao G, Fischer RS, and Jensen RA, (1992) A monofunctional prephenate dehydrogenase created by cleavage of the 5' 109 bp of the tyrA gene from *Erwinia herbicola*. *J Gen Microbiol* 138:1309-1316.
32. Chiu HJ, Abdubek P, Astakhova T, Axelrod HL, Carlton D, Clayton T, Wilson IA, (2010) The structure of *Haemophilus influenzae* prephenate dehydrogenase suggests unique features of bifunctional TyrA enzymes. *Acta Crystallographica* F66:1317-1325.
33. Dun BQ, Wang XJ, Lu W, Zhao ZL, Hou SN, Zhang BM, Li GY, Evans TC Jr, Xu MQ and Lin M, (2007) Reconsititution of glyphosate resistance from a split 5-enolpyruvyl shikimate-3-phosphate synthase gene in *Escherichia coli* and transgenic tobacco. *Appl Environ Microbiol* 73:7997-8000.
34. Byng GS, Berry A, and Jensen RA, (1985) Evolutionary implications of features of aromatic amino acid biosynthesis in the genus *Acinetobacter*. *Arch Microbiol* 143:122-129.
35. Sampathkumar P and Morrison JF, (1982) Chorismate mutase-prephenate dehydrogenase from *Escherichia coli*. Purification and properties of the bifunctional enzyme. *Biochim Biophys Acta* 702:204-211.
36. Hassounah, S, (2009) MSc Thesis. Concordia University, Montréal.
37. Hudson GS, and Davidson BE, (1984) Nucleotide sequence and transcription of the phenylalanine and tyrosine operons of *Escherichia coli* K12. *J Mol Biol* 180:1023-1051.
38. Koch GL, Shaw DC and Gibson F, (1972) Studies on the relationship between the active sites of chorismate mutase-prephenate dehydrogenase from *Escherichia coli* or *Aerobacter aerogenes*. *Biochim Biophys Acta* 258:719-730.
39. Heyde E, (1979) Chorismate mutase-prephenate dehydrogenase from *Aerobacter aerogenes*: evidence that the two reactions occur at one active site. *Biochemistry* 18: 2766-2775.
40. Turnbull J and Morrison JF, (1990) Chorismate mutase-prephenate dehydrogenase from *Escherichia coli*. 2. Evidence for two different active sites. *Biochemistry* 29:10255-10261.

41. Turnbull J, Cleland WW and Morrison JF, (1991) pH dependency of the reactions catalyzed by chorismate mutase-prephenate dehydrogenase from *Escherichia coli*. *Biochemistry* 30:7777-7782.
42. Christendat D and Turnbull JL, (1996) Identification of active site residues of chorismate mutase-prephenate dehydrogenase from *Escherichia coli*. *Biochemistry* 35:4468-4479.
43. Vincent S, Chen S, Wilson D, and Ganem B, (2002) Probing the overlap of chorismate mutase and prephenate dehydrogenase sites in the *Escherichia coli* T-protein: a dehydrogenase-selective inhibitor. *Bioorganic & Medicinal Chemistry Letters* 12: 929-931.
44. Christopherson RI, Heyde E and Morrison JF, (1983) Chorismate mutase-prephenate dehydrogenase from *Escherichia coli*: spatial relationship of the mutase and dehydrogenase sites. *Biochemistry* 22:1650-1656.
45. Christendat D and Turnbull JL, (1999) Identifying groups involved in the binding of prephenate to prephenate dehydrogenase from *Escherichia coli*. *Biochemistry* 38: 4782-4793.
46. Ismoyo F, (2000) PhD Thesis. McGill University, Montréal.
47. Chen S, Vincent S, Wilson DB and Ganem B, (2003) Mapping of chorismate mutase and prephenate dehydrogenase domains in the *Escherichia coli* T-protein. *Eur J Biochem* 270:757-763.
48. Duggleby RG, Sneddon MK, Morrison JF, (1978) Chorismate mutase-prephenate dehydratase from *Escherichia coli*: active sites of a bifunctional enzyme. *Biochemistry* 17:1548-1554.
49. Quashie PK, (2010) MSc Thesis. Concordia University, Montréal.
50. Hur S and Bruice TC, (2002) The mechanism of catalysis of the chorismate to prephenate reaction by the *Escherichia coli* mutase enzyme. *Proc Nat Acad Sci* 99:1176-1181.
51. Copley SD, and Knowles JR, (1987) The conformational equilibrium of chorismate in solution: implications for the mechanism of the non-enzymic and the enzyme-catalyzed rearrangement of chorismate to prephenate. *J Am Chem Soc* 109: 5008-5013.

52. Guildford WJ, Copley SD and Knowles JR, (1987) On the mechanism of the chorismate mutase reaction. *J Am Chem Soc* 109: 5013-5019.
53. Zhang X, Zhang X, and Bruice TC, (2005) A definitive mechanism for chorismate mutase. *Biochemistry* 44: 10443-10448.
54. Christendat D, Saridakis VC, and Turnbull JL, (1998) Use of site-directed mutagenesis to identify residues specific for each reaction catalyzed by chorismate mutase-prephenate dehydrogenase from *Escherichia coli*. *Biochemistry* 37:15703-15712.
55. Chook YM, Ke H, and Lipscomb WN, (1993) Crystal structures of the monofunctional chorismate mutase from *Bacillus subtilis* and its complex with a transition state analog. *Proc Nat Acad Sci* 90:8600-8603.
56. Liu DR, Pastor RM and Schultz PG, (1996) Analysis of active site residues in Escherichia coli chorismate mutase by site-directed mutagenesis. *J Am Chem Soc* 118:1789-1790.
57. Zhang S, Kongsaeree P, Clardy J, Wilson DB, and Ganem B, (1996) Site-directed mutagenesis of monofunctional chorismate mutase engineered from the *E. coli P-protein. Bioorg Med Chem* 4:1015-1020.
58. Kast P, Asif-Ullah M, Jiang N and Hilvert D, (1996) Exploring the active site of chorismate mutase by combinatorial mutagenesis and selection: the importance of electrostatic catalysis. *Proc Natl Acad Sci* 93:5043-5048.
59. Hermes JD, Tipton PA., Fisher MA, O'Leary MH., Morrison JF and Cleland WW, (1984) Mechanisms of enzymatic and acid-catalyzed decarboxylations of prephenate. *Biochemistry* 23: 6263-6275.
60. Hudson GS, Howlett GJ, and Davidson BE, (1983) The binding of tyrosine and $NAD^+$ to chorismate mutase/prephenate dehydrogenase from *Escherichia coli* K12 and the effects of these ligands on the activity and self-association of the enzyme. Analysis in terms of a model. *J Biol Chem* 258:3114-3120.
61. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ and Higgins DG, (2007) *ClustalW version 2. Bioinformatics* 23: 2947-2948.

62. Konstantin Okonechnikov, Olga Golosova, Mikhail Fursov, the UGENE team, (2012) Unipro UGENE: a unified bioinformatics toolkit, *Bioinformatics* 28:1166-1167.
63. Sun W, Shahinas D, Bonvin J, Hou W, Kimber MS, Turnbull J and Christendat D, (2009) The crystal structure of *Aquifex aeolicus* prephenate dehydrogenase reveals the mode of tyrosine inhibition. *J Biol Chem* 284:13223-13232.
64. Delano WL, (2002) The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.
65. Osuna J, Flores H, Saab-Rincon G, (2012) The β1 domain of protein G can replace the chorismate mutase foamin of the T-protein. *FEBS letters*, 586:466-471.
66. Shirley BA (Eds.). (1995). *Methods in Molecular Biology (Vols. 40). Protein stability and Folding: Theory and Practice*. Totowa, NJ: Humana Press Inc.
67. England JL, and Haran G, (2011) Role of solvation effects in protein denaturation: from thermodynamics to single molecules and back. *Ann Rev Phys Chem* 62: 257-277.
68. Kelly SM, Jess TJ, and Price NC, (2005) How to study proteins by circular dichroism. *Biochim Biophys Acta* 1751:119-39.
69. Kelly SM, and Price NC, (2000) The use of circular dichroism in the investigation of protein structure and function. *Curr Protein Pept Sci* 1: 349-384.
70. Lobley A, Whitmore L, and Wallace BA, (2001) DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra. *Bioinformatics* 18:211-212.
71. Whitmore L, and Wallace BA, (2007) Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers* 89:392-400.
72. Sreerama N, and Woody RW, (2000) Estimation of protein secondary structure from circular dichrosim spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem* 287: 252-260.
73. Lakowicz JR, (2006) *Principles of Fluorescence Spectroscopy* (3rd Ed.). New York, NY: Springer.
74. Haque T, (2001) MSc Thesis. McGill University, Montréal.
75. Stuart BH, (2004), *Infrared Spectroscopy: Fundamentals and Applications*. New York, NY: Wiley.

76. Noda I, (1993) Generalized two-dimensional correlation method applicable to infrared, raman, and other types of spectroscopy. *Appl Spectrosc* 47: 1329-1336.
77. Rieger CE, and Turnbull J, (1996) Small scale biosynthesis and purification of gram quantities of chorismic acid. *Prep Biochem Biotechnol* 26:67-76.
78. Dudzinski PK and Morrison JF, (1976) The preparation and purification of sodium prephenate. *Prep Biochem Biotechnol* 6:113-121.
79. Dawson CMR, Elliott CD, Elliott HW, and Jones MK, (1986), *In Data for Biochemical Research.* Oxford Science Publications, Clarendon Press. Oxford.
80. Turnbull J, (1988) PhD Thesis. Australian National University, Canberra.
81. Chung CT, Niemela SL, and Miller RH, (1989) One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. *Proc Nat Acad Sci* 86: 2172-2175.
82. Sambrook J, Russell D, (2001) *Molecular Cloning,* 3rd ed., Cold Spring Harbor Laboratory Press, Cold Spring, New York.
83. NCCLS Approved Standard ACS-1 (1979) National Committee for Clinical Laboratory Standards: Villanova, PA.
84. Swiss Institute of Bioinformatics (SIB), ExPASy (Expert Protein Analysis System), http://ca.expasy.org.
85. Nozaki Y, (1972) *In Methods in Enzymology*, vol 26 (Hirs, C.H.W. and Timasheff, S.N., eds.), Academic, New York, pp 43-50.
86. Lehman JW, (1998) Operational Organic Chemistry: a problem-solving approach to the laboratory course, 3rd ed., Prentice Hall, New York.
87. Gallagher W, (2007) FTIR Analysis of Protein Structure, http://www.chem.uwec.edu/Chem455_S05/Pages/Manuals/FTIR_of_proteins.pdf
88. Janes RW, (2008) Reference datasets for protein circular dichroism and synchrotron radiation circular dichroism spectroscopic analyses. In *Modern Techniques for Circular Dichrosim and Synchrotron Radiation Circular Dichroism Spectroscopy* (Wallace BA and Janes RW, Eds.).
89. Amiri A, (2012) Honors Thesis. Concordia University, Montréal.

90. Zelent B, Kusba J, Gryczynski I, Johnson ML, and Lakowicz JR, (1998) Time-resolved and steady-state fluorescence quenching of N-acetyl-L-tryptophanamide by acrylamide and iodide. *Biophy Chem* 73:53-75.
91. Tsaprailis G, Chan DWS, and English AM, (1998) Conformational states in denaturants of cytochrome c and horseradish peroxidases examined by fluorescence and circular dichroism. *Biochemistry* 37: 2004-2016.
92. Bonvin J, (2008) PhD Thesis. Concordia University, Montréal.
93. Robbins RJ, Fleming GR, Beddard GS, Robbinson GW, Thistlethwaite PJ, and Woolfe GJ, (1980) Photophysics of Aqueous Tryptophan: pH and Temperature Effects. *J Am Chem Soc* 102: 6271-6279.
94. Tsaprailis G, (1997) PhD Thesis. Concordia University, Montréal.
95. Surewicz WK, Mantsch HH, and Champman D, (1993) Determination of protein secondary structure by fourier transform infrared spectroscopy: a critical assessment. *Biochemistry* 32: 389-394.
96. Carta R, Tola G, (1996) Solubilities of L-cystine, L-tyrosine, L-leucine, and glycine in aqueous solutions at various pHs and NaCl concentrations. *J Chem Eng Data*. 41: 414-417.
97. Turnbull J, Cleland WW and Morrison JF, (1990) Chorismate mutase-prephenate dehydrogenase from *Escherichia coli*. 1. Kinetic characterization of the dehydrogenase reaction by use of alternative substrates. *Biochemistry* 29:10245-10254.
98. Aponte, (2002) MSc Thesis. Concordia University, Montréal.
99. Dong AC, Prestrelski SJ, Allison SD, and Carpenter JF, (1995) Infrared spectroscopic studies of lyophilization-induced and temperature-induced protein aggregation. *J Pharm Sci*. 84: 415-424.
100. Dill KA, (1990) Dominant forces in protein folding. *Biochemistry* 29:7133-7155.
101. Chi EY, Krishnan S, Randolph TW, Carpenter JF, (2003) Physical stability of proteins in aqueous solution: mechanism and driving forces in nonnative protein aggregation. *Pharm Res* 20:1325-1336.
102. Miles EW, Yutani K, and Ogasahara K, (1982) Guanidine hydrochloride induced unfolding of the α subunit of tryptophan synthase and of the two α proteolytic

fragments: evidence for stepwise unfolding of the two α domains. *Biochemistry* 21:2586-2592.

103. Garrido F, Taylor JC, Alfonso C, Markham GD and Pajares MA, (2012) Structural basis for the stability of a thermophilic methionine adenosyltransferase against guanidinium chloride. *Amino Acids* 42: 631-637.

104. Lumry R, and Eyring H, (1954) Conformation changes of proteins. *J Phys Chem* 58:110-120.

105. Timasheff SN, (1998) Control of protein stability and reactions by weakly interacting cosolvents: the simplicity of the complicated. *Adv Protein Chem* 51:355-432.

106. Ma BG, Chen LL, Zhang HY, (2007) What determines protein folding type? An investigation of intrinsic structural properties and its implications for understanding folding mechanisms. *J Mol Biol* 370: 439-448.

107. Vamvaca K, Butz M, Walter KU, Taylor SV, Hilvert D, (2005) Simultaneous optimization of enzyme activity and quaternary structure by directed evolution. *Protein Sci* 14:2103-2114.

108. Ross PD, Subramanian S, (1981) Thermodynamics of protein association reactions: Forces contributing to stability. *Biochemistry* 20: 3096-3012.

## Appendix 1

| | | | | |
|---|---|---|---|---|
| **1** | **MSFMEALKDL** | **RSEIDSLDRE** | **LIQLFAKRLE** | **LVSQVGKVKH** |
| **41** | **QHGLPIYAPE** | **REIAMLQARR** | **LEAEKAGISA** | **DLIEDVLRRF** |
| **81** | **MRESYANENQ** | **FGFKTINSDI** | **HKIVIVGGYG** | **KLGGLFARYL** |
| **121** | **RASGYPISIL** | **DREDWAVAES** | **ILANADVVIV** | **SVPINLTLET** |
| **161** | **IERLKPYLTE** | **NMLLADLTSV** | **KREPLAKMLE** | **VHTGAVLGLH** |
| **201** | **PMFGADIASM** | **AKQVVVRCDG** | **RFPERYEWLL** | **EQIQIWGAKI** |
| **241** | **YQTNATEHDH** | **NMTYIQALRH** | **FSTFANGLHL** | **SKQPINLANL** |
| **281** | **LALSSPIYRL** | **ELAMIGRLFA** | **QDAELYADII** | **MDKSENLAVI** |
| **321** | **ETLKQTYDEA** | **LTFFENNDRQ** | **GFIDAFHKVR** | **DWFGDYSEQF** |
| **361** | **LKESRQLLQQ** | **ANDLKQG** | | |

**Figure 38:** ***H. influenzae*** **TyrA sequence.** CM (1-100 residues) and PD (101 to 377 residues) domains are colored in blue and in red, respectively. Sequence of monofunctional PD expressed is underlined (81-377 residues).

## Appendix 2A

Synchronous maps

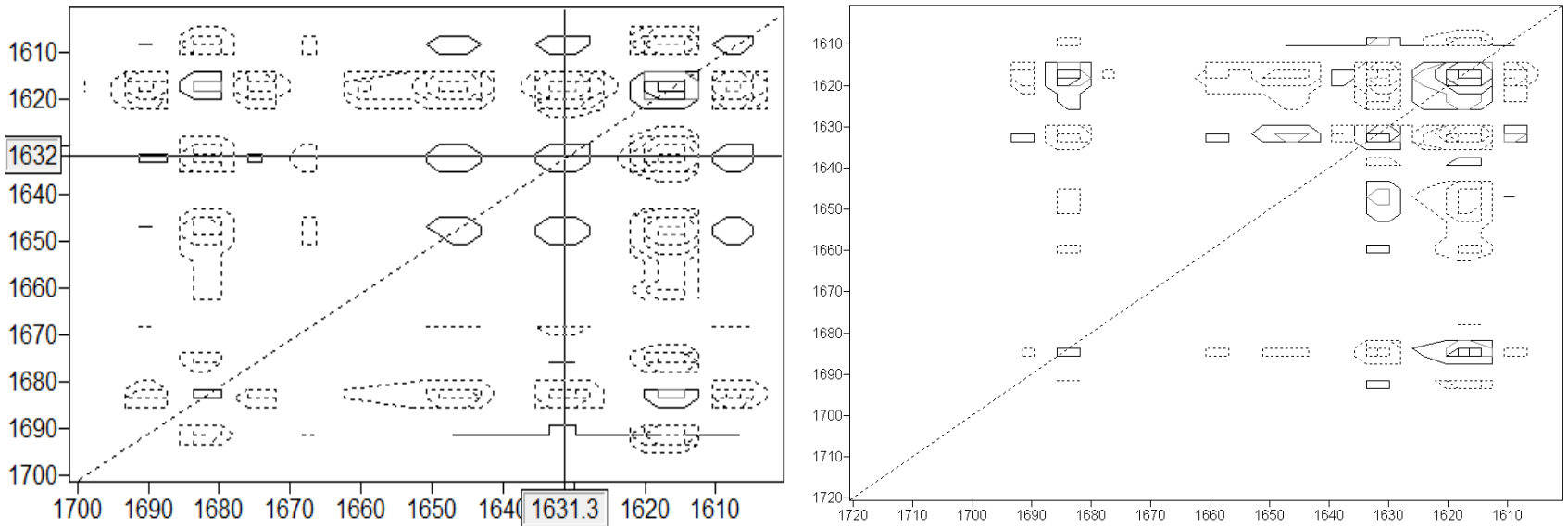



**Figure 39: Synchronous maps for FL (left) and Δ80 CM-PD (right) without 24 h incubation in $D_2O$.**

## Appendix 2B

Synchronous maps

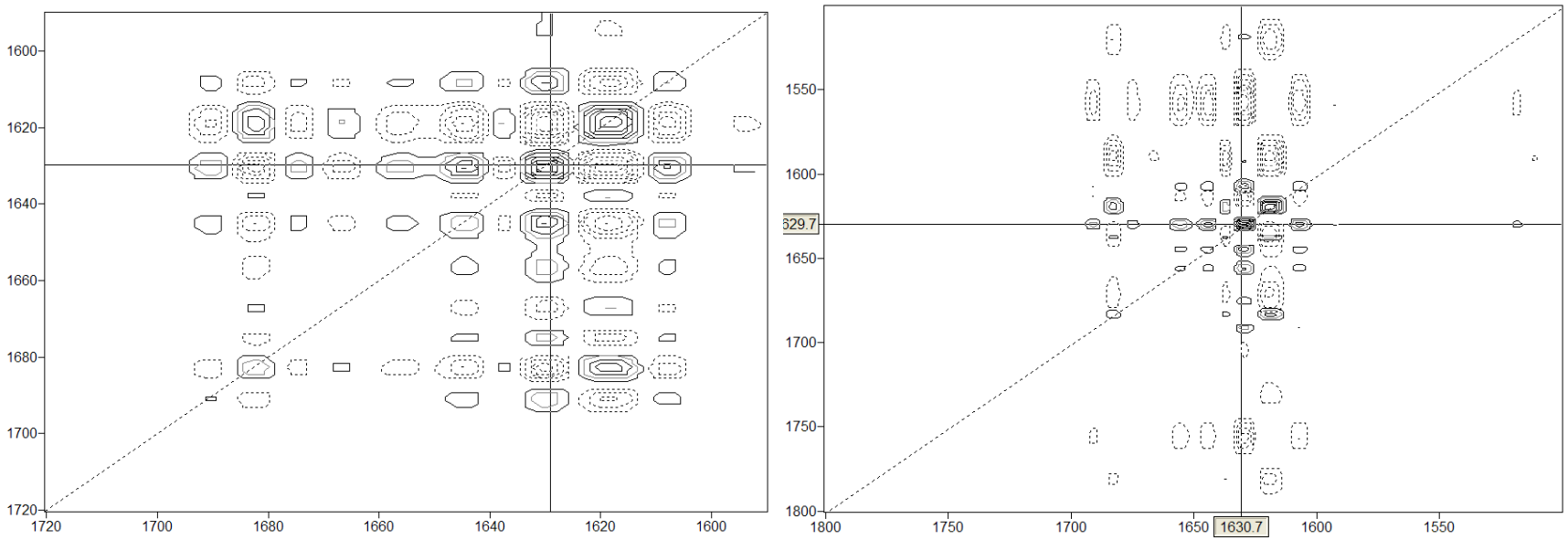
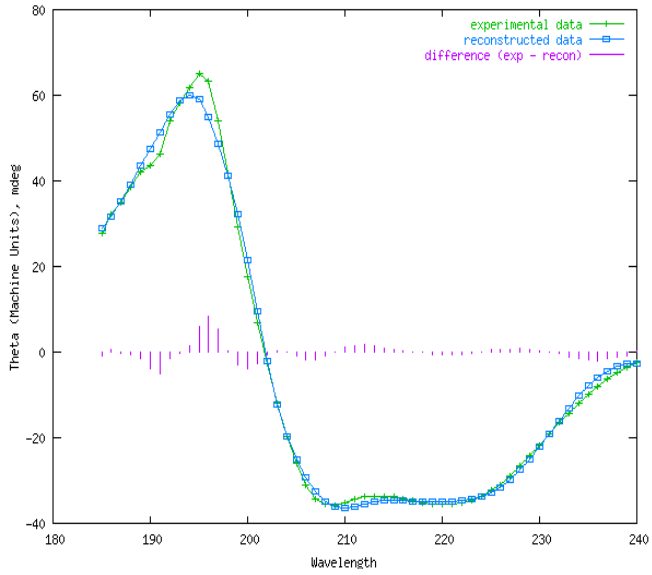



**Figure 40: Synchronous maps for FL and Δ80 CM-PD after 24 h incubation in $D_2O$.**

## Appendix 3A

A




B

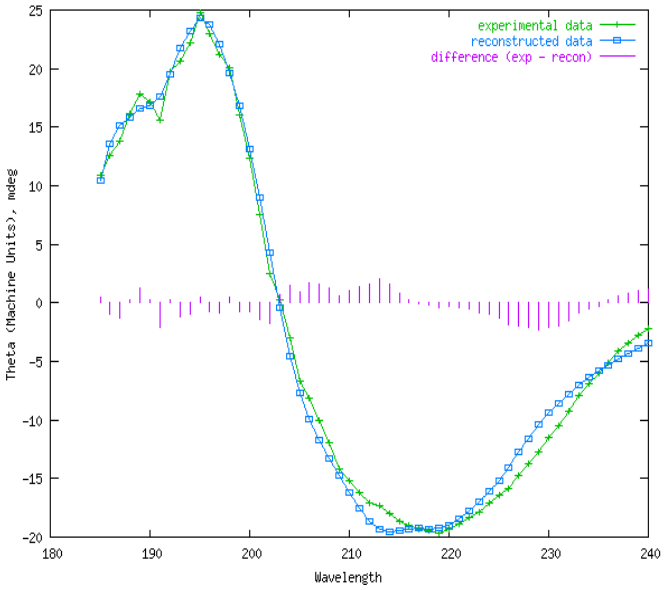



**Figure 41: Far-UV CD spectra of *H. influenzae* FL CM-PD at a high protein concentration at (A) 25°C and (B) 95°C.** Protein samples were prepared as Fig 13, measured on a Jasco 815 Spectropolarimeter using a 12 µm spacer. The curves were the results of the average of 5 scans performed at a rate of 50 nm/min. Reconstructed data was from Dichroweb, CONTIN algorithm (22). (A) has NRMSD of 0.063 and (B) has 0.081.

## Appendix 3B

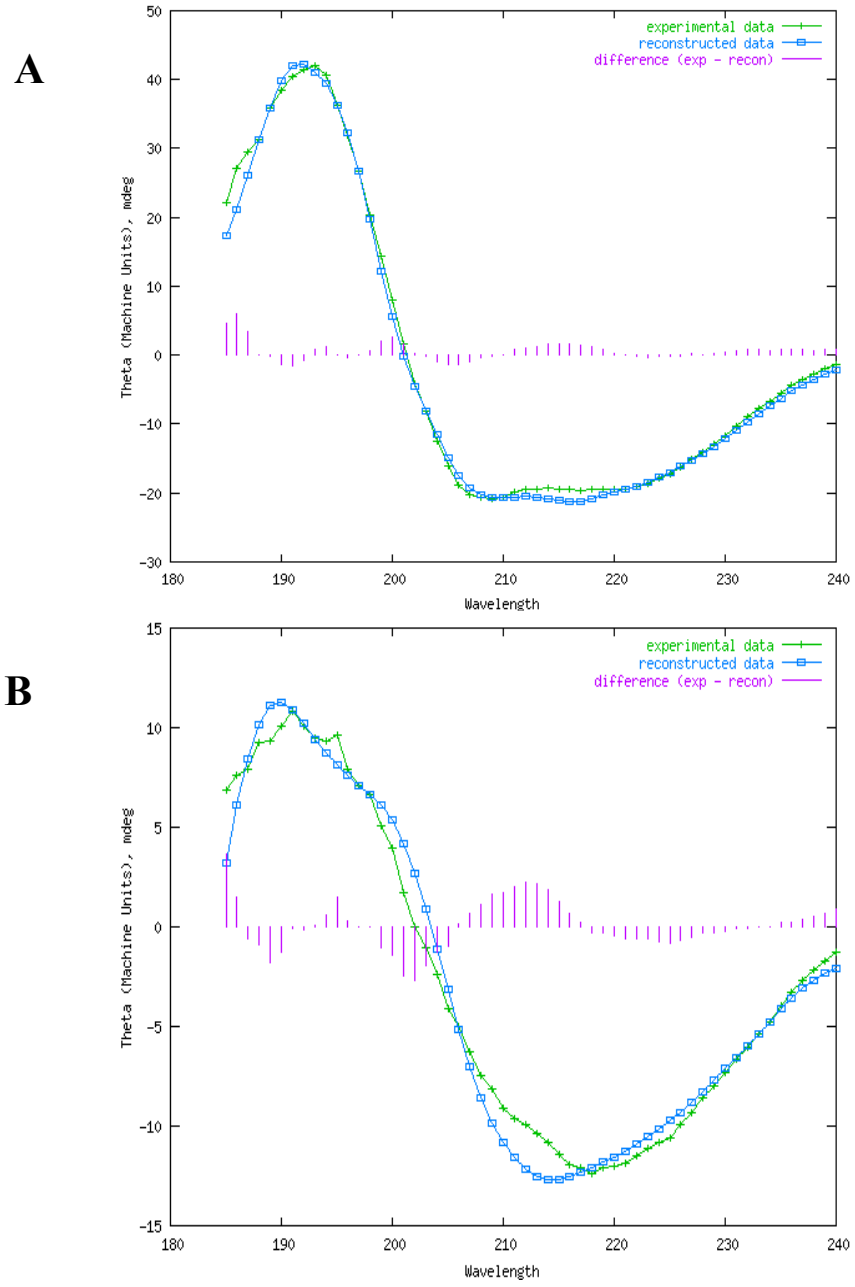



**Figure 42: Far-UV CD spectra of *H. influenzae* Δ80 CM-PD at a high protein concentration at (A) 25°C and (B) 95°C.** Protein samples were prepared as Fig 13, and measured on a Jasco 815 Spectropolarimeter using a 12 µm spacer. The curves were the results of the average of 5 scans performed at a rate of 50 nm/min. Reconstructed data was from Dichroweb, CONTIN algorithm (22). (A) has NRMSD of 0.068 and (B) has 0.143.

## Appendix 4

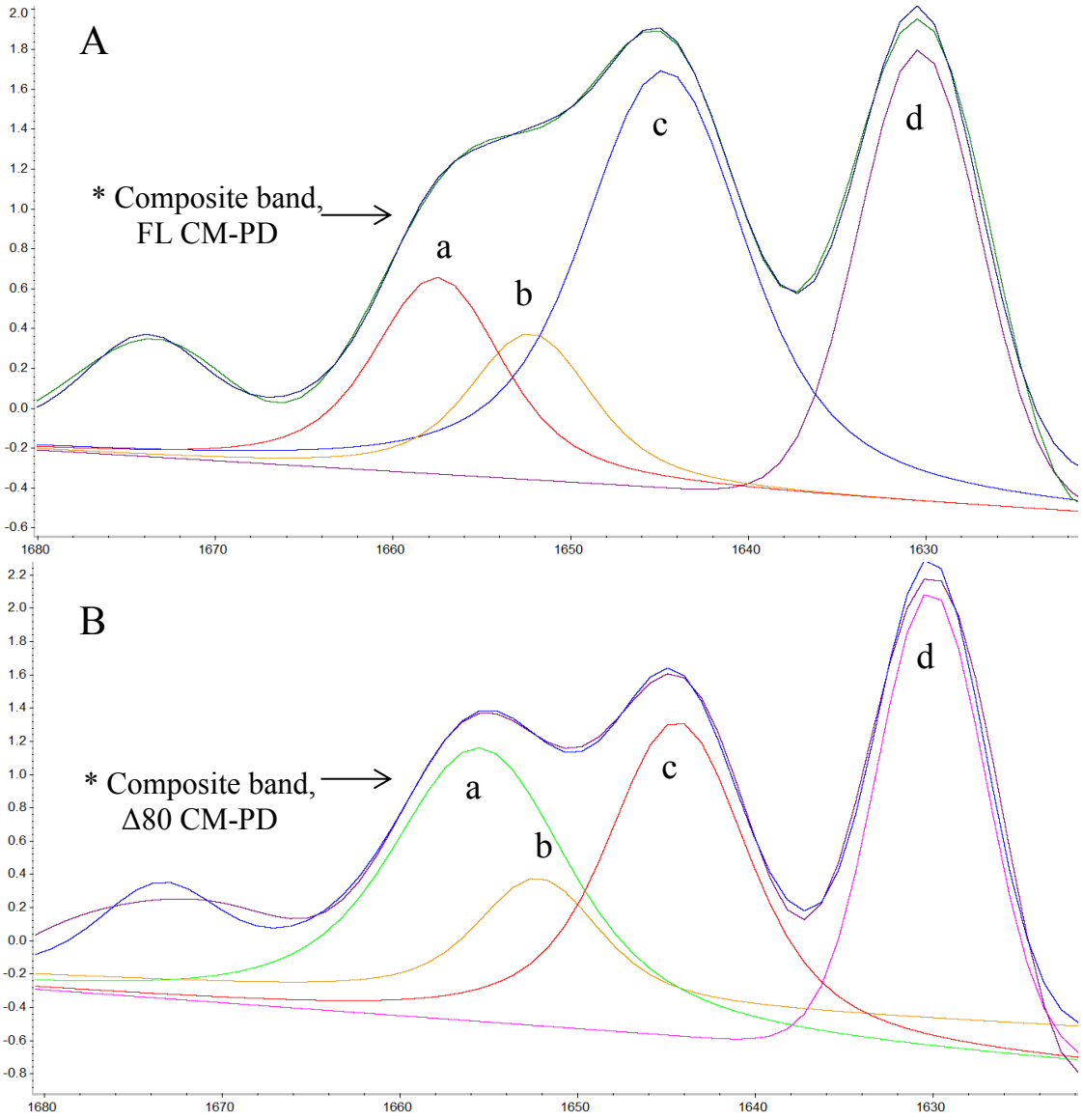



**Figure 43: Fourier self- deconvoluted (A) FL CM-PD and (B) Δ80 CM-PD at 25°C with computed composite and component peaks**. Fourier self- deconvoluted using bandwidth 19.2 and enhancement 3.0. The component peaks were found using Peak Resolution from OMNIC software analysis function. Resolved peaks were: a) α-helix (a, 1657cm$^{-1}$), b) α-helix (b, 1650 cm$^{-1}$), c) random coil (c, 1644 cm$^{-1}$), and d) intramolecular β-sheet (d, 1630 cm$^{-1}$). * Note the computed composite spectrum generated based on newly resolved peaks were super imposable with FL and Δ80 CM-PD spectra.