INVESTIGATIONS ON ENZYMES OF THE ENOLASE SUPERFAMILY

by

Oliver P. Kuehm

Submitted in partial fulfilment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
December 2020

DEDICATION

I dedicate this work to my Grandmother, for without her intense drive, passion, and

motivation, I would not be where I am today. My MSc has allowed me to become

incredibly close with her, and for that, I will be forever thankful.

TABLE OF CONTENTS

CHAPTER 3 THERMAL STABILITY OF ACTIVE SITE
VARIANT PROTEINS WITHIN THE MANDELATE

LIST OF TABLES

LIST OF FIGURES

LIST OF SCHEMES

ABSTRACT

Enzymes of the enolase superfamily (ENS) of share a high level of structural similarity and catalyze a conserved partial reaction yet display an overall mechanistic divergence. Mandelate racemase (MR) serves as a paradigm for the enzyme-catalyzed abstraction of an a-proton from a carbon acid substrate. This work described the characterization of a series of potent chloro- and fluoro-substituted phenylboronic acid (PBA) inhibitors of MR, including 3,4-dichloroPBA, for which MR exhibited the highest binding affinity observed to-date (K_d = 13.8 nM). Investigations on the conserved KxK motif of the MR subgroup, revealed that the KxK residue was essential for catalysis and for the stability of MR and L-tartrate dehydratase. The K164M and K164R MR variants displayed a change in thermal melting temperature (ΔT_m) of –4.03 and –8.36 °C, respectively. Finally, approaches towards production of human reverse thymidylate synthase, an ENS enzyme, using precursor proteins bearing two protease specific sites were described.

LIST OF ABBREVIATIONS USED

ATCC	american type culture collection
BSA	bovine serum albumin
BzH	benzohydroxamate
CD	circular dichroism
CfN	cupferron
CoA	coenzyme A
DFT	density functional theory
DSC	differential scanning calorimetry
DSF	differential scanning fluorimetry
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
ENS	enolase superfamily of enzymes
FdUMP	5-fluorodeoxyuridine 5′-monophosphate
FPLC	fast protein liquid chromatography
5-FU	5-fluorouracil
FucD	L-fuconate dehydratase
GC-MS	gas chromatography-mass spectrometry
αHBP	α-hydroxybenzylphosphonate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HFA	N-hydroxyformanilide
IPTG	isopropyl β-D-1-thiogalactopyranoside
ITC	isothermal titration calorimetry
LB	lysogeny broth
MBP	maltose binding protein
MLE	muconate lactonizing enzyme
MR	mandelate racemase
MW	molecular weight
MTX	methotrexate
NADH	nicotinamide adenine dinucleotide – reduced form
NDSB	non-detergent sulfobetaine

NMR	nuclear magnetic resonance spectroscopy
OD	optical density
PBA	phenylboronic acid
αPGA	phenylglycolate
PLP	pyridoxal phosphate
PMP	pyridoxamine phosphate
PVDF	polyvinylidene difluoride
rTS	reverse thymidylate synthase
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SSHB	short strong hydrogen bond
TarD	D-tartrate dehydratase
TB	terrific broth
TEV	tobacco etch virus
TGD	L-talarate-galactarate dehydratase
TIM	triosephosphate isomerase
Tris	tris(hydroxymethyl)aminomethane
TS	transition state
UV	ultra-violet
WT	wild-type

ACKNOWLEDGEMENTS

First and foremost, I must extend my most sincere thanks to Dr. Stephen Bearne. Dr. Bearne has unrivaled patience, work ethic, and care that infectiously spreads in his lab group. When I first began my MSc here at Dalhousie, I was unsure, unconfident, and lost in a field of science that I had no experience in. Dr. Bearne provided me with the guidance that I needed, teaching me a love for science and research that has helped shape me into the student and person I am today. He is a role model and a friend.

I would like to sincerely thank my supervisory committee members Dr. Cathy Too and Dr. Jan Rainey for their time and care in guiding me through my MSc. Their patience and effort have not gone unnoticed.

Many thanks to the members of the Bearne lab during my time here: Amar, Tom, Noa, Krishna, Josh, Chris, Himank, Laura, and Lillian. All of whom have made my MSc more productive and enjoyable and I am lucky to call them friends. I'd like to specifically thank Dr. Andy Yan-Song for his knowledge, patience, and help throughout my time here. From teaching me the intricacies of molecular biology and protein science, to training me on the ITC, he never hesitated to help me. I would like to thank Dr. Himank Kumar for his help in my first semester here and for training me on the DSC. Thanks to Dr. David Langelaan for the conversations in the hall and in his office that helped me so much my difficulties in purifying protein. And many thanks to all of the graduate students in the department for making me feel welcome from day one.

I am extremely grateful to have been able to complete a master's degree while being fully funded. I would like to thank the Killam family and NSERC-CREATE BioActives for funding my research and giving me exposure to many more opportunities

along the way. Similarly, being able to work as a teaching assistant throughout the program has given me exposure to many wonderful experiences with the undergraduate students. I am incredibly appreciative of Paul Briggs and Heidi MacKinnon for how much work they put in to maintain a smooth operation, and also for how caring and supportive they are of their students and TAs. With that said, the department staff, Roisin, Barb, and Brenda have been with me every step of the way. Their support through every step of my MSc has been integral to any successes that I have had.

I would like to give special thanks for my friends from StFX and from high school, who have supported me throughout my entire degree. A special thank you to my roommate of two years here, Sheumais MacLeod.

I must extend my profound appreciation to my wonderful girlfriend, Haileigh, who has been there every step of the way. Her work ethic, caring, and everyday passion as a friend and student is inspiring. She motivates those around her to succeed, and I am lucky to have had her along for this journey.

Lastly, I would like to thank my Mom and Dad for their endless and undying support. I am so grateful to have the support system that they provide for me. They have never let me down, and I am forever indebted to them.

CHAPTER 1 INTRODUCTION

1.1 ENZYMATIC ABSTRACTION OF PROTONS FROM CARBON ACIDS

The catalytic proficiency of an enzyme, or the virtual association constant for the enzyme and the altered substrate in the transition state (TS), has been shown to reach values as high as ~10^{23} M^{-1} (Miller & Wolfenden, 2002). This special ability of enzymes to stabilize the transition state of energetically difficult reactions is applicable to enzyme-catalyzed abstraction of protons from carbon acids.

Enzymatic abstraction of a proton from a carbon atom adjacent to a carboxylic acid group is highly unfavourable yet occurs ubiquitously throughout the enolase superfamily of enzymes (ENS) (Babbitt *et al.*, 1996). Enzymes in the ENS catalyze the abstraction of α-protons from carbon acid substrates with high pK_a values causing the formation of an *aci*-carboxylate intermediate (**Scheme 1.1**) (Babbitt *et al.*, 1996; Gerlt, *et al.*, 2012). Within the ENS, mandelate racemase (MR) has been thoroughly characterized as a model enzyme to understand how enzymes abstract an α-proton, a difficult step in catalysis (Bearne & St. Maurice, 2017). As an indication of the difficulty of this α-proton abstraction, the pK_a of the α-proton of mandelate is ~29 in solution and the rate of nonenzymatic racemization of (*R*)-mandelate (k_{non}) is 3×10^{-13} s^{-1} (Chiang, *et al.*, 1990; Bearne & Wolfenden, 1997). However, MR catalyzes the racemization of (*R*)- and (*S*)-mandelate with a rate enhancement (k_{cat}/K_m) of 15 orders of magnitude (Bearne & Wolfenden, 1997).

Scheme 1.1 Conserved partial reaction catalyzed by ENS enzymes

aci-carboxylate intermediate
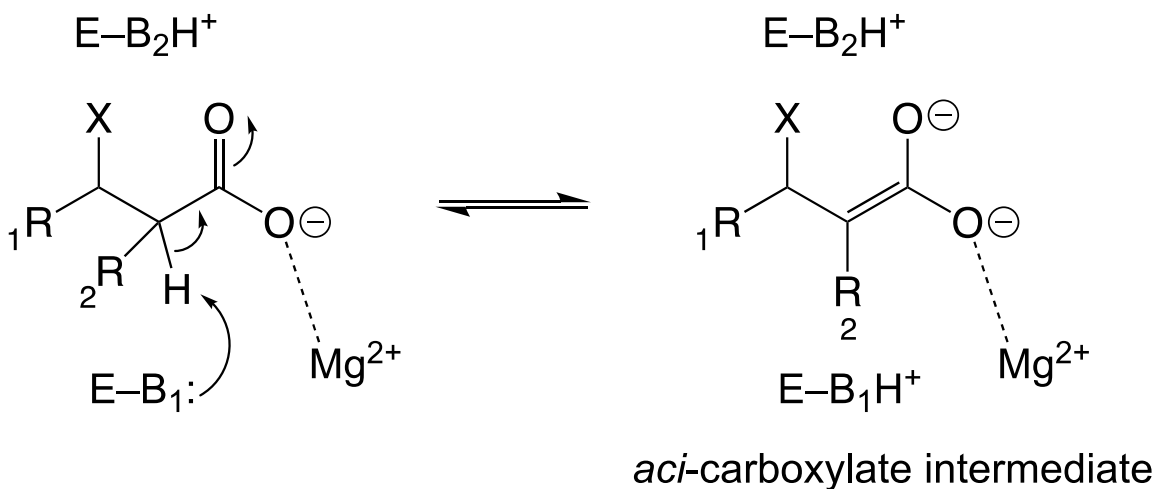
To catalytically facilitate this dramatic rate enhancement of such a difficult reaction, two energetic barriers on the enzyme must be overcome, which can be illustrated by Marcus theory. Marcus theory addresses the energetic barriers to catalysis and dictates that there is both a thermodynamic and an intrinsic barrier to enzymatic α-proton abstraction from a carbon acid (Gerlt & Gassman, 1992, 1993a; Gerlt, *et al.*, 1997). The thermodynamic barrier is attributed to the energetic instability associated with negative charge buildup on the enol or enolate intermediate (Gerlt, *et al.*, 1997; Thibblin, Jencks, 1979). The intrinsic barrier is kinetic in nature and arises from how abstraction of an α-proton from a carbon acid substrate occurs at a slower rate than that of a heteroatom or normal acid of equal acidity (Albery & Knowles, 1982; Bernasconi, 1992; Eigen, 1964). This difference in acidity is proposed to arise from changes in the arrangement of electron density around the carbon acid during proton abstraction and the requirement for re-orientation of solvent dipoles, which often lags behind the electronic rearrangement (Bernasconi, 1992; Guthrie, Peter, 1998). Enzymes overcome the intrinsic barrier to

catalysis by largely omitting solvent molecules from entering the active site and

positioning specific "solvating" amino acids to stabilize (formed) transition states and

intermediates (Gerlt & Gassman, 1993a; Gerlt & Gassman, 1993b).

From the thermodynamic perspective, it is the extremely high pK_a values for

carbon acids (pK_a > 50 for alkanes, ~ 22 for (R)-/(S)-mandelic acid, and ~29 for the (R)-

/(S)-mandelate anion) that pose the thermodynamic challenge for enzymes (Gerlt &

Gassman, 1992). To overcome this challenge, enzymes can either cause a decrease in the

pK_a of the carbon acid substrate or an increase in the pK_a of the catalytic basic residue(s)

(Toney, 2019). As a method of decreasing the pK_a of the substrate, enzymes can utilize

coenzymes and cofactors to enhance the reaction favourability by stabilizing the reaction

intermediates (Gerlt, 2007; Richard & Amyes, 2001). In the cases of amino acid

racemases, aminotransferases, and acyl-CoA dehydrogenases, coenzymes such as

pyridoxal phosphate (PLP), pyridoxamine phosphate (PMP), and flavin derivatives are

used as coupled "electron sinks" to delocalize and stabilize the α-anionic intermediates

formed (Mozzarelli & Bettati, 2006). Other enzymes that do not employ these coenzymes

may still catalyze α-proton abstractions from carbon acid substrates with high pK_a values

by using alternative methods of stabilizing intermediates (Toney, 2019).

In some cases, significant electrostatic stabilization can be achieved by use of a

metal cation, which often acts as a powerful electrophilic catalyst at the active site of

enzymes (Guthrie & Kluger, 1993). The positive charge of metal cations presumably

provides stabilization to the growing negative charge on the transition state that arises

during the deprotonation reaction, favouring intermediate formation (Babbitt, *et al.*,

1996). As such, an active-site divalent metal cation is required for catalysis in the case of

the ENS to electrostatically stabilize the *aci*-carboxylate intermediate formed during the

conserved partial reaction (**Figure 1.1**) (Babbitt, *et al.*, 1996).
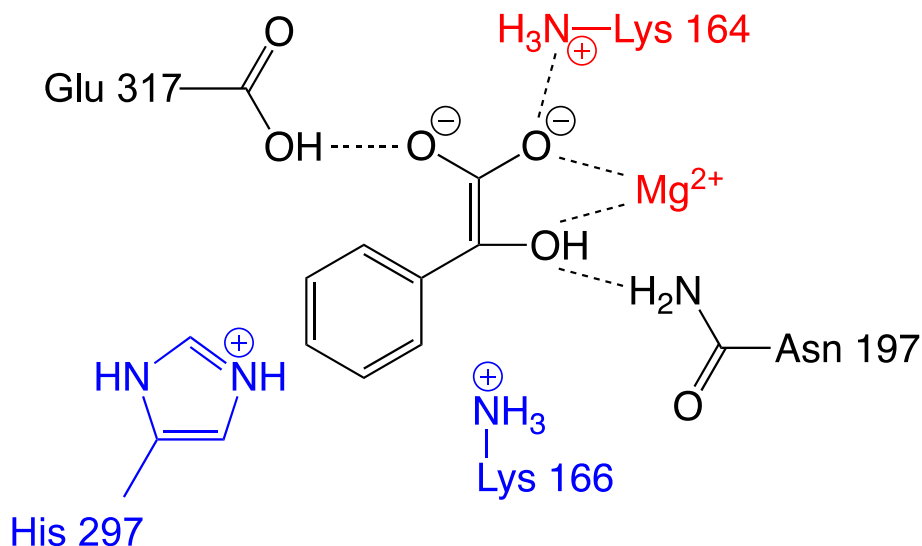
Figure 1.1. Electrostatic stabilization of the *aci*-carboxylate intermediate formed in MR catalysis. Highlighting the electrostatic stabilization provided by the Mg^{2+} ion and the positively charged Lys 164 to the *aci*-carboxylate intermediate. Also shown are Glu 317 and Asn 197, which form hydrogen bonds to the intermediate. The Brønsted acid-base catalysts His 297 and Lys 166 are also shown. (Bearne & St. Maurice, 2017).

The MR-catalyzed racemization of mandelate enantiomers provides an excellent

example for the use of electrophilic catalysis in enzymatic α-proton abstraction. The pK_a

value for mandelate at the active site of MR was estimated to be 7.4, which is close to the

reported pK_a for both Brønsted acid/base catalyst residues His 297 and Lys 166 in MR

(~6.4) (Gerlt, *et al.*, 1991; Landro, *et al.*, 1991). The steep increase in α-proton acidity of

carbon acid substrates is also seen in enzymes that employ hydrogen bonding and metal

coordination in their catalysis (Gerlt, *et al.*, 1991). In fact, available X-ray crystal

structures for proton-abstracting enzymes reveal that electrophilic catalysts are common and are always located proximal to the carbonyl or carboxylic acid groups of the substrates (Gerlt, *et al.,* 1991). Further, Guthrie and Kluger proposed that the electrostatic stabilization that exists in a low polarity medium, such as at the active site of MR, contributes significantly to the increased acidity of the substrate (Guthrie & Kluger, 1993).

As an alternative to reduction of substrate pK_a, the basicity of the proton abstracting residue can be modified, while simultaneously stabilizing the deprotonated intermediate to aid abstraction of the proton. Increased basicity of a catalytic residue has been seen with the enzyme triosephosphate isomerase (TIM), where the pK_a of catalytic Brønsted base Glu 165 is raised from 3 in the ground state to 10 in the transition state (Zhai, *et al.*, 2018). In the case of MR, the pK_a values of the catalytic basic residue in both directions of the racemization reaction are perturbed to 6.4, which is a change of ~1.4 units for Lys 166 and ~1.1 units for His 297 (Landro, *et al.*, 1994; Fetter, *et al.*, 2019). Altogether, MR makes use of both substrate pK_a reduction and modification of the pK_a of catalytic Brønsted base residues to overcome the thermodynamic barrier for α-proton abstraction.

1.2 THE ENOLASE SUPERFAMILY (ENS)

1.2.1 DISCOVERY AND CLASSIFICATION

The discovery of a pronounced level of similarity between the secondary, tertiary, and quaternary protein structure of muconate lactonizing enzyme (MLE) and MR sparked modern efforts towards classification of enzyme superfamilies (Neidhart, *et al.*, 1990).

Though MLE and MR catalyze different overall reactions, it was determined that they share a partial reaction (Babitt, *et al.*, 1995). Taken together, the notable similarities between structure and function of two distinctly different enzymes indicated the presence of divergent evolution from a common ancestor (Gerlt, *et al.*, 2012). Later analyses linked a large number of genes encoding homologous enzymes with conserved domain architecture, Brønsted base catalytic residues, and metal ion binding residues to the well-characterized glycolytic enzyme enolase. Resultingly, the group of homologous enzymes was denoted the "enolase superfamily" of enzymes (Babbitt, *et al.*, 1996). Currently, the ENS superfamily contains over 49000 member enzymes classified into seven subgroups based on the identity of the Bronsted acid-base catalysts and metal ion ligand residues (Akiva, *et al.*, 2014).

The ENS was the first mechanistically diverse superfamily reported, sharing a minimum of 35% amino acid sequence homology, significant structural homology, and a partial reaction amongst all members: the divalent metal-cation assisted, Bronsted base-catalyzed abstraction of a proton from a carbon acid substrate to form an enolic intermediate (**Scheme 1.1**) (Babbit, *et al.*, 1995; Gerlt, *et al.*, 2012). Conserved structural features include a $(\beta/\alpha)_7\beta$-barrel domain containing the enzyme active site and an $\alpha + \beta$-capping domain that provides many determinants for substrate recognition (Neidhart, *et al.*, 1990). Catalysis in the ENS occurs within a structually conserved active site, residing at the interface between the barrel and capping domains. At the active sites, a number of residues coordinate a divalent metal, often Mg^{2+}. A Brønsted base catalyst residue abstracts the α-proton of the substrate, forming the enolate intermediate that is stabilized by the Mg^{2+}. Brønsted acid catalysis then guides the intermediate to the reaction product,

which can be effected by the conjugate base of the Brønsted base catalyst or another catalytic residue (Gerlt, *et al.*, 2012).

The variation in the Brønsted acid-base catalyst residues and metal ion ligand residues between members give rise to the seven subgroups of the ENS, which are capable of catalyzing a variety of reactions, including racemization and epimerization (1,1-proton transfer), β-elimination, or intramolecular addition/elimination reactions (Babbitt, *et al.*, 1996; Gerlt, *et al.*, 2012, 2005). The seven subgroups, named after the first described enzyme in each subgroup, are the enolase, MR, MLE, β-methylaspartate ammonia lyase, D-mannonate dehydratase, D-glucarate dehydratase, and galactarate dehydratase subgroups (Akiva, *et al.*, 2014).

1.2.2 THE MR SUBGROUP

The MR subgroup of the ENS contains ~12 000 mechanistically distinct members and is named for its archetype enzyme MR (Akiva, *et al.*, 2014). Though the reactions catalyzed by enzymes within the MR subgroup are diverse in mechanism, after the conserved partial reaction takes place, either one of two fates can arise for the enolic intermediate: racemization/epimerization (i.e. 1,1-proton transfer) or β-elimination (i.e. dehydration) Overall, the vast majority of MR subgroup enzymes catalyze a dehydration reaction (Bearne & St. Maurice, 2017).

ENS enzymes are assigned to their constituent subgroups on a sequential and structural basis rather than a mechanistic one. Enzymes of the MR subgroup share a plenitude of structural features up to the quaternary structural level, which is octameric in nature but most accurately described as a tetramer of dimers (Fee, Hegeman, & Kenyon,

1974; Tsou, *et al.*, 1989). The overall structural organization remains characteristic to the ENS, with the active site lying between the capping ($\alpha + \beta$) and barrel ((β/α)$_7\beta$) domains. An additional binding determinant and/or structural support lies in the interdigitating loop that emanates from the neighboring subunit of the dimer and inserts into the distal portion of the active site (Bearne, 2017). Meanwhile, catalysis is afforded most directly by the Brønsted acid-base catalytic machinery responsible for catalyzing the initial deprotonation, and subsequent partial reactions. In the majority of cases in the MR subgroup, these Brønsted acid-base residues are lysine and histidine. MR subgroup enzymes contain a conserved Lys-X-Lys (KxK) motif at the end of the second β-strand where the latter lysine is implicated in catalysis as one of the two Bronsted acid-base catalytic residues (Yew, *et al.*, 2007). In the β-elimination reaction pathway within the MR subgroup, the conserved Bronsted base lysine is responsible for initial deprotonation preceding the elimination reaction (Yew, *et al.*, 2006a, 2006b; Yew, *et al.*, 2007). The Brønsted acid-base histidine is part of a His-Asp dyad. Both the dyad and the KxK motif play a role in modulating the pK_a of the Brønsted acid-base catalysts at the active site (Landro, *et al.*, 1991; Fetter, *et al.*, 2019).

Racemization/epimerization is not the predominant reaction catalyzed by the enzymes of the MR subgroup. However, racemization also occurs as part of the reaction catalyzed by L-talarate galactarate dehydratase (TGD), and only minimally in the case of L-fuconate dehydratase (FucD) (Yew, *et al.*, 2006a; Yew, *et al.*, 2007). Racemization and epimerization reactions within the MR subgroup occur by direct protonation of the *aci*-carboxylate intermediate by a Bronsted acid catalytic residue. In the case of MR, which follows a two-base mechanism, the enolate intermediate is formed following

8

deprotonation by the Bronsted base catalyst (Lys 166 or His 297, depending on the

substrate) and is followed by enantiospecific protonation by the conjugate acid of the

second Bronsted basic residue (Powers, *et al.*, 1991). Two-base enzymatic mechanisms

used in racemization/epimerization reactions within the MR subgroup require the second

base to be positioned antipodal to the deprotonating Brønsted base residue, which permits

enantiospecific product formation.

On the other hand, β-elimination is the reaction pathway followed by the vast

majority of MR subgroup enzymes, which involves elimination of a β-hydroxyl group

from an acid sugar substrate (Gerlt, *et al.*, 2012). Hence, the majority of MR subgroup

members are acid sugar dehydratases.

1.2.3 REPRESENTATIVE ENZYMES OF THE MR SUBGROUP

To conduct the studies on the MR subgroup for this work, four homologous

enzymes were selected that had previously been studied and characterized in depth, both

structurally and functionally, by Gerlt and co-workers: MR (*Pseudomonas putida*), L-

fuconate dehydratase (FucD from *Xanthomonas campestris*), L-talarate galactarate

dehydratase (TGD from *Salmonella typhimurium*), and D-tartrate dehydratase (TarD from

Bradyrhizobium japonicum) (Yew, *et al.*, 2006a; Yew, *et al.*, 2006b; Yew, *et al.*, 2007).

FucD catalyzes the dehydration of L-fuconate prior to stereospecific ketonization of its

enol product to 2-keto-3-deoxy-L-fuconate. TGD catalyzes the competing dehydration

and epimerization of L-talarate and galactarate to form 5-keto-4-deoxy-D-glucarate. TarD

catalyzes the dehydration of D-tartrate followed by a nonenzymatic keto-enol
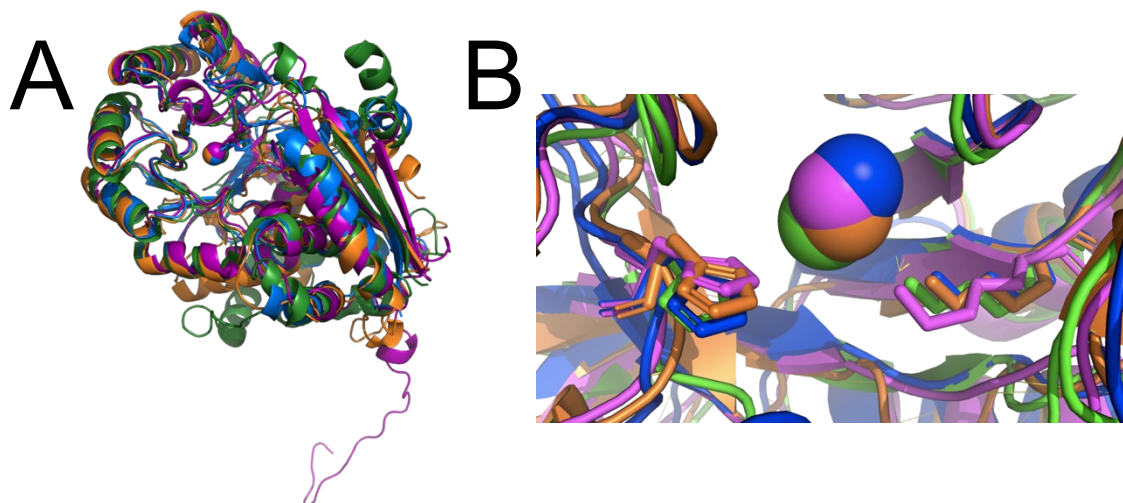
Figure 1.2. Structure overlay of *Pp*MR, *Xc*FucD, *Bj*TarD, and *St*TGD. The X-ray crystal structures of wild-type MR, FucD, TarD, and TGD monomers are shown to highlight the conserved structure of the MR subgroup (A). The active sites of the superposed wild-type MR, FucD, TarD, and TGD are shown, highlighting the conserved Mg^{2+} ion and the Brønsted acid-base His and Lys residues (B). MR is shown in blue, FucD is shown in green, TarD is shown in orange, and TGD is shown in purple. The Mg^{2+} ion is shown in space filling orientation.

tautomerization of the product to yield oxaloacetate. Hence, FucD, TGD, and TarD encompass the diversity of reactions typical of the MR subgroup.

Indeed, the representative MR subgroup enzymes share structural similarities down to the level of active site architecture. Superposing the structures of MR and the three acid sugar dehydratases shows the similarities in the positioning of the acid-base Lys residue at the end of the second β-strand, as well as the His-Asp dyad at the end of the sixth and seventh β-strands, respectively (Yew, *et al.*, 2007). Thus, the level of homology in the active site architecture of enzymes that are distinctly different in mechanism and substrate suggests that enzymes in the MR subgroup are "hard-wired" for acid-base chemistry (Gerlt, Babbitt, & Rayment, 2005).

As briefly mentioned above, apart from MR, TGD is the only other enzyme of the studied MR subgroup enzymes that facilitates a 1,1-proton transfer as well as, the dehydration of the epimeric substrates L-talarate and galactarate (Yew, *et al.*, 2007). This occurs such that when L-talarate is the substrate for dehydration, concomitant epimerization to galactarate occurs, though in the opposing situation, when galactarate is the substrate for dehydration, only minimal epimerization to L-talarate occurs. Hence, TGD affords a mechanistic linkage between the 1,1-proton transfer- and β-elimination-catalyzing enzymes. In the present studies, this group of representative enzymes were used to explore the role of the catalytic residues in protein stability within the MR subgroup (Chapter 3).

1.3 MANDELATE RACEMASE

Mandelate racemase from *Pseudomonas putida* is an enzyme of the ENS and the enzyme for which the MR subgroup is named. MR catalyzes a Mg^{2+}-assisted, Brønsted base-catalyzed 1,1-proton transfer that results in the interconversion of the enantiomers of mandelate (**Scheme 1.2**).

Scheme 1.2. The MR-catalyzed racemization of (*R*)- and (*S*)-mandelate
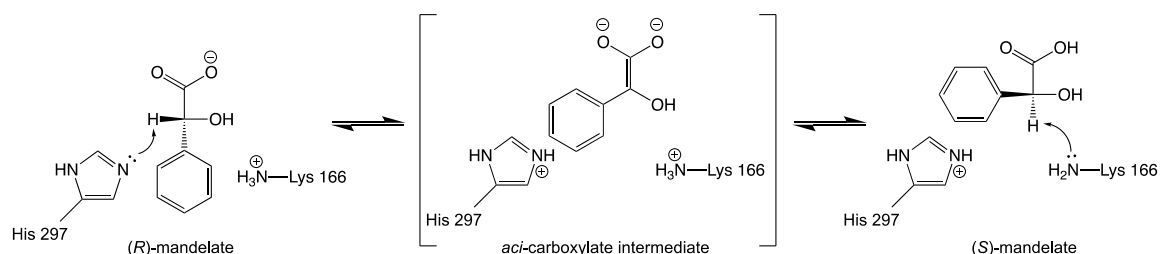
1.3.1 DISCOVERY AND ROLE

The discovery of MR in 1953 came after it was observed that *P. putida* strain A.3.12 (ATCC 12633) was able to grow on plates containing minimal media with either (*R*)- or (*S*)-mandelate as the sole carbon source (Gunsalus, Stanier, & Gunsalus, 1953; Gunsalus, Gunsalus, & Stanier, 1953). Both (*R*)- and (*S*)-mandelic acid are naturally occurring α-hydroxycarboxylic acids produced from plant origin (Bhat & Vaidyanathan, 1976). The discovery of MR led to an operon analysis of *P. putida*, which revealed that the gene encoding MR resided among several genes encoding proteins in the catabolic pathway of mandelic acid. While the *P. putida* strain A.3.12 is one of 10 total mandelate racemases, other microorganisms including bacteria and *Aspergillus niger* have been shown capable of using either (*R*)- or (*S*)-mandelate as a carbon source due to the

expression of (*R*)-/(*S*)-stereospecific mandelate dehydrogenase (Bhat & Vaidyanathan, 1976; Zhou, *et al.*, 2018). Later discovery of the ENS in the 1990s put the spotlight on MR as the prototype for understanding enzymes catalyzing heterolytic cleavage of C–H bonds, leading to its appointment as archetype of the MR subgroup.

1.3.2 MECHANISM OF MR CATALYSIS

Mechanistic studies of MR began in 1970 when it was proposed that racemization was initiated by heterolytic cleavage of the α-C–H bond, which generates an unstable carbanionic intermediate (Kenyon & Hegeman, 1970). However, uncertainty remained with respect to the mechanism by which MR facilitated this reaction. It was suggested that in order to catalyze energetically disfavoured reactions such as the heterolytic cleavage of a C–H bond, enzymes must either immediately stabilize the carbanionic intermediate generated in a stepwise reaction or utilize a concerted mechanism to avoid its formation altogether (Thibblin & Jencks, 1979). In the early 1990s, a two-base mechanism was proposed and confirmed through a variety of experimental techniques. This mechanism occurs via deprotonation and reprotonation reactions catalyzed by Brønsted acid-base residues located on opposing sides of the chiral α-carbon (Neidhart, *et al.*, 1991).

In an attempt to identify the catalytic Brønsted acid-base residues, Gerlt and colleagues conducted deuterium exchange experiments (Powers, *et al.*, 1991). High-resolution gas chromatography paired with mass spectrometry (GC-MS) was used to observe the substrate and products at early time points in the MR-catalyzed reaction conducted in D_2O solvent. It was found that on a short time scale, the α-proton of product

was derived from the solvent D_2O, consistent with a two-base mechanism. Further, in observing the remaining substrate from the reaction, no solvent exchange was seen in the (R)-mandelate substrate, but significant solvent exchange occurred when (S)-mandelate was the substrate. This indicated that MR utilized enantiospecific Brønsted base catalytic residues to facilitate catalysis. More specifically, these results showed that the base responsible for deprotonation of (R)-mandelate was monoprotic, and that the base responsible for deprotonation of (S)-mandelate was polyprotic.

The first X-ray crystal structures of MR showed two residues proximal to the α-carbon of the substrate (Neidhart, *et al.*, 1991). His 297 was poised in a position to deprotonate (R)-mandelate and Lys 166 was poised to deprotonate (S)-mandelate, consistent with the aforementioned isotope exchange experiments (Powers, *et al.*, 1991). Additionally, it was shown that the H297N MR variant, while catalytically inactive, was able to incorporate solvent-derived deuterium into (S)-mandelate (Landro, *et al.*, 1991). To support these mechanistic hypotheses, it was shown that α-phenylglycidate (αPGA), an alkylating epoxide analog of mandelate, only formed an irreversible adduct with MR when the (R)-αPGA was used (i.e., where the epoxide replaces the α-proton position of (S)-mandelate) (Landro, *et al.*, 1994). This suggested that Lys 166 acted as the deprotonating base for (S)-mandelate. Further, X-ray crystallographic studies showed the corresponding (R)-αPGA forming an adduct to Lys 166, supporting its role as a Brønsted base in MR catalysis.

To help demonstrate the formation of a reaction intermediate, Kozarich, Neidhart, and Kenyon (Landro, *et al.*, 1991) showed that the H297N variant of MR could catalyze the stereospecific elimination of the bromide ion from (S)-4-(bromomethyl)mandelate to

yield 4-methyl-benzoylformate. The remaining unreacted substrate was entirely in the (*R*)-conformation. This strongly suggested that intermediate formation also occurs in the wild-type enzyme.

1.3.3 STRUCTURE OF MR

Early X-ray crystallographic studies showed that MR exists as a tightly packed homooctamer, which, as a result of the intimate interfacing between symmetric pairs of monomers, has been deemed a tetramer of dimers (**Figure 1.3**) (Neidhart, *et al.*, 1991). Further, the dimeric pairs of MR subunits constitute the catalytic form mostly as a result of the previously discussed interdigitating loop (residues 92-96). The interdigitating loop emanates from the N-terminal capping domain of one monomer and inserts into the active site of the symmetry adjacent monomer, providing structural support and a direct binding determinant (Leu 93) (Bearne, 2017). MR has also been shown to exist as a tetramer under gel-filtration conditions (Sagy & Bearne, unpublished).

The catalytic subunits of MR are bidomain in nature, consistent with the conserved structure of ENS superfamily enzyme (**Figure 1.4**) (Neidhart, *et al.*, 1991; Babbitt, *et al.*, 1996). Early studies revealed that the active site of MR was located at the interface of the two domains in the X-ray crystal structure, determined by the position of europium (III), which is known to bind to divalent metal cation sites in proteins, and manganese, with which MR is active and the electron density differential from Mg^{2+} permits position determination (Neidhart, *et al.* 1991). In addition, the X-ray crystal structure of MR with 4-iodomandelate bound revealed the location of the substrate within
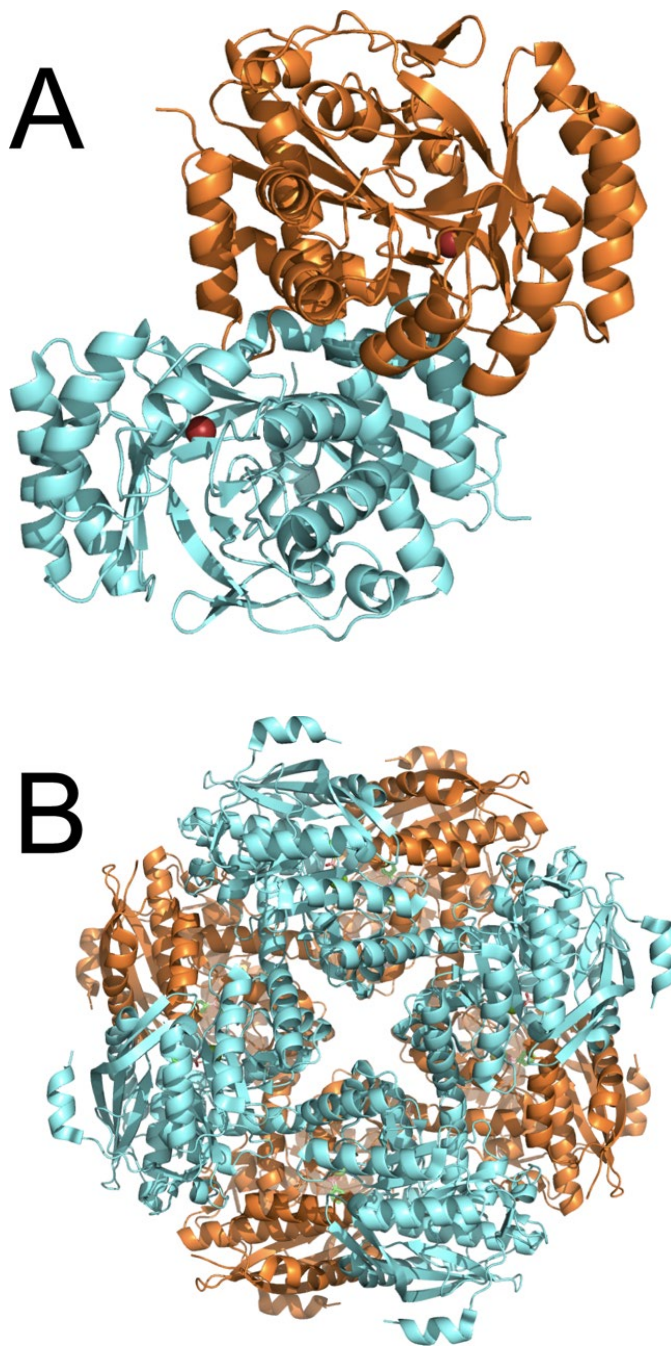
Figure 1.3. Quaternary structure of MR. Shown is the tightly associated MR catalytic dimer (A) and the homooctomeric structure (B) comprising a tetramer of dimers. Alternating monomeric subunits are shown in orange and blue to highlight the symmetry in the quaternary structure of MR. The bound Mg^{2+} ions are shown in red (space filling model) (PDB entry 6VIM).

the enzyme as a result of the electron dense iodo-substituent. A number of X-ray crystal structure determinations for MR variant proteins with various ligand molecules have permitted further elaboration of the roles of various amino acids residues in the catalytic mechanism (Landro *et al.*, 1994; Kallarakal, *et al.*, 1995; Landro *et al.*, 1994; Mitra, *et al.*, 1995; Schafer, *et al.*, 1996).

The two domains of MR are an N-terminal $\alpha + \beta$ capping domain and a C-terminal $(\beta/\alpha)_7\beta$ barrel. The essential 20s (residue 17-30) and 50s (residue 51-55) loops, exisiting in the N-terminal domain, link the regions of secondary structure in the capping domain and interact with the distal portion of substrate or ligand in the hydrophobic binding region (Bourque & Bearne, 2008). The 20s and 50s loops vary in sequence and position in the MR subgroup, but influence ligand binding by providing a number of side chain binding determinants and also dictate the polarity of the active site. The C-terminal domain is a commonly conserved α/β barrel structure, resembling that of TIM (α_8/β_8) but lacking the eighth α-helix of the domain (Banner, *et al.*, 1975). Overall, the C-terminal domain contributes many highly conserved elements to MR catalysis, such as the Mg^{2+}-binding and catalytic residues. The N-terminal capping domain contributes residues with side chains that interact with the distal portion of the substrate, conferring substrate specificity to the enzyme (Neidhart, *et al.*, 1991).
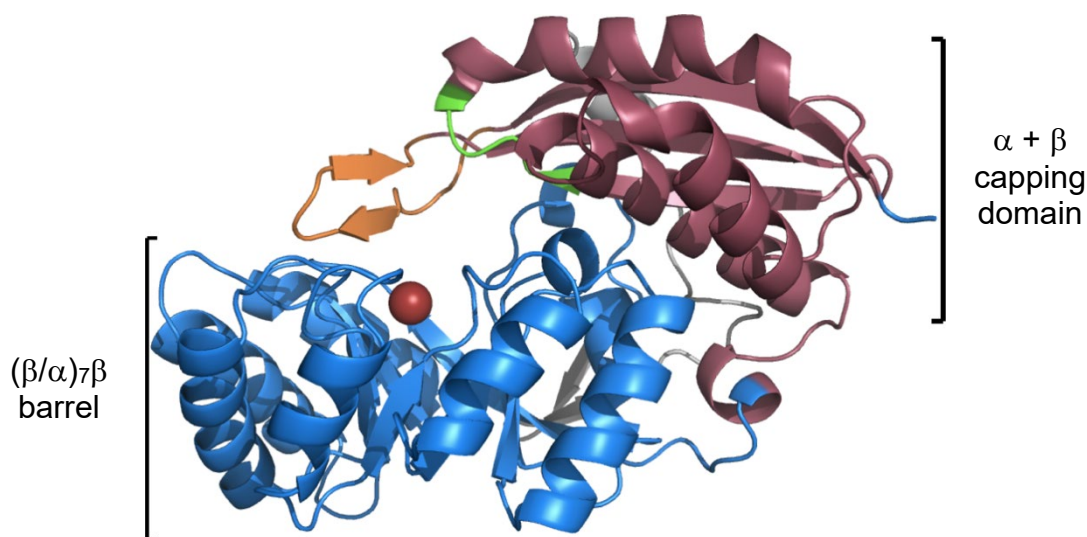
Figure 1.4. Structure of the monomeric subunits of MR. Shown are highlighted features of the MR subunits. The $(\beta/\alpha)_7\beta$ barrel is shown in blue as labelled, the N-terminal $\alpha + \beta$ capping domain is shown in light red, and the Mg^{2+} ion is shown in red (space filling representation). The 20s loop is shown in orange and the 50s loop is shown in green.

With regard to the substrate binding determinants, a number of specific hydrogen bonds and electrostatic interactions comprise the interactions between MR and its substrates that permit catalysis. Glu 317 and Lys 164 each interact with one of the carboxylate oxygens, with Lys 164 forming a hydrogen bond. Ser 139 forms a hydrogen bond to the same oxygen as Glu 317, while Asn 197 interacts with the α-hydroxyl group of the substrate, both of which play a role in intermediate stabilization (St. Maurice & Bearne, 2000). Remaining binding determinants include the Brønsted acid-base catalysts Lys 166 and His 297, which also plausibly interact with the phenyl ring of substrate through cation-π/NH$_2$-π interactions (Nagar & Bearne, 2015). Most recently, the triad of Tyr 137, Lys 164, and Lys 166 has been found to interact in an intimate fashion through hydrogen bonding and electrostatic repulsion (i.e., Lys 164–NH$_3^+$ - $^+$H$_3$N–Lys 166) to

lower the pK_a of the Lys 166 for catalysis (Fetter, *et al.*, 2019). Finally, an important

structural feature that supports catalysis is the hydrophobic binding pocket. The grouping

of residues in this pocket act as a binding determinant for the hydrophobic phenyl ring of

the substrate. Overall, these residues include the N-terminal domain residues (Leu 18,

Val 22, Val 29, Phe 52, and Tyr 54), barrel domain residues (Leu 298, Leu 319, and Leu

321), and Leu 93 from the interdigitating loop (Neidhart, *et al.*, 1991; Bearne, 2017). The

flexibility/plasticity of the hydrophobic binding pocket accounts for the observed ability

of MR to bind a plethora of substrates and ligands (St. Maurice & Bearne, 2004).

Fee *et al.* determined that the active-site divalent magnesium cation is required for

catalysis (1974). It plays a substantial role in ENS catalysis by furnishing electrostatic

stabilization of the enolate intermediate (Guthrie, Kluger, 1993; Babbitt, *et al.*, 1996).

Gerlt and co-workers observed an octahedral coordination of the Mg^{2+} in the MR active

site (Neidhart, *et al.*, 1991). The highly conserved ligands of the Mg^{2+} in the MR

subgroup serve as a partial identifier in classification of the ENS. Specifically, an aspartic

acid and two glutamic acids from the C-terminal domain coordinate to the Mg^{2+} in a

meridional coordination (Bearne, 2017). The coordination sites of Mg^{2+} are occupied by

one of oxygen of the carboxylate of the substrate, the α-hydroxyl group of the substrate,

and an ordered water molecule. The requirement for a divalent metal ion is conserved

across the ENS as an essential aspect of catalysis, which allows for subgroup

classification based on the specific ligands of the metal ion (Akiva, *et al.*, 2014).

1.3.4 INHIBITION OF MR

1.3.4.1 Transition State Stabilization

Like all catalysts, enzymes facilitate catalysis of a typically difficult or slow

reaction by decreasing the energy of activation by stabilization of the transition state

(Wolfenden, 1972). This explanation of enzyme catalysis is known as transition state

(TS) stabilization (Pauling, 1948). In the case of MR catalysis, the observed lifetime of

the deprotonated reaction intermediate was at least as long as the time taken for rotation

of the polyprotic Brønsted acid-base catalyst Lys 166 (C–N$^\zeta$H$_3{}^+$) (Landro, *et al.*, 1991).

As a result, MR must provide stabilization to the enolic intermediate in order to sustain

its lifetime. In support of this, Wolfenden and Bearne showed a massive rate

enhancement of the MR catalyzed reaction, 15 orders of magnitude greater than the

nonenzymatic counterpart (Bearne & Wolfenden, 1997).

Wolfenden's 1972 model of TS stabilization permits determination of a

hypothetical equilibrium constant describing the dissociation of the enzyme·substrate

complex via the TS to yield free enzyme and the TS (K_{tx}) using the thermodynamic cycle

shown in **Scheme 1.3** (Wolfenden, 1972). Wolfenden and Bearne obtained an estimation

of the upper limit for K_{tx} for MR catalysis from the ratio of the rate of the nonenzymatic

reaction (k_{non}) and the catalytic efficiency of the enzyme (k_{cat}/K_m), which yielded a value

of 2×10^{-19} M. This value corresponds to a transition state free energy stabilization of

~26 kcal/mol (**Figure 1.5**). In comparison to the affinity of MR for substrate ($K_m = K_S \approx$

1×10^{-3} M), it is evident that the enzyme exhibits preferential binding for the transition

state, which affords catalysis at the enormous proficiency seen (St. Maurice & Bearne,

2002). It should be noted that the reciprocal of K_{tx} is defined as the proficiency of an

enzyme (i.e., the efficiency of the enzymatic reaction divided by the rate of the

nonenzymatic reaction) (Radzicka & Wolfenden, 1995).

Scheme 1.3 Thermodynamic cycle linking the enzyme catalyzed and nonenzymatic reactions.

$$
\begin{array}{ccc}
\begin{matrix} S \\ + \\ E \end{matrix} & \xrightarrow{\ \ K_{non}^{\ddagger}\ \ } & \begin{matrix} S^{\ddagger} \\ + \\ E \end{matrix} \\[2ex]
K_S \Big\Updownarrow & & K_{tx} \Big\Updownarrow \quad E + P \\[2ex]
ES & \xrightleftharpoons{\ \ K_{cat}^{\ddagger}\ \ } & ES^{\ddagger}
\end{array}
$$
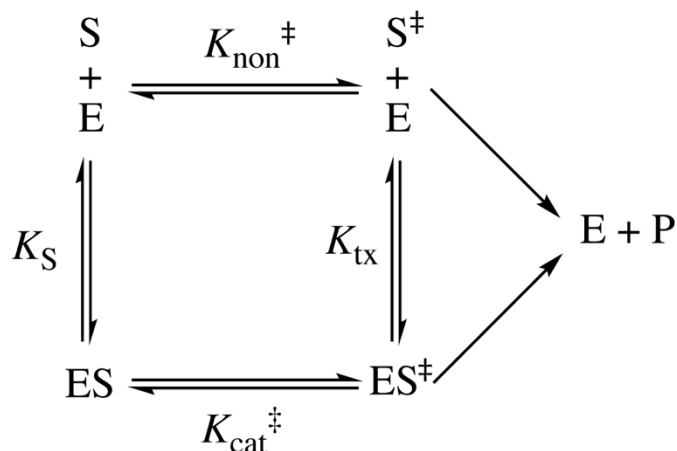
In order to afford proficient catalysis, MR and enzymes of the ENS must also

stabilize the *aci*-carboxylate intermediate and the resulting accumulation of negative

charge in the active site. This falls in accord with the Hammond postulate, which states

that the transition state of a reaction resembles the ground state species that it is more

close to in energy, either substrate or product (Hammond, 1955). Immediately

surrounding the site of deprotonation exists a localized region of positive charge

consisting of Lys 166, Lys 164, and the Mg^{2+} ion. In addition to the electrostatic

stabilization of the intermediate afforded by the metal ion (Guthrie & Kluger, 1993;

Babbitt, *et al.*, 1996), Lys 164 in the KxK motif forms a hydrogen bond to one of the

carboxylate oxygens of substrate, which may further stabilize the negative charge in the

intermediate and TS (Babbitt, *et al.*, 1996).
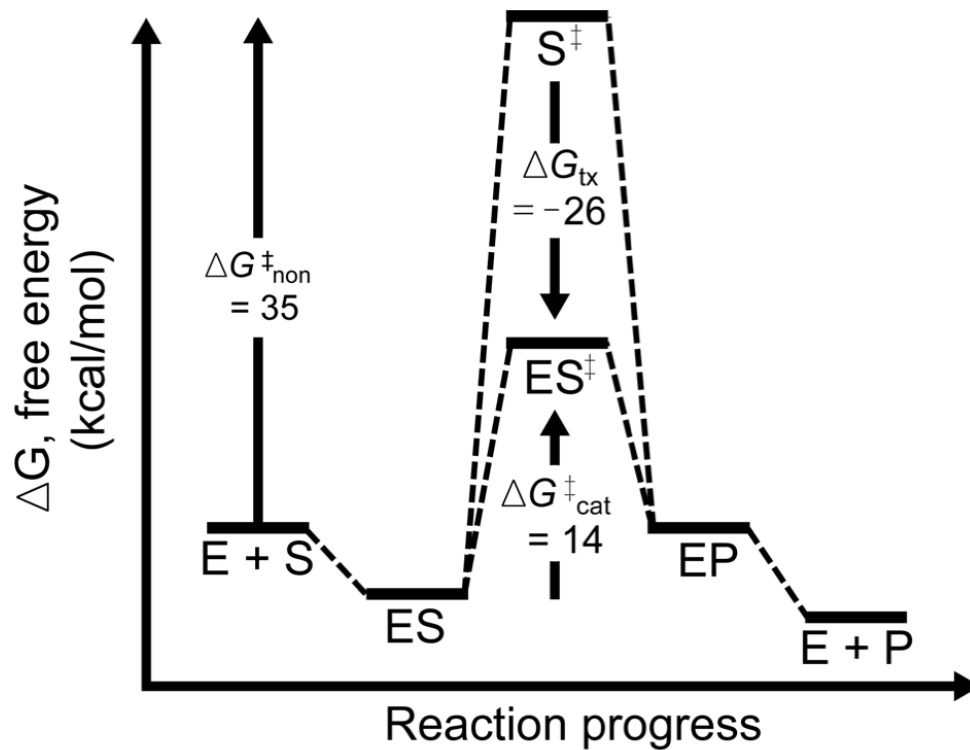
Figure 1.5 Transition state stabilization conferred by MR. The free energy change associated with forming the transition state ($\Delta G^{\ddagger}_{non}$) of the nonenzymatic reaction, the catalyzed reaction ($\Delta G^{\ddagger}_{cat}$), and the free energy stabilization in the ES complex ($\Delta G^{\ddagger}_{tx}$) are shown. In order to facilitate catalysis, enzymes (E) exhibit preferential and enhanced binding to the TS ($S^{\ddagger}$) in the enzyme-substrate complex (ES, $ES^{\ddagger}$), which aids in reaction turnover to product (EP, P). Adapted from Bearne & Wolfenden (1997).

Additional TS stabilization arises from Glu 317, which provides either general acid catalysis or electrophilic catalysis. Upon deprotonation of the substrate, Glu 317 protonates the carboxylate group of mandelate, which forms a stable enol instead of the unstable enolate species (Mitra, *et al.*, 1995). Gerlt and Gassman proposed that the formed "enolic" intermediate becomes stabilized by formation of a strong "low-barrier" hydrogen bond to the γ-carboxylic acid functionality of Glu 317 (Gerlt, *et al.*,1991). Low-barrier or short strong hydrogen bonds (SSHB) occur when two heteroatoms with approximately equal pK_a values come into closer proximity than the sum of their van der Waals radii. The proximity of the two species reduces the energy barrier between the two hydrogen bonds such that the proton is equally shared between both heteroatoms (Cleland & Kreevoy, 1994). In combination, the use of electrostatic stabilization in a low-polarity environment and use of general acid-base catalysis can function to reduce the pK_a of substrate to partially account for the proficiency of MR (Gerlt & Gassman, 1993a, 1993b; Guthrie & Kluger, 1993).

Finally, the Brønsted acid-base residues Lys 166 and His 297 play a substantial role outside of their role in direct deprotonation of substrate. The role of His 297 and Lys 166 as binding determinants through the formation of cation-π/NH_2-π interaction with the phenyl ring was mentioned above. However, these interactions may play a more substantial role in transition state stabilization. Bearne and Nagar observed that the high-affinity binding of MR to the TS analogue inhibitor benzohydroxamate (BzH) was decreased for the K166M and H297N MR variants (Nagar & Bearne, 2015). These changes in binding affinity predominantly implicated Lys 166 in binding, since the K166M MR variant resulted in 443-fold reduced affinity. Their observations also

suggested that additional cation-π/NH$_2$-π interactions with the hydroxamate/hydroximate moiety of BzH were also present (Nagar & Bearne, 2015). These findings imply that the developing π system of the *aci*-carboxylate intermediate is stabilized by Lys 166 and His 297.

1.3.4.2 Inhibitors of MR

Transition state analogue inhibitors mimic the structural and electronic properties of the TS, which, in an ideal situation would yield a high-affinity interaction with the enzyme close to that of the actual TS (i.e., K_{tx}) (Wolfenden, 1972). Extensive research on MR has yielded a number of transition state analogues that inhibit MR with binding affinities 2-3 orders of magnitude greater than the substrate binding affinity (St. Maurice, *et al.*, 2003; Burley & Bearne, 2005; Bourque, Burley, & Bearne, 2007; Lietzan, *et al.*, 2012).

In design of TS analogues for the reaction catalyzed by MR, both structural and electronic features have been utilized to mimic the *aci*-carboxylate intermediate. Potent reversible competitive inhibition for a series of TS analogues has been observed, including benzohydroxamate (K_i = 9.3 µM), cupferron (CfN, K_i = 2.7 µM), *N*-hydroxyformanilide (HFA, K_i = 2.8 µM), and α-hydroxybenzylphosphonate (α-HBP, K_i = 4.7 µM) (St. Maurice & Bearne, 2000; Bourque, *et al.*, 2007). Structures for the TS analogue inhibitors of MR are given in **Figure 1.6**. Additionally, BzH was confirmed as a TS analogue of MR by correlation of the catalytic efficiency (k_{cat}/K_m) with the competitive kinetic inhibition constant (K_i) for a series of MR variants (Lietzan, *et al.*, 2012). Hydroxamates are structurally analogous to the intermediate resembling the sp^2

hybridization of the α-carbon but sacrifice true intermediate mimicry by loss of an atom in place of the *aci*-carboxylate oxygens. Phosphonates mimic the dianionic electrostatic character of the *aci*-carboxylate but sacrifice structural mimicry in their tetrahedral geometry.
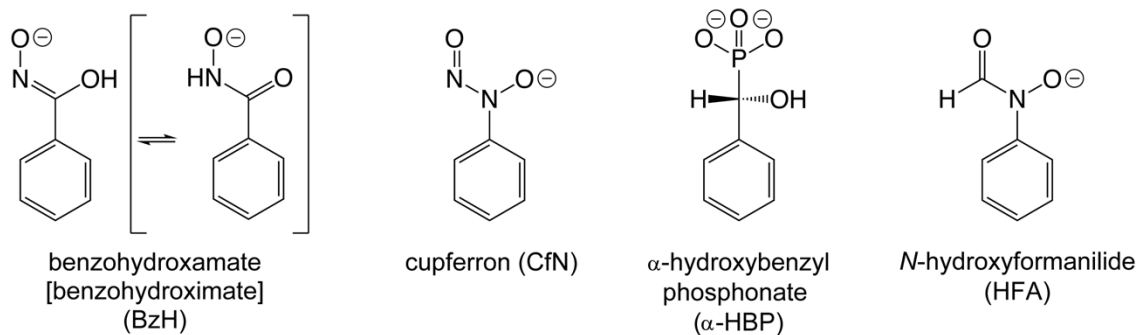
Figure 1.6. Structures of characterized TS analogue inhibitors of MR. Characterized transition state analogue inhibitors of MR are shown. Hydroxamates mimic the structural character of the *aci*-carboxylate intermediate, while phosphonates mimic the electronic character.

X-Ray crystallographic studies of MR with bound BzH and CfN revealed that the protein assumed a more compact conformation (Lietzan, *et al.*, 2012). Notably, the hydrophobic binding pocket was constricted about the phenyl ring of the TS analogues, consistent with a regional contribution to TS stabilization. Coordination to the active site Mg^{2+} was also tighter, consistent with additional stabilization of the TS. The observation of an altered conformation of MR upon binding TS analogues agrees with reports by Schramm, that dynamic changes in protein structure are linked to catalysis (Schramm, 2005). As a result, it appears as if TS analogue inhibitors of MR truly are bound in a TS-like fashion, inducing changes in enzyme structure which may resemble those that accompany catalysis.

1.4 REVERSE THYMIDYLATE SYNTHASE

1.4.1 DISCOVERY AND ROLE

The number of nonredundant sequences within the UniProtKB/TrEMBL and UniProtKB/SwissProt databases is greater than 54 million and climbing (Ghasempur, *et al.*, 2014). A notable challenge in the field of protein science lies in that less than 1% of the annotated proteins in databases have been functionally characterized (Ghasempur, *et al.*, 2014). As an example within the MR subgroup, an open reading frame denoted *rTS* was found to have unassigned function. The open reading frame was assigned to the MR subgroup on the basis of similarity in encoded amino acid sequence, active site metal ion ligand, and the presence of a histidine residue in a homologous location with the catalytic His 297 of MR (Babbitt *et al.*, 1996).

Reverse thymidylate synthase (rTS) is a protein encoded by an mRNA from *Homo sapiens.* This protein accumulates in tumour cells and is overexpressed in cell lines resistant to the chemotherapeutic agents methotrexate (MTX) and 5-flurouracil (5-FU) (Dolnick, 1993; Srimatkandada, *et al.*, 1983; Dolnick, *et al.*, 1996; Dolnick, *et al.*, 1997; Kuo, *et al.*, 2008; Lin & Chow 2010). While establishing a quantitative PCR assay for the gene (*TS*) encoding human thymidylate synthase (TS), partial overlap of the *TS* gene was discovered at the chromosome location 18p.1.32 with another gene, which was designated reverse thymidylate synthase (*rTS*) (Dolnick, 1993). rTS is transcribed in the reverse direction of TS, hence its name. The *rTS* gene has marginal overlap between its 3′-untranslated region and the last intron of *TS* (Dolnick, 1993). The *rTS* gene has received a considerable amount of clinical interest within the field of cancer research over the past 20-30 years because of its overexpression in MTX- and 5-FU- resistant cell

lines, yet its function remains largely unknown (Dolnick, 1993, Kuo, *et al.*, 2008; Srimatkandada, *et al.*, 1983; Dolnick, *et al.*, 1996; Dolnick, *et al.*, 1997; Lin & Chow 2010). MTX is an antifolate inhibitor of dihydrofolate reductase (DHFR), binding with 400-fold greater affinity than its substrate 7,8-dihydrofolate, while 5-FU is a prodrug that is converted intracellularly to 5-fluorodeoxyuridine 5′-monophosphate (FdUMP), an irreversible inactivator of TS (Rajagopalan, *et al.*, 2002; Rose, Farrell, & Schmitz, 2002). The metabolic cycle connecting TS and DHFR, as well as the corresponding interplay of MTX and 5-FU, is shown in **Figure 1.7**. Both DHFR and TS are essential for DNA biosynthesis, specifically in the *de novo* synthesis of 2′-deoxythymidine-5′-monophosphate (dTMP), making them attractive targets for chemotherapy agents.

Since the initial discovery of the *rTS* gene, three protein isoforms have been identified as products of alternative mRNA splicing, which are denoted as *Hs*rTSα, β, and γ (Dolnick & Black, 1996; Dolnick, *et al.*, 1997). Structurally, all three isoforms of rTS have an identical 341 C-terminal amino acid sequence and differ only in their N-terminal sequences and lengths. At the N-terminus, rTSα contains an additional 20 residues, rTSβ has 75 additional residues, and rTSγ has a further 27 residues added to the N-terminal end of the rTSβ isoform. The primary structures are shown schematically in **Figure 1.8** (Wichelecki, *et al.*, 2014).
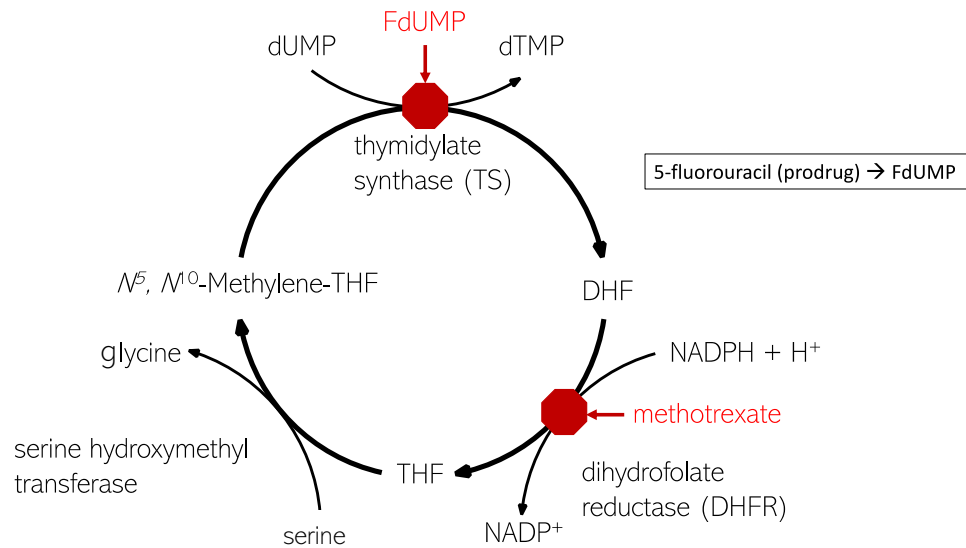
Figure 1.7 Metabolic cycle linking dihydrofolate reductase and thymidylate synthase. dTMP is produced via TS catalysis during DNA biosynthesis. The effects of the competitive inhibitor MTX and the irreversible inactivator FdUMP on DHFR and TS catalytic activity, respectively are shown (Voet, Voet, 2004).
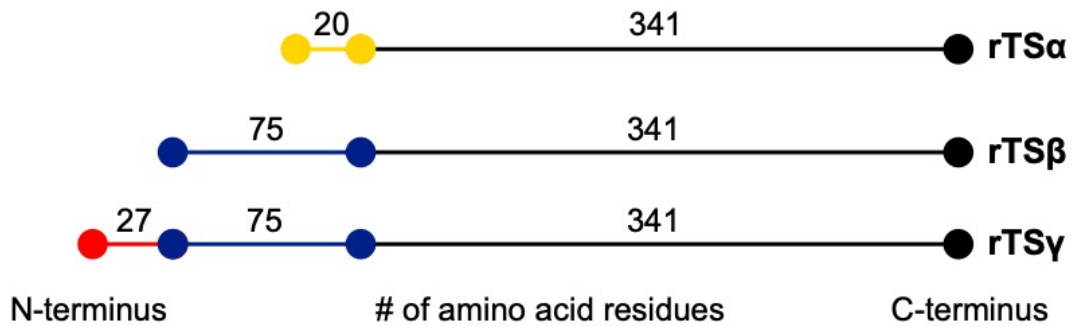
Figure 1.8 rTS isoforms as derived from splice variants and their relative amino acid sequence lengths. Different splice variants of the *rTS* gene yielding the HsrTS isoforms α, β, and γ.

1.4.2 FUNCTION OF rTS

1.4.2.1 Clinical Observations and Associations

The rTS protein from *H. sapiens* has garnered interest in cancer research (Dolnick, 1993, Kuo, *et al.*, 2008; Srimatkandada, *et al.*, 1983; Dolnick, *et al.*, 1996; Dolnick, *et al.*, 1997; Kuo, 2008; Lin & Chow, 2010). Dolnick, who pioneered much of the research into the *rTS* gene, discovered the first splice variants of rTS, denoting them as rTSα and rTSβ (Dolnick & Black, 1996). At that point, the rTSβ isoform attracted the most attention due to the finding that rTSβ accumulates in the H630 colon tumour cell line (Dolnick & Black 1996). Further, rTSβ was overexpressed 40-70-fold in a 5-FU-resistant H630 subline, which is far greater than that observed for the other rTS isoforms and TS, which were overexpressed only 7-8-fold (Dolnick & Black, 1996). Indeed, the authors concluded that the elevation of rTSα, rTSγ, and TS was proportional to gene activation. It was also found that rTS RNA is present in a large variety of cultured normal and tumor human cell lines. Furthermore, rTS was overexpressed in the human K562 B1A cell line that was resistant to MTX (Dolnick, *et al.*, 1996).

A number of studies over the past 20 years elaborated upon the clinical relevance of the rTS proteins, while enzymology studies were not pursued. Notably, nuclear *rTSβ* expression was found to correlate with 5-FU resistance, suggesting that it might serve as a potential biomarker for the resistance phenotype in both breast and colon cancer (Kuo, *et al.*, 2008; Lin & Chow, 2010). In clinical studies, rTSβ was shown to be expressed in breast cancer tissue, but not the surrounding normal tissues (Kuo, *et al.*, 2008). Other clinical studies showed a statistically significant correlation between rTSβ expression levels and a decrease in the five-year survival rate of colon cancer patients (Lin & Chow, 2010). Dolnick and colleagues went as far as to indicate that the effect of rTSβ is phosphorylation state-dependent (e.g., at Ser 121) in the chemotherapy resistant phenotype. A 10-fold decrease in the phosphorylation at Ser 121 in the resistance phenotype of rTSβ was shown, suggesting that the protein could serve as a potential biomarker (Dolnick, *et al.*, 2005). Interestingly, for colon and breast cancer patients, an association between rTSβ expression and poor prognoses was reported (Lin & Chow, 2010). Overall, these observations prompted subsequent investigation of the function of rTS.

1.4.2.2 Cell Biology Studies

In early non-enzymatic functional studies, TS was shown to be physically associated with proteins of the same molecular weight as rTSα and rTSβ (Dolnick, *et al.*, 1996). Thus, it was postulated that the rTS isoforms may function as modulators of TS activity by physical association *in vivo*. Further hypotheses pointed towards the importance of the stoichiometric ratio of rTSα to rTSβ in conferring alternative functions

to TS, such as increased sensitivity to certain TS inhibitors and resistance to others (Dolnick, *et al.*, 1996). Dolnick hypothesized that the rTS isoforms have the potential to form protein-protein interactions, which is similar to the observed tetrameric/octameric quaternary structure of MR (Fee, Hegeman, & Kenyon, 1974; Tsou, *et al.*, 1989; Sagy & Bearne, unpublished). Alternative functions of rTSβ were then proposed assuming the potential for protein-protein interactions with TS. Notably, the hypothesized ability of rTSβ to modulate folate binding to TS through protein–protein interactions resulting in a downregulation of TS function (Dolnick, *et al.*, 1997). Research on rTSβ to this point suggested that the protein had a regulatory effect on TS activity, yet the precise function remained undetermined.

In the early 2000s, Dolnick and co-workers proposed the novel function of rTSβ in the enzymatic synthesis of lipophilic metabolites derived from methionine (Dolnick, *et al.*, 2003). Indeed, a functional assay demonstrated that rTSβ could catalyze the transfer of the carboxyl carbon of methionine from S-adenosylmethionine to a lipophilic acceptor molecule *in vitro* (Dolnick, *et al.*, 2003). This proposal of a novel function for rTSβ arose from the abundance of lipophilic compounds derived from methionine present in rTSβ-overexpressing cells (Dolnick, *et al.*, 2003). Consequently, rTSβ was proposed to be involved in the regulation of TS through a novel methionine-based signaling pathway. The signaling molecules were proposed to share properties with the acyl-homoserine lactones, a group of molecules used in bacteria for regulation of gene expression (Dolnick, *et al.*, 2003). As a result, Dolnick proposed that these molecules produced by rTSβ could downregulate *TS* gene expression (Dolnick, *et al.*, 2005).

Finally, the additional 27 residues at the N-terminus of the rTSγ isoform were proposed to constitute a mitochondrial signaling sequence. The hypothesis that the only difference between rTSβ and rTSγ is the N-terminal mitochondrial localization signaling sequence would indicate that the β and γ isoforms of rTS have similar enzymatic functions (Liang, *et al.*, 2005). This hypothesis was promptly discarded with identification of the 27 N-terminal residues of rTSγ as part of the highly conserved first β-strand of the N-terminal capping domain and the "20s" loop of the MR subgroup (Wichelecki, *et al.*, 2014; Bearne & St. Maurice, 2017; Yew, *et al.*, 2006a; Yew, *et al.*, 2007; Rakus *et al.*, 2008). In fact, it was found that the stability conferred by the 27 N-terminal amino acids is essential for production of soluble protein (Wichelecki, *et al.*, 2014). Thus, the 27 additional residues at the N-terminus of rTSγ are a necessary structural feature of the enzyme architecture and likely confer FucD activity to the enzyme.

1.4.2.3 Structural and Mechanistic Studies

In an effort to establish the functional role of the rTS proteins, Gerlt and colleagues conducted an in-depth characterization of the rTSγ isoform, which was determined to have L-fuconate dehydratase (FucD) activity (Wichelecki, *et al.*, 2014). A 71% amino acid sequence similarity and 52% identity was observed between rTSγ and the canonical FucD from *Xanthomonas campestris* (Wichelecki, *et al.*, 2014). **Figure 1.9** shows the X-ray crystal structure of rTSγ at 1.74-Å resolution (PDB: 4A35) and highlights the 27 N-terminal residues appended to the γ isoform, which are absent in the β isoform. Previous investigations on enzymes within the ENS revealed that enzymes

with up to 79% identity can catalyze different reactions (Wichelecki, *et al.,* 2014).

Therefore, an experimental characterization of the actual reaction catalyzed by rTSγ was

required. Based on the X-ray crystal structure, Gerlt *et al.* found that rTSγ shares many

features with FucD (Wichelecki, *et al*., 2014). First, it has a similar sequence within the

$(\beta/\alpha)_7\beta$-barrel core domain. Like FucD, a conserved KxK motif, consisting of Lys 220-

Val 221-Lys 222 at the end of the second β-strand of the barrel domain is used for base-

catalyzed proton abstraction. The acidic residues Asp 250, Glu 276, and Glu 305 at the

ends of the third, fourth, and fifth β-strands of the barrel, respectively, provide ligands for

the Mg^{2+}. Finally, a His-Asp dyad (H355-D328) exists at the end of the seventh and sixth

β-strands of the barrel domain, respectively, which acts as the general acid catalyst for

dehydration of the enediolate intermediate (Wichelecki, *et al.*, 2014; Neidhart, *et al.*,

1991).

In functional studies, rTSγ was screened against 72 acid sugars for dehydration

activity, which revealed that rTSγ was most active as a FucD, though it also showed

activity with substrates L-galactonate, L-arabinarate, D-arabinonate, and D-ribonate

(Wichelecki, *et al.*, 2014). The FucD activity of rTSγ, however, was 10-fold lower than

that of the canonical *Xc*FucD. The difference was attributed to the modest sequence

identity shared between the two and possibly differences in expression of recombinant

rTSγ in *E. coli*, such as the lack of required post-translational modifications (Wichelecki,

et al., 2014). Further, the proposed function of rTSβ in transferring a carboxylate group

from S-adenosylmethionine to a lipophilic acceptor was negated by the fact that MR

subgroup members function in abstraction of protons alpha to a carboxylate group, which

is not a partial reaction known to result in decarboxylation (Wichelecki, *et al.*, 2014). In

the experimental work conducted by Gerlt and coworkers, rTSγ was produced as a

soluble hexahistidine-tagged construct at levels usable for experimental characterization.

On the other hand, rTSβ was largely insoluble with only a minor amount of soluble

protein being obtained as a "tagless" construct. The lack of soluble rTSβ is the main

obstacle prohibiting discovery of function for this isoform. Nonetheless, functional

assignment to rTSγ was a key finding in suggesting possible function of the rTS isoforms.

In pursuit of the function of the clinically relevant HsrTSβ, similar substrate

screening assays were conducted with the β isoform. However, no positive results were

obtained using the library of 72 acid sugars, indicating that rTSβ is likely devoid of

catalytic activity (Wichelecki, $et\ al.$, 2014). Clearly, the 27 N-terminal amino acids are

required for catalysis. Without the N-terminal region, it is possible that conformation of

the active site of rTSβ could be dramatically altered, and/or the hydrophobic active site

may be solvent exposed, leading to protein aggregation. The dependence on the N-

terminal region was further illustrated by Gerlt $et\ al.$, who showed that truncated variants

of rTSγ lacking only 5, 8, or 12 residues from the N-terminus were all insoluble

(Wichelecki, $et\ al.$, 2014). The lack of stability resulting from the loss of the N-terminal

sequence yields an alternative hypothesis for the function of the rTSβ isoform. Possibly,

the solvent exposed active site could allow for alternative function in metabolism or

protein-protein interactions.

To explore the stability of the isoforms in more detail, differential scanning

fluorimetry (DSF) was used to characterize both rTSγ and rTSβ. These studies showed

that rTSβ was more thermolabile than rTSγ ($\Delta T_m = 5\ °C$), which was in accord with the

observed loss in stability and low protein production for the β isoform (Wichelecki, $et\ al.$,
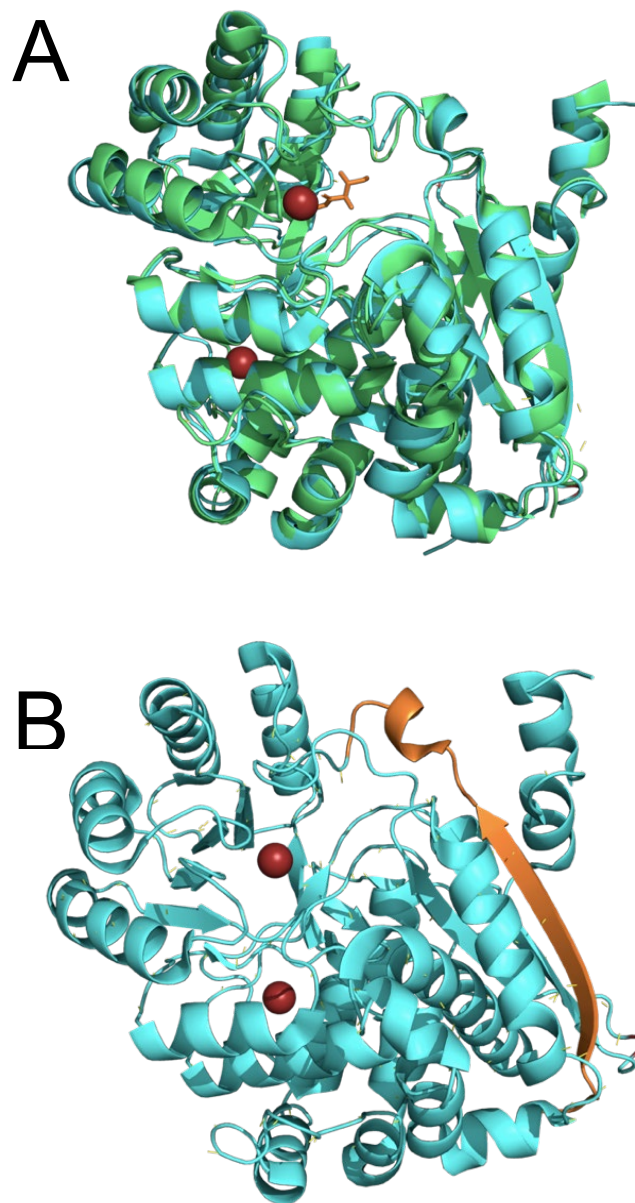
Figure 1.9. Structure of rTSγ. (A) An overlay of the structures of rTSγ (PDB 4A35, blue) and *Xc*FucD bound to the substrate analogue L-erythronohydroxamate (PDB 2HXT, green), is shown, which highlights structural similarities. (B) The X-ray crystal structure of rTSγ (blue) is shown. The 27 N-terminal residues of rTSγ that are absent in the β isoform are highlighted in orange (PDB entry 4A35). Active site Mg^{2+} ions are shown in sphere representation (red). Both FucD and rTSγ have bound Mg^{2+} ions in the active site (the upper sphere in (A) and (B)), while rTSγ has an additional bound Mg^{2+} ion at a distant site.

2014). In the presence of D-erythronohydroxamate, an intermediate analogue, an increase

in the thermal melting temperature of rTSγ was observed; however, rTSβ did not show

thermal stabilization in the presence of D-erythronohydroxamate (Wichelecki, *et al.*,

2014). Typically, the use of an active site ligand increases the thermal stability of the

enzyme-ligand complex, which yields an increased melting temperature (Niesen,

Berglund, & Vadadi, 2007). Thus, the active site has been altered in some way that

obviates binding of the typical ligands (Wichelecki, *et al.*, 2014). These findings indicate

that the N-terminal 27 amino acids are necessary for enzymatic activity, both in binding

and catalysis (Wichelecki, *et al.*, 2014).

The association of rTSβ overproduction and resistance to the chemotherapy

agents MTX and 5-FU suggests a functional role for rTSβ. Consequently, Gerlt and co-

workers screened MTX and 5-FU with rTSβ as possible substrates using various analysis

techniques (Wichelecki, *et al.*, 2014). Dolnick and colleagues proposed that the glutamate

side chain of MTX has a proton alpha to the carboxylate group may serve as a site for

deprotonation by rTSβ. Also, the exposed active site of rTSβ could accommodate MTX,

which has roughly twice the molecular weight of L-fuconate. However, in the presence of

D_2O, no exchange of the α-proton was observed. Further, no rTSβ-catalyzed

defluorination of 5-FU as an alternative substrate was observed, indicating that rTSβ is

catalytically inactive with 5-FU (Wichelecki, *et al.*, 2014). However, these studies had

not explored the possibility that rTSβ could bind MTX, 5-FU, or other proteins, such as

TS or DHFR *in vivo* without enzymatic activity. Since TS was shown to physically

associate with proteins of the same molecular weight as rTSα and rTSβ by

coimmunoprecipitation (Dolnick, *et al.*, 1996), and also possibly with DHFR (Dolnick, *et*

al., 1996), there is the possibility that such binding interactions play a role in the function of rTSβ. However, to investigate the role of rTSβ in such interactions, a method of producing the protein in soluble form and in sufficient amounts is required. In Chapter 4, I describe attempts to develop constructs of rTSγ that, upon protease treatment, yield improved amounts of rTSβ.

1.5 OVERVIEW OF THIS WORK

The highly conserved MR subgroup of the ENS has garnered a considerable amount of attention over the last 50 years. Primarily, the MR subgroup offers a highly characterized group of enzymes for which further characterization can yield deeper insights into how active site achieve high proficiency for catalyzing heterolytic C–H bond cleavage reactions.

MR is extremely proficient at catalyzing the interconversion of its enantiomeric substrates (*R*)- and (*S*)-mandelate, a reaction that has high thermodynamic and kinetic barriers. MR binds the enolate-like transition state of its catalyzed reaction 15 orders of magnitude greater than it binds substrate. The characterization of MR inhibitors has long been studied, but most recently culminated in the finding that phenylboronic acid (PBA) and its *p*-halogen-substituted derivatives have unprecedented potency in binding to MR (Sharma, *et al.*, 2020). These MR inhibitors bind to MR in a novel fashion with dative bond formation between the catalytic Brønsted base His 297 and the boron atom. To build upon these findings, I explored a series of chloro- and fluoro-substituted PBAs, ascertaining a structure-activity relationship that can be used for inhibitor design in the

future. Interestingly, I identified inhibitors that bind to MR with 100-fold greater affinity than typically observed for previous TS analogue inhibitors of MR (Chapter 2).

Enzymes of the MR subgroup in the ENS catalyze a conserved partial reaction: the Mg^{2+}-assisted, Brønsted base-catalyzed abstraction of the α-proton from a carbon acid substrate, forming an enolate intermediate (Gerlt, *et al.*, 2012). The buildup of negative electrostatic character accompanying the formation of the reactive intermediate is counteracted by a localized region of positive charge at the active site. This region comprises both lysines of the conserved KxK motif and the divalent metal Mg^{2+}. Recently, the Bearne lab has shown that the Brønsted acid-base lysine *destabilizes* the protein as a whole (Kumar & Bearne, unpublished). To further investigate this phenomenon, I generated a series of proteins variants at the KxK position in MR (K164M, K164R), FucD (K218M), TarD (K182M), and TGD (K195M) and observed the thermal stability (Chapter 3). It was found that the KxK residue *stabilizes* the proteins of the MR subgroup in contrast to the destabilizing effect of the adjacent KxK residue.

A gene encoding three human MR subgroup protein isoforms (rTSα, β, and γ) was found to partially overlap the gene encoding DNA biosynthetic enzyme TS (Dolnick, 1993). Of these proteins, the rTSβ isoform was found to accumulate in cancer cell lines resistant to the chemotherapeutic agents MTX and 5-FU (Dolnick *et al.*, 1996; Wichelecki, *et al.*, 2014). The isoforms of rTS have closely related structures as the β and γ isoforms differ only by 27 amino acids at the N-terminus of the protein. Interestingly, the γ form is catalytically active and classified as a FucD, while the aberrantly expressed β isoform lacks catalytic function and is highly unstable. In this work, I explore approaches to obtain rTSβ using various approaches including engineering protease

recognition sites into the relatively stable rTSγ and the addition of a protein solubilizing

fusion tag MBP (Chapter 4). These studies suggest approaches that might afford better

production of rTSβ for future analyses.

CHAPTER 2 INHIBITION OF MANDELATE RACEMASE BY CHLORO- AND FLUORO-SUBSTITUTED PHENYLBORONIC ACIDS

2.1 INTRODUCTION

2.1.1 USE OF BORON IN INHIBITOR DESIGN

Boron-containing inhibitors have a rich history of use for both inhibiting enzymes and binding to proteins with profound potency, sparking recent interest for their use in clinical inhibitor design. Boron shows great potential for use in clinical applications because of its relatively low abundance in nature and its low toxicity when used for anticancer, antibacterial, antiviral, and other clinical applications (Trippier & McGuigan, 2010; Baker, *et al.*, 2009; Whyte, Vilar, & Woscholski, 2013). The design of boron-containing inhibitors relies on the stability of boron-containing complexes, the known reversible complex formation with other biomolecules, and the ability to use boron as a bioisostere of carbon (Trippier, & McGuigan, 2010).

The most advantageous use of boron in inhibitor design arises from its strong Lewis acidity. The electron-vacant p orbital of boron can readily accept electron donation from Lewis base groups (typically amines and alcohols in a biological setting) to form a coordinate covalent (dative) bond (Frenking, 2015). Formation of a dative bond to boron causes a conversion in hybridization state from a neutral sp^2-hybridized state to an anionic sp^3-hybridized tetrahedral adduct (Frenking, 2015). In consideration of boron-containing inhibitors binding to proteins, studies have shown that dative boron adducts occur with serine, threonine, lysine, or histidine residues (**Scheme 2.1**) (Tsilikounas, Kettner, & Bachovchin, 1992; Tsilikounas, Kettner, & Bachovchin, 1993; Tulinsky &

Blevins, 1987; Zervosen, *et al.*, 2011). Specifically, boronic acids have shown particular promise for use in enzyme inhibition since the hydroxyl groups covalently linked to the boron center can form hydrogen bonds with other functional groups at the enzyme active site to enhance molecular recognition (Windsor, *et al.*, 2018).

Scheme 2.1. Common adducts formed between side chains of amino acids in proteins and boronic acid ligands.
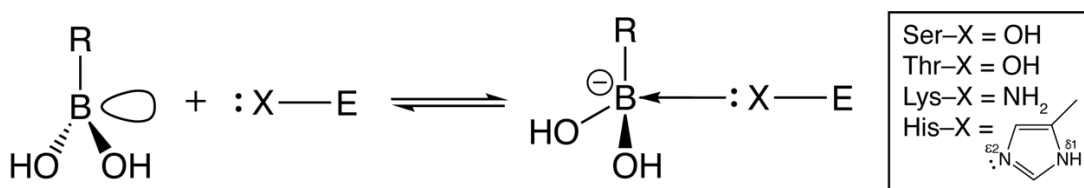
The coordinate covalent bond formed with the boron center has been shown to have sufficient strength to be effectively irreversibly bound to the protein (Whyte, Vilar, & Woscholski, 2013). This prominent form of binding was shown in inhibitor design studies for the thoroughly characterized serine hydrolases, where dative bond formation occurs between the serine Lewis base at the active site and the boronic acid ligand (Whyte, Vilar, & Woscholski, 2013; Tsilikounas, Kettner, & Bachovchin, 1992; Lanier, *et al.*, 2017). The formation of the dative bond and concurrent conversion to an sp^3-hybridized anionic adduct mimics the properties of the transition state of the hydrolytic enzyme-catalyzed reaction, which can further bolster the binding affinity for the ligand.

The chemical environments that yielded the dative bond-driven high affinity binding seen in the aforementioned studies were hypothesized to apply to the enzymes catalyzing heterolytic cleavage of a C–H bond. Enzymes of the ENS, as described in Chapter 1, may offer the exact environment needed to achieve the high inhibitor binding

affinity, with specific coordination by the Brønsted acid-base residues that catalyze deprotonation at the α-C of the substrate (Landro, *et al.*, 1991; Landro, *et al.*, 1994). Furthermore, design of inhibitors where the boronic acid is positioned as a carboxylate bioisostere at the position of proton abstraction in these enzymatic systems may promote advantageous interactions that yield greater binding affinity. To assess this, Bearne and colleagues applied this boronic acid inhibitor design strategy to the ENS enzyme MR using phenylboronic acid (PBA) (Sharma, *et al.*, 2020).

2.1.2 INHIBITION OF MR BY PBA

Preliminary research found PBA to potently inhibit MR with similar affinity to known transition state analog inhibitors (BzH, Cfn, and α-HBP) (Sharma, *et al.*, 2020; Nagar & Bearne, 2015). Variation of the substituent at the *para*-position revealed that small electron withdrawing substituents increased the affinity of the interaction by ~1-2 orders of magnitude. Furthermore, isothermal titration calorimetry (ITC) and X-ray crystallographic studies showed that MR bound PBA similarly to the transition state analogues such that the Brønsted acid-base catalyst His 297 was coordinated to the electrophilic boron center (**Figure 2.1**). The two hydroxyl groups of PBA were found to form hydrogen bonds or electrostatic interactions with catalytic residues His 297, Lys 166, Lys 164, and Glu 317, while also occupying the fifth and sixth coordination sites of the divalent magnesium cation.

Most interestingly, through X-ray crystallography and DFT calculations PBA was shown to take part in a dative bond formed from the His 297–$N^{\varepsilon 2}$ to the boron center in
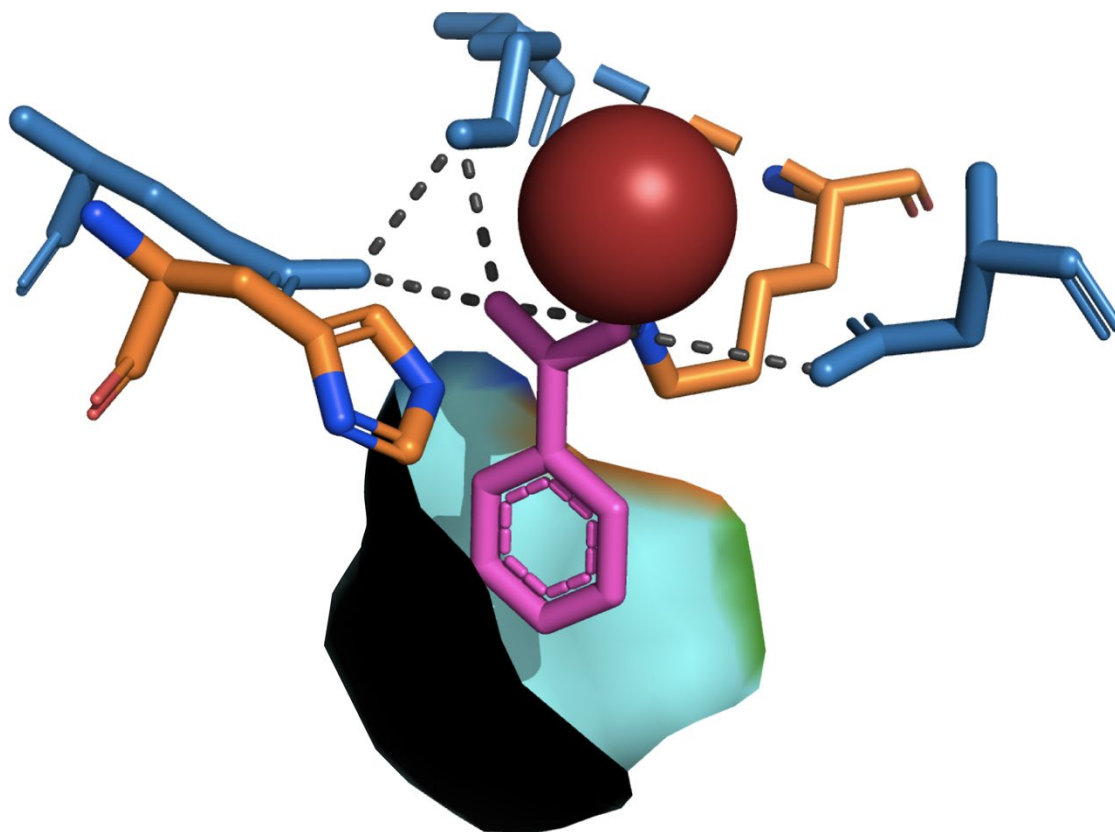
Figure 2.1. Active site architecture of MR with bound PBA. Representative view of PBA (magenta) interacting (dotted lines) with active site residues (stick representation) and Mg^{2+} (sphere) (PDB entry 6VIM). Brønsted acid-base catalysts Lys 166 and His 297 (stick representations) are shown in orange. Select residues of the hydrophobic pocket of MR (Leu 18, Val 22, Val 29, Phe 52, Tyr 54, Leu 298, Leu 319, and Leu 321, surface representation) are shown interacting with the phenyl ring of PBA.

50% occupancy of the eight active sites of the MR octamer (**Figure 2.2**) (Sharma, *et al.*, 2020). This finding was supported by ^{11}B NMR studies, where a strong upfield change in chemical shift was observed, which is characteristic of additional electron density on the boron consistent with a change from an sp^2 to an sp^3 conformation (Tsilikounas, Kettener, & Bachovchin, 1993; Baldwin, *et al.*, 1991; Zhong, *et al.*, 1991). The potential dative bond interaction was proposed, in part, to contribute to the high affinity binding observed for PBA and its substituted derivatives, as well as numerous H-bonding interactions with the two hydroxyl groups on the boron. DFT calculations suggested that PBA is recognized as a transition state analogue because of the protonated state of Lys 164 and Lys 166 ($R–NH_3^{+)}$), and the conjugate base form of Glu 317 ($R–CO_2^-$) expected for the reaction intermediate formed during MR catalysis (Sharma, *et al.*, 2020).

It is most likely that PBA binding to MR capitalizes on multiple interactions accompanying the binding of TS analogues, while the electrophilic boron center allows for enhanced binding potency (Sharma, *et al.*, 2020). Additionally, the binding mode observed in X-ray crystallographic studies indicates that the phenyl ring of PBA binds to the hydrophobic binding pocket of residues in MR (Nagar, *et al.*, 2015; Bearne, 2017; St. Maurice & Bearne, 2004). The studies on substituent effects that uncovered the unprecedented inhibition of MR by 4-chloroPBA similarly found that all *p*-halogen-substituted PBAs increased the potency of binding to MR (Sharma, *et al.*, 2020). Through kinetic experiments with (*R*)- and (*S*)-4-chloromandelate, it was found that the enhanced binding of 4-chloroPBA did not simply arise from fortuitous interactions in the active site, but instead likely arose from the electron-withdrawing properties of the substituent. Taken together, it is likely that the electron-withdrawing effects of a *p*-substituent on
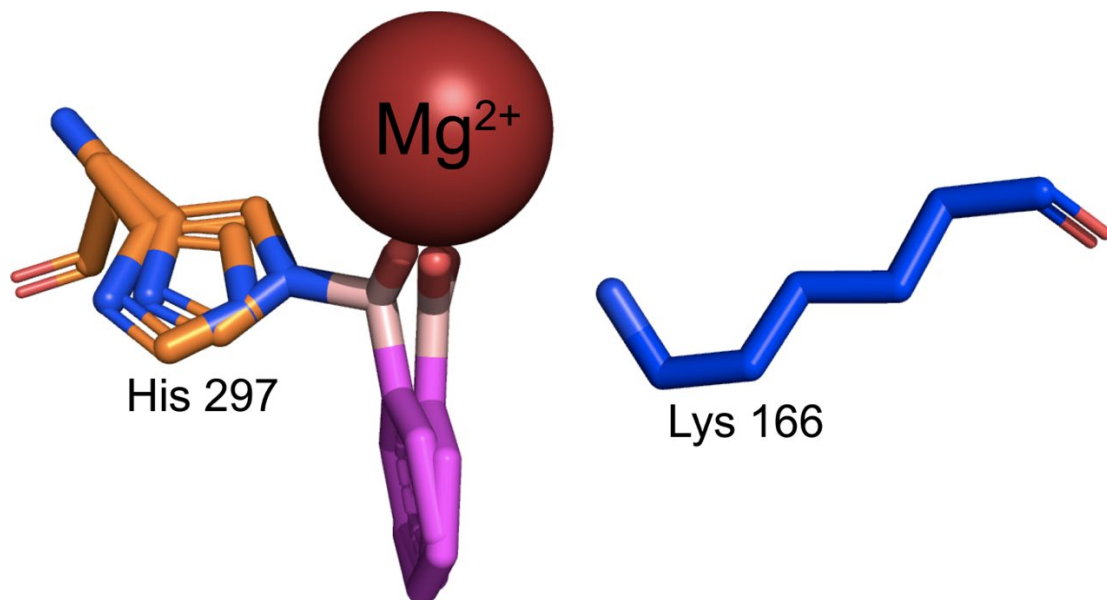
Figure 2.2. Formation of a dative bond in MR with bound PBA. Shown is a representative X-ray crystal structure of MR with PBA bound (PDB file 6VIM). The observed trigonal planar and tetrahedral conformation of PBA (magenta, stick representation) is observed bound to MR observed at certain active sites of the MR octamer. Brønsted acid-base residues Lys 166 and His 297 are shown in blue and orange, respectively (stick representation). The Mg^{2+} ion is shown in red (space filling representation).

PBA potentiate binding to MR by increasing the electrophilicity at the boron center, as well as affording enhanced binding at the hydrophobic pocket.

Based on these findings, determination of a structure-activity relationship for halogen-substituted PBAs may uncover additional binding factors that contribute to the design of more potent MR inhibitors. Further, given the inhibitory potency observed for PBA and its derivatives characterized by Bearne and colleagues (Sharma, *et al.*, 2020), further investigations regarding formation of an active site dative bond to boronic acids in MR could yield promising results. In this work, a series of chloro- and fluoro-substituted PBAs have been characterized to determine a structure-function relationship for PBA inhibition. Of the various chloro- and fluoro-substituted PBAs, 3,4-dichloroPBA was found to inhibit MR with an affinity ~3 orders of magnitude greater than TS analogues, and 5-fold greater than 4-chloroPBA itself. Complementarily, ^{11}B NMR studies supported presence of a dative bond formed from MR to the liganded 3,4-dichloroPBA. Overall, inhibition of MR by chloro- and fluoro-substituted PBAs offers contribution to the enzymologist's toolkit as a method for designing inhibitors of enzymes of the ENS and those catalyzing heterolytic C–H cleavage.

2.2 MATERIALS AND METHODS

2.2.1 GENERAL

All reagents were purchased for Sigma-Aldrich Canada Ltd. (Oakville, ON, Canada), unless otherwise stated. Chloro- and fluoro-substituted phenylboronic acid compounds were obtained commercially from Acrotein ChemBio Inc. (Hoover, AL). Circular dichroism (CD) spectra were collected using a JASCO J-810 spectropolarimeter

(Jasco Inc., Easton, MI). UV-visible and OD_{600} spectrophotometric readings were obtained using a HP 8453 UV-visible spectrophotometer. Nuclear magnetic resonance (^{1}H and ^{11}B NMR) spectra were obtained using a Bruker AV 500 MHz spectrometer at the Dalhousie University Nuclear Magnetic Resonance Research Resource Centre (NMR-3).

2.2.2 PROTEIN PRODUCTION

Recombinant wild-type MR from *Pseudomonas putida* was produced in *E. coli* BL21 Star™(DE3) cells transformed with a pET-52b(+) plasmid (Novagen, Madison, WI) containing the MR open reading frame, as described by Narmandakh and Bearne (2010). The resulting vector encodes the MR gene product as a fusion protein with an N-terminal StrepII-tag. Small cultures (5 mL) of *E. coli* BL21 Star™(DE3) cells were grown overnight at 37 °C from glycerol stocks in sterile LB media containing ampicillin (100 µg/mL). Large-scale production cultures (2×1 L LB media, 100 µg/mL ampicillin) were inoculated with the overnight cultures at a volume of 10 mL per litre of LB media and incubated at 37 °C with shaking (225 rpm) for 8 h. Cells were harvested using centrifugation ($3795 \times g$, 10 min, 4 °C), and washed with wash buffer (100 mM Tris-Cl, 150 mM NaCl, 1 mM EDTA, pH 7.5). Cell pellets were stored at –20 °C for future use. Frozen cell pellets were thawed and resuspended in cold wash buffer (~35 mL). The cell suspension was kept on ice and sonicated for 6×30 s using a Branson Sonifier 250 with 1 min "resting" periods between sonication intervals (setting 5.5, 1-s bursts). The cell lysate was clarified by ultracentrifugation ($146550 \times g$, 35 min, 4 °C) and applied to a column containing StrepTactin Superflow affinity resin (10 mL) (IBA Life Sciences,

Göttingen, Germany) connected to an ÄKTA fast protein liquid chromatography (FPLC) system (GE Healthcare, Baie-Durfé, QC). After washing the column with wash buffer (15 × column volume), enzyme was eluted by addition of wash buffer containing desthiobiotin (2.5 mM). Eluted enzyme was dialyzed over 3 × 8 h against MR assay buffer (100 mM HEPES, 3.3 $MgCl_2$, pH 7.5), with each 8-h dialysis using assay buffer (500 mL). Protein was aliquoted (1 mL) and stored at –20 °C for future use.

2.2.3 PROTEIN CHARACTERIZATION AND QUANTIFICATION

Protein concentrations were determined using Bradford assays conducted according to the manufacturer's directions (Bio-Rad Laboratories, Missisauga, ON) with bovine serum albumin standards. Alternatively, the protein concentration was determined using UV absorbance at 280 nm using an extinction coefficient of 53 400 $M^{-1}cm^{-1}$ as determined by the ProtParam web tool from ExPASy (Gasteiger, *et al.*, 2003) for MR under the assumption that all cysteine residues were reduced, as can be seen in the X-ray crystal structures. Protein purity was assessed using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (10%) with Coomassie brilliant blue (R-250) staining.

2.2.4 KINETIC ASSAYS

MR kinetic assays were conducted using a CD-based assay as previously described (Narmandakh & Bearne, 2010; Sharp, Hegeman, & Kenyon, 1979). MR was diluted in HEPES buffer (100 mM, pH 7.5) containing $MgCl_2$ (3.3 mM) (assay buffer) with 0.01% w/v bovine serum albumin (BSA) on ice. Substrate samples were incubated

at 25 °C prior to initiation of the enzymatic reaction. At each assay point, 100 μL MR

was added to a 1900 μL solution of substrate ((R)- or (S)- mandelate) in a quartz cuvette

with a 1-cm light path, mixed by inversion, and subsequently analyzed using CD

spectroscopy. Substrate samples (0.5-15.0 mM) were incubated at 25 °C. The change in

ellipticity at 262 nm was measured over a time course of three minutes to observe

reaction progress. A final MR concentration of 3.64 nM was used for assays.

2.2.5 DATA ANALYSIS OF ENZYMATIC REACTION KINETICS

Values for the maximum reaction velocity (V_{max}) and the Michaelis constant (K_m)

were determined from plots of initial velocity (v_i) versus substrate concentration ([S]).

The Michaelis-Menten equation (**eqn. 2.1**) was fit to the initial velocity data using non-

linear regression analysis and the *KaleidaGraph* v. 4.02 software (Synergy Software,

Reading, PA). All kinetic data were obtained in triplicate and average values are reported.

The reported error is the standard deviation from the triplicate determinations. Values of

k_{cat} were determined by dividing the V_{max} values by the total enzyme concentration ([E]$_T$).

$$v_i = \frac{V_{max}[S]}{K_m + [S]} \qquad\qquad (2.1)$$

2.2.6 IC$_{50}$ EXPERIMENTS

The concentrations of inhibitor that yield an enzymatic reaction velocity that is

50% of the uninhibited reaction velocity (IC$_{50}$) were obtained by conducting an inhibitor

dose-response experiment using the previously described CD-based assay (Section 2.2.4)

with 1 mM (R)- mandelate ([S] ≈ K_m). **Eqn. 2.2** was fit to initial velocity data using

Kaleidagraph v. 4.02 software (where *n* is the Hill number, indicating the cooperativity in binding, binding is independent if *n* =1). IC_{50} values are related to inhibition constants (K_i) by **eqn. 2.3,** assuming competitive inhibition. The inhibitor concentration ranges used for IC_{50} experiments are as follows: 0-40 nM 3,4-dichloroPBA, 0-200 nM 3,4-dichloro-2-fluoroPBA, 0-1 μM 4-chloro-2-fluoroPBA, 0-800 nM 4-chloro-2,3-difluoroPBA, 0-150 nM 4-chloro-3-fluoroPBA, 0-1 μM 2,4-dichloroPBA, 0-1.2 μM 3,4,5-trichloroPBA, 0-800 nM 3-chloroPBA, 0-50 μM 2-chloroPBA, 0-200 nM 4-chloro-3,5-difluoroPBA, 0-2.5 mM 4-chloro-2,6-difluoroPBA.

$$\frac{v_i}{v_0} = \frac{IC_{50}^n}{IC_{50}^n + [I]^n} \tag{2.2}$$

$$K_i = \frac{IC_{50}}{1 + \left(\frac{[S]}{K_m}\right)} \tag{2.3}$$

2.2.7 INHIBITION STUDIES

The inhibition constants (K_i) for 3,4-dichloroPBA and 4-chloro-2,6-difluoroPBA were obtained using the previously described CD-based assay (Section 2.2.4). Assays were conducted under varying concentrations of 3,4-dichloroPBA (15.0, 30.0, and 45.0 nM), and 4-chloro-2,6-difluoroPBA (0.4 and 0.8 mM). Competitive inhibition constants (K_i) for the inhibition of MR by these PBAs were determined from plots of the apparent K_m/V_{max} values versus inhibitor concentration according to **eqn. 2.4** using *Kaleidagraph* v. 4.02 software.

$$v_i = \frac{V_{max}[S]}{K_m\left(1 + \frac{[I]}{K_i}\right) + [S]} \tag{2.4}$$

2.2.8 ISOTHERMAL TITRATION CALORIMETRY

Solutions of each of the chloro- and fluoro-substituted PBAs were prepared in the buffer from the final protein dialysis to eliminate heat signals that could arise from buffer mismatch. The enzyme and ligand solutions were degassed (Microcal Thermovac) for 15 min before loading into the sample cell (1.46 mL) and injection syringe (297 μL), respectively. The cell contained wild-type MR at a concentration of 40 μM and the injection syringe contained ligand at a concentration of 500 μM. Titrations of ligand into MR were conducted at 20 ºC. A total of 29 injections of 10 μL were made at 240 s intervals. The heat released from the initial injection (5 μL) was excluded from data analyses to minimize the effect of titrant diffusion from the syringe tip during equilibration of the system. Titrations of ligand solutions (500 μM) into the sample cell containing only buffer were conducted as a control to correct titration data for the heat of dilution and mixing. Nonenzymatic ligand control isotherms were subtracted from the binding isotherms prior to curve fitting. All ITC titrations were carried out in triplicate and the average reading is reported with the standard deviation as the error.

Association constants (K_a) and thermodynamic parameters for ligand binding (ΔH and ΔS) were obtained by fitting the calorimetric data to a single-site model using *Origin* v. 7.0 software (OriginLab, Northhampton, MA). Dissociation constants (K_d) were obtained from their reciprocal relationship to the K_a values obtained (**eqn. 2.5**) for the association equilibrium of M + L $\rightleftharpoons$ ML.

$$K_a = \frac{[ML]}{[M] + [L]} = \frac{1}{K_d} \tag{2.5}$$

2.2.8.1 DETERMINATION OF BINDING AFFINITY FOR A HIGH-AFFINITY LIGAND USING COMPETITION-BASED ITC

The thermodynamic and binding affinity parameters for the binding of 3,4-dichloroPBA and 4-chloro-3,5-dichloroPBA to MR were obtained by the following ITC experiments. MR (40 μM) in solution with a saturating concentration of low-affinity ligand PBA (500 μM, designated L_2) was titrated with high-affinity ligand (500 μM, designated L_1), which competes for the same active site. Binding affinity and thermodynamic data for PBA were previously determined by ITC ($K_d = K_2 = 1.7 \times 10^{-6}$ M, $\Delta H = -0.961$ kcal/mol) (Sharma, *et al.*, 2020). The apparent dissociation constant (K_{app}) was obtained from the titration of 3,4-dichloroPBA into the sample cell containing MR and PBA. Binding affinity of MR for PBA (K_2) is related to the K_{app} according to **eqn. 2.6**, allowing for calculation of the binding affinity of MR for the high-affinity ligands (K_1) using **eqn. 2.7**. For competition-based ITC experiments, the high-affinity ligand was titrated into the sample cell containing only the low-affinity ligand in buffer to correct for the heat of dilution and mixing. The corresponding isotherms were subtracted from competition ITC data prior to curve fitting.

$$K_{app} = \frac{K_1}{1 + (K_2)(L_{2\ sat})} \tag{2.6}$$

$$K_2 = \left(\frac{K_1}{K_{app}} - 1\right)\left(\frac{1}{L_{2\,sat}}\right) \qquad (2.7)$$

2.2.9 DETERMINATION OF HEAT CAPACITY CHANGE FOR WILD-TYPE MR UPON BINDING OF PBA

Isothermal titrations were conducted as above (Section 2.2.8) at a series of temperatures (10, 15, 20, 25, and 30 ºC) to determine the heat capacity change (ΔC_p) upon binding of MR to ligand. The enthalpies of binding (ΔH) were plotted as a function of temperature, and the ΔC_p was obtained from the slope of the resulting linear plot.

2.2.10 CONCENTRATION CORRECTION FOR 3,4-DICHLOROPBA, 2,4-DICHLOROPBA, AND 4-CHLORO-2,6-DIFLUOROPBA

^{1}H NMR spectroscopy was used to correct the concentration of ligand solutions for presence of boric acid contaminant in the commercial samples. Solutions containing a known mass of ligand and internal standard ethylene carbonate (88.06 g/mol, 99% purity) were prepared in 1 mL of dimethyl sulfoxide-d$_6$ (DMSO). ^{1}H NMR spectra were obtained and integrated signal intensities corresponding to the aromatic protons of the substituted PBAs were compared to the signal intensity of the four identical alkyl protons of ethylene carbonate located at 4.20 ppm. The accurate concentrations of the substituted PBAs in solution were then calculated from the ratio of signal intensities and the known concentration of ethylene carbonate using *MestReNova* software v. 14.1.0-24037, Mestrelab Research S. L. (Escondido, CA).

2.2.11 ^{11}B NMR SPECTROSCOPY

All NMR spectra were recorded at 25 °C in MR assay buffer containing 10% D_2O

in 5 mm quartz NMR tubes (final volume 750 µL) (Sigma-Aldrich Canada Ltd.,

Oakville, ON). Quartz tubes were used to minimize background signal arising from the

boron in borosilicate glass. Chemical shifts corresponding to the ^{11}B nuclei are reported

relative to a $BF_3 \cdot OEt_2$ external standard ($\delta = 0.00$ ppm). Spectra were obtained at 250 µM

in the case of both the free 3,4-dichloroPBA and the equimolar MR and 3,4-dichloroPBA

samples. Free ligand spectra were obtained for all compounds in the studies series to

assess contamination by boric acid. Boric acid contamination was only observed for 3,4-

dichloroPBA, 2,4-dichloroPBA, and 4-chloro-2,6-difluoroPBA. Samples were prepared

in DMSO-d$_6$. The background signal arising from the borosilicate glass of the

spectrometer probe was smoothed by Whittaker smoothing using *MestReNova* software

v. 14.1.0-24037, Mestrelab Research S. L. (Escondido, CA).

2.3 RESULTS

2.3.1 INHIBITION STUDIES

The structures of the series of chloro- and fluoro-substituted PBA ligands

characterized in this work are shown in **Scheme 2.2.** Preliminary kinetic analysis was

completed by obtaining IC$_{50}$ values for each inhibitor **Table 2.1**. The IC$_{50}$ values

indicated that all inhibitors in the series would be suitable for ITC-based analysis with the

exception of 4-chloro-2,6-difluoroPBA that had an IC$_{50}$ value of 0.856 ± 0.025 mM. In

general, ITC studies are most amenable to determination of binding constants on the

order of 10^{-8} M to 10^{-4} M but may be extended to 10^{-12} M using competition-based

methods (Menèndez, 2020). Hence, determination of the binding affinity for this compound is beyond the range of detection for ITC. Classical inhibition studies yielded a K_i value of 0.497 ± 0.086 mM (**Figure 2.4**), which was near the affinity of MR for its substrate (R)-mandelate ($K_m \approx 1$ mM) and far from the potency seen for the remaining inhibitors in the series (St. Maurice & Bearne, 2002). Further, a classical inhibition study was conducted for the inhibition of MR by 3,4-dichloroPBA, which was the most potent inhibitor in the series as confirmed by both the determined IC_{50} and K_d values. The determined K_i value for 3,4-dichloroPBA was 22.91 ± 3.45 nM (**Figure 2.5**).
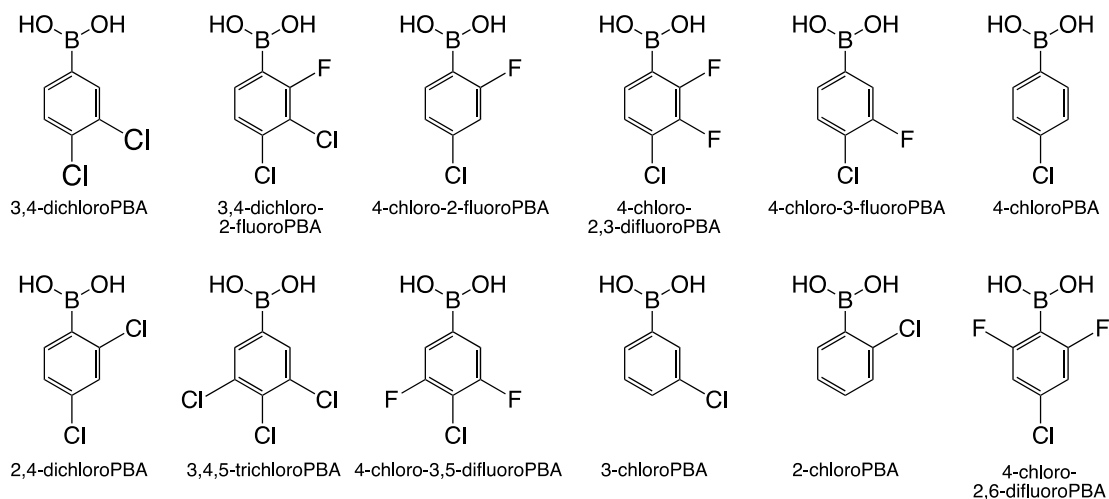
Figure 2.3. Structures of compounds in the series of chloro- and fluoro-substituted PBAs characterized for inhibition of MR.

Table 2.1. IC_{50} values for inhibition of MR by a series of chloro- and fluoro-substituted PBAs. Inhibitors are listed in order of decreasing binding potency (K_d) to MR.

Inhibitor	IC_{50} (nM)
3,4-dichloroPBA	20.5 ± 1.1
3,4-dichloro-2-fluoroPBA	93.1 ± 1.5
4-chloro-2-fluoroPBA	435 ± 12
4-chloro-2,3-difluoroPBA	341 ± 17
4-chloro-3-fluoroPBA	57.5 ± 2.7
4-chloroPBA	149 ± 6[a]
2,4-dichloroPBA	436 ± 19
3,4,5-trichloroPBA	552 ± 11
3-chloroPBA	272 ± 25
2-chloroPBA	9540 ± 848
4-chloro-3,5-difluoroPBA	54.0 ± 2.4
4-chloro-2,6-difluoroPBA	$8.26\ (\pm\ 0.25) \times 10^5$

[a]Value is from Sharma, *et al.* (2020)
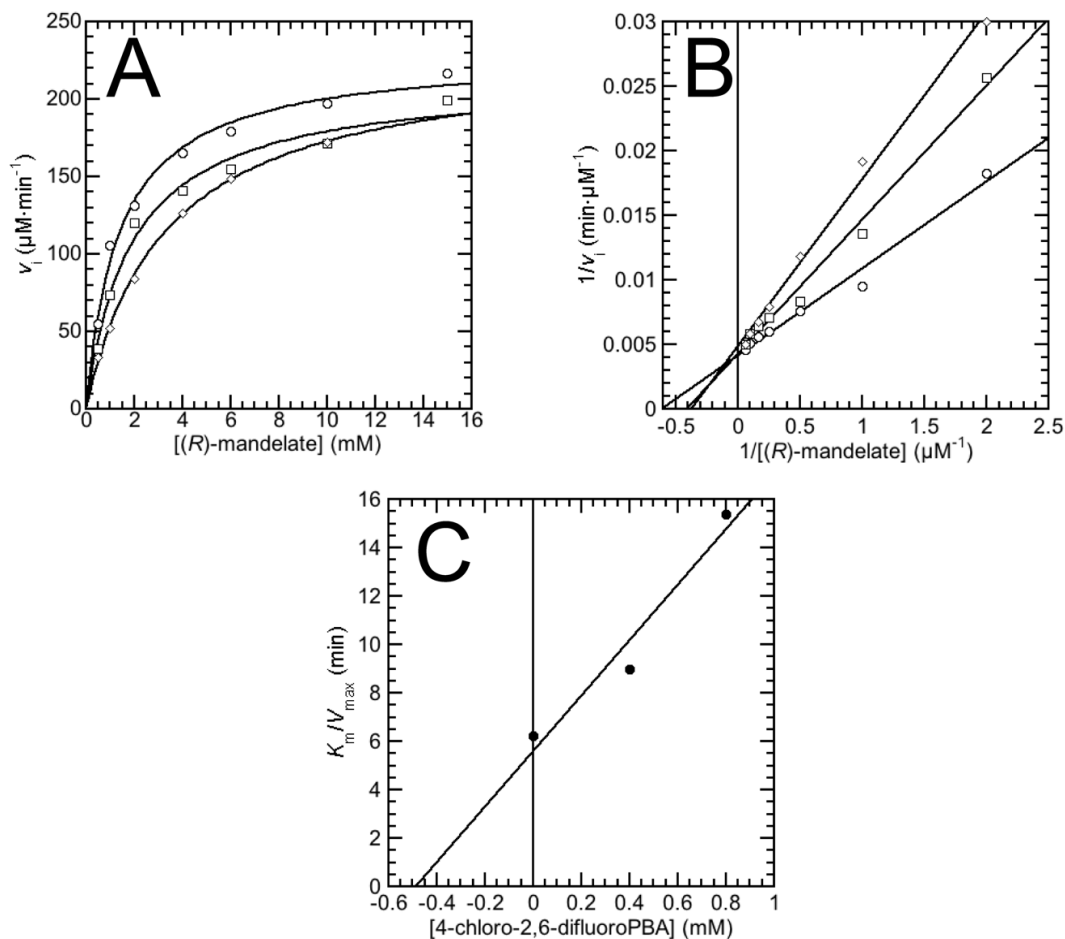
Figure 2.4. Competitive inhibition of MR by 4-chloro-2,6-difluoroPBA. A
representative Michaelis-Menten plot (A) and a representative Lineweaver-Burk plot (B)
showing competitive inhibition of MR by 3,4-dichloroPBA are depicted. The
concentrations of 4-chloro-2,6-difluoroPBA used were 0 (○), 0.4 (□), and 0.8 (◇) mM.
The concentrations of (*R*)-mandelate used were 0.5, 1.0, 2.0, 4.0, 6.0, 10.0, and 15.0 mM
and the concentration of MR was 150 ng/mL. A representative replot (C) of the apparent
K_m/V_m (obtained from fitting the Michaelis-Menten equation to initial velocity data) vs.
inhibitor concentration is shown. The K_i value (determined in triplicate) for inhibition of
MR by 4-chloro-2,6-difluoro PBA is 0.497 ± 0.086 mM.
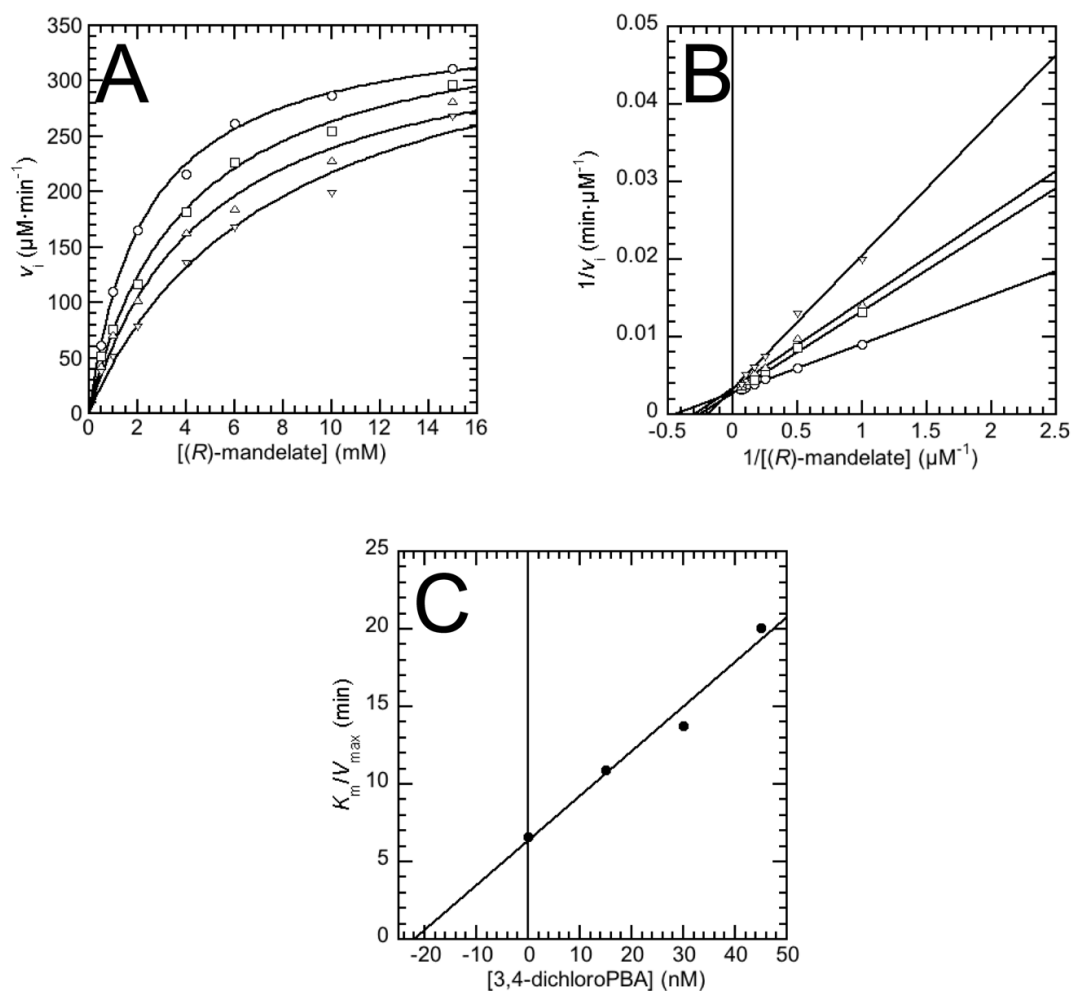
Figure 2.5. Competitive inhibition of MR by 3,4-dichloroPBA. A representative Michaelis-Menten plot (A) and a representative Lineweaver-Burk plot (B) showing competitive inhibition of MR by 3,4-dichloroPBA are depicted. The concentrations of 3,4-dichloroPBA used were 0 ($\circ$), 15 ($\square$), 30 ($\triangle$), and 45 ($\triangledown$) nM. The concentrations of (R)-mandelate used were 0.5, 1.0, 2.0, 4.0, 6.0, 10.0, and 15.0 mM and the concentration of MR was 150 ng/mL. A representative replot (C) of the apparent K_m/V_m (obtained from fitting the Michaelis-Menten equation to initial velocity data) vs. inhibitor concentration is shown. The K_i value (determined in triplicate) for inhibition of MR by 3,4-dichloroPBA is 22.91 ± 3.45 nM.

2.3.2 ISOTHERMAL TITRATION STUDY OF MR BINDING A SERIES OF CHLORO- AND FLUORO-SUBSTITUTED PBAs

Isothermal calorimetric titrations are shown in **Figure 2.6** for the series of substituted PBAs binding to MR. The figure depicts the binding isotherms (top panel), where the heat signature of the binding event for a given ligand was observed. Each peak in the binding isotherm resulted from a single injection of the ligand solution such that positive peaks corresponded to an endothermic event, while negative peaks corresponded to an exothermic binding event. The binding isotherm as plotted represents the power applied to the sample cell to return the system to baseline from the temperature change arising from each ligand injection. From this, a plot of the enthalpy of binding against the evolving molar ratio of [ligand]/[MR] was generated (bottom panel), from which binding (K_a and K_d), stoichiometric (n), and thermodynamic (ΔG, ΔH, and ΔS) information could be obtained as in **eqn. 2.8**, where M is the enzyme, and L is the ligand.

$$\Delta G° = \Delta H° - T\Delta S° \ = \ -RT \ln \left(\frac{[ML]}{[M][L]} \right) = \ -RT \ln K_a = RT \ln K_d \qquad (2.8)$$

ITC requires that the affinity of the enzyme-ligand system of interest falls within a range where (1) the added ligand must reach saturation by the end of the titration and (2) the enthalpy change associated with the interaction is detectable by the instrument (Menèndez, 2020). For complexes with a 1:1 stoichiometry, the shape of the titration curve is dependent upon the dimensionless parameter c, which is the product of the association constant (K_a) and the number of binding sites in the sample cell (eqn. 2.9), where high values of c correspond to a greater sigmoidicity in the titration curve.
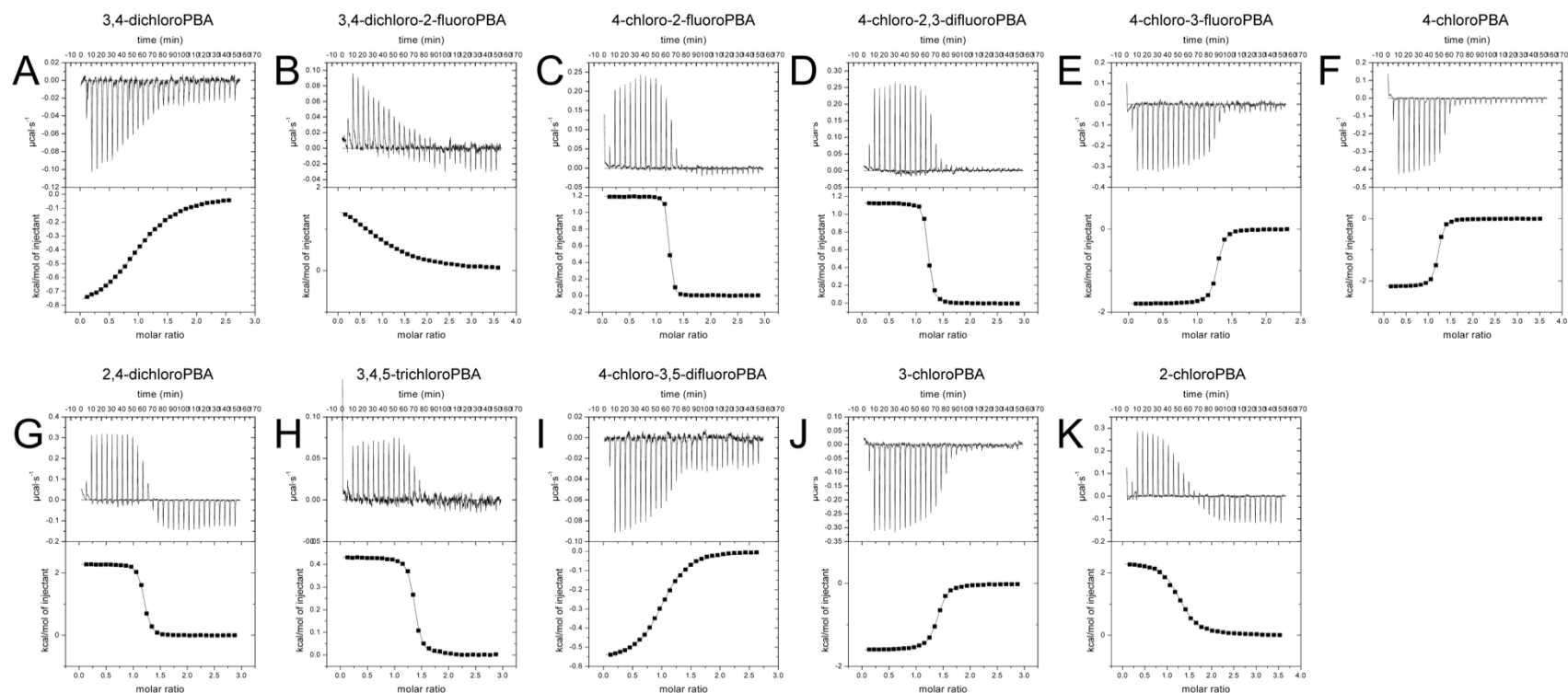
Figure 2.6. Representative ITC binding isotherms for MR binding a series of chloro- and fluoro-substituted PBAs. Isothermal calorimetric titration of 3,4-dichloroPBA (A), 3,4-dichloro-2-fluoroPBA (B), 4-chloro-2-fluoroPBA (C), 4-chloro-2,3-difluoroPBA (D), 4-chloro-3-fluoroPBA (E), 4-chloroPBA (F), 2,4-dichloroPBA (G), 3,4,5-trichloroPBA (H), 4-chloro-3,5-difluoroPBA (I), 3-chloroPBA (J), and 2-chloroPBA (K) into MR. All titrations were conducted using a concentration of 40 µM MR in the sample cell and 500 µM ligand in the injector. Titrations for MR with 3,4-dichloroPBA and 4-chloro-3,5-difluoroPBA were done using the competitive ITC experiment detailed in section 2.2.8. The titrations involved 30 injections of 10 µL at 240 s intervals and were conducted at 25 °C. The first injection was 5 µL and was omitted in the analysis.

Practical values of c fall in the range of 1-1000, typically requiring an enzyme-ligand system where affinity (K_d) is in the low micromolar to low nanomolar range (Wiseman, *et al.*, 1989; Ward & Holdgate, 2001). The previously determined affinities for PBA and 4-chloroPBA indicated that a structure-activity study for the series of chloro- and fluoro-substituted PBAs would fall into the ideal range of ITC prerequisite conditions (Sharma, *et al.*, 2020). This is especially important given the tendency for MR to precipitate at the higher enzyme concentrations (>90 µM) needed for a lower-affinity binding interaction and constant stirring in the ITC cell (Nagar, *et al.*, 2015).

$$c = K_a \cdot [enzyme]_{total} \cdot n \tag{2.9}$$

Calorimetric data are given in **Table 2.2**, providing both affinity and thermodynamic data for the series of MR ligands studied in this work. The binding constants for MR binding the inhibitors fell in the range of ~1 mM and 1 µM with the exception of 4-chloro-2,6-difluoroPBA, for which the binding constant (K_i) was 0.497 ± 0.086 mM. Binding isotherms were either exothermic or endothermic in the ΔH range of ±1-2 kcal/mol, except in the case of 3,4,5-trichloroPBA and 4-chloro-3,5-dichloroPBA, which had weaker enthalpies of 0.469 ± 0.053 kcal/mol and –0.577 ± 0.213 kcal/mol, respectively. Based on X-ray crystallographic studies, DFT calculations, and ligand binding studies for PBA and 4-chloroPBA, MR has a 1:1 stoichiometry with PBA and its chloro-/fluoro-substituted derivatives. To reflect this, the stoichiometry parameter n was fixed to a value of 1.0 in performing analysis on the calorimetric data.

Table 2.2. Isothermal titration calorimetry data for MR binding a series of chloro- and fluoro-substituted PBAs at 20 °C.

Compound	IC$_{50}$, nM[a]	K_d, nM	ΔH, cal·mol^{-1}	ΔS, cal·mol^{-1}·K^{-1}	ΔG, kcal·mol^{-1}
3,4-dichloroPBA	20.5 ± 1.1	13.8 ± 1.4	-1744 ± 32	30.0 ± 0.1	-10.5 ± 0.1
3,4-dichloro-2-fluoroPBA	93.1 ± 1.5	14.9 ± 6.4	1907 ± 427	42.4 ± 0.9	-10.5 ± 0.3
4-chloro-2-fluoroPBA	435 ± 12	23.4 ± 3.1	1072 ± 110	38.6 ± 0.6	-10.2 ± 0.1
4-chloro-2,3-difluoroPBA	341 ± 17	61.6 ± 2.2	1229 ± 146	32.7 ± 0.1	-6.5 ± 0.1
4-chloro-3-fluoroPBA	57.5 ± 2.7	65.0 ± 2.2	-2084 ± 175	25.8 ± 0.6	-9.6 ± 0.01
4-chloroPBA	149 ± 6	71.6 ± 27.3	-2238 ± 84	24.4 ± 1.6	-9.4 ± 0.4
2,4-dichloroPBA	436 ± 19	74.8 ± 17.8	2017 ± 227	40.0 ± 1.1	-9.6 ± 0.1
3,4,5-trichloroPBA	552 ± 11	96.4 ± 9.0	469 ± 53	33.7 ± 0.4	-6.3 ± 0.1
3-chloroPBA	272 ± 25	222 ± 10	-1734 ± 279	24.5 ± 0.9	-8.9 ± 0.02
2-chloroPBA	9540 ± 848	965 ± 165	2134 ± 380	34.8 ± 1.0	-8.1 ± 0.1
4-chloro-3,5-difluoroPBA	54.0 ± 2.4	148 ± 39	-576.7 ± 213	29.3 ± 0.4	-9.2 ± 0.2
4-chloro-2,6-difluoroPBA	8.26 (± 0.25) × 10^5	6.76 (± 1.1) × 10^5	–	–	–

[a]Values obtained by kinetic determinations.

2.3.3 DETERMINATION OF THE CHANGE IN THE SPECIFIC HEAT CAPACITY ACCOMPANYING MR BINDING PBA

The change in the specific heat capacity (ΔC_p) of binding is an informative parameter for an enzyme-ligand system, as it can suggest the types of ligand-binding interactions at the binding site (Holdgate & Ward, 2005). The specific heat capacity is defined as the change in the enthalpy of binding as a function of temperature, as given in equation 2.10. The change in specific heat capacity of binding was studied for PBA by conducting ITC experiments over a range of temperatures from 10 °C to 30 °C and the results are shown in **Figure 2.7**. The determined ΔC_p value for the binding of PBA to MR was -395 ± 33 cal·mol^{-1}K^{-1}.

$$\Delta C_\text{p} = \frac{\delta \Delta H}{\delta T} \tag{2.10}$$

2.3.4 ^{11}B NMR SPECTROSCOPY STUDIES FOR MR BINDING TO A SERIES OF CHLORO- AND FLUORO-SUBSTITUTED PBAs

Bearne and co-workers have shown the potential for formation of a dative bond between the His 297–N$^{\varepsilon 2}$ and the boron center of PBA, which was supported by crystallographic evidence, DFT calculations, and ^{11}B NMR spectroscopy studies (Sharma, *et al.*, 2020). The hybridization state of boron in bound PBA exhibits a chemical shift change from the free ligand form (28.2 ppm) to the bound form (0.97 ppm), which was characteristic for formation of an anionic tetrahedral boron center (sp^3 hybridization) (Sharma, *et al.*, 2020). Consequently, I obtained the ^{11}B NMR spectrum for the most potent PBA derived inhibitor of the series, 3,4-dichloroPBA. Using
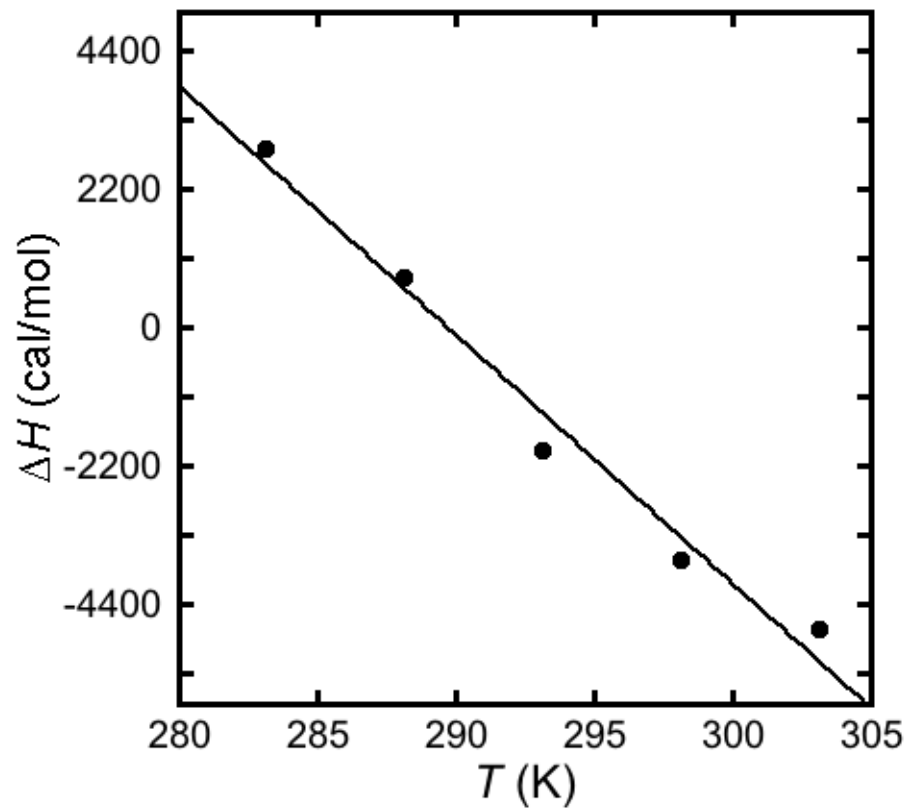
Figure 2.7. Change in the specific heat capacity accompanying MR binding PBA.
The binding enthalpy for PBA was measured in isothermal titrations at 10, 15, 20, 25, and 30 °C. The slope of the line gives the ΔC_p value, which was determined to be -395 ± 33 cal·mol^{-1}K^{-1}. The R^2 value for the linear line is 0.979.

an analogous approach, spectra of the free ligand (250 µM 3,4-dichloroPBA), and bound

ligand (250 µM equimolar MR and 3,4-dichloroPBA) were obtained (**Figure 2.8**). In the

presence of PBAs, MR precipitates readily above a concentration of 90 µM as previously

described (Sharma, *et al.*, 2020). ^{11}B NMR spectra have conventionally been obtained

using MR concentrations of ~300-400 µM in the presence of PBA derivatives (Sharma, *et

al.*, 2020; Douglas & Bearne, unpublished). In the present case, MR precipitated in the

presence of equimolar 3,4-dichloroPBA at concentrations above 250 µM. At these

relatively low concentrations, the signal-to-noise ratio becomes problematic, but I was

able to obtain reasonable spectra. The spectral data for 3,4-dichloroPBA (**Figure 2.8**)

were similar to that of PBA, which showed a change in chemical shift from 16.42 ppm

for free ligand to 0.90 ppm in the bound species, suggesting formation of a dative bond in

a similar manner to PBA (Sharma, *et al.*, 2020).

There is a notable signal at 19.12 ppm in the free ligand ^{11}B NMR spectrum of

3,4-dichloroPBA. This peak likely arises from boric acid contamination (chemical shift δ

$\approx$ 20 ppm) present in the commercial compounds, which are typically 97% pure (Maki,

Ishihara, & Yamamoto, 2006; Kitamura, *et al.*, 2011). Although the boric acid peak

appears very large relative to the 3,4-dichloroPBA peak in the ^{11}B NMR, I have shown

that the purity of the 3,4-dichloroPBA compound is 95.7% using an ethylene carbonate

internal standard. Similar determinations were completed for 4-chloro-2,6-difluoroPBA

and 2,4-dichloroPBA, for which the purity was found to be 93.4% and 97.0%,

respectively. None of the other compounds studied in this chapter were found to have

such contamination. Signals in ^{11}B NMR are broad as a result of relatively rapid

relaxation times of quadrupolar nuclei, which can confound relative concentration determinations as in this case (Bendel, 2004).

The free ligand spectra showed the main signal for 3,4-dichloroPBA at 16.05 ppm, which is strongly upfield shifted from the pH 3 control (entirely in the trigonal planar form) (**Figure 2.9**). This phenomenon arises from the acid-base equilibrium of 3,4-dichloroPBA at the pH for which the ^{11}B NMR spectra was taken. It was observed that the resulting upfield chemical shift in the enzyme-containing sample (0.69 ppm) was greater than the value for 3,4-dichloroPBA at pH 9 (6.0 ppm, entirely in tetrahedral conformation, **Figure 2.9**) and that of the imidazole·PBA complex (5.7 ppm) (Chen, *et al.*, 2016).
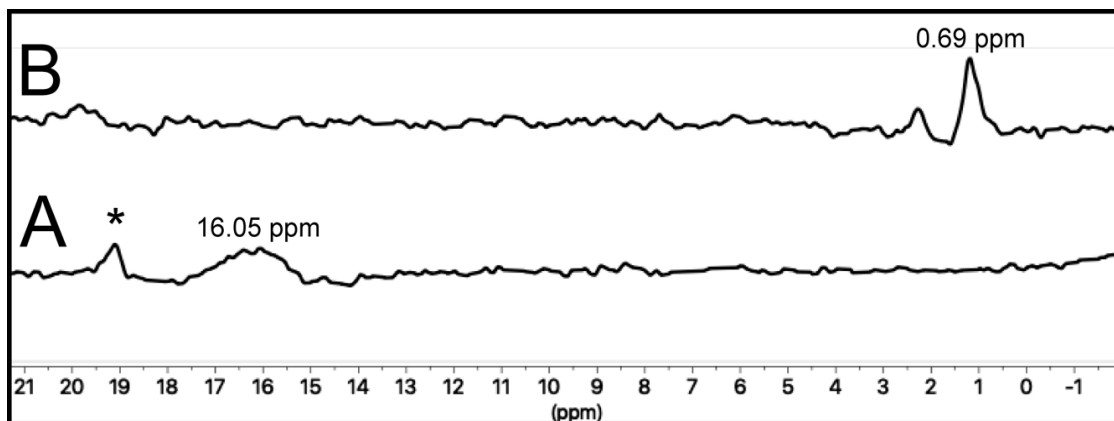
Figure 2.8. ^{11}B NMR spectra of free 3,4-dichloroPBA and 3,4-dichloroPBA bound to wild-type MR. Shown are the observed ^{11}B NMR spectra of (A) 250 μM 3,4-dichloroPBA and (B) equimolar 250 μM 3,4-dichloroPBA and MR (32000 scans per spectrum) taken at pH 7.5. The disappearance of the major peak at 16.05 and appearance of a new peak at 0.69 ppm in the enzyme-containing sample is indicative of a change to a sp^3-hybridized boron and consistent with formation of a dative bond between MR and 3,4-dichloroPBA. The signal at 19.12 ppm (*) in spectrum A likely corresponds to boric acid contamination. Boric acid does not inhibit MR (Sharma, *et al.*, 2020).
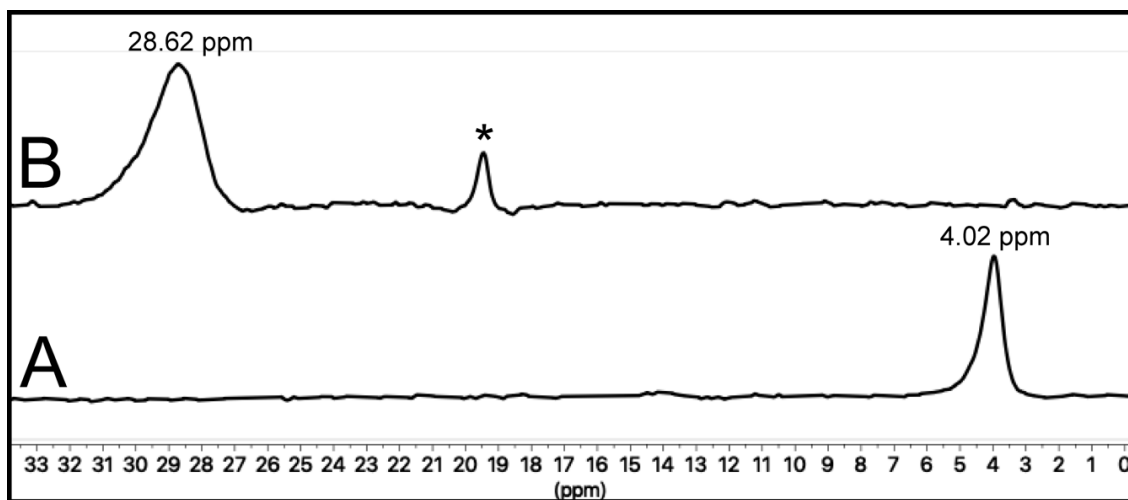
Figure 2.9. ^{11}B NMR spectra of free 3,4-dichloroPBA at pH 3.0 and 9.0. Shown are ^{11}B NMR spectra for 3,4-dichloroPBA at (A) pH 9.0, where the boronic acid is entirely in the tetrahedral anionic sp^3 conformation. Spectra were obtained at 2.5 mM 3,4-dichloroPBA. (B) pH 3.0, where the boronic acid is entirely in the trigonal planar sp^2 conformation, and

2.4 DISCUSSION

2.4.1 THE STRUCTURE-ACTIVITY RELATIONSHIP FOR MR BINDING TO A SERIES OF CHLORO- AND FLUORO- SUBSTITUTED PBAs

Building on the work of Bearne and colleagues identifying PBA and *p*-halogen-substituted PBAs as potent reversible inhibitors of MR, I have characterized a series of chloro- and fluoro-substituted PBAs (Sharma, *et al.*, 2020). The vast majority of the inhibitors in the series bind to MR with affinities greater than that of PBA, which agrees with the initially observed trend, suggesting that small electron withdrawing substituents increase the binding affinity of the PBA-based ligands (Sharma, *et al.*, 2020). Bearne and co-workers showed that the increased affinity of MR for 4-chloroPBA relative to PBA arose from the electron withdrawing effects of the Cl- substituent rather than fortuitous interactions with the enzyme (Sharma, *et al.*, 2020). This left open the possibility that additional substituents with similar properties to the 4-chloro-substituent may further increase the binding affinity.

As opposed to the classical kinetic inhibition studies used for MR, ITC-based analysis offered a superior technique for determination of the binding affinities for this series of inhibitors due to its robustness in the range of affinities determined from the IC$_{50}$ analyses. In addition to directly measuring binding affinities, ITC provides thermodynamic data for the enzyme-inhibitor system (Menèndez, 2020).

Interestingly, 3,4-dichloro-substituted PBAs are not only the most potent inhibitors of MR in this series but also the most potent inhibitors of MR described to date. This builds upon the trend initially shown by Bearne and co-workers in determining that 4-chloroPBA increases affinity more than other *para*-substituted PBAs because of its electron withdrawing character (Sharma, *et al.*, 2020). The observed trend of *m*/*p*-chloro-

substituents improving affinity greater than identical *m/p*-fluoro-substituents in the series is caused at least in part by the corresponding electron withdrawing effects. Both the Hammett and hydrophobicity substituent constants are greater for *meta*- and *para*-Cl-substituents than F-substituents (**Table 2.3**), which appears to be a desirable property towards MR affinity (Hansch, Leo, & Taft, 1991; Fujita, Hawasa, & Hansch, 1964).

Table 2.3. Hammett and hydrophobicity constants for fluoro- and chloro-substituents.

Substituent	σ_p[a]	σ_m[a]	π[b]
F	0.34	0.06	0.14
Cl	0.37	0.23	0.71

[a]Values are from Hansch, Leo, Taft. (1991)
[b]Values are from Fujita, Hawasa, Hansch. (1964)

In order to make additional conclusions regarding the structure-activity relationship for this series of inhibitors the thermodynamics of binding must be considered (*vide infra*). As per equation 2.8, energetically favoured binding of ligand to the protein is characterized by a negative free energy of binding (ΔG). For each of the chloro- and fluoro-substituted PBAs studied, the obtained ΔG values were negative on the order of ~ –8.1 to –10 kcal/mol (excluding the 3,4,5-trisubstituted PBAs, which are discussed below). These binding free energies are more favourable than the values of –7.7 kcal/mol observed for PBA and –7.49 kcal/mol observed for MR binding BzH, as well as –4.36 and –4.80 kcal/mol for (*S*)- and (*R*)-atrolactate, respectively (**Table 2.2**) (Sharma, *et al.*, 2020; Nagar & Bearne, 2015). This marked increase in binding affinity for the chloro- and fluoro-substituted PBAs indicates significantly changed interactions

between MR and the substituted PBAs. For structure-activity conclusions to be made

relating the increased binding affinity associated with the substituted PBAs, only

comparisons with the binding affinity of unsubstituted PBA will be considered since

these will reflect only those structural differences arising from the chloro- and fluoro-

substituents.

The ΔC_p value for PBA binding to MR was -395 ± 33 cal·mol^{-1}K^{-1}, which is in

excellent agreement for the ΔC_p value of -358 ± 3 cal·mol^{-1}K^{-1} accompanying BzH

binding to MR (Nagar & Bearne, 2015). When considered with the major entropic

contribution to binding, these values of ΔC_p are characteristic of hydrophobic interactions

(Lin, Schwarz, & Eisenstein, 1995; Guan, *et al.*, 2011; Parker, *et al.*, 1999).

Complimentarily, through X-ray crystallography it has been shown that the phenyl ring

for both BzH and PBA binds in the hydrophobic pocket of MR, which is lined with

hydrophobic residues from the N-terminal domain, the barrel domain, and Leu 93 from

the interdigitating loop of the adjacent monomer (see **Figure 2.1**) (Sharma, *et al.*, 2020;

Nagar & Bearne 2015; Bearne 2017).

The aforementioned variance in the ΔH values for the series of Cl- and F-substituted

PBAs (~ -2 to $+2$ kcal/mol) may arise from variations in inhibitor structure. All of the

ortho-substituent-containing PBA inhibitors exhibited endothermic binding with MR.

Conversely, all other inhibitors in the series (solely *meta*- and *para*-substituted) exhibited

exothermic binding with MR, with the exception of 3,4,5-trichloroPBA, which fits an

isolated trend discussed below. Upon closer examination, this trend is also reflected in

the entropy of binding for these compounds. The binding of *ortho*-substituent-containing

PBAs in this series is marked by ΔS values ranging from 32.7 to 42.4 cal/mol (37.7

cal/mol·K average) at 20 °C, as opposed to the ΔS for the solely *meta-* and *para-*substituted inhibitor values ranging from 24.4 to 33.7 cal/mol·K (29.9 cal/mol average). A comparison of the two trends can be seen in **Figure 2.10**, where the entropic contribution to binding is shown as $T\Delta S$. While the observed ΔS values are quite similar across the series, there is a slight reduction in ΔS for the solely *meta-* and *para-*substituted series. This reduction is offset by a negative ΔH, resulting in a minimal difference in the K_d and ΔG values across the series. This phenomenon is common to structure-activity relationship studies within biological systems, where enthalpic gain from structural changes to a ligand increases the strength or number of binding-site interactions corresponds to an entropic penalty, minimizing $\Delta\Delta G$ between inhibitors in the same series (Ward & Holdgate, 2001).

The observed less favourable ΔH accompanying binding of the *ortho-*substituted PBAs may arise from rotation of the B–C bond, shifting the hydroxyl groups of the boronic acid out of plane with the phenyl ring. This conformation would differ from the structure of bound PBA, in which the trigonal planar form has the hydroxyl groups of the boronic acid in plane with the phenyl ring. In the tetrahedral form, the boron is coordinated to His 297 in a dative bond, while the hydroxyl groups of PBA are angled towards Lys 166 and Lys 164 to which hydrogen bonds are formed (Sharma, *et al.*, 2020). Large enthalpic changes (>2 kcal/mol) have been linked to changes in the mode of binding, which is commonly seen through a decrease in the number or strength of the hydrogen bonds from protein to ligand (Velasquez-Campoy, Kiso, & Freire, 2001; Ward, Holdgate & 2001; Holdgate, *et al.*, 1997). Induced rotation of the $B(OH)_2$ group in the MR-bound PBA species has the potential to alter or remove the hydrogen bonds from the

Table 2.4. Comparison of isothermal titration calorimetry data for various inhibitors of MR at 20 °C

Type of inhibitor	Ligand/Inhibitor	K_d (μM)	ΔH (kcal/mol)	ΔS (cal/mol·K)	ΔG (kcal/mol)
boronic acid	3,4-dichloroPBA	0.0138 ± 0.0014	−1.74 ± 0.03	30.0 ± 0.1	−10.5 ±0.1
boronic acid	3,4-dichloro-2-fluoroPBA	0.015 ± 0.0064	1.91 ± 0.43	42.4 ± 0.9	−10.5 ± 0.3
boronic acid	PBA[a]	1.7 ± 0.3	−0.96 ± 0.07	23.1 ± 0.5	−7.7 ± 0.1
hydroxamate	BzH[b] (TS analogue)	2.90 ± 0.40	4.06 ± 0.07	39.22 ± 0.92	−7.42 ± 0.08
glycolate	(S)-atrolactate[b] (S/P analogue)	557 ± 25	7.85 ± 0.18	40.97 ± 0.55	−4.36± 0.02
glycolate	(R)-atrolactate[b] (S/P analogue)	259 ± 24	−1.29 ± 0.04	11.78 ± 0.27	−4.80 ± 0.06

[a]Values are from Sharma, *et al.*, (2020).
[b]Values are from Nagar & Bearne (2015).

hydroxyl groups to active site residues or alter coordination to the Mg^{2+} ion, which could yield the large $\Delta\Delta H$ observed between the *o*-substituted and solely *m/p*-substituted PBAs by ITC. Of course, care must be taken in drawing conclusions strictly from thermodynamic data without X-ray crystal structure data on both the complex and the free partners (Holdgate & Ward, 2005; Ward & Holdgate, 2001). For this reason, future work on the structure-activity relationship of Cl- and F-substituted PBAs binding MR should include structural determinations for both *o*-substituent containing PBAs and solely *m/p*-substituted PBAs.

Excluded from these two trend-groups are 3,4,5-trichloroPBA and 4-chloro-3,5-difluoroPBA, the only 3,5-disubstituted compounds in the series. The 3,4,5-trichloroPBA and 4-chloro-3,5-difluoroPBA inhibitors both fall into the lower potency range of this series of inhibitors. The enthalpy for the binding 3,4,5-trichloroPBA and 4-chloro-2,6-difluoroPBA to MR were found to be 0.469 ± 0.053 kcal/mol and -0.576 ± 0.213 kcal/mol, respectively, while the corresponding entropies of binding were 33.7 cal/mol·K and 29.3 cal/mol·K. While both inhibitors fall into the lower ΔS range observed for the solely *m/p*-substituted PBAs, the ΔH values are entirely intermediary within the set, not fitting into the strongly exothermic or endothermic groups. Altogether, the mode of binding in the case of *m,m,p*-trisubstituted PBAs may be altered by steric constraints arising from the additional *m*-substitution. It is unlikely that the hydrogen bonding network or interactions between the Brønsted acid-base catalysts Lys 166 and His 297 and the boron center are altered because the substitution occurs at a position that is remote from these sites. Moreover, based on the observation that the inhibitory potency increases from 4-chloroPBA to the 3,4-dichloro-substituted PBAs, one would
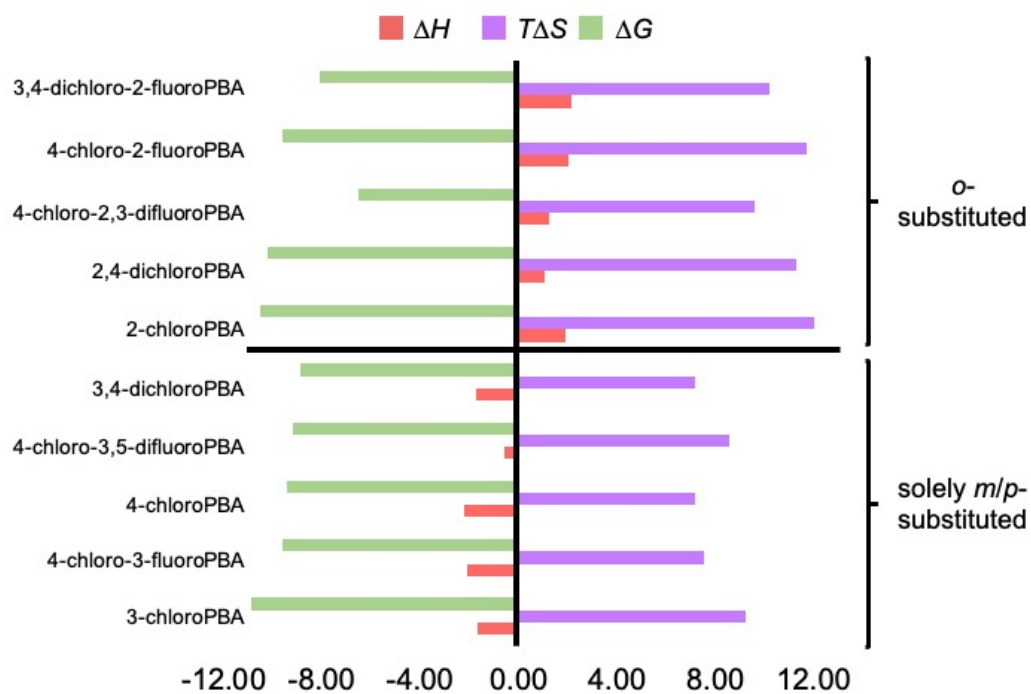
Figure 2.10. Comparison of the thermodynamic parameters for a series of chloro-and fluoro-substituted PBAs binding to MR. Shown are the thermodynamic parameters of binding (ΔH, $T\Delta S$, and ΔG) for the series of inhibitors studied binding to MR obtained from ITC studies. Red bars represent ΔH, purple bars represent ΔS, and green bars represent ΔG for ligand binding to MR. The plot is split into groups: *o*-substituted PBAs and solely *m*/*p*-substituted PBAs based on the sign of the ΔH of binding.

hypothesize that 3,4,5-trichloroPBA would exhibit even greater binding affinity. Such is clearly not the case and the additional *m*-substituent appears to cause detrimental steric interactions.

Alternatively, it should be noted that inhibitors within the series that did not have the 4-chloro-substituent were in the lower echelon of binding affinities. This is consistent with the need for a small, *p*-substituted EWG for potent inhibition of MR by PBA derivatives (Sharma, *et al.*, 2020).

Finally, it was found that 4-chloro-2,6-difluoroPBA had much reduced potency as an inhibitor of MR ($K_i = 0.676 \pm 0.11$ mM), indicating that the active site of MR does not prefer two *o*-substitutions, even when the substituent is a small fluoro group. Unfortunately, the extremely low affinity of MR for this inhibitor prohibited ITC-based studies. Hence, I was not able to determine the thermodynamic parameters accompanying binding. It would be expected that such a dramatic change in affinity would exhibit a different thermodynamic signature. The affinity data were obtained using classical kinetic studies, which showed that 4-chloro-2,6-difluoroPBA was a reversible competitive inhibitor of MR.

2.4.2 DETERMINATION OF THE MODE OF 3,4-DICHLOROPBA BINDING TO MR

In order to observe the presence of a dative bond between His 297–N$^{\varepsilon 2}$ and the boron center, ^{11}B NMR was employed in a similar manner as was utilized for PBA binding to MR (Sharma, *et al.*, 2020). ^{11}B NMR spectroscopy has been used to characterize the hybridization state of bound boronic acid inhibitors in a number of studies involving enzymes including the aforementioned serine hydrolases (Baldwin, *et*

al., 1991a; London & Gabel, 2001; Adebodun & Jordan, 1988). A change from a trigonal planar conformation (B(OH)$_2$) to a bound tetrahedral adduct results in a large upfield chemical shift in the ^{11}B NMR spectrum due to the increased electron density at the boron center.

The observed spectral changes accompanying the binding of 3,4-dichloroPBA to MR were in excellent agreement with the ^{11}B NMR data obtained by Bearne and colleagues (Sharma, *et al.*, 2020), indicating similar presence of an N–B dative bond between His 297 and the boron center for both PBA and 3,4-dichloroPBA. The enzyme-containing sample yielded an ^{11}B NMR signal far sharper than that for the free ligand spectra, which was consistent with both studies on PBA and MR and for aryl boronic acids bound to chymotrypsin (Sharma, *et al.*, 2020; Baldwin, *et al.*, 1991b). The pK_a of 3,4-dichloroPBA must be considered in this case where the pK_a value is predicted to be similar to the buffer (pH 7.5). Based on the observation that the ^{11}B NMR chemical shifts for the trigonal (pH 3.0) and tetrahedral (pH 9.0) species are 28.62 and 6.0 ppm, respectively (**Figure 2.9**), the adduct may have additional charge density at the boron center.

2.5 CONCLUSIONS AND FUTURE WORK

A series of Cl- and F-substituted phenylboronic acids were assessed as potential high affinity ligands for MR. ITC analysis permitted direct determination of direct dissociation constants for MR-ligand binding. Within the series of inhibitors, 3,4-dichloro-substituted PBAs were found to be the most potent inhibitors of MR seen to date, with the highest affinity corresponding to 3,4-dichloroPBA ($K_d = 13.8 \pm 1.4$ nM). A

notable change in thermodynamic parameters accompanying binding of the *o*-substituted

PBAs within the series likely relates to alteration of the hydrogen bonding network

between the enzyme and the substituted PBA. The binding process was entropy-driven,

probably due to liberation of ordered active site water molecules and binding of the

phenyl ring of the PBAs at the hydrophobic pocket of MR. These entropically dominated

aspects of PBA binding were enhanced by Cl- and F-*m*/*p*-substitutions and hindered by *o*-

substituents. Finally, it was suggested through ^{11}B NMR spectroscopic studies that for

the most potent inhibitor in the series, 3,4-dichloroPBA, a dative bond was formed

between the boron center and the His 297–N$^{\varepsilon 2}$.

As a result of the trends observed within the series of substituted PBAs, structural

studies should be undertaken for representative inhibitors for both the *o*-substituent

containing PBA and the *m*- and *p*-substituted PBA (3,4-dichloroPBA and 3,4-dichloro-2-

fluoroPBA). This will allow for elucidation of the complex trends observed in the

thermodynamic data obtained. Additionally, the enhanced affinity for 3,4-dichloroPBA

relative to the fluoro-substituted analogues indicates a tolerance for additional bulk at

these positions. Future characterization of bromo-substituted PBAs may allow for full

capitalization of the binding mode and further enhance the potency of inhibitors. Full

understanding of the molecular interactions and specific enzymatic binding determinants

for the substituted PBAs will give a broader understanding of boronic acid inhibitor

design for both MR and enzymes catalyzing heterolytic cleavage of C–H bonds.

CHAPTER 3 THERMAL STABILITY OF ACTIVE SITE VARIANT PROTEINS WITHIN THE MANDELATE RACEMASE SUBGROUP OF THE ENOLASE SUPERFAMILY

3.1 INTRODUCTION

As discussed in Chapter 1, catalysis by enzymes in the MR subgroup relies on Brønsted acid-base chemistry effected by a conserved histidine residue at the end of the seventh β-strand and a commonly conserved lysine at the C-terminal end of the second β-strand (Kallakaral, *et al.*, 1995; Landro, *et al.*, 1994, Powers, *et al.*, 1991). In addition, a lysine exists two-residues away from the lysine where Brønsted acid-base chemistry occurs, which constitutes a conserved "KxK" motif in the MR subgroup, where "x" is a non-conserved amino acid residue. This first lysine of the motif is part of a group of polar residues that "solvate" ligands in the active site, which contributes to enthalpic stabilization of intermediates, reducing the $\Delta G^{\ddagger}$ of the reaction (Neidhart, *et al.*, 1991). In MR, Glu 217 and Lys 164 interact with the carboxyl group of mandelate, Asp 197 interacts with the α-hydroxyl group of mandelate. The conserved Mg^{2+} interacts with the α-hydroxyl and carboxylate of the substrate and also stabilizes the additional negative charge built up in the intermediate and TS of the reaction (**Figure 3.1**) (Mitra, *et al.*, 1995; Neidhart *et al.*, 1991; St. Maurice & Bearne, 2000; Fee *et al.*, 1974; Guthrie & Kluger, 1993). Additionally, the proximity of Lys 164, Lys 166, and the active site Mg^{2+} comprises a region of the active site with a high local positive charge. This region is conserved within the MR subgroup, which raises two questions: (1) Is this charge state being maintained for a catalytic role, and (2) is it energetically destabilizing to the enzymes?
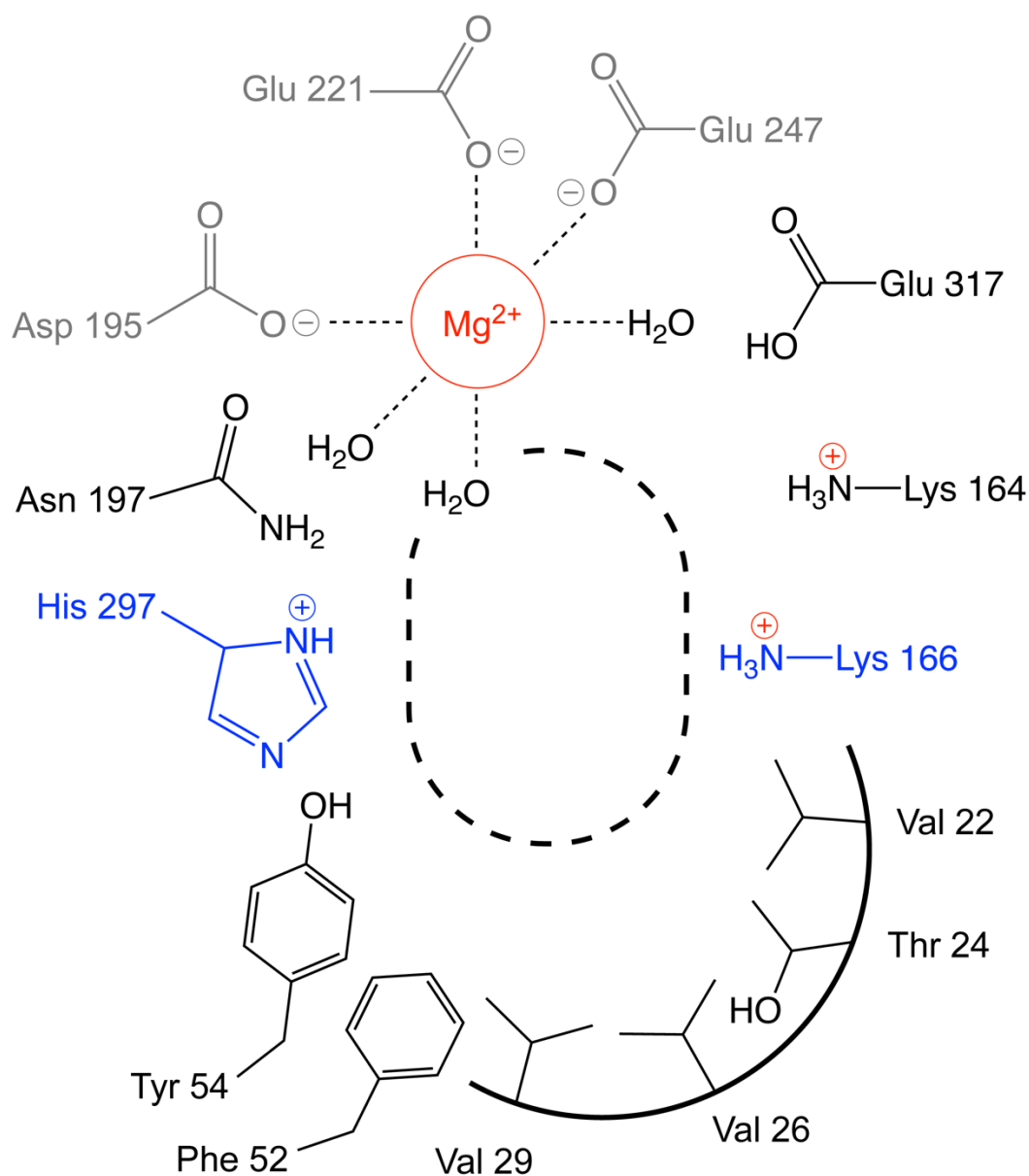
Figure 3.1. Two-dimensional reconstruction of the free MR active site.
Catalytic residues discussed within this body of work are shown, as well as the relevant active site Mg^{2+}. The localized highly positively charged region is highlighted in red, comprising Lys 166, Lys 164, and Mg^{2+}. The ligand binding position is represented by a ellipse. In the absence of substrate, three water molecules coordinate the Mg^{2+} ion in addition to the three liganded residues (figure adapted from Bearne & St Maurice, 2017).

Four representative members of the MR subgroup, *Pp*MR, *Xc*FucD, *Bj*TarD, and *St*TGD, were chosen to broaden the scope of study. These four enzymes represent the MR subgroup well, showcasing both the incredible level of structural similarity (**Figure 3.2** and **Figure 1.2**) and the mechanistic diversity (**Scheme 3.1**) present within the subgroup. Additionally, these enzymes have all been both functionally and structurally characterized using activity assays and X-ray crystallography, respectively, making analysis considerably more feasible (Landro, *et al.*, 1993; Yew, *et al.*, 2006a; Yew, *et al.*, 2006b; Yew, *et al.*, 2007). While the prototypical MR catalyzes a 1,1-proton transfer reaction that racemizes (*R*)- and (*S*)-mandelate, the majority of enzymes in the MR subgroup are actually classified as acid sugar dehydratases, catalyzing a β-elimination reaction following the conserved enolization partial reaction (Bearne & St. Maurice, 2017). FucD catalyzes the dehydration of L-fuconate followed by a stereospecific ketonization of the enol product (Yew, *et al.*, 2006a), TGD catalyzes the competing dehydration and epimerization of L-talarate and galactarate (Yew, Ws, *et al.*, 2007), and TarD catalyzes the dehydration of D-tartrate, which is followed by non-enzymatic ketonization of the enol product (Yew, *et al.*, 2006b). Because TarD catalyzes the formation of the enol of oxaloacetate to be ketonized in solution, it has been suggested that it furnishes a link between the 1,1-proton transfer reaction catalyzed by MR and the acid sugar dehydration reactions catalyzed by the other sugar acid dehydratases of the MR subgroup (Yew, *et al.*, 2006b).
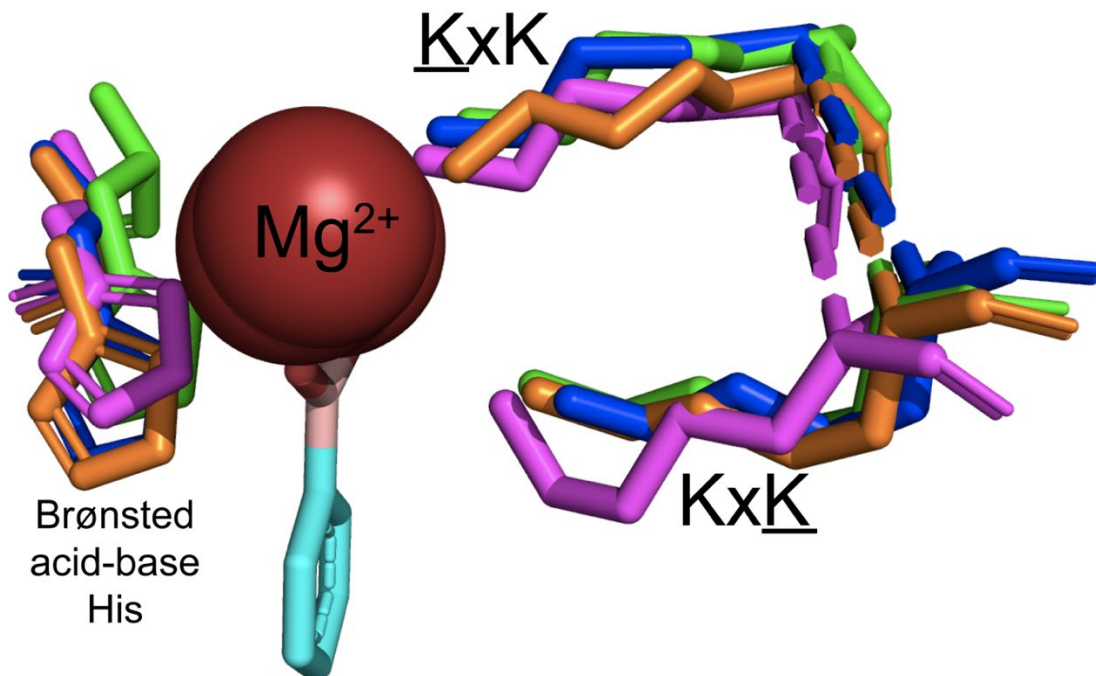
Figure 3.2. Structure overlay of *Pp*MR, *Xc*FucD, *Bj*TarD, and *St*TGD. Select active site residues studied within this chapter are shown for MR (magenta, PDB file 6VIM), FucD (green, PDB file 2HXT), TGD (orange, PDB file 2PP0), and TarD (blue, PDB file 2DW7). The residues shown are the Brønsted acid-base histidine and lysine (KxK), as well as the adjacent KxK lysine. The Mg^{2+} ions (red) are shown in space filling orientation. PBA bound at the MR active site is shown as a reference ligand.

Scheme 3.1. Catalytic mechanism of *Xc*FucD, *St*TGD, and *Bj*TarD. Adapted from Bearne (2017).
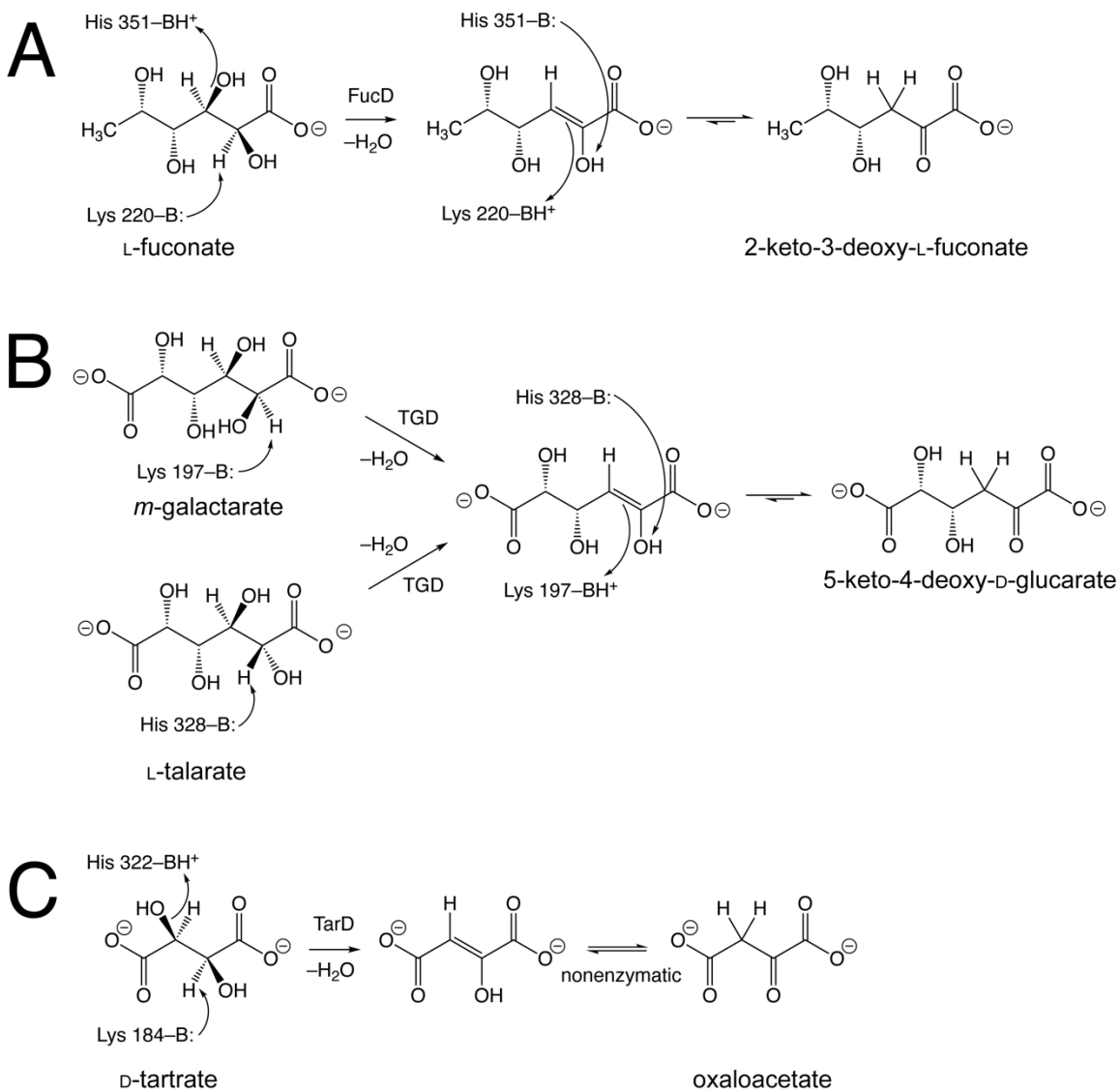
3.1.1 USING DIFFERENTIAL SCANNING CALORIMETRY TO DETERMINE THE RELATIVE STABILITY OF PROTEINS

The stability of the representative MR subgroup enzymes and variants was characterized using differential scanning calorimetry (DSC), a technique that profiles the molar heat capacity of samples as a function of temperature (Durowoju, *et al.*, 2017). DSC thermograms provide a complete thermodynamic profile of the thermal denaturation process, indicating the energy required (enthalpy, ΔH) to disrupt interactions at the tertiary and even sometimes quaternary structural levels. The thermal transition temperature (melting temperature, T_m) is the temperature where a maximum of the $C_p(T)$ function occurs and provides a direct measure of protein stability, which can be compared to protein variants (Sturtevant, 1987). Previous unpublished work by H. Kumar from the Bearne lab revealed that wild-type MR had a T_m of 63.6 °C (**Figure 3.3**).

In DSC, the excess C_p of the sample is reported, as samples are analyzed in reference to the buffer obtained after dialysis. The representative thermograms shown in **Figure 3.3** show a clear T_m peak in the $C_p(T)$ function for each of the wild-type MR subgroup proteins being studied. The MR subgroup proteins were shown to follow a two-state model with irreversible thermal denaturation under kinetic control (native (N) $\rightleftharpoons$ unfolded (U) $\rightarrow$ denatured (D)) (Kumar & Bearne, unpublished). This was primarily and most obviously observed by visible precipitation of the protein in the sample after cooling. The thermogram appears to show a steep decrease in C_p after the thermal denaturation event has occurred, which is a phenomenon observed in certain cases of proteins that follow a two-state denaturation. This phenomenon arises from aggregation and subsequent precipitation of denatured protein. The heavily negative ΔC_p DSC signal after the T_m is a sum of the endothermic
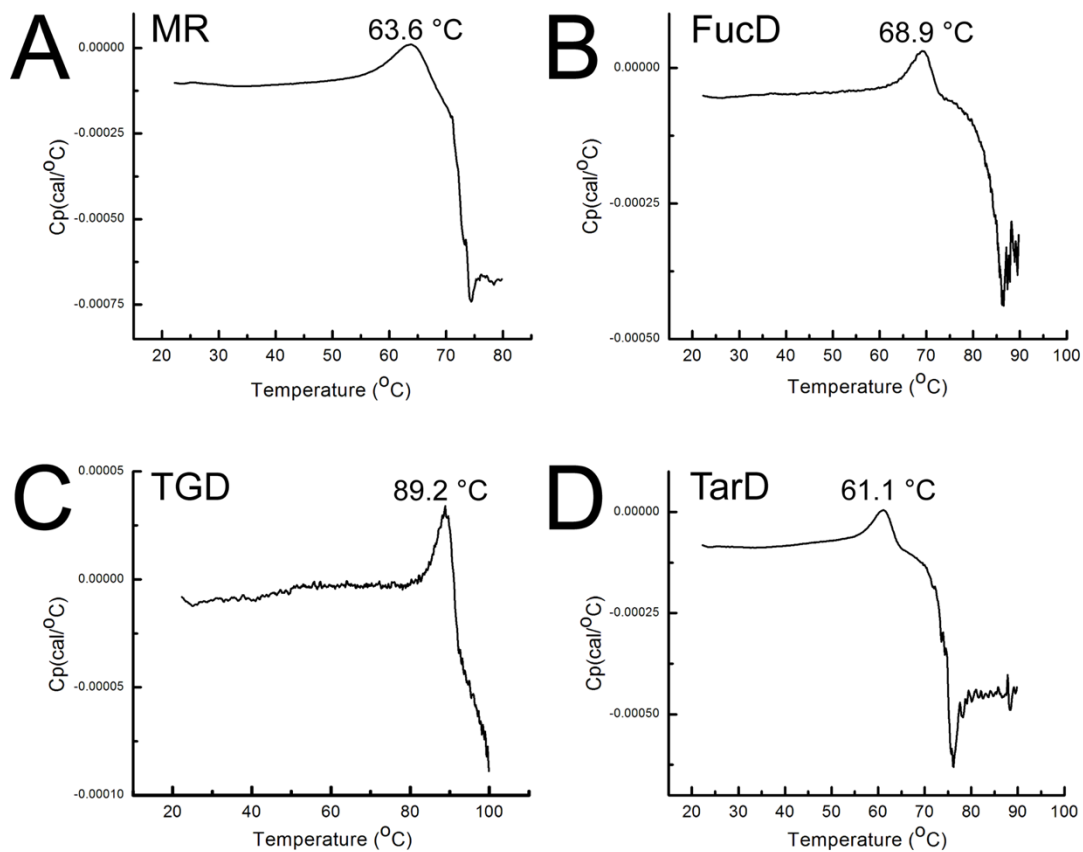
Figure 3.3. Representative DSC thermograms for wild-type *Pp*MR, *Xc*FucD, *Bj*TarD, and *St*TGD. Thermal melting data are obtained from the observed maximum in the shown $C_p(T)$ functions obtained for wild type MR, FucD, TarD, and TGD. The peaks correspond to T_m values of (A) 63.63 °C for MR, (B) 68.93 °C for FucD, (C) 89.15 °C for TGD, and (D) 61.06 °C for TarD (Kumar, Bearne, unpublished).

protein unfolding event and the relatively high exothermicity of the aggregation event driven by the hydrophobic effect (Privalov & Gill, 1988).

Thermodynamic parameters for the unfolding transition (N $\rightleftharpoons$ U) and kinetic parameters for the irreversible step (U $\rightarrow$ D) determined for MR show that the rate of inactivation is sufficiently fast such that the native and unfolded states were no longer in equilibrium (Kumar & Bearne, unpublished). Thus, the native state irreversibly proceeds to the denatured state at the thermal transition temperature (N $\rightarrow$ D). This was demonstrated by obtaining DSC thermograms at variable scan rates ranging from 15 °C/h and 90 °C/h, which showed a dependence of the T_m on the scan rate. The irreversible thermal transition of MR was confirmed by obtaining a reheating thermogram that showed no significant reproduced events (Kumar & Bearne, unpublished). As a result of the inherent irreversibility of the MR system, accurate thermodynamic parameters (e.g. $\Delta H°$) for the MR subgroup proteins cannot be determined using DSC

A common checkpoint for protein determinations by DSC is completed by comparison of the indirectly calculated van't Hoff heat change (ΔH_{vH}) and the calorimetric heat change (ΔH_{cal}), which can be obtained from the integration of the $C_p(T)$ function according to **equation 3.1**, where T_0 and T_1 are the beginning and ending temperatures of the thermal transition (Johnson, 2013). The ratio of the ΔH_{vH} to ΔH_{cal} is unity for a two-state protein system, while a ratio over one is indicative of unfolding of a self-association state (dimer, trimer, etc.) and a value less than one is indicative of protein unfolding via one or more intermediate states (Johnson, 2013). DSC data obtained for MR show a $\Delta H_{vH}/\Delta H_{cal}$ ratio of ~2 for each of the studied scan rates as is given in previously observed data obtained by H. Kumar (unpublished). This is unsurprising given

that in solution, MR exists as tightly interacting dimers arranged in either a tetrameric or octameric state (Fee, Hegeman, & Kenyon, 1974; Tsou, *et al.*, 1989; Sagy & Bearne, unpublished).

$$\Delta H_{cal} = \int_{T_0}^{T_1} C_p dT \qquad (3.1)$$

3.1.2 EFFECT OF SUBSTITUTING THE BRØNSTED BASE CATALYSTS ON THE THERMOSTABILITY OF MR SUBGROUP VARIANTS

Previous unpublished work done by the Bearne Lab group was conducted to investigate the effect that manipulation of the active site environment would have on the thermal stability of MR and the subgroup enzymes (Kumar & Bearne, unpublished). For MR, the K166M variant was generated to observe the effect that the positively charged Brønsted acid-base lysine has on the stability of MR (**Figure 3.4**). Most interestingly, Lys 166 destabilizes the wild-type enzyme, which is likely a result of the positively charged electrostatic character of the protonated residue in its free form with no ligand bound. Further, to assess the effect that the active site divalent Mg^{2+} cation on the thermal stability of MR, the apo-enzyme was generated by overnight dialysis against EDTA-containing assay buffer and was found to reduce the T_m by –9.48 °C (Kumar & Bearne, unpublished). Taken together, these observations suggest that the positive charge of the protonated Lys 166 gives rise to the destabilization of the protein. It is doubtful that the K166M substitution significantly alters the structure of the protein to a point where thermal stability is altered, given that the X-ray crystal structure of Lys 164 MR variant proteins were found to maintain wild-type structure (Kallarakal, *et al.*, 1995).

Substitution of the active site Brønsted acid-base residues (i.e. His 297 and Lys 166 in MR) throughout the representative MR subgroup members were generated to observe the effect of substitution on the thermal stability (Kumar & Bearne, unpublished). Wild-type FucD, TGD, and TarD have T_m values of 68.93, 89.15, and 61.06 °C, respectively. The results shown in **Figure 3.4** (Kumar & Bearne, unpublished) indicate that all variant proteins had greater thermal stability upon substitution of the active-site Brønsted acid-base lysine residue throughout the subgroup. The active-site Brønsted acid-base histidine did not show such a trend in thermal stability, except for the H297N MR variant, which was relatively destabilized. Taken together, these results suggest that the active-site Brønsted acid-base lysine is protonated in the free enzyme form and that the subsequent positive charge of Lys 166 and its MR subgroup equivalents destabilizes the enzymes.

The highly enhanced thermal stability of TGD arises from additional quaternary structure relative to MR, FucD, and TarD. X-ray crystal structures of TGD reveal that the first 19 residues of the N-terminus of the protein wrap around a monomer of an adjacent dimer, which is not seen in the other observed members of the MR subgroup (Yew, *et al.*, 2007).

3.1.3 PROTECTION AGAINST THERMAL DENATURATION BY LIGANDS

Typically, ligands can protect proteins against thermal denaturation such that T_m is increased (Holdgate & Ward, 2005). Ligand binding is associated with a release of free energy, which stabilizes the enzyme·ligand complex. This stabilization is observed by an increase in T_m for the complex. For MR, Kumar and Bearne observed that greater thermal
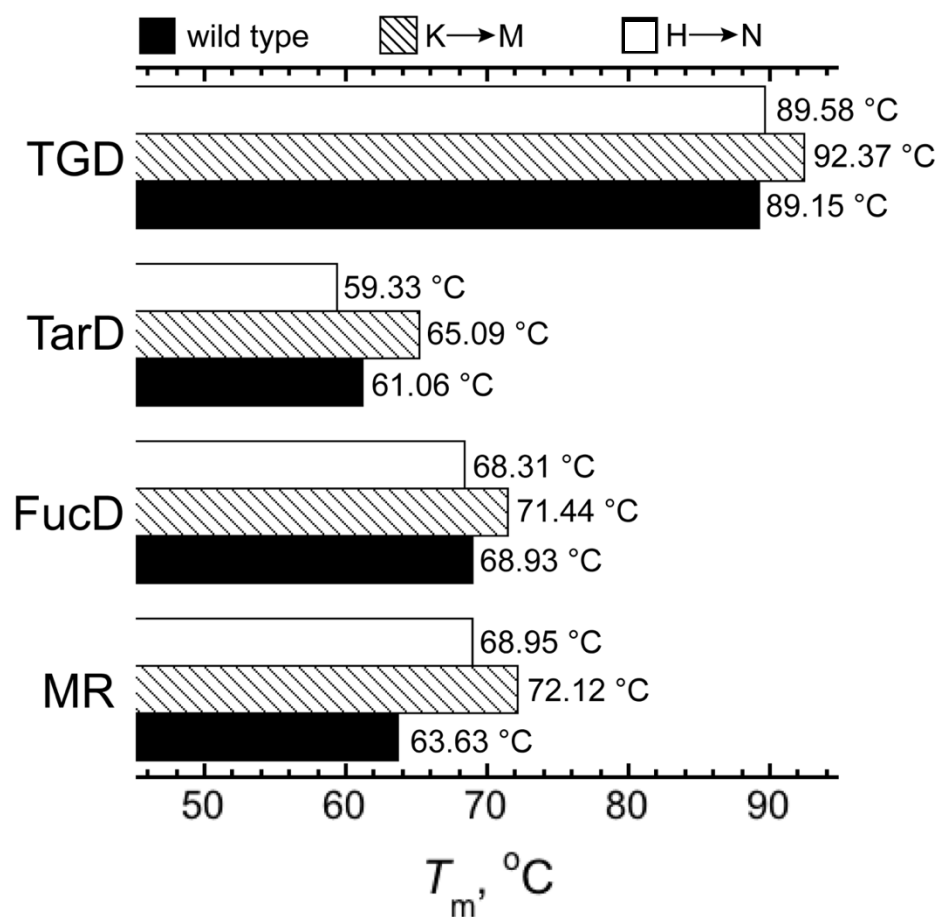
Figure 3.4. Comparison of the thermal stability of KxK and catalytic Brønsted acid-base His point mutation variants of MR subgroup members. Shown are the T_m values for comparison of the contribution to stability that the active-site Brønsted acid-base residues provide to the protein.

stabilization of MR was achieved when (*S*)-atrolactate ($K_d = 557 \pm 25$ μM, $T_m^{max} = 73.2 \pm$ 0.2 °C) was present in comparison to (*R*)-atrolactate ($K_d = 259 \pm 24$ μM, $T_m^{max} =$ 70.5 ± 0.3 °C) (Nagar, *et al.*, 2015). Much more pronounced protection against thermal denaturation was observed in the presence of the transition state analogue BzH ($K_d = 2.90 \pm 0.40$ μM, $T_m^{max} = 75.4 \pm 0.8$ °C). These findings suggest that more potent inhibitors of MR and/or inhibitors that bind similarly to the TS yield greater ΔT_m values.

3.1.4 FUNCTION OF THE CONSERVED <u>K</u>XK RESIDUE OF THE MR SUBGROUP

Originally, the energetic contributions of the active site Brønsted acid-base residues to enzyme stability were explored because they were expected to contribute to the highly positive character of the active site. The finding that the second lysine of the Kx<u>K</u> motif destabilizes the free enzyme form of MR subgroup enzymes evoked exploration of the role of the adjacent <u>K</u>xK residue in stability.

Lys 164 in MR plays an important role in catalysis. X-ray crystal structures of MR revealed that Lys 164 participates in a hydrogen bonding network with Lys 166 and Tyr 137, forming a triad of residues (Neidhart, *et al.*, 1991). This triad of residues has also been shown throughout the MR subgroup for FucD, TarD, TGD, and other members (Yew, *et al.*, 2006a; Yew, *et al.*, 2006b; Yew, *et al.*, 2007). Lys 164 and its equivalent residues across the subgroup are denoted a "first shell" residue due to its direct interaction with substrate in the form of a hydrogen bond with the negatively charged carboxylate oxygen of the bound substrate.

Using pH-rate profiles for the enzymatic reaction of MR, Bearne and colleagues determined that the hydrogen bonding interaction between Lys 164 and Tyr 137

decreased the pK_a of Lys 166 by ~1.4 units and of His 297 by ~1.1 units (Fetter, *et al.*, 2019). This reduction in pK_a contributes to the overall pK_a perturbation of Lys 166 from ~10.5 in folded proteins (Harris & Turner, 2002; Grimsley, Scholtz, & Pace, 2009) to the observed value of 6.4 required for catalysis in MR (as observed in the MR·(*S*)-mandelate complex) (Fetter, *et al.*, 2019). Thus, it is evident that Lys 164 and its equivalent residues in the MR subgroup do not only affect the reaction being catalyzed but impact the entire electrostatic character of the active site.

Overall, Lys 164 plays a major role in reduction of the pK_a of Lys 166 to a value that can affect catalysis of the enzymatic reaction (Fetter, *et al.*, 2019). The H-bond formed between Tyr 137 and Lys 164 can serve to stabilize the protonated ε–NH$_3^+$ group form of Lys 164, while simultaneously positioning the side chain of Lys 164 proximal to the adjacent ε–NH$_3^+$ group of Lys 166. The resulting destabilization of the protonated Lys 166 results in the observed downward perturbation of the pK_a (Fetter, *et al.*, 2019; Neidhart, *et al.*, 1991). Thus, we anticipated that, contrary to the destabilization caused by Lys 166 in the free enzyme, Lys 164 and its equivalent residues across the MR subgroup actually serve to stabilize the free enzyme.

To test this hypothesis, the KxK variant proteins of the MR subgroup members MR, FucD, TarD, and TGD were generated and characterized in order to explore the energetic contribution of the residue across the subgroup. DSC of the variants allowed for determination of the thermal stability and comparison with the corresponding wild-type enzymes. Changes in the thermal stability (ΔT_m) of the variants relative to the wild-type proteins were then used to determine what contribution that the conserved KXK residue provides to the proteins.

3.2 MATERIALS AND METHODS

3.2.1 GENERAL

All reagents were purchased for Sigma-Aldrich Canada Ltd. (Oakville, ON,

Canada) unless otherwise stated. Circular dichroism (CD) spectra were collected using a

JASCO J-810 spectropolarimeter (Jasco Inc., Easton, MI). UV-visible and OD_{600}

spectrophotometric readings were obtained using a HP 8453 UV-visible

spectrophotometer (Agilent Technologies, Inc., Santa Clara, CA).

3.2.2 EXPRESSION AND PURIFICATION OF *Pp*MR VARIANT PROTEINS

Recombinant variants of *Pp*MR (wild-type, K164R, and K164M) were

overexpressed in *E. coli* BL21 (DE3) cells transformed with a pET-52b(+) plasmid

(Novagen, Madison, WI) containing the open reading frames encoding the *Pp*MR

variants as described by Narmandakh and Bearne (2010). The resulting vector encodes

the MR gene product as a fusion protein with an N-terminal StrepII-tag. Small scale

cultures (5 mL) of *E. coli* BL21 (DE3) cells were grown overnight at 37 °C from glycerol

stocks in sterile LB media containing ampicillin (100 µg/mL). Large scale expression

cultures (2 L LB media, 100 µg/mL ampicillin) were inoculated with the overnight

cultures at a volume of 10 mL per litre of LB media. The culture producing wild-type

MR was incubated at 37 °C with shaking (225 rpm) for 8 h, while cultures producing the

K164M and K164R MR variants were incubated at 37 °C with shaking (225 rpm) for 2 h

and then incubated at room temperature (22 °C) with shaking (225 rpm) for 24 h. Cells

were harvested using centrifugation (3795 × *g*, 10 min, 4 °C), and washed with wash

buffer (100 mM Tris-Cl, 150 mM NaCl, 1 mM EDTA, pH 7.5). Cell pellets were stored at –20 °C for future use.

Frozen cell pellets were thawed and resuspended in ice-cold wash buffer (~35 mL). The cell suspension was kept on ice and sonicated for 6×30 s using a Branson Sonifier 250 with 1 min "resting" periods in between sonication (setting 5.5, 1-s bursts). The cell lysate was clarified by ultracentrifugation ($146550 \times g$, 35 min, 4 °C) and applied to a column containing StrepTactin Superflow affinity resin (10 mL) (IBA Life Sciences, Göttingen, Germany) connected to an ÄKTA fast protein liquid chromatography (FPLC) system (GE Healthcare, Baie-Durfé, QC). After washing the column with wash buffer ($15 \times$ column volume), protein was eluted by addition of wash buffer containing desthiobiotin (2.5 mM). Eluted enzyme was dialyzed over 3×8 h against MR assay buffer (100 mM HEPES, 3.3 $MgCl_2$, pH 7.5), with each 8-h dialysis using fresh assay buffer (500 mL). Protein was aliquoted (1 mL) and stored at –20 °C for future use. In all cases, the protein fusion tag was not removed. Greater than 97% purity of the wild type, K164M, and K164R MR variants using this protocol was previously demonstrated using SDS-PAGE (Douglas, *et al.*, unpublished).

3.2.3 EXPRESSION AND PURIFICATION OF *Xc*FucD VARIANT PROTEINS

Recombinant variants of *Xc*FucD (wild-type and K218M) were overexpressed in *E. coli* BL21 (DE3) cells transformed with a pET-52b(+) plasmid (Novagen, Madison, WI) containing the open reading frames encoding the three *Xc*FucD variants. The resulting vector encodes the FucD gene product as a fusion protein with an N-terminal StrepII-tag. Small scale cultures (5 mL) of *E. coli* BL21 (DE3) cells were grown

overnight at 37 °C from glycerol stocks in sterile LB media containing ampicillin (100 μg/mL). Large scale expression cultures (2 L LB media, 100 μg/mL ampicillin) were inoculated with the overnight cultures at a volume of 10 mL per litre of LB media. The cultures were incubated until an OD_{600} of 0.5-0.6 was obtained, at which point expression was induced by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.1 mM. The cultures were then incubated at 22 °C for 24 h. Cells were harvested using centrifugation (3795 × g, 10 min, 4 °C), and washed with wash buffer (100 mM Tris-Cl, 150 mM NaCl, 1 mM EDTA, pH 7.5). Cell pellets were stored at –20 °C for future use.

Protein purification conditions for all FucD variants mimic those previously described in section 3.2.2 for proteins bearing a Strep-II fusion tag with the following exceptions. Frozen cell pellets were resuspended in lysis buffer (100 mM Tris, 5 mM $MgCl_2$, pH 7.5). Eluted enzyme was dialyzed against FucD assay buffer (50 mM Tris, 10 mM $MgCl_2$, pH 7.5). The purity of obtained protein was assessed using SDS-PAGE.

3.2.4 EXPRESSION AND PURIFICATION OF *Bj*TarD VARIANT PROTEINS

Recombinant variants of *Bj*TarD (wild-type and K182M) were overexpressed in *E. coli* BL21 (DE3) cells transformed with a pET-15b(+) plasmid (Novagen, Madison, WI) containing the open reading frames encoding the three *Bj*TarD variants as fusion proteins bearing an N-terminal hexahistidine (His)$_6$-tag. Small scale cultures (5 mL) of *E. coli* BL21 (DE3) cells were grown overnight at 37 °C from glycerol stocks in sterile LB media containing ampicillin (100 μg/mL). Large scale expression cultures (2 L LB media, 100 μg/mL ampicillin) were inoculated with the overnight cultures at a volume of

10 mL per litre of LB media. The cultures were incubated at 37 °C for 2 h with shaking (225 rpm), and then incubated at 22 °C for 24 h with shaking (225 rpm). Cells were harvested using centrifugation (3795 × g, 10 min, 4 °C), and washed with wash buffer (100 mM Tris-Cl, 150 mM NaCl, 1 mM EDTA, pH 7.5). Cell pellets were stored at –20 °C for future use.

Frozen cell pellets were thawed and resuspended in cold wash buffer (~35 mL). The cell suspension was kept on ice and sonicated for 6 × 30 s using a Branson Sonifier 250 with 1 min "resting" periods in between sonication (setting 5.5, 1-s bursts). The cell lysate was clarified by ultracentrifugation (146550 × g, 35 min, 4 °C) and applied to a column containing Ni^{2+}-charged His·Bind affinity resin (2.5 mL) (Novagen, Madison, WI). After washing the column with binding buffer (25 mL) (20 mM Tris-HCl, 500 mM NaCl, and 5 mM imidazole, pH 7.9), wash buffer (15 mL) (20 mM Tris-HCl, 500 mM NaCl, and 60 mM imidazole, pH 7.9), protein was eluted by addition of strip buffer (7 mL) (20 mM Tris-HCl, 500 mM NaCl, and 100 mM EDTA, pH 7.9). Eluted enzyme was dialyzed over 3 × 8 h against assay buffer (50 mM Tris-HCl, 10 $MgCl_2$, pH 8.0), with each 8-h dialysis using fresh assay buffer (500 mL). Protein was aliquoted (1 mL) and stored at –20 °C for future use. In all cases the fusion tag was not removed. The purity of the protein was assessed using SDS-PAGE.

K182M TarD was highly unstable and resulted in precipitated protein aggregates upon purification under the conditions outlined above. Further, the protein samples obtained by His·Bind affinity chromatography were contaminated by other proteins of approximately 70 and 90 kDa (possibly Hsp70, Hsp90, ArnA, and/or SlyD). To overcome the problem of contamination by background proteins, an *E. coli* low

background strain (LOBSTR, Kerafast, Inc., Boston, MA) was transformed by the appropriate pET-15b(+)-TarD plasmid and used for protein production (Andersen, Leksa, & Schwartz, 2013). For K182M TarD, after the clarified lysate was applied to the column, the column was washed with 25 mL of binding buffer, followed by 10 mL of TarD assay buffer containing 500 mM NaCl and 150 mM imidazole to remove background expressed proteins. The protein was then eluted in 15 mL of TarD assay buffer containing 500 mM NaCl and 250 mM imidazole. No dialysis was performed, and any experimental analyses were carried out in the 250 mM imidazole elution buffer.

In addition, wild-type TarD was also produced, purified, and experimentally analyzed using the identical purification method described above as a control for the K182M TarD. The purity of the protein was assessed using SDS-PAGE.

3.2.5 EXPRESSION AND PURIFICATION OF *St*TGD VARIANT PROTEINS

Recombinant variants of *St*TGD (wild-type and K195M) were overexpressed in *E. coli* BL21 (DE3) cells transformed with a pET-15b(+) plasmid (Novagen, Madison, WI) containing the open reading frames encoding each of the two *St*TGD variants as a fusion protein bearing an N-terminal hexahistidine (His)$_6$-tag. Small scale cultures (5 mL) of *E. coli* BL21 (DE3) cells were grown overnight at 37 °C from glycerol stocks in sterile LB media containing ampicillin (100 μg/mL). Large scale expression cultures (2 L LB media, 100 μg/mL ampicillin) were inoculated with the overnight cultures at a volume of 10 mL per litre of LB media. The cultures were incubated at 37 °C for 2 h with shaking (225 rpm), and then incubated at 22 °C for 24 h with shaking (225 rpm). Cells were harvested using centrifugation (3795 × *g*, 10 min, 4 °C), and washed with wash buffer

(100 mM Tris-Cl, 150 mM NaCl, 1 mM EDTA, pH 7.5). Cell pellets were stored at –20 °C for future use.

Protein purification conditions for all TGD variants mimic those previously described in section 3.2.4 for proteins bearing an N-terminal (His)$_6$ fusion tag. The purity of the obtained protein was assessed using SDS-PAGE.

3.2.6 SITE-DIRECTED MUTAGENESIS

Plasmids encoding the recombinant wild-type MR subgroup proteins (pET-52b(+) for MR and FucD and pET-15b(+) for TarD and TGD) were extracted from *E. coli* DH5α cells (New England Biolabs, Ipswich, MA) using a Qiagen miniprep kit (Qiagen, Toronto, ON) following the manufacturer's instructions. The plasmid DNA was used as the template DNA for site-directed mutagenesis (Hemsley, *et al.*, 1989). The primers used for mutagenesis (Integrated DNA Technologies Coralville, IA) inserted an altered codon at the mutation site and are listed in **Table 3.1**. Phusion high fidelity polymerase was used for all amplification reactions according to manufacturer's instructions (New England Biolabs, Ipswich, MA). The thermocycle "PCR" conditions used were 98 °C for a 5-min initial denaturation, followed by 25 cycles of 30 s at 98 °C, 30 s at the annealing temperature, and 3 min at 72 °C for elongation. After the temperatures were cycled 25 times, a final extension was conducted at 72 °C and was followed by storage at 12 °C. The reaction product was treated with DpnI restriction enzyme overnight at 25 °C to digest wild-type methylated template DNA according to manufacturer's instructions (New England Biolabs, Ipswich, MA). DNA products were then separated using agarose gel electrophoresis (0.8% agarose, 100 V, 400 mA, 60 min) and subsequently isolated

using a Qiagen Gel Extraction kit (Qiagen, Toronto, ON) according to manufacturer's

instructions. Isolated DNA products were used to transform chemically competent DH5α

cells using the heat shock method (Froger & Hall, 2007) and grown on agar plates (1.5%)

containing ampicillin (100 μg/mL) overnight. Transformants were cultured in 5 mL

volumes of sterile LB media with ampicillin (100 μg/mL) and DNA was extracted

(Miniprep, Qiagen, Toronto, ON). The open reading frames were then sequenced

commercially (Robarts Research Institute, London, ON) using Sanger sequencing to

verify that only selected mutations to the open reading frame were introduced. Plasmids

with the correct sequences were used to transform *E. coli* BL21 (DE3) cells using the

heat shock method (Froger & Hall, 2007). Transformants were cultured in LB media (5

mL) containing ampicillin (100 μg/mL) and stored in 15% glycerol stocks at –80 °C.

These cells were then used for protein expression.

3.2.7 CD SPECTRA OF MR SUBGROUP VARIANT PROTEINS

CD spectra of MR subgroup variant proteins were obtained using a protein

concentration of 100 μg/mL in 10 mM HEPES buffer at pH 7.5, containing 3.3 mM

MgSO$_4$. This buffer composition was chosen to minimize contributions of buffer

components to the background absorbance at low wavelengths. A 0.1-cm quartz cuvette

was used to collect spectra from 190-260 nm. Three buffer spectra were averaged and

subtracted from the three averaged protein spectra to obtain a final CD spectrum. CD

spectra for wild type, K164M, and K164R MR were previously obtained (Douglas &

Bearne, unpublished). Protein concentrations were determined using UV

spectrophotometry at 280 nm using extinction coefficients determined by the ProtParam

Table 3.1. Primers used for site-directed mutagenesis of the ORF encoding *Pp*MR, *Xc*FucD, *Bj*TarD, and *St*TGD. The DNA oligomer sequence of primers used to generate K164M MR, K164R MR, K218M FucD, K195M TGD, and K182M TarD is shown. All primers are shown in the 5′ to 3′ direction. The altered bases are shown in boldface and the codon encoding the point-mutation is underlined.

Variant	Primer	DNA Sequence
K164M MR[a]	Forward	5′-GGATTCCGGGCGGTT**AT**G**ACCAAGATCG-3′
	Reverse	5′-GCCGATCTTGGT**CA**TAACCGCCCGGAAT-3′
K164R MR[b]	Forward	5′-GGATTCCGGGCGGTT**AGG**ACCAAGATCGGC-3′
	Reverse	5′-GCCGATCTTGGT**CCT**AACCGCCCGGAATCC-3′
K218M FucD	Forward	5′-TTCCGCACCATC**AT**GCTCAAGGTCGGC-3′
	Reverse	5′-AAGCCATCGGCCACCGCTTCCTTG-3′
K195M TGD	Forward	5′-AATGGCGGTATT**AT**GTTGAAAGTCGGAC-3′
	Reverse	5′-ATGCCATTTTCGCGGGAAATCACC-3′
K182M TarD	Forward	5′-TACAACGTCGTGA**T**GATGAAGATCGGC-3′
	Reverse	5′-TAGCCGCGGTCGAGATAGCCACG-3′

[a]K164M MR was generated by Fetter, *et al.* (2019)
[b]K164R MR was generated by Douglas & Bearne (unpublished)

web tool from ExPASy under the assumption that all cysteine residues were reduced (Gasteiger, *et al.*, 2005). The obtained CD spectroscopic data (mdeg) was converted to MRE using eqn. 3.2.

$$MRE = \frac{MW(\frac{g}{mol}) \times \theta(mdeg) \times (0.001)(\frac{deg}{mdeg})}{b(cm) \times c(\frac{\mu g}{mL}) \times (10^{-6})(\frac{g}{\mu g}) \times 10(\frac{dmol}{mol}) \times (no.\,AA - 1)} \quad (3.2)$$

Due to the specialized conditions required for purification of the K182M TarD variant, it was not possible to obtain CD spectral data due to the high concentration of imidazole. Furthermore, the protein could not be dialyzed into the typical buffer used for CD spectroscopy since the protein precipitated under these conditions.

3.2.8 ACTIVITY DETERMINATION FOR MR SUBGROUP VARIANT PROTEINS

3.2.8.1 *Pp*MR

K164M and K164R MR were assessed for activity with (*R*)-mandelate as the substrate. Activity was assessed using the CD-based assay described in section 2.2.4. Activity was assayed by observing the change in ellipticity at 262 nm using a 0.5-cm pathlength. All assays were conducted in MR assay buffer containing 0.01% w/v BSA at 25 °C. Assays were conducted using a higher enzyme concentration (520 µg/mL, 12.1 µM) for the variant proteins than for wild type. The concentration of (*S*)- and (*R*)-mandelate used for assaying K164M ranged from 3-20 mM and from 0.5-15 mM for

K164R. If no enzymatic velocity was observed in measurements (i.e., no significant difference from the buffer control rate) then the variant enzyme was deemed to have too low an activity to be measured by the CD-based assay (data not shown).

3.2.8.2 *Xc*FucD

K218M FucD was assessed for activity with L-fuconate as the substrate. Activity was assessed using a CD-based assay analogous to the one used for MR. Initial velocity values were obtained from the negative linear slopes of the observed ellipticity at 216 nm over the time of the enzymatic reaction. Reactions were initiated by addition of FucD with 0.005% BSA to a solution of substrate in assay buffer. Reaction progress was observed by following the ellipticity at 216 nm for 300 s.

A constant substrate concentration of 4 mM L-fuconate was used (~8-times larger than the K_m), which gives an easily measurable initial enzymatic reaction velocity with wild-type FucD. The maximum assayed enzyme concentration was 100 μg/mL. A 0.2-cm pathlength quartz cuvette was used to minimize voltage due to absorbance at high enzyme concentrations. If no enzymatic velocity was observed in measurements up to the maximum enzyme concentration (i.e., no significant difference from the buffer control rate) then the variant enzyme was deemed to have too low an activity to be measured by the CD-based assay.

3.2.8.3 *Bj*TarD

K182M TarD was assessed for activity with D-tartrate as the substrate. Activity was assayed by conducting a continuous coupled spectrophotometric assay using malate

dehydrogenase to reduce the oxaloacetate product of the TarD catalyzed reaction to L-malate (Yew, *et al.*, 2006b). Reaction rates were obtained by quantifying the oxidation of NADH through an observed decrease in absorbance at 340 nm. The assay was conducted in a 1-cm pathlength quartz cuvette (1 mL final reaction volume) at 25 °C. Enzyme (12 µg/mL final concentration) was added to the assay mixture containing D-tartrate (0.2-10 mM), 0.16 mM NADH, and 10 units of L-malate dehydrogenase in TarD assay buffer (50 mM HEPES, 10 mM $MgCl_2$, at pH 7.5). If no enzymatic velocity was observed in measurements up to the maximum enzyme concentration (i.e., no significant difference from the buffer control rate) then the variant enzyme was deemed to have too low an activity to be measured by the assay.

3.2.8.4 *St*TGD

K195M TGD was assessed for activity with *m*-galactarate as the substrate. Activity was assessed using a CD-based assay analogous to the one used for MR, as detailed by Bearne and coworkers (2017). Activity was assayed by observing the change in ellipticity at 323 nm using a 0.5-cm pathlength following the conversion of *m*-galactarate to 5-keto-4-deoxy-D-glucarate. All assays were conducted in TGD assay buffer (50 mM Tris, 10 mM $MgCl_2$ at pH 8.0) at 25 °C. Assays were conducted by adding K195M TGD (1608 µg/mL) to a solution of *m*-galactarate in assay buffer at a final volume of 1 mL. The concentration of *m*-galactarate used for assaying K195M ranged from 0.1-6.0 mM. If no enzymatic velocity was observed in measurements (i.e., no significant difference from the buffer control rate) then the variant enzyme was deemed to have too low an activity to be measured by the CD-based assay.

3.2.9 DIFFERENTIAL SCANNING CALORIMETRY

A Malvern Panalytical (Microcal VP-DSC) Differential Scanning Calorimeter was used (Malvern Instruments Inc., Northhampton, VA). Scans were analyzed using Malvern *Origin* v. 7.0 software. All scans were conducted from 20-100 °C at a scan rate of 45 °C/h. The DSC was calibrated using assay buffer (100 mM HEPES, 3.3 mM $MgCl_2$, pH 7.5) for a minimum of three scans prior to the protein scan. Protein scans were completed using a protein concentration of 7.3 µM in assay buffer. The corresponding thermograms were plotted after subtraction of the scanned baseline control and the thermal melting temperature T_m was obtained directly from the plot at the maximum of the two-state transition.

3.2.10 LIGAND STABILIZATION STUDIES

Ligand stabilization studies were conducted using the previously described DSC method. DSC thermal denaturation data for protein samples (7.3 µM) in the presence of the boron-based inhibitors phenylboronic acid, 4-chlorophenylboronic acid, and 3,4-dichlorophenylboronic acid (0-200 µM) were obtained. The T_m values were plotted as a function of inhibitor concentration and fit to a rectangular hyperbola with a y-offset value equal to the T_m value of unliganded MR using non-linear regression analysis through *KaleidaGraph* software (eqn 3.3) (v. 4.02, Synergy Software, Reading, PA). The saturation value of T_m is reported as the $T_{m, max}$ and the K_d^m value is the concentration of inhibitor (I) that yielded a T_m value that is one-half the value of the $T_{m, max}$.

$$T_m = T_m^0 + \frac{T_m^{max}[I]}{K_d^m + [I]} \qquad (3.3)$$

3.3 RESULTS

3.3.1 GENERATION AND CHARACTERIZATION OF MR SUBGROUP VARIANT PROTEINS

To assess the energetic contribution of the conserved KxK residue to the enzyme stability, variant proteins for representative MR subgroup member proteins were generated using site-directed mutagenesis. The purity of the produced proteins was found to be >95% using SDS-PAGE (**Figure 3.5**) and the effects of mutation at the KxK residue on activity and secondary structure were assessed by enzymatic assays and CD spectroscopy (**Figure 3.6**), respectively. Enzymatic activity determinations showed a complete loss in activity for all of the KXK → MXK variant proteins (data not shown). K164R MR had reduced activity compared to wild-type with K_m values of 1.4 ± 0.2 mM ($R \rightarrow S$) and 1.2 ± 0.2 mM ($S \rightarrow R$), as well as k_{cat}/K_m values of 3.3 ± 0.4 M^{-1}s^{-1} ($R \rightarrow S$) and 3.5 ± 0.6 M^{-1}s^{-1} ($S \rightarrow R$) (Douglas & Bearne, unpublished).

It should be noted that the T_m for K182M TarD had to be obtained using a modified purification protocol. It was found that K182M TarD was highly unstable in the TarD assay buffer typically used for DSC analyses on the TarD variants. K182M TarD produced insoluble aggregates in the assay buffer. To reduce the contaminating proteins present in K182M TarD preparations, the *E. coli* LOBSTR cell strain was used. LOBSTR cells carry genomically modified copies of native *E. coli* genes *arnA* and *slyD*, such that

the protein products exhibit reduced affinity for Ni^{2+} affinity chromatography resin

(Andersen, Leksa, & Schwartz, 2013). The resulting purified protein is visualized by

SDS-PAGE in **Figure 3.5**.

3.3.2 ANALYSIS OF THE THERMAL STABILITY OF THE <u>K</u>XK VARIANT PROTEINS OF THE MR SUBGROUP

The contribution of the Lys 164 residue of MR and the analogous residues of the

representative members of the MR subgroup were probed by mutation to either

methionine or arginine (K164M MR, K164R MR, K218M FucD, K182M TarD, and

K195M TGD). The impact of the substituted residue on enzyme stability was observed

using differential scanning calorimetry, wherein ΔT_m indicates the impact that the

substitution had on the stability of the protein relative to the wild-type T_m. The thermal

melting data (**Figure 3.7, 3.8**) are listed in **Table 3.2** along with the T_m values for the

wild-type proteins for comparison. Substitution of the <u>K</u>xK residue by a methionine

yielded a negative ΔT_m for both MR and TarD, while having a negligible impact on T_m

for both FucD and TGD. The K164R MR variant had a ΔT_m of 8.36 °C that is nearly two-

fold greater than that of K164M in the negative direction (**Figure 3.8**).
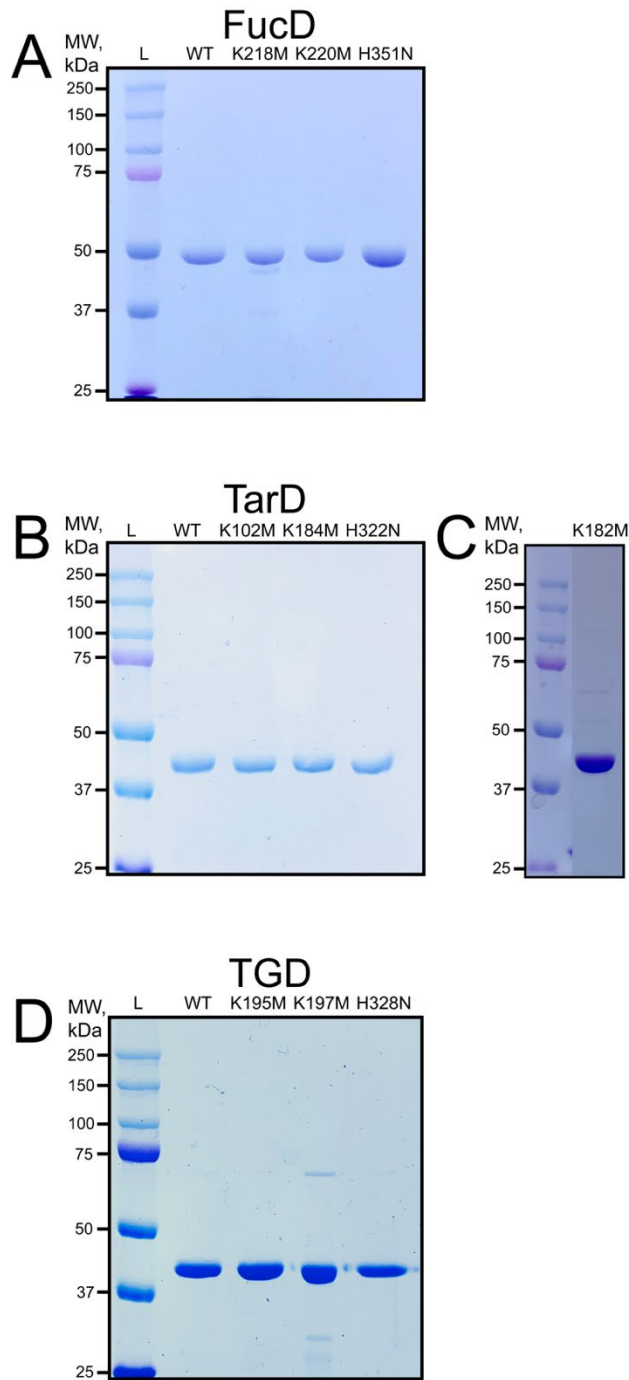
Figure 3.5. Representative 10% SDS-PAGE electrophoretograms assessing the purity of variant MR subgroup proteins. Lane "L" corresponds to the molecular weight ladder in each. Lanes are labelled with the identity of each variant for FucD (A), TarD (B and C), and TGD (D).
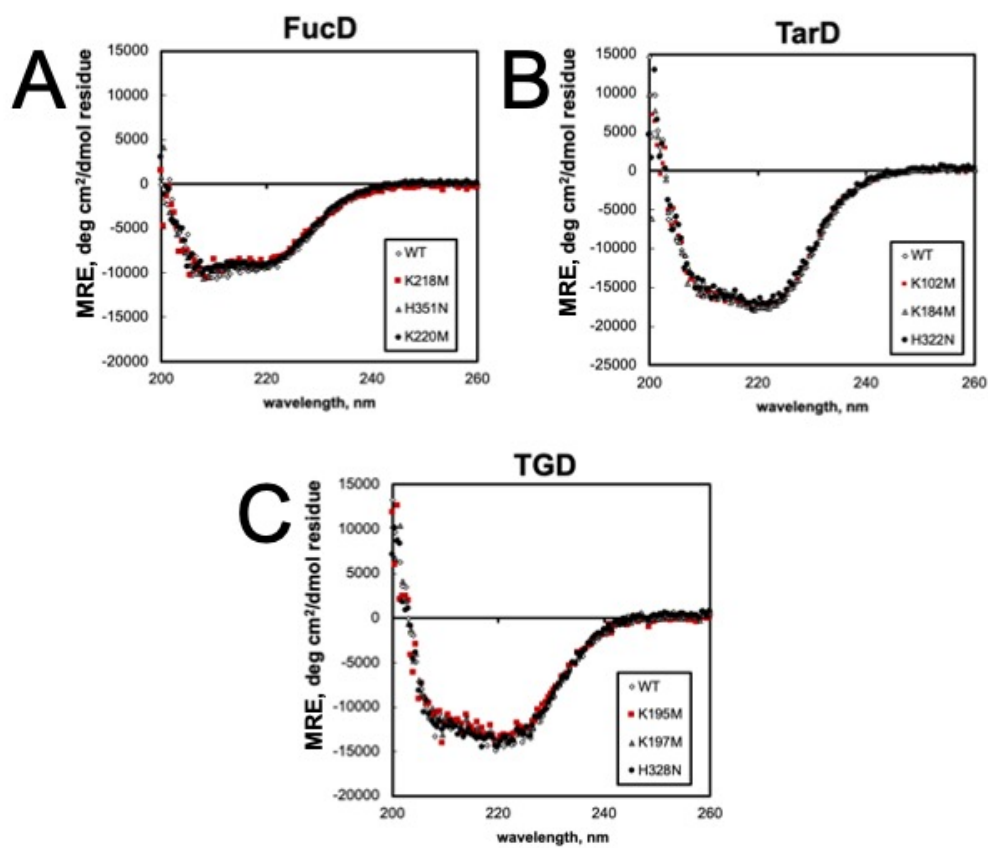
Figure 3.6. Representative CD spectra for generated variant proteins of MR subgroup enzymes. Overlaid CD spectra of (A) wild-type, K218M, K220M, and H351N FucD, (B) wild-type, K102, K184, and H322N TarD, and (C) wild-type, K195M, K197M, and H328N TGD. For CD spectra of MR variants K164M, K164R, K166M, and K166R see Douglas, *et al.*, unpublished.

Table 3.2. Comparison of thermal melting temperatures for MR subgroup $\underline{K}$xK variant proteins. The reported T_m value of TarD was obtained using the alternative purification protocol, in which the sample buffer contained 250 mM imidazole.

Variant	T_m, °C	ΔT_m, °C
WT MR	63.6 ± 0.2[a]	–
K164M MR	59.6 ± 0.3	–4.03
K164M MR	55.27 ± 0.07	–8.36
WT FucD	68.93 ± 0.03[a]	–
K218M FucD	69.10 ± 0.04	0.17
WT TGD	89.2 ± 0.3[a]	–
K195M TGD	88.8 ± 0.3	–0.33
WT TarD	61.1 ± 0.2[a]	
	55.7 ± 0.1[b]	–
K182M TarD	47.5 - 55.4[b]	–0.28 - –8.23

[a]Value from Kumar & Bearne (unpublished)
[b]Values obtained using assay buffer containing 250 mM imidazole and 500 mM NaCl
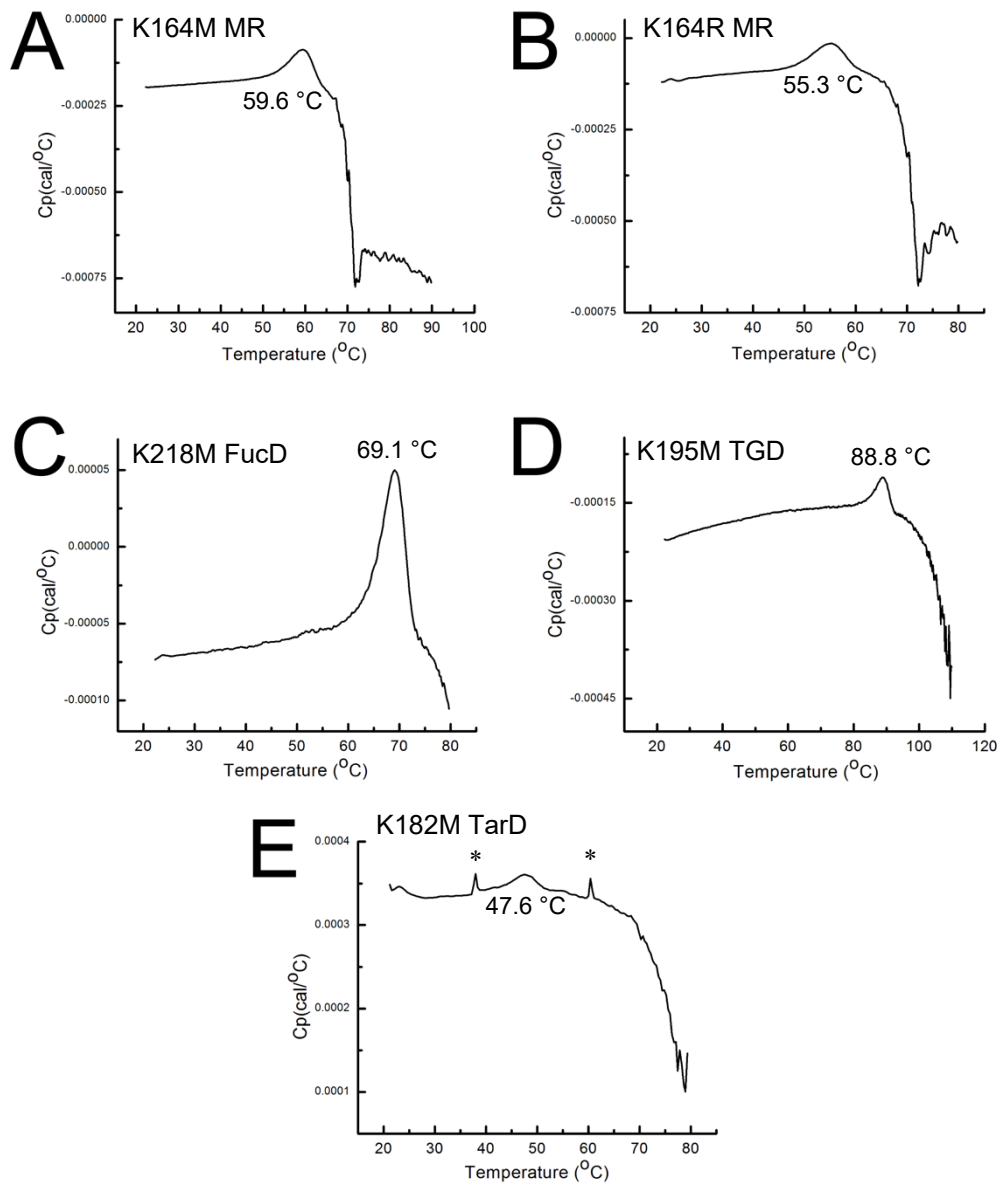
Figure 3.7. Representative DSC thermograms for the K164M MR, K164R MR, K218M FucD, K195M TGD, and K182M TarD variants. T_m values were obtained from the observed maximum in the shown $C_p(T)$ functions obtained for (A) K164M MR (59.6 °C), (B) K164R MR (55.3 °C), (C) K218M FucD (69.1 °C), (D) K195M TGD (88.8°C), and (E) K182M TarD (47.6 °C) variants. (Note that two anomalous peaks (*) arising from electronic noise are present in (E)).
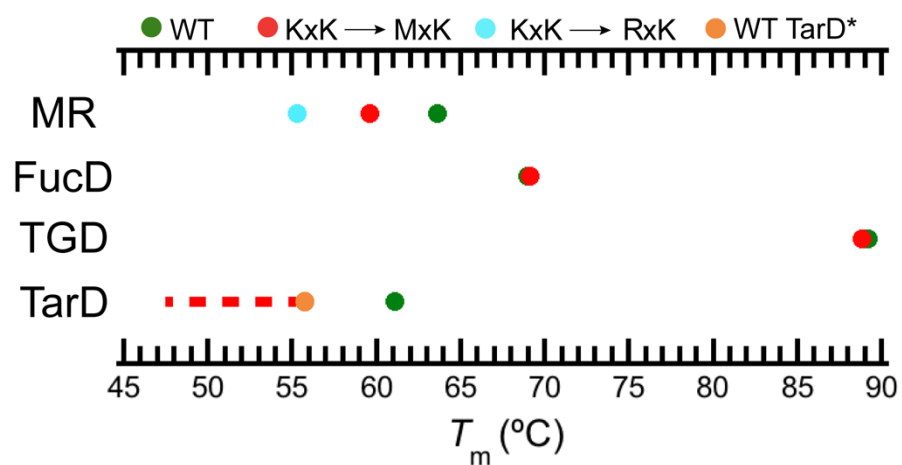
Figure 3.8. Comparison of thermal melting temperatures for MR subgroup KxK variant proteins. The T_m for wild-type protein is shown in green, KxK → MxK variants are shown in red, KxK → RxK variants are shown in cyan, and the T_m corresponding to wild type TarD purified under alternative conditions is shown in orange. The red dashed line corresponds to the range of T_m values obtained for K182M TarD.

3.3.3 THERMAL STABILIZATION OF MR BY LIGAND BINDING

The thermal stabilization effected by the binding of the potent boronic acid-based inhibitors discussed in Chapter 2 (PBA, 4-chloroPBA, and 3,4-dichloroPBA) was assessed using DSC. The present study showed that PBA, 4-chloroPBA, and 3,4-dichloroPBA increased the thermal melting temperature of MR from 63.63 °C in the absence of ligand to $T_{m, max}$ values of 75.68 ± 0.24 °C, 80.07 ± 0.33 °C, 82.98 ± 1.05 °C, respectively (**Figure 3.9**).
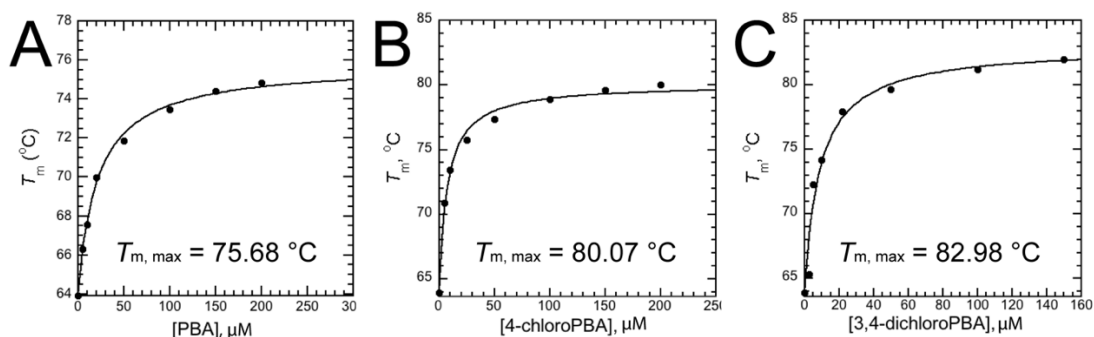
Figure 3.9. Thermal stabilization of wild-type MR by PBA, 4-chloroPBA, and 3,4-dichloroPBA. Shown are plots depicting the hyperbolic dependence of T_m for wild-type MR on the concentration of boron-based inhibitor. The changes in thermal stability correspond to thermal denaturation of wild-type MR in the presence of (A) PBA, (B) 4-chloroPBA, and (C) 3,4-dichloroPBA. For wild-type MR, $T_m = 63.63$ °C.

3.4 DISCUSSION

3.4.1 CONTRIBUTION OF THE <u>K</u>XK RESIDUE IN THE MR SUBGROUP TO PROTEIN THERMAL STABILITY

Interest in the Lys 164 residue of MR and the conserved <u>K</u>xK residue in the MR subgroup initially arose to better understand the effect that its positive charge had on the pK_a value of the acid-base catalyst Lys 166. Bearne and co-workers found that Lys 164 is essential to MR catalysis, affecting the catalyzed reaction by destabilizing the positively charged $\varepsilon\text{-}NH_3^+$ side chain of Lys 166 through electrostatic interactions, thus reducing the pK_a of the residue (**Figure 3.10**) (Fetter, *et al.*, 2019). In the present work, I investigated whether the positive charge on Lys 164 would have a destabilizing effect on MR and other members of the MR subgroup, similar to that observed for the second lysine of the Kx<u>K</u> motif in the MR subgroup. DSC data for the <u>K</u>xK series of variants of the MR subgroup revealed that only the thermal stability of K164M MR and K182M TarD exhibited a significant change. That said, all <u>K</u>xK → <u>M</u>xK variant proteins were shown to be catalytically inactive. Such a dependence on the <u>K</u>xK residue for catalysis is characteristic for a "first shell" residue that directly interacts with ligands bound to the active site. The selective dependence of the thermal stability on the identity of the <u>K</u>xK residue is interesting because of its similarity in position and identity to the Kx<u>K</u> residue but contrasting contribution to enzyme stability.
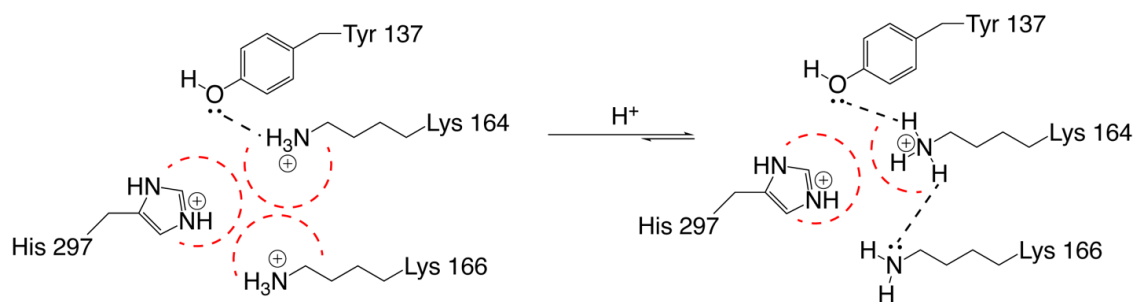
Figure 3.10 Mechanism of the Tyr 137-and Lys 164-modulated reduction of the pK_a of Lys 166. Hydrogen bonding is represented by a black dashed line and positive-positive charge repulsions are represented by red dashed semicircles (adapted from Fetter, *et al.*, 2019).

In order to discuss the stabilizing effect of Lys 164 and the KxK residue of the MR subgroup, two functional aspects must be covered: its modulatory effects on the pK_a of the KxK residue and its function as a substrate/ligand binding determinant. For this discussion I will focus on MR because more extensive studies have been conducted on this enzyme (Bearne & St. Maurice, 2017). Lys 164 is believed to have the strongest effect on modulating the pK_a of Lys 166. The K164C and K164M MR variants exhibited a marked reduction in catalytic efficiency, hindering kinetic studies (Fetter, *et al.*, 2019). Consequently, binding data were obtained directly using ITC to measure the binding affinity of K164C MR for BzH. The binding affinity of K164C MR for BzH (K_d) was found to be 94.6 ± 9.6 μM, roughly 26-fold weaker than the affinity of wild-type MR for BzH, yet 17- and 32-fold greater than the affinity of K166M and K166C MR. This likely arises from the fact that H-bonding interactions are lost in the K166C MR variant, but affinity is retained from the ε-NH$_3^+$–π interactions between the protonated Lys 166 and

the phenyl and hydroxamate/hydroximate functionalities of BzH (Fetter, *et al.*, 2019; Nagar & Bearne, 2015).

Of the MR variants studied in this work (K164M and K164R MR), the substitution of the lysine served to assess the ability of the enzyme to tolerate changes in either the electrostatic or steric properties at that position. In the case of K164M MR, the steric properties of the lysine side chain were conserved, but the positive charge was removed. The resulting ΔT_m for K164M MR was -4.03 °C relative to wild-type MR, indicating that, at the very least, the positive electrostatic character of Lys 164 contributed to the thermal stability of the protein. Secondly, a lysine to arginine substitution at the K̲xK position (K164R MR) increased the bulk of the residue at position 164, while maintaining the positive charge of the residue. In this case, the ΔT_m was -8.35 °C, which was more destabilizing than the K164M mutation. Thus, position 164 appears to be highly sensitive to added steric bulk. Of course, synergistic effects between Lys 164 and nearby residues that stabilize the enzyme may also be adversely affected upon substitution. Notable interactions between residues that are shown to increase thermal stability include hydrogen bonding, salt-bridges, and cation-π interactions (Vogt, Woell, & Argos, 1997; Chakravarty & Varadarajan, 2002). Of these interactions, Lys 164 is involved in a hydrogen bonding network with Lys 166 and Tyr 137. Furthermore, ε-NH_3^+ is near the phenyl group of Tyr 137 giving rise to a potential cation-π interaction (Fetter, *et al.*, 2019).

Substitution of Lys 164 could result in the loss of the Lys 164-Tyr 137 hydrogen bond, any potential cation-π interactions between Lys 164 and Tyr 137, and loss of an essential binding determinant. As a result, the destabilization arising from substitutions of

Lys 164 relative to substitutions of Lys 166 may simply be caused by loss of additional interactions and their corresponding thermal stabilization. Nevertheless, the change in thermal stability (ΔT_m) observed in the case of the Lys 164 variant proteins is common for removal of first-shell residues (Sayer & Louis, 2008; Yakovlev, *et al.*, 2009).

A number of studies involving thermal stability analysis of site-directed mutant proteins have shown similar ΔT_m values for active site mutations. Research done by Louis and co-workers on the interactions of active site residues of HIV-1 protease (160 kDa homodimer) with inhibitors showed ΔT_m values of +7.3 °C and +6.7 °C for two aspartate residues responsible for intra- and inter-subunit hydrogen bonds to stabilize the protease fold and inhibitor binding (Sayer & Louis, 2008). Thermal denaturation was irreversible in this case. Meanwhile, Makarov and colleagues observed varying effects for substitutions of active site residues of ribonuclease Sa (96 residues) in the ΔT_m range of +0.1 to +7.9 °C for a reversible denaturation (Yakovlev, *et al.*, 2009). Through site-directed mutagenesis studies, it was observed that residues directly involved in catalysis ("first-shell" residues) contributed stability to the variant proteins generated, while residues that indirectly impacted catalysis were found to either destabilize the enzyme or have no effect on the thermal stability. This may have occurred as a result of the smaller size of the protein, requiring essential residues to have a stabilizing effect on the enzyme (Yakovlev, *et al.*, 2009). Interestingly, I observed the same trend.

It has long been thought that residues that participate in enzyme catalysis are optimized in such a way that promotes and prioritizes catalysis over contributions to stability of the enzyme (Williams, 1972; Warshel, 1978). This idea falls in line with the induced fit model of substrate binding to an enzyme, where conformational flexibility of

the free enzyme promotes substrate binding, which is closely followed by a

conformational change by the enzyme to its catalytically active conformer (Koshland,

1994). Though one could expect the thermal stability to increase with the size of a protein

due to the burying of hydrophobic regions of backbone and side chains, this does not, in

fact, occur. As the size of the protein increases, the relative content of polar and

hydrophobic residues increases to a lesser extent than the content of charged amino acids,

which are far more likely to destabilize the protein (Kajander, *et al.*, 2000). This allows

the potential for enzymes to bury charged residues to accomplish greater catalysis, while

concurrently decreasing the thermal stability.

Studies on *E. coli* ribonuclease HI (a 155 residue, reversibly denaturing protein)

by Liu and co-workers have helped to elucidate the energetic values that correspond to

changes is T_m due to perturbation of residue interactions by site-directed mutagenesis

(Kanaya, Oobatake, & Liu, 1996). Primarily, it was shown that for negatively charged

residues at the active site that are held proximal to each other, the resulting repulsive

strain of the interaction destabilized the enzyme. This was attributed to the free enzyme

requiring increased conformational freedom to retain optimal function (Kanaya,

Oobatake, & Liu, 1996). Liu and colleagues observed similar ΔT_m values to our data for

(1) stabilization by addition of cofactor Mg^{2+} (3.0 kcal/mol, ΔT_m = 10.3 °C) and (2)

stabilization by removal of a repulsive charged interaction (1.07-2.25 kcal/mol, ΔT_m =

3.7-7.4 °C) (Kanaya, Oobatake, & Liu, 1996). A similar study by Grimsley and

colleagues on the contribution of hydrogen-bonding to protein stability showed that the

destabilization arising from removal of an H-bonding interaction for eight variants of

ribonuclease T1 (a 130 residue, reversibly denaturing protein) was on average –9.5 °C

and –2.92 kcal/mol (Pace, *et al.*, 2014). Although, it should be noted that the stabilization provided by hydrogen bonding interactions is heavily context dependent.

Energy values were obtained from the thermal denaturation process according to **equations 3.3-3.5**, under the assumption of a reversible system that follows a two-state denaturation mechanism and where all thermodynamic values are those for the protein unfolding process (Kanaya, Oobatake, & Liu, 1996). Of course, the enthalpy values for this reversible system can be compared to the MR system studied in this body of work under the caveat that MR undergoes irreversible thermal denaturation and enthalpy values cannot be accurately obtained.

$$\Delta G(T) = \Delta H(T) - \Delta T \Delta S(T) \tag{3.3}$$

$$\Delta H(T) = \Delta H_m + \Delta C_p(T - T_m) \tag{3.4}$$

$$\Delta S(T) = \Delta S_m + \Delta C_p \ln\left(\frac{T}{T_m}\right) \tag{3.5}$$

It remains a possibility that the positive-positive charge repulsion between the KxK residues of the MR subgroup induce a strain to afford a conformational freedom for the enzyme that prioritizes catalysis over stability. This aligns with the explanation of the protonation state of the free enzyme, which suggests that the majority form of free enzyme involves a protonated KxK residue with the Brønsted acid-base histidine remaining in its basic (deprotonated) form. Under these conditions, it is likely that

resolution of the destabilization from the positive charge repulsion state can arise upon binding of ligand. It so happens that in the presence of a saturating amount of (R)-atrolactate (a substrate analogue for which Lys 166 of MR acts as the Brønsted acid towards), the T_m of the MR·(R)-atrolactate complex is nearly identical to that of K166M MR. Of course, the resulting similarity in thermal stabilization could be fortuitous.

3.4.2 THERMAL STABILIZATION OF MR USING HIGH-AFFINITY LIGANDS

Binding of ligand to protein can only occur if there is a release of free energy, which stabilizes the protein-ligand complex against thermal denaturation (Holdgate & Ward, 2005). The extent of the thermal stabilization depends on the magnitude of the binding energy. Given that the MR subgroup of enzymes denature irreversibly, estimation of this binding energy is not reliably obtained using DSC. However, a series of inhibitors of MR were assessed for their thermal stabilization potential in binding to MR. Benzohydroxamate (K_d = 3.7 ± 0.1 μM) stabilized MR by 11.5 °C relative to the non-liganded counterpart. (S)- and (R)-atrolactate, which are analogues of (S)- and (R)-mandelate, stabilized MR by 9.3 and 6.5 °C, respectively (Kumar & Bearne, unpublished). The thermal denaturation data (T_m) with bound ligand was plotted as a function of inhibitor concentration, and a rectangular hyperbola (eqn. 3.2) was fit to the data with a y-offset equal to the T_m for the non-liganded MR (63.63 °C) to obtain the saturation $T_{m,\,max}$. These data led to the hypothesis that the PBA series of inhibitors described in Chapter 2 may provide additional thermal stability to the MR·ligand complex. Though the predictive quality of the data is hindered by the irreversibility of the

system, thermal stabilization from ligand binding can elucidate valuable information about the ligand-protein complex (Holdgate & Ward, 2005).

It became immediately obvious that the stabilization offered by PBA, 4-chloroPBA, and 3,4-dichloroPBA was far greater than that of any of the previously characterized inhibitors, with 3,4-dichloroPBA showing the largest increase in thermal stability (**Figure 3.9**) with a $T_{m, max}$ value of 82.98 °C. This stabilization likely arises from the same active site interactions that afford such potent inhibition as described in Chapter 2. Unfortunately, comparison of the thermal stabilization from inhibitor to inhibitor is hindered by the inability to obtain true values of thermodynamic data from the DSC analysis. Furthermore, affinity estimates and thermodynamic data obtained from DSC measurements should be used carefully. The magnitude of ΔT_m observed for different ligands has been shown to differ greatly even between ligands with very similar affinities for the protein (Holdgate & Ward, 2005). Also, changes in ΔT_m for compounds with similar affinities depends on both entropic and enthalpic contributions from the binding energy, where greater T_m shifts are often observed for entropically driven binding processes.

Therefore, when comparing MR inhibitors PBA, 4-chloroPBA, and 3,4-dichloroPBA, the high level of structural similarity may allow for speculative conclusions to be drawn from their effects on the stability of MR, even with the large differences in binding affinity. The largest change in $T_{m, max}$ value is observed with the addition of the 4-chloro substituent on the phenyl ring, which has been shown to increase the potency of inhibition through electron withdrawing effects (Sharma, *et al.*, 2020). Additionally, the negative ΔC_p value obtained for PBA in Chapter 2 (-395 cal·mol^{-1}K^{-1}) (Lin, Schwarz, &

Eisenstein, 1995) paired with the entropically driven binding process that was discovered in ITC binding studies indicates that hydrophobic interactions occur between PBA and MR. Altogether, the marked increase in thermostability of MR complexed with the series of PBA inhibitors likely arises from the entropically driven binding process, hydrophobic interactions at the active site, and the binding energy for the series of potent inhibitors.

3.5 CONCLUSIONS AND FUTURE WORK

The positively charged <u>K</u>xK residues of MR (Lys 164) and TarD (Lys 182) have been shown to contribute to the thermal stability of the free enzymes. As shown for MR, Lys 164 affects catalysis by participating as both a binding determinant for the carboxylate functionality of bound substrate and as a member of the hydrogen-bonding network comprising the "catalytic triad" of Lys 164, Lys 166, and Tyr 137 that serves to lower the pK_a of the Brønsted acid-base catalyst Lys 166 (Fetter, *et al.*, 2019). Towards contributing to the thermal stability, however, Lys 164 provides a hydrogen bond and has a potential cation-π interaction between the ε-NH$_3^+$ and the aromatic ring of Tyr 137. The significant stabilization provided by these interactions would account for the ΔT_m of – 4.03 °C for K164M MR relative to wild-type MR. K164R MR caused a further decrease in thermal stability (ΔT_m = -8.36 °C), which can likely be attributed to steric effects as well as loss of hydrogen bonding interactions.

In the case of TarD, altered purification conditions were required to obtain pure and stable protein. K182M TarD had observed T_m values ranging from 47.3-55.2 °C relative to the T_m of wild-type TarD at 55.5 °C. The range in values was attributed to the instability of the variant protein, which could have been obtained in solution in a batch-

dependent misfolded state. In any case, Lys 182 is essential to TarD catalysis since

K182M TarD was inactive. The corresponding lack of change in the thermal stability of

FucD and TGD may be due to the increased overall thermal stability of these proteins,

especially TGD. This insensitivity of the thermal stability to residue substitution was also

observed with substitution of the Brønsted base catalytic His, and both lysines of the

KxK motif, which was paired with heightened overall thermal stability of the free

enzymes (i.e., both TarD and TGD have T_m values significantly greater than those of MR

and TarD).

To elaborate upon the destabilization seen by site-directed mutagenesis of the

KxK residue in the MR subgroup, ITC binding studies using the variant protein with

substrate and transition state analogue inhibitors should be done to obtain binding and

thermodynamic data. Binding thermodynamic studies may give rise to information on the

loss of catalytic function in the variants corresponding to changes in conformational

stability due to substitution of the KxK. Overall, elucidation of the total function of the

KxK residue of the MR subgroup may allow for a deeper understanding of MR subgroup

catalysis.

CHAPTER 4. APPROACHES TOWARDS PRODUCTION OF REVERSE THYMIDYLATE SYNTHASE

4.1 INTRODUCTION

From the perspective of an enzymologist, the rTS isoforms pose an intriguing problem. rTSγ is functionally active as an L-fuconate dehydratase within the MR subgroup of the ENS, while rTSβ, lacking only the 27 N-terminal residues of rTSγ, is catalytically inactive (Wichelecki, *et al.*, 2014). Additionally, rTSβ appears to be associated with chemotherapy resistance since its overexpression is reported to be induced by the presence of MTX and 5-FU, making this a protein of interest from a clinical perspective (Dolnick, 1993, Kuo, *et al.*, 2008; Srimatkandada, *et al.*, 1983; Dolnick, *et al.*, 1996a; Dolnick, *et al.*, 1997; Lin & Chow, 2010). Experimental characterization of rTSβ has largely been hindered by lack of soluble protein production in *E. coli*, which is likely due to the absence of the 27 amino acids at the N-terminus that, as discussed previously, are known to confer stability to the protein. The rTSγ isoform has not proven to be problematic in this regard, since published reports clearly show its soluble production using a pNIC26-Bsa4 vector in *E. coli* BL21(DE3)-R3-pRARE2 cells (Wichelecki, *et al.*, 2014).

The present chapter describes various attempts towards discovering an effective method for the production of soluble rTSβ in amounts sufficient to pursue functional characterization. Since soluble rTSγ can be obtained in ample quantities, my approach has been to produce rTSγ variants with an engineered protease recognition site. This would allow for post-translational cleavage by a given protease, yielding soluble rTSβ

(**Table 4.1**). In addition, attempts to obtain soluble rTSβ using the well-known solubility enhancing maltose binding protein (MBP) as an N-terminal protein tag to increase overall yields is described (see **Figure 4.1).**

Production of soluble rTSβ is essential in order to facilitate experimental analyses in search of a functional role for the protein. Only once soluble production of the protein is obtained, can methods of substrate or binding partner determination be carried out. Research by Gerlt and colleagues provided a precedent to follow in determination of rTSβ function. The use of differential scanning fluorimetry was shown to be a viable method of analysis for monitoring binding of ligands, as is seen with rTSγ and D-erythronohydroxamate (Wichelecki, *et al.*, 2014). Therefore, utilization of differential scanning calorimetry, a method widely used in the Bearne lab, may prove useful in experimental characterization of rTSβ towards discovery of a binding partner. Further, attempts to characterize rTSβ from this point forward should primarily focus on binding partners, and not necessarily substrates for catalysis. This point arises from the observed lack of catalytic activity of rTSβ with any of the sugar acid-based substrates for the longer rTSγ isoform and the potential *in vivo* interactions between rTSβ and TS/DHFR.

Table 4.1. Common proteases considered for the rTS purification systems. | indicates the main cleavage site of the protease (Waugh, 2011).

Cleavage Tag	Sequence and Cleavage	Enzyme Name	Notes
Factor Xa cleavage tag	IEGR|	Factor Xa	Factor Xa cleaves after the Arg residue but can also cleave less frequently at secondary basic sites. Its most common secondary cleavage site is between the Gly and Arg residues in its own recognition site, although the frequency of these events is protein specific.
TEV (tobacco etch virus) cleavage tag	ENLYFQ|G	Tobacco etch virus protease	Cleavage occurs between the Gln and Gly residues. TEV is often reported to have better specificity for its recognition site compared to enterokinase, Thrombin, or Factor Xa.
Thrombin cleavage tag	LVPR|GS	Thrombin	Thrombin cleaves preferentially between the Arg and Gly residues. Off target cleavage can occur at nonspecific sites, normally from contaminating proteases.
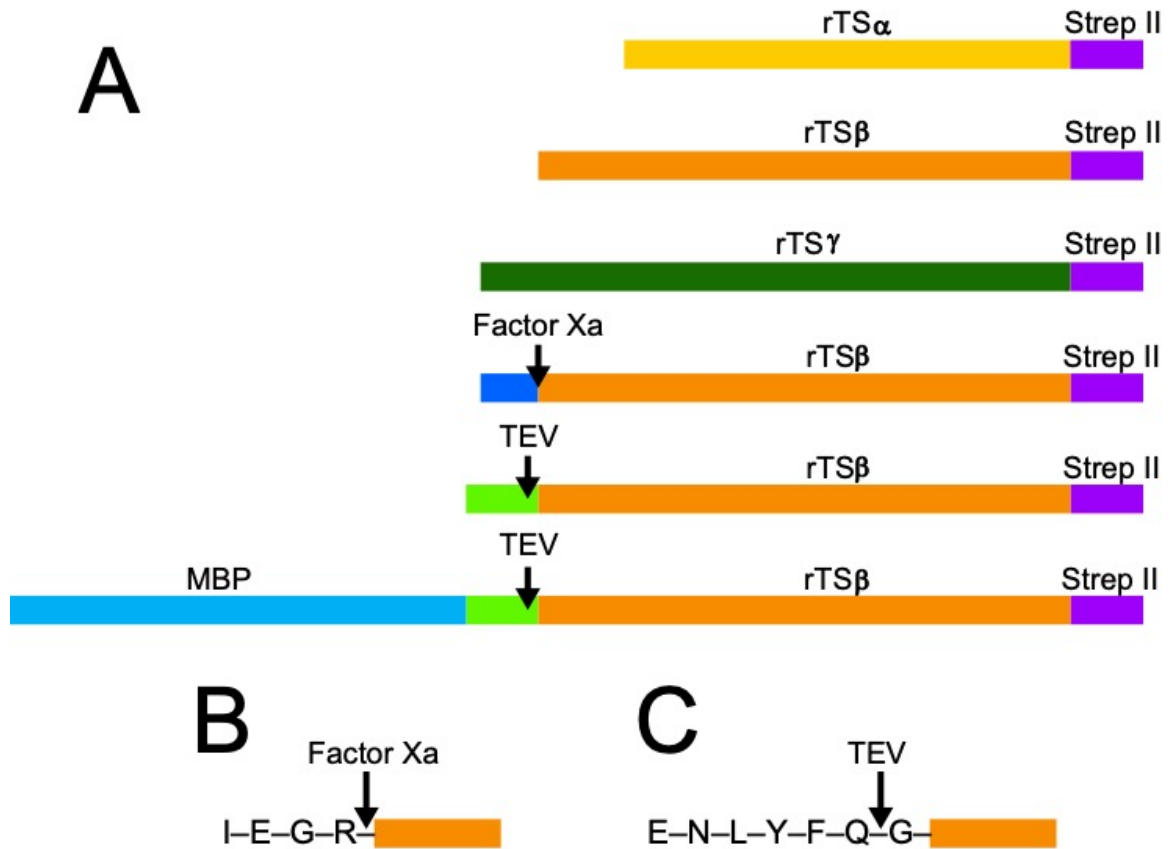
Figure 4.1. Schematic representation of various protein constructs and variants used within this chapter. The series of protein constructs discussed in this work are shown (A). The recognition sequence for Factor Xa protease is I-E-G-R where cleavage is catalyzed after the Arg residue (B). The recognition sequence for TEV protease is E-N-L-Y-F-Q-G, where cleavage is catalyzed after the Gln residue, leaving a Gly residue appended to the cleavage product (C).

4.2 MATERIALS AND METHODS

4.2.1 GENERAL

All reagents were purchased for Sigma-Aldrich Canada Ltd. (Oakville, ON, Canada) unless otherwise stated. Circular dichroism (CD) spectra were collected using a JASCO J-810 spectropolarimeter (Jasco Inc., Easton, MI). UV-visible and OD_{600} spectrophotometric readings were obtained using a HP 8453 UV-visible spectrophotometer. Proton nuclear magnetic resonance (^1H NMR) spectra were obtained using a Bruker AV 500 MHz spectrometer at the Dalhousie University Nuclear Magnetic Resonance Research Resource Centre (NMR-3).

4.2.2 MOLECULAR CLONING

Commercially synthesized open reading frames encoding *Hs*rTSα, *Hs*rTSβ, and *Hs*rTSγ (BioBasic Inc., Markham, ON) were subcloned into pPSG-IBA03 vectors (IBA Life Sciences, Göettingen, Germany), which incorporated a C-terminal Strep-II fusion tag in the proteins. After digestion of the synthesized genes was completed using the restriction enzyme Esp3I, the products were ligated into Esp3I-digested pPSG-IBA3 plasmids. The *rTSα, rTSβ*, and *rTSγ* genes in pPSG-IBA3 were expressed in *E. coli* BL21 (DE3) cells.

4.2.3 PROTEIN EXPRESSION, PURIFICATION, AND QUANTIFICATION

4.2.3.1 General protocol for overexpression of open reading frames in a pPSG-IBA3 plasmid in *E. coli*

Small scale cultures (5 mL) of *E. coli* BL21 Star™(DE3) (Thermo Fisher Scientific, Inc., Waltham, MA) cells were grown overnight at 37 °C from glycerol stocks

in sterile LB media containing ampicillin (100 µg/mL). Large scale expression cultures (2 L LB media, 100 µg/mL ampicillin) were inoculated with the overnight cultures at a volume of 10 mL per litre of LB media and incubated at 37 °C with shaking (225 rpm). The cultures were incubated until an OD_{600} of 0.5-0.6 was obtained, at which point expression was induced by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.1 mM. The cultures were then shaken and incubated at 22 °C for 24 h. Cells were harvested using centrifugation (3795 × g, 10 min, 4 °C), and washed with buffer W (100 mM Tris-Cl, 150 mM NaCl, 1 mM EDTA, pH 8.0). Cell pellets were stored at –20 °C for future use.

4.2.3.2 General protocol for purification of Strep-tagged fusion proteins using StrepTactin® XT affinity chromatography resin

Frozen cell pellets were thawed and resuspended in cold buffer W (~35 mL). The cell suspension was kept on ice and sonicated for 6 × 30 s using a Branson Sonifier 250 with 1 min "resting" periods in between sonication (setting 5.5, 1-s bursts). The cell lysate was clarified by ultracentrifugation (146,550 × g, 35 min, 4 °C) and applied to a gravity column (10 mL) containing 1 mL of Strep-Tactin® XT affinity resin (IBA Life Sciences, Göttingen, Germany). Protein was eluted by addition of buffer W containing D-biotin (1 mM). Eluted samples were dialyzed over 3 × 8 h against assay buffer (50 mM Tris, 10 mM $MgCl_2$, pH 8.0), with each 8-h dialysis using assay buffer (500 mL). Protein was aliquoted (1 mL) and stored at –20 °C for future use. All protein constructs in this work are schematically represented in **Figure 4.1**.

4.2.3.3 rTSβ

Non-denaturing purification

Wild-type rTSβ was expressed and purified according to the protocol described by IBA Life Sciences for *E. coli* cells containing the pPSG-IBA03 plasmid (IBA Life Sciences, Göttingen, Germany). Refer to section 4.2.3.1 for overexpression and culturing conditions. Protein purification conditions for rTSβ followed the outlined procedure in section 4.2.3.2.

Denaturing purification

rTSβ was expressed and purified according to the protocol described by IBA Life Sciences for *E. coli* cells containing the pPSG-IBA03 plasmid (IBA Life Sciences, Göttingen, Germany). Overexpression and culturing conditions are given in section 4.2.3.1.

Frozen cell pellets were thawed and resuspended in cold wash buffer (~35 mL) containing urea (6 M) to denature and solubilize protein (100 mM Tris, 150 mM NaCl, 1 mM EDTA, pH 8.0). The resuspended cells were gently stirred on ice using a magnetic stir bar until a thick translucent solution was produced. The cell suspension was kept on ice and sonicated 6×30 s on a Branson Sonifier 250 with 1 min "resting" periods between sonications (setting 5.5, 1-s bursts). The cell lysate was clarified by ultracentrifugation ($146,550 \times g$, 35 min, 4 °C) and applied to a gravity column containing 1 mL Strep-Tactin® XT affinity resin. Protein was eluted with addition of wash buffer containing D-biotin (1 mM) and urea (6 M). Eluted samples were dialyzed against assay buffer (50 mM Tris, 10 mM $MgCl_2$, pH 8.0). The dialysis solution was

changed five times over the course of 24 h, each time the diffusate was replaced with fresh assay buffer (500 mL) containing successively lower concentrations of urea with each change (6 M, 4 M, 2 M, 1 M, and 0 M). Protein was aliquoted (1 mL) and stored at –20 °C for future use.

Purification from inclusion bodies

An inclusion body isolation was completed for rTSβ, starting from the pelleted cells described above in section 4.2.3.1. Frozen cell pellets were thawed and resuspended in cold wash buffer (~35 mL, 100 mM Tris, 150 mM NaCl, 1 mM EDTA, pH 8.0). The cell suspension was kept on ice and sonicated 6×15 s using a Branson Sonifier 250 with 1 min "resting" periods between sonication periods (setting 5.5, 1-s bursts). The cell lysate was clarified by ultracentrifugation ($146{,}550 \times g$, 35 min, 4 °C) and the clarified lysate was discarded. The pellet was then resuspended in a similar volume of wash buffer. This solution was kept on ice and sonicated 6×15 s using a Branson Sonifier 250 with 1 min "resting" periods in between sonication periods (setting 5.5, 1-s bursts). Similarly, this lysed solution was clarified by ultracentrifugation ($146{,}550 \times g$, 35 min, 4 °C) and the clarified lysate was discarded.

The pellets (containing inclusion bodies) were then resuspended in cold wash buffer (~35 mL) containing urea (6 M) to denature and solubilize protein (100 mM Tris, 150 mM NaCl, 1 mM EDTA, pH 8.0). This solution was applied to a gravity column containing Strep-Tactin® XT affinity resin (1 mL). Protein was eluted with addition of wash buffer containing D-biotin (1 mM) and urea (6 M). Eluted samples were dialyzed against assay buffer (50 mM Tris, 10 mM MgCl₂, pH 8.0). The dialysis solution was

changed five times over the course of 24 h, each time the diffusate was replaced with fresh assay buffer (500 mL) containing successively lower concentrations of urea with each change (6 M, 4 M, 2 M, 1 M, and 0 M). Protein was aliquoted (1 mL) and stored at –20 °C for future use.

4.2.3.4 rTSγ

Wild-type rTSγ was expressed and purified according to the protocol described by IBA Life Sciences for *E. coli* cells containing the pPSG-IBA03 plasmid (IBA Life Sciences, Göttingen, Germany). Refer to section 4.2.3.1 for overexpression and culturing conditions. Protein purification conditions for rTSγ followed the outlined procedure in section 4.2.3.2.

4.2.3.5 rTSγ-Xa

An rTSγ variant (rTSγ-Xa) containing a Factor Xa protease recognition sequence (see Section 4.2.5.1) was expressed and purified according to the protocol described by IBA Life Sciences for *E. coli* cells containing the pPSG-IBA03 plasmid (IBA Life Sciences, Göttingen, Germany). Refer to section 4.2.3.1 for overexpression and culturing conditions. Protein purification conditions for rTSγ-Xa followed the outlined procedure in section 4.2.3.2.

Protease-catalyzed cleavage of the rTSγ-Xa variant was completed by use of Factor Xa protease (New England Biolabs, Ipswich, MA), which recognizes the amino acid sequence Ile-Glu/Asp-Gly-Arg (IEGR) and catalyzes cleavage of the protein after the arginine residue in its recognition sequence. Cleavage reactions were conducted

according to the manufacturer's protocols. Factor Xa protease was added in a 1:50 ratio to the protein of interest (rTSγ-Xa) and carried out in rTSγ assay buffer (50 mM Tris, 10 mM MgCl$_2$, pH 8.0) for 4 h at 25 °C. Progress of the cleavage reaction was assessed using SDS-PAGE (10%) with Coomassie brilliant blue (R-250) staining.

4.2.3.6 rTSγ-TEV

An rTSγ variant (rTSγ-TEV) containing a TEV protease recognition sequence (see Section 4.2.5.2) was expressed and purified according to the protocol described by IBA Life Sciences for *E. coli* cells containing the pPSG-IBA03 plasmid (IBA Life Sciences, Göttingen, Germany). Refer to section 4.2.3.1 for overexpression and culturing conditions. Protein purification conditions for rTSγ-TEV followed the outlined procedure in section 4.2.3.2.

Protease-catalyzed cleavage of rTSγ-TEV and MBP-TEV-rTSβ was completed using TEV protease kindly provided by Dr. David Langelaan (see section 4.2.3.8). The TEV protease was produced as a His$_6$-tagged construct (His-TEV S219V). TEV protease has an amino acid recognition sequence of Glu-Asn-Leu-Tyr-Phe-Gln-(Gly/Ser) (ENLYFQG) and catalyzes cleavage of the protein between the glutamine and glycine/serine residues in its recognition sequence. Protease-catalyzed cleavage reactions were completed at 4 °C with mixing over the course of 8-10 h, using an initial aliquot of TEV protease (250 μL, 1.0 mg/mL), followed by subsequent aliquots of protease added every two hours over the time-course of the experiment. Reactions were performed under reducing conditions in the presence of dithiothreitol (1 mM). Products of the cleavage

reaction were assessed using SDS-PAGE (10%) with Coomassie brilliant blue (R-250) staining.

4.2.3.7 MBP-TEV-rTSβ

An rTSβ construct (MBP-TEV-rTSβ) containing an N-terminal maltose binding protein (MBP) fusion tag and a TEV protease recognition sequence (see Section 4.2.5.1) was expressed and purified according to the protocol described by IBA Life Sciences for *E. coli* cells containing the pPSG-IBA03 plasmid (IBA Life Sciences, Göttingen, Germany). Overexpression and culturing conditions are shown in section 4.2.3.1. Large scale expression cultures contained "terrific broth" (TB) media, while culturing conditions remained otherwise identical to the previously discussed conditions.

Protein purification conditions for the MBP-TEV-rTSβ-StrepII construct followed the outlined procedure in section 4.2.3.2 with minor alterations. Protein was eluted with addition of wash buffer containing D-biotin (1 mM) and dithiothreitol (DTT) (1 mM). Eluted samples were subjected to TEV protease-catalyzed cleavage over the course of 8 h as described above. The reaction mixture was then applied to a gravity column containing amylose affinity resin (New England Biolabs, Ipswich, MA) for the purpose of separating cleaved MBP and uncleaved full-length MBP-TEV-rTSβ from the protein of interest, rTSβ, via affinity purification. The flow-through from the column was then dialyzed over 3×8 h against assay buffer (50 mM Tris, 10 mM MgCl$_2$, pH 8.0), with each 8 h dialysis using fresh assay buffer (500 mL). Finally, the dialyzed mixture was applied to the gravity column containing 1 mL Strep-Tactin® XT affinity resin. Protein was eluted with addition of wash buffer containing D-biotin (1 mM). Eluted samples were dialyzed over 3

× 8 h against assay buffer (50 mM Tris, 10 mM MgCl$_2$, pH 8.0), with each 8 h dialysis using fresh buffer (500 mL). Protein was aliquoted (1 mL) and stored at –20 °C for future use.

4.2.3.8 TEV PROTEASE

TEV protease (His-TEV S219V) was kindly provided by Dr. David Langelaan (Dalhousie University, Halifax, NS) as a glycerol cell stock, which was expressed and purified according to the manufacturer's protocol for *E. coli* cells transformed with the pET-21b plasmid (Millipore Sigma, Burlington, ON). Overexpression and culturing conditions are shown in section 4.2.3.1 with minor alterations. Cells were harvested by centrifugation (3795 × *g*, 10 min, 4 °C), followed by washing with lysis buffer (20 mM Tris, 100 mM NaCl, 5 mM β-mercaptoethanol, 10% glycerol, pH 8.0). Cell pellets were stored at –20 °C for future use.

Protein purification conditions for TEV protease followed the outlined procedure in section 4.2.3.2 with minor alterations. Clarified lysate was applied to a gravity column (15 mL) containing IMAC Sepharose^{TM} 6 Fast Flow Ni^{2+} affinity resin (GE Healthcare, Chicago, IL). Protein was eluted with addition of lysis buffer containing imidazole (300 mM). The eluted proteins further purified using a gravity column (15 mL) containing SP Sepharose^{TM} Fast Flow ion exchange resin (GE Healthcare, Chicago, IL). Protein was eluted from the ion exchange column with addition of buffer containing 500 mM NaCl (20 mM MES, 5 mM β-mercaptoethanol, 10% glycerol, pH 6.0). Eluted sample was aliquoted (250 µL) and stored at –20 °C for future use.

4.2.4 PROTEIN CHARACTERIZATION AND QUANTIFICATION

Protein concentrations were determined using Bradford assays conducted according to the manufacturer's directions (Bio-Rad Laboratories, Missisauga, ON) with bovine serum albumin standards. Alternatively, the protein concentration was determined using UV absorbance at 280 nm using extinction coefficients determined by the ProtParam web tool from ExPASy under the assumption that all cysteine residues were reduced, as can be seen in **Table 4.2** (Gasteiger, *et al.*, 2005). Protein purity was assessed by SDS-PAGE (10%) with Coomassie brilliant blue (R-250) staining.

Table 4.2. Calculated molecular weight and estimated molar extinction coefficients of the rTS proteins and variants. Molar extinction coefficient values are reported assuming all disulfide bonds are reduced (Gasteiger, *et al.*, 2005).

Protein Name	Molecular Weight (g/mol)	ε (M^{-1}cm^{-1})
rTSβ	48219.61	79870
rTSγ	51112.93	79870
rTSγ-Xa	51254.14	79870
rTSγ-TEV	51339.30	81360
MBP-TEV-rTSβ	91079.98	147710

4.2.5 SITE-DIRECTED MUTAGENESIS

4.2.5.1 rTSγ-Xa

A pPSG-IBA03 plasmid (IBA Life Sciences, Göttingen, Germany) encoding wild-type rTSγ from *Homo sapiens* bearing a C-terminal Strep-tag® II was extracted from *E. coli* DH5α cells using a Qiagen miniprep kit (Qiagen, Toronto, ON) following the manufacturer's instructions. The plasmid DNA was used as the template DNA for site-directed mutagenesis (Hemsley, *et al.*, 1989). The primers used for mutagenesis contained a DNA insert encoding the Factor Xa protease amino acid recognition sequence IEGR and are shown in **Table 4.1** (Integrated DNA Technologies Coralville, IA). Phusion high fidelity polymerase was used for all amplification reactions according to manufacturer's instructions (New England Biolabs, Ipswich, MA). The thermocycle "PCR" conditions used were 98 °C for a 5-min initial denaturation, followed by 25 cycles of 30 s at 98 °C, 30 s at the annealing temperature, and 3 min at 72 °C for elongation. After the temperatures were cycled 25 times, a final extension was conducted at 72 °C and was followed by storage at 12 °C. The reaction product was treated with DpnI restriction enzyme overnight at 25 °C to remove wild-type methylated template DNA according to manufacturer's instructions (New England Biolabs, Ipswich, MA). PCR products were then separated by agarose gel electrophoresis (0.8% agarose, 100 V, 400 mA, 60 min) and subsequently isolated by using a Qiagen Gel Extraction kit (Qiagen, Toronto, ON) according to manufacturer's instructions. Isolated PCR products were used to transform chemically competent DH5α cells using the heat shock method (Froger & Hall, 2007) and grown on agar plates (1.5%) containing ampicillin (100 μg/mL) overnight. Transformants were cultured in 5-mL volumes of sterile LB media with

ampicillin (100 μg/mL) and DNA was extracted (miniprep, Qiagen, Toronto, ON). The

open reading frames were then sequenced commercially (Robarts Research Institute,

London, ON) using automated Sanger sequencing to verify that only selected mutations

to the *rTS* gene were introduced. Plasmids with the correct sequences were used to

transform *E. coli* BL21 Star™(DE3) (Thermo Fisher Scientific, Inc., Waltham, MA) cells

using the heat shock method (Froger & Hall, 2007). Transformants were cultured in LB

media (5-mL) containing ampicillin (100 μg/mL) and stored in 15% glycerol stocks at –

80 °C. These cells were then used for protein expression.

4.2.5.2 rTSγ-TEV

A pPSG-IBA03 plasmid (IBA Life Sciences, Göttingen, Germany) encoding

wild-type rTSγ from *Homo sapiens* bearing a C-terminal Strep-tag® II was extracted

from *E. coli* DH5α cells using a Qiagen miniprep kit (Qiagen, Toronto, ON) following

the manufacturer's instructions. The plasmid DNA was used as the template DNA for

site-directed mutagenesis (Hemsley, *et al.*, 1989). The primers used for mutagenesis

contained a DNA insert encoding the TEV protease amino acid recognition sequence

ENLYFQGS are shown in **Table 4.1** (Integrated DNA Technologies Coralville, IA).

Phusion high fidelity polymerase was used for all amplification reactions according to

manufacturer's instructions (New England Biolabs, Ipswich, MA). The thermocycle

"PCR" conditions used were 98 °C for a 5-min initial denaturation, followed by 25 cycles

of 30 s at 98 °C, 30 s at the annealing temperature, and 3 min at 72 °C for elongation.

After the temperatures were cycled 25 times, a final extension was conducted at 72 °C

and was followed by storage at 12 °C. The reaction product was treated with DpnI

restriction enzyme overnight at 25 °C to remove wild-type methylated template DNA

according to manufacturer's instructions (New England Biolabs, Ipswich, MA). PCR

products were then separated by agarose gel electrophoresis (0.8% agarose, 100 V, 400

mA, 60 min) and subsequently isolated by using a Qiagen Gel Extraction kit (Qiagen,

Toronto, ON) according to manufacturer's instructions. Isolated PCR products were used

to transform chemically competent DH5α cells using the heat shock method (Froger &

Hall, 2007) and grown on agar plates (1.5%) containing ampicillin (100 µg/mL)

overnight. Transformants were cultured in 5-mL volumes of sterile LB media with

ampicillin (100 µg/mL) and DNA was extracted (miniprep, Qiagen, Toronto, ON). The

open reading frames were then sequenced commercially (Robarts Research Institute,

London, ON) using automated Sanger sequencing to verify that only selected mutations

to the *rTS* gene were introduced. Plasmids with the correct sequences were used to

transform *E. coli* BL21 Star™(DE3) (Thermo Fisher Scientific, Waltham, MA) cells

using the heat shock method (Froger & Hall, 2007). Transformants were cultured in LB

media (5 mL) containing ampicillin (100 µg/mL) and stored in 15% glycerol stocks at –

80 °C. These cells were then used for protein expression.

4.2.5.1 MBP-TEV-rTSβ

A pMAL-p2x plasmid (New England Biolabs, Ipswich, MA) bearing the *malE*

gene that encodes for a characteristic N-terminal maltose binding protein (MBP) when an

exogenous gene is present was extracted from *E. coli* DH5α cells using a Qiagen

miniprep kit (Qiagen, Toronto, ON) according to manufacturer's instructions. The

extracted pMAL-p2x vector was used as the template DNA for amplification of the MBP

gene using primers listed in **Table 4.2**. The pPSG-IBA03 plasmid DNA encoding the

rTSγ-TEV variant was extracted from *E. coli* DH5α cells. The pPSG-IBA3 plasmid was

used as the template for amplification of a linear product that is primed for restriction

digestion and insertion of the *malE* gene. The resulting DNA construct would encode for

the N-terminal MBP fusion tag, a short linking region, followed by a TEV protease

recognition site, the rTSβ isoform, and a C-terminal Strep-tag® II (Integrated DNA

Technologies, Coralville, IA).

Forward and reverse primers contained restriction sites for NdeI and AvaI (New

England Biolabs, Ipswich, MA) restriction enzymes, respectively, and were used to

amplify regions of the pMAL-p2x and pPSG-IBA03 plasmids (**Figure 4.2**). Phusion high

fidelity polymerase was used for all amplification reactions according to manufacturer's

instructions (New England Biolabs, Ipswich, MA). The thermocycle "PCR" conditions

used were 98 °C for a 5-min initial denaturation, followed by 25 cycles of 30 s at 98 °C,

30 s at the annealing temperature, and 3 min at 72 °C for elongation. After the

temperatures were cycled 25 times, a final extension was conducted at 72 °C and was

followed by storage at 12 °C. The reaction product was treated with DpnI restriction

enzyme for 2 h at 37 °C to remove wild-type methylated template DNA according to

manufacturer's instructions. PCR products were then separated by agarose gel

electrophoresis (0.8% agarose, 100 V, 400 mA, 60 min) and the desired product was then

isolated by using a Qiagen Gel Extraction kit (Qiagen, Toronto, ON) according to

manufacturer's instructions. Extracted PCR product DNA of both the amplified genes

from the pPSG-IBA03 and pMAL-p2x plasmids was subjected to a double digest with

NdeI and AvaI restriction enzymes according to the manufacturer instructions (New

England Biolabs, Ipswich, MA). The resulting mixture was subjected to a DNA extraction using a Qiagen Gel Extraction kit (Qiagen, Toronto, ON) according to manufacturer's instructions. Finally, the extracted digestion products from the PCR reactions were used in an enzymatic ligation reaction using T4 DNA ligase and ATP (10 mM) as per manufacturer's instructions (Sigma-Aldrich Canada Ltd., Toronto, ON).

DNA ligation products were used to transform chemically competent DH5α cells using heat shock (Froger & Hall, 2007) and grown on agar plates (1.5%) containing ampicillin overnight (100 μg/mL). Transformants were cultured in 5-mL volumes of sterile LB media with ampicillin (100 μg/mL) and DNA was extracted (miniprep, Qiagen, Toronto, ON). The open reading frames were then sequenced commercially (Robarts Research Institute, London, ON) using automated Sanger sequencing to verify that only selected mutations to the *rTS* gene were introduced. Plasmids with the correct sequences were used to transform *E. coli* BL21 Star™(DE3) (Thermo Fisher Scientific, Inc., Waltham, MA) cells using the heat shock method (Froger & Hall, 2007). Transformants were cultured in LB media (5 mL) containing ampicillin (100 μg/mL) and stored in 15% glycerol stocks at –80 °C. These cells were then used for protein expression.

Figure 4.2 Simplified representation of the subcloning process for the MBP-TEV-rTSβ-Strep-II construct. The amplified *malE* gene, as well as the insertion region within the pPSG-IBA3 plasmid are shown.

Table 4.3 Primers for constructing ORFs encoding the rTS isoform variants. All primers are shown in the 5′ to 3′ direction. Altered DNA corresponding to the inserted protease recognition sites is shown in boldface and underlined (not applicable for the MBP-TEV-rTSβ primers).

Construct/Variant	Primer Name	DNA Sequence
rTSγ-Xa	Forward	5'- **GGCCGC**ATGCACACCGACCCGGACTAC-3'
	Reverse	5'-**TTCAAT**GTGACCACCCAGGGAGGTCGGG-3'
rTSγ-TEV	Forward	5'-**TATTTTCAGGGC**ATGCACACCGACCCGGACTAC-3'
	Reverse	5'-**CAGGTTTTCCAG**GGAGGTCGGGAAACGAACGTC-3'
MBP-TEV-rTSβ	pPSG-IBA3 Reverse	5'-GCTTAATTCATATGTATATCTCCTTCTTAAAGTTAAAC-3'
	pPSG-IBAS3 Forward	5'-CGCGTCTCGGGGAAAACCTGTATTTTCAGGGCATGCACACCG-3'
	pMAL-p2X Forward	5'-CGTTAACCCATATGAAAATCGAAGAAGGTAAACTGG-3'
	pMAL-p2X Reverse	5'-GAATTCCCGAGGTTGTTGTTATTGTTATTGTTGTTGTTGTTCG-3'

4.2.6 N-TERMINAL SEQUENCING OF PROTEASE CLEAVAGE PRODUCTS

Protein samples of rTSγ-Xa variant were subjected to Factor Xa protease cleavage under conditions described previously. After the cleavage reaction was complete, products were assessed using SDS-PAGE (10%) with Coomassie brilliant blue (R-250) staining. After SDS-PAGE was complete, samples were transferred to a polyvinylidene fluoride (PVDF) membrane through electroblotting at 300 mA in TAE buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA, at pH 8.0). N-Terminal sequencing samples of the cleavage reaction products were prepared and sent for analysis at SPARC BioCentre (SickKids, Toronto, ON).

4.2.7 DIFFERENTIAL SCANNING CALORIMETRY

A Malvern Panalytical (Microcal VP-DSC) Differential Scanning Calorimeter was used (Malvern Instruments Inc., Northhampton, VA). Scans were analyzed using Malvern *Origin* v. 7.0 software. All scans were conducted from 20-80 °C as a scan rate of 45 °C/h. The DSC was calibrated using assay buffer (50 mM Tris, 10 mM $MgCl_2$, pH 8.0) for a minimum of three scans prior to the protein scan. Protein scans were completed using a protein concentration of 7.3 μM in assay buffer. The corresponding thermograms were plotted after subtraction of the scanned baseline control and the thermal melting temperature T_m was obtained from the plot.

4.2.8 CD SPECTROSCOPY

A CD spectrum of rTSβ was obtained using a protein concentration of 100 μg/mL in 10 mM HEPES buffer at pH 7.5, containing 3.3 mM $MgSO_4$. This buffer composition

was chosen to minimize contributions of buffer components to the background absorbance at low wavelengths. A 0.1-cm quartz cuvette was used to collect spectra from 190-260 nm. Three buffer spectra were averaged and subtracted from the three averaged protein spectra to obtain a final CD spectrum. Protein concentrations were determined using UV spectrophotometry at 280 nm using extinction coefficients determined by the ProtParam web tool from ExPASy under the assumption that all cysteine residues were reduced (Gasteiger, *et al.*, 2005). The obtained CD spectroscopic data (mdeg) was converted to MRE using eqn. 3.2.

4.3 RESULTS AND DISCUSSION

4.3.1 PRODUCTION AND CHARACTERIZATION OF rTSγ

To initiate my research on the rTS proteins, I cloned the synthetic *rTSα*, *rTSβ*, and *rTSγ* ORFs into separate pPSG-IBA03 plasmids encoding a recombinant fusion protein bearing C-terminal Strep II affinity tag. A C-terminal affinity tag was chosen for application to the rTS proteins due to the identity in the C-terminal 341 amino acids, while the N-terminus is varied between isoforms.

High yields of soluble rTSγ were obtained by expression in *E. coli* followed by purification using a StrepTactin XT (IBA Life Sciences, Göettingen, Germany) affinity column (**Figure 4.3**) . Previously, Gerlt and colleagues had identified rTSγ as a member of the MR subgroup via structural and sequential homology and functionally determined it to be active as a FucD (Wichelecki, *et al.*, 2014). Of course, the β and γ isoforms were of specific experimental interest based on a literature precedent that classified rTSγ as a L-fuconate dehydratase. The same work found rTSβ to be catalytically inactive as an
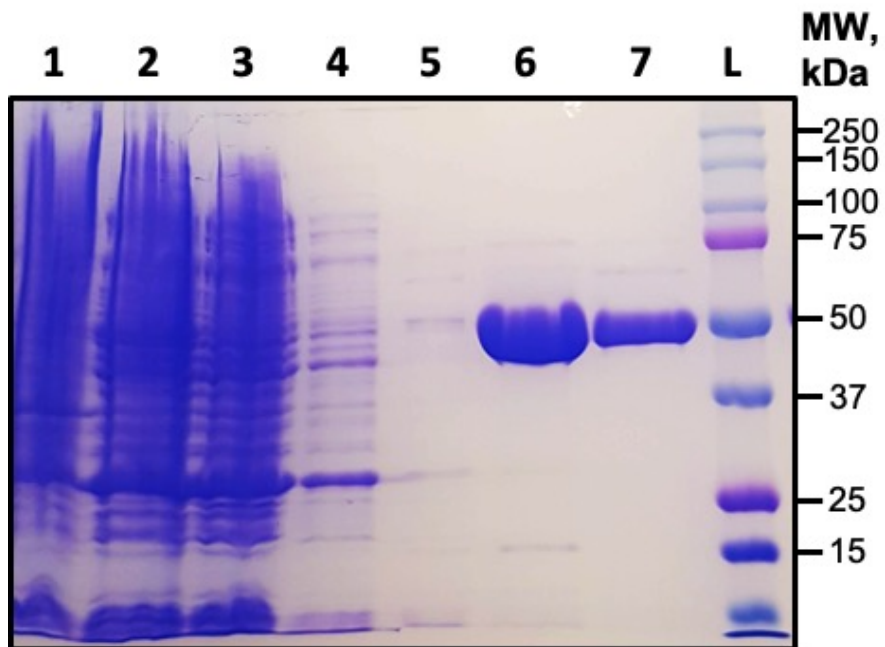
Figure 4.3. Representative SDS-PAGE electrophoretogram of rTSγ. Shown is a 10% SDS polyacrylamide gel. The lanes of the gel correspond to ladder = L, cell pellet = 1, clarified lysate = 2, column flow through = 3, column wash = 4, elution fraction 1 = 5, elution fraction 2 = 6, elution fraction 3 = 7.

ENS member enzyme, while being a clinically relevant protein (Wichelecki, *et al.*, 2014).

4.3.2 PRODUCTION AND CHARACTERIZATION OF rTSβ

Because previous researchers had shown rTSβ to be unstable and poorly expressed (Wichelecki, *et al.*, 2014), I investigated alternative methods for producing the protein. rTSβ lacks stability relative to the longer rTSγ isoform, yet the isoforms only differ by the 27 N-terminal amino acids as a result of alternative splicing of mRNA. To assess the stability of rTS proteins, variants of rTSγ lacking the N-terminal 5, 8, and 12 residues were generated by Gerlt and colleagues. Each of the variant proteins were insoluble, which displays the instability of rTSβ (Wichelecki, *et al.*, 2014).

In the early stages of experimentation on rTSβ, it became evident that the protein is highly unstable with the expression system used. Production of rTSβ was very low relative to rTSγ, which is indicated by the lack of protein observed at 51 kDa (**Figure 4.4**). These findings corroborated those of Gerlt and co-workers (Dolnick, *et al.*, 2003; Wichelecki, *et al.*, 2014) in which, even when soluble protein was produced, only low yields of soluble rTSβ were obtained under non-denaturing conditions. Other *in vitro* studies involved production of rTSβ in the presence of urea (6 M) in the purification process, which denatures and solubilizes protein. These studies utilized a protein refolding process that entailed a gradual removal of denaturant in dialysis but have historically lacked characterization of properly renatured protein (Dolnick, *et al.*, 2003; Dolnick, *et al.*, 2005). The failure to obtain adequate amounts of rTSβ obviated the possibility of conducting activity measurements for FucD activity.
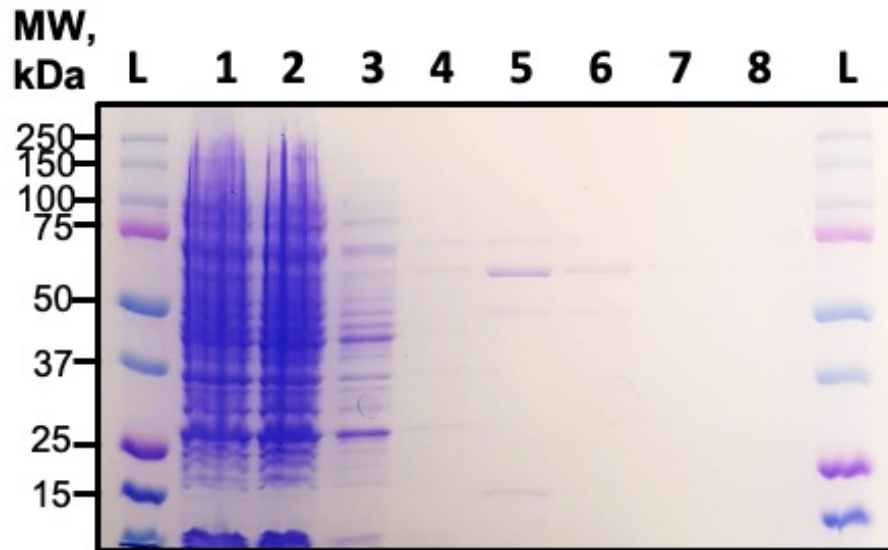
Figure 4.4. Representative SDS-PAGE electrophoretogram monitoring the purification of rTSβ under non-denaturing conditions. Shown is a 10% SDS polyacrylamide gel. The lanes of the gel correspond to ladder = L, cell pellet = 1, clarified lysate = 2, column flow through = 3, column wash = 4, elution fraction 1 = 5, elution fraction 2 = 6, elution fraction 3 = 7, elution fraction 4 = 8. It should be noted that little to no soluble protein was observed at the expected molecular weight (51 kDa) in column elution lanes 5-8.

Considering these results and those of the earlier work by Gerlt and colleagues, efforts to produce soluble rTSβ required alternative methods. The identity between rTSβ and rTSγ offers a unique experimental situation for obtaining soluble rTSβ. Since soluble rTSγ could be obtained in abundance, we decided to use this as a protein scaffold for obtaining pure soluble rTSβ. Hence, we altered the *rTSγ* ORF using site-directed mutagenesis to encode a site-specific protease recognition sequence that would allow rTSβ to be obtained post-translationally by proteolytic cleavage of the corresponding variant of rTSγ. The gene alteration involved engineering a short DNA sequence encoding the protease recognition site into the *rTSγ* ORF, which was inserted immediately preceding the 5′-end of *rTSβ* in the pPSG-IBA03 vector.

Proteases are commonly used for post-translational cleavage and removal of fusion protein domains in experimental studies. Fusion protein constructs often include domains that afford solubility or stability, affinity tags, and/or signaling sequences for protein localization (Jenny, Mann, & Lundblad, 2003). An ideal protease for the cleavage of rTSγ to obtain rTSβ would be highly specific, have mild cleavage reaction conditions, and be cost effective in its availability. It was also imperative for the protease to cleave its recognition site on the C-terminal side such that no additional amino acids remained appended to rTSβ post-cleavage. A list of typical proteases is given in **Table 4.1**. Of the common options, Factor Xa protease offered many of the aforementioned advantages in its use, which made it a promising candidate for cleavage of the rTSγ variant.

The protease recognition sequence for Factor Xa consists of amino acids Ile-Glu/Asp-Gly-Arg| (I-E/D-G-R|) and cleavage occurs after the arginine. Factor Xa is

commercially available and has a desirable protease to protein ratio for experimental use (approximately 1:50).

4.3.3 GENERATION, PRODUCTION, AND CHARACTERIZATION OF rTSγ-Xa

The rTSγ variant with an encoded Factor Xa recognition site (rTSγ-Xa) was obtained via site-directed mutagenesis. PCR primers were designed such that a DNA insert that encodes for the Factor Xa recognition site was placed within the *rTSγ* ORF at a position that would allow post-translational cleavage of variant rTSγ to rTSβ. Once the rTSγ-Xa variant was produced, the cleavage reaction conditions were optimized. The expected products for Factor Xa-catalyzed cleavage of rTSγ-Xa included the rTSβ-StrepII construct (48 kDa), the dissociated monomers of Factor Xa protease (27 and 16 kDa), and any remaining uncleaved rTSγ-Xa construct (51 kDa).

As shown in **Figure 4.5**, Factor Xa protease was relatively efficient at catalyzing the cleavage of rTSγ-Xa at the specified recognition site. Products of the cleavage reaction were observed to be rTSβ (48 kDa), and two alternative products at 40 and 39 kDa, respectively. Uncleaved rTSγ-Xa (51 kDa) was also observed in the mixture. Cleavage time trial results indicated that Factor Xa-catalyzed cleavage was most efficient after a period of 4 h at 25 °C (**Figure 4.5**). No change in cleavage efficiency was observed when the ratio of protease to fusion protein was increased to 1:25 and 1:10. Hence, the optimal protease to protein ratio was left at 1:50.

Significant amounts of secondary products with masses of 40 and 39 kDa were observed in the product mixture. These likely arose from non-specific cleavage activity by the Factor Xa protease. Non-specific cleavage of fusion proteins is a common phenomenon that is detrimental to experimental use of Factor Xa protease (Jenny, Mann,
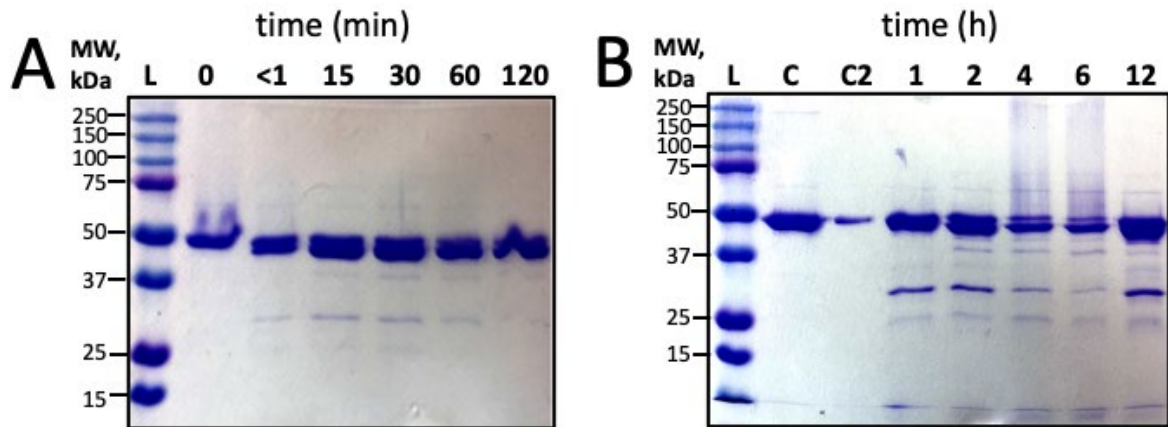
Figure 4.5. Representative SDS-PAGE electrophoretograms monitoring short and long scale time trials of Factor Xa-catalyzed cleavage of rTSγ-Xa. Shown in both panels are 10% SDS polyacrylamide gels. (A) The lanes of the gel correspond to Factor Xa-catalyzed cleavage of rTSγ-Xa and the time of cleavage is indicated above each lane in minutes. The molecular weight ladder is labelled "L." (B) The lanes of the gel correspond to Factor Xa-catalyzed cleavage of rTSγ-Xa and the time of cleavage is indicated above each lane in hours. The molecular weight ladder is labelled "L," rTSγ control is labelled "C," and rTSγ-Xa uncleaved control is labelled "C2." All cleavage reactions were conducted in a 1:50 ratio of protease to fusion protein.

& Lundblad, 2003). For Factor Xa, secondary cleavage sites often precede basic residues, most commonly arginine (Jenny, Mann, & Lundblad, 2003).

Unfortunately, the formation of non-specific cleavage products as well as uncleaved rTSγ-Xa construct presents a problem for purification, especially since the desired product, the uncleaved construct, and several of the non-specific cleavage products all bear the C-terminal Strep-II tag. To obtain a better understanding of the nature of the non-specific cleavage products we sought to identify them using N-terminal amino acid sequencing by Edman degradation.

The products visualized by SDS-PAGE (**Figure 4.6**) were transferred to a PVDF membrane using electroblotting. The bands containing non-specific cleavage products were excised from the PVDF membrane and subjected to commercial Edman degradation.

The first cleavage product identified by N-terminal sequencing was the gel band corresponding to a molecular weight of 40 kDa on the electrophoretogram (**Figure 4.6**). N-Terminal sequencing results indicated several possibile amino acid identities at each amino acid position, but the most likely sequence of the first product was I-G-P-E-K. This amino acid sequence appears twice in the amino acid sequence of rTSβ. The N-terminal-most occurrence of the I-G-P-E-K sequence would yield a cleavage product of ~40 kDa, which aligns with the electrophoretic analysis. Cleavage at the other IGPEK site within the protein sequence would yield products with molecular weights of 26.8 kDa and 24.9 kDa, which were not observed using SDS-PAGE analysis.

The second cleavage product identified via N-terminal sequencing was the gel band corresponding to a molecular weight of 39 kDa on the electrophoretogram (**Figure**
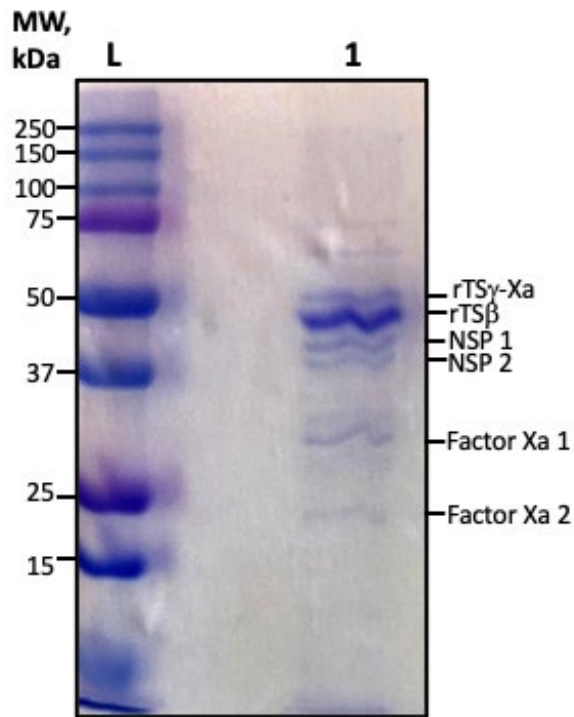
Figure 4.6. Representative SDS-PAGE electrophoretogram showing optimized Factor Xa-catalyzed cleavage of rTSγ-Xa. Shown is a 10% SDS polyacrylamide gel. The lanes of the gel correspond to ladder = L, 1 = cleaved rTSγ-Xa reaction mixture. Reaction products are labelled within the figure. Non-specific cleavage products are labelled "NSP1" and "NSP2," and the monomeric subunits of Factor Xa are labelled "Factor Xa 1" and "Factor Xa 2." This electrophoretogram highlights non-specific cleavage products at 39 and 40 kDa, as well as uncleaved rTSγ-Xa construct at 51 kDa.

4.6). The N-terminal amino acid sequence that was identified was M-H-T-D-P-D, which was unambiguous in the analysis. The sequence M-H-T-D-P-D is the N-terminal sequence of rTSβ indicating that Factor Xa-catalyzed cleavage occurred at both the IEGR primary recognition site and an unknown location at the C-terminus of the protein. This is supported by the observed purification of the mixture of rTSβ and uncleaved rTSγ-Xa from the non-specific cleavage products (**Figure 4.7**). Thus, only one cleavage site was obtained by N-terminal sequencing.

The Factor Xa-catalyzed cleavage reaction products were subjected to purification via use of a StrepTactin® XT chromatography resin. Since this resin has an affinity for proteins with a Strep tag, the Factor Xa protease should not be retained by the resin, permitting purification of the products that were derived from rTSγ-Xa remaining in the mixture. As shown in **Figure 4.7**, the StrepTactin® XT-purified mixture contained rTSγ-Xa and rTSβ. Indeed, purification of primary cleavage products from non-specific cleavage products may be possible using StrepTactin XT resin upon further experimentation. However, separation of rTSγ-Xa from rTSβ in the reaction mixture is essential for Factor-Xa catalyzed cleavage of rTSγ-Xa to be utilized effectively, yet is a difficulty within our current system.

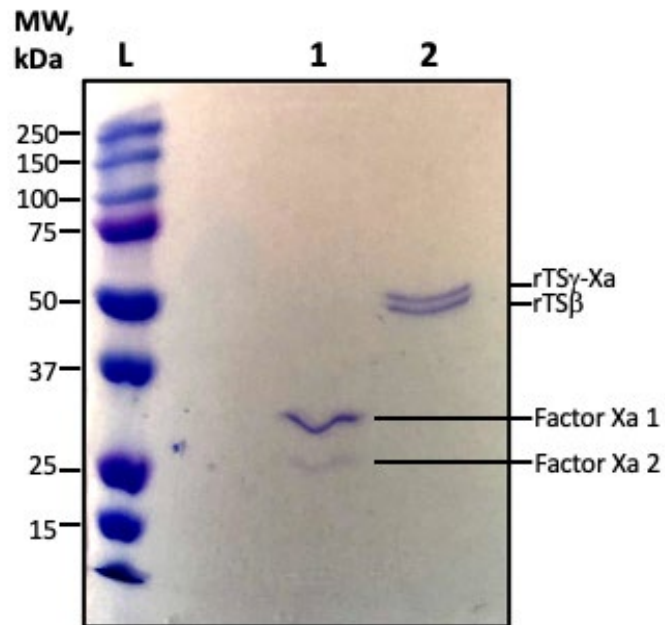**Figure 4.7. Representative SDS-PAGE electrophoretogram showing StrepTactinXT purification of the reaction mixture from Factor Xa-catalyzed cleavage of rTSγ-Xa.** Shown is a 10% SDS polyacrylamide gel. The lanes of the gel correspond to ladder = L, 1 = column flow through, 2 = column elution.

4.3.4 GENERATION, PRODUCTION, AND CHARACTERIZATION OF rTSγ-TEV

Indeed, the use of Factor Xa protease for cleavage of the rTSγ-Xa variant appeared to suffer from a major problem: the promiscuous nature of its protease activity. Consequently, we turned our attention to tobacco etch virus (TEV) protease. TEV protease is not only highly robust in its cleavage specificity and efficiency, it is easily produced as a recombinant fusion protein construct in *E. coli* cells (Cesaretto, Burrone, & Petris, 2016). TEV protease has a highly specific recognition sequence-ENLYFQ|G/S. Although, the cleavage product to the C-terminus of the recognition sequence will retain the glycine or serine residue required for cleavage, this was deemed not to be a major impediment to employing a TEV-protease-based approach.

A variant of rTSγ (rTSγ-TEV) containing a TEV protease recognition sequence was generated using site-directed mutagenesis in an analogous method to the production of rTSγ-Xa. A sequence encoding for the TEV protease-specific recognition sequence was inserted at a position within the *rTSγ* ORF such that proteolytic cleavage would result in an rTSβ product with an N-terminal Gly or Ser. rTSγ-TEV production levels were similar to rTSγ-Xa in *E. coli* BL21 (DE3) cells.

As shown in **Figure 4.8**, the TEV protease-catalyzed cleavage of the fusion construct was highly inefficient despite increasing the TEV protease concentration, the temperature, or the time of reaction. This observation suggested that the TEV protease was being physically occluded from its recognition site on the recombinant rTSγ-TEV.
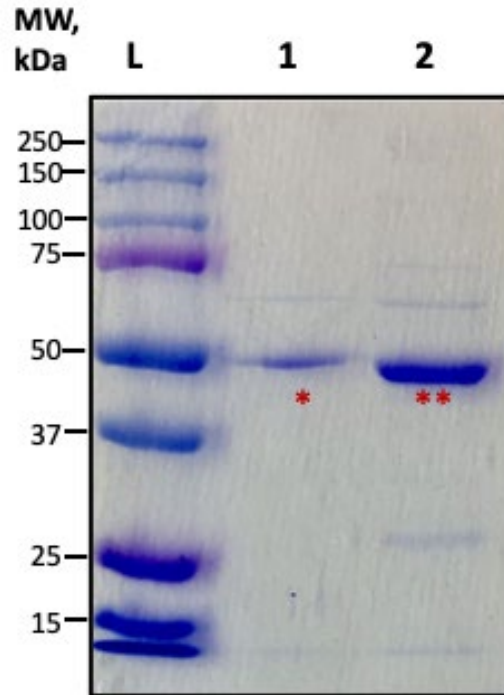
Figure 4.8. Representative SDS-PAGE electrophoretogram showing TEV protease-catalyzed cleavage of rTSγ-TEV. Shown is a 10% SDS polyacrylamide gel. The lanes of the gel correspond to ladder = L, 1 = rTSγ-TEV uncleaved control, 2 = 30 min of cleavage. TEV protease can be seen at 27 kDa in lane 2, while * and ** indicate the rTSγ-TEV uncleaved control and rTSγ-TEV after 30 min TEV protease-catalyzed cleavage, respectively. Slight differences in molecular weight can be attributed to concentration differences between the sample lanes, not TEV protease-catalyzed cleavage. The apparent increase in the mass of the protein in the major band in lane two arose from increased protein loading relative to lane 1, not from actual changes in the protein mass.

Assuming that physical occlusion of TEV protease from its recognition site on rTSγ-TEV was occurring, we conducted the protease digestion in the presence of semi-denaturing concentrations of urea (1-4 M) (**Figure 4.9**). TEV protease is known to retain its activity in a variety of buffering systems, as well as in the presence of up to 2 M urea (Sun, *et al.*, 2012). Experimental optimization of the reaction catalyzed by TEV protease with rTSγ-TEV under varying concentrations of urea is shown in **Figure 4.9**. In the presence of urea (2 M), cleavage of rTSγ-TEV was optimal. The reaction products were visualized using SDS-PAGE, which revealed prominent levels of protein corresponding to the full construct rTSγ-TEV (51 kDa), rTSβ (48 kDa), and TEV protease (26 kDa). Notably, no non-specific cleavage of rTSγ-TEV was observed, which was an improvement over the Factor Xa-catalyzed cleavage of rTSγ-Xa. The finding of improved cleavage in the presence of urea supported the notion that rTSγ-TEV was physically excluding TEV from its cleavage recognition site.

It is evident that TEV protease is superior to Factor Xa in the rTSβ purification system, with only occlusion from the recognition site reducing cleavage activity. To overcome the physical barrier precluding TEV protease activity, alternative rTSγ protein constructs were explored. Primarily, the DNA that encodes for the aforementioned 27 N-terminal amino acid region that distinguishes rTSγ from rTSβ was investigated as a scaffold for manipulation to improve the purification system. To aid in the purification of protease cleavage products, we opted to include an N-terminal tag with affinity properties that would permit separation of cleavage product rTSβ from any uncleaved construct, under the assumption of the presence of non-negligible amount of uncleaved construct. We also anticipated that insertion of an N-terminal tag that differs from the native rTSγ
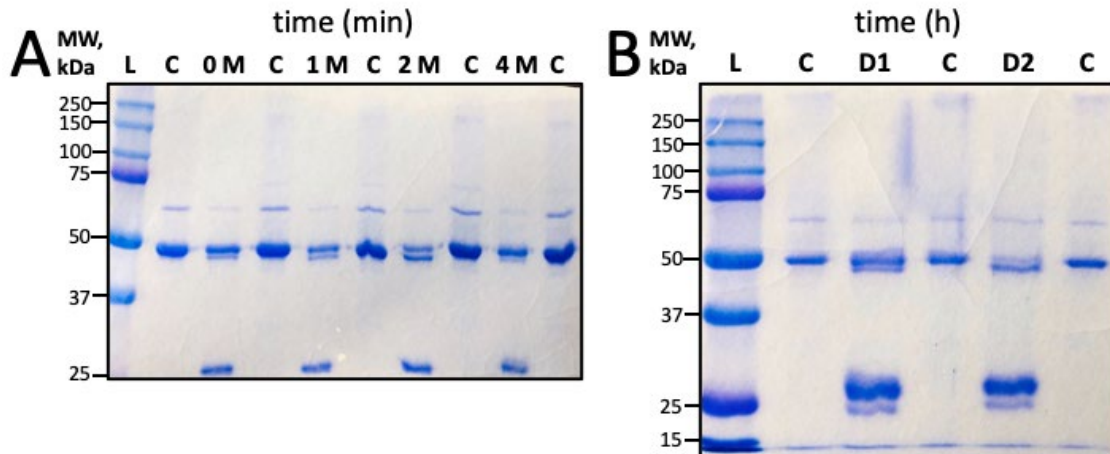
Figure 4.9. Representative SDS-PAGE electrophoretograms monitoring TEV protease-catalyzed cleavage of rTSγ-TEV under varied conditions. Shown are 10% SDS polyacrylamide gels. (A) The lanes of the gel correspond to ladder = L, C = rTSγ-TEV uncleaved control, and the remaining lanes show TEV protease-catalyzed cleavage of rTSγ-TEV in the presence of the indicated concentration of urea. All cleavage reactions took place over 8 h. (B) The electrophoretogram follows the reaction progression of TEV protease on rTSγ-TEV over 15 h with multiple doses of TEV protease (doses were of 1 mg/mL TEV). The lanes of the gel correspond to ladder = L, C = rTSγ-TEV uncleaved control, D1 = dose 1 of TEV with 5 h cleavage, D2 = dose 2 of TEV with an additional 10 h of cleavage. The minor band observed at ~60 kDa in all lanes is likely heat shock protein 60 (Hsp60), a background protein produced in *E. coli.*

N-terminal region would likely expose the protease recognition site to TEV protease, especially if the fusion tag was bulkier. Finally, an affinity tag that increased the solubility of the fusion protein was desired (Lichty, *et al.*, 2005).

Consequently, we chose to employ a maltose binding protein (MBP) "tag" fused to the N-terminus of the rTSγ-TEV construct. MBP is highly soluble and has been shown to exhibit intrinsic chaperone activity towards its associated fusion protein (Costa, *et al.* 2014). MBP is more effective when used as a N-terminal tag as opposed to a C-terminal tag since it promotes proper folding of its fusion partner by binding to the C-terminus and occluding its self-association (Sachdev & Chirgwin, 2000). The passive role that MBP plays in promoting folding of its fusion partner is correlated with a large hydrophobic surface on the protein (Costa, *et al.*, 2014). Above all, using MBP as an N-terminal tag offers a unique upside in its affinity interaction with amylose, which is commercially available in the form of a chromatographic resin.

4.3.5 GENERATION, PRODUCTION, AND CHARACTERIZATION OF MBP-TEV-rTSβ

In order to optimize assembly of a fusion protein construct for rTSβ purification, the DNA sequence encoding an N-terminal MBP affinity tag was subcloned into the pPSG-IBA03 plasmid encoding rTSγ-TEV. The gene encoding MBP (*malE*) was obtained from a pMAL-p2x vector by PCR amplification (New England Biolabs, Ipswich, MA). Primers were designed such that recognition sequences for the AvaI and NdeI restriction enzymes were inserted onto the amplified DNA products to be used for subcloning into the pPSG-IBA03 vector. A second set of primers were designed such that

restriction sites for both AvaI and NdeI were inserted into the region of DNA that encodes the 27 amino acids absent in rTSβ, effectively using the region as a multiple cloning site for *malE*. A visual depiction of the subcloning process is shown in **Figure 4.2**.

As anticipated, the newly generated construct gave much improved TEV protease-catalyzed cleavage as shown in **Figure 4.10**, indicating that TEV protease can access its recognition site. However, it was also evident that production of MBP-rTSγ-TEV protein was markedly reduced, likely due to the large size of the MBP fusion tag (42 kDa) resulting in the MBP-rTSγ-TEV protein having an overall molecular weight of 93 kDa. Proteins fused to tags of this size are not always produced well in *E. coli* BL21 Star™(DE3) cells (Costa, *et al.*, 2013). This fact paired with the known instability of the rTSβ fusion partner contributes to the low yields for this protein construct. That said, the purification system designed for MBP-rTSγ-TEV yielded the highest purity of rTSβ (**Figure 4.10**). To counter the low yields of protein, bacterial cultures were grown in "terrific broth" (TB) media, an enriched media for cultivation of bacteria, which permitted a "brute force" method for obtaining protein. By using large scale expression and enriched media, yields were suitable for purification purposes.

The MBP-TEV-rTSβ protein was first purified using StrepTactin XT column chromatography. TEV protease-catalyzed cleavage of the purified protein was then conducted over an extended period of time, using a protease to protein ratio in excess of typical amounts to ensure complete cleavage of the MBP-TEV-rTSβ protein. Cleavage of the full construct was far more efficient than TEV protease-catalyzed cleavage of rTSγ-TEV but a significant amount of uncleaved construct still remained (**Figure 4.10**).
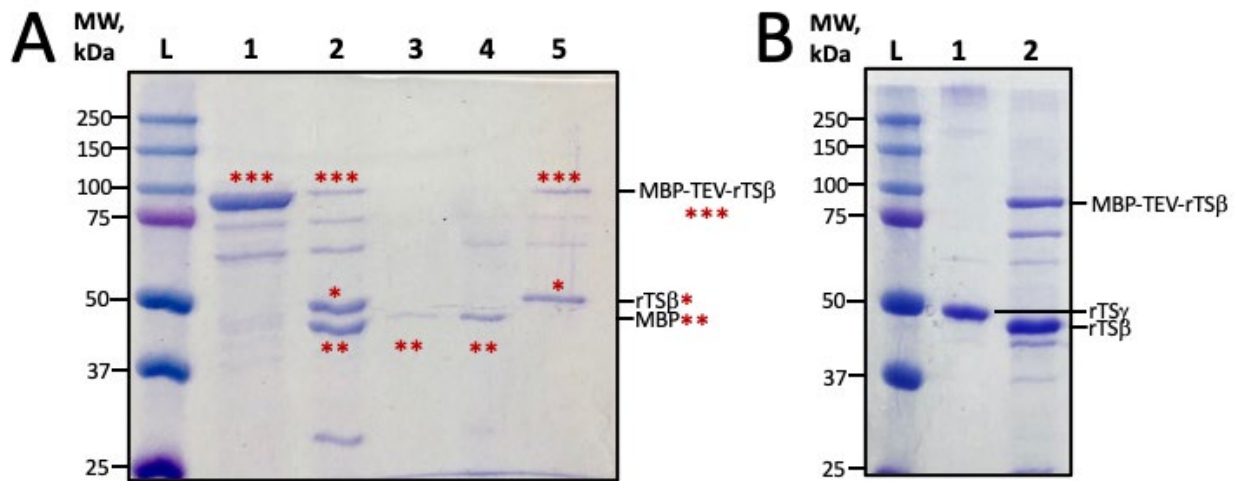
Figure 4.10. Representative SDS-PAGE electrophoretograms showing the purification of MBP-TEV-rTSβ. Shown are 10% SDS polyacrylamide gels. (A) The lanes of the gel correspond to ladder = L, 1 = uncleaved MBP-TEV-rTSβ, 2 = TEV protease cleaved mixture, 3 = amylose column elution, 4 = StrepTactinXT column flow through, 5 = StrepTactinXT column elution. (*) indicates the presence of rTSβ, (**) indicates the presence of MBP, and (***) indicates the presence of MBP-TEV-rTSβ. (B) The lanes of the gel correspond to ladder = L, 1 = rTSγ control, 2 = Semipure rTSβ from MBP-TEV-rTSβ construct purification. The flow-through from the amylose column chromatography step in (A) is not shown, but contained similar amounts of both MBP-TEV-rTSβ and rTSβ to lane 5.

159

Although the affinity of the N-terminal MBP tag for amylose chromatography resin could permit the separation of uncleaved protein from rTSβ in theory, attempts to purify the cleavage product rTSβ from the full MBP-TEV-rTSβ construct, TEV protease, and cleaved MBP on amylose resin were largely unsuccessful. This is evident by the large amount of MBP-TEV-rTSβ present in the StrepTactinXT elution (**Figure 4.10**).

In all the approaches utilizing protease-catalyzed cleavage of a soluble rTSγ-construct to yield rTSβ, the resulting mixture of proteins arising from the proteolytic cleavage reactions remained a problem for purification. In the case of the MBP construct, the hinderance to purification of rTSβ in this case likely arises from the oligomeric state of rTSβ in solution. The oligomeric state of MR subgroup members is tetrameric or octameric in solution (Fee, Hegeman, & Kenyon, 1974; Tsou, *et al.*, 1989; Sagy & Bearne, unpublished). Thus, if the reaction mixture contains both rTSβ and MBP-rTSγ-TEV, the formation of higher order structures (**Figure 4.11**) could result in affinity interactions that would hinder purification. Thus, the use of the MBP construct suffers from purification problems, as well as expression of the relatively large MBP-rTSγ-TEV fusion construct. Indeed, oligomeric structures could be disrupted by altering the buffering conditions (i.e., salt concentrations or pH), though purification may only be inhibited by these interactions in part.
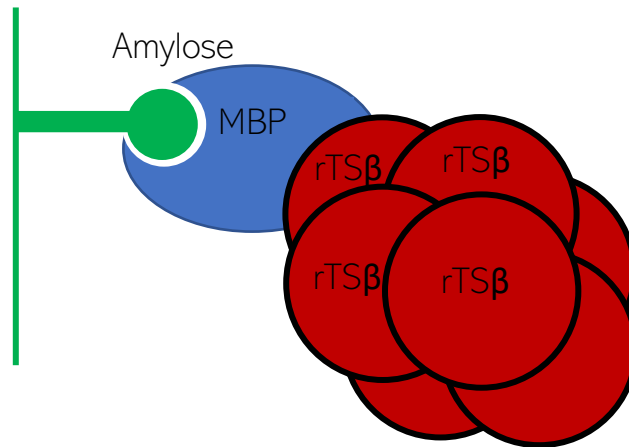
Figure 4.11. Schematic representation of the oligomeric state of MBP-TEV-rTSβ that could inhibit purification of rTSβ on amylose. Based on the tetrameric or octameric state of MR subgroup proteins in solution, if even one oligomer retained the MBP tag, purification of the cleavage products on the amylose column would be hindered.

4.3.6 DENATURING PURIFICATION AND REFOLDING STUDIES ON rTSβ

Methods of protein purification under protein-denaturing conditions are not as preferred as native methods. The major downfall to using denaturing methods of protein purification is in the steps involving refolding of the isolated protein. It is extremely difficult to guarantee that refolding has been fully achieved or that it is correctly folded in a fashion identical to the protein *in vivo*. In the case of rTSβ, however, a denaturing method of purification was worth exploring. Work done by Nicole Easton of the Bearne lab demonstrated the ability to obtain low levels of pure rTSβ by use of high concentrations of urea (6 M) to solubilize the protein as shown in **Figure 4.12A**. To refold the protein, protein was dialyzed against buffers containing incrementally lower concentrations of urea until 0 M urea was obtained. Structural characterization of the protein by CD spectroscopy indicated that folding had occurred to some extent (**Figure 4.12B**).

Because of the low yield of rTSβ, the volume of bacterial culture was scaled up to 4L using "terrific broth" culturing media. Although this approach produced a higher yield of protein, upon sequential dialysis to achieve non-denaturing buffering conditions, the protein aggregated and precipitated. However, addition of 10% glycerol stabilized the protein in otherwise normal buffering conditions. DSC analysis (**Figure 4.13**) afforded a DSC thermogram that did not exhibit the shape characteristic for protein unfolding. Instead, a wide thermal melting event was observed, occurring over a broad temperature range. This likely arises from an aggregation of protein during the refolding process, which yields a signal arising from a multitude of misfolded states of the protein. Overall,

this method of rTSβ purification shows promise, but has significant limitations in the refolding process.

There is a literature precedent for an experimental method of protein refolding from denaturing conditions using the non-detergent sulfobetaine family of compounds (Vuillard, Rabilloud, & Goldber, 1998). The use of NDSB 201 (3-(1-pyridinio)-1-propanesulfonate) is reported in refolding experiments for many proteins, inluding rTSβ (Dolnick, *et al.*, 2003; Dolnick, *et al.*, 2005). NDSBs are known stabilizing and solubilizing agents that act in the early stages of the protein folding process and inhibit abortive interactions that can cause aggregation and precipitation of a protein (Vuillard, Rabilloud, & Goldber, 1998). The methodologies described by Dolnick and colleagues in 2003 and 2005 for rTSβ were applied to the protein construct detailed in this work. Unfortunately, rTSβ aggregated and precipitated immediately in the presence of the NDSB 201-containing buffer. The aggregation was unlikely to have occurred as a result of added NDSB 201, but instead was likely due to the instability of rTSβ. Further investigation into the use of NDSB compounds with recombinant rTSβ may still be warranted.
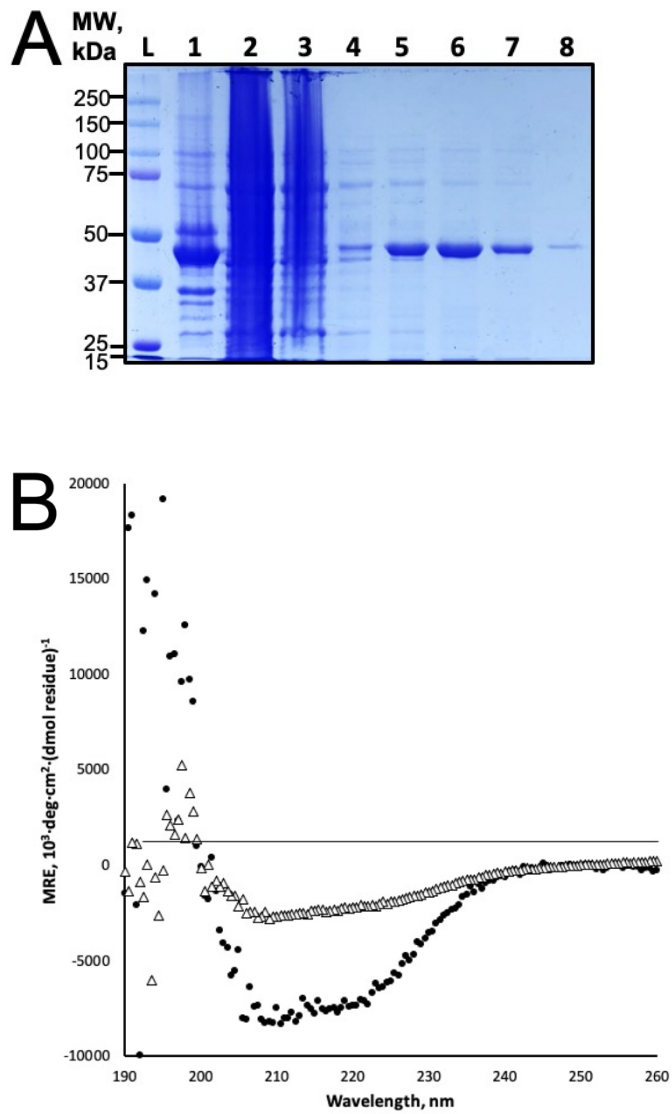
Figure 4.12. Evaluation of the purification of rTSβ using denaturing conditions. (A) Shown is a representative 10% SDS polyacrylamide gel showing the purification of rTSβ. The lanes of the gel correspond to ladder = L, 1 = cell pellet, 2 = clarified cell lysate, 3 = StrepTactinXT column flow through, 4 = column wash, 5 = elution fraction 1, 6 = elution fraction 2, 7 = elution fraction 3, 8 = column regeneration. (B) Shown is a CD spectrum of rTSβ (△) in comparison to that of wild-type *Xc*FucD (●) measured from 220 nm to 190 nm. The protein concentration was 100 μg/mL.
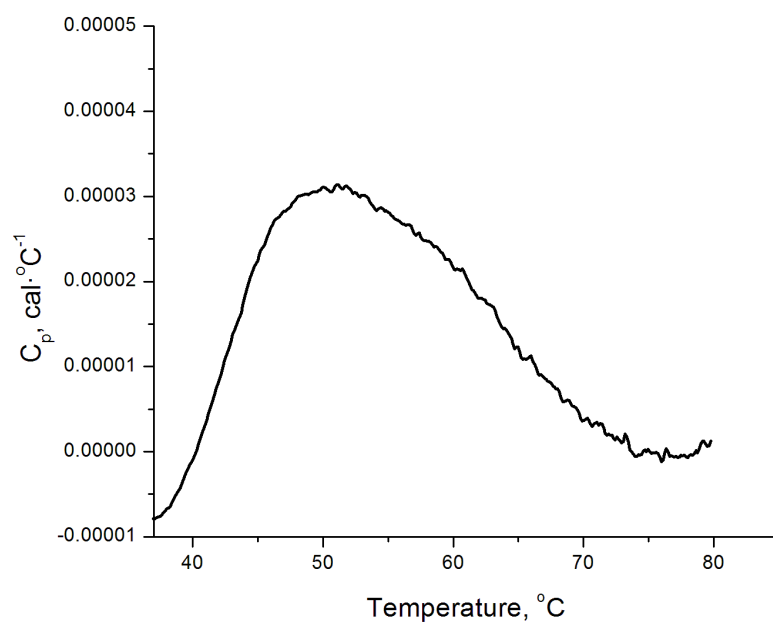
Figure 4.13. Representative DSC thermogram monitoring the thermal denaturation of refolded rTSβ. The thermogram was taken over a temperature range of 20 °C to 80 °C at a scan rate of 45 °C/h. Observed is a large peak indicating a thermal denaturation event occurring over a broad temperature range. The protein concentration was 7.3 μM. A selected region of the thermogram is shown to highlight the thermal denaturation.

4.3.7 PURIFICATION OF rTSβ FROM INCLUSION BODIES

The final approach I explored for the purification of rTSβ followed a protocol typically used for proteins that are expressed insolubly in inclusion bodies within the expression host cells. The cellular stress resulting from a high concentration of inducer and expression under strong promotors often forces recombinant proteins to be localized into inclusion bodies, especially if the protein is inherently unstable (Singh, *et al.*, 2015). The process of recovering protein from inclusion bodies in *E. coli* is similar to that of a standard protein purification under denaturing conditions. Insoluble inclusion bodies are isolated from soluble cell components during clarification of cell lysate. It is typical to incorporate a detergent wash step into the protocol to disrupt membranes within the cell pellet. In experimentation, rTSβ was successfully isolated from inclusion bodies (**Figure 4.14**). However, the use of inclusion body isolation did not show any specific advantage over the purification of rTSβ under denaturing conditions.
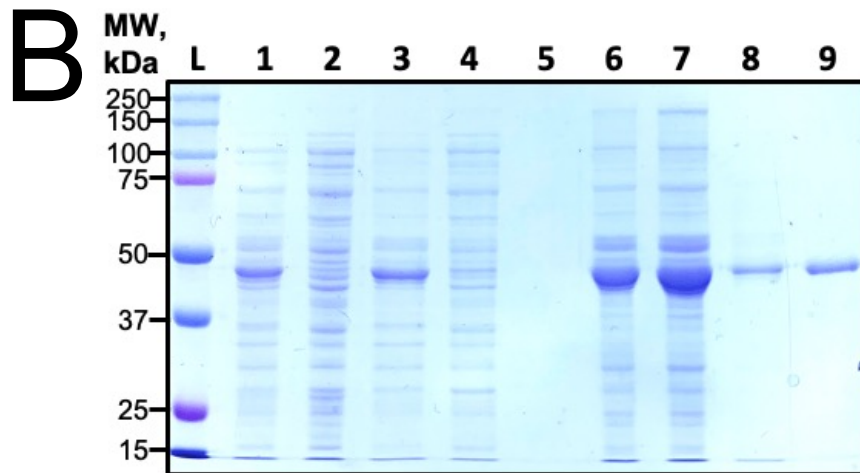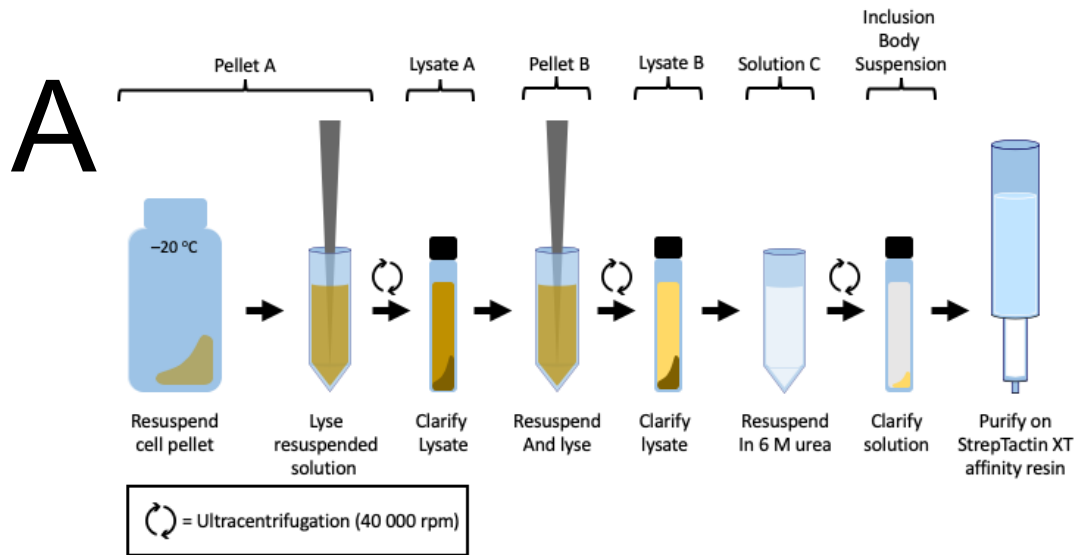
Figure 4.14. Representative SDS-PAGE electrophoretogram showing the purification of rTSβ by isolation from inclusion bodies. (A) The workflow for purification of rTSβ by isolation of inclusion bodies. (B) Shown is a 10% SDS polyacrylamide gel. The lanes of the gel correspond to ladder = L, 1 = cell pellet A, 2 = clarified lysate A, 3 = pellet B, 4 = clarified lysate B, 5 = resuspended solution C, 6 = inclusion body suspension, 7 = StrepTactinXT column flow-through, 8 = column wash, 9 = column elution (where A refers to the cell resuspension, B refers to the first pellet wash, and C refers to the second pellet wash).

4.4 CONCLUSIONS AND FUTURE WORK

In consideration of the findings of this work, it is evident that rTSβ is a difficult protein to express and purify for *in vitro* experimentation. A number of different methods were examined in this work for obtaining recombinant rTSβ from *E. coli*, but limitations accompanied each method. The methods that showed the most promise were those involving purification under denaturing conditions, given that the yield of pure protein was the greatest by far. Future work on the purification and analysis of rTSβ should focus on optimization of expression host and cell strain if use of *E. coli* is desired. Further, if denaturing conditions are used for purification, careful design of protein refolding conditions are required. Analysis of protein folding in the case of rTSβ is especially difficult considering the lack of a known functional assayn. Protein refolding is a topic at the forefront of protein science research, and new methods and protocols are developed at a high rate, giving rise to potential for a viable method for obtaining soluble rTSβ. Based on the classification of the rTS proteins as members of the MR subgroup, it is plausible that rTSβ may play a functional role in the oligomerization state of rTSγ *in vivo*, and should be assessed in future study.

Refolding assays have garnered attention in the field of modern protein science to solubilize proteins of interest that are otherwise insoluble when produced. These assays commonly involve the use of dyes to monitor thermal denaturation of proteins by fluorescence detection, which enables determination of protein stability using real-time PCR (Huynh & Partch, 2015). The dye molecules fluoresce with a high quantum yield when bound to a hydrophobic region that becomes exposed during denaturation. Common dyes include 1,8-ANS (1-anilinonaphthalene-8-sulfonate), 2,6-TNS

(naphthalene-6-sulfonic acid), and SYPRO orange (Huynh & Partch, 2015). SYPRO orange has fluorescence properties that are most compatible with common real-time PCR instruments. Therefore, thermal denaturation assays using this method may permit determination of protein refolding conditions by fluorescence output. This system should be applied to pure denatured rTSβ in future experimentation, allowing observation of a high number of buffering conditions for optimization of protein refolding and stability.

It was observed through experimental optimization of both the MBP-rTSβ-TEV fusion construct and attempts to purify rTSβ in non-denaturing conditions that pure rTSβ is highly unstable and results in insoluble aggregates in standard buffering conditions. Previous hypotheses partially attribute the instability of rTSβ *in vitro* to the absence of high order mammalian protein expression mechanisms, especially post-translational modifications (Wichelecki, *et al.*, 2014). Altogether, non-denaturing purification methods of rTSβ via expression in *E. coli* may be futile. Future research regarding rTSβ should investigate the viability of higher order expression hosts and/or specialized strains of *E. coli*.

CHAPTER 5 CONCLUSIONS AND FUTURE WORK

The remarkable structural homology and conserved partial reaction of the ENS enzymes has served as a paradigm for understanding the relationship between enzyme structure and divergently evolved functions of homologous enzymes (Gerlt, *et al.*, 2012). As such, MR and its closely related homologues from the MR subgroup have been the subject of much research over the past three decades (Bearne & St. Maurice, 2017). In this thesis, I have investigated several topics with respect to the ENS, including the inhibition of MR by chloro- and fluoro-substituted PBAs, the contributions of the conserved KxK motif to the thermal stability of MR subgroup enzymes, and development of approaches towards production of the clinically relevant human MR subgroup member rTSβ for i*n vitro* functional studies.

Inhibition of MR by chloro- and fluoro-substituted PBAs

The catalytic proficiency of MR and corresponding affinity for the TS of its catalyzed reaction has been the subject of numerous inhibition studies, wherein inhibitor molecules have been rationally designed to mimic structural or electronic properties of the *aci*-carboxylate intermediate (Lietzan, *et al.*, 2012). MR binds TS analogue inhibitors with an affinities ~2-3 orders of magnitude greater than its affinity for the substrates, (*R*)- and (*S*)-mandelate (St. Maurice & Bearne, 2000; Bourque, *et al.*, 2007; Lietzan, *et al.*, 2012). Recently, however, the inhibition of MR by substituted PBAs has yielded binding affinities ~50-fold greater than the TS analogue inhibitors (Sharma, *et al.*, 2020). These enhanced affinities arise from multiple hydrogen bonding interaction, as well as the dative bond between His 297 and the boron atom of the inhibitors. In this work, I have

characterized the affinity and thermodynamics for MR binding a series of chloro- and fluoro-substituted PBAs using ITC. I have shown that various patterns of halogen substitution on the phenyl ring enhance the binding affinity by a further 2-3 orders of magnitude relative to the TS analogue inhibitors of MR. Specifically, 3,4-dichloroPBA, the most potent inhibition of MR reported to-date, was observed to have a dissociation constant of 13.8 ± 1.4 nM for binding to MR.

Thermodynamic analysis of the binding events revealed that the chloro- and fluoro-substituted PBAs bind in an entropy-driven manner. As a result, we hypothesize that the enhanced binding to MR exhibited by the halogenated PBAs arises from both the electron-withdrawing properties and increased interactions with the hydrophobic pocket of MR (St. Maurice & Bearne, 2004). The observed negative value of the change in specific heat capacity (ΔC_p) accompanying MR binding PBA is consistent with such a hydrophobic binding process. The thermodynamic analysis also suggested that o-substituted PBAs alter or weaken the hydrogen bonding network for PBA binding at the MR active site but do not significantly alter the potency of binding.

In the future, determination of ΔC_p for MR binding both 3,4-dichloroPBA and 3,4-dichloro-2-fluoroPBA may afford additional insights with respect to the structure-activity relationship discussed in this work. Furthermore, X-ray crystallography studies on both 3,4-dichloroPBA and 3,4-dichloro-2-fluoroPBA should be completed to provide atomistic insights into the thermodynamic trends observed for binding. Additional investigations on MR binding alternative halogen substituted PBAs (i.e., bromo- and/or iodo-) would also add valuable information to the structure-activity relationship. Overall,

this work provides additional knowledge about the application of boronic acid inhibitors as an inhibitor design strategy targeting enzymes that catalyze C–H cleavage reactions.

Thermal stability of KxK variant proteins within the MR subgroup of the ENS

As for all enzymes, electrostatic stabilization of reactive intermediates is required for catalysis in the MR subgroup of the ENS. To achieve this, MR subgroup enzymes feature a region of localized positive charge at the active site arising from the Mg^{2+} cation and the lysines of the KxK motif (Guthrie & Kluger, 1993; Kumar & Bearne, unpublished). As such, enzymes often prioritize catalytic function and proficiency over stability (Williams, 1972; Warshel, 1978). Recently, it was observed that the second lysine of the KxK motif is destabilizing to the enzyme, motivating an investigation into the contributions to enzyme stability made by the first lysine of the KxK motif in MR subgroup enzymes (Kumar & Bearne, unpublished). In this work I have assessed the thermal stability of a series of KxK variant proteins of the MR subgroup. I found that the KxK residue is essential for catalysis in the MR subgroup, as well as for the stability of MR and TarD. This stability is thought to arise several localized hydrogen bonds to the lysine residue that are disrupted upon conversion of the lysine to a methionine residue. Additionally, catalysis is entirely interrupted by the loss of the KxK binding determinant.

To further expand this work, ITC binding studies may provide useful thermodynamic data that could indicate whether the contributions of the lysine arise from entropic or enthalpic contributions. Detailed structural determinations of the variant proteins may also aid in developing a complete understanding of the calorimetric result.

Approaches to production of rTS

The classification of enzymes in the ENS is based on structural and sequential information as opposed to functional or mechanistic considerations (Gerlt, *et al.*, 2012). Consequently, the isoforms of the human protein rTS were identified as members of the ENS based on their sequential homology and active-site architecture compared with other members of the ENS (Wichelecki, *et al*., 2014; Dolnick, *et al.*, 2003). While Gerlt and colleagues functionally characterized the γ isoform as a FucD, the β isoform was found to be unstable and catalytically inactive, while only differing from the sequence of rTSγ in first 27 residues of the N-terminus. Furthermore, rTSβ has been studied for more than 25 years due to its aberrant expression in certain cancer cell lines. The presumed lack of catalytic function provoked research into the linkage between this protein and the MR subgroup (Wichelecki, *et al*., 2014).

In this work I described a number of approaches towards obtaining the rTSβ isoform for *in vitro* functional studies. Given the enhanced stability and ease of production of rTSγ, I hypothesized that incorporation of site-specific protease recognition sequences into rTSγ could allow for ample production of the full construct followed by post-purification cleavage to yield rTSβ. As such, rTSγ variants containing a Factor Xa or TEV protease site were generated and the extent of proteolytic cleavage catalyzed by these proteases was assessed. These production methods were mostly hindered by the inability to separate cleaved rTSβ from the uncleaved construct. Therefore, an N-terminal MBP fusion tag was incorporated to the construct (MBP-TEV-rTSβ), which could allow for separation of protease-cleaved and uncleaved construct by affinity chromatography on amylose. Unfortunately, the oligomeric state of MR subgroup enzymes

(tetrameric/octameric) likely led to the incorporation of uncleaved protein in the oligomeric complex, which hindered the purification.

Future studies should be focused on obtaining rTSβ using denaturing methods to solubilize the protein during the purification process. This requires both a robust method of protein refolding, as well as a method of assessing the extent of refolding. Furthermore, future studies may succeed if alternative expression hosts or specialized strains of *E. coli* are used.

REFERENCES

Adebodun, F., & Jordan, F. (1988). Boron-11 nuclear magnetic resonance studies of the structure of the transition-state analogue phenylboronic acid bound to chymotrypsin. *Journal of the American Chemical Society*, *110*(1), 309–310. https://doi.org/10.1021/ja00209a060

Akiva, E., Brown, S., Almonacid, D. E., Barber, A. E., Custer, A. F., Hicks, M. A., Huang, C. C., Lauck, F., Mashiyama, S. T., Meng, E. C., Mischel, D., Morris, J. H. Sunil, O., Schnoes, A. M., Stryke, D., Yunes, J. M., Ferrin, T. E., Holliday, G. L., Babbitt, P. C. (2014). The structure-function linkage database. *Nucleic Acids Research*, *42*(D1). https://doi.org/10.1093/nar/gkt1130

Albery, W. J., & Knowles, J. R. (1986). Energetics and mechanism of proline racemase. *Biochemistry*, *25*(9), 2572–2577. https://doi.org/10.1021/bi00357a043

Andersen, K. R., Leksa, N. C., & Schwartz, T. U. (2013). Optimized *E. coli* expression strain LOBSTR eliminates common contaminants from His-tag purification. *Proteins: Structure, Function and Bioinformatics*, *81*(11), 1857–1861. https://doi.org/10.1002/prot.24364

Babbitt, P. C., & Gerlt, J. A. (1997). Understanding enzyme superfamilies. Chemistry as the fundamental determinant in the evolution of new catalytic activities. *Journal of Biological Chemistry*, *272*(49), 30591–30594. https://doi.org/10.1074/jbc.272.49.30591

Babbitt, P. C., Hasson, M. S., Wedekind, J. E., Palmer, D. R. J., Barrett, W. C., Reed, G. H., Rayment, I., Ringe, D., Kenyon, G. L., Gerlt, J. A. (1996). The enolase superfamily: A general strategy for enzyme-catalyzed abstraction of the α-protons of carboxylic acids. *Biochemistry*, *35*(51), 16489–16501. https://doi.org/10.1021/bi9616413

Baker, S. J., Ding, C. Z., Akama, T., Zhang, Y. K., Hernandez, V., & Xia, Y. (2009). Therapeutic potential of boron-containing compounds. *Future Medicinal Chemistry*, *1*(7), 1275–1288. https://doi.org/10.4155/fmc.09.71

Baldwin, J. E., Claridge, T. D. W., Derome, A. E., Schofield, C. J., & Smith, B. D. (1991a). ^{11}B NMR studies of an aryl boronic acid bound to chymotrypsin and subtilisin. *Bioorganic and Medicinal Chemistry Letters*, *1*(1), 9–12. https://doi.org/10.1016/S0960-894X(01)81080-5

Baldwin, J. E., Claridge, T. D. W., Derome, A. E., Smith, B. D., Twyman, M., & Waley, S. G. (1991b). Direct observation of a tetrahedral boronic acid-β-lactamase complex using 11B NMR spectroscopy. *Journal of the Chemical Society, Chemical Communications*, (8), 573–574. https://doi.org/10.1039/C39910000573

Banner, D. W., Bloomer, A. C., Petsko, G. A., Phillips, D. C., Pogson, C. I., Wilson, I. A., Corran, P. H., Furth, A. J., Milman, J. D., Offord, R. E., Priddle, J. D., Waley, S. G. (1975). Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5Å resolution: Using amino acid sequence data. *Nature*, *255*(5510), 609–614. https://doi.org/10.1038/255609a0

Bearne, S. L. (2017). The interdigitating loop of the enolase superfamily as a specificity binding determinant or 'flying buttress.' *Biochimica et Biophysica Acta - Proteins and Proteomics*, *1865*(5), 619–630. https://doi.org/10.1016/j.bbapap.2017.02.006

Bearne, S. L., & St. Maurice, M. (2017). A Paradigm for C–H bond cleavage: structural and functional aspects of transition state stabilization by mandelate racemase. *Advances in Protein Chemistry and Structural Biology* (Vol. 109, pp. 113–160). Academic Press Inc. https://doi.org/10.1016/bs.apcsb.2017.04.007

Bearne, S. L., St. Maurice, M., & Vaughan, M. D. (1999). An assay for mandelate racemase using high-performance liquid chromatography. *Analytical Biochemistry*, *269*(2). https://doi.org/10.1006/abio.1999.4018

Bearne, S. L., & Wolfenden, R. (1997). Mandelate racemase in pieces: effective concentrations of enzyme functional groups in the transition state[†]. *Biochemistry*, *36*(7). https://doi.org/10.1021/bi9620722

Becker, D. J., & Lowe, J. B. (2003). Fucose: Biosynthesis and biological function in mammals. *Glycobiology*, *13*(7). https://doi.org/10.1093/glycob/cwg054

Bendel, P. (2005). Biomedical applications of ^{10}B and ^{11}B NMR. *NMR in Biomedicine*, *18*(2), 74–82. https://doi.org/10.1002/nbm.886

Bernasconi, C. F. (1992). The principle of nonperfect synchronization: more than a qualitative concept? *Accounts of Chemical Research*, *25*(1). https://doi.org/10.1021/ar00013a002

Bertino J. R., Carman M. D., Weiner H. L., Cashmore A., Moroson B. A., Srimatkandada S., Schornagel J. H., Medina W. D., Dube S. K. (1983). Gene amplification and altered enzymes as mechanisms for the development of drug resistance. *Cancer Treatment*, *67*(10), 901–904.

Bhat, S. G., & Vaidyanathan, C. S. (1976). Purification and properties of l-4-hydroxymandelate oxidase from *Pseudomonas convexa*. *European Journal of Biochemistry*, *68*(2). https://doi.org/10.1111/j.1432-1033.1976.tb10818.x

Black, A. R., & Dolnick, B. J. (1996). Expression of rTS correlates with altered growth regulation of thymidylate synthase. *Cancer Research*, *56*(4), 700–705.

Bourque, J. R., & Bearne, S. L. (2008). Mutational analysis of the active site flap (20s loop) of mandelate racemase. *Biochemistry*, *47*(2), 566–578. https://doi.org/10.1021/bi7015525

Bourque, J. R., Burley, R. K. M., & Bearne, S. L. (2007). Intermediate analogue inhibitors of mandelate racemase: *N*-hydroxyformanilide and cupferron. *Bioorganic and Medicinal Chemistry Letters*, *17*(1), 105–108. https://doi.org/10.1016/j.bmcl.2006.09.079

Brandt, D. S., & Chu, E. (1997). Future challenges in the clinical development of thymidylate synthase inhibitor compounds. *Oncology Research*, *9*(8), 403–410.

Cesaratto, F., Burrone, O. R., & Petris, G. (2016). Tobacco Etch Virus protease: A shortcut across biotechnologies. *Journal of Biotechnology*, *231*, 239–249. https://doi.org/10.1016/j.jbiotec.2016.06.012

Cleland, W., & Kreevoy, M. (1994). Low-barrier hydrogen bonds and enzymic catalysis. *Science*, *264*(5167). https://doi.org/10.1126/science.8009219

Chen, Z. J., Tian, Z., Kallio, K., Oleson, A. L., Ji, A., Borchardt, D., Jiang, D. E., Remington, S. J., Ai, H. W. (2016). The N-B interaction through a water bridge: understanding the chemoselectivity of a fluorescent protein based probe for peroxynitrite. *Journal of the American Chemical Society*, *138*(14), 4900–4907. https://doi.org/10.1021/jacs.6b01285

Chiang, Y., Kresge, A. J., Pruszynski, P., Schepp, N. P., & Wirz, J. (1990). The enol of mandelic acid, detection, acidity in aqueous solution, and estimation of the keto-enol equilibrium constant and carbon acidity of mandelic acid. *Angewandte Chemie International Edition*, *29*(7). https://doi.org/10.1002/anie.199007921

Costa, S. J., Almeida, A., Castro, A., Domingues, L., & Besir, H. (2013). The novel Fh8 and H fusion partners for soluble protein expression in Escherichia coli: A comparison with the traditional gene fusion technology. *Applied Microbiology and Biotechnology*, *97*(15), 6779–6791. https://doi.org/10.1007/s00253-012-4559-1

Costa, S., Almeida, A., Castro, A., & Domingues, L. (2014). Fusion tags for protein solubility, purification, and immunogenicity in *Escherichia coli*: The novel Fh8 system. *Frontiers in Microbiology*, *5*(FEB). https://doi.org/10.3389/fmicb.2014.00063

Dolnick, B. J., & Black, A. R. (1996). Alternate splicing of the rTS gene product and its overexpression in a 5-fluorouracil resistant cell line. *Cancer Research*, *56*(14), 3207–3210.

Dolnick, B. J. (1993). Cloning and characterization of a naturally occurring antisense RNA to human thymidylate synthase mRNA. *Nucleic Acids Research*, *21*(8), 1747–1752. https://doi.org/10.1093/nar/21.8.1747

Dolnick, B. J., Angelino, N. J., Dolnick, R., & Sufrin, J. R. (2003). A Novel Function for the rTS gene. *Cancer Biology & Therapy*, *2*(4). https://doi.org/10.4161/cbt.2.4.424

Dolnick, B. J., Black, A. R., Winkler, P. M., Schindler, K., & Hsueh, C. T. (1996). rTS gene expression is associated with altered cell sensitivity to thymidylate synthase inhibitors. *Advances in Enzyme Regulation*, *36*, 165–180. https://doi.org/10.1016/0065-2571(95)00009-7

Dolnick, B. J., Lu, K., Yin, M. B., & Rustum, Y. M. (1997). Recent advances in the study of rTS proteins. rTS expression during growth and in response to thymidylate synthase inhibitors in human tumor cells. *Advances in Enzyme Regulation*, *37*, 95–109. https://doi.org/10.1016/S0065-2571(96)00007-6

Dolnick, R., Wu, Q., Angelino, N. J., Stephanie, L. V., Chow, K. C., Sufrin, J. R., & Dolnick, B. J. (2005). Enhancement of 5-fluorouracil sensitivity by an rTS signaling mimic in H630 colon cancer cells. *Cancer Research*, *65*(13), 5917–5924. https://doi.org/10.1158/0008-5472.CAN-05-0431

Durowoju, I. B., Bhandal, K. S., Hu, J., Carpick, B., & Kirkitadze, M. (2017). Differential scanning calorimetry - A method for assessing the thermal stability and conformation of protein antigen. *Journal of Visualized Experiments*, *2017*(121). https://doi.org/10.3791/55262

Easton, N. M., Aboushawareb, S. A. E., & Bearne, S. L. (2018). A continuous assay for L-talarate/galactarate dehydratase using circular dichroism. *Analytical Biochemistry*, *544*, 80–86. https://doi.org/10.1016/j.ab.2017.12.015

Eigen, M. (1964). Proton transfer, acid-base catalysis, and enzymatic hydrolysis. Part I: elementary processes. *Angewandte Chemie International Edition*, *3*(1), 1–19. https://doi.org/10.1002/anie.196400011

Fee, J. A., Hegeman, G. D., & Kenyon, G. L. (1974). mandelate racemase from *Pseudomonas putida.* subunit composition and absolute divalent metal ion requirement. *Biochemistry*, *13*(12), 2528–2532. https://doi.org/10.1021/bi00709a008

Felfer, U., Strauss, U. T., Kroutil, W., Fabian, W. M. F., & Faber, K. (2001). Substrate spectrum of mandelate racemase Part 2. (Hetero)-aryl-substituted mandelate derivatives and modulation of activity. *Journal of Molecular Catalysis*, *15*(4–6), 213–222. https://doi.org/10.1016/S1381-1177(01)00035-2

Fetter, C. M., Morrison, Z. A., Nagar, M., Douglas, C. D., & Bearne, S. L. (2019). Altering the Y137-K164-K166 triad of mandelate racemase and its effect on the observed pK_a of the Brønsted base catalysts. *Archives of Biochemistry and Biophysics*, *666*, 116–126. https://doi.org/10.1016/j.abb.2019.03.011

Fisher, L. M., Albery, W. J., & Knowles, J. R. (2538). Energetics of proline racemase: tracer perturbation experiments using [^{14}C] proline that measure the interconversion rate of the two forms of free enzyme. *Biochemistry*, *25*, 2538–2542.

Fitch, C. A., Platzer, G., Okon, M., Garcia-Moreno, B. E., & McIntosh, L. P. (2015). Arginine: Its pK_a value revisited. *Protein Science*, *24*(5), 752–761. https://doi.org/10.1002/pro.2647

Frenking, G. (2015). Peculiar boron startles again. *Nature*, *522*(7556), 297–298. https://doi.org/10.1038/522297a

Froger, A., & Hall, J. E. (2007). Transformation of plasmid DNA into *E. Coli* using the heat shock method. *Journal of Visualized Experiments*, (6). https://doi.org/10.3791/253

Fujita, T., Iwasa, J., & Hansch, C. (1964). A new substituent constant, π, derived from partition coefficients. *Journal of the American Chemical Society*, *86*(23), 5175–5180. https://doi.org/10.1021/ja01077a028

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. *The Proteomics Protocols Handbook*. Totowa, NJ: Humana Press. https://doi.org/10.1385/1-59259-890-0:571

Gerlt, J. A., & Gassman, P. G. (1995). Mandelate racemase (MR,1 EC 5.1.2.2) from *Pseudomonas putida* ATCC 12633 catalyzes the equilibration of the (*R*)-and (*S*)-enantiomers of mandelate (Kenyon & Hegeman. *J. Am. Chem. Soc*, *34*, 11943–11952.Gerlt, J. A., Kozarich, J. W., Kenyon, G. L., & Gassman, P. G. (1991). Electrophilic catalysis can explain the unexpected acidity of carbon acids in enzyme-catalyzed Reactions. *Journal of the American Chemical Society*, *113*(25), 9667–9669. https://doi.org/10.1021/ja00025a039

Gerlt, J. A., Babbitt, P. C., Jacobson, M. P., & Almo, S. C. (2012). Divergent evolution in enolase superfamily: Strategies for assigning functions. *Journal of Biological Chemistry*, *287*(1), 29–34. https://doi.org/10.1074/jbc.R111.240945

Gerlt, J. A., Babbitt, P. C., & Rayment, I. (2005). Divergent evolution in the enolase superfamily: The interplay of mechanism and specificity. *Archives of Biochemistry and Biophysics*, *433*(1), 59–70. https://doi.org/10.1016/j.abb.2004.07.034

Gerlt, J. A., & Gassman, P. G. (1993). Understanding the rates of certain enzyme-catalyzed reactions: Proton abstraction from carbon acids, acyl-transfer reactions, and displacement reactions of phosphodiesters. *Biochemistry*, *32*(45), 11943–11952. https://doi.org/10.1021/bi00096a001

Gerlt, J. A., & Gassman, P. G. (1993). An explanation for rapid enzyme-catalyzed proton abstraction from carbon acids: importance of late transition states in concerted mechanisms. *Journal of the American Chemical Society*, *115*(24). https://doi.org/10.1021/ja00077a062

Gerlt, J. A., & Gassman, P. G. (1992). Understanding enzyme-catalyzed proton abstraction from carbon acids: Details of stepwise mechanisms for β-elimination reactions. *Journal of the American Chemical Society*, *114*(15), 5928–5934. https://doi.org/10.1021/ja00041a004

Gerlt, J. A., Kozarich, J. W., Kenyon, G. L., & Gassman, P. G. (1992). Additions and corrections: electrophilic catalysis can explain the unexpected acidity of carbon acids in enzyme-catalyzed reactions. *Journal of the American Chemical Society*, *114*(10), 4016. https://doi.org/10.1021/ja00036a083

Gerlt, J. A., Kreevoy, M. M., Cleland, W. W., & Frey, P. A. (1997). Understanding enzymic catalysis: The importance of short, strong hydrogen bonds. *Chemistry and Biology*, *4*(4), 259–267. https://doi.org/10.1016/S1074-5521(97)90069-7

Ghasempur, S., Eswaramoorthy, S., Hillerich, B. S., Seidel, R. D., Swaminathan, S., Almo, S. C., & Gerlt, J. A. (2014). Discovery of a novel L-lyxonate degradation pathway in *Pseudomonas aeruginosa* PAO1. *Biochemistry*, *53*(20), 3357–3366. https://doi.org/10.1021/bi5004298

Grimsley, G. R., Scholtz, J. M., & Pace, C. N. (2009). A summary of the measured pK values of the ionizable groups in folded proteins. *Protein Science*, *18*(1), 247–251. https://doi.org/10.1002/pro.19

Guan, R., Ho, M. C., Brenowitz, M., Tyler, P. C., Evans, G. B., Almo, S. C., & Schramm, V. L. (2011). Entropy-driven binding of picomolar transition state analogue inhibitors to Human 5′-methylthioadenosine phosphorylase. *Biochemistry*, *50*(47), 10408–10417. https://doi.org/10.1021/bi201321x

Gunsalus, C. F., Stanier, R. Y., & Gunsalus, I. C. (1953). The enzymatic conversion of mandelic acid to benzoic acid III. Fractionation and properties of the soluble enzymes. *Journal of Bacteriology*, *66*(5). https://doi.org/10.1128/JB.66.5.548-553.1953

Gunsalus, I. C., Gunsalus, C. F., & Stanier, R. Y. (1953). The enzymatic conversion of mandelic acid to benzoic acid I. Gross fractionation of the system into soluble and particulate components. *Journal of Bacteriology*, *66*(5). https://doi.org/10.1128/JB.66.5.538-542.1953

Guthrie, J. P. (1997). Prediction of the rate constants for proton abstraction from carbon acids, using a simple model and multidimensional marcus theory. *Journal of the American Chemical Society*, *119*(5). https://doi.org/10.1021/ja9629367

Guthrie, J. P., & Kluger, R. (1993). Electrostatic stabilization can explain the unexpected acidity of carbon acids in enzyme-catalyzed reactions. *Journal of the American Chemical Society*, *115*(24), 11569–11572. https://doi.org/10.1021/ja00077a063

Hammond, G. S. (1955). A correlation of reaction rates. *Journal of the American Chemical Society*, *77*(2), 334–338. https://doi.org/10.1021/ja01607a027

Hansch, C., Leo, A., & Taft, R. W. (1991). A survey of hammett substituent constants and resonance and field parameters. *Chemical Reviews*, *91*(2), 165–195. https://doi.org/10.1021/cr00002a004

Harris, T. K., & Turner, G. J. (2002). Structural basis of perturbed pK_a values of catalytic groups in enzyme active sites. *IUBMB Life*, *53*(2), 85–98. https://doi.org/10.1080/15216540211468

Kenyon, G. L., & Hegeman, G. D. (1970). Mandelic acid racemase from Pseudomonas putida. Evidence favoring a carbanion intermediate in the mechanism of action. *Biochemistry*, *9*(21), 4036–4043. https://doi.org/10.1021/bi00823a002

Hemsley, A., Arnheim, N., Toney, M. D., Cortopassi, G., & Galas, D. J. (1989). A simple method for site-directed mutagenesis using the polymerase chain reaction. *Nucleic Acids Research*, *17*(16), 6545–6551. https://doi.org/10.1093/nar/17.16.6545

Holdgate, G. A., Tunnicliffe, A., Ward, W. H. J., Weston, S. A., Rosenbrock, G., Barth, P. T., Taylor, I.W.F., Pauptit, R. A., Timms, D. (1997). The entropic penalty of ordered water accounts for weaker binding of the antibiotic novobiocin to a resistant mutant of DNA gyrase: A thermodynamic and crystallographic study. *Biochemistry*, *36*(32), 9663–9673. https://doi.org/10.1021/bi970294+

Holdgate, G. A., & Ward, W. H. J. (2005). Measurements of binding thermodynamics in drug discovery. *Drug Discovery Today*, *10*(22), 1543–1550. https://doi.org/10.1016/S1359-6446(05)03610-X

Huynh, K., & Partch, C. L. (2015). Analysis of protein stability and ligand interactions by thermal shift assay. *Current Protocols in Protein Science*, *79*(1), 28.9.1-28.9.14. https://doi.org/10.1002/0471140864.ps2809s79

Jenny, R. J., Mann, K. G., & Lundblad, R. L. (2003). A critical review of the methods for cleavage of fusion proteins with thrombin and Factor Xa. *Protein Expression and Purification*, *31*(1), 1–11. https://doi.org/10.1016/S1046-5928(03)00168-2

Johnson, C. M. (2013). Differential scanning calorimetry as a tool for protein folding and stability. *Archives of Biochemistry and Biophysics*, *531*(1–2), 100–109. https://doi.org/10.1016/j.abb.2012.09.008

Kajander, T., Kahn, P. C., Passila, S. H., Cohen, D. C., Lehtiö, L., Adolfsen, W., Warwicker, J., Schell, U., Goldman, A. (2000). Buried charged surface in proteins. *Structure*, *8*(11), 1203–1214. https://doi.org/10.1016/S0969-2126(00)00520-7

Kallarakal, A. T., Mitra, B., Kozarich, J. W., Gerlt, J. A., Clifton, J. G., Petsko, G. A., Kenyon, G. L., Landro, J. A. (1995). Mechanism of the reaction catalyzed by mandelate racemase: structure and mechanistic properties of the K166R mutant. *Biochemistry*, *34*, 2788-2797. https://doi.org/10.1021/bi00009a007

Kanaya, S., Oobatake, M., & Liu, Y. (1996). Thermal stability of *Escherichia coli* ribonuclease HI and its active site mutants in the presence and absence of the Mg^{2+} Ion. *Journal of Biological Chemistry*, *271*(51). https://doi.org/10.1074/jbc.271.51.32729

Kaur, J., Kumar, A., & Kaur, J. (2018). Strategies for optimization of heterologous protein expression in *E. coli:* Roadblocks and reinforcements. *International Journal of Biological Macromolecules*, *106*, 803–822. https://doi.org/10.1016/j.ijbiomac.2017.08.080

Kitamura, M., Suzuki, T., Abe, R., Ueno, T., & Aoki, S. (2011). ^{11}B NMR sensing of d-block metal ions *in vitro* and in cells based on the carbon-boron bond cleavage of phenylboronic acid-pendant cyclen. *Chemistry*, *50*(22), 11568–11580. https://doi.org/10.1021/ic201507q

Koshland, D. E. (1995). The key–lock theory and the induced fit theory. *Angewandte Chemie International Edition*, *33*(23–24), 2375–2378. https://doi.org/10.1002/anie.199423751

Kuo, S. J., Wang, H. C., Chow, K. C., Chiou, S. H., Chiang, S. F., Lin, T. Y., Lin, T. Y., Chiang, I. P., Chen, D. R. (2008). Expression of rTSβ as a 5-fluorouracil resistance marker in patients with primary breast cancer. *Oncology Reports*, *19*(4), 881–888. https://doi.org/10.3892/or.19.4.881

Landro, J. A., Gerlt, J. A., Kozarich, J. W., Koo, C. W., Shah, V. J., Kenyon, G. L., Neidhart, D. J., Fujita, S., Petsko, G. A. (1994). The role of lysine 166 in the mechanism of mandelate racemase from *Pseudomonas putida*: Mechanistic and crystallographic evidence for stereospecific alkylation by (R)-α-phenylglycidate. *Biochemistry*, *33*(3), 635–643. https://doi.org/10.1021/bi00169a003

Landro, J. A., Kallarakal, A. T., Ransom, S. C., Gerlt, J. A., Kozarich, J. W., Neidhart, D. J., & Kenyon, G. L. (1991). Mechanism of the reaction catalyzed by mandelate racemase. 3. Asymmetry in reactions catalyzed by the H297N mutant. *Biochemistry*, *30*(38), 9274–9281. https://doi.org/10.1021/bi00102a020

Lanier, M., Cole, D. C., Istratiy, Y., Klein, M. G., Schwartz, P. A., Tjhen, R., Jennings, A., Hixon, M. S. (2017). Repurposing Suzuki coupling reagents as a directed fragment library targeting serine hydrolases and related enzymes. *Journal of Medicinal Chemistry*, *60*(12), 5209–5215. https://doi.org/10.1021/acs.jmedchem.6b01224

Li, R., Powers, V. M., Kenyon, G. L., & Kozarich, J. W. (1995). Racemization of vinylglycolate catalyzed by mandelate racemase. *Journal of Organic Chemistry*, *60*(11), 3347–3351. https://doi.org/10.1021/jo00116a017

Liang, P., Nair, J. R., Lei, S., McGuire, J. J., & Dolnick, B. J. (2005). Comparative genomic analysis reveals a novel mitochondrial isoform of human rTS protein and unusual phylogenetic distribution of the rTS gene. *BMC Genomics*, *6*. https://doi.org/10.1186/1471-2164-6-125

Lichty, J. J., Malecki, J. L., Agnew, H. D., Michelson-Horowitz, D. J., & Tan, S. (2005). Comparison of affinity tags for protein purification. *Protein Expression and Purification*, *41*(1), 98–105. https://doi.org/10.1016/j.pep.2005.01.019

Lietzan, A. D., Nagar, M., Pellmann, E. A., Bourque, J. R., Bearne, S. L., & St. Maurice, M. (2012). Structure of mandelate racemase with bound intermediate analogues benzohydroxamate and cupferron. *Biochemistry*, *51*(6), 1160–1170. https://doi.org/10.1021/bi2018514

Lin, Y. L., & Chow, K. C. (2010). rTSβ as a novel 5-fluorouracil resistance marker of colorectal cancer: A preliminary study. *Annals of the Academy of Medicine Singapore*, *39*(2), 107–111.

Lin, Z., Schwarz, F. P., & Eisenstein, E. (1995). The hydrophobic nature of GroEL-substrate binding. *Journal of Biological Chemistry*, *270*(3), 1011–1014. https://doi.org/10.1074/jbc.270.3.1011

London, R. E., & Gabel, S. A. (2001). Development and evaluation of a boronate inhibitor of γ-glutamyl transpeptidase. *Archives of Biochemistry and Biophysics*, *385*(2), 250–258. https://doi.org/10.1006/abbi.2000.2169

Maki, T., Ishihara, K., & Yamamoto, H. (2006). 4,5,6,7-Tetrachlorobenzo[d][1,3,2]dioxaborol-2-ol as an effective catalyst for the amide condensation of sterically demanding carboxylic acids. *Organic Letters*, *8*(7), 1431–1434. https://doi.org/10.1021/ol060216r

Maurice, M. S., Bearne, S. L., Lu, W., & Taylor, S. D. (2003). Inhibition of mandelate racemase by α-fluorobenzylphosphonates. *Bioorganic and Medicinal Chemistry Letters*, *13*(12), 2041–2044. https://doi.org/10.1016/S0960-894X(03)00311-1

Maurice, M. S., & Bearne, S. L. (2002). Kinetics and thermodynamics of mandelate racemase catalysis. *Biochemistry*, *41*(12), 4048–4058. https://doi.org/10.1021/bi016044h

Mazurenko, S., Kunka, A., Beerens, K., Johnson, C. M., Damborsky, J., & Prokop, Z. (2017). Exploration of protein unfolding by modelling calorimetry data from reheating. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-16360-y

Menéndez, M. (2020). Isothermal titration calorimetry: principles and applications. in *eLS* (pp. 113–127). Wiley. https://doi.org/10.1002/9780470015902.a0028808

Michal G.Rose, Michael P.Farrell, & John C.Schmitz. (2002). Thymidylate synthase: a critical target for cancer chemotherapy. *Clinical Colorectal Cancer*, *1*(4), 220–229.

Miller, B. G., & Wolfenden, R. (2002). Catalytic proficiency: The unusual case of OMP decarboxylase. *Annual Review of Biochemistry*, *71*(1), 847–885. https://doi.org/10.1146/annurev.biochem.71.110601.135446

Mitra, B., Kallarakal, A. T., Kozarich, J. W., Gerlt, J. A., Clifton, J. G., Petsko, G. A., & Kenyon, G. L. (1995). Mechanism of the reaction catalyzed by mandelate racemase: importance of electrophilic catalysis by glutamic acid 317. *Biochemistry*, *34*(9), 2777–2787. https://doi.org/10.1021/bi00009a006

Mozzarelli, A., & Bettati, S. (2006). Exploring the pyridoxal 5′-phosphate-dependent enzymes. *Chemical Record*, *6*(5), 275–287. https://doi.org/10.1002/tcr.20094

Nagar, M., & Bearne, S. L. (2015). An additional role for the Brønsted acid-base catalysts of mandelate racemase in transition state stabilization. *Biochemistry*, *54*(44), 6743–6752. https://doi.org/10.1021/acs.biochem.5b00982

Nagar, M., Lietzan, A. D., St. Maurice, M., & Bearne, S. L. (2014). Potent inhibition of mandelate racemase by a fluorinated substrate-product analogue with a novel binding mode. *Biochemistry*, *53*(7), 1169–1178. https://doi.org/10.1021/bi401703h

Nagar, M., Narmandakh, A., Khalak, Y., & Bearne, S. L. (2011). Redefining the minimal substrate tolerance of mandelate racemase. Racemization of trifluorolactate. *Biochemistry*, *50*(41), 8846–8852. https://doi.org/10.1021/bi201188j

Neidhart, D. J., Howell, P. L., Petsko, G. A., Powers, V. M., Li, R., Kenyon, G. L., & Gerlt, J. A. (1991). Mechanism of the reaction catalyzed by mandelate racemase. 2. crystal structure of mandelate racemase at 2.5-Å resolution: Identification of the active site and possible catalytic residues. *Biochemistry*, *30*(38), 9264–9273. https://doi.org/10.1021/bi00102a019

Neidhart, D. J., Kenyon, G. L., Gerlt, J. A., & Petsko, G. A. (1990). Mandelate racemase and muconate lactonizing enzyme are mechanistically distinct and structurally homologous. *Nature*, *347*(6294), 692–694. https://doi.org/10.1038/347692a0

Niesen, F. H., Berglund, H., & Vedadi, M. (2007). The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature Protocols*, *2*(9), 2212–2221. https://doi.org/10.1038/nprot.2007.321

Pace, C. N., Fu, H., Fryar, K. L., Landua, J., Trevino, S. R., Schell, D., Thurlkill, R. L., Imura, S., Scholtz, J. M., Gajiwala, K., Sevcik, J., Urbanikova, L., Myers, J. K., Takano, K., Hebert, E. J., Shirley, B. A., Grimsley, G. R. (2014). Contribution of hydrogen bonds to protein stability. *Protein Science*, *23*(5), 652–661. https://doi.org/10.1002/pro.2449

Parker, M. H., Lunney, E. A., Ortwine, D. F., Pavlovsky, A. G., Humblet, C., & Brouillette, C. G. (1999). Analysis of the binding of hydroxamic acid and carboxylic acid inhibitors to the stromelysin-1 (matrix metalloproteinase-3) catalytic domain by isothermal titration calorimetry. *Biochemistry*, *38*(41), 13592–13601. https://doi.org/10.1021/bi991222g

Pauling, L. (1948). Nature of forces between large molecules of biological interest. *Nature*, *161*(4097), 707–709. https://doi.org/10.1038/161707a0

Powers, V. M., Koo, C. W., Kenyon, G. L., Gerlt, J. A., & Kozarich, J. W. (1991). Mechanism of the reaction catalyzed by mandelate racemase. 1. Chemical and kinetic evidence for a two-base mechanism. *Biochemistry*, *30*(38), 9255–9263. https://doi.org/10.1021/bi00102a018

Privalov, P. L., & Gill, S. J. (1988). Stability of protein structure and hydrophobic interaction. *Advances in Protein Chemistry*, *39*(C), 191–234. https://doi.org/10.1016/S0065-3233(08)60377-0

Radzicka, A., & Wolfenden, R. (1995). A proficient enzyme. *Science*, *267*(5194), 90–93. https://doi.org/10.1126/science.7809611

Rajagopalan, P. T. R., Zhang, Z., McCourt, L., Dwyer, M., Benkovic, S. J., & Hammes, G. G. (2002). Interaction of dihydrofolate reductase with methotrexate: Ensemble and single-molecule kinetics. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(21), 13481–13486. https://doi.org/10.1073/pnas.172501499

Rajagopalan, P. T. R., Zhang, Z., McCourt, L., Dwyer, M., Benkovic, S. J., & Hammes, G. G. (2002). Interaction of dihydrofolate reductase with methotrexate: Ensemble and single-molecule kinetics. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(21), 13481–13486. https://doi.org/10.1073/pnas.172501499

Richard, J. P., & Amyes, T. L. (2001). Proton transfer at carbon. *Current Opinion in Chemical Biology*, *5*(6), 626–633. https://doi.org/10.1016/S1367-5931(01)00258-7

Sayer, J. M., & Louis, J. M. (2009). Interactions of different inhibitors with active-site aspartyl residues of HIV-1 protease and possible relevance to pepsin. *Proteins: Structure, Function and Bioinformatics*, *75*(3), 556–568. https://doi.org/10.1002/prot.22271

Schramm, V. L. (2005). Enzymatic transition states and transition state analogues. *Current Opinion in Structural Biology*, *15*(6), 604–613. https://doi.org/10.1016/j.sbi.2005.10.017

Sharma, A. N., Grandinetti, L., Johnson, E. R., St Maurice, M., & Bearne, S. L. (2020). Potent inhibition of mandelate racemase by boronic acids: Boron as a mimic of a carbon acid center. *Biochemistry*, *59*(33), 3026–3037. https://doi.org/10.1021/acs.biochem.0c00478

Sharp, T. R., Hegeman, G. D., & Kenyon, G. L. (1979). A direct kinetic assay for mandelate racemase using circular dichroic measurements. *Analytical Biochemistry*, *94*(2), 329–334. https://doi.org/10.1016/0003-2697(79)90368-3

Siddiqi, F., Bourque, J. R., Jiang, H., Gardner, M., St. Maurice, M., Blouin, C., & Bearne, S. L. (2005). Perturbing the hydrophobic pocket of mandelate racemase to probe phenyl motion during catalysis. *Biochemistry*, *44*(25), 9013–9021. https://doi.org/10.1021/bi0473096

Singh, A., Upadhyay, V., Upadhyay, A. K., Singh, S. M., & Panda, A. K. (2015). Protein recovery from inclusion bodies of *Escherichia coli* using mild solubilization process. *Microbial Cell Factories*, *14*(1). https://doi.org/10.1186/s12934-015-0222-8

Srimatkandada, S., Medina, W. D., Cashmore, A. R., Whyte, W., Engel, D., Moroson, B. A., Franco, C. T., Dube, S. K., Bertino, J. R. (1983). Amplification and organization of dihydrofolate reductase genes in a human leukemic cell line, K-562, resistant to methotrexate. *Biochemistry*, *22*(25), 5774–5781. https://doi.org/10.1021/bi00294a015

St. Maurice, M., & Bearne, S. L. (2000). Reaction intermediate analogues for mandelate racemase: Interaction between Asn 197 and the α-hydroxyl of the substrate promotes catalysis. *Biochemistry*, *39*(44), 13324–13335. https://doi.org/10.1021/bi001144t

St. Maurice, M., & Bearne, S. L. (2004). Hydrophobic nature of the active site of mandelate racemase. *Biochemistry*, *43*(9), 2524–2532. https://doi.org/10.1021/bi036207x

Sturtevant, J. M. (1987). Biochemical applications of differential scanning calorimetry. *Annual Review of Physical Chemistry*, *38*(1), 463–488. https://doi.org/10.1146/annurev.pc.38.100187.002335

Sun, C., Liang, J., Shi, R., Gao, X., Zhang, R., Hong, F., Yuan, Q., Wang, S. (2012). Tobacco etch virus protease retains its activity in various buffers and in the presence of diverse additives. *Protein Expression and Purification*, *82*(1), 226–231. https://doi.org/10.1016/j.pep.2012.01.005

Thibblin, A., & Jencks, W. P. (1979). Unstable carbanions. general acid catalysis of the cleavage of 1-phenylcyclopropanol and 1 -phenyl-2-arylcyclopropanol anions. *Journal of the American Chemical Society*, *101*(17), 4963–4973. https://doi.org/10.1021/ja00511a028

Toney, M. D. (2019). Carbon acidity in enzyme active sites. *Frontiers in Bioengineering and Biotechnology*, *7*(FEB). https://doi.org/10.3389/fbioe.2019.00025

Trippier, P. C., & McGuigan, C. (2010). Boronic acids in medicinal chemistry: Anticancer, antibacterial and antiviral applications. *MedChemComm*, *1*(3), 183–198. https://doi.org/10.1039/c0md00119h

Tsilikounas, E., Kettner, C. A., & Bachovchin, W. W. (1993). ^{11}B NMR spectroscopy of peptide boronic acid inhibitor complexes of α-lytic protease. Direct evidence for tetrahedral boron in both boron-histidine and boron-serine adduct complexes. *Biochemistry*, *32*(47), 12651–12655. https://doi.org/10.1021/bi00210a013

Tsilikounas, E., Kettner, C., & Bachovchin, W. W. (1992). Identification of serine and histidine adducts in complexes of trypsin and trypsinogen with peptide and nonpeptide boronic acid inhibitors by proton NMR spectroscopy. *Biochemistry*, *31*(51), 12839–12846. https://doi.org/10.1021/bi00166a019

Tsou, A. Y., Ransom, S. C., Gerlt, J. A., Powers, V. M., & Kenyon, G. L. (1989). Selection and characterization of a mutant of the cloned gene for mandelate racemase that confers resistance to an affinity label by greatly enhanced production of enzyme. *Biochemistry*, *28*(3), 969–975. https://doi.org/10.1021/bi00429a008

Tulinsky, A., & Blevins, R. A. (1987). Structure of a tetrahedral transition state complex of alpha-chymotrypsin dimer at 1.8-A resolution. *Journal of Biological Chemistry*, *262*(16), 7737–7743. https://doi.org/10.2210/pdb6cha/pdb

Velazquez-Campoy, A., Kiso, Y., & Freire, E. (2001). The binding energetics of first- and second-generation HIV-1 protease inhibitors: Implications for drug design. *Archives of Biochemistry and Biophysics*, *390*(2), 169–175. https://doi.org/10.1006/abbi.2001.2333

Velazquez-Campoy, A., Kiso, Y., & Freire, E. (2001). The binding energetics of first- and second-generation HIV-1 protease inhibitors: Implications for drug design. *Archives of Biochemistry and Biophysics*, *390*(2), 169–175. https://doi.org/10.1006/abbi.2001.2333

Voet, D., Voet, J. G. (2004). *Biochemistry* (3rd ed.). John Wiley & Sons, Inc.

Vogt, G., Woell, S., & Argos, P. (1997). Protein thermal stability, hydrogen bonds, and ion pairs. *Journal of Molecular Biology*, *269*(4), 631–643. https://doi.org/10.1006/jmbi.1997.1042

Vuillard, L., Rabilloud, T., & Goldberg, M. E. (1998). Interactions of non-detergent sulfobetaines with early folding intermediates facilitate in vitro protein renaturation. *European Journal of Biochemistry*, *256*(1), 128–135. https://doi.org/10.1046/j.1432-1327.1998.2560128.x

Ward, W. H. J., & Holdgate, G. A. (2001). Isothermal titration calorimetry in drug discovery. *Progress in Medicinal Chemistry*, *38*(C), 309–376. https://doi.org/10.1016/S0079-6468(08)70097-3

Warshel, A. (1978). Energetics of enzyme catalysis. *Proceedings of the National Academy of Sciences of the United States of America*, *75*(11), 5250–5254. https://doi.org/10.1073/pnas.75.11.5250

Waugh, D. S. (2011). An overview of enzymatic reagents for the removal of affinity tags. *Protein Expression and Purification*, *80*(2), 283–293. https://doi.org/10.1016/j.pep.2011.08.005

Wen, S. Y., Fedorov, A. A., Fedorov, E. V., Almo, S. C., & Gerlt, J. A. (2007). Evolution of enzymatic activities in the enolase superfamily: L-talarate/galactarate dehydratase from Salmonella typhimurium LT2. *Biochemistry*, *46*(33), 9564–9577. https://doi.org/10.1021/bi7008882

Wen, S. Y., Fedorov, A. A., Fedorov, E. V., Almo, S. C., & Gerlt, J. A. (2007). Evolution of enzymatic activities in the enolase superfamily: L-talarate/galactarate dehydratase from Salmonella typhimurium LT2. *Biochemistry*, *46*(33), 9564–9577. https://doi.org/10.1021/bi7008882

Whyte, G. F., Vilar, R., & Woscholski, R. (2013). Molecular recognition with boronic acids-applications in chemical biology. *Journal of Chemical Biology*, *6*(4), 161–174. https://doi.org/10.1007/s12154-013-0099-0

Wichelecki, D. J., Balthazor, B. M., Chau, A. C., Vetting, M. W., Fedorov, A. A., Fedorov, E. V., Lukk, T., Patskovsky, Y. V., Stead, M. B., Hillerich, B. S., Seidel, R. D., Almo, S. C., Gerlt, J. A. (2014). Discovery of function in the enolase superfamily: D-mannonate and D-gluconate dehydratases in the d-mannonate dehydratase subgroup. *Biochemistry*, *53*(16), 2722–2731. https://doi.org/10.1021/bi500264p

Wichelecki, D. J., Froese, D. S., Kopec, J., Muniz, J. R. C., Yue, W. W., & Gerlt, J. A. (2014). Enzymatic and structural characterization of rTSγ provides insights into the function of rTSβ. *Biochemistry*, *53*(16), 2732–2738. https://doi.org/10.1021/bi500349e

Williams, R. J. P. (1971). Catalysis by metallo-enzymes: The entatic state. *Inorganica Chimica Acta Reviews*, *5*(C), 137–155. https://doi.org/10.1016/0073-8085(71)80016-5

Windsor, I. W., Palte, M. J., Lukesh, J. C., Gold, B., Forest, K. T., & Raines, R. T. (2018). Sub-picomolar inhibition of HIV-1 protease with a boronic acid. *Journal of the American Chemical Society*. https://doi.org/10.1021/jacs.8b07366

Wiseman, T., Williston, S., Brandts, J. F., & Lin, L. N. (1989). Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Analytical Biochemistry*, *179*(1), 131–137. https://doi.org/10.1016/0003-2697(89)90213-3

Wolfenden, R. (1972). Analog approaches to the structure of the transition state in enzyme reactions. *Accounts of Chemical Research*, *5*(1), 10–18. https://doi.org/10.1021/ar50049a002

Yakovlev, G. I., Mitkevich, V. A., Shaw, K. L., Trevino, S., Newsom, S., Pace, C. N., & Makarov, A. A. (2009). Contribution of active site residues to the activity and thermal stability of ribonuclease Sa. *Protein Science*, *12*(10), 2367–2373. https://doi.org/10.1110/ps.03176803

Yew, W. S., Fedorov, A. A., Fedorov, E. V., Almo, S. C., & Gerlt, J. A. (2007). Evolution of enzymatic activities in the enolase superfamily: L-talarate/galactarate dehydratase from *Salmonella typhimurium*. *Biochemistry*, *46*(33), 9564–9577. https://doi.org/10.1021/bi7008882

Yew, W. S., Fedorov, A. A., Fedorov, E. V., Rakus, J. F., Pierce, R. W., Almo, S. C., & Gerlt, J. A. (2006a). Evolution of enzymatic activities in the enolase superfamily: L-fuconate dehydratase from *Xanthomonas campestris*. *Biochemistry*, *45*(49), 14582–14597. https://doi.org/10.1021/bi061687o

Yew, W. S., Fedorov, A. A., Fedorov, E. V., Wood, B. M. K., Almo, S. C., & Gerlt, J. A. (2006b). Evolution of enzymatic activities in the enolase superfamily: D-tartrate dehydratase from *Bradyrhizobium japonicum. Biochemistry*, *45*(49), 14598–14608. https://doi.org/10.1021/bi061688g

Yuen, R., & Schachter, H. (1972). L-Fucose metabolism in mammals. I. Port liver L-fuconate hydro-lyase. *Canadian Journal of Biochemistry*, *50*(7), 798–806. https://doi.org/10.1139/o72-111

Zervosen, A., Herman, R., Kerff, F., Herman, A., Bouillez, A., Prati, F., Pratt, R. F., Frère, J. M., Joris, B., Luxen, A., Charlier, P., Sauvage, E. (2011). Unexpected tricovalent binding mode of boronic acids within the active site of a penicillin-binding protein. *Journal of the American Chemical Society*, *133*(28), 10839–10848. https://doi.org/10.1021/ja200696y

Zhong, S., Jordan, F., Polgar, L., & Kettner, C. (1991). Observation of tightly bound [11]B nuclear magnetic resonance signals on serine proteases. direct solution evidence for tetrahedral geometry around the boron in the putative transition-state analogues. *Journal of the American Chemical Society*, *113*(25), 9429–9435. https://doi.org/10.1021/ja00025a001

Zhou, M., Tang, C., Xu, J., & Yu, H. (2018). Genome mining and characterization of a new mandelate racemase. *Chinese Journal of Biotechnology*, *34*(6), 897–905.

APPENDIX A

ORF DNA ENCODING rTS PROTEIN CONSTRUCTS

rTSα

5´-ATGCAGAAAATGGAATCTCGTGGTGTTGAACTGCCGTCCCTGTGGGAAAA
AGCGCTGAAACTGATCGGTCCGGAAAAAGGTGTTGTGCACCTGGCTACCGCG
GCAGTTCTGAACGCGGTTTGGGACCTGTGGGCGAAACAGGAAGGTAAACCGG
TTTGGAAACTGCTGGTTGATATGGATCCGCGTATGCTGGTATCTTGCATCGAT
TTCCGTTATATCACTGACGTTCTGACCGAAGAAGACGCTCTGGAAATCCTGCA
GAAAGGCCAGATCGGTAAAAAGAACGCGAAAACAGATGCTGGCACAGGG
CTACCCGGCGTACACCACCTCCTGCGCTTGGCTGGGTTACTCTGACGATACTC
TGAAACAGCTGTGTGCACAGGCGCTGAAAGATGGTTGGACCCGTTTCAAAGT
TAAAGTTGGTGCTGATCTGCAGGATGACATGCGTCGTTGTCAGATTATTCGTG
ACATGATTGGTCCTGAAAAAACTCTGATGATGGATGCGAACCAGCGTTGGGA
CGTTCCGGAAGCAGTTGAATGGATGAGCAAACTGGCTAAATTCAAACCGCTG
TGGATTGAAGAACCGACCTCTCCGGATGACATTCTGGGCCACGCTACCATCA
GCAAAGCTCTGGTTCCGCTGGGTATCGGCATCGCAACCGGTGAACAGTGCCA
CAACCGTGTTATCTTCAAACAGCTGCTGCAGGCGAAAGCGCTGCAGTTCCTG
CAGATTGACTCCTGCCGTCTGGGCTCTGTTAACGAAACCTGAGCGTGCTGCT
GATGGCTAAAAAATTCGAAATCCCGGTTTGTCCGCACGCTGGCGGTGTTGGC
CTGTGCGAACTGGTTCAGCACCTGATTATCTTCGATTATATCTCTGTGTCTGCT
TCTCTGGAAACCGTGTTTGCGAATACGTTGATCACCTGCATGAACACTTCAA
ATATCCGGTAATGATCCAGCGTGCGTCCTACATGCCGCCAAAAGATCCGGGC
TATAGCACTGAAATGAAAGAAGAATCCGTTAA-3´

rTSβ

5´-ATGCACACTGACCCGGACTACTCTGCGGCGTACGTTGTTATCGAAACCGA
CGCTGAAGACGGTATCAAAGGCTGCGGTATCACCTTCACCCTGGGTAAAGGT
ACCGAAGTTGTTGTGTGCGCTGTTAACGCTCTGGCGCACCACGTTCTGAACAA
AGATCTGAAAGACATCGTTGGTGATTTCCGTGGTTTCTATCGTCAGCTGACCT
CTGATGGTCAGCTGCGTTGGATCGGCCCGGAAAAAGGTGTTGTTCACCTGGC
TACCGCGGCTGTTCTGAACGCAGTTTGGGACCTGTGGGCTAAACAGGAAGGT
AAACCGGTTTGGAAACTGCTGGTAGATATGGATCCGCGTATGCTGGTTTCTTG
TATTGACTTCCGTTACATCACCGATGTTCTGACCGAAGAAGATGCACTGGAA
ATCCTGCAGAAAGGTCAGATCGGTAAAAAGAACGTGAAAACAGATGCTG
GCTCAGGGCTACCCGGCTTACACCACTTCCTGCGCATGGCTGGGTTACTCTGA
TGACACTCTGAAACAGCTGTGCGCGCAGGCTCTGAAAGATGGTTGGACTCGT
TTCAAAGTTAAAGTTGGTGCTGATCTGCAGGATGACATGCGTCGTTGCCAGAT
CATTCGTGATATGATCGGTCCTGAAAAAACCCTGATGATGGATGCTAACCAG

190

CGCTGGGACGTTCCTGAAGCTGTTGAATGGATGAGCAAACTGGCTAAATTCA
AACCACTGTGGATCGAAGAACCGACCTCTCCGGATGATATCCTTGGCCACGC
AACTATCTCTAAAGCGCTGGTTCCGCTGGGTATCGGTATTGCTACCGGTGAAC
AGTGCCACAACCGTGTTATCTTTAAACAACTGCTGCAAGCAAAGGCACTGCA
GTTCCTGCAGATCGACTCCTGCCGTCTGGGTAGCGTTAACGAAAACCTGTCTG
TTCTGCTGATGGCGAAAAAATTCGAAATCCCGGTTTGCCCGCATGCTGGCGGT
GTTGGTCTGTGTGAACTGGTTCAGCACCTGATTATCTTTGACTACATCTCCGTT
TCTGCTTCTCTGGAAAACCGTGTTTGCGAATATGTTGACCACCTGCACGAACA
CTTCAAATACCCGGTTATGATCCAGCGCGCGTCTTACATGCCGCCGAAAGAT
CCGGGTTACTCTACCGAAATGAAAGAAGAATCTGTGAAAAAACACCAGTACC
CGGATGGTGAAGTTTGTAA-3´

rTSγ

5´-ATGGTCCGTGGTCGTATCTCTCGTCTGTCTGTTCGTGACGTTCGTTTCCCGA
CCTCCCTGGGTGGTCACGGTGCGGACGCAATGCACACCGACCCGGACTACTC
TGCGGCGTATGTTGTTATCGAAACCGACGCAGAAGACGGCATCAAAGGCTGC
GGTATCACCTTCACCCTGGGTAAAGGCACCGAAGTGGTTGTTTGCGCTGTTAA
CGCACTGGCGCACCACGTACTGAACAAAGATCTGAAAGATATCGTTGGCGAT
TTCCGTGGTTTCTACCGTCAGCTGACCTCTGACGGCCAGCTGCGTTGGATCGG
CCCGGAAAAAGGCGTTGTTCACCTGGCTACCGCGGCAGTGCTGAACGCGGTT
TGGGATCTGTGGGCGAAACAGGAAGGTAAACCGGTTTGGAAACTGCTGGTAG
ACATGGACCCGCGTATGCTGGTTAGCTGCATCGACTTCCGTTACATCACCGAC
GTTCTGACTGAAGAAGACGCCCTGGAAATCCTGCAGAAAGGTCAGATCGGCA
AAAAGAACGTGAAAAACAGATGCTGGCTCAGGGTTACCCGGCCTACACCA
CCTCCTGTGCGTGGCTGGGCTACTCTGATGACACCCTGAAACAGCTGTGCGC
GCAGGCGCTGAAAGATGGTTGGACCCGTTTCAAAGTGAAAGTTGGTGCGGAT
CTGCAGGACGATATGCGTCGTTGTCAGATCATCCGTGACATGATCGGTCCGG
AAAAAACCCTGATGATGGACGCGAACCAGCGCTGGGATGTTCCGGAAGCGGT
TGAATGGATGTCCAAACTGGCTAAATTCAAACCGCTGTGGATCGAAGAACCG
ACCTCTCCAGACGACATCCTGGGTCACGCGACCATCTCCAAAGCTCTGGTTCC
GCTGGGCATCGGCATCGCGACCGGTGAACAGTGCCACAACCGTGTTATCTTC
AAACAGCTGCTGCAGGCGAAAGCACTGCAGTTCCTGCAGATTGATAGCTGCC
GCCTGGGCTCCGTTAACGAAAACCTGTCTGTGCTGCTGATGGCTAAAAAATTC
GAAATCCCGGTTTGTCCGCACGCGGGTGGCGTGGGTCTGTGCGAACTGGTTC
AGCACTTGATCATCTTCGATTACATCTCTGTGTCTGCTAGTCTTGAAAACCGT
GTCTGTGAATATGTTGACCACCTGCATGAACACTTTAAATATCCGGTGATGAT
CCAACGCGCGTCGTACATGCCGCCGAAAGATCCAGGCTATTCTACTGAAATG
AAAGAAGAGTCTGTTAA-3´

rTSγ-Xa

5´-ATGGTCCGTGGTCGTATCTCTCGTCTGTCTGTTCGTGACGTTCGTTTCCCGA
CCTCCCTGGGTGGTCACATTGAAGGCCGCATGCACACCGACCCGGACTACTC
TGCGGCGTATGTTGTTATCGAAACCGACGCAGAAGACGGCATCAAAGGCTGC

GGTATCACCTTCACCCTGGGTAAAGGCACCGAAGTGGTTGTTTGCGCTGTTAA
CGCACTGGCGCACCACGTACTGAACAAAGATCTGAAAGATATCGTTGGCGAT
TTCCGTGGTTTCTACCGTCAGCTGACCTCTGACGGCCAGCTGCGTTGGATCGG
CCCGGAAAAAGGCGTTGTTCACCTGGCTACCGCGGCAGTGCTGAACGCGGTT
TGGGATCTGTGGGCGAAACAGGAAGGTAAACCGGTTTGGAAACTGCTGGTAG
ACATGGACCCGCGTATGCTGGTTAGCTGCATCGACTTCCGTTACATCACCGAC
GTTCTGACTGAAGAAGACGCCCTGGAAATCCTGCAGAAAGGTCAGATCGGCA
AAAAAGAACGTGAAAAACAGATGCTGGCTCAGGGTTACCCGGCCTACACCA
CCTCCTGTGCGTGGCTGGGCTACTCTGATGACACCCTGAAACAGCTGTGCGC
GCAGGCGCTGAAAGATGGTTGGACCCGTTTCAAAGTGAAAGTTGGTGCGGAT
CTGCAGGACGATATGCGTCGTTGTCAGATCATCCGTGACATGATCGGTCCGG
AAAAAACCCTGATGATGGACGCGAACCAGCGCTGGGATGTTCCGGAAGCGGT
TGAATGGATGTCCAAACTGGCTAAATTCAAACCGCTGTGGATCGAAGAACCG
ACCTCTCCAGACGACATCCTGGGTCACGCGACCATCTCCAAAGCTCTGGTTCC
GCTGGGCATCGGCATCGCGACCGGTGAACAGTGCCACAACCGTGTTATCTTC
AAACAGCTGCTGCAGGCGAAAGCACTGCAGTTCCTGCAGATTGATAGCTGCC
GCCTGGGCTCCGTTAACGAAAACCTGTCTGTGCTGCTGATGGCTAAAAAATTC
GAAATCCCGGTTTGTCCGCACGCGGGTGGCGTGGGTCTGTGCGAACTGGTTC
AGCACTTGATCATCTTCGATTACATCTCTGTGTCTGCTAGTCTTGAAAACCGT
GTCTGTGAATATGTTGACCACCTGCATGAACACTTTAAATATCCGGTGATGAT
CCAACGCGCGTCGTACATGCCGCCGAAAGATCCAGGCTATTCTACTGAAATG
AAAGAAGAGTCTGTTAAAAAACACCAGTACCCGGATGGTGAAGTTTGGAAA
AAACTGCTGCCGGCGCAGGAAAACGCCGGGAGCGCTTGGAGCCACCCGCAG
TTCGAAAAATAA-3′

rTSγ–TEV

5′-ATGGTCCGTGGTCGTATCTCTCGTCTGTCTGTTCGTGACGTTCGTTTCCCGA
CCTCCCTGGGTGGTCACGAAAACCTGTATTTTCAGGGCATGCACACCGACCC
GGACTACTCTGCGGCGTATGTTGTTATCGAAACCGACGCAGAAGACGGCATC
AAAGGCTGCGGTATCACCTTCACCCTGGGTAAAGGCACCGAAGTGGTTGTTT
GCGCTGTTAACGCACTGGCGCACCACGTACTGAACAAAGATCTGAAAGATAT
CGTTGGCGATTTCCGTGGTTTCTACCGTCAGCTGACCTCTGACGGCCAGCTGC
GTTGGATCGGCCCGGAAAAAGGCGTTGTTCACCTGGCTACCGCGGCAGTGCT
GAACGCGGTTTGGGATCTGTGGGCGAAACAGGAAGGTAAACCGGTTTGGAA
ACTGCTGGTAGACATGGACCCGCGTATGCTGGTTAGCTGCATCGACTTCCGTT
ACATCACCGACGTTCTGACTGAAGAAGACGCCCTGGAAATCCTGCAGAAAGG
TCAGATCGGCAAAAAGAACGTGAAAAACAGATGCTGGCTCAGGGTTACCCC
GGCCTACACCACCTCCTGTGCGTGGCTGGGCTACTCTGATGACACCCTGAAA
CAGCTGTGCGCGCAGGCGCTGAAAGATGGTTGGACCCGTTTCAAAGTGAAAG
TTGGTGCGGATCTGCAGGACGATATGCGTCGTTGTCAGATCATCCGTGACATG
ATCGGTCCGGAAAAAACCCTGATGATGGACGCGAACCAGCGCTGGGATGTTC
CGGAAGCGGTTGAATGGATGTCCAAACTGGCTAAATTCAAACCGCTGTGGAT
CGAAGAACCGACCTCTCCAGACGACATCCTGGGTCACGCGACCATCTCCAAA
GCTCTGGTTCCGCTGGGCATCGGCATCGCGACCGGTGAACAGTGCCACAACC

GTGTTATCTTCAAACAGCTGCTGCAGGCGAAAGCACTGCAGTTCCTGCAGATT
GATAGCTGCCGCCTGGGCTCCGTTAACGAAAACCTGTCTGTGCTGCTGATGGC
TAAAAAATTCGAAATCCCGGTTTGTCCGCACGCGGGTGGCGTGGGTCTGTGC
GAACTGGTTCAGCACTTGATCATCTTCGATTACATCTCTGTGTCTGCTAGTCTT
GAAAACCGTGTCTGTGAATATGTTGACCACCTGCATGAACACTTTAAATATCC
GGTGATGATCCAACGCGCGTCGTACATGCCGCCGAAAGATCCAGGCTATTCT
ACTGAAATGAAAGAAGAGTCTGTTAAAAAACACCAGTACCCGGATGGTGAA
GTTTGGAAAAACTGCTGCCGGCGCAGGAAAACGGGGAGCGCTTGGAGCCA
CCCGCAGTTCGAAAAATAA-3′

MBP–TEV–rTSγ

5′-ATGAAAATAAAAACAGGTGCACGCATCCTCGCATTATCCGCATTAACGAC
GATGATGTTTTCCGCCTCGGCTCTCGCCAAAATCGAAGAAGGTAAACTGGTA
ATCTGGATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGA
AATTCGAGAAAGATACCGGAATTAAAGTCACCGTTGAGCATCCGGATAAACT
GGAAGAGAAATTCCCACAGGTTGCGGCAACTGGCGATGGCCCTGACATTATC
TTCTGGGCACACGACCGCTTTGGTGGCTACGCTCAATCTGGCCTGTTGGCTGA
AATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGGAT
GCCGTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTT
ATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA
GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAGGTAAGAGCGCGCTG
ATGTTCAACCTGCAAGAACCGTACTTCACCTGGCCGCTGATTGCTGCTGACGG
GGGTTATGCGTTCAAGTATGAAAACGGCAAGTACGACATTAAAGACGTGGGC
GTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCTGGTTGACCTGATTA
AAAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAAGCTGCCTT
TAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAA
CATCGACACCAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAG
GGTCAACCATCCAAACCGTTCGTTGGCGTGCTGAGCGCAGGTATTAACGCCG
CCAGTCCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGAC
TGATGAAGGTCTGGAAGCGGTTAATAAAGACAAACCGCTGGGTGCCGTAGCG
CTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATCCACGTATTGCCGCCACTA
TGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCGCAGATGTCCGC
TTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTCGTCAGA
CTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCGAACAACAA
CAACAATAACAATAACAACAACCCCGGGAAAACCTGTATTTTCAGGGCATGC
ACACCGACCCGGACTACTCTGCGGCGTATGTTGTTATCGAAACCGACGCAGA
AGACGGCATCAAAGGCTGCGGTATCACCTTCACCCTGGGTAAAGGCACCGAA
GTGGTTGTTTGCGCTGTTAACGCACTGGCGCACCACGTACTGAACAAAGATCT
GAAAGATATCGTTGGCGATTTCCGTGGTTTCTACCGTCAGCTGACCTCTGACG
GCCAGCTGCGTTGGATCGGCCCGGAAAAAGGCGTTGTTCACCTGGCTACCGC
GGCAGTGCTGAACGCGGTTTGGGATCTGTGGGCGAAACAGGAAGGTAAACC
GGTTTGGAAACTGCTGGTAGACATGGACCCGCGTATGCTGGTTAGCTGCATC
GACTTCCGTTACATCACCGACGTTCTGACTGAAGAAGACGCCCTGGAAATCC
TGCAGAAAGGTCAGATCGGCAAAAAGAACGTGAAAAACAGATGCTGGCTC

AGGGTTACCCGGCCTACACCACCTCCTGTGCGTGGCTGGGCTACTCTGATGAC
ACCCTGAAACAGCTGTGCGCGCAGGCGCTGAAAGATGGTTGGACCCGTTTCA
AAGTGAAAGTTGGTGCGGATCTGCAGGACGATATGCGTCGTTGTCAGATCAT
CCGTGACATGATCGGTCCGGAAAAAACCCTGATGATGGACGCGAACCAGCGC
TGGGATGTTCCGGAAGCGGTTGAATGGATGTCCAAACTGGCTAAATTCAAAC
CGCTGTGGATCGAAGAACCGACCTCTCCAGACGACATCCTGGGTCACGCGAC
CATCTCCAAAGCTCTGGTTCCGCTGGGCATCGGCATCGCGACCGGTGAACAG
TGCCACAACCGTGTTATCTTCAAACAGCTGCTGCAGGCGAAAGCACTGCAGT
TCCTGCAGATTGATAGCTGCCGCCTGGGCTCCGTTAACGAAAACCTGTCTGTG
CTGCTGATGGCTAAAAAATTCGAAATCCCGGTTTGTCCGCACGCGGGTGGCG
TGGGTCTGTGCGAACTGGTTCAGCACTTGATCATCTTCGATTACATCTCTGTG
TCTGCTAGTCTTGAAAACCGTGTCTGTGAATATGTTGACCACCTGCATGAACA
CTTTAAATATCCGGTGATGATCCAACGCGCGTCGTACATGCCGCCGAAAGAT
CCAGGCTATTCTACTGAAATGAAAGAAGAGTCTGTTAA-3′