Purification of A-Raf and

Structural Studies of Mannitol Dehydrogenase

A Thesis Submitted to the College of

Graduate Studies and Research

in Partial Fulfillment of the Requirements

for the Degree of Master of Science

in the Department of Biochemistry

University of Saskatchewan

Saskatoon

By

Jennifer Puttick

© Copyright Jennifer Puttick, September 2008. All rights reserved.

Copyright Note

I hereby grant the University of Saskatchewan and/or its agents the non-exclusive license to archive and make accessible, under the conditions specified below, my thesis, dissertation, or project report in whole or in part in all forms of media, now or for the duration of my copyright ownership. I retain all other ownership rights to the copyright of the thesis, dissertation or project report. I also reserve the right to use in future works (such as articles or books) all or part of this thesis, dissertation, or project report.

I hereby certify that, if appropriate, I have obtained and attached hereto a written permission statement from the owner(s) of each third party copyrighted matter that is included in my thesis, dissertation, or project report, allowing distribution as specified below. I certify that the version I submitted is the same as that approved by my advisory committee.

Abstract

The work herein describes the research of two separate projects: the purification of A-Raf and the crystallization and X-ray diffraction of *Thermotoga maritima* mannitol dehydrogenase (TmMtDH).

A-Raf is a one of three Raf isoforms of serine/threonine kinases involved in the mitogen-activated protein kinase (MAPK) pathway, a cell proliferation pathway that has been associated with many cancers. In addition, only the A-Raf isoform can uniquely bind to the regulatory subunit of phosphatidylinositol-3-kinase (PI3K), which is part of the Akt/PI3K pathway and is another important signaling molecule deregulated in human cancers. Therefore, the main focus of this study was to purify and crystallize this protein in order to characterize what makes A-Raf structurally unique from the other two Raf isoforms.

Several portions of A-Raf were purified throughout this study, but most research concentrated on the conserved region 2 and 3 (CR2 and CR3) domains of A-Raf and the full-length protein. The CR2/CR3 domains and full-length A-Raf were purified by affinity chromatography on a glutathione Sepharose column and column fractions were analyzed by SDS-PAGE. Two different bands measuring approximately 75 kDa and 66 kDa resolved on the SDS-PAGE gel of full-length A-Raf while three bands measuring approximately 75 kDa, 66 kDa and 45 kDa resolved on the SDS-PAGE gel of the CR2/CR3 domains of A-Raf. The CR2/CR3 domains and full-length A-Raf were also extensively studied by mass spectrometry but results were inconclusive. Western blot analysis was also performed on the CR2/CR3 domains and full-length A-Raf. Results indicated that multiple bands were present and that degradation of the protein had taken place. A-Raf was thus deemed unsuitable for crystallization trials and the project was terminated.

Mannitol is an acyclic polyalcohol and is used commercially for several purposes including acting as an osmoregulatory compound in several pharmaceuticals and as an artificial sweetener in products targeted for diabetic patients. Commercially, mannitol is produced by the hydrogenation of 50% fructose/50% glucose syrup at high temperatures. However, the product of this process yields an excess of sorbitol and therefore the mannitol requires further purification. Mannitol dehydrogenase catalyzes the conversion of D-fructose to D-mannitol and has therefore been targeted for studies to produce a commercial mannitol bioreactor.

The aims of this study included crystallization of the hyperthermophilic *Thermotoga maritima* (TmMtDH) mannitol dehydrogenase and subsequent X-ray diffraction and structure analysis. Dr. Claire Vieille at Michigan State University provided purified protein for crystallization trials. Two conditions produced diffraction quality crystals of TmMtDH. Condition 1 crystals grew in a solution containing 30% 2-methyl-2,4-pentanediol (MPD) plus 0.1 M HEPES-Na at pH 7.5. Condition 2 crystals grew in a solution containing 15-20% (w/v) polyethylene glycol (PEG) 4000 or 8000 plus 0.1 M sodium citrate at pH 4, 0.2 M sodium bromide and 10% glycerol. Crystals were flash cooled in liquid nitrogen and diffracted on the in-house diffractometer at the Saskatchewan Structural Sciences Center and at beamline 08ID-1 at the Canadian Light Source. Data were collected to 3.3 Å for the crystal that grew in condition 1 but the structure could not be solved before the completion of this project. The space group of the condition 1 crystal was $P2_12_12_1$ with unit cell dimensions a = 83.43 Å, b = 120.61 Å, c = 145.76 Å.

Acknowledgements

I would like to sincerely thank Dr. Louis Delbaere for his guidance and support throughout my years working for him, first as a summer student, then as a research associate and finally as a graduate student. Thank you to all of the people in Dr. Delbaere's lab that helped answer my many questions and helped me learn the ins and outs of crystallography, including but not limited to Sanjukta Aich, Yvonne Leduc, Lata Prasad, and Kurt Nienaber. I would also like to thank the members of my committee for their patience and support in letting me choose a new career path in veterinary medicine while also trying to finish my degree in biochemistry. I would especially like to thank Dr. Deborah Anderson for her guidance with the A-Raf product and her determination to help me finish my project and thesis.

I would also like to thank my family and friends for listening to me while I was frustrated (even if they didn't completely understand what I was talking about) and for cheering with me when I finally found protein crystals (even if they didn't really think it was that exciting). Your enthusiasm and advice was always much appreciated!

Dedication

I dedicate this work to my husband Colin, the formatting genius, without whom this thesis would be a complete disaster. Thank you for putting up with my grumpy moods and helping me put everything back together – I love you! I also dedicate this work to my parents for their enormous gift of education. Thank you for your support, encouragement and guidance, and for teaching me the importance of honour.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AbMtDH	*Agaricus bisporus* mannitol dehydrogenase
BCR	Breakpoint cluster region
CLS	Canadian Light Source
CR	Conserved region
CRD	Cysteine-rich domain
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
GST	Glutathione S-transferase
LDR	Long-chain dehydrogenase reductase
MAPK	Mitogen-activated protein kinase
MDR	Medium-chain dehydrogenase reductase
MEK	Mitogen-activated protein kinase/extracellular regulated kinase kinase
MPD	2-methyl-2,4-pentanediol
MS	Mass spectrometry
MtDH	Mannitol dehydrogenase
PDB	Protein Data Bank
PDK	Phosphoinositide-dependent kinase
PEG	Polyethylene glycol
PfMtDH	*Pseudomonas fluorescens* mannitol dehydrogenase
PH	Pleckstrin homology
PI3K	Phosphatidylinositol 3-kinase
PTB	Phosphotyrosine-binding
RBD	Ras binding domain
SDR	Short-chain dehydrogenase reductase
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SH2	Src homology 2
SH3	Src homology 3
SSSC	Saskatchewan Structural Sciences Centre
TmMtDH	*Thermotoga maritima* mannitol dehydrogenase

1.0 Introduction

1.1 A-Raf

1.1.1 Cell Proliferation Pathways

Information from external stimuli is transduced through cellular pathways that control cell growth, differentiation and apoptosis (Bivona and Philips, 2003). Receptors that regulate these processes span the plasma membrane, binding extracellular polypeptides such as growth factors, and transduce this information into the cells through a series of intracellular phosphorylation events (Schlessinger, 2000).

In general, growth signals must be transduced from the external cellular environment, across the plasma membrane to the cytoplasm and then to target organelles like the nucleus. Growth signals are transduced by growth factors, which are regulatory proteins that stimulate cell division and trigger intracellular signaling events. Cell proliferation pathways begin when growth factors in the external cellular environment bind to their receptor on the surface of the cell. This binding then triggers a cascade of phosphorylation events involving numerous proteins. In many instances proteins in these phosphorylation cascades are inactive when they are dephosphorylated and active in their phosphorylated state (Schlessinger, 2000). The protein activation from the phosphorylation cascade then eventually translocates to the target organelle inside the cell, e.g. the regulatory proteins in the nucleus that affect cellular growth (Chong *et al.,* 2003).

Cell proliferation pathways are therefore extremely complex. If genes for proteins in cell proliferation pathways contain mutations that result in their deregulation, this can cause uncontrolled growth of cells and give rise to cancers. It is for this reason that cell proliferation pathways have been studied considerably.

1.1.2 The MAPK and Akt/PI3K Cell Proliferation Pathways

Two cell proliferation pathways are of particular interest for this project: the mitogen-activated protein kinase (MAPK) pathway and the phosphatidylinositol 3-kinase (PI3K)/Akt pathway.

In general, these two cell proliferation pathways are initiated when growth factors are bound to their specific growth factor receptor that span the plasma membrane (Bivona & Philips, 2003; Chong *et al.*, 2003). These activated receptors then autophosphorylate on

tyrosine residues and bind a number of signaling proteins including protein kinases, lipid kinases, lipases, and adapter proteins. Through activation of the receptor kinase, the pathway then cycles through a cascade of phosphorylation events that eventually leads to a broad range of cellular events.

The MAPK pathway is initiated when growth factors are bound to their specific growth factor receptor, like protein tyrosine kinase receptors (PTKRs) that span the plasma membrane (Bivona & Philips, 2003). For example, epidermal growth factor (EGF) binds to epidermal growth factor receptor (EGFR), which is a tyrosine kinase (Cioffi *et al.,* 1997; Schlessiner, 2000). These activated receptors autophosphorylate on tyrosine residues and then bind Grb2, the adapter protein, through its Src homology 2 (SH2) domain (Burke *et al.*, 2001; Pawson, 2002). Grb2 is composed of a SH2 domain and two Src homology region 3 (SH3) domains. SH2 domains are composed of approximately 100 amino acid residues, are non-catalytic and bind to tyrosine phosphorylated sites on proteins (Sonyang *et al.*, 1993). SH3 domains are composed of approximately 50 amino acid residues, are non-catalytic and bind to proline-rich sequences (Yu *et al.,* 1994). Alternatively, Shc can act as an intermediate adaptor protein. Shc binds to the activated receptor through its phosphotyrosine-binding (PTB) domain, which causes the activated receptor to phosphorylate Shc on a tyrosine residue (Schlessinger and Lemmon, 2003). Consequently, as with direct binding to activated receptors, Grb2 then binds to phosphorylation sites on the intermediate adapter protein, Shc (Figure 1.1).

Ras is a low molecular weight GTPase that cycles between inactive (Ras-GDP) and active (Ras-GTP) forms. Specifically, Ras must be isoprenylated at the C-terminal end at a CAAX sequence in order to associate with the plasma membrane (Hancock *et al.*, 1990; Willumsen *et al.*, 1984; Stokoe *et al.,* 1994). SOS is a guanine nucleotide exchange factor for Ras that is constitutively associated with Grb2, the adapter protein previously described. The binding of Grb2 to a phosphorylated site on either EGFR or Shc at the plasma membrane therefore allows the localization of SOS to the plasma membrane. SOS then causes Ras to release GDP and bind GTP, resulting in its activation. Activated Ras can then bind effector molecules like Raf and PI3K (Hall, 1994; Rodriguez-Viciana *et al*., 1994).

The Raf kinases, A-Raf, B-Raf and Raf-1 (also known as c-Raf), are a family of serine/threonine kinases. Activated Ras-GTP binds the Ras binding and cysteine-rich domains

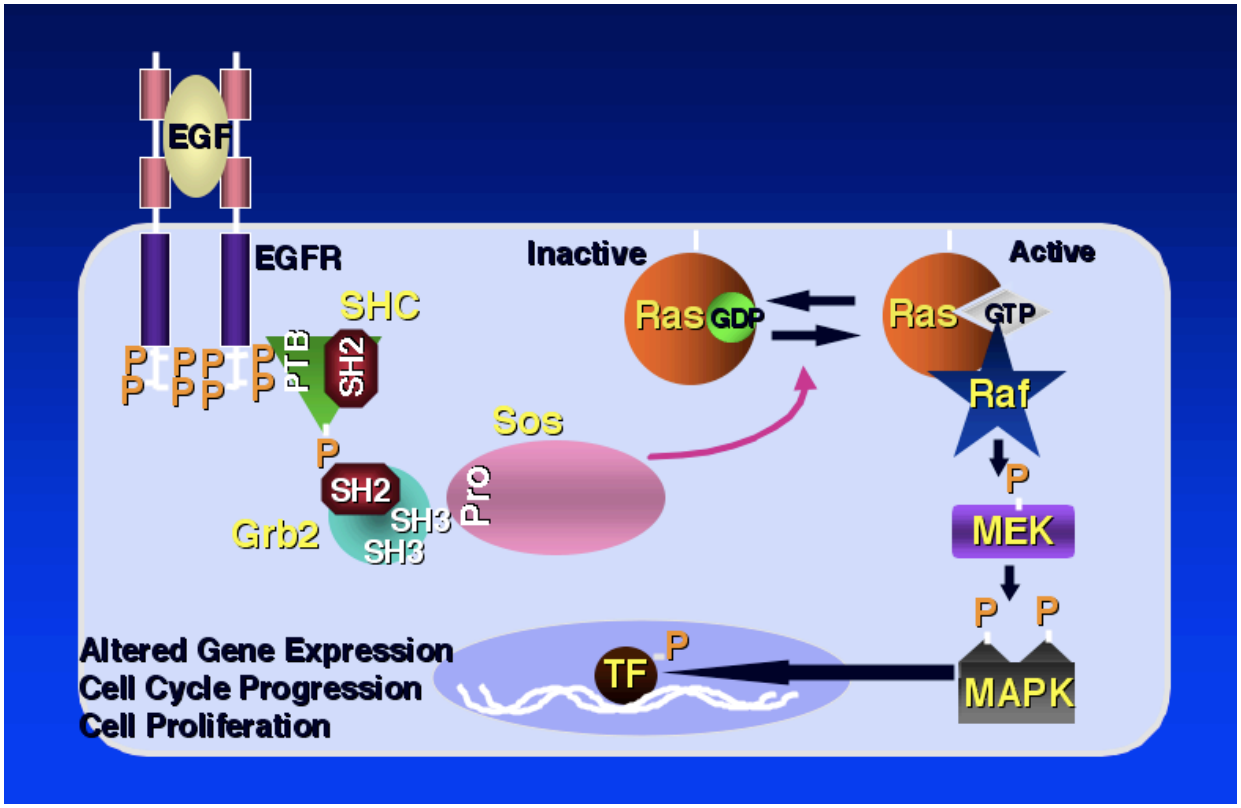


Figure 1.1 Schematic representation of the mitogen-activated protein kinase pathway. Binding of epidermal growth factor (EGF) to its specific receptor, epidermal growth factor receptor (EGFR), promotes the receptor autophosphorylation of tyrosine residues (Cioffi *et al.*, 1997). The phosphotyrosine-binding (PTB) domain of the intermediate adaptor protein Shc bind to the phosphotyrosine residues of the receptor, which in turn promotes the phosphorylation of tyrosine residues on Shc. The adapter protein Grb2 then binds Shc through its Src homology 2 (SH2) domains (Yu *et al.*, 1994). SOS, constitutively associated with the Src homology 3 (SH3) domains of Grb2, is now localized at the plasma membrane and can activate Ras-GDP to Ras-GTP (Schlessinger and Lemmon, 2003). Activated Ras-GTP then binds effector molecules like Raf, which can then phosphorylate mitogen-activated protein/extracellular regulated kinase-kinase (MEK) on serine residues (Daum *et al.*, 1994). MEK then phosphorylates mitogen-activated protein kinase (MAPK), allowing translocation into the nucleus for phosphorylation of transcription factors (TF) and target gene expression, cell cycle progression and cell proliferation (Schlessinger and Bar-Sagi, 1994). Image courtesy of Dr. Deborah Anderson, Saskatchewan Cancer Agency.

of Raf proteins (Marais *et al*., 1997; Pritchard *et al.,* 1995). Raf is activated by phosphorylation on tyrosine, serine and/or threonine residues by several kinases, to fully activate its kinase activity (Stokoe *et al.,* 1994). Mitogen-activated protein kinase/extracellular regulated kinase kinase (MEK) is also a protein kinase and is the substrate for Raf. Activated Raf can phosphorylate MEK1 and MEK2 on serine residues, which in turn phosphorylate MAPK on threonine and tyrosine residues (Gardner *et al.* 1994; Marais and Marshall, 1996; Yan and Templeton, 1994). Like MEK, MAPK is a general term for two kinases, MAPK1 and MAPK2. After translocation into the nucleus, the active phosphorylated MAPK then phosphorylates and activates various transcription factors required for growth, thereby completing the signaling cascade from extracellular signal to target organelle (Anderson, 2006; Schlessinger and Bar-Sagi, 1994; Kolch, 2000). A schematic representation of the MAPK pathway can be viewed in Figure 1.1.

The PI3K/Akt pathway is also initiated when growth factors, like EGF, bind to their specific growth factor receptors, like EGFR (Bjorge *et al.,* 1990). This leads to receptor activation and autophosphorylation on tyrosine residues, which then signals for the recruitment of PI3K to the plasma membrane (Cantley, 2002).

PI3K is a heterodimeric lipid kinase that is composed of an 85 kDa regulatory subunit (p85) and a 110 kDa catalytic subunit (p110). The p85 subunit mediates binding between activated tyrosine kinases and p110, relocalizing p110 to the plasma membrane. The p85 protein also has an inhibitory effect on p110 kinase activity in the absence of growth factor stimulation that is relieved upon stimulation (Luo and Cantley, 2005). The p85 subunit contains the following five domains: an SH3 domain, a breakpoint cluster region (BCR) homology domain flanked by two proline-rich regions, and two SH2 domains separated by an inter-SH2 (iSH2) domain (Figure 1.2). The BCR domain is a G protein binding domain (Chamberlain *et al*., 2004). Similar to the SH2 and SH3 domains of Grb2 as previously discussed, the SH2 and SH3 domains of p85 are involved in mediating protein-protein interactions. Previous studies have demonstrated that the SH3 domain of p85 can bind the proline-rich sequences flanking the BCR domain (Zheng *et al.,* 1994) and that the iSH2 domain of p85 binds tightly to the adapter binding domain (ABD) of the p110 catalytic domain (Yu *et al.*, 1998).

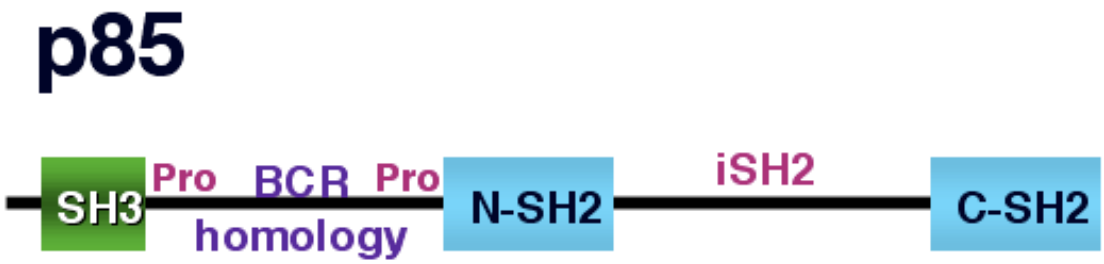


Figure 1.2 The domains of the p85 regulatory subunit of phosphatidylinositol 3-kinase. The p85 regulatory subunit of phosphatidylinositol 3-kinase (PI3K) contains several different important domains. The Src homology 3 (SH3) domain of p85 binds proline-rich sequences, including the proline-rich sequences flanking the breakpoint cluster region (BCR) homology domains. The BCR homology domain is a G-protein binding site (Musacchio *et al.*, 1996). The N-terminal and C-terminal Src homology 2 (SH2) domains bind phosphorylated tyrosine sites and help to relocalize the catalytic p110 subunit to the receptors at the plasma membrane (Daum *et al.*, 1994). There also exists a p110 binding site in the inter-SH2 (iSH2) region. Image courtesy of Dr. Deborah Anderson, Saskatchewan Cancer Agency.

The p110 catalytic subunit of PI3K is responsible for phosphorylation of phosphatidylinositol lipids at the D3 position and for the serine phosphorylation of proteins (Paez *et al.*, 2001). This PI3K phosphorylation results in the recruitment of pleckstrin homology (PH) domain-containing proteins to the plasma membrane. One important example of a PH domain-contain protein is Akt, a serine/threonine kinase. Once Akt has been recruited to the membrane, it is phosphorylated by phosphoinositide-dependent kinase 1 and 2 (PDK1 and PDK2). The resulting activated form of Akt can then phosphorylate many different targets and enhance protein synthesis and/or cell survival (Figure 1.3).

1.1.3 Raf Kinase Family

The Raf family of serine/threonine kinases consists of three isoforms: Raf-1, A-Raf and B-Raf (Daum *et al.*, 1994). These isoforms contain high sequence identity with each other but possess unique functions, including interactions with various proteins such as regulatory proteins, adapters, chaperones, and substrates. A sequence alignment of the three Raf isoforms can be found in Figure 1.4.

All Raf kinases share three highly conserved regions (CR) termed CR1, CR2 and CR3. CR1, consisting of amino acid residues 25 to 154 in A-Raf, is the most N-terminal domain. CR1 contains a Ras binding domain (RBD) and a cysteine-rich domain (CRD) with a zinc-finger motif capable of binding two molecules of zinc (Chong *et al.*, 2003). CR2, consisting of amino acid residues 210 to 225 in A-Raf, is a serine and threonine-rich domain in which many of these residues are thought to have regulatory functions (Pritchard *et al.,* 1995). The specific function of the CR2 domain remains unclear, although studies have confirmed that various protein-protein interactions through the CR2 domain seem to affect the activation and localization of the particular Raf (Chong *et al.*, 2002). CR3, consisting of amino acid residues 292 to 588 in A-Raf, is the most C-terminal domain and is also known as the catalytic kinase domain. The CR1 and CR2 domains regulate the function of the catalytic CR3 domain (Garnett *et al.*, 2005; Daum *et al.*, 1994).

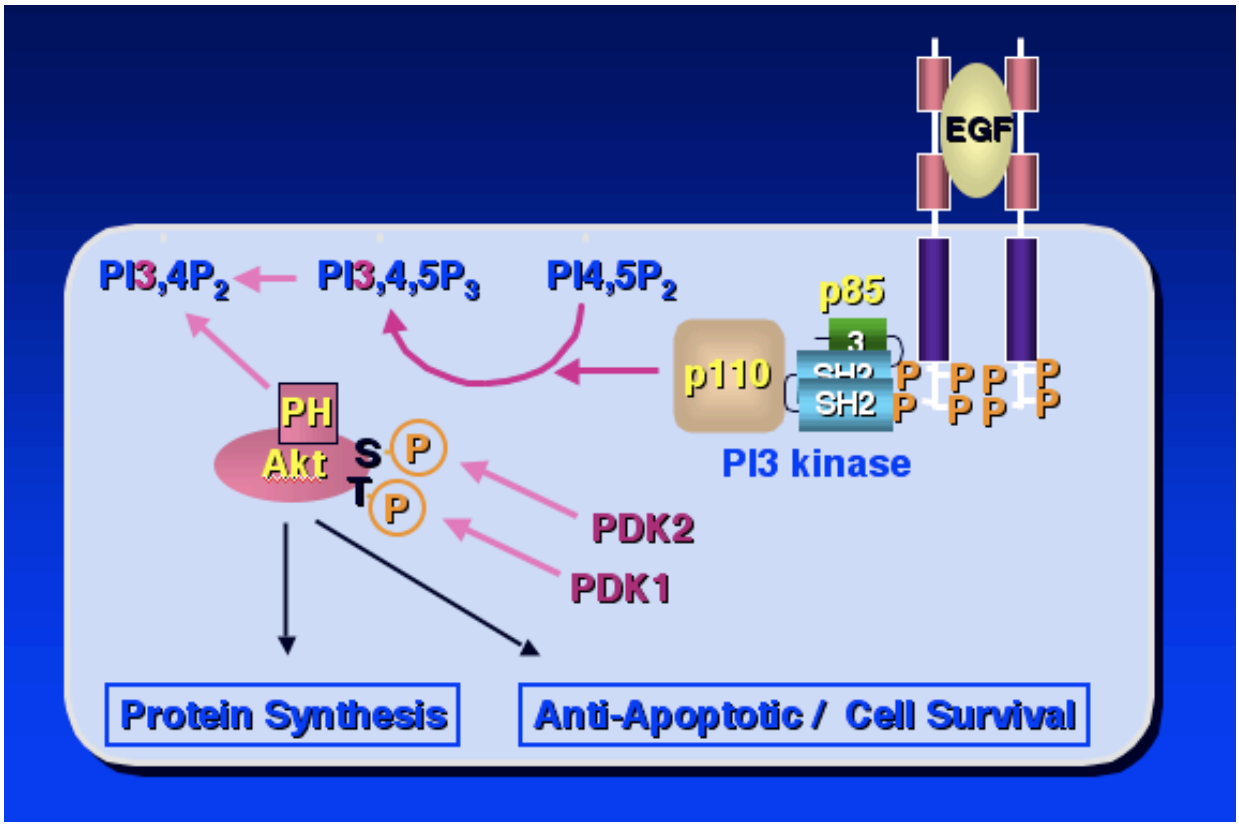


Figure 1.3 Schematic representation of the phosphatidylinositol 3-kinase/Akt pathway. Binding of epidermal growth factor (EGF) to the epidermal growth factor receptor (EGFR), promotes the autophosphorylation of tyrosine residues. The autophosphorylated residues act to recruit phosphatidylinositol 3-kinase (PI3K), which is composed of a regulatory p85 domain and a catalytic p110 domain, to the plasma membrane. The p110 subunit phosphorylates phosphatidylinositol lipids at the D3 position of the inositol ring. These lipids then recruit pleckstrin homology (PH) domain-containing proteins like Akt to the plasma membrane. Once at the membrane, Akt is phosphorylated by phosphoinositide-dependent kinase (PDK) and can then go on to activate effector molecules, resulting in protein synthesis and cell survival. Image courtesy of Dr. Deborah Anderson.

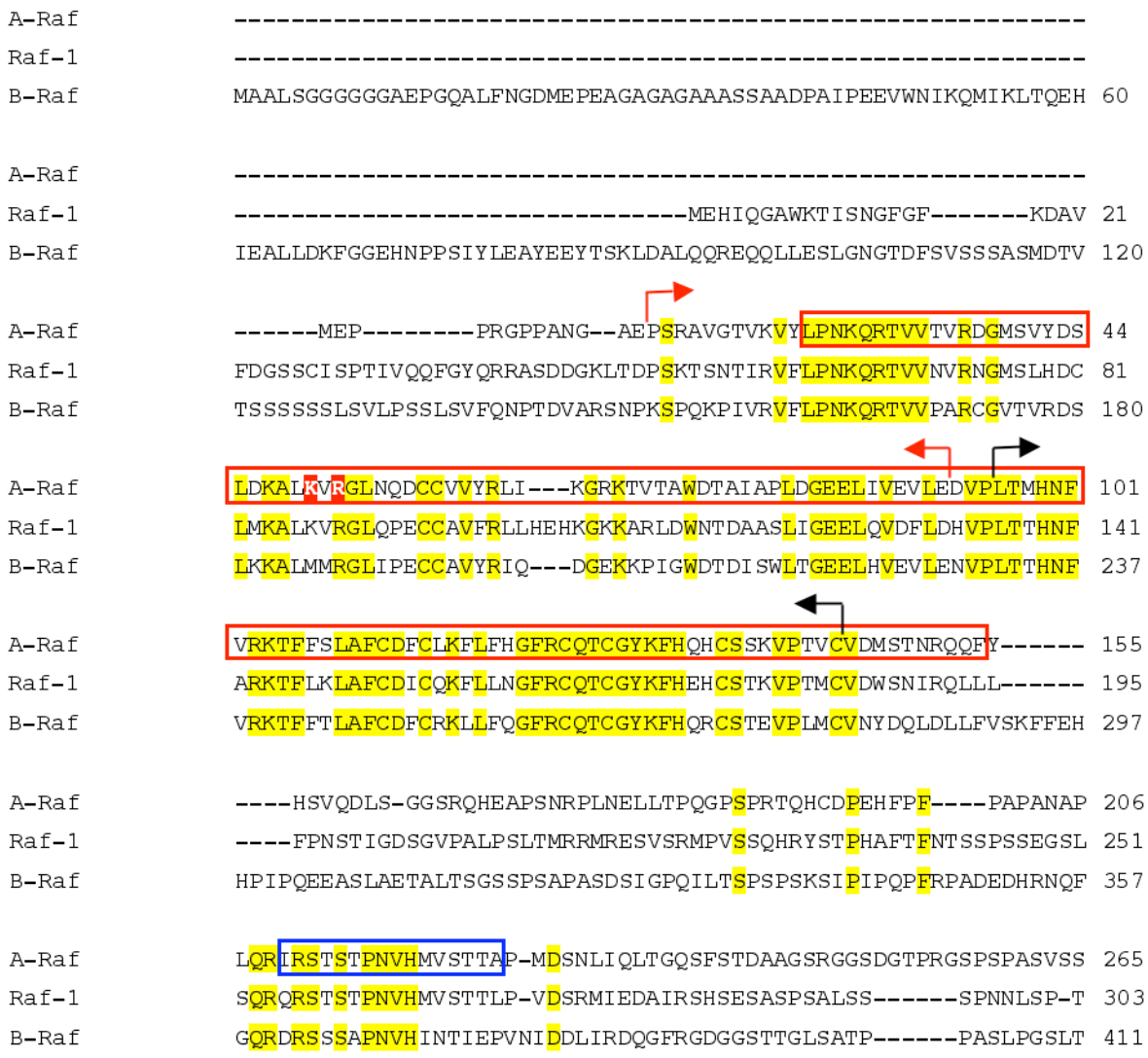

A-Raf ----------
Raf-1 ----------
B-Raf MAALSGGGGGGAEPGQALFNGDMEPEAGAGAGAAASSAADPAIPEEVWNIKQMIKLTQEH 60
A-Raf ----------
Raf-1 ----------MEHIQGAWKTISNGFGF-------KDAV 21
B-Raf IEALLDKFGGEHNPPSIYLEAYEEYTSKLDALQQREQQLLESLGNGTDFSVSSSASMDTV 120
A-Raf ------MEP--------PRGPPANG--AEPSRAVGTVKVYLPNKQRTVVTVRDGMSVYDS 44
Raf-1 FDGSSCISPTIVQQFGYQRRASDDGKLTDPSKTSNTIRVFLPNKQRTVVNVRNGMSLHDC 81
B-Raf TSSSSSSLSVLPSSLSVFQNPTDVARSNPKSPQKPIVRVFLPNKQRTVVPARCGVTVRDS 180
A-Raf LDKALKVRGLNQDCCVVYRLI---KGRKTVTAWDTAIAPLDGEELIVEVLEDVPLTMHNF 101
Raf-1 LMKALKVRGLQPECCAVFRLLHEHKGKKARLDWNTDAASLIGEELQVDFLDHVPLTTHNF 141
B-Raf LKKALMMRGLIPECCAVYRIQ---DGEKKPIGWDTDISWLTGEELHVEVLENVPLTTHNF 237
A-Raf VRKTFFSLAFCDFCLKFLFHGFRCQTCGYKFHQHCSSKVPTVCVDMSTNRQQFY------ 155
Raf-1 ARKTFLKLAFCDICQKFLLNGFRCQTCGYKFHEHCSTKVPTMCVDWSNIRQLLL------ 195
B-Raf VRKTFFTLAFCDFCRKLLFQGFRCQTCGYKFHQRCSTEVPLMCVNYDQLDLLFVSKFFEH 297
A-Raf ----HSVQDLS-GGSRQHEAPSNRPLNELLTPQGPSPRTQHCDPEHFPF----PAPANAP 206
Raf-1 ----FPNSTIGDSGVPALPSLTMRRMRESVSRMPVSSQHRYSTPHAFTFNTSSPSSEGSL 251
B-Raf HPIPQEEASLAETALTSGSSPSAPASDSIGPQILTSPSPSKSIPIPQPFRPADEDHRNQF 357
A-Raf LQRIRSTSTPNVHMVSTTAP-MDSNLIQLTGQSFSTDAAGSRGGSDGTPRGSPSPASVSS 265
Raf-1 SQRQRSTSTPNVHMVSTTLP-VDSRMIEDAIRSHSESASPSALSS------SPNNLSP-T 303
B-Raf GQRDRSSSAPNVHINTIEPVNIDDLIRDQGFRGDGGSTTGLSATP------PASLPGSLT 411

Figure 1.4 Sequence alignment of the three Raf isoforms.
Sequence alignment of the three Raf isoforms, A-Raf (ACC P10398), B-Raf (ACC P15056) and Raf-1 (ACC P04049). Residues highlighted in yellow are conserved in all three Raf isoforms. The conserved regions (CR) of A-Raf are shown in the red box (CR1), blue box (CR2) and green box (CR3)(Beck *et al.*, 1986). The black arrows denote the cysteine-rich domain of CR1 and the red arrows denotes the Ras-binding domain of CR1. Amino acid residues K50 and R52 (highlighted in red) are important for determining lipid binding specificity of A-Raf toward PI(3,4)P_2 (Johnson *et al.*, 2005). The acidic amino acids that raise the basal activity of B-Raf, D448 and D449 are highlighted in purple (Bosch *et al.*, 1997). The sequence alignment was performed by ClustalW (Thompson *et al.*, 1994).

Raf-1 is perhaps the best characterized of the three Raf isoforms. Raf-1 is a 73-kDa protein with 648 amino acid residues (Ghosh *et al.*, 1996; Pritchard *et al.,* 1995). B-Raf is an 84 kDa protein with 766 amino acid residues. A-Raf is a 68-kDa serine/threonine kinase with 606 amino acid residues (Beck *et al.,* 1987; Lee *et al.,* 1994). Like Raf-1, mutations in A-Raf have been characterized as oncogenic (Beck *et al.* 1987). However, A-Raf has been studied herein because of its ability to bind p85, the regulatory subunit of PI3K from the PI3K/Akt pathway (King *et al.*, 2000; Fang *et al.,* 2002). Further unique characteristics of A-Raf will be explored in subsequent sections.

1.1.4 Oncogenicity of B-Raf

B-Raf can become oncogenic and more than 30 mutations of the B-Raf gene have been implicated in human cancers (Wan *et al.*, 2004). An analysis of 22 of these mutations determined that while most mutations elevated the kinase activity of B-Raf towards its direct downstream target MEK *in vivo*, two mutations of the gene lowered kinase activity (Wan *et al.*, 2004). Studies have also shown that while B-Raf can be fully activated by oncogenic Ras alone, A-Raf and Raf-1 require synergistic activation of oncogenic Ras and other tyrosine kinases like Src (Marais *et al.,* 1997).

B-Raf is also unique in its regulation by tyrosine phosphorylation. A-Raf and Raf-1 contain conserved tyrosine residues within the CR3 domain that need to be phosphorylated by activated tyrosine kinases in order to become negatively charged and thus activated (Bosch *et al.*, 1997). Previous studies have shown that mutation of either of the corresponding tyrosine residues within both Raf-1 and A-Raf to aspartate residues can significantly increase the inactivities by mimicking the effects of phosphorylation (Fabian *et al*., 1993). B-Raf is the only Raf isoform that has been associated with human cancers and the potential for B-Raf to become oncogenic is due to 2 key residues: D448 and D449 (Bosch *et al.*, 1997; Figure 1.4). The acidic aspartate residues at these two positions allow B-Raf to maintain higher basal activity in an unphosphorylated state than A-Raf or Raf-1. Thus, the role of tyrosine phosphorylation in both A-Raf and Raf-1 appears to be crucial for activation and regulation, whereas tyrosine phosphorylation does not seem to be as important in B-Raf regulation.

1.1.5 Unique Characteristics of A-Raf

The specific roles of the three Raf isoforms are still unclear (Yuryev *et al.*, 2002). A-Raf is a 68-kDa kinase found ubiquitously in the body, except in the brain, but in the highest levels in urogenital tissues (Storm *et al.,* 1990; Mahon *et al*., 2005). A-Raf is the weakest Raf activator of MEK, the downstream target of Raf (Marais *et al*., 1997; Pritchard *et al*., 1995). A-Raf can localize to mitochondria and has therefore been implicated in apoptosis (Yuryev *et al.*, 2000). Unlike B-Raf, mutations in A-Raf and Raf-1 have not been detected in human cancers (Sridhar *et al.*, 2005). However, recent studies have demonstrated a unique role for A-Raf in cell proliferation (King *et al.*, 2000; Mahon *et al.,* 2005).

It was recently discovered that p85, the regulatory subunit of PI3K, could uniquely bind to A-Raf in a phosphorylation-independent manner through its two SH2 domains and one SH3 domain (King *et al*., 2000; Fang *et al*., 2002). Phosphorylation-dependent interactions with SH2 domains requires a pY-X-X-M target sequence, where pY denotes a phosphorylated tyrosine, X can be any amino acid residue, and M denotes a methionine residue (Mayer *et al.,* 1992; Eck *et al.,* 1993; Nantel *et al.,* 1998). Studies performed by Fang *et al.* (2002) demonstrated that p85 is still able to bind A-Raf if the Y500 residue, the only Y-X-X-M motif present in A-Raf, was mutated to phenylalanine and therefore could not be phosphorylated. This suggests A-Raf may have a role in both the MAPK and the PI3K/Akt cell proliferation pathways.

A-Raf also has unique lipid binding properties. A-Raf and Raf-1 can bind to certain monophosphorylated phosphatidylinositides (PI3P, PI4P, PI5P) and to phosphatidylinositol 3,5-bisphosphate ($PI3,5P_2$) (Johnson *et al.,* 2000). However, A-Raf also uniquely binds to two other phosphatidylinositol bisphosphates, $PI4,5P_2$, $PI3,4P_2$, whereas Raf-1 does not (Klarlund *et al.,* 1997; Johnson *et al.,* 2005). The binding of these lipids also increases the activity of A-Raf when assayed towards a GST-MEK substrate protein in the presence of micelles containing different lipid compositions (Klarlund *et al.*, 1997). Mutagenesis studies determined that the basic residues K50 and R52 (Figure 1.4) are important for determining lipid binding specificity of A-Raf toward $PI3,4P_2$ (Johnson *et al.*, 2005). This suggests that although three Raf isoforms exist in the body, A-Raf has unique functions.

1.1.6 Raf Regulation by Phosphorylation

The activity of all three Raf isoforms is regulated positively and negatively by different kinases. For example, activated tyrosine kinases have positive effects on Raf-1 and A-Raf but have no effect on B-Raf (Marais *et al*., 1995; Marais *et al*., 1997; Bosch *et al*., 1997). Protein kinase A has negative effects on Raf-1 by phosphorylating amino acid residues Ser43 and Ser233 (Damaz *et al*., 2002) but has positive effects on B-Raf and little or no effect on A-Raf (Schramm *et al*., 1994; Mischak *et al*., 1996). Finally, Akt has negative effects on Raf-1 and B-Raf (Zimmermann and Moelling, 1999; Reusch *et al*., 2001; Guan *et al*., 2000).

These kinases therefore phosphorylate the Raf isoforms on both stimulatory and inhibitory phosphorylation sites. For example, two inhibitory sites in Raf-1 include phosphorylation of Ser249 and Ser621, which allows the docking of 14-3-3 proteins and therefore stabilization of a closed and inactive Raf conformation (Yaffe *et al.*, 1997). On the other hand, phosphorylation of Ser471 in Raf-1 is critical for the docking of MEK. There also exists an activation segment in the activation loop of Raf-1 consisting of the amino acid sequence DFGLATVKSR. In this activation segment, amino acid residues T491 and S494 in Raf-1 are analogous to T599 and S602 in B-Raf (Zhang and Guan, 2000; Chong *et al*, 2001). The importance of analogous amino acid residues T452 and T455 in A-Raf are unknown.

Less is known about the regulation of A-Raf by phosphorylation. Like B-Raf and Raf-1, there also exists a potential docking site for 14-3-3 proteins in A-Raf consisting of phosphorylated amino acid residues Ser214 and Ser582 (Baljuls *et al.*, 2008). Phosphorylation of Ser432 is essential for the binding of MEK and therefore A-Raf signaling. The phosphorylation of the activation segment that is vital for Raf-1 and B-Raf activation appears to be important only for Ras-mediated stimulation of A-Raf and consists of the phosphorylation of Thr452 and Thr455 (Baljuls *et al.*, 2008). Finally, seven novel phosphorylation sites between amino acid residues 248 to 267 have recently been associated with positive regulation. These residues are called the isoform-specific hinge segment and involve the phosphorylation of Ser257, Ser262 and Ser264 (Baljuls *et al*., 2008).

1.1.7 Previously Characterized Raf Structures

Several NMR and X-ray crystallographic structures of Raf-1 have been characterized (Nassar, 1995; Mott *et al.*, 1996; Terada *et al.,* 1999). Emerson *et al.* (1995) solved the structure of the Ras-binding domain of human Raf-1 (amino acid residues 55-132) by NMR

techniques. The fold of the Ras-binding domain consists of a five-stranded beta sheet, a twelve-residue alpha-helix and an one-turn helix. The Ras-binding site is contained in the N-terminal beta-hairpin and the C-terminal end of the alpha-helix (Figure 1.5).

Terada *et al.* (1999) solved the structure of the Ras-binding domain of rat Raf-1 in complex with mutant Ras proteins by NMR techniques. These studies determined that amino acid residues 51-131 of Raf-1 and Glu31 of the Ras protein form a tight salt bridge that holds the complex in place. The authors also determined that Glu31 of the RBD of Raf is vitally important for the binding efficiency of Ras to Raf-1, as an E31K mutation of Ras reduced binding efficiency to Raf-1. The biological importance of this mutation is unknown.

B-Raf has also been crystallized with two different inhibitors: BAY43-9006 (purchased commercially and termed the ZM inhibitor in these studies) (Figure 1.6) and SB-590085. Using the BAY43-9006 inhibitor, structure analysis indicated a link between B-Raf and Raf-1, in that B-Raf mutants with decreased activity stimulate MEK by activation of endogenous Raf-1, possibly through an allosteric transphosphorylation mechanism (Wan *et al.*, 2004). However, studies involving B-Raf crystallized with the SB-590085 inhibitor suggested that this inhibitor stabilizes the kinase domain of B-Raf in an active conformation allowing for selective inhibition of the MAPK pathway (King *et al*., 2006).

1.1.6 Aims and Objectives

Although the three Raf isoforms contain high sequence identity and three highly conserved regions, each isoform seems to be unique. Therefore, the question remains as to why three isoforms are present in the body.

Several NMR and X-ray crystallographic structural studies have characterized the Ras binding domain of Raf-1 and B-Raf with and without selective inhibitors (Nassar, 1995; Mott *et al.*, 1996; Wan *et al.*, 2004). These structural studies have given insights into the distinct properties of Raf-1 and B-Raf, but at the start of these studies no structural data were available for A-Raf.

The studies reported herein were initiated in order to structurally characterize A-Raf and to shed light on its unique protein and lipid binding characteristics and regulation. The first aim of this study was to purify full-length A-Raf and various portions of single or multiple conserved regions of A-Raf for crystallization. A purification protocol for A-Raf was established based on previous studies and but was altered to obtain the 95% protein purity that

is best crystallization studies. The second aim of the study was to crystallize A-Raf. The final aim of the study was to diffract crystals of A-Raf at either an in-house X-ray source or at a synchrotron in order to obtain high quality data for structural determination.

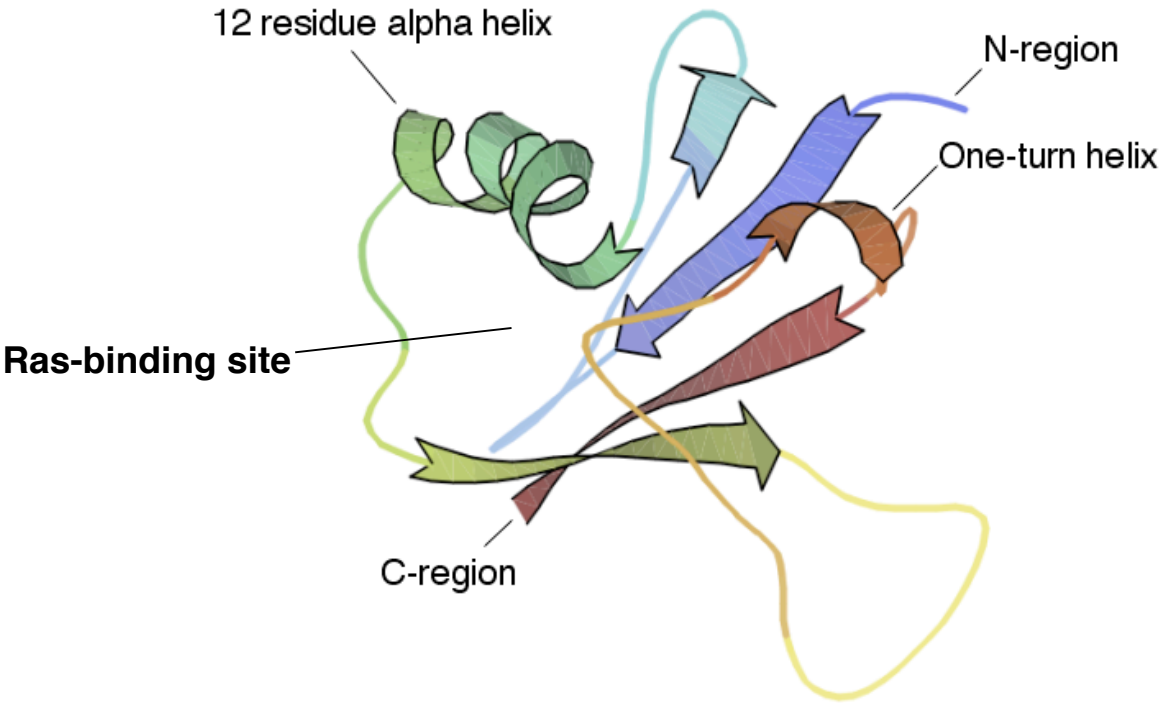


Figure 1.5 Structure solution of the Ras-binding domain of human Raf-1.
Using NMR techniques, Emerson *et al.* (1995) solved the structure of the Ras-binding domain (amino acid residues 55-132) of human Raf-1. The Ras-binding domain consists of a five-stranded beta sheet, a twelve residue alpha helix and a one-turn helix. The Ras-binding site, where the E31K mutation is located, is contained in the N-terminal beta-hairpin and the C-terminal end of the alpha helix (Protein Data Bank accession number 1RFA).

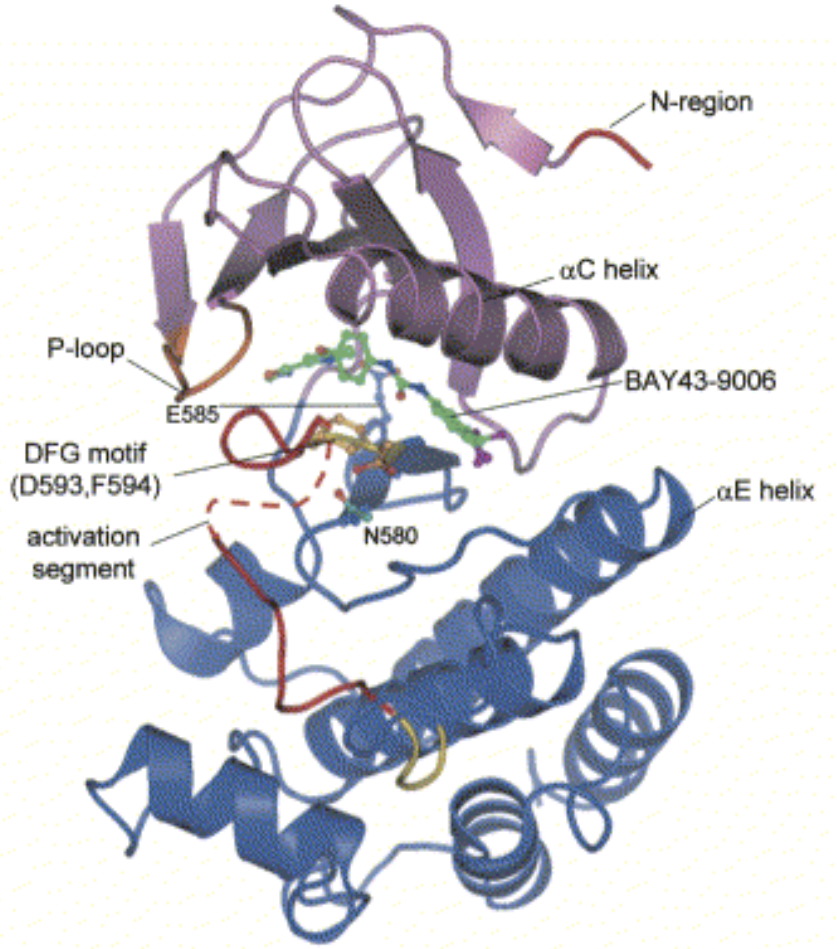


Figure 1.6 Structure of the kinase domain of wild-type B-Raf in complex with the Bay43-9006 inhibitor.
Wan *et al.* (2004) co-crystallized wild-type B-Raf in complex with a Raf inhibitor known as Bay43-9006. Analysis of the structure revealed an Asp-Phe-Gly (DFG) motif composed of amino acid residues D593, F594 and G595. Residues Asn580 and Glu585 are highlighted because they were mutated in other B-Raf studies by the authors. The activation segment of B-Raf is in red, the P loop is in orange and the N terminal and C terminal ends are in magenta and marine respectively (Protein Data Bank accession number 1UWH).

1.2 Mannitol Dehydrogenase

1.2.1 Commercial Applications of Mannitol

Mannitol, the alcohol form of mannose, is an acyclic six-carbon polyalcohol and one of the most abundant sugar alcohols in nature. Mannitol is found naturally in bacteria, algae, lichens, fungi and many plants. The natural functions of mannitol include acting as an osmoregulatory compound, acting as an efficient respiratory source during postharvest development, and accumulating during times of stress (Horer *et al.*, 2001).

During mannitol oxidation, NADH or NADPH is produced and can be used to synthesize ATP via the citric acid cycle. The oxidation of a NADH molecule results in the production of three molecules of ATP in the citric acid cycle. This results in a more efficient energy production system than in organisms that metabolize sugars such as sucrose because glycolysis generates only a net of two molecules of ATP.

Mannitol also has many commercial applications in food, pharmaceuticals, medicine and chemistry. For example, mannitol is absorbed poorly by the body and contains half the calories of sugar, making it a useful alternative sweetener for diabetic patients (Weymarn *et al.*, 2003). As a pharmaceutical, mannitol is used to reduce intracranial pressure, to treat patients with renal failure, and to promote the excretion of substances like aspirin and barbiturates in overdose situations (Lee *et al.*, 2003). Mannitol is further used as a specialty chemical in plastic manufacturing (Soetaert *et al.*, 1995)

An average of 50,000 tons of mannitol are produced annually for commercial uses. Commercially, mannitol is produced by the hydrogenation of 50% fructose/50% glucose syrup at high pressures and temperatures using an activated nickel catalyst (Kulbe *et al.*, 1987; Soetaert *et al.*, 1999). However, during this process sorbitol is produced in three-fold excess to mannitol, causing the need for further expensive downstream processing of the product (Kaup *et al.*, 2004). Therefore, engineered enzymes that can increase the rate of reaction and mannitol selectivity would be extremely useful to lower production costs and increase mannitol production (Kaup *et al.,* 2005). Consequently, many bacterial species utilizing the enzyme mannitol dehydrogenase are now being investigated for commercial applications.

1.2.2 Mannitol Dehydrogenase

Mannitol Dehydrogenase (MtDH) catalyzes the conversion of D-Fructose to D-Mannitol (Figure 1.7).

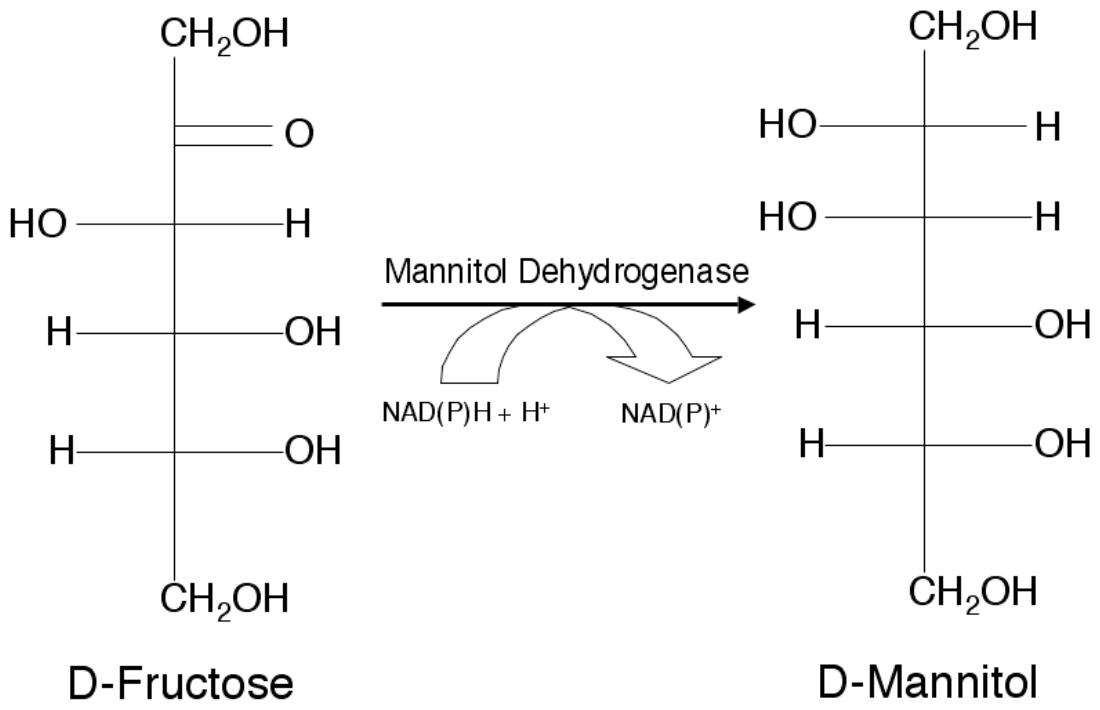


Figure 1.7 The reaction catalyzed by mannitol dehydrogenase.
Mannitol dehydrogenase catalyzes the conversion of D-Fructose to D-Mannitol using either NADP(H) or NAD(H) as a coenzyme.

In general, there exist three categories of MtDHs based on sequence analysis data, protein size and coenzyme-binding motifs: short-chain dehydrogenase reductases (SDRs), medium-chain dehydrogenase reductases (MDRs), and long-chain dehydrogenase reductases (LDRs). Fungal MtDHs belong to the SDR group and contain approximately 250 amino acid residues and do not require metals for catalysis (Horer *et al.,* 2001; Persson *et al.,* 1991). Many bacterial MtDHs belong to the MDR group. MDRs contain approximately 350 amino acid residues and are often zinc-dependent (Nordling *et al.,* 2002; Sasaki *et al.,* 2005). LDRs are generally monomeric, contain approximately 350-560 amino acid residues and do not require metals for catalysis (Kavanagh *et al.,* 2002; Klimacek & Nidetzky, 2002). Although it has been possible to roughly group the known MtDH species, little sequence identity exists between SDRs, MDRs and LDRs (Figure 1.8)

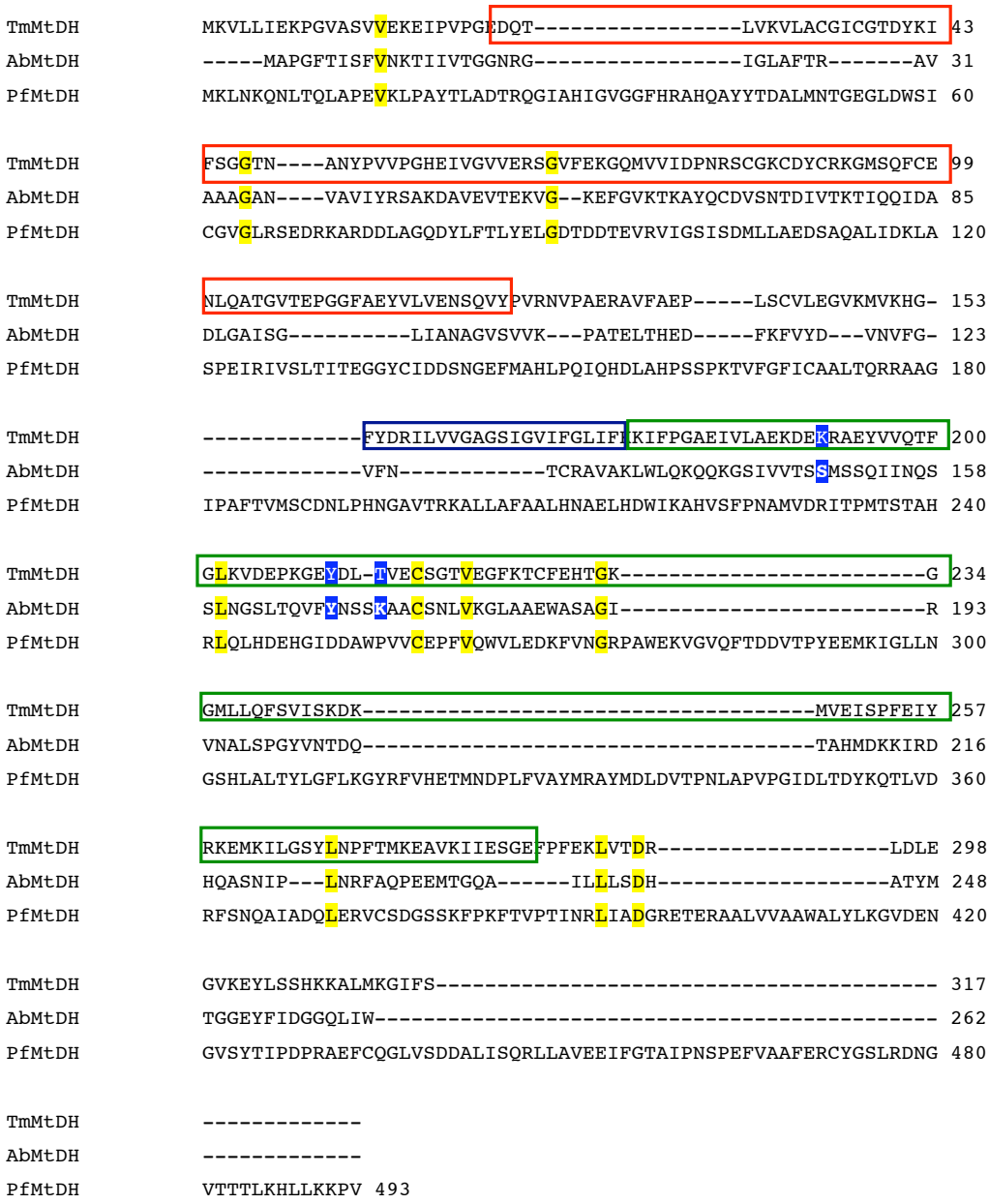


Figure 1.8 Sequence alignment of SDR, MDR and LDR groups.
Shown is a sequence alignment of one member from each of the short, medium and long-chain dehydrogenase reductase (SDR, MDR, LDR) groups. *Thermotoga maritima* mannitol dehydrogenase (TmMtDH, ACC Q9W9D4) belongs to the MDR group, *Agaricus bisporus* mannitol dehydrogenase (AbMtDH, ACC A93868) belongs to the SDR group and *Pseudomonas fluorescens* mannitol dehydrogenase (PfMtDH, ACC O18355) belongs to the LDR group. Sequence alignment was performed using ClustalW (Thompson *et al.*, 1994). Residues highlighted in yellow are common to TmMtDH, AbMtDH and PfMtDH. Pfam (Bateman *et al.,* 2004) predicts that TmMtDH will contain three structural domains: an alcohol dehydrogenase domain (red box), a transmembrane domain (blue box) and an alcohol dehydrogenase-zinc domain (green box). Structural analysis by X-ray crystallography of the SDR AbMtDH revealed a catalytic triad vital for catalysis involving amino acid residues Ser149, Tyr169 and Lys173 (bolded in blue).

For example, *Pseudomonas fluorescens* MtDH has been identified as an LDR by sequence similarity and contains 493 amino acid residues with a molecular mass of 54 kDa (Kavanagh *et al.,* 2002). LDR MtDHs were originally of significant interest because they had several unique potential uses as commercial mannitol producers. For example, clinically one can quantify the mannitol concentration in blood and urine to determine if a person has signs of diabetes (Kaup *et al.,* 2005) and LDR MtDHs could potentially produce mass quantities of mannitol for these studies. In addition, the transgenic expression of MtDHs in plants results in improved salt tolerance and resistance to oxidative stress, increasing their stability in drought conditions (Kavanagh *et al.*, 2002). Now it is understood that these unique properties extend to MDRs like *Thermotoga maritima* mannitol dehydrogenase (TmMtDH) and thus there has been a recent surge in bacterial MtDH characterization for these commercial applications.

On the other hand, many fungal MtDHs belong to the SDR group. MtDH from *Agaricus bisporus* is classified as a SDR. This enzyme contains only 262 amino acid residues and has a molecular mass of 28 kDa. TmMtDH is part of the MDR group and contains 317 amino acid residues with a molecular mass of 34 kDa. As previously mentioned, TmMtDH is a zinc-containing enzyme and, like other mannitol dehydrogenases, requires either NAD(H) or NADP(H) for activity. The unique characteristics of TmMtDH will be explored in subsequent sections but a sequence alignment of *Agaricus bisporus*, *Pseudomonas fluorescens* and *Thermotoga maritima* MtDH can be found in Figure 1.8.

1.2.3 Mannitol Dehydrogenase as a Commercial Bioreactor

Currently, there are three biological systems used to commercially produce mannitol: fermentation, membrane cell-recycle bioreactors and whole cell biotransformation (Klimacek *et al*., 2003). One of the laboratory processes for mannitol production to date uses the bacterium *Leuconostoc mesenteroides* which catalyzes the following reaction by fermentation:

glucose + 2 fructose ➜ 2 mannitol + lactic acid + acetic acid + CO_2

Figure 1.9 The fermentation reaction catalyzed by *Leuconostoc mesenteroides* mannitol dehydrogenase.

L. mesenteroides can produce up to 150 g/L of mannitol (Soetaert, 1999). The bacterium *Leuconostoc pseudomesenteroides* has been used to produce mannitol commercially in a membrane cell-recycle bioreactor system (Von Weymarn, 2002), however the mannitol yield during this process is quite low. A third commercial system used to produce mannitol involves a whole cell biotransformation process. During this process, extracellular glucose isomerase catalyzes the isomerization of glucose to fructose. Subsequently, a recombinant *Escherichia coli* strain expressing MtDH catalyzes the reduction of fructose to mannitol (Kaup, 2005). Unfortunately, this system requires precise pH control and the expensive petrochemical formate to recycle the nucleotide cofactor.

More recently, Baumchen and Bringer-Meyer (2007) discovered a recombinant oxidation/reduction cycle for the conversion of D-fructose to D-mannitol that allows for the recycling of the reduced nucleotide cofactor. In this method, the authors used *Corynebacterium glutamicium* cells transfected with an *fdh* gene from *Mycobacterium vaccae* N10 formate dehydrogenase to allow constant supply of NADH cofactor and thus mannitol production at a constant rate. This method was also successful when *fdh* was coexpressed in *Bacillus megaterium* (Baumchen *et al*., 2007).

Many believe that it is theoretically possible to convert 100% fructose into 100% mannitol using an immobilized enzyme system, which is how fructose syrup is commercially produced. Theoretically, mannitol could be produced by a similar system with the addition of an MtDH bioreactor that is thermostable and employs a nucleotide cofactor that is inexpensive and can be recycled.

1.2.4 *Thermotoga maritima* Mannitol Dehydrogenase

Thermotoga maritima Mannitol Dehydrogenase (TmMtDH) is part of the medium-chain dehydrogenase reductase group with a molecular mass of 34870 Da and 317 amino acid residues. TmMtDH is classified as an alcohol dehydrogenase (primary accession number Q9WYD4) and was chosen for study because of its thermostable characteristics. As previously discussed, an MtDH bioreactor used for the commercial production of mannitol must be thermostable and employ an inexpensive nucleotide cofactor that can be easily recycled. *Thermotoga maritima* is a hyperthermophilic bacterium that can resist 15-minute heat treatments at 85 ˚C (Figure 1.10). In addition, TmMtDH can also use either NAD(H) or

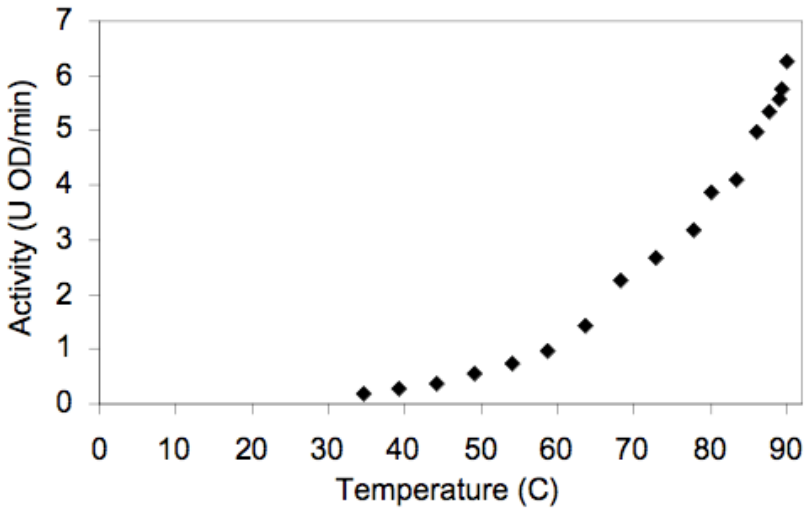


Figure 1.10 Thermostability of *Thermotoga maritima* mannitol dehydrogenase (TmMtDH) activity.
Dr. Claire Vieille, Michigan State University, performed a preliminary analysis of the activity of TmMtDH under various temperatures from 20 to 90 ˚C. The activity of TmMtDH was measured between 20 ˚ and 90 ˚C on a Cary 300 Bio UV/visible spectrophotometer (Varian Instruments, Walnut Creek, CA) equipped with a Peltier system. The fructose reduction reaction (1 mL) contained 50 mM sodium phosphate buffer at pH 6.1, 200 mM fructose, 0.3 mM NADH, and 5.46 μg TmMtDH. The reaction was initiated by adding the enzyme to the reaction mixture that had been preheated to 20 – 90 ˚C for 15 minutes. The mannitol oxidation reaction (1 mL) contained 50 mM glycine buffer at pH 8.3, 0.5 mM NAD^+, 200 mM mannitol, and 5.46 μg TmMtDH. Activity was recorded at 340 nm for 1 minute. One unit of activity was defined as the amount of MtDH needed to consume (fructose reduction) or produce (mannitol oxidation) 1 μmol/min of NADH. Reduction reactions with substrates other than fructose (200 mM D-xylose, D-xylulose, D-tagatose, L-sorbose, L-arabinose, L-threonine, acetaldehyde, or 2-butanone) were performed in the same conditions as for fructose. Oxidation reactions with substrates other than mannitol (with 200 mM sorbitol, xylitol, ethanol, or 2-butanol) were performed in the same conditions as for mannitol.

NADP(H) as a cofactor. NADP(H) is an expensive substrate and thus should not be used in a commercial setting if possible.

Sequence analyses by Pfam and InterPro (Bateman *et al.,* 2004; Apweiler *et al.,* 2001) predict that TmMtDH contains two domains: an NAD-binding domain and an alcohol dehydrogenase (ADH)-zinc domain (Figure 1.8). Pfam predicts that amino acid residues 24 to 126 will encode the ADH domain, amino acid residues 154 to 174 encode for a transmembrane domain, and residues 174 to 282 encode for an ADH-zinc domain. Pfam defines the ADH domain as the catalytic domain of the alcohol dehydrogenase. Many ADH domains, like that predicted for TmMtDH, contain an inserted zinc-binding domain and exhibit GroES-like structures (Jornvall *et al.,* 1987). GroES is chaparonin whose structure consists of a heptameric ring of identical 10 kDa monomers (Hunt *et al.*, 1996). Many alcohol dehydrogenases contain similar ring-like structures and are therefore called GroES-like. In general, zinc-containing enzymes are usually dimeric or tetrameric and generally bind two zinc atoms per subunit (Jornvall *et al.,* 1987; Sun & Plapp, 1992).

The domain predictions made by Pfam and InterPro are consistent with the predicted domain characteristics made by the sequence analysis that classified TmMtDH as a medium-chain dehydrogenase reductase. In addition, these predicted domain structures also agree with preliminary characterizations of TmMtDH that have been done by Dr. Claire Vieille at Michigan State University. In addition to analyzing the effect of temperature on this hyperthermophilic source of MtDH, the zinc content was also assessed. To determine the effect of zinc on TmMtDH activity, the enzyme was treated with 10 mM EDTA for 20 minutes at 37 ˚C and it was determined that the EDTA reduced enzyme activity by 96%. In addition, adding 20 mM zinc chloride back into the EDTA-treated TmMtDH restored activity back to 80% of the native enzyme activity (C. Vieille and colleagues, unpublished results).

1.2.5 Analysis of Previously Characterized Mannitol Dehydrogenase Structures

Currently, only three structures of MtDHs have been published: the SDR *Agaricus bisporus* MtDH (Horer *et al.,* 2001) and two different complexes of the LDR *Pseudomonas fluorescens* MtDH (Kavanagh *et al.,* 2002). *P. fluorescens* mannitol 2-dehydrogenase is a 54 kDa long chain dehydrogenase reductase and is characterized as a secondary alcohol dehydrogenase. The binary (PDB 1LJ8) and ternary (PDB 1M2W) complexes of *P.*

fluorenscens MtDH with NAD^+ and NAD^+ and D-mannitol have been solved by X-ray crystallography to a resolution of 1.7 Å and 1.8 Å respectively (Kavanagh *et al.,* 2002; Bubner *et al*., 2008; Figure 1.11). The structures confirm biochemical data that LDRs are monomers and do not require a metal for catalysis. Two domains form the overall structure of *P. fluorescens* MtDH with the active site located at the base of the cleft between the domains. The majority of residues that bind NAD^+ are from the N-terminal domain while mostly residues from the C-terminal domain are responsible for mannitol binding. NAD^+ is bound with the adenine *anti* and the nicotinamide *syn* between the two domains in a Rossman fold motif, which is an NAD-binding motif that consists of parallel beta strands and alpha helices in a beta-alpha-beta-alpha-beta order (Rao and Rossman, 1973).

In addition, the structure of *P. fluorescens* in complex with NAD+ published by Bubner *et al*. (2008) demonstrated that Asp69 is important in determining binding preference for NAD+ over NADP+. The authors used mutational studies to demonstrate that the cofactor preference could be changed by mutating Asp69 to an alanine residue, resulting in the mutated enzyme using NADP(H) almost as well as the wild-type enzyme used NAD(H). An enzyme used as a commercial biorector for mannitol production would need to use the less expense NAD(H) over the more expensive NADP(H) and these findings demonstrate that it is possible to alter cofactor preference.

As previously discussed, sequence analysis had classified *Agaricus bisporus* MtDH as an SDR. In general SDRs contain approximately 250 residues and do not require metals for catalysis. The *Agaricus bisporus* MtDH crystal structure in complex with $NADP^+$ was solved to a resolution of 1.5 Å (Figure 1.12). Prior to crystallization, *Agaricus bisporus* MtDH was characterized as a homotetramer in solution by dynamic light scattering analysis (Sassoon *et al.*, 2002). Crystal data agree with this tetrameric structure, as three tetramers are present per asymmetric unit. A further structural analysis of other members of the SDR family indicates that three residues, called the catalytic triad, are vital for catalysis. These residues have been identified as Ser149, Tyr169 and Lys173 (Figure 1.8). These residues do not appear to be conserved in the LDR family but may be conserved in TmMtDH.

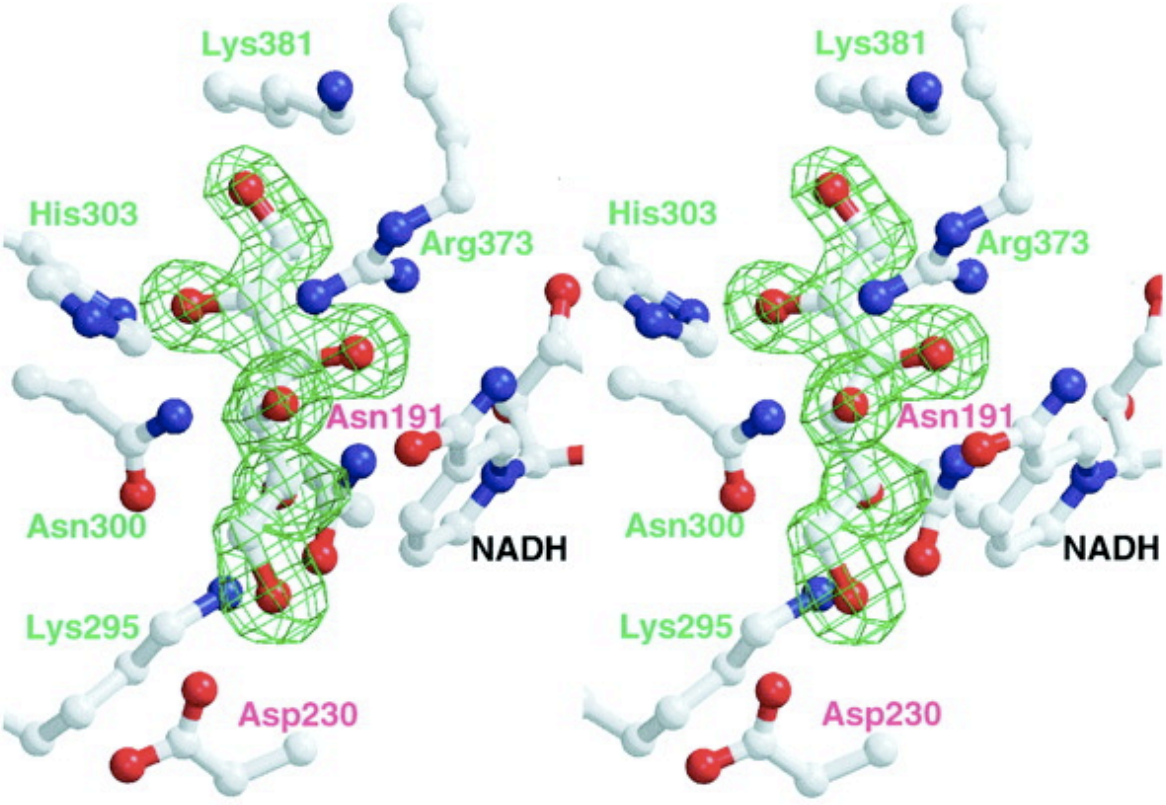


Figure 1.11 The structure of *Pseudomonas fluorescens* mannitol dehydrogenase bound to NAD and mannitol.
The important interactions between *Pseudomonas fluorescens* mannitol dehydrogenase and substrates NAD and mannitol are shown through an overlay of the binary and ternary complexes. Representative electron density for mannitol (shown in green) is contoured at 4 σ. From Kavanagh *et al*. (2002).

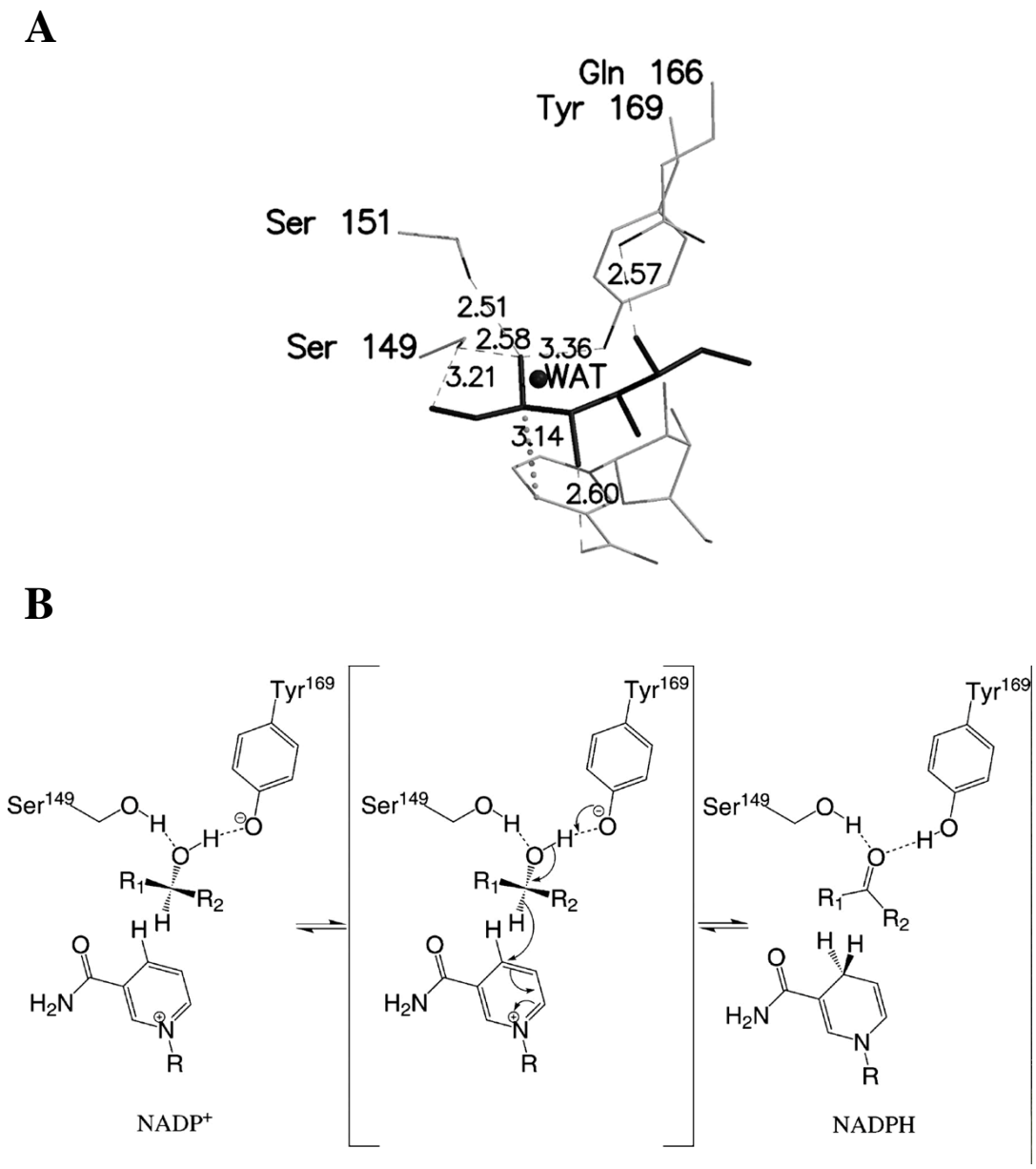


Figure 1.12 Binding of mannitol and proposed catalytic mechanism of *Agaricus bisporus* mannitol dehydrogenase

Three important residues (Ser149, Tyr169 and Lys173) have been identified as being vital for *Agaricus bisporus* mannitol dehydrogenase (AbMtDH) catalysis of mannitol and have been termed the catalytic triad. The binding of mannitol to important residues in AbMtDH is shown in A, along with a water atom (WAT) that is within 0.88 Å to the O_2 oxygen atom of mannitol and is theorized to be expelled during mannitol binding. B shows the proposed catalytic mechanism of AbMtDH involving two of the three residues of the catalytic triad, Ser149 and Tyr169. The proposed reaction mechanism suggests that Tyr169 is depronated, facilitated by the positive environment from Lys173 and $NADP^+$, and that this tyrosine interacts directly with the hydroxyl group of mannitol. From Horer *et al* (2001).

Although little overall sequence identity exists between *Pseudomonas fluorescens* MtDH, *Agaricus bisporus* MtDH and TmMtDH, the previously solved structures of *Pseudomonas fluorescens* MtDH and *Agaricus bisporus* MtDH shed light on what the crystal structure of TmMtDH may look like. All three enzymes contain a nicotinamide-binding domain and thus this domain of TmMtDH is likely to be of a similar Rossman fold as the previously solved structure. In addition, previous studies have also characterized TmMtDH as a tetramer in solution so an overall crystallographic symmetry similar to that of *Agaricus bisporus* could be expected (Figure 1.12).

1.2.6 Aims and Objectives

Preliminary characterization of MtDH from the hyperthermophilic source, *Thermotoga maritima*, has suggested that this enzyme is a promising target for commercial applications. Although crystallographic structures of the SDR, *Agaricus bisporus* MtDH, and the LDR, *Pseudomonas fluorescens* MtDH, are available, no structures are available for any of the bacterial MDR like TmMtDH. Due to the low sequence identity between the SDR *Agaricus bisporus* MtDH and TmMtDH (20%), the LDR *Pseudomonas fluorescens* MtDH and the MDR TmMtDH, X-ray crystallographic data of TmMtDH is necessary to deduce the unique characteristics of this enzyme.

The primary aim of the project reported herein was to crystallize TmMtDH. The enzyme was purified and characterized by Dr. Claire Vieille at Michigan State University and samples were provided for crystallization screening. Once crystallization conditions were deduced, these were optimized to produce high quality crystals for diffraction experiments. The next objective was to diffract and solve the structure of TmMtDH at an X-ray source, either in-house or synchrotron. The structure was to be solved by molecular replacement using models of alcohol dehydrogenases available in the protein data bank or a combination of structures to improve phase calculations for molecular replacement. If molecular replacement was not successful, MAD data collection at the Canadian Light Source (CLS) would likely be employed.

2.0 Materials and Methods

2.1 Protein Purification

2.1.1 Purification of A-Raf

In general, most buffers and common laboratory chemicals were purchased from Sigma-Aldrich (Oakville, Ontario) and BioShop (Burlington, Ontario).

All A-Raf plasmids were created according to the protocol in Johnson *et al.* (2005). Although attempts were made to purify many different A-Raf expression plasmids, most efforts concentrated on the purification of full-length human A-Raf (amino acid residues 1-606) and the CR2/CR3 domains (amino acid residues 200-606) of A-Raf. All full-length or portions of A-Raf contained a glutathione S-transferase (GST) tag for purification on a glutathione Sepharose column.

For protein expression, each plasmid was expressed in one of two different *Escherichia coli* strains: BL21 (genotype F^- *ompT* $hsdS_B$ (r_B^- m_B^-) *gal dcm* (DE3), catalogue number 70235) and Rosettagami (genotype Δ(ara–leu)7697 ΔlacX74 ΔphoA PvuII phoR araD139 ahpC galE galK rpsL (DE3) F'[lac+ lacI q pro] gor522::Tn10 trxB pRARE2 (CamR, KanR, StrR, TetR), catalogue number 71351), both supplied by Novagen (Gibbstown, NJ). Colonies were selected by plating on Luria broth (LB)/agar plates with 100 μg/mL ampicillin before being cultured for 24 hours at 37 ˚C in LB growth media plus 100 μg/mL ampicillin (BL21 cells) or 100 μg/mL ampicillin, 10 μg/mL chloramphenicol, 25 μg/mL kanamycin, and 15 μg/mL tetracycline (Rosettagami cells) for 24 hours. Once an OD_{600} of 0.6 had been reached, 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) plus 400 μL of zinc chloride was added and cells were grown for 1 hour to overnight at 20 ˚C. Cells were centrifuged in a BR4i Jouan centrifuge (Analytical Instruments, Golden Valley, MN) at 7800 x g and the supernatant was discarded. The pellet was then resuspended in 30 mL of ice cold 1 x PBS buffer (0.15 M NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4 at pH 7.3) with 1 % Triton X-100, 1 mM dithiothreitol (DTT), 1 mM phenylmethylsulfonylfluoride (PMSF) (in 100% isopropanol) and 10 μg/mL benzamidine-HCl.

Resuspended cells were then lysed by either sonication or French Press. In the sonication method, cells were lysed on ice six times in 1 minute bursts with an interval of 6 minutes at 50 W power. In the French Press method, cells were lysed with two passes through the French pressure cell operated at 1400 psi. Cells were chilled for 10 minutes at 4 ˚C after

each pass. The sample was centrifuged at 20,000 x g for 20 minutes at 4 ˚C. The supernatant was filtered through a 0.8 μm Acrodisc syringe filter (Pall Corporation, Mississauga, Ontario) and the pellet was discarded. Protein concentration was determined spectrophotometrically at a wavelength of 280 nm and approximately 5 μL were saved for future SDS-PAGE analysis. The rest of the sample was saved for further purification and 5 mM DTT was added to keep the sample environment reduced.

Two different column systems containing glutathione Sepharose 4-B resin were used for A-Raf purification: a gravity column run at 4 ˚C and the high-pressure AKTA Fast Protein Liquid Chromatography (FPLC, Amersham Biosciences, Piscataway, NJ) system run at room temperature. The gravity column and AKTA FPLC pre-packed column contained 3 mL and 1 mL of glutathione Sepharose beads, respectively. The wash buffer solution contained 50 mM Tris-HCl at pH 7.0, 0.15 M NaCl and 1 mM DTT. The column was washed with 20 times the bed volume (600 mL for gravity column) of wash buffer before the sample was added. The sample was added to the resin and allowed to equilibrate for 24 hours at 4 ˚C. The column was again washed 2 times with 150 mL of a buffer containing 50 mM Tris-HCl at pH 7.0, 0.15 M NaCl and 1 mM DTT), or until the OD_{280} no longer showed any absorbance. A solution of 300 μL of PreScission protease (Amersham Biosciences, Piscataway, NJ) in 4.7 mL of 50 mM Tris-HCl at pH 7.0, 0.15 M NaCl and 1 mM DTT was then added slowly to the column and incubated for 48 hours at 4 ˚C.

After the addition of GST-PreScission protease, the column was washed with a volume of approximately 200 mL of wash buffer (50 mM Tris-HCl at pH 7.0, 0.15 M NaCl, 1 mM DTT) or until the OD_{280} no longer showed any absorbance. Approximately 1 mL fractions were collected as the column was washed. The presence of protein in eluted column fractions was detected spectrophotometrically at a wavelength of 280 nm. Approximately 10 μL of sample from each column fraction were applied to a 12% sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) with a low molecular weight protein marker (Amersham Biosciences, product number 17-0446-01, molecular weights: 97.0, 66.0, 45.0, 30.0, 20.1, 14.4 kDa, Piscataway, NJ) for reference to check for the presence of protein. Fractions containing a band corresponding to the approximate molecular weight of A-Raf were concentrated by centrifugation at 3700 x g in a 30 kDa Vivaspin centrifugal concentrator (Sigma-Aldrich, Oakville, Ontario). Pure concentrated fractions of A-Raf containing 10% (v/v) glycerol were

then stored at –80 ˚C. The final concentration of the purified fraction was verified by the Bradford method (Sigma-Aldrich, product number B6916, Oakville, Ontario) of protein quantitation (Bradford, 1976).

2.1.2 Purification of TmMtDH

The following purification protocol was supplied by Dr. Claire Vieille of Michigan State University. The *Thermotoga maritima* gene was cloned into the NdeI and XhoI sites of pET24a, yielding the plasmid pTmMtDH. The pTmMtDH was expressed with a C-terminal His tag in *Escherichia coli* (*E. coli*) BL21 (DE3) cells. The *E. coli* cells were grown in 1 L super broth (24 g yeast extract, 12 g tryptone, 13 mL glycerol, 1 mM $MgSO_4$, 15.3 g K_2HPO_4 and 1.7 g KH_2PO_4 per litre) containing 50 μg/mL kanamycin at 37 ˚C. Expression was induced by adding 0.6 mM IPTG and then cells were allowed to grow for 16 hours. The cells were centrifuged at 4000 x *g* for 10 min and the pellet resuspended in 50 mM 3-(N-morpholino) propanesulfonic acid (MOPS) at pH 7.0 containing 5 mM β-mercaptoethanol and protease inhibitors. The *E. coli* cells were disrupted twice using a French pressure cell (96 MPa), followed by centrifugation at 25 000 x *g* for 30 min. The supernatant was heated at 85 ˚C for 20 min and centrifuged again at 20 000 x *g* for 20 min. The extract was loaded onto a nickel-nitrilosacetic acid agarose column (Qiagen Technologies, Mississauga, Ontario). The column was washed according to the Gibco BRL procedure for Protein Expression System, pRoEX-1 vector. The purified protein was dialyzed against 50 mM MOPS at pH 7.0 containing 5 mM β-mercaptoethanol. The protein sample was stored at –80 ºC and sent for studies for this project.

Upon receipt of purified TmMtDH, the protein was first dialyzed against a storage buffer solution containing 50 mM Tris at pH 8.5. TmMtDH was then concentrated to 40 mg/mL, 10% glycerol was added and the protein was stored at –80 ˚C. During later experiments, TmMtDH was dialyzed against a storage buffer solution containing 50 mM Na_2HPO_4 at pH 7.0 and 0.2 M NaCl before concentration and storage.

2.1.3 SDS-PAGE Analysis

The 12% separating gel was prepared by mixing 9.4 mL of 40% acrylamide stock (38.9 g acrylamide plus 1.07 g *bis*-acrylamide in 100 mL distilled water), 12.3 mL of distilled water, 7.5 mL of 1 M Tris-HCl at pH 8.8, 0.3 mL of 10% SDS, 20 μL of N,N,N',N'-

Tetramethylethylenediamine (TEMED), and 0.15 mL 10% ammonium persulfate. The separating gel was poured into the gel apparatus and allowed to polymerize. The 4.5% stacking gel was prepared by mixing 1.7 mL of 40% acrylamide stock, 10.8 mL of distilled water, 1.9 mL of 1 M Tris-HCl at pH 6.8, 0.15 mL of 10% SDS, 20 μL of TEMED, and 15 μL of 10% ammonium persulfate. The stacking gel was applied on top of the separating gel and a lane comb was inserted before the gel was allowed to polymerize. A running buffer solution of 196 mM glycine, 0.1% SDS and 50 mM Tris-HCl at pH 8.8 was then poured over the gel apparatus. Samples were prepared by mixing approximately 10-20 μL of protein solution (depending on concentration) with 10 μL of buffer containing 125 mM Tris-HCl at pH 6.8, 10% 2-mercaptoethanol, 10% SDS, 10% glycerol, and a small amount of bromophenol blue and were then boiled over a water bath for 5 minutes. Samples were applied to the gel using a micropipette. The power was then turned on to 40 mA and then gel was run until the bromophenol blue reached the bottom of the gel.

The gel was disassembled from the gel apparatus and placed in a container and stained with 0.25% (w/v) Coomassie Brilliant Blue R 250 in methanol-water-glacial acetic acid (v/v) (5-5-1). The gel was incubated for 2 to 4 hours while on a rocking incubator. The gel was then washed in 7% (v/v) acetic acid. Gels were scanned using an Epson-Twain HP scanner to produce images for documentation.

2.2 Western Blot Analysis of A-Raf

A 12% SDS-PAGE gel was run according to previously described protocols. A piece of 0.2 μm Immuno-Blot PVDF membrane (BioRad, Hercules, CA) and Whatmann 3 MM gel blot filter paper (BioRad, Hercules, CA) were cut to fit the size of the gel. The SDS-PAGE gel was removed from the gel plates and the stacking gel was removed. The overnight transfer apparatus was prepared and all air bubbles were removed. The sandwich of support pad, 2 pieces of Whatmann filter paper, gel, PVDF membrane, 2 pieces of Whatmann filter paper, and support pad was inserted into the transfer tank filled with transfer buffer (20 mM Tris base at pH 8.3, 150 mM glycine, 20% methanol) and the apparatus was run at 45 V overnight.

The transfer sandwich was then disassembled and the membrane was blocked overnight on a shaker with a skim milk blocking solution containing 5% skim milk powder, 0.05 % Tween-20 and 25 mL of 0.1 M PBS at pH 7.4. The membrane was then incubated for 1 hour on

a shaker with diluted primary antibodies in 20 mL of 0.1 M PBS at pH 7.4 plus 0.05 % Tween-20.

Two different primary antibodies were used for A-Raf Western blots. The first primary antibody, anti-A-Raf AP1, was raised in rabbits against human GST-A-Raf and recognizes the CR2/CR3 domains of A-Raf (amino acid residues 200-606). The AP1 antibody was generated and affinity purified by staff in Dr. Anderson's laboratory (Johnson *et al.,* 2004). The second primary antibody, C-20 (Santa Cruz Biotechnology, sc-408, Santa Cruz, CA) recognizes the 20 amino acid residues at the C-terminus of A-Raf. The concentrations of AP1 and C-20 for all Western blots were 3 μg/mL and 2 μg/mL respectively. A 0.4 μg/mL concentration of goat anti-rabbit IgG horseradish peroxidase (Amersham Biosciences, Piscataway, NJ) was used as a secondary antibody for all Western blots.

After incubation with primary antibodies, the membrane was washed three times for 5, 15 and 5 minutes with 50 mL of the PBS/Tween mixture (0.05 % Tween-20 plus 25 mL of 0.1 M PBS at pH 7.4). The membrane was then incubated for 90 minutes with secondary antibodies diluted in the PBS/Tween buffer solution. The PVDF membrane was washed 4 additional times with 0.05 PBS/Tween buffer for 15 minutes each. The substrate solution for horseradish peroxidase was prepared in accordance with the Amersham ECL kit protocol. The PVDF was incubated with this substrate solution for 1 minute. Excess solution was drained off and the membrane was placed on a 3 MM filter with the protein side up. The membrane and filter were wrapped in Saran wrap and placed in an X-ray cassette. A sheet of Kodak-X-Omat Blue XB-1 film was placed on top of the membrane for 7-60 seconds. The film was developed in an Amersham Hyperprocessor Automatic Film Processor (Amersham Biosciences, Baie D'Urfe, Quebec).

2.3 Mass Spectrometry Analysis

Samples of purified A-Raf and TmMtDH were resolved on a 12% SDS-PAGE gel. Bands of protein were excised from the gel and delivered to Doug Olson at the National Research Council (Saskatoon, SK) for trypsin digestion (Shevchenko *et al*., 2006) and Matrix Assisted Laser Desorption/Ionization Time-of-Flight (MALDI-TOF) or Quadrupole Time-of-Flight (Q-TOF) mass spectrometery. Doug Olson analyzed the resulting peaks using Mascot and FindMod programs.

2.4 Gel Filtration Studies

Gel filtration was performed on A-Raf and TmMtDH samples using the AKTA FPLC system (Amersham Biosciences) with Unicorn 5.01 software at room temperature. A HiPrep 16/60 pre-packed column with a column volume of 120 mL (GE Healthcare Lifesciences, Baie d'Urfe, Quebec) was washed with 2 column volumes (240 mL) of degassed water and 2 column volumes (240 mL) of PBS buffer before 400 μL of 4 mg/mL A-Raf or 1.2 mL of 15 mg/mL TmMtDH were injected into the system. PBS buffer was applied to the column and 1 mL fractions were collected and protein peaks monitored via UV detection. The flow rate for the duration of the experiment was 0.5 mL/min.

TmMtDH was also analyzed by gel filtration on a calibrated HiPrep 16/60 pre-packed column. The same protocol used for gel filtration on the uncalibrated column was used for calibrated studies.

2.5 Crystallization Trials

2.5.1 Crystallization of A-Raf

Various commercially available sparse matrix screening kits, like Crystal Screen, Crystal Screen 2, PEG/Ion 2 Screen, and Crystal Screen Lite from Hampton Research (Viejo, CA), the Heavy + Light Twin Pack HT-96 from Molecular Dimensions (Apopka, FL), and Nextal Classics 1 and II and Nextal PEGs from Qiagen (Mississauga, Ontario) were used to try and find appropriate crystallization conditions for native full-length A-Raf and the CR2 and CR3 domains of A-Raf. Screens were also repeated to include 0.01 to 0.1 mM 3-(Dimethylamino)-N-[3-[(4-hydroxybenzoyl)-amino]-4-methylphenyl]benzamide (product name ZM 336372, Tocris Biosciences, Hornby, Ontario), according to the protocol in Wan *et al.* (2004). Crystallization screens were performed using the microbatch-under-oil and hanging drop vapour diffusion methods (Chayen, 1999). In microbatch experiments, 0.5 – 1 μL of 8 – 10 mg/mL protein was added to 0.5 – 1 μL of precipitant solution in a 1:1 ratio of protein to precipitant under 2.5 – 5 mL of either 100% paraffin oil (mineral oil), or a 50% paraffin: 50% silicon oil mixture.

Any large crystals produced from a crystallization screen were soaked in a cryo-solution containing either 40% (v/v) ethylene glycol or glycerol and the mother liquor and were then

flash cooled in liquid nitrogen. These crystals were screened on the in-house diffractometer at the SSSC at the University of Saskatchewan. Izit dye (Hampton Research) was added to a drop containing any crystals that were not large enough to be screened at the SSSC in a volume of 10% of the drop volume to try to determine if the crystals were protein or salt.

2.5.2 Crystallization of TmMtDH

TmMtDH was also screened against many commercially available sparse matrix screening kits using both the microbatch-under-oil and hanging drop vapour diffusion methods. For both crystallization methods, 1 μL of 4-10 mg/mL TmMtDH in a storage buffer containing either 50 mM Tris at pH 8.5 or 50 mM Na_2HPO_4 at pH 7.0 and 0.2 M NaCl was added to 1 μL of precipitant solution from the screening kit. During microbatch-under-oil experiments, the final drop volume of 2 μL was covered with 5 mL of 100% mineral oil. During hanging drop vapour diffusion experiments, 2 μL drops containing the protein and precipitant solution were placed on plastic cover slips and were inverted over 1 mL of precipitant solution. Both hanging drop vapour diffusion and microbatch-under-oil experiments were incubated at 4 ˚C, room temperature or 30 ˚C.

A solution containing 600 μL of 10 mg/mL TmMtDH in a buffer containing 50 mM Tris at pH 8.5 was also sent to the Hauptman-Woodward Institute in Buffalo, New York for high-throughput screening of 1536 crystallization conditions by the microbatch-under-oil method. Images of the protein and precipitant solution drops were viewed using MacroScope software and promising conditions were repeated in-house using larger scale microbatch-under-oil and hanging drop vapour diffusion methods.

Screening kit conditions that produced promising results were optimized by varying components such as pH, salt concentration and precipitant concentration. Co-crystallization of TmMtDH was also attempted by adding 5 mM NAD, 5 mM NADP or 10 μM $ZnCl_2$ to 10 mg/mL TmMtDH in a storage buffer containing either 50 mM Tris at pH 8.5 or 50 mM Na_2HPO_4 at pH 7.0 and 0.2 M NaCl.

To ensure that the crystals formed in the above conditions were protein and not salt, 0.2 μL (10% of the drop volume) of Izit Dye (Hampton Research) was added to the protein and precipitant solution drops. Protein crystals were flash cooled in liquid nitrogen for diffraction experiments. No extra cryo-protectant was added prior to flash cooling for crystals formed

under condition 1. Crystals formed by condition 2 were soaked in a cryo-solution containing either 40% glycerol or 40% ethylene glycol and the mother liquor before flash cooling.

2.6 Diffraction of TmMtDH Crystals

TmMtDH crystals were diffracted either at the SSSC or on the 08ID-1 protein crystallography beam line at the CLS. Diffraction data at 0.5^{o} scans were collected for a TmMtDH crystal grown in Condition 1 at the CLS at a detector distance of 200 mm, for 180^{o}, with 7-second exposure, and at a wavelength of 1.331 Å. Intensity data were indexed, integrated and scaled by Michel Fodje using X-Ray Detector Software (XDA; Kabsch, 1993).

2.7 Molecular Replacement Model Search for TmMtDH

A molecular replacement model for the diffraction data set from the TmMtDH crystal was found using the 3D-PSSM search engine (Kelley *et al.,* 2000). The structure of ketose reductase (sorbitol dehydrogenase) from Silverleaf Whitefly (PDB 1E3J) and a second structure from NADP(+)-Dependent Alcohol Dehydrogenase from *Bacillus stearothermophilus* (PDB 1RJW) were the top two results. Using these two models, attempts to calculate initial phases using the program PHASER (Storoni *et al.*, 2004) were made. More recently, the program 3D-Jigsaw (Bates *et al.*, 2001) was used to locate another potential molecular replacement model, the structure of L-Threonine Dehydrogenase from *Pyrococcus horikoshii* (PDB 2DFV). In addition domains of previously solved MtDH structures, including the NAD-binding domain and the alcohol dehydrogenase domain could be used for molecular replacement.

3.0 Results

3.1 A-Raf

3.1.1 Purification of A-Raf

The laboratory of Dr. Anderson at the Saskatchewan Cancer Agency provided several different plasmids encoding full-length or portions of A-Raf for purification experiments. Attempts were made to purify full-length A-Raf, CR2/CR3 domains (amino acid residues 200-606), the CR2 domain (amino acid residues 200-290), and the CR3 domain (amino acid residues 290-606) of A-Raf from BL21 cells and Rosettagami cells, but most efforts were concentrated on the purification of full-length A-Raf and the CR2/CR3 domains of A-Raf.

The original purification protocol used BL21 cells for expression. BL21 cells (Novagen) are all-purpose cells for high-level protein expression that are deficient in *OmpT* and *Lon* proteases that may compromise the expression level of recombinant proteins (Studier, 1990). Later purification protocols also used Rosettagami cells for protein expression, as these cells contain a trxB/*gor* mutation to enhance disulfide bond formation and contain tRNA genes used in eukaryotic protein expression that are rarely used in *E. coli* protein expression. These cells are derived from a *lac*ZY mutant of BL21 to enable control of expression levels by adjusting IPTG concentration, and include the *lon* and *ompT* deficiencies of BL21 to increase protein stability (www.emdbiosciences.com, accessed June 24, 2008).

Initially, Dr. Sanjukta Aich, a research associate in the laboratory of Dr. Louis Delbaere, created an optimized purification protocol for the expression of A-Raf and its domains in BL21 cells based on a previous purification protocol from the laboratory of Dr. Anderson. Dr. Aich tested concentration of IPTG, temperature variation and induction time period to obtain optimal induction of the *E. coli* BL21 cells. Thus in the original purification protocol used for these studies, a 1 L culture of BL21 cells containing the plasmid encoding full-length GST-A-Raf was induced with 5 μM IPTG for 24 hours at 20 ˚C. Also according to the original purification protocol, after the cultures were induced, the bacteria were lysed by sonication and bound to glutathione Sepharose beads in a gravity column. All A-Raf products were tagged with GST for purification on a glutathione Sepharose column. Therefore, the masses of the GST-A-Raf were as follows: full-length GST-A-Raf 94.5 kDa, GST-CR2/CR3 71.5 kDa. After the cleavage of the 26.5 kDa GST tag by PreScission protease from either the full-length A-Raf or the CR2/CR3 domains of A-Raf, the product was analyzed spectrophotometrically at a

wavelength of 280 nm to determine which samples contained protein and by 12% SDS-PAGE to determine the purity of the sample. It should also be noted that the GST-PreScission enzyme used to cleave the A-Raf protein from the GST while on the glutathione Sepharose column is approximately 46 kDa. A concentration of 10% glycerol was added to the purified protein before storage at –80 ˚C to avoid freezing damage.

Unfortunately, full-length A-Raf and the CR2/CR3 domains of A-Raf proved to be difficult to completely purify by this original purification protocol as shown by the multiple bands resolved on the SDS-PAGE gel (Figures 3.1, 3.2). The expected molecular weights of full-length A-Raf and the CR2/CR3 domains of A-Raf are approximately 68 kDa and 45 kDa respectively. Most noticeably two bands of approximately 75 kDa and 66 kDa were continually present during the purification of both full-length A-Raf and the CR2/CR3 domains of A-Raf. An additional band of approximately 45 kDa was present when the CR2/CR3 domains of A-Raf were purified (Figure 3.1). In addition, the yield of semi-purified A-Raf produced from a 1 L induction was only approximately 0.4 mg protein. This indicates that perhaps in addition to instability issues there were also problems with very low efficiency of induction or lysis.

To attempt to resolve the purification problems of A-Raf, variations to the expression protocol were attempted. Firstly, it was believed that inducing a larger volume of culture might produce more purified protein in a single batch as the 1 L induction yielded only 1.4 mg of CR2/CR3, or 0.4 mg of full-length A-Raf. Thus, inductions of 5 L of culture were also attempted. Unfortunately, the 5 L induction appeared to yield approximately the same amount of protein as the 1 L induction.

Secondly, based on the studies performed by Ghosh *et al.* (1995), the concentration of IPTG was decreased to 1 μM from 5 μM and the time for induction decreased to 1 hour from 24 hours. It was theorized that by changing the concentration of IPTG and induction time that even though a smaller final yield of protein would be produced, that more full-length protein would be produced. In reality, by decreasing the time for induction and the concentration of IPTG, the yield of A-Raf was reduced to approximately 0.25 mg of CR2/CR3, or 0.05 mg of full-length A-Raf. It was therefore difficult to analyze whether the quality of the resulting protein sample was improved, due to the extremely low concentration of A-Raf in the sample.

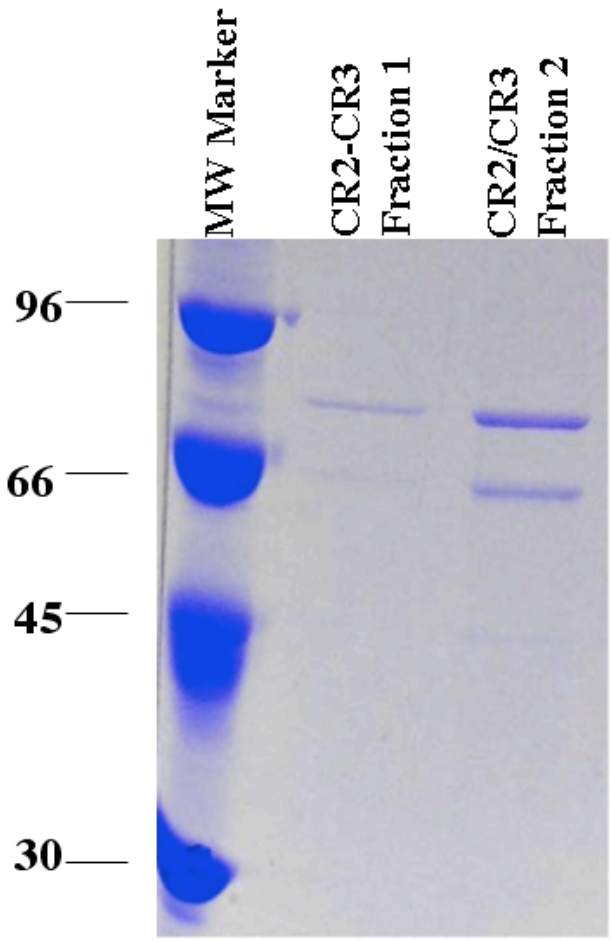


Figure 3.1 SDS-PAGE gel of the purification products of the conserved region 2 and 3 (CR2/CR3) domains of A-Raf.
The CR2/CR3 domains of A-Raf were expressed as a glutathione S-transferase (GST) fusion protein and affinity purified by a gravity column containing 3 mL of glutathione Sepharose beads. After the addition of 300 μL of PreScission Protease, CR2/CR3 A-Raf was washed from the column and collected in 200 1 mL fractions. Fractions were concentrated by centrifugation in a protein concentrator and the presence of protein was detected spectrophotometrically at a wavelength of 280 nm. Fractions 1 and 2 consist of the first (fraction 1) and second (fraction 2) 1 mL that were eluted from the column after the addition of PreScission protease. Fractions 1 and 2 were separated by 12% SDS-PAGE and stained with Coomassie blue. The expected molecular masses of the CR2/CR3 domains are 45 kDa, while the GST-CR2/CR3 has an expected molecular mass of approximately 71.5 kDa and GST-PreScission has an expected molecular mass of 46 kDa.

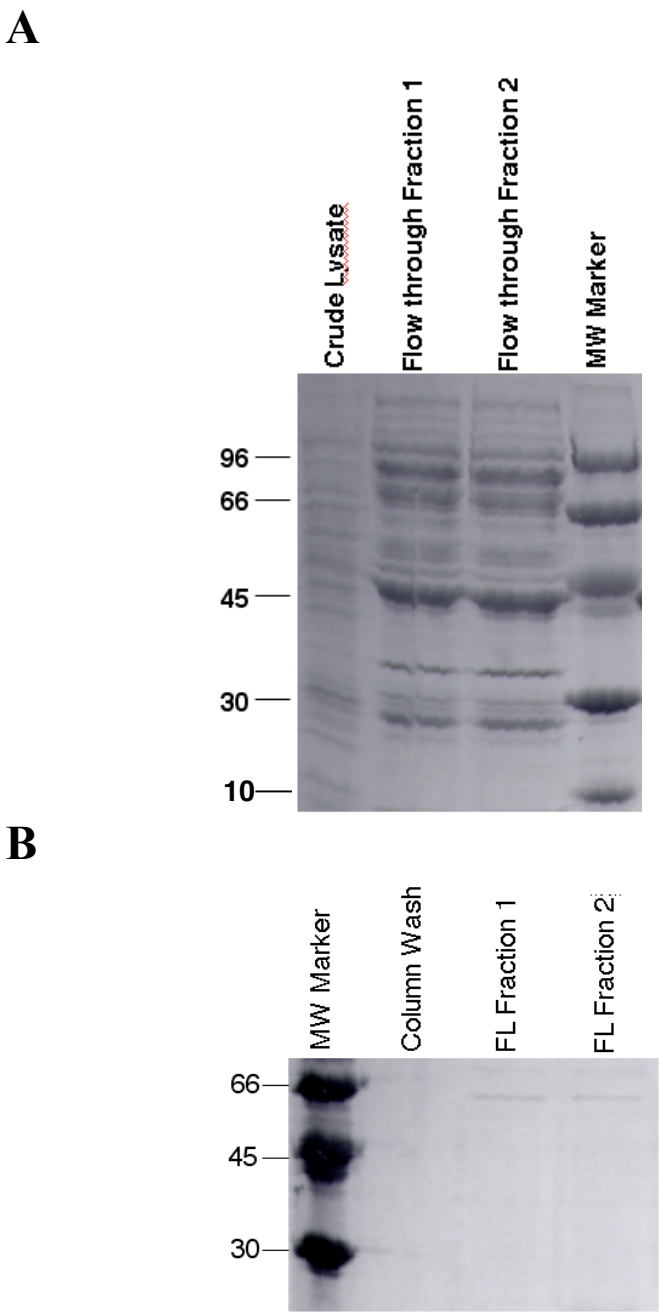


Figure 3.2 SDS-PAGE of purification products of full-length A-Raf
Full-length A-Raf was expressed as a glutathione S-transferase (GST) fusion protein and affinity purified by gravity column containing 3 mL of glutathione Sepharose beads. **A**. After the addition of the crude lysate, the column was washed with 300 mL of wash buffer (50 mM Tris-HCl at pH 7.0, 0.15 M NaCl and 1 mM DTT) and 1 mL fractions were collected. The first two 1 mL fractions (Fractions 1 and 2) of this column wash before the addition of PreScission protease are shown. **B**. The column was washed (Column Wash) before full-length (FL) A-Raf was liberated from the column after the addition of PreScission protease to cleave it from the GST. The first two 1 mL fractions after the addition of PreScission protease (FL Fractions 1 and 2) are shown. All fractions were separated by 12% SDS-PAGE and stained with Coomassie blue. The low molecular weight (MW) marker is from Amersham Biosciences.

As well, the bands seen in previous purifications of A-Raf at higher IPTG concentrations and longer induction times remained despite this specific protocol change.

Thirdly, it was theorized that perhaps the 14 cysteine residues present in A-Raf, an extremely large number compared to an average protein, were causing the instability issues of the protein. Thus the original purification protocol was again altered so that 1 mM dithiothreitol (DTT), a reducing agent, was added to every step of purification while using BL21 cells. In addition, a purification protocol utilizing Rosettagami cells was used in an attempt to purify A-Raf under oxidizing conditions since it was theorized that perhaps an environment that was favourable towards the formation of disulfide bonds might help to increase the stability of A-Raf. However, no noticeable difference was observed in either strongly reducing or oxidizing environments when the protein samples were purified, as a similar yield of protein was produced (data not shown).

In retrospect, the "contaminating" bands may have been explained by GST-fusion protein, as the expected molecular mass of the GST-CR2/CR3 is approximately 71.5 kDa and could correspond with the larger of the bands, 75 kDa, consistently noted on SDS-PAGE. Therefore, problems could have been more due to poor yield or poor ability to cleave the GST tag than with stability. The approximate molecular mass of GST is 26.5 kDa and this does not seem to correspond to any of the bands consistently seen on SDS-PAGE analysis. On the other hand, the molecular mass of GST-PreScission is 46 kDa and this is close to the approximately 45 kDa band seen with CR2/CR3 purification studies. Perhaps at this stage it might have been more useful to use Western blots as a mean to detect GST and A-Raf products than to use mass spectrometry analysis, but at the time of study primary antibodies were not available.

Next, the method by which the bacteria were lysed was altered. The original protocol used sonication to lyse the bacterial cells, but the high temperature created during the procedure was thought to be too harsh for the protein. Thus, a French press method was used instead. Unfortunately, this change in protocol again seemed to have little effect on the outcome of the purification as there were still contaminating bands present in the SDS-PAGE gel of the purification products.

Finally, attempts were also made to purify A-Raf using a high-pressure glutathione Sepharose-binding column (GE Healthcare Lifesciences, product number 17-0757-01, Piscataway, NJ) in the AKTA FPLC system. Unfortunately, it appeared that the GST-tagged A-Raf was not binding to the resin in the high-pressure column because all protein remained in

the unbound fraction prior to the addition of glutathione as confirmed by the Bradford Method protein quantification and a Coomassie blue stained SDS-PAGE gel of the purification products. Attempts were made to increase the binding of the protein to the high-pressure glutathione Sepharose column by altering flow rate but purification but this method remained unsuccessful. The reason why the GST-A-Raf bound to glutathione Sepharose beads in the gravity column method but not in the FPLC system remains unclear.

3.1.2 A-Raf Mass Spectrometry Analysis

Regardless of whether the full-length A-Raf or the CR2/CR3 domains of A-Raf were purified and by which purification strategy, two bands of approximately 75 kDa and 66 kDa continued to hinder progress. Matrix assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) was thus performed to determine the identity of the 75 kDa and 66 kDa protein bands resolved by SDS-PAGE from the purification of the full-length A-Raf. Initially, Dr. Sanjukta Aich from the laboratory of Dr. Louis Delbaere performed the first purification analysis of full-length A-Raf. She also noted that two bands of approximately 75 kDa and 66 kDa were resolved by SDS-PAGE analysis on the purified A-Raf products. Dr. Aich excised the 66 kDa band from the SDS-PAGE gel because she believed it was closer to the expected molecular weight of full-length A-Raf and sent the sample for trypsin digestion and MALDI-TOF MS studies by Doug Olson at the National Research Council in Saskatoon, Saskatchewan. MALDI-TOF MS analysis identified this 66 kDa band as full-length A-Raf. However, at this time the 75 kDa band was not assessed and thus its identity remained unknown.

After subsequent purification analyses reported herein, the 75 kDa and 66 kDa bands continued to appear in the purification products of full-length A-Raf. Thus, further MS analyses were performed to try to determine the identity of these extra protein products. A second MALDI-TOF MS analysis was performed by Doug Olson of these two protein bands produced from the purification of full-length A-Raf. Mascot, the program used to determine the most likely identity of the MS analysis trypsin-digested peaks based on composition, matched the 75 kDa band to the heat shock protein dnaK and the 66 kDa band to mopA, a 60-kDa chaperonin 1. Thus, it was at this point that the purification protocol was first altered so that the concentration of IPTG was decreased to 1 μM and the induction time limited to 1 hour as suggested by Ghosh *et al.* (1995).

As previously mentioned, after modifying the concentration of IPTG and induction time, the 75 kDa and 66 kDa bands continued to appear in the products of the full-length A-Raf purification. MALDI-TOF MS was repeated in hopes that the changes to the purification protocol had helped rid the final purified sample of the heat shock protein and that either the 75 kDa or the 66 kDa protein band would be identified as full-length A-Raf. Unfortunately, the change in IPTG concentration seemed to have little effect as this time. Mascot matched the 75 kDa band to the heat shock protein hsp70 and the 66 kDa band to an unknown protein product from *E. coli*.

Similar attempts at MALDI-TOF MS analysis were made for the 75 kDa, 66 kDa and 45 kDa bands that resulted from the purification of the CR2/CR3 domains of A-Raf (Table 3.1). As with the full-length A-Raf purification products, none of the excised bands could be matched by Mascot to the CR2/CR3 domains of A-Raf. Mascot uses databases like Expasy to identify the peptide mass fingerprint so it was theorized that Mascot could not recognize the CR2/CR3 domains of A-Raf because it was not the full-length protein that was found in Expasy. Thus, the sequence of the peptides produced from the trypsin digest was compared to the sequence of the CR2/CR3 domains of A-Raf using the program ClustalW (Thompson *et al.*, 1994). Unfortunately, the result indicated very little sequence identity between the CR2/CR3 domains and the trypsin-digested peptides.

3.1.3 Western Blot Analysis of A-Raf

Most efforts were focused on purifying the full-length A-Raf protein and the CR2/CR3 domains of A-Raf but the presence of a 75 kDa and a 66 kDa protein band resolved by SDS-PAGE hindered progress. After MALDI-TOF MS experiments, Western blot analysis was performed on the purified full-length A-Raf and the CR2/CR3 domains of A-Raf to try to determine if any of the protein bands were fragments of A-Raf. Two different antibodies were used for blotting: C-20, which binds to the C-terminal 20 amino acids of A-Raf, and AP1, which binds to residues in the CR2 and CR3 domains of A-Raf. In the Western blot of the full-length A-Raf, the C-20 antibody detected two protein bands corresponding to approximately 65 kDa and 48 kDa (Figure 3.3A). Since the C-20 antibody binds to the C-terminal 20 amino acid residues of A-Raf, this suggests that both of these bands contain A-Raf fragments and that degradation of the protein or incomplete formation of full-length protein had likely occurred.

Table 3.1 Peptides of *Escherichia coli* GroEL-like protein matching the trypsin digested purification product of the CR2/CR3 domains of A-Raf.
The CR2/CR3 domains of A-Raf were expressed as a glutathione S-transferase (GST)-fusion protein and affinity purified by gravity column containing glutathione Sepharose beads. Purified protein was eluted from the column after the addition of PreScission protease and analyzed by SDS-PAGE. Three bands were separated by SDS-PAGE and the smallest band measuring approximately 45 kDa was excised and sent for trypsin digest and MALDI-TOF mass spectrometry analysis at the National Research Council. The Mascot search engine determined that the sequence of the peptides from the protein evaluated matched most closely to a GroEL-like protein from *Escherichia coli*. The peptides determined by mass spectrometry that most closely match peptides from *E. coli* GroEL-like protein are shown above. A score above 53 indicates identity or extensive homology.

Peptide	Score
LAGGVAVIK	36
LIAEAMDK	30
SFGAPTITK	29
GVNVLADAVK	45
AVTAAVEELK	30
VGAATEVEMK	43
ATLEDLGQAK	34
EMLPVLEAVAK	21
GQNEDQNVGIK	50
EMLPVLEAVAK	29
AVAAGMNPMDLK	46
AAVEEGVVAGGGVALIR	56
QQIEEATSDYDREK	21
AIAQVGTISANSDETVGK	61
DTTTIIDGVGEEAAIQGR	50
QIVLNCGEEPSVVANTYVK	72
NDAADLGAAGGMGGMGGMGGMM	16
EIELEDKFENMGAQMVK	21
ANDAAGDGTTTATVLAQAIITEGLK	75
AMLQDIATLTGGTVISEEIGMELEK	20
GGDGNYGYNAATEEYGNMIDMGILDPTK	22
GYLSPYFINKPETGAVELESPFILLADK	65

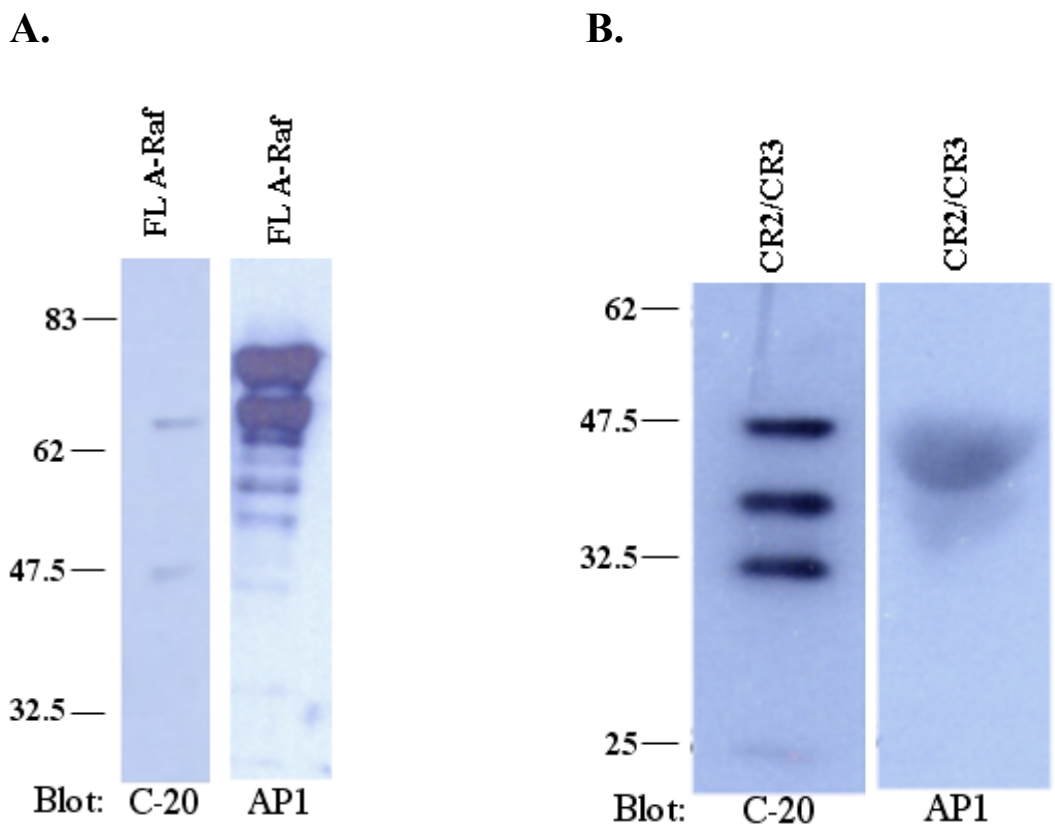


Figure 3.3 Western blot of full-length A-Raf using C-20 and AP1 primary antibodies. Full length (FL) GST-A-Raf were expressed, purified and cleaved from GST as previously described herein. Purification products were resolved by 12% SDS-PAGE and analyzed by Western blot. Two different primary antibodies were used for analysis: 2 μg/mL C-20 (Santa Cruz Biotechnology), which recognizes the C-terminal 20 amino acids of A-Raf, and 3 μg/mL AP1 (provided by Dr. Anderson's laboratory, Johnson *et al.,* 2004), which recognizes the CR2/CR3 domains of A-Raf (amino acid residues 200-606). A 0.4 μg/mL concentration of goat anti-rabbit IgG horseradish peroxidase (Amersham Biosciences) was used as a secondary antibody for all Western blots. Western blot to FL A-Raf containing C-20 and AP1 antibodies were exposed for 60 seconds and 30 seconds respectively (A), while blots to CR2/CR3 A-Raf containing C-20 and AP1 antibodies were exposed for 7 seconds and 30 seconds respectively (B). Note that the prestained molecular marker (Amersham Biosciences, product code 17-0446-01) for the FL A-Raf Western blot and the CR2/CR3 A-Raf Western blot do not align.

However, the AP1 antibody detected several bands, most noticeably protein bands corresponding to approximately 75 kDa and 65 kDa (Figure 3.3A). The AP1 antibody recognizes the CR2 and CR3 domains of A-Raf, suggested that the 75 kDa and 65 kDa bands both contain A-Raf fragments and that degradation has likely occurred.

In the Western blot of the CR2/CR3 domains of A-Raf, the AP1 antibody detected two protein bands of approximately 45 kDa and 38 kDa (Figure 3.3B), while the C-20 antibody bound to four protein bands of approximately 47.5, 40, 30, and 24 kDa (Figure 3.3B). These results are interesting because it would seem more likely that the AP1 antibody would bind to more CR2/CR3 fragments than the C-20 antibody. However, the fact that both antibodies bound to protein bands in the Western blot suggests that CR2/CR3 fragments are indeed present in the purification product.

3.1.4 Gel Filtration of A-Raf

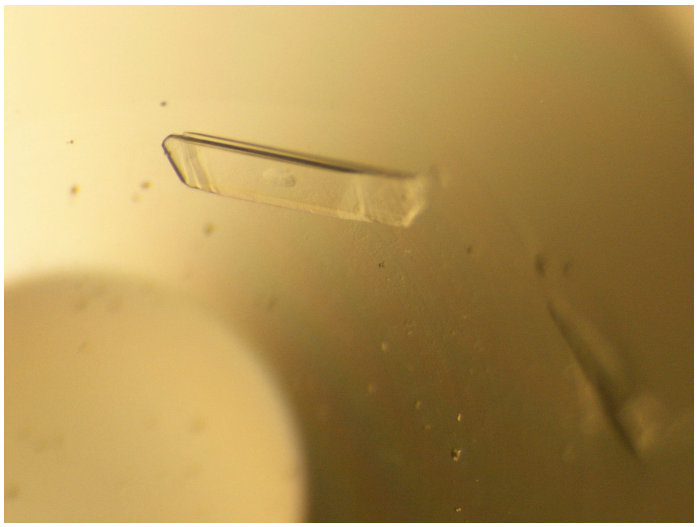
Gel filtration studies of A-Raf using a calibrated column were attempted in order to more accurately determine the molecular weight of the resulting filtration peaks. Unfortunately, because the expression and purification protocol produced such a low yield of both A-Raf and CR2/CR3, gel filtration studies were inconclusive and no peaks resolved.

3.1.5 Crystallization of A-Raf

Prior to the Western blot studies, crystallization attempts were made on the purified samples of the CR2/CR3 domains of A-Raf. Crystals of approximately 0.4 x 0.15 x 0.05 mm in diameter were produced from a screening kit made by Yvonne Leduc, a technician in the laboratory of Dr. Louis Delbaere, based on forty cocktails from Hampton Research Crystallization Screen (Figure 3.4A). These crystals were grown by microbatch under 5 mL of 50% paraffin: 50% silicon oil in 20% polyethylene glycol (PEG) 4000, 20% 2-propanol and 0.1 M sodium citrate at pH 5.5.

The crystals were first diffracted on the in-house x-ray source at the SSSC. The initial diffraction studies resulted in no discernable diffraction pattern, and the crystals were saved for subsequent diffraction studies. In addition, crystals of approximately 0.4 x 0.15 x 0.1 mm in dimension that formed from a precipitant solution of 70% 2-methyl-2,4-pentanediol (MPD), and 0.1 M HEPES at pH 7.5 (Nextal Classics, Figure 3.4B) and crystals measuring roughly 0.3

A



B

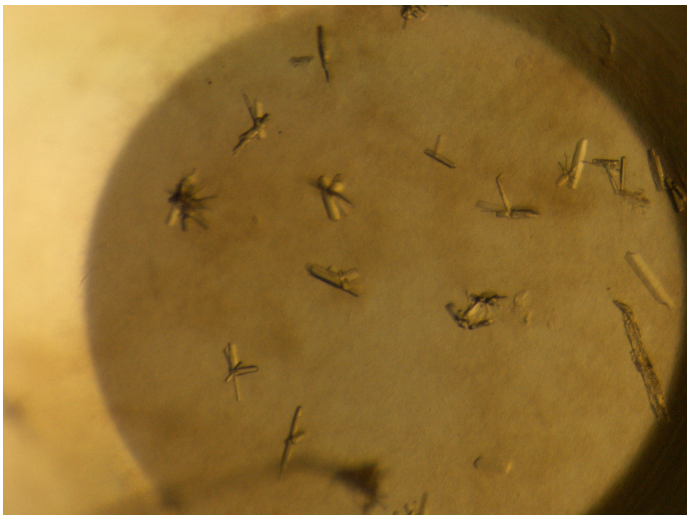


Figure 3.4 Images of potential conserved region 2 and 3 (CR2/CR3) A-Raf crystals.

A. Crystals grown in 20% polyethylene glycol (PEG) 4000, 20% 2-propanol and 0.1 M sodium citrate at pH 5.5, measuring approximately 0.4 x 0.15 x 0.05 mm in diameter.

B. Crystals grown in 70% 2-methyl-2,4-pentanediol (MPD), and 0.1 M HEPES at pH 7.5, measuring roughly 0.3 x 0.15 x 0.05 mm in dimension. Images are magnified 40X and were taken on a Nikon SMA1000 camera microscope.

x 0.15 x 0.05 mm in dimension from a precipitant solution of 3.4 M 1,6-hexanediol, 0.1 M Tris at pH 8.5 and 0.2 M $MgCl_2$ (Nextal Classics) were also screened on the SSSC. Initially, the crystals did not produce a diffraction pattern but the crystals were re-screened and this time produced characteristic salt diffraction patterns.

Also of note, other crystals formed using different crystallization screening kits with the addition of ZM Raf inhibitor to the crystallization cocktail. The ZM inhibitor is an inhibitor that has been studied extensively with B-Raf (Wan *et al.*, 2004). The ZM inhibitor was used in co-crystallization studies with B-Raf (PDB 1UWH) and seemed to stabilize B-Raf enough for protein crystal formation and subsequent X-ray diffraction data collection. It was hoped that by adding the ZM inhibitor to A-Raf crystallization attempts that A-Raf would also be stabilized and perhaps protein crystals could form. Unfortunately, any crystals that formed under these co-crystallization conditions proved to be salt after diffraction at the SSSC.

3.2 *Thermotoga maritima* Mannitol Dehydrogenase

3.2.1 Gel Filtration of TmMtDH

Samples of TmMtDH were purified by protocols described herein and shipped on dry ice for study. Once protein samples had been received, 10% glycerol was added and the samples were stored at –80 ˚C. TmMtDH was first analyzed by gel filtration after removal of the 10% glycerol by dialysis to avoid unwanted extra viscosity in the gel filtration column (Figure 3.5). The peak was collected into fractions, concentrated and resolved by a 12% SDS-PAGE gel. Despite the appearance of only one peak in the gel filtration software, two bands of approximately 40 kDa and 30 kDa resolved by SDS-PAGE (Figure 3.6), suggesting that degradation of the protein had occurred during shipping.

Secondly, to determine the approximate molecular weight of the sample TmMtDH was analyzed by a calibrated gel filtration column. The gel filtration column had previously been calibrated by analyzing proteins of known size and determining the volume of buffer required to elute the proteins from the column. TmMtDH eluted at a column volume, which corresponds to a molecular weight of approximately 135 kDa. These data suggest that TmMtDH is a tetramer in solution.

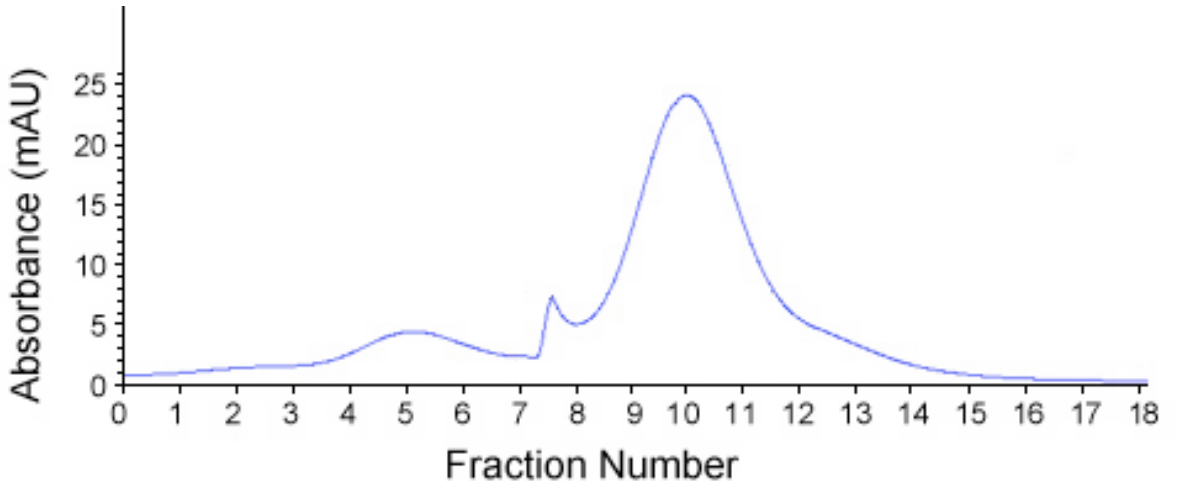


Figure 3.5 Gel filtration elution profile of *Thermotoga maritima* mannitol dehydrogenase (TmMtDH).
Gel filtration elution profile of TmMtDH as captured by the Unicorn 5.01 Software (Amersham Biosciences). A volume of 1.2 mL of 15 mg/mL TmMtDH was injected into the system for analysis by an uncalibrated HiPrep 16/60 gel filtration column. After wash with 1.5 times column volume (180 mL) the protein eluted as a single peak. Approximately 1 mL fractions were collected an analyzed by 12% SDS-PAGE gel. Figure 3.6 shows SDS-PAGE results of Fractions 8 to 11.

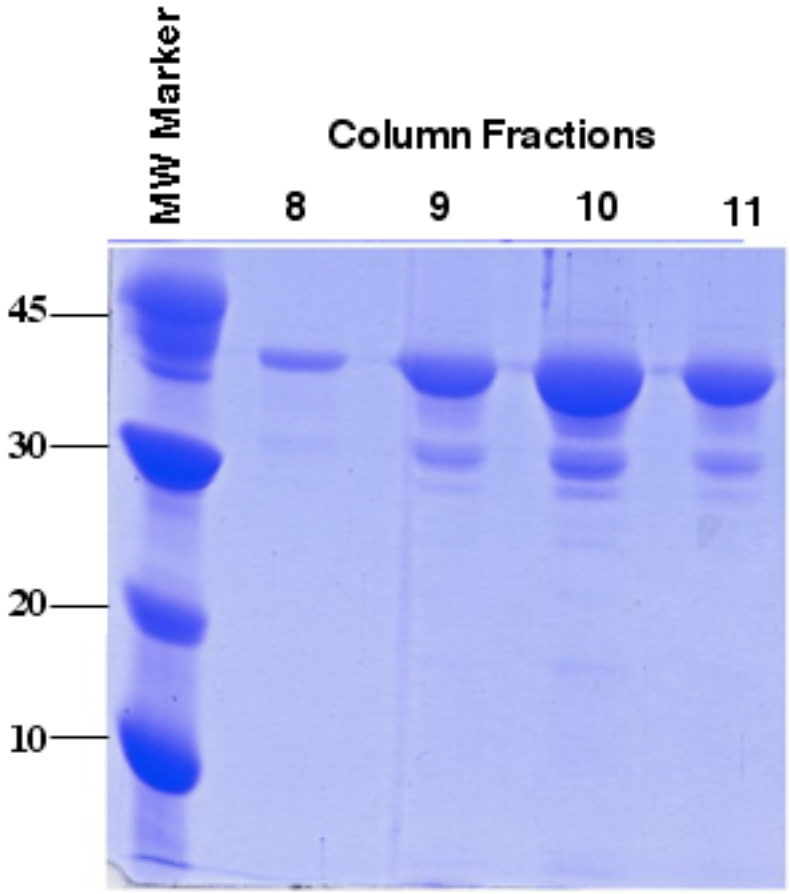


Figure 3.6 SDS-PAGE gel of *Thermotoga maritima* mannitol dehydrogenase (TmMtDH) fractions collected from gel filtration.
A sample of 15 mg/mL purified TmMtDH was analyzed on a 16/60 HiPrep column (GE Healthcare LifeSciences). Four 1 mL fractions were collected as the protein was eluted from the column in a single peak. Fractions were then resolved by 12% SDS-PAGE gel and stained with Coomassie Blue. The molecular weight (MW) marker was purchased from Amersham Biosciences. Fractions numbers 8 through 11 correspond to the fractions labeled in Figure 3.5.

3.2.2 Crystallization of TmMtDH

Samples of TmMtDH were thawed from –80 ˚C to 4 ˚C on ice for crystallization experiments. In vapour diffusion experiments, 1 μL of protein was added to 1 μL of mother liquor on a coverslip and then inverted over a well containing 1000 μL of mother liquor. In microbatch experiments, 1 μL of protein was added to 1 μL of mother liquor and then covered with a combination of silicone and paraffin oils. Crystallization experiments were then incubated at either room temperature or 4 ˚C.

3.2.2.1 Crystallization Condition 1

Two conditions produced crystals large enough for diffraction experiments. The first condition, Condition 1, contained 30% MPD and 0.1 M HEPES-Na at pH 7.5 and was part of the Nextal MPD Crystallization screening kit. A single crystal of TmMtDH formed in approximately 14 days at room temperature by the microbatch-under-oil method. This Condition 1 crystal was rectangular-shaped and measured approximately 0.1 x 0.05 x 0.05 mm (Figure 3.7).

Due to the high concentration of MPD in crystallization Condition 1, it was believed that no additional cryo-protection was required to protect the crystal that formed in the MPD screening kit prior to flash cooling in liquid nitrogen. To test this theory, a Hampton's Research crystal loop was filled with the solution of Condition 1. The solution in the loop remained clear and did not freeze upon flash cooling in liquid nitrogen, indicating that the concentration of MPD in the crystallization condition should be sufficient to act as a cryo-protectant. Thus, the Condition 1 crystal was mounted on a Hampton's Research 18 mm pin and flash cooled for diffraction experiments without additional cryo-protectant.

Due to the fact that a commercially available screening kit originally produced the Condition 1 crystal, further optimization was performed to attempt to reproduce the crystallization manually because the original crystals produced were not of diffraction quality. Preliminary attempts to reproduce the identical condition that first formed this crystal failed. It was originally believed that perhaps the method of making the precipitant solution commercially was different than the method of making the solutions in the laboratory. Thus a copy of the production sheets was obtained from Qiagen. However, the repeated condition did not result in the crystallization of TmMtDH.

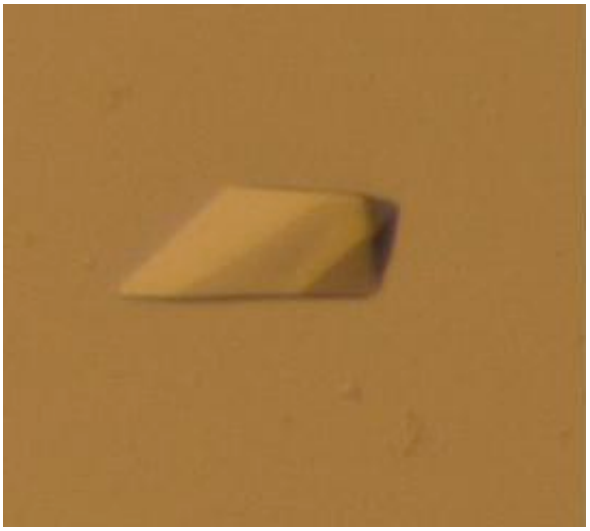


Figure 3.7 Crystals of *Thermotoga maritima* mannitol dehydrogenase (TmMtDH) grown under condition 1.
Crystal resulting from 10 mg/mL TmMtDH grown by the microbatch-under-oil method in Condition 1 containing 30% 2-methyl-2,4-pentanediol (MPD) and 0.1 M Na-HEPES at pH 7.5. The image was taken at 40X magnification with a Nikon SMA1000 microscope camera.

It should be noted that the original crystal that formed in the MPD screening kit was first noticed when the mother liquor had almost completely evaporated. The original crystal formed under the microbatch-under-oil method, and was covered with 5 mL of 50% mineral oil and 50% silicon oil. This complete evaporation may have finally allowed the solution to supersaturate sufficiently for nucleation to occur. Attempts were made to capitalize on this finding. Different ratios of mineral oil and silicon oil were used to cover the microbatch crystallization plate to experiment with evaporation time. Water was added to a rim around the crystallization plate to slow evaporation, and different temperatures were tested. As before, despite modifications to speed up or slow down evaporation time, the original crystallization condition could not be manually repeated.

Further optimizations were attempted using different concentrations of MPD, different pHs of the HEPES-Na buffer, different buffers and by adding salt and glycerol to the crystallization condition. Trials were also performed in both microbatch-under-oil and vapour diffusion crystallization plates and at various temperatures but TmMtDH could not be crystallized.

As a final effort to reproduce the crystals that formed in Condition 1, further microbatch-under-oil plates were set up using the original Nextal MPD screening kit. It should be noted that this screening kit contained only a limited volume of precipitant solution; thus the kit could not be used directly to form vapour diffusion plates. Again, the screens were optimized with different ratios of oils, different temperatures and water around the rim of the plates but TmMtDH could not be crystallized again. Evaporation of the solutions in the screening kit may have occurred by this time and thus perhaps the concentration of the precipitants was no longer optimal for crystal formation for this protein.

3.2.2.2 Crystallization Condition 2

In addition to using commercially available screening kits to determine the crystallization condition of TmMtDH, samples of TmMtDH were also sent to the high-throughput crystallization screening facility at the Hauptmann-Woodward Institute in Buffalo, New York. This high-throughput facility utilizes a robot to dispense very small protein volumes for crystallization trials against 1536 crystallization conditions. The drops formed by the robot are then digitally photographed at various time periods for the researcher to view at a later date.

The second condition that produced TmMtDH crystals in the laboratory, Condition 2 (15-20% polyethylene glycol 1000 (w/v), 0.1 M sodium citrate at pH 4, 0.2 M sodium bromide and 10% glycerol), is based upon a condition from the Hauptmann-Woodward Institute that formed TmMtDH crystals. Unlike the optimization of the Condition 1 crystal, Condition 2 could be manually repeated. Crystals had to be re-grown locally because the original crystals produced at the Hauptmann-Woodward Institute were not returned. Condition 2 crystals grew best in a 30 ˚C environment by the vapour diffusion method.

Condition 2 was able to crystallize native TmMtDH as well as co-crystallize TmMtDH with either NADP or zinc chloride substrates. Whether substrates were present or not, Condition 2 produced bipyramidal-shaped crystals (Figures 3.8, 3.9). Native TmMtDH Condition 2 crystals formed in approximately 5 days and measured approximately 0.15 x 0.15 x 0.1 mm. TmMtDH and zinc chloride crystals also formed in about 5 days and measured 0.1 x 0.075 x 0.075 mm. TmMtDH and NADP crystals formed in approximately 2 to 3 days and measured 0.075 x 0.075 x 0.05 mm.

Condition 2 was further optimized by changing the molecular weight of the polyethylene glycol, the pH and buffer of the solution, and the concentration and type of salt used in precipitant solution. In addition to polyethylene glycol 1000, 15 to 20% polyethylene glycol 4000 and 6000 also produce TmMtDH crystals. The crystals formed faster in the higher molecular weight polyethylene glycol and grew to a larger size than the crystals that formed in the original Condition 1. It was also noted that no crystals formed in variations of Condition 2 that used a different buffer or a pH outside the range of 4 to 4.5.

Several different cryo-protectants were tested with the mother liquor of Condition 2 and the higher molecular weight polyethylene glycol variations before crystals were flash cooled in liquid nitrogen. Both 40% ethylene glycol and 40% glycerol added to the mother liquor seemed to react the best after flash cooling. Thus, either 40% ethylene glycol or 40% glycerol in a solution of mother liquor were added to the protein drops. The crystals were then mounted on either an 18 mm Hampton's Research loop or a Mitegen loop.

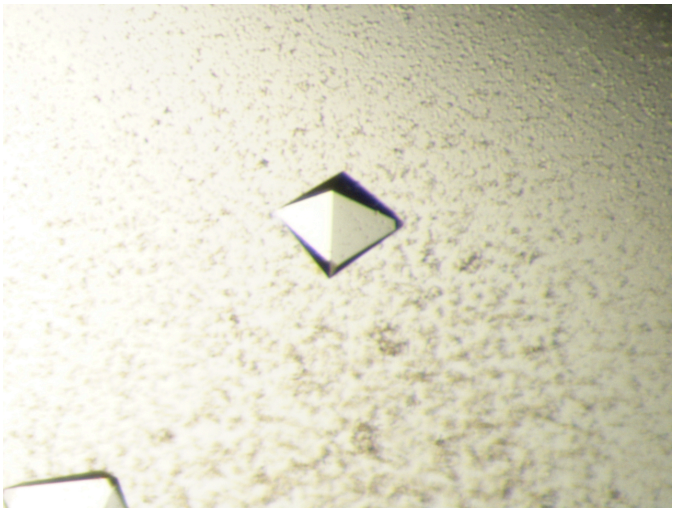


Figure 3.8 Crystals of *Thermotoga maritima* mannitol dehydrogenase (TmMtDH) grown under condition 2.
Crystal of native 10 mg/mL TmMtDH grown by the vapour diffusion method in Condition 2 containing 17.5% polyethylene glycol (PEG) 1000 (w/v), 0.1 M sodium citrate at pH 4, 0.2 M sodium bromide and 10% glycerol. The longest dimension of the crystal shown in the centre of the figure is approximately 0.15 mm. The image was taken at 40X magnification with a Nikon SMA1000 camera microscope.

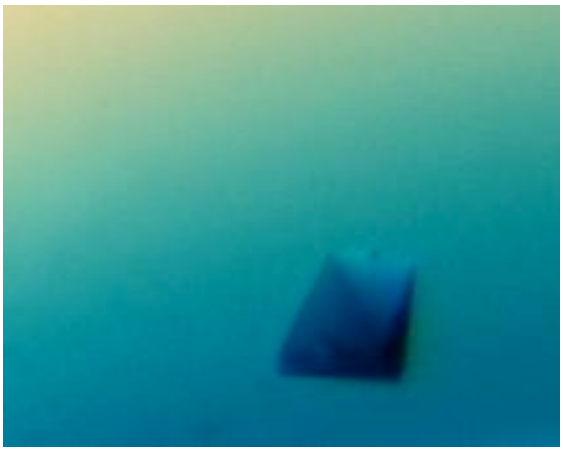


Figure 3.9 Crystal of *Thermotoga maritima* mannitol dehydrogenase (TmMtDH) grown under condition 2 and dyed with Izit dye.
Crystal produced from 10 mg/mL TmMtDH grown by hanging drop vapour diffusion in crystallization Condition 2 (17.5% polyethylene glycol 1000 (w/v), 0.1 M sodium citrate at pH 4, 0.2 M sodium bromide and 10% glycerol). Crystals were dyed with 10% of the drop volume (0.2 μL) of Izit dye (Hampton Research). The image was taken at 40X magnification with a Nikon SMA1000 camera microscope.

3.2.3 Diffraction of TmMtDH

3.2.3.1 X-Ray Diffraction at the Structural Sciences Centre

Prior to diffraction studies, crystals were tested to determine whether they were protein or salt by adding Izit dye to the mother liquor. Protein crystals generally contain approximately 50% water. Due to the high content of water in a protein crystal, the Izit dye will diffuse through and turn the protein crystal blue. A salt crystal however has very little water content and will remain colorless upon the addition of Izit dye to the mother liquor. Condition 1 crystals were not able to be tested with Izit dye because only a single crystal grew in the original solution and this crystal was saved for diffraction experiments. However, Condition 2 crystals were easily grown in a short period of time and thus extra crystals could be tested with Izit dye. Condition 2 crystals dyed blue with Izit dye, indicating a strong likelihood that they were protein crystals (Figure 3.9).

The diffractometer at the SSSC has the ability to diffract crystals at either room temperature or under a cryo-stream at 100 K. The ability to experiment with diffraction at room versus low temperature allows the researcher to deduce if the cryo-protectant or cooling process in liquid nitrogen has affected the diffraction ability of the crystal. Poor cryo-protectants will not protect against the ice ring formation that will appear in the diffraction pattern of a crystal during low temperature experiments. Since Condition 2 crystals formed relatively easily, these crystals were tested under both low and room temperatures in diffraction experiments. Again, because only a single crystal formed by Condition 1, this crystal was saved for low temperature experiments only.

Occasionally during low temperature experiments, ice rings formed as demonstrated by the characteristic circular pattern in the diffraction. Flash annealing was performed on these icy crystals by covering the cryo-stream with a piece of cardboard for 2 to 3 seconds. This procedure allowed the crystal to rapidly warm and then cool again once the cryo-stream was released. Flash annealing was able to rid the diffraction pattern of ice rings. Thus, it was deduced that the cryo-protectant added to the Condition 2 crystals prior to flash cooling in liquid nitrogen was likely sufficient to prevent ice ring formation, and that the ice rings likely formed sometime during the transfer of a selected few crystals from the dewar of liquid nitrogen to the magnetic mount on the goniometer.

In general, diffraction experiments at the SSSC were carried out by a similar method. The crystal was manually centered in the X-ray beam using the visualization software. The detector distance was set to 7 cm and the exposure time varied from 30 to 90 seconds. Five to ten frames were collected for 0 and 90˚ to ensure that the crystal was centered in the beam and so that the frames could be used to determine the crystal system and space group if diffraction occurred.

Whether experiments were conducted at room temperature or under the cryo-stream, no diffraction was noted for any of the Condition 2 crystals mounted on the in-house diffractometer. Perhaps this was due to the small size of the crystals and the lower intensity of the X-ray source compared to that of a synchrotron source. The Condition 1 crystal that was tested under low temperatures diffracted to a very low resolution with only a minutely discernible protein diffraction pattern. Again, this is likely due to the relatively low intensity of the beam and the small crystal size.

3.2.3.2 X-Ray Diffraction at the Canadian Light Source

Both the Condition 1 crystal and several of the native crystals grown in the original Condition 2 were diffracted on beamline 08ID-1 at the Canadian Light Source (CLS). However, the crystals that formed under Condition 2 with TmMtDH that were co-crystallized with substrates and the crystals that formed with higher molecular weight polyethylene glycol in the variation of Condition 2 were not diffracted at the CLS because the crystals had not yet formed when the synchrotron beam time was available. Also, at the time of diffraction, the 08ID-1 protein crystallography beamline at the CLS was being commissioned so several software and hardware problems were noted at the time that the diffraction experiments took place.

Native Condition 2 crystals diffracted to approximately 5 to 6 Å resolution. It appeared that the crystals grown by the vapour diffusion method diffracted slightly better than the microbatch method. The vapour diffusion crystals were also less likely to have ice rings in the diffraction pattern and thus generally did not require flash annealing. Due to the low resolution of the Condition 2 diffraction, no data were collected for these crystals.

However, the crystal that formed in Condition 1 diffracted to a resolution of 3.3 Å. Data were collected under the following conditions: 0.5˚ frame width for 180˚, 7 second exposure

time, detector distance of 200 mm, and 1.331 nm wavelength. The data were then processed and merged together by Michel Fodje using the program XDS (Kabsch, 1993).

3.2.4 Space Group and Diffraction Statistics of TmMtDH Data

As previously mentioned, data were collected for the Condition 1 crystal at the CLS. Data were scaled and integrated with the program XDS by Michel Fodje at the CLS. Data statistics are summarized in Table 3.2. The crystal was orthorhombic and belongs to the space group $P2_12_12_1$, which indicated a two-fold screw axis about each of the three orthogonal crystal axes. The unit cell dimensions were a = 83.43 Å, b = 120.61 Å, c = 145.76 Å. Unfortunately, the structure was not solved by the completion of this project.

Table 3.2. Data collection statistics for the crystal of TmMtDH grown in condition 1. Values in parentheses are for the highest resolution shell.

Temperature (K)	150
Beamline	08ID-1, Canadian Light Source
Detector	MAR 225 CCD
Space Group	$P2_12_12_1$
Unit-cell parameter (Å)	a = 83.43, b = 120.61, c = 145.76
Matthews Coefficient ($Å^3Da^{-1}$)	2.64
Solvent content (%)	53
Unit-cell volume ($Å^3$)	1470000
Molecular Mass (Da)	34870 [317 amino-acid residues]
No. of molecules in the asymmetric unit	4
No. of measured reflections	314666
Total No. of unique reflections	47727
Resolution Range (Å)	7.0–3.3 (3.8–3.3)
Completeness	99 (92.4)
Redundancy	7.20 (6.59)
R_{merge}*	0.024 (0.027)
$\langle I/\sigma(I)\rangle$	8.46 (6.55)

* Rmerge = $|I - \langle I\rangle|/\Sigma\langle I\rangle$, where $\langle I\rangle$ is the average intensity over symmetry-related reflections and I is the measured intensity.

4.0 Discussion

4.1 A-Raf

4.1.1 Purification of A-Raf

After observing the doublet of bands continuously present after purification of full-length A-Raf, several hypotheses were made. The predicted molecular weight of full-length A-Raf is approximately 68 kDa. It was interesting that the doublet of bands present in all SDS-PAGE gels were of a slightly higher and slightly lower molecular weight that the predicted molecular weight of full-length A-Raf according to Coomassie staining. The 75 kDa and 66 kDa bands could have been samples of full-length GST-A-Raf that had been degraded to different amounts during purification, although it is difficult to assess the accurate mass of these bands by SDS-PAGE methods. Further analysis in the form of Western blots and mass spectrometry analysis were therefore required to obtain more information, because the sensitivity of Western blots is much greater than the sensitivity of SDS-PAGE analysis. It is interesting to note that these two bands of 75 kDa and 66 kDa were also present in all SDS-PAGE gels of purification products of the CR2/CR3 domains of A-Raf. The predicted molecular weight of the CR2/CR3 domains of A-Raf alone is approximately 45 kDa so the third band present in these gels measuring approximately 45 kDa could then reasonably be attributed to the desired purification product.

At this point, several hypotheses were made as to why the A-Raf samples were not being completely purified and the scientific method was changed accordingly. Several steps were taken in earlier steps of the protein expression protocol to try to improve the yield of full-length protein before the column step was performed. Two different bacterial cell types were induced to create different environments for A-Raf, one completely oxidized and the other reduced. Neither cell type seemed to alter the result that was seen with the original purification protocol. It was then hypothesized that perhaps decreasing the concentration of IPTG used for induction and decreasing the time for induction might produce better quality, full-length protein. Unfortunately, in reality this change also failed because so little protein was produced that it was difficult to determine if the protein was in fact of better quality. In addition, because of slight changes to protein quality that might occur when different batches of protein were produced, it was deemed unlikely that any crystals that might grow from this batch of protein might never be repeated on subsequent batches. Finally, an attempt to purify A-Raf by high-

pressure affinity chromatography was also unsuccessful because the GST-tagged protein would not bind to the high-pressure glutathione Sepharose column.

The low purification yield (approximately 1.4 mg CR2/CR3 or 0.4 mg full-length A-Raf from a 1 L induction) made A-Raf very difficult to work with. Subsequent work on the protein, including Western blot and mass spectrometry studies allowed further characterization and it was hoped that information from these new analyses could provide additional important information on how the purification protocol might further be altered finally to completely purify the protein of interest. Again, in retrospect perhaps purification problems were associated more with low yields of induction and poor cleavage of the GST tag than stability. In addition, the extra bands on SDS-PAGE may have been due to uncleaved yet partially degraded GST-A-Raf or GST-CR2/CR3 or perhaps even GST-PreScission or GST itself.

4.1.2 A-Raf Mass Spectrometry Analysis

Due to the fact that the identity of the 75 kDa and 66 kDa protein bands resolved by SDS-PAGE from the purification products of full-length A-Raf and the CR2/CR3 domains of A-Raf could not be determined solely by SDS-PAGE, attempts at mass spectrometry analysis were made. During this procedure, the protein bands of interest were first excised from the SDS-PAGE gel and digested by the enzyme trypsin. Trypsin cleaves only on the carbonyl side of lysine and arginine residues and produces peptides that are characteristic of the protein from which they were cleaved. These peptides can then be analyzed by MALDI-TOF MS to produce a unique peptide mass fingerprint that can be used to search through databases to identify a protein.

It was thought that perhaps trying to analyze only the CR2/CR3 domains of A-Raf instead of the full-length protein could have been problematic because the CR2/CR3 domains alone are not present in the sequence analysis searched used by the Mascot programs. Thus, manual attempts at peptide fragment sequence alignments were attempted using ClustalW. However, these trypsin-digested peptide fragments could not be manually aligned to peptides in the CR2/CR3 sequence.

As previously discussed, mass spectrometry results have been difficult to interpret and understand. Initial mass spectrometry results from the work of Dr. Sanjukta Aich seemed promising at the time. Dr. Aich resolved the same two bands by SDS-PAGE from purification products, but was able to characterize the 66 kDa band as A-Raf from mass spectrometry

results. Further mass spectrometry results proved less successful as each time the trypsin-digested peptides matched to a different protein in the database, despite changes to the purification protocol. In retrospect, further use of Western blots at an earlier stage of this project may have provided more useful information than repeated mass spectrometry analysis.

It was suggested that perhaps too many attempts at MALDI-TOF MS analysis were made in the quest to determine the identity of the 75 kDa and 66 kDa protein bands. So many attempts at MS analysis were made in essence because the first MALDI-TOF MS analysis was successful in proving that full-length A-Raf was indeed one of the products present after A-Raf purification. It was hoped that if slight modifications were made in the purification protocol that A-Raf could again be identified in the purification products and that eventually the purification protocol could be modified enough so that a single product formed: the full-length A-Raf protein. Unfortunately, after the numerous MALDI-TOF MS studies, this was not the case and further MS studies and purification modifications were abandoned.

In terms of implications for the study of A-Raf by X-ray crystallography, the fact that full-length A-Raf could only be identified from mass spectrometry results after the initial purification procedure performed by Dr. Aich, suggested that this protein was most likely not suitable for crystallographic studies. It is important that proteins targeted for crystallographic studies be of the highest purity possible. During the mass spectrometry studies, the fact that the excised bands from the SDS-PAGE gels of purification products could neither be matched to A-Raf or even to the same protein in the database with each analysis confirmed the fact that A-Raf was very difficult to purify and would be unlikely to crystallize.

4.1.3 Western Blot Analysis of A-Raf

Western blot analyses are useful when trying to identify a protein when no activity assay is available. In the case of the Raf kinases, activity is assayed by analyzing ability to phosphorylate MEK. A-Raf is a much weaker activator of MEK in comparison to the other Raf kinases and thus is hard to analyze through this type of activity assay (Chong *et al.*, 2003). Western blots use antibodies that are specific for a protein of interest. In this case, the AP1 antibody is specific for residues in the CR2/CR3 domains of A-Raf and the C-20 antibody is specific for the C-terminal 20 amino acids of A-Raf. Thus, it was hoped that analysis by Western blot would provide more information about the identity of the unknown 75 kDa and 66 kDa purification products that hindered progress.

The SDS-PAGE gel of the CR2/CR3 domains of A-Raf after purification showed fewer bands compared to the numerous antibody-detected bands in the Western blot of the same sample. It should be noted that all Western blots were performed on protein that had been stored at –80 ˚C with 10 % glycerol for a week prior to being used in the blot. The fact that the protein might have further degraded during the incubation period between purification and the blot or during thawing from –80 ˚C to 4 ˚C might explain how the antibodies bound to more protein bands in the blot or these additional bands could be explained by the increased sensitivity of Western blot over SDS-PAGE analysis.

As previously mentioned, in the SDS-PAGE gel of the purification products, two protein bands were resolved measuring 75 kDa and 66 kDa. The AP1 antibody, which recognizes the CR2 and CR3 domains of A-Raf, bound to two bands measuring approximately 75 kDa and 66 kDa. The C-20 antibody, which recognizes the C-terminal 20 amino acids of A-Raf, bound to two bands measuring approximately 66 kDa and 47.5 kDa. These results suggest that perhaps both the 75 kDa and 66 kDa bands of the original SDS-PAGE analysis contained full-length A-Raf to some degree.

In the Western blot of the CR2/CR3 domains of A-Raf, AP1 antibodies bound to two protein bands measuring approximately 45 kDa and 38 kDa. The C-20 antibodies bound to four protein bands measuring approximately 47.5 kDa, 40 kDa, 30 kDa, and 24 kDa. The fact that the AP1 and C-20 antibodies bound to these protein bands suggests that the original 47 and/or 45 kDa band resolved by SDS-PAGE was most likely composed of the CR2/CR3 domains of A-Raf but that the protein was degraded. In addition, the SDS-PAGE gel performed after purification always resolved two higher molecular weight bands of 75 kDa and 66 kDa but neither of these bands bound to the AP1 or C-20 antibody during the Western blots. This suggests that the 75 kDa and 66 kDa bands are composed of contaminating proteins, instead of GST-CR2/CR3 fusion proteins as previously predicted.

4.1.4 Crystallization of A-Raf

Despite all evidence indicating that A-Raf was likely not pure enough for crystallization studies, commercial crystallization screening kits were still performed on the little purified protein that accumulated from many purification trials. It was believed that the next step should be taken and screening should be tried just to see what would happen. Several crystals did form during these crystallization trails and protocols for harvesting and flash cooling the crystals

were followed. Unfortunately, all crystals were characterized as salt after diffraction at the Saskatchewan Structural Sciences Centre.

These results are not surprising when considering the previous characterizations made during purification, mass spectrometry and Western blot analyses. All results suggested that A-Raf was not completely purified and was therefore unlikely to crystallize. In addition, proteins targeted for crystallographic studies should be 95% pure to ensure that the protein of interest is the protein that is being crystallized and that the sample remains homogenous and monodisperse. Again, previous results suggested that A-Raf was not at least 95% pure and thus unlikely to crystallize, although perhaps multiple purification steps, including exchange columns, affinity columns, gel filtration, and large-scale production, may have yielded different results.

4.2 *Thermotoga maritima* Mannitol Dehydrogenase

4.2.1 Gel Filtration of TmMtDH

Previous studies on enzymes from similar hyperthermophilic sources have shown that it is common for the enzyme to exist in both a dimeric and a tetrameric state (Karlstrom *et al.,* 2006). Specifically, the tetrameric form has been shown to be involved in thermal stability. In addition, TmMtDH has also been previously classified by sequence analysis as a medium-chain dehydrogenase reductase and, in general, MDRs are either dimers or tetramers in solution. The calibrated gel filtration analysis of TmMtDH confirmed the predicted tetrameric structure.

In primary gel filtration studies, the enzyme eluted as one peak that resolved into two bands on the SDS-PAGE gel. The two bands were also present on gels of purification products received from Dr. Claire Vieille at Michigan State University. This problem was solved by the addition of a higher salt concentration to the storage buffer of TmMtDH to improve enzyme stability. During further calibrated gel filtration analysis, it was again confirmed that the enzyme was a tetramer, as the enzyme eluted at a gel filtration volume that corresponded to the molecular weight of a tetramer of TmMtDH.

Previous structures of similar alcohol dehydrogenases also showed a tetrameric form in solution and in the crystal structure. For example, the structure of L-threonine dehydrogenase from the hyperthermophilic source *Pyrococcus horikoshii* (PhTDH) was recently characterized by Ishikawa *et al.* (2007). PhTDH shares many unique characteristics and 33% sequence identity with TmMtDH. PhTDH has also been characterized as a medium chain alcohol

dehydrogenase that consists of an NAD(H) binding domain and a catalytic domain that requires zinc for activity. Finally, PhTDH is a homo-tetramer in solution, similar to results for TmMtDH.

4.2.2 Crystallization of TmMtDH

Purified TmMtDH was received from Dr. Claire Vieille at Michigan State University. Dr. Vieille designed a protocol that purified TmMtDH so that only two protein bands resolved by SDS-PAGE. MALDI-TOF MS studies confirmed that both bands were TmMtDH with slightly different modifications. Dr. Vieille tried several different storage buffers to enhance the stability of TmMtDH and decided that a buffer containing 50 mM Tris at pH 8.5 was best. Therefore, 10 mg/mL TmMtDH in this 50 mM Tris at pH 8.5 storage buffer was initially screened against many commercially available screening kits. After further evaluation, Dr. Vieille realized that a higher salt content in a storage buffer would probably help to stabilize the hyperthermophilic protein. Dr. Vieille thus used a storage buffer containing 50 mM Na_2HPO_4 at pH 7.0 and 0.2 M NaCl. Later analyses of TmMtDH revealed that the choice Dr. Vieille made in changing the buffer was a correct one, as SDS-PAGE gels of purification products in the higher salt storage buffer resolved a single purified protein band. Thus, the higher salt storage buffer was used for later crystallization trials.

It has been theorized that enzymes from hyperthermophilic sources are easier to crystallize because of their high stability (Mukherjee & Guptasarma, 2005). TmMtDH seemed to follow this reasoning as the enzyme crystallized in several conditions found in different commercially available screening kits. The technique of microbatch-under-oil was used for most screening kits to preserve the solutions purchased from manufacturers. In microbatch-under-oil, a small volume of the protein solution is placed under oil. A small volume of precipitant solution is then mixed with protein drop. In theory, as evaporation through the oil occurs, the precipitant and protein solution should reach a state of supersaturation where nucleation and crystal formation occurs. The extent of evaporation can be controlled somewhat by altering the concentration and type of oil used during microbatch-under-oil experiments. Mineral oil, or liquid paraffin, is slow to evaporate and thus increases the time until supersaturation could theoretically be reached. A mixture of silicon oil and paraffin oil will speed up evaporation and decrease the time until supersaturation could theoretically be reached. In the dry Saskatchewan environment, most microbatch-under-oil TmMtDH experiments were

executed with 100% mineral oil because a mixture of oils evaporated too quickly and supersaturation of the protein solution could not be reached.

As an alternative to microbatch-under-oil, hanging drop vapour diffusion was used for several crystallization trials. In this method, the precipitant solution and protein solution are mixed in a 1:1 ratio on a plastic cover slip before being inverted over the precipitant solution. The concentration difference between the precipitant solution in the well and the precipitant solution mixed with the protein solution causes supersaturation and theoretically, nucleation and thus crystal formation. Hanging drop vapour diffusion was not used as often as microbatch-under-oil, especially for initial crystallization screening attempts, because it uses a larger volume of precipitant solution. In addition, because supersaturation generally occurs much later in hanging drop vapour diffusion experiments than in microbatch-under-oil experiments, the composition of the drop can change quite dramatically over time. These changes can include important components to crystallization like precipitant concentration and pH.

For all protein samples, 10% glycerol was added to avoid freezing damage to the protein, as glycerol is a cryo-protectant. After purification, glycerol did not seem to affect the quality of protein or its ability to form crystals as demonstrated by the crystallization of TmMtDH in both Condition 1 and Condition 2 and the ability to collect data on Condition 1 crystals.

4.2.2.1 Crystallization Condition 1

TmMtDH first crystallized most successfully in a solution based on condition G4 of the Nextal MPD Screening Kit (Qiagen Biotechnologies). A solution containing 10 mg/mL TmMtDH in a storage buffer of 50 mM Tris at pH 8.5 crystallized in a precipitant solution containing 30% MPD and 0.1 M Na-HEPES at pH 7.5 in the microbatch-under-oil method at room temperature. As previously discussed, this crystallization condition produced a single crystal and could not be reproduced despite many attempts. Several theories have been previously suggested as to why this occurred, including the change in storage buffer of the protein and possible concentration changes to the precipitant solution due to screening kit evaporation over time (see section 3.2.2.1). In addition, different purified batches of protein can behave slightly differently in crystallization conditions. The crystal that initially formed was from the first batch of protein received from Michigan while subsequent attempts to re-

crystallize the protein under the same condition were completed with a separate batch of protein. However, despite trying to repeat the crystallization conditions either manually or again with the Nextal MPD Screening Kit, the one original crystal that grew was of sufficient quality to diffract and collect data.

4.2.2.2 Crystallization Condition 2

Similar to the discovery of Condition 1, Condition 2 was a variation of a screening solution that was one of 1536 conditions present in the Hauptmann-Woodward Institute High-throughput Crystallization Screening Facility. This condition was easy to reproduce manually in the laboratory, unlike Condition 1. Condition 2 crystals also seemed to grow under different slight variations in incubation temperature, polyethylene glycol molecular weight and concentration, and incubation time. In addition, unlike difficulties associated with Condition 1, Condition 2 crystals grew from two different purification batches of TmMtDH protein and in two different storage buffer solutions. The original Condition 2 TmMtDH crystals grew from 10 mg/mL TmMtDH in a storage buffer solution of 50 mM Tris at pH 8.5 while the later Condition 2 crystals grew from 10 mg/mL TmMtDH in a storage buffer solution of 50 mM Na_2HPO_4 at pH 7.0 and 0.2 M NaCl in the same precipitant solution.

4.2.3 Diffraction of TmMtDH

Unfortunately, despite the fact that Condition 2 crystals grew to a larger size and under a variety of slightly different crystallization conditions with apparent ease, these crystals did not diffract to the same high resolution and to the same quality as the one Condition 1 crystal that formed from the Nextal MPD Screening Kit. A well-formed crystal is characterized by well-defined borders and no cracks. Also, a well-formed crystal is a single crystal and not twinned. A twinned crystal is one in which two crystals grow together from the same nucleation point. Both Condition 1 and Condition 2 crystals had well defined borders and did not appear to be cracked or twinned. It must have therefore been a difference in the internal packing of the crystal that made the Condition 1 crystal diffract to a higher resolution than the Condition 2 crystal. For more information on crystal packing and how this affects diffraction quality, please see Appendix 1.

4.2.4 Characterization of TmMtDH Crystal Data

Dr. Lata Prasad of the laboratory of Dr. Louis Delbaere attempted to solve the TmMtDH data but she was unsuccessful before the completion of this project. Although the structure of TmMtDH could not be solved, the data that were collected were used for a primary characterization of TmMtDH crystals. The space group, unit cell parameters, Matthews coefficient, solvent content, unit cell volume and number of molecules in the asymmetric unit all revealed important structural information for the primary structural characterization of TmMtDH.

The space group was determined to be $P2_12_12_1$ with a unit cell dimensions a = 83.43 **Å**, b = 120.61 **Å**, c = 145.76 **Å**. This is an orthorhombic space group with three two-fold screw axes present. In addition, systematic absences due to crystallographic symmetry exist under the following conditions for this space group: h00 present only when h=2n, 0k0 present only when k=2n and 00l present only when l=2n. Therefore, this space group was confirmed from the point group 222 by the presence of these systematic absences in the reflection data file. More information on the determination of space groups and symmetry can be found in Appendix 1.

Matthews first defined the Matthews Coefficient (V_M) as the crystal volume per unit of protein molecular weight. Matthews also showed that there is a relationship between V_M and the fractional volume of solvent in a crystal (Matthews, 1969). Thus, V_M is useful in estimating the number of molecules in the unit cell because crystals generally contain a solvent content of approximately 50%, although this value may not be accurate for higher molecular weight proteins because they generally include a higher solvent content. The molecular weight of TmMtDH is 34 kDa and is therefore well within the range of molecular weights that conform to the predictions of the Matthews coefficient. The specific volume of TmMtDH was calculated to be approximately 1470000 $Å^3$ based on the measurements from the unit cell. The Matthews coefficient was subsequently calculated to be 2.64 $Å^3$/Da. These results again suggest that TmMtDH is a tetramer, as was expected from previous studies and the characterization of similar enzymes, like PhTDH and *Agaricus bisporus* MtDH.

Unfortunately, it was determined that although the Condition 1 crystal of TmMtDH was of sufficient quality for diffraction experiments and primary structural characterization, the data collected were not of adequate quality to solve the structure. This was in part due to the fact that at the time the structure was to be solved, the only previously solved structures available in the Protein Data Bank contained minimal sequence identity with TmMtDH, although the NAD-

binding domain and the alcohol dehydrogenase domain of previously solved MtDH structures could have been used to solve portions of TmMtDH by molecular replacement. The search engine 3D-PSSM (Kelley *et al.,* 2000) was used to help identify structures in the protein data bank that might be useful models for molecular replacement.

The top two results from 3D-PSSM were the structures of ketose reductase from Silverleaf Whitefly (PDB 1E3J) and of an NAD(+)-dependent alcohol dehydrogenase from *Bacillus stearothermophilus* (PDB 1RJW). The ketose reductase contained 24% sequence identity to TmMtDH while the alcohol dehydrogenase contained 23% identity with TmMtDH. The alcohol dehydrogenase did share some structural characteristics with TmMtDH, including a zinc-binding domain. It was thought that perhaps the low sequence identity could be overcome by combining the models from the two structures to elucidate initial phases for the TmMtDH structure. Unfortunately this attempt was unsuccessful.

Later, the search engine 3D-Jigsaw (Bates *et al.*, 2001) was used to find another potential model for molecular replacement once new TmMtDH data was collected. The program identified the structure of L-Threonine Dehydrogenase from *Pyrococcus horikoshii* (PhTDH) that had only recently been deposited in the protein data bank (PDB 2DFV). Due to the many structural domain similarities and the homo-tetrameric nature of PhTDH and TmMtDH, it is hoped that PhTDH will prove to be a better model for molecular replacement when better data can be collected for TmMtDH.

Although the crystal structure of TmMtDH was not solved, it is expected that the enzyme will exhibit a tetrameric structure in the asymmetric unit. The X-ray structure of PhTDH revealed overall crystallographic symmetry showing a homo-tetrameric structure. It is therefore likely, based on the similarity between characterization results and gel filtration results comparing TmMtDH to PhTDH and *Agaricus bisporus* MtDH, that TmMtDH should exhibit similar tetrameric structural characteristics as these previously solved crystallographic structures.

In addition, both *Agaricus bisporus* MtDH and PhTDH contain dinucleotide-binding domains. The nucleotide-binding domains in these previously solved structures exhibit the typical Rossman fold, alternating alpha helices and beta sheets, as do other enzymes containing this specific domain. Therefore, it is likely that the nucleotide-binding domain of TmMtDH will contain a similar fold.

The structure of *Agaricus bisporus* MtDH (Horer *et al.*, 2001) also identified three residues termed the catalytic triad that are vital for catalysis: Ser149, Tyr169 and Lys173. A sequence alignment of *Agaricus bisporus* MtDH and TmMtDH (Figure 1.8) showed that only Tyr169 was conserved. However, manual adjustment of the sequence revealed two lysine residues, K188 and K191, at positions analogous to AbMtDH Ser149 and a threonine residue, T214, near AbMtDH Lys173 suggesting that perhaps a similar catalytic triad may be present in TmMtDH.

Finally, like TmMtDH, PhTDH is a zinc-containing enzyme. Previous results indicated that zinc could be important for TmMtDH activity, as a 20-minute incubation of EDTA with the enzyme reduced activity by 96%. In general, most NAD(H)-dependent alcohol dehydrogenases contain two zinc ions in each monomer: one zinc is termed the functional catalytic zinc and is present at the active site, while the other zinc is termed structural and is located in a loop that is important for tetramer formation (Ishikawa *et al.*, 2007). Only one zinc ion was present in the structure of PhTDH, the structural zinc, and this ion is coordinated by four cysteine residues, Cys97, Cys100, Cys103, and Cys111. A sequence alignment between PhTDH and TmMtDH reveals that these cysteine residues are conserved with Cys84, Cys87, Cys90, and Cys98 in TmMtDH (Figure 4.1). This suggests the TmMtDH enzyme may also coordinate a zinc ion using the analogous four cysteine residues and thus have a similar structure in this region of the protein.

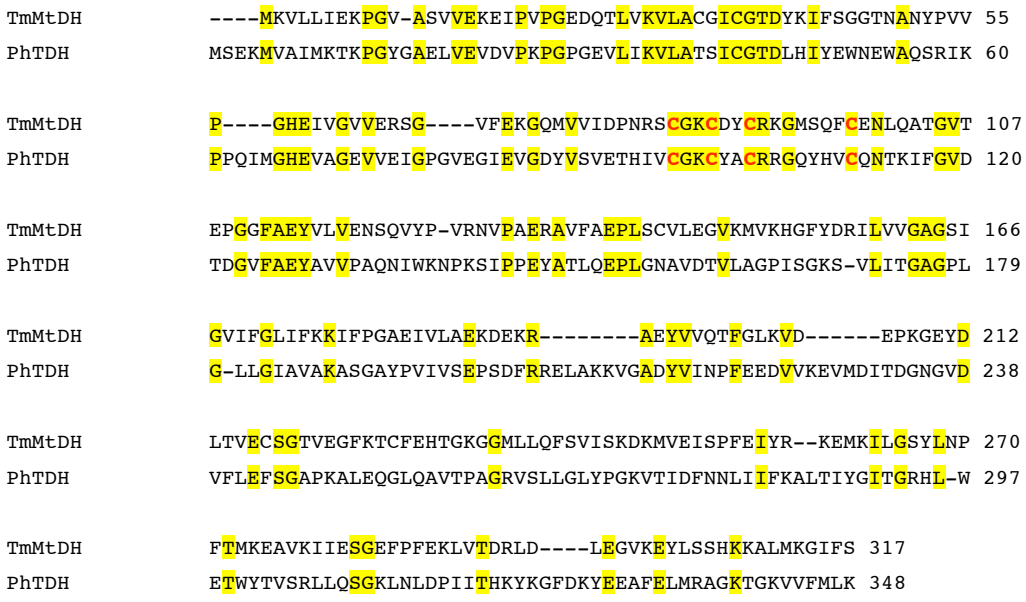


Figure 4.1 Sequence alignment of *Thermotoga maritima* mannitol dehydrogenase (TmMtDH) and *Pyrococcus horikoshii* L-threonine dehydrogenase (PhTDH).
There exist four cysteine residues (Cys84, 87, 90, and 98) shown to be involved in coordination of the structural zinc ion in PhTDH (shown in red). The alignment was performed using the program ClustalW (Thompson *et al.*, 1994). Amino acid residues conserved in TmMtDH and PhTDH are highlighted in yellow.

5.0 Overall Summary, Future Work and Conclusions

5.1 A-Raf

The initial objectives of the A-Raf project were to purify, crystallize and solve the structure of full-length A-Raf or a combination of the conserved regions of A-Raf. Most work concentrated on the purification aspect of this project; theory and practice suggest that a protein of 95% purity is ideal for crystallization trials to be successful. Unfortunately, despite alterations to the induction and expression protocol, contaminating proteins continued to hinder progress. An SDS-PAGE analysis of the purification products showed a high excess of these contaminating proteins in comparison to the expected A-Raf product. In addition, numerous attempts to determine the identity of these contaminating bands were made through MALDI-TOF MS analyses. The original positive result indicating that one of the bands resolved by SDS-PAGE of the purification products of full-length A-Raf could not be repeated despite purification protocol changes.

In retrospect, further alterations to the purification protocol could have been attempted, including using multiple affinity column steps to achieve the desired 95% protein purity for crystallization studies. The extra bands that separated on SDS-PAGE gel could likely have been GST-fusion proteins that were uncleaved by the PreScission protease. In addition, the low expression yield made analysis difficult.

It was thus determined that A-Raf was not a suitable target for crystallization studies and the project was halted. Since the termination of the A-Raf project, a structure of the N-terminal Ras Binding Domain of A-Raf has been deposited in the protein data bank (PDB 1WXM). Zhao *et al.* (to be published, 2005) from the Riken Structural Genomics/Proteomics Initiative solved the structure through NMR techniques. The A-Raf Ras binding domain has been characterized as an alpha-beta roll, homologous with the catalytic subunit of phosphatidylinositol 3-kinase (Figure 5.1). Little other information is known at this point because the primary citation listed in the protein data bank has not yet been published.

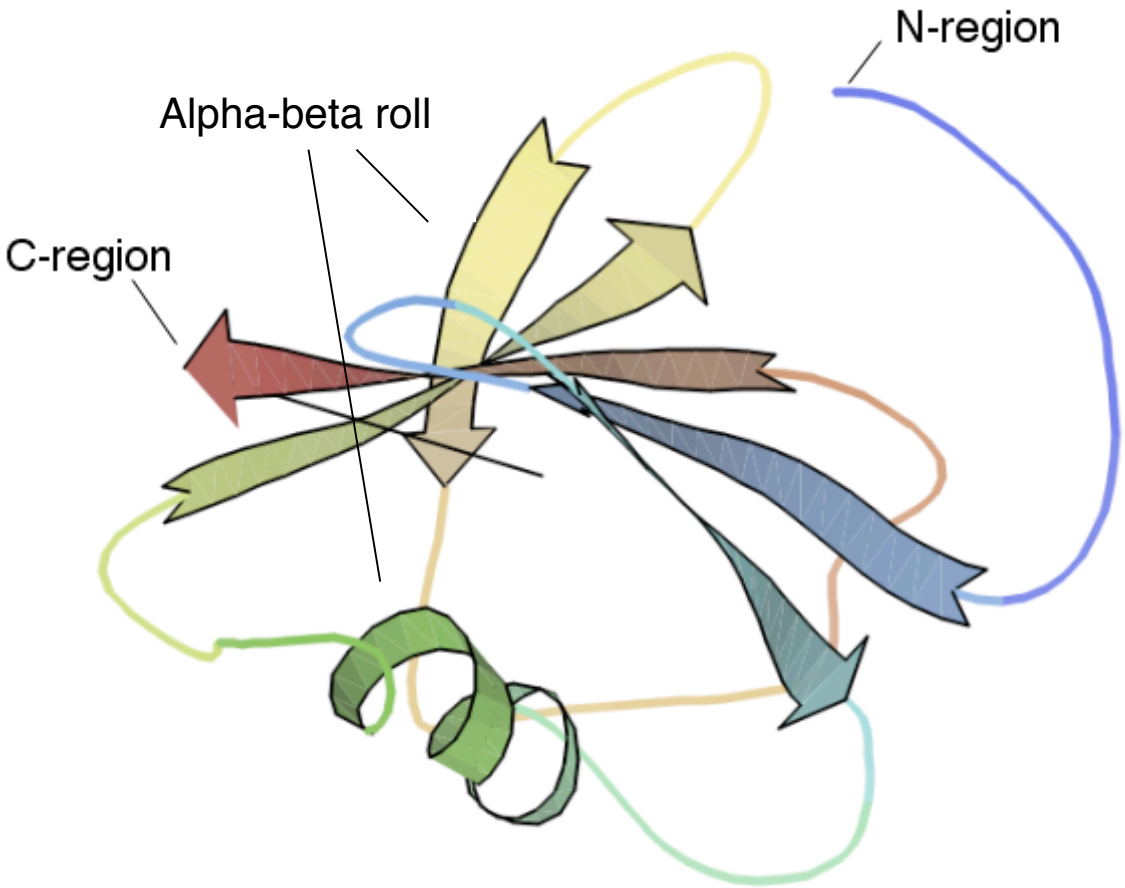


Figure 5.1 NMR structure of the Ras-binding domain of A-Raf.
The structure of the Ras-binding domain of human A-Raf has been characterized as an alpha-beta roll, homologous with the catalytic subunit of phosphatidylinositol 3-kinase (PDB accession number 1RFA).

5.2 *Thermotoga maritima* Mannitol Dehydrogenase

The primary objectives of the TmMtDH project reported herein were to crystallize, diffract and solve the structure of TmMtDH. The protein provided for study was originally purified sufficiently by Dr. Claire Vieille at Michigan State University, but alterations to the storage buffer by the addition of more salt increased the stability of the protein prior to crystallization studies.

Crystallization trials originally concentrated on finding conditions from commercially available sparse matrix crystallization screens. The protein was also sent to the Hauptmann-Woodward high-throughput crystallization facility. Two crystallization conditions formed crystals large enough for diffraction experiments: Condition 1 (30% MPD and 0.1 M Na-HEPES at pH 7.5) was based on condition G4 from the Nextal MPD Screening kit and Condition 2 (17.5% polyethylene glycol 1000 (w/v), 0.1 M sodium citrate at pH 4, 0.2 M sodium bromide and 10% glycerol) was based on one of the 1536 conditions from the high-throughput screen.

Manual attempts to repeat crystal formation from crystallization Condition 1 were unsuccessful despite many attempts, while experiments to repeat crystal formation from crystallization Condition 2 were successful. It was also determined that Condition 2 could be slightly optimized by varying polyethylene glycol molecular weight and incubation temperature to produce larger crystals in a smaller amount of time. Finally, Condition 2 was also used to co-crystallize TmMtDH with substrates like NADP and zinc chloride.

Crystals from Condition 1 and Condition 2 were screened at the in-house diffractometer at the Saskatchewan Structural Sciences centre. No crystals from either crystallization condition produced a discernible diffraction pattern on this source, likely due to the fact that it is not as intense an X-ray source as a synchrotron source. The crystals from both crystallization conditions were also diffracted at the Canadian Light Source. Condition 2 crystals diffracted to approximately 5 to 6 Å resolution while the Condition 1 crystal diffracted to approximately 3.3 Å resolution. Data were subsequently collected for the Condition 1 crystal and preliminary characterization of TmMtDH was possible. The crystal belonged to the space group $P2_12_12_1$ with unit cell dimensions a = 83.43 Å, b = 120.61 Å, c = 145.76 Å.

Future work includes growing better quality crystals for diffraction experiments, including trying new crystallization conditions, change cryo-conditions and altering growth conditions like temperature. A better model for the calculation of initial phases by molecular

replacement has been identified using the program 3D-Jigsaw. The structure of L-threonine dehydrogenase from *Pyrococcus horikoshii* contains 33% sequence identity to TmMtDH and shares multiple domain similarities. Thus, when better quality data can be collected for TmMtDH, it should be possible to solve the structure by molecular replacement. Further characterization of the enzyme should then be possible in order to answer questions about substrate specificity and the potential ability of the protein to act as a bioreactor for mannitol production in a commercial setting. Structural information will give information about the function of TmMtDH, giving insight for structural manipulations to allow the enzyme to produce increased amount of mannitol for commercial use.

6.0 References

Anderson, D. (2006) Role of lipids in the MAPK signaling pathway. Prog. Lip. Res. *45*, 102-119.

Apweiler R, Attwood TK, Bairoch A, Bateman A, Birney E, Biswas M, Bucher P, Cerutti L, *et al*. (2001) The InterPro database, an integrated documentation resource for protein families, domains and functional sites. Nucl. Acids Res. *29*, 37-40.

Baljuls, A., Schmitz, W., Mueller, T., Zahedi, R.P., Sickmann, A., Hekman, M., and Rapp, U.R. (2008) Positive regulation of A-Raf by phosphorylation of isoform-specific hinge segment and identification of novel phosphorylation sites. J. Biol. Chem. M801782200V1.

Bateman, A., Coin, L., Durbon, R., Finn R., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E.L.L., Studholme, D.J., Yeats, C., and Eddy, S.R. (2004) The Pfam protein families database. Nucl. Acids Res. *32*, D138-141.

Bates, P.A., Kelley, L.A., MacCallum, R.M., and Sternberg, M.J. (2001) Enhancement of protein modeling by human intervention in applying the automatic programs 3D-JIGSAW and 3D-PSSM. Proteins *5*, 39-46.

Baumchen, C. and Bringer-Meyer, S. (2007) Expression of glf Z.m. increases D-mannitol formation in whole cell biotransformation with resting cells of *Corynebacterium glutamicum*. Appl. Microbiol. Biotechnol. *76*, 545-552.

Baum, C., Roth, A.H., Biedendieck, R., Malten, M., Follmann, M., Sahm, H., Bringer-Meyer, S., and Jahn, D. (2007) D-mannitol production by resting state whole cell biotrans-formation of D-fructose by heterologous mannitol and formate dehydrogenase gene expression in *Bacillus megaterium*. Biotechnol. J. *2*, 1408-1416.

Beck, T.W., Huleihel, M., Gunnell, M., Bonner, T.I., and Rapp, U.R. (1987) The complete coding sequence of the human A-raf-1 oncogene and transforming activity of a human A-raf carrying retrovirus. Nucl. Acids Res. *15*, 595-609.

Bivona, T. and Phillips, M. (2003) Ras pathway signaling on endomembranes. Cur. Opinion Cell Biol. *15*, 136-142.

Bjorge, J.D., Chan, T.O., Antczak, M., Kung, H.J., and Fujita, D.J. (1990) Activated type I phosphatidylinositol kinase is associated with the epidermal growth factor (EGF) receptor following EGF stimulation. Proc. Natl. Acad. Sci. USA *87*, 2816-3820.

Bogoyevitch, M.A., Marshall, C.J., and Sugden, P.H. (1995) Hypertrophic agonists stimulate the activities of the protein kinases c-Raf and A-Raf in cultured ventricular myocytes. J. Biol. Chem. *270*, 26303-26310.

Bosch, E., Cherwinski, H., Peterson, D., and McMahon, M. (1997) Mutations of critical amino acids affect the biological and biochemical properties of oncogenic A-Raf and Raf-1. Oncogene *15*, 1021-1033.

Bradford, M.M. (1976) A rapid and sensitive test for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. *72*, 248-254.

Burke, P., Schooler, K. and Wiley, H.S. (2001) Regulation of epidermal growth factor receptor signaling by endocytosis and intracellular trafficking. Mol. Biol. Cell *12*, 1897-1910.

Chamberlain, M.D., Berry, T.R., Paster, M.C., and Anderson, D.H. (2005) The p85 subunit of phosphatidylinositol 3'-kinase binds to and stimulates GTPase activity of Rab proteins. J. Biol. Chem. *279*, 48607-48614.

Chayen, N.E. (1999). Crystallization with oils: a new dimension in macromolecular crystal growth. J. Crystal Growth *196*, 434-441.

Chong, H., Vikis, H. and Guan, K. (2003) Mechanisms of regulating the Raf kinase family. Cell. Signal. *15*, 463-469.

Cioffi, C.L., Garay, M., Johnston, J.F., McGraw, K., Boggs, R.T., Hreniusk D., and Monia, B.P. (1997) Selective inhibition of A-Raf and C-Raf mRNA expression by antisense oligodeoxynucleotides in rat vascular smooth muscle cells: role of A-Raf and C-Raf in serum-induced proliferation. Mol. Pharm. *51*, 383-389.

Daum, G., Eisenmann-Tappe, I., Fries, H.W., Troppmair, J., and Rapp, U.R. (1994) The ins and outs of Raf kinases. Trends Bioch. Sci. *19*, 474-480.

Dumaz, N., Light, Y., and Marais, R. (2002) Mol Cell Biol 22, 3717-3728

Eck, M.J., Shoelson, S.E. and Harrison, S.C. (1993) Recognition of a high-affinity phosphotyrosyl peptide by the Src homology-2 domain of p56lck. Nature *363*, 87-91.

Emerson, S.D., Madison, V.S., Palermo, R.E., Waugh, D.S., Scheffler, J.E., Tsao, K.L., Kiefer, S.E., Liu, S.P., and Fry, D.C. (1995) Solution structure of the Ras-binding domain of c-Raf-1 and identification of its Ras interaction surface. Biochemistry *34*, 6911-6918.

Fabian, J.R., Daar, I.O. and Morrison, D.K. (1993) Critical tyrosine residues regulate the enzymatic and biological activity of Raf-1 kinase. Mol. Cell Biol. *13*, 7170-7179.

Fang, Y., Johnson, L., Mahon, E., and Anderson, D. (2002) Two phosphorylation-independent sites on the p85 SH2 domains bind A-Raf kinase. Biochem. Biophys. Res. Commun. *290*, 1267-1274.

Gardner, A.M., Vaillancourt, R.R., Lange-Carter, C.A., and Johnson, G.L. (1994) MEK-1 phosphorylation by MEK kinase, Raf, and mitogen-activated protein kinase: analysis of phosphopeptides and regulation of activity. Mol. Biol. Cell *5*, 193-201.

Garnett, M., Rana, S., Paterson, H., Barford, D., and Marais, R. (2005) Wild-type and mutant B-RAF activate C-RAF through distinct mechanisms involving heterodimerization. Mol. Cell. *20*, 963-969.

Ghosh, S., Basu, S., Strum, J., Base, S., and Bell, R. (1995) Identification of conditions that facilitate the expression of GST fusions as soluble, full-length proteins. Anal. Biochem. *225*, 376-378.

Gorti, S., Forsythe, E.L. and Pusey, M.L. (2003) Measurable characteristics of lysozyme crystal growth. Acta Cryst. D *61*, 837-843.

Guan, K.L., Figueroa, C., Brtva, T.R., Zhu, T., Taylor, J., Barber T.D., *et al*. (2000) Negative regulation of the serine/threonine kinase B-Raf by Akt. J. Biol. Chem. *275*, 27354-27359.

Hall, A. (1994) A biochemical function for ras -- at last. Science *264*, 1413-1414.

Hancock, J.F., Paterson, H., and Marshall, C.J. (1990) A polybasic domain or palmitoylation is required in addition to the CAAX motif to localize $p21^{ras}$ to the plasma membrane. Cell *98*, 69-80.

Helliwell, J.R. (2005) Protein crystal perfection and its application. Acta Cryst. D *61*, 793-798.

Hindley, A and Kolch, W (2002) Extracellular signal regulated kinase (ERK)/mitogen activated protein kinase (MAPK)-independent functions of Raf kinases. J. Cell Sci. *115*, 1575-1581.

Horer, S., Stoop, J., Mooibroek, H., Baumann, U., and Sassoon, J. (2001) The Crystallographic Structure of the Mannitol 2-Dehydrogenase $NADP^+$ Binary Complex from *Agaricus bisporus*. J. Biol. Chem. *276*, 27555-27561.

Hoy, R.S. (2002) <www.pha.jhu.edu/~robhoy/teaching/crystallattices.html> John Hopkins University, accessed November 17, 2005.

Huleihel, M., Goldsborough, M., Cleveland, J., Gunnell, M., Bonner, T., and Rapp, U.R. (1986) Characterization of murine A-raf, a new oncogene related to the v-raf oncogene. Mol. Cell Biol. *6*, 2655-2662.

Hunt, J.F., Weaver, A.J., Landry, S.J., Gierasch, L., and Deisenhofer, J. (1996) The crystal structure of the GroES co-chaperonin at 2.8 Å resolution. Nature *379*, 37-43.

Ishikawa, K., Higashi, N., Nakamura, T., Matsuura, T., and Nakagawa, A. (2007) The first crystal structure of L-threonine dehydrogenase. J. Mol. Biol. *366*, 857-867.

Janin, J. (1975) Radiocrystallographic analysis of protein structures. Biochimie *57*, 505-514.

Johnson, L., James, K., Chamberlain, M., and Anderson, D. (2005) Identification of Key Residues in the A-Raf Kinase Important for Phosphatidylinositol Lipid Binding Specificity. Biochemistry *44*, 3432-3440.

Jornvall, H., Persson, B. and Jeffery, J. (1987) Characteristics of alcohol/polyol dehydrogenases. The zinc-containing long-chain alcohol dehydrogenases. Eur. J. Biochem. *167*, 197-201.

Kabsch, W. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J. Appl. Crystallogr. *26*, 795-800.

Karlstrom, M., Steen, I.H., Madern, D., Fedo, A.E., Birkeland, N.K., and Ladenstein, R. (2006) The crystal structure of a hyperthermostable subfamily II isocitrate dehydrogenase from *Thermotoga maritima.* FEBS J. *273*, 2851-2868.

Kaup, B., Bringer-Meyer, S. and Sahm, H. (2005) D-Mannitol formation from D-glucose in a whole-cell biotransformation with recombinant *Escherichia coli*. Appl. Microbiol. Biotechol. *69*, 397-403.

Kavanagh, K. L., Klimacek, M., Nidetzky, B., and Wilson, D.K**.** (2002) Crystal structure of *Pseudomonas fluorescens* mannitol 2-dehydrogenase binary and ternary complexes. Specificity and catalytic mechanism. J. Biol. Chem. *277,* 43433–43442.

Kelley, L.A., MacCallum, R.M. and Sternberg, M.J.E. (2000) Enhanced genome annotation using structural profiles in the program 3D-PSSM enhanced. J. Mol. Biol. *299*, 499-520.

King, T., Fang, Y., Mahon, E., and Anderson, D. (2000) Using a phage display library to identify basic residues in A-Raf required to mediate binding to the Src homology 2 domains of the p85 Subunit of phosphatidylinositol 3'-kinase. Cell. Signal. *265*, 36450-36453.

Klimacek, M., Kavanagh, K., Wilson, D., and Nidetzky, B. (2003) On the role of Bronsted catalysis in *Pseudomonas fluorescens* mannitol 2-dehydrogenase. Biochem. J. *375*, 141-149.

Kolch, W. (2000) Meaningful relationships: the regulation of the Ras/Raf/MEK/ERK pathway by protein interactions. Biochem. J. *351*, 289-305.

Kulbe, K.D., Schwab, U. and Gudernatsch, W. (1987) Enzyme-catalyzed production of mannitol and gluconic acid. Product recovery by various procedures. Ann. N.Y. Acad. Sci. *506*, 552-568.

Lee, J., Koo, B., Kim, S., and Hyun, H. (2003) Purification and characterization of a novel mannitol dehydrogenase from a newly isolated strain of *Candida magnoliae*. Appl. Environ. Micro. *69*, 4438-4447.

Liu, W., Shen, X., Yang, Y., Xianglei, Y., Xie, J., Yan, J., Jiang, J., Liu, W., Wang, H., Sun, M., Zheng, Y., and Gu, J. (2004) Trihydrophobin 1 is a new negative regulator of A-Raf kinase. J. Biol. Chem. *279*, 10167-10175.

Luo, J. and Cantley, L.C. (2005) The negative regulation of phosphoinositide 3-kinase signaling by p85and its implications in cancer. Cell Cycle *4*, 1309-1312.

Marais, R., Light, Y., Paterson, H.F., and Marshall, C.J. (1995) Ras recruits Raf-1 to the plasma membrane for activation by tyrosine phosphorylation. EMBO J. *14*, 3136-3145.

Marais, R., Light, Y., Paterson, H.F., Mason, C.S., and Marshall, C.J. (1997) Differential regulation of Raf-1, A-Raf and B-Raf by oncogenic Ras and tyrosine kinases. Am. Soc. Biochem. Molec. Biol. Inc. *272,* 4378-4383.

Marais, R and Marshall, C.J. (1996) Control of the ERK MAP Kinase cascade by Ras and Raf. Cancer Surv. *27*, 101-125.

Matthews, B.W. (1969) Solvent content of protein crystals. J. Mol. Biol. *33*, 491-497.

Mayer, B.J., Jackson, P.K., Van Etten, R.A., and Baltimore, D. (1992) Point mutations in the abl SH2 domain coordinately impair phosphotyrosine binding *in vitro* and transforming activity *in vivo*. Mol. Cell. Biol. *12*, 609-618.

Mischak, H., Seitz T., Janosch, P., Eulitz, M., Steen H., Scherller, M., *et al*. (1996) Negative regulation of Raf-1 by phosphorylation of serine 621. Mol. Cell Biol. *16*, 5409-5418.

Mott, H.R., Carpenter, J.W., Zhong, S., Ghosh, S., Bell, R.M., and Campbell, S.L. (1996) The solution structure of the Raf-1 cysteine-rich domain: a novel Ras and phospholipids binding site. Proc. Natl. Acad. Sci. USA *93***,** 8312-8317.

Mukherjee, S. and Guptasarma, P. (2005) Direct proteolysis-based purification of an overexpressed hyperthermophile protein from *Escherichia coli* lysate: a novel exploitation of the link between structural stability and proteolytic resistance. Prot. Exp. Purif. *40*, 71-76.

Musacchio, A., Cantley, L. and Harrison, S. (1996) Crystal structure of the breakpoint cluster region-homology domain from phosphoinositol 3-kinase p85 α subunit. Proc. Natl. Acad. Sci. USA *93*, 14373-14378.

Nantel, A., Mohammad-Ali, K., Sherk, J., Posner, B.I., and Thomas, D.Y. (1998) Interaction of the Grb10 adapter protein with the Raf1 and MEK1 kinases. J Biol. Chem. *273*, 10475-10484.

Nassar, N (1995) The 2.2 A crystal structure of the Ras-binding domain of the serine/threonine kinase c-Raf1 in complex with Rap1A and a GTP analogue. Nature *375,* 554-560

Nordling, E., Jörnvall, H., and Persson, B**.** (2002) Medium-chain dehydrogenases/reductases (MDR). Eur. J. Biochem. *269***,** 4267–4276.

Paez, J. and Sellers, W. (2001) PI3K/PTEN/AKT PATHWAY: A critical mediator of oncogenic signalling. Cancer Treat. Res. *115*, 145-146.

Pawson, T. (2004) Specificity in signal transduction: from phosphotyrosine-SH2 domain interactions to complex cellular systems. Cell *110*, 669-672.

Persson, B., Krook, M. and Jörnvall, H. (1991). Characteristics of short-chain alcohol dehydrogenases and related enzymes. Eur. J. Biochem. *200***,** 537–543.

Pritchard, C.A., Samuels, M.L., Bosch, E., and McMahon, M. (1995) Conditionally oncogenic forms of the A-Raf and B-Raf protein kinases display different biological and biochemical properties in NIH 3T3 cells. Mol. Cell. Biol. *15*, 6430-6442.

Puttick, J., Vieille, C., Song, S.H., Fodje, M.N., Grochulski, P., and Delbaere, L.T.J. (2007) Crystallization, preliminary X-ray diffraction and structure analysis of *Thermotoga maritima* mannitol dehydrogenase. Acta Cryst. F *63*, 350-352.

Rao S and Rossmann M (1973). Comparison of super-secondary structures in proteins. J. Mol. Biol. *76*, 241–56

Reusch, H.P., Zimmermann, S., Schaefer, M., Paul, M., and Moelling, K. (2001) Regulation of Raf by Akt controls growth and differentiation in vascular smooth muscle cells. J. Biol. Chem. *6*, 6-11.

Rodriguez-Viciana, P., Warne, P.H., Dhand, R., Vanhaesebroeck, B., Gout, I., Fry, M.J., Waterfield, M.D., and Downward, J. (1994) Phosphatidylinositol-3-OH kinase as a direct target of Ras. Nature *370*, 527-532

Saha, B. (2004) Purification and Characterization of a Novel Mannitol Dehydrogenase from *Lactobacillus intermedius*. Biotechnol. Prog. *20*, 537-542.

Sasaki, Y., Laivenieks, M. and Zeikus, G. (2005) *Lactobacillus reuteri* ATCC 53608 *mdh* gene cloning and recombinant mannitol dehydrogenase characterization. App. Microbiol. Biotechnol. *68*, 36-41.

Sassoon, J., Horer, S., Stoop, J., Mooibroek, H., and Baumann, U. (2001) Crystallization and preliminary crystallographic analysis of Mannitol dehydrogenase (MtDH) from the common mushroom *Agaricus bisporus*, Acta Cryst. D. *57*, 711-713.

Schlessinger, J. (2000) Cell signaling by receptor tyrosine kinases. Cell *103*, 211-225.

Schelessinger, J. and Bar-Sagi, D., (1994) Activation of Ras and other signaling pathways by receptor tyrosine kinases. Cold Spring Harb. Quant. Biol. *59*, 173-179.

Schlessinger, J. and Lemmon, M.A. (2003) SH2 and PTB domains in tyrosine kinase signaling. Sci. STKE *2003*, RE12.

Schramm, K., Niehof, M., Radziwill, G., Rommel, C., and Moelling, K. (1994) Phosphorylation of c-Raf-1 by protein kinase A interferes with activation. Biochem. Biophys. Res. Commun. *201*, 740-747.

Shen, Q. (2003) Improving triplet-phase accuracy by symmetry observations in reference-beam diffraction measurements. Acta Cryst. A. *59*, 335-340.

Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V., and Mann, M. (2006). In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nature Protocols *1*, 2856-2860.

Soetaert, W., Buchholz, K. and Vandamme, E.J. (1995) Production of D-mannitol and D-lactic acid by fermentation with *Leuconostoc mesentoides*. Agrofood Hi-Tech. *6*, 41-44.

Soetaert, W., Vanhooren, P. and Vandamme, E. (1999) Production of mannitol by fermentation methods. Biotechnology *10*, 261-275.

Songyang, Z., Shoelson, S.E., Chaudhuri, M., Gish, G., Pawson, T. Haser, W.G., King, F., Roberts, T., Ratnofsk, S., and Lechleider, R.J. (1993) SH2 domains recognize specific phosphopeptide sequences. Cell *72*, 767-778.

Sridhar, S., Hedley, D. and Siu, L. (2005) Raf kinase as a target for anticancer therapeutics. Molec. Cancer Therap. *4*, 677-685.

Steelman, L.S., Pohnert, S.C., Shelton, J.G., Franklin, R.A., Bertrand, F.E., and McCubrey, J.A. (2004) JAK/STAT, Raf/MEK/ERK, PI3K/Akt and BCR-ABL in cell cycle progression and leukemogenesis. Leukemia *18*, 189-218.

Stokoe, D., Macdonald, S.G., Cadwallader, K., Symons, M., and Hancock, J.F. (1994) Activation of Raf as a result of recruitment to the plasma membrane. Science *264*, 1463-1467.

Storoni, L.C., McCoy, A.J. and Read, R.J. (2004) Likelihood-enhanced fast rotation functions. Acta Cryst. D. *60*, 432-438.

Studier, W.F., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. Meth. Enzymol. *185*, 60-89.

Sunn, H.W. and Plapp, B.V. (1992) Progressive sequence alignment and molecular evolution of the Zn-containing alcohol dehydrogenase family. J. Mol. Evol. *34*, 522-535.

Sutor, S.L., Vroman, B.T., Armstrong, E.A., Abraham, R.T., and Karnitz, L.M. (1999) A phosphatidylinositol 3-kinase-dependent pathway that differentially regulates c-Raf and A-Raf. J. Biol. Chem. *274*, 7002-2010.

Tavares, F.W., Bratko, D., Striolo, A., Blanch, H.W., and Prausnitz, J.M. (2004) Phase behavior of aqueous solutions containing dipolar proteins from second-order perturbation theory. J. Chem. Phys. *120*, 9859-9869.

Terada, T., Ito, Y., Shirouzu, M., Tateno, M., Hashimoto, K., Kigawa, T., Ebisuzaki, T., Takio, K., Shibata, T., Yokoyama, S., Smith, B.O., Laue, E.D., and Cooper, J.A. (1999) Nuclear

magnetic resonance and molecular dynamics studies on the interactions of the Ras-binding domain of Raf-1 with wild-type and mutant Ras proteins. J. Mol. Biol. *286*, 219-232.

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl. Acids Res. *22*, 4673-4680.

Vivanco, I. and Sawyers, C. (2002) The phosphatidylinositol 3-kinase-AKT pathway in human cancers. Nature Rev. Cancer. *2*, 489-501.

Wan, P.T., Garnett, M.J., Roe, S.M., Lee, S., Niculescu-Duvaz, D., Good, V.M., Jones, C.M., Marshall, C.J., Springer, C.J., Barford, D., and Marais, R. (2004) Mechanism of activation of the RAF-MAPK signaling pathway by oncogenic mutations of B-RAF. Cell *116*, 855-967.

Weymarn, F., Kiviharju, K., Jaaskelainen, S., and Leisola, M. (2003) Scale-up of a New Bacterial Mannitol Production Process. Biotechnol. Prog. *19*, 815-821.

Willumson, B.M., Christensen, A., Hubbert, N.L., Pagageorge A.G., and Lowy, D.R. (1984) The p21 Ras C-terminus is required for transformation and membrane association. Nature *310*, 583-586).

Yaffe, M. B., Rittinger, K., Volinia, S., Caron, P. R., Aitken, A., Leffers, H., Gamblin, S. J., Smerdon, S. J., and Cantley, L. C. (1997) Cell 91, 961-971.

Yan, M. and Templeton, D.J. (1994) Identification of 2 serine residues of MEK-1 that are differentially phosphorylated during activation by raf and MEK kinase. J. Biol. Chem. *269*, 19067-19073.

Yu, H., Chen, J.K., Feng, S., Dalgarno, D.C., Brauer, A.W., and Schreiber, SL. (1994) Structural basis for the binding of proline-rich peptides to SH3 domains. Cell *76*, 933-945.

Yu, J., Wjasow, C. and Backer, J.M. (1998) Regulation of the p85/p110alpha phosphatidylinositol 3'-kinase. Distinct roles for the n-terminal and c-terminal SH2 domains. J. Biol. Chem. *273*, 30199-30203.

Zamliny, Z. (2002) < http://www.chembio.uoguelph.ca/educmat/chm729/recip/vlad.htm> University of Guelph, accessed November 17, 2005.

Zhao, C., Tochio, N., Koshiba, S., Inoue, M., Kigawa, T., and Yokoyama, S. (2005) Solution structure of the N-terminal Ras-binding domain (RBD) in human A-Raf kinase. Protein Data Bank Accession 1WXM

Zimmermann, S., and Moelling, K. (1999) Phosphorylation and regulation of Raf by Akt (protein kinase B) Science *286*, 1741-1744.

Zheng, Y., Bagrodia, S. and Cerione, R. (1994) Activation of phosphatidylinositol 3-kinase activity by Cdc42Hs binding to p85. J. Biol. Chem. *269*, 18727-18730.

Appendix A

A.1 Theory of Crystallization

The rate-limiting step in X-ray crystallography is obtaining the protein or small molecule crystals required for diffraction. However, before beginning any crystallization trials, it is best if the protein of interest is purified to at least 95% purity. A protein of interest can be obtained by many means, such as being collected directly from its source, cloned into a bacterial or mammalian source, or from a collaborator.

In theory, all proteins can be crystallized. In practice however, this is not always the case. The crystallization process begins when homogenous solids in liquid begin to dissolve until an equilibrium point, the saturation point, is reached. At this saturation point, there is no net change in the number of molecules entering or leaving the solid phase but this is not the phase where crystals will grow. To drive crystal growth, one needs to push the system into supersaturation, where a drive in the solution system kinetics towards the solid phase creates a critical first nucleation point. This nucleation point is theoretically caused by the random contacts between quasi-stable aggregates. Because the hydrophobic interior of the aggregate is more favored, water will be pushed away to increase the drive of surface molecules to the interior of the aggregate. The volume of the aggregate will continue to grow until it stabilizes and becomes a nucleus. The farther the system is driven towards supersaturation, the more nucleation events that can occur (Figures A.1, A.2).

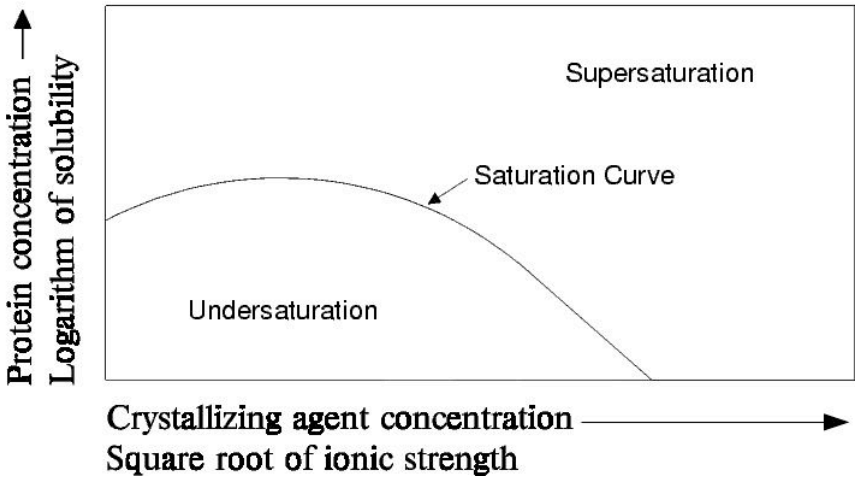


Figure A.1 Schematic representation of the relationship between protein concentration and crystallizing agent concentration and their effects on supersaturation
(From www-structmed.cimr.cam.ac.uk/Course/Crystals/Theory/phases.html).

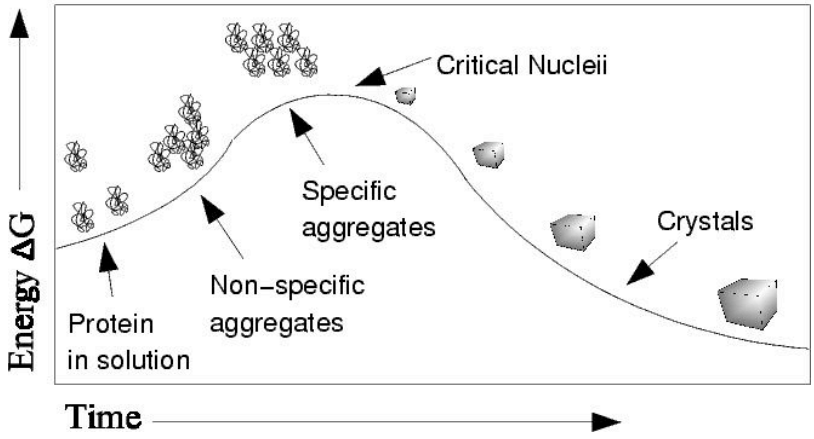


Figure A.2 Schematic representation of the relationship between free energy and time.
As the protein in solution reaches a state of supersaturation, aggregates form causing the formation of critical nuclei. These critical nuclei form the base for crystal growth (From www-structmed.cimr.cam.ac.uk/Course/Crystals/Theory/phases.html).

Crystal growth can occur by four methods: normal growth, which occurs only for small molecule crystals; 2-dimensional growth, where growth occurs as plates along the crystal surface; spiral growth dislocations, where the crystal plates grow in a spiral shape due to lower supersaturation levels; and 3-dimensional growth, where multiple nucleation events occur in basically the same space to form one crystal (Gorti, 2005). No matter how the crystal grows, it is important that surface defects, impurities and general crystal disorders are minimized in order to obtain the best possible crystals for diffraction.

To find appropriate crystallization conditions for a specific protein, most researchers will use commercially available crystallization screens, available from companies such as Hampton Research (www.hamptonresearch.com) and Sigma Aldrich Research (www.sigmaaldrich.com). Most commercial screens are sparse matrix screens that are based on results from previously crystallized proteins, or grid screens that test two precipitant solutions concurrently. Crystallization robots and commercial facilities, such as the Hauptman-Woodward Screening Facility in Buffalo, New York will perform crystallization screens for you using very small volumes of protein solution. Once a crystal or crystal-like condition has been determined from a screen, the condition must be optimized to obtain the best crystal possible for diffraction (Helliwell, 2005).

The most common crystallization technique is vapour diffusion. In principle, vapour diffusion begins when the protein solution is mixed with usually equal amounts of a precipitant solution called the well solution. The well solution can contain any combination of precipitants, salts, buffers, or other additives of a researcher's choice. In hanging drop vapour diffusion, the well solution is generally mixed with the protein solution in a 1:1 ratio and then suspended over the well. A time-dependent increase in the saturation level will ensue until enough water has diffused out of the drop to cause supersaturation and, hopefully, crystal growth.

The microbatch-under-oil technique is also a commonly used crystallization technique. In this method, a small volume consisting of a 1:1 ratio of protein to precipitant solution is dispensed in a microbatch plate and then covered with oil. Unlike the vapour diffusion method, the rate of evaporation of the experiment and thus rate of supersaturation is affected by the choice of oil, whether silicon, paraffin or a mixture.

A.2 Theory of X-Ray Diffraction

X-rays are very high-energy waves emitted between 1 and 10 nm in wavelength. X-ray diffraction occurs when X-rays strike the electrons in a crystal. The electrons in the crystal become secondary emitters, which scatter as waves that are proportional to the square of acceleration of the particle. Electrons from an atom will scatter in phase when they are experiencing the same electric field, causing portionive interference. The intensity of the X-ray scattering from a single atom is proportional to the square of the atomic number.

In 1913, Sir Bragg and his son derived Bragg's law to explain why the cleavage faces of crystals appear to reflect X-ray beams at certain angles of incidence. Bragg's Law equation is $n\lambda=2d\sin\theta$, where λ denotes the wavelength, and d is equal to the spacing between the parallel of the crystal. Bragg's law was derived in terms of an incident and reflected ray or beam to describe the resulting reflection of X-rays. As a consequence, the reflected rays make an angle of 2θ with the direct beam. The directions of portionive interference depend upon the size and shape of the unit cell, and the intensities of the reflections depend on the types of atoms present and their locations in the unit cell (Janin, 1975).

The unit cell is the smallest unit that is repeated by translation in three dimensions to make up a crystal. It consists of three sides (*a, b, c*) and three angles between the sides (α, β, γ). There are 14 possible Bravais lattices, which are arrangements of points in space (Figure A.3). There are seven different crystal systems: cubic, tetragonal, orthorhombic, rhombohedral, hexagonal, monoclinic and triclinic. There also exist four different lattice types: primitive (P), body-centered (I), face-centered (F) or side-centered (C). The 14 possible Bravais lattices are thus a combination of the seven crystal systems and the four lattice types. In addition, there exist added crystallographic symmetries like rotation, inversion and mirror planes that make up the 32 point groups. Finally, once glide planes and screw axes are added into the symmetry, there exist 230 possible space groups.

For example, $P2_12_12_1$ is an orthorhombic space group. This space group is of the primitive lattice type (P) and belongs to the point group 222. The $P2_12_12_1$ space group contains three two-fold screw axes that translate half of the unit cell and rotate by 180^0. An orthorhombic point group is defined by three perpendicular two-fold axes and/or perpendicular mirror planes with a unit cell of $a \neq b \neq c$ and $\alpha = \beta = \gamma = 90^0$.

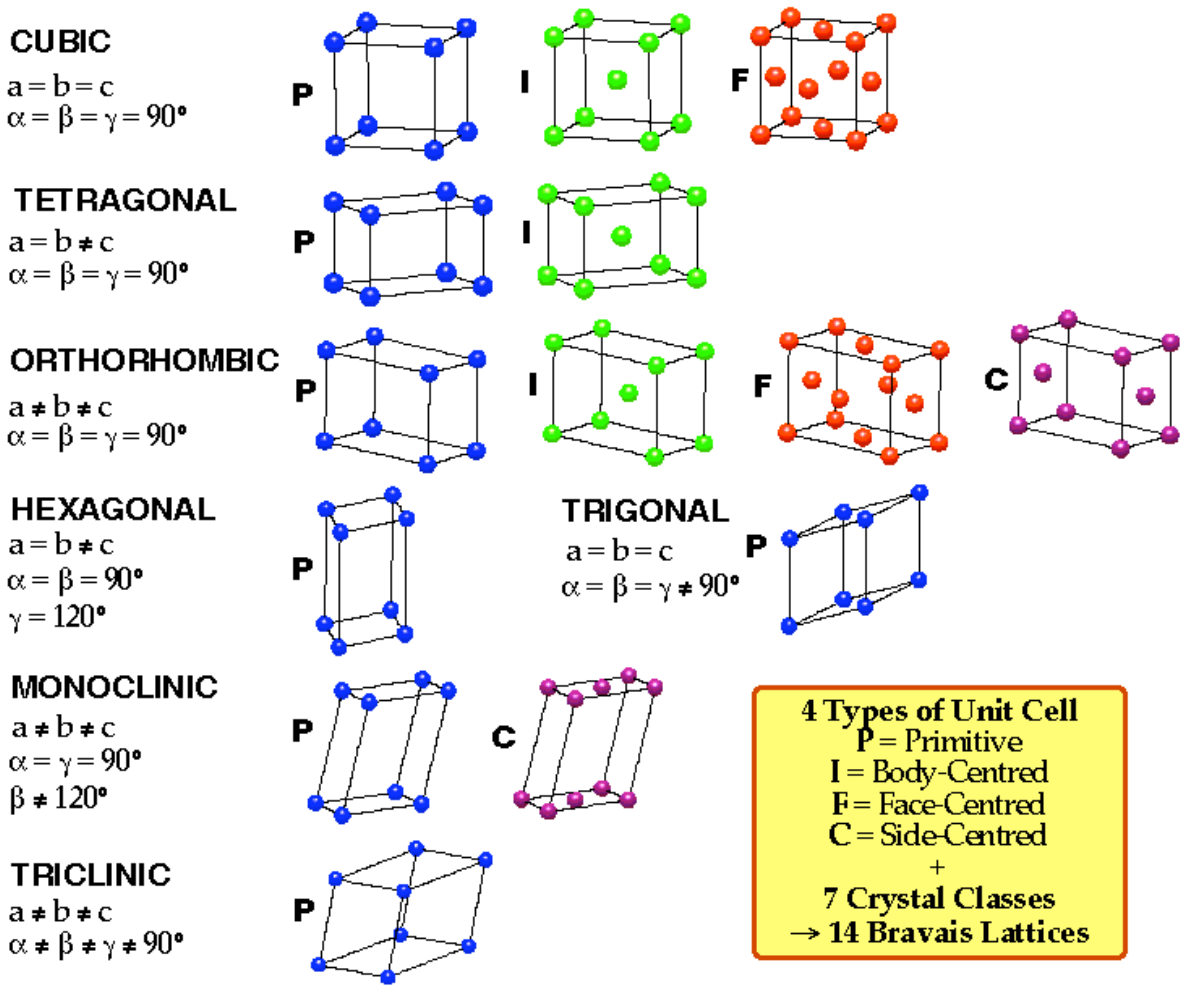


Figure A.3 Schematic representation of the 14 Bravais lattices.
The 14 Bravais lattices are composed of seven crystal systems and four lattice types (from http://www.ncl.ox.ac.uk).

A crystal form is an assemblage of equivalent faces generated by the symmetry of the crystal. The atoms in the crystal will also be contained in three-dimensional lattice planes, which are planes intersecting at least three noncollinear lattice points. A family of lattice planes is a set of parallel, equally spaced planes, which together contain all the points of a three-dimensional Bravais lattice (Hoy, 2002). Miller indices quantify the intercepts of the lattice planes in the main crystallographic axes in terms of unit cell dimensions. These indices are determined by the reciprocals of the intercepts along the axes x, y and z.

For example, the space group $P2_12_12_1$ contains four equivalent positions. The co-ordinates of the equivalent positions are: x ,y, z; 1/2 – x, -y, 1/2 +z; 1/2 + x, 1/2 – y, -z; -x, 1/2 + y, 1/2 –z. The equivalent reflections and the relationship between the phase of the prime reflection and the phase of each of the equivalent reflections can thus be derived:

For m = 1:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$(h\,k\,l)\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = (h\,k\,l)$$

$$S_h = 2\pi(h\,k\,l)\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = 0$$

Therefore, $\phi\,h\,k\,l = \phi\,h\,k\,l$

Following the same formula for the other equivalent positions, we see that for m = 2, ϕ hkl = ϕ -h –k -l, + π(-h + k); for m = 3, ϕ hkl = ϕ h –k –l + π(h + -k); and for m = 4, ϕ hkl = ϕ -h k –l + π(-hkl). Friedel's law states that the intensities of the h k l and –h –k –l reflections are equal. Therefore, if Friedel's law holds true, then only one octant ($|F_{hkl}|$) needs to be collected to achieve a unique data set for this space group because all other reflections are equivalent.

Another important concept in X-ray diffraction is that of the reciprocal lattice. The reciprocal space lattice is a set of imaginary points portioned in such a way that the direction of a vector from one point to another coincides with the direction of a normal to the real space planes. The separation of those points is equal to the reciprocal of the real interplanar distance (Zamliny, 2002). The reciprocal lattice contains repeat periods, the smallest unit of translation along an axis in the unit cell, and d spacing, the perpendicular distance between parallel planes in direct or reciprocal space. The reciprocal lattice spacing, d*, is inversely proportional to a direct lattice repeat period, and a direct lattice spacing, d, is inversely proportional to a reciprocal lattice repeat period. The principal axes of the reciprocal lattice x*, y*, z* are the normals from the origin to the principal planes of the unit cell, (100), (010), and (001) respectively.

From the reciprocal lattice, we can visualize the important concept of the Sphere of Reflection, or the Ewald Sphere, named for German physicist and crystallographer Paul Peter Ewald. To find the Ewald sphere, it is easiest to start by imagining a crystal rotating in real space, which causes the rotation of the reciprocal lattice in reciprocal space. The Ewald sphere is a sphere whose radius is the inverse of the wavelength, and contains a surface that passes through the origin of reciprocal space and the centre of the reciprocal lattice (Janin, 1975). Bragg's law can only be satisfied when a reciprocal lattice point touches the surface of the Ewald sphere. As the reciprocal lattice is rotated, most points on the three-dimensional Ewald sphere can be located and their intensities can be recorded on the detector. Based on the Ewald sphere, it is clear that each reciprocal lattice vector can only give rise to a diffraction point if the reciprocal lattice is moved successively through different orientations by moving the crystal relative to the incident beam (Shen, 2003)