2013-07-19

Structure-based Targeting of Transcriptional Regulatory Complexes Implicated in Human Disease: A Dissertation

Brendan J. Hilbert
University of Massachusetts Medical School

STRUCTURE-BASED TARGETING OF TRANSCRIPTIONAL REGULATORY COMPLEXES IMPLICATED IN HUMAN DISEASE

A Dissertation Presented

By

Brendan Jay Hilbert

Submitted to the Faculty of the
University of Massachusetts Graduate School of Biomedical Sciences, Worcester
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

July 19, 2013

Biochemistry and Molecular Pharmacology

STRUCTURE-BASED TARGETING OF TRANSCRIPTIONAL REGULATORY COMPLEXES IMPLICATED IN HUMAN DISEASE

A Dissertation Presented

By

Brendan Jay Hilbert

The signatures of the Dissertation Defense Committee signify completion and approval as to the style and content of the Dissertation.

Celia A. Schiffer, Ph.D., Thesis Advisor

William Royer Jr., Ph. D., Thesis Advisor

David G. Lambright, Ph.D., Member of Committee

Brian A. Kelch, Ph.D., Member of Committee

Katherine A. Fitzgerald, Ph.D., Member of Committee

Steven R. Grossman, M.D., Ph.D., Member of Committee

The signature of the Chair of the Committee signifies that the written dissertation meets the requirements of the Dissertation Committee.

Daniel N. Bolon, Ph.D., Chair of Committee

The signature of the Dean of the Graduate School of Biomedical Sciences signifies that the student has met all graduation requirements of the school.

Anthony Carruthers, Ph.D.
Dean of the Graduate School of Biomedical Sciences

Biochemistry and Molecular Pharmacology Program
July 19, 2013

Dedicated to the memory of Dr. Kai Lin.

ACKNOWLEDGEMENTS

I must recognize my mentors and colleagues for all of their help and encouragement while pursuing my doctorate. I have been extremely fortunate to have two of the greatest mentors I could have hoped for. Dr. Celia Schiffer has cultivated an amazing lab environment in which I am fortunate enough to have worked. She always provided guidance and encouragement, especially during the times when my research, or confidence in my work, was lacking. My other mentor Dr. William Royer Jr., has provided exceptional support throughout my tenure at UMass. Whether accompanying me on trips to the synchrotron, or having an impromptu conversion in his office, I always felt reassured and inspired to tackle the difficult problems in front of me. Beyond my two mentors, every member of my lab has contributed to my successes in some form. Many I must recognize for their direct impact on the results presented in this dissertation. I am indebted to Dr. Shivender Shandilya for taking the time to teach me to use most of software I used for solving and analyzing crystal structures. Similarly, I must thank Dr. Ayşegül Özen and Kristina Prachanronarong for their help in analyzing my data. My conversations with Dr. Akbar Ali were essential for identifying new compounds that advanced my research. Ellen Nalivaika was always helpful not just with protocols, reagents, and equipment, but also in keeping our lab coordinated and functioning. I would also like to thank Suvana Lam for her advice and help, especially with teaching me protein purification as I began my graduate career. All of the Schiffer, Royer, and Lin lab current and former members have proved essential to my successes. Their talent and enthusiasm for science leave me humbled. Simply put, I could not have dreamed of a

better lab family.

Beyond the Schiffer and Royer lab bays, the Biochemistry and Molecular Pharmacology Department promotes an environment collaboration and solidarity. At one point or another I have discussed my research, my successes, and my frustrations with practically every member of the department. Positive attitudes and friendly faces always addressed my questions with enthusiasm and support. Every conversation was met with a smile, and for that I am truly thankful.

I cannot forget the role my family and friends have played, unconditionally loving and supporting me throughout my time at UMass. My parents always provided a place for me to get away for a weekend, allowing me to unwind so I could face my challenges refreshed. My siblings, Bryan, Josh, and Laura, while providing encouragement, always reminded me what was truly important in life. Every member of my family has given unwavering support to me for my entire academic career, helping me reach my goals. My friends have played a similar role, encouraging me in my work, while helping me find relief and relaxation outside the lab. And finally, I have to thank my girlfriend, Melissa Fulham for her love and reassurance. I’m extremely grateful for the role my friends and family have had in my time here at UMass.

ABSTRACT

Transcriptional regulatory complexes control gene expression patterns and permit cellular responses to stimuli. Deregulation of complex components upsets target gene expression and can lead to disease. This dissertation examines proteins involved in two distinct regulatory complexes: C-terminal binding protein (CtBP) 1 and 2, and Interferon Regulatory Factors (IRF) 3 and 5. Although critical in developmental processes and injury response, CtBP transcriptional repression of cell adhesion proteins, pro-apoptotic factors, and tumor suppressors has been linked to the pathogenesis of multiple forms of cancer. IRFs function in the immune system and have been implicated in autoimmune disorders.

Understanding IRF activation is critical to treating pathogens that target IRF function or for future autoimmune disease therapies. We attempted to determine crystal structures that would provide the details of IRF activation, allowing insight into mechanisms of pathogen immune evasion and autoimmune disorders. Although no new structures were solved, we have optimized expression of C-terminal IRF-3 / co-activator complexes, as well as full-length IRF3 and IRF5 constructs. Modifying the constructs coupled with new crystal screening will soon result in structures which detail IRF activation, advancing understanding of the roles of IRF family members in disease.

Through structural and biochemical characterization we sought to identify and develop inhibitors of CtBP transcriptional regulatory functions. High concentrations of CtBP substrate, 4-Methylthio 2-oxobutyric acid (MTOB), have been shown in different cancer models to interfere with CtBP transcriptional regulation. We began the process of

structure based drug design by solving crystal structures of both CtBP family members bound to MTOB. The resulting models identified critical ligand contacts and unique active site features, which were utilized in inhibitor design. Potential CtBP inhibitors were identified and co-crystallized with CtBP1. One such compound binds to CtBP more than 1000 times more tightly than does MTOB, as a result of our structure-based inclusion of a phenyl ring and a novel pattern of hydrogen bonding. This molecule provides a starting point for the development of compounds that will both bind more tightly and interfere with transcriptional signaling as we progress towards pharmacologically targeting CtBP as a therapy for specific cancers.

TABLE OF CONTENTS

Title Page i

Signature Page ii

Dedication iii

Acknowledgements iv

Abstract vi

Table of Contents viii

List of Tables xii

List of Figures xiii

Preface xvi

Chapter I: Introduction

Transcriptional Regulatory Complexes 2

The CtBP Family

CtBP Genes and Isoforms 4

CtBP Domain Arrangement 4

Localization 7

Oligomerization 8

Nuclear Functions 9

Cytoplasmic Functions 14

Dehydrogenase Activity 14

Roles in Cancer 17

Targeting CtBP to Treat Cancer 21

Scope of Thesis 22

Interferon Regulatory Factors

The IRF Family and Domain Arrangement 24

The Distinct Functions of IRF3 and IRF5 25

Innate Immunity: Activating Interferon Gene Transcription 29

IRF3 Activation 32

IRF5 Activation 33

IRFs and Disease 34

IRF Structures 36

Scope of Thesis 39

Chapter II: Characterizing Interferon Regulatory Factor Activation

Introduction 43

Materials and Methods 48

Results

Dimeric IRF3 / IAD complex 53

Full-length IRF3 60

vIRF3 and IRF5 65

Discussion 70

Chapter III: Crystal Structures of Human CtBP in Complex with Substrate MTOB Reveal Active Site Features Useful for Inhibitor Design

Introduction 75

Materials and Methods 81

Results

CtBP1 and CtBP2 MTOB co-crystal structures 86

MTOB placement 91

CtBP – MTOB interactions 100

The unique hydrophilic cavity of the CtBP active site 115

Discussion 123

Chapter IV: Structure-Guided Design of CtBP inhibitors

Introduction 129

Materials and Methods 131

Results

Exploiting CtBP – MTOB contacts to identify new inhibitors 133

Small Molecule CtBP1 Co-crystallization and Structure 140

CtBP1 HIPP co-crystal structure 146

Experimental analysis of MTOB, phenylpyruvate, and HIPP binding affinity 162

Discussion 170

Chapter V: Discussion

Implications from CtBP co-crystal structures 175

Future directions 182

Future inhibitor design 183

Concluding remarks 193

Appendix

The effects of MTOB binding on CtBP conformation 196

MTOB placement in the CtBP2 active site 211

MTOB van der Waal analysis with repulsive terms 216

Bibliography 220

LIST OF TABLES

Table 3.1 – X-ray data collection and refinement statistics 87

Table 3.2 – Hydrogen bond network in MTOB – CtBP structures 106

Table 4.1 – X-ray data collection and refinement statistics 147

LIST OF FIGURES

Figure 1.1 – CtBP1 and CtBP2 domain arrangement 6

Figure 1.2 – CtBP binding proteins use common binding motifs to form histone-modifying complexes 13

Figure 1.3 – The D-2-hydroxyacid dehydrogenase enzymatic reaction 16

Figure 1.4 – IRF domain arrangement 27

Figure 1.5 – The enhanceosome mediates transcriptional activation 31

Figure 1.6 – Activation and dimerization in the IRF family 38

Figure 2.1 – IRF3 5D / CBP crystallization 55

Figure 2.2 – IRF3 5D / CBP fiber diffraction 59

Figure 2.3 – Full-length IRF3 2D crystals 62

Figure 2.4 – Full-length IRF3 2D crystal diffraction 64

Figure 2.5 – vIRF3 and IRF5 expression 69

Figure 3.1 – Sequence alignment of CtBP1 and 2 with various D-isomer specific 2-hydroxyacid dehydrogenases 77

Figure 3.2 – Architecture of the CtBP ternary complexes 89

Figure 3.3 – The binding location of MTOB relative to other D2-HDH substrates 93

Figure 3.4 – MTOB placement and conformations in the CtBP1 active site 96

Figure 3.5 – Occupants of the CtBP1 active site 98

Figure 3.6 – Positioning MTOB in the CtBP2 active site 102

Figure 3.7 – Substrate hydrogen bond network 105

Figure 3.8 – The mobility of the CtBP1 Arg97 guanidinium group 110

Figure 3.9 – MTOB van der Waals interactions 114

Figure 3.10 – Comparing other CtBP water networks to the CtBP1 ternary complex 118

Figure 3.11 – A small CtBP residue side chain allows formation of the water network 122

Figure 4.1 – Compounds docked into the CtBP1 active site 135

Figure 4.2 – Docking software incorrectly docks MTOB 138

Figure 4.3 – Candidates chosen for structural and biochemical characterization 142

Figure 4.4 – The 2.5 Å D-mandelic acid / CtBP1 complex 145

Figure 4.5 – The conformation of the HIPP structure matches the binary and MTOB bound CtBP1 structures 150

Figure 4.6 – Positioning HIPP in the CtBP1 active site 152

Figure 4.7 – The HIPP omit map confirms the novel conformation 155

Figure 4.8 – HIPP interactions in the CtBP1 active site 158

Figure 4.9 – The HIPP – CtBP1 hydrogen bond network 160

Figure 4.10 – The effect of HIPP on the CtBP water network 164

Figure 4.11 – HIPP affinity for CtBP2 is higher than phenylpyruvate or MTOB 168

Figure 5.1 – The unique cavity in the CtBP active site will provide specificity for inhibitors 178

Figure 5.2 – MTOB and HIPP fill only a small portion of the active site 181

Figure 5.3 – Recent docking with new compounds 187

Figure 5.4 – Docked compounds that would displace the water network 189

Figure A.1 – Comparing the asymmetric unit of the binary CtBP2 structure to the ternary structure 198

Figure A.2 – The CtBP2 binary complex subunits are identical 201

Figure A.3 – MTOB does not induce conformational changes in CtBP2 dimers 203

Figure A.4 – The range of small differences between individual CtBP2 ternary structure monomers 208

Figure A.5 – Comparing MTOB bound CtBP1 and CtBP2 structures 210

Figure A.6 – MTOB $2F_o$-F_c density in the CtBP2 active site 213

Figure A.7 – MTOB omit map F_o-F_c density in the CtBP2 active site 215

Figure A.8 – Examining the repulsive van der Waals forces 219

PREFACE

The results presented in Chapter III have been formatted as a manuscript and will be submitted to ACS Chemical Biology. The results presented in Chapter IV will be submitted for publication upon completion of key experiments.

CHAPTER I

Introduction

Transcriptional Regulatory Complexes

Regulation of Gene Expression

Transcriptional regulatory complexes often use chromatin modification to modulate gene expression. In response to stimuli, transcriptional activators or repressors and chromatin modifying enzymes form large multi-subunit complexes to activate or repress specific gene expression. The chromatin modifying enzymes lend fine control to the complex by tailoring specific effects with different types of covalent modifications. The specific regulatory complex members recruited to DNA controls modifications including histone phosphorylation, acetylation, methylation, and ubiquitination, permitting control of gene transcription (*1*). These modifications of DNA-binding histones alter chromatin architecture by changing histone-histone and DNA-histone interactions and by recruiting activator or repressor proteins to histone moieties. The diversity of chromatin modification function in concert with ubiquitous and specific transcriptional complex subunits to tightly control transcription of both individual and groups of genes, maintaining homeostasis. By tightly controlling gene expression patterns, transcriptional regulation allows cells to respond to millions of stimuli. However, improper gene regulation disturbs cell function, possibly contributing to disease states.

This thesis focuses on C-terminal Binding Proteins (CtBP) and Interferon Regulatory Factors (IRF): repressor and activator proteins involved in independent transcriptional complexes. CtBP proteins are co-transcriptional regulatory factors that associate with specific transcriptional factors to impact expression of a large number of

genes. Our interest in the CtBP family stems from the apparent roles these proteins have in the development and progression of cancer. Targeting CtBP appears to relieve transcriptional repression of proteins involved in apoptosis, tumor suppression, and cell adhesion. IRFs are specific transcriptional factors that function in the immune response, immune cell development, and tumor suppression. Critical for activating interferon (IFN) transcription, the IRFs are frequently targeted for inhibition by pathogens, whereas elevated IRF activity can lead to autoimmune disorders. Understanding IRF inhibition and activation are therefore critical to treat various infections and autoimmune disorders. This thesis therefore focused on elucidating mechanisms of activation and inhibition of both the CtBP and IRF families through structural and biochemical studies in order to develop future treatments for human disease.

The CtBP Family

CtBP Genes and Isoforms

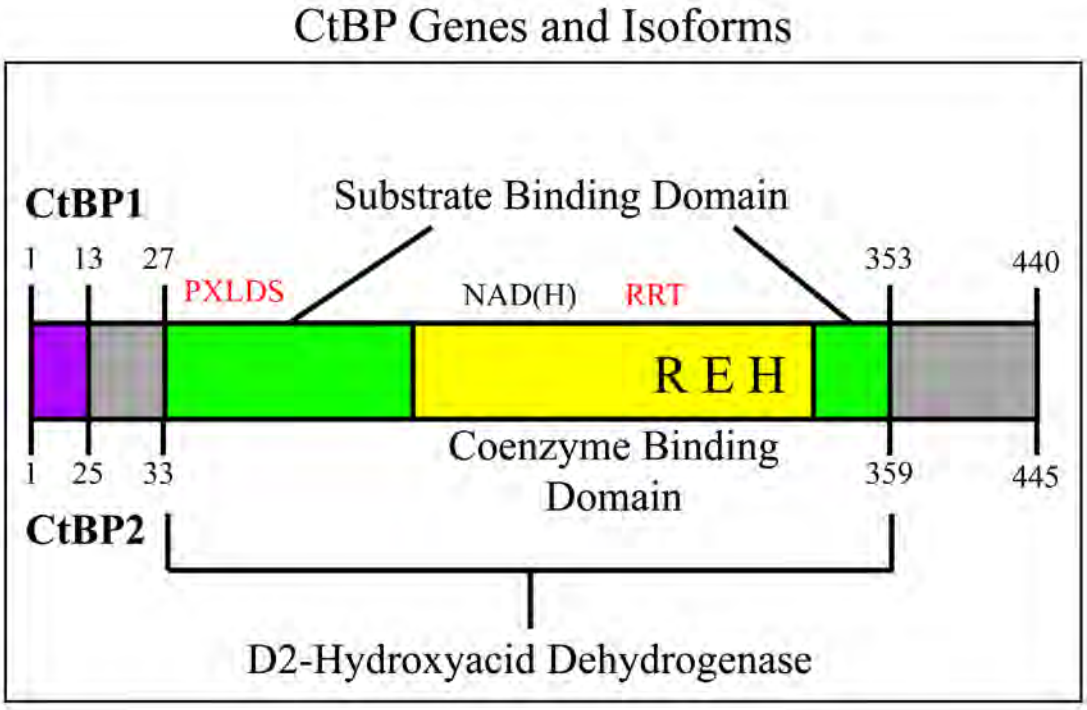
C-terminal Binding Protein 1 *(CtBP1)* (441 amino acids) was discovered for the ability to bind the C-terminus of the Adenovirus E1A protein and repress oncogenic transformation (*2, 3*). Unlike invertebrates, which possess only a single CtBP locus (*4*), the vertebrate genome contains a second CtBP gene, *CtBP2* (445 amino acids), which shares 83% amino acid similarity to *CtBP1* (*5*). Both vertebrate loci each code for multiple CtBP isoforms, which include long (CtBP-L) and short (CtBP-S) RNA splice variants. The CtBP1-S truncation lacks the first 13 N-terminal amino acids present in CtBP1-L (*6*), while CtBP2-S lacks the first 25 N-terminal amino acids present in CtBP2-L (*7, 8*). **(Figure 1.1)** Additionally, a tissue specific promoter in ribbon synapses expresses an additional related protein, RIBEYE, composed of an N-terminal domain fused to CtBP2 (20-445) (*9*).

CtBP Domain Arrangement

The CtBP proteins share high homology with the D-isomer specific 2-Hydroxy Acid Dehydrogenase (D2-HDH) family in both sequence and structure (*2, 10*). The core dehydrogenase domain possesses a common dinucleotide or Rossman fold typical of many NAD(H) dependent dehydrogenases (*11*). This dehydrogenase domain, consisting approximately of residues 30-350, can be subdivided into two domains: a substrate

Figure 1.1: CtBP1 and CtBP2 domain arrangement. The CtBP-L isoforms possess an the full CtBP sequence, whereas the CtBP-S isoforms possess an N-terminal truncation (purple). In CtBP2-S, the loss of the first 25 amino acids eliminates a NLS, shifting localization to the cytoplasm. The D2-HDH domain consists of a substrate binding domain (green) and a coenzyme binding domain (yellow) connected by small hinges. In addition to the NAD(H) binding site, the conenzyme binding domain contains the catalytic residues R, E, H. Both domains contain binding sites utilized in complex formation. The substrate binding domain has the PXDLS binding site, whereas the coenzyme binding domain contains the RRT binding site.

Figure 1.1


CtBP Genes and Isoforms
CtBP1
Substrate Binding Domain
1
13
27
PXLDS
NAD(H)
RRT
353
440
R E H
1
25
33
Coenzyme Binding
Domain
359
445
CtBP2
D2-Hydroxyacid Dehydrogenase

binding domain and coenzyme binding domain **(Figure 1.1)**. The two globular domains connect via a hinge region, forming a cleft at their convergence (*10*). The CtBP family dehydrogenase domain retains the catalytic residues present in other D2-HDHs (*2*). The C-terminal ~90 amino acids after the dehydrogenase domain are unstructured (*12*).

Localization

The CtBP isoforms localize to the cytoplasm or the nucleus based on a combination of protein sequence and post-translational modifications. CtBP2-L, possessing an NLS sequence in the N-terminal 25 amino acids, localizes to the nucleus (*7, 13*). The CtBP2-S isoform, however, lacks this NLS, localizing instead to the cytoplasm (*7*). The CtBP2 isoform, RIBEYE, also localizes to the cytoplasm, aggregating in large protein complexes at the active zone of the ribbon synapse (*9, 14*). Interestingly, CtBP2-L dimer interface mutants localize to cytoplasm, suggesting factors other than the N-terminal NLS contribute to localization (*15*).

Signals dictating CtBP1 localization are less defined and as such vary throughout the literature. Verger et al. observed CtBP1-L and CtBP1-S in both the nucleus and cytoplasm (*7*), consistent with the general observations of CtBP1 function in both the cytoplasm and nucleus. Conflicting reports from studies may be due to the effects of post-translational modification on CtBP1 localization. For example, phosphorylation of S158 in the dehydrogenase domain redistributes CtBP1 into the cytoplasm (*16*). Conversely, SUMOylation of the extreme C-terminus at K428 may maintain CtBP1 in the nucleus (*17*), although these results were not reproduced (*18*). Inhibiting CtBP1

dimerization appears to impair nuclear accumulation, suggesting oligomerization may play a role in localization (*18*). It is also possible that CtBP1 imports into the nucleus via binding partners, including CtBP2, which possess an NLS (*7*).

Oligomerization

Typical of the D2-HDH family, CtBP forms dimers. Oligomerization exists in all the structures solved for different D2-HDH members, including both bacterial and human enzymes. Structural data for bacterial members includes, but is not limited to, two distinct D-Lactate Dehydrogenases (D-LDH) (*19, 20*), Hydroxyisocaproate Dehydrogenase (D-HicDH) (*21*), Formate Dehydrogenase (FDH) (*22*), and D-Glycerate Dehydrogenase (DGDH) (*23*). Although fewer related human proteins exist, the two most highly related dehydrogenases have also been crystallized. Structures of Glyoxylate Reductase/Hydroxypyruvate Reductase (GRHPR) (*24*) and D-3 Phosphoglycerate Dehydrogenase (PHGDH) (*25*) were both solved by X-ray crystallography. Roughly 90% of the dimer interface exists between the well conserved coenzyme binding domains of each monomer (*21*). The variable substrate binding domain therefore minimally contributes to the interface, suggesting dimerization is evolutionarily conserved in this family.

By preferentially binding different redox states of NADH, CtBP acts as a redox sensor, modulating its activity in response to metabolic changes. CtBP possesses a 100 fold higher affinity for NADH than for NAD^+ (*26*), making CtBP sensitive to perturbations in the NADH/NAD^+ ratio. Binding of either NADH or NAD^+ induces CtBP

dimerization (*27*), a structural change necessary for full CtBP activity (*15*, *28-30*). Stimuli that increase cellular NADH, such as hypoxia and high extracellular glucose, induce higher CtBP activity. These observations suggest the level of coenzyme binding, dependent on the NADH/NAD^+ ratio, modulates CtBP oligomerization and formation of competent transcriptional regulatory complexes.

Dimerization provides a stable platform for other transcriptional complex members to bind. Each CtBP monomer contains a PXDLS domain that possesses a common binding site utilized by critical CtBP binding partners; dimerization increases the available PXDLS sites (*15*). Evidence suggests having only a single monomer available results in competition for binding between different complex members, disrupting transcriptional regulation (*15*). Only rarely will monomeric CtBP be required in specific regulatory complexes, indicating the context of the transcriptional regulation may determine CtBP oligomeric state (*31*). Overall, oligomerization appears critical for CtBP to properly recruit the necessary transcriptional regulatory machinery.

Nuclear Functions

Studies in the last 15 years have focused on CtBP transcriptional roles in development as well as in cancer. The role in regards to development will be discussed here, while CtBP's function in cancer is discussed below. Hildebrand et al. demonstrated that CtBP activity is critical in murine development. Mutation in *Ctbp2* results in embryonic lethality; E10.5 embryos exhibit defects in heart and neural development (*32*). Although viable, *Ctbp1*-null mice are smaller in size than littermates containing one or

two functional copies of *Ctbp1*. The *Ctbp1*-null mice exhibit a 23% mortality rate within 20 days of birth (*32*). Reducing the dosage of *Ctbp1* in a *Ctbp2*-null background, and vice versa, severely impedes development in the embryo, demonstrating the overlapping roles of CtBP proteins (*32*). Additional roles in development were revealed through studies of CtBP in *D. melanogaster*. The single *CtBP* gene possesses regulatory roles in wing (*33*), eye (*34, 35*), sensory organ (*36*), as well as more generalized roles in development (*37, 38*). The role CtBP in each of these developmental processes is related to CtBP's function in transcriptional regulation.

CtBP achieves transcriptional regulation through formation of large DNA-bound chromatin remodeling complexes **(Figure 1.2)**. DNA binding proteins recruit CtBP by binding the PXDLS domain, discovered from the interaction with the adenovirus E1A protein (*2*). The PXDLS domain exists in the unique Substrate Binding Domain of the N-terminus of CtBP (*39*). Many DNA-binding proteins, including the most commonly cited DNA-binding CtBP partners, zinc finger proteins ZEB 1/2, act as the bridge between promoters and the CtBP PXDLS domain (*4, 15, 40-43*). An additional RRT binding motif (RRTGXPPXL) discovered in the CtBP coenzyme binding domain also binds cofactors involved in repression (*44*). DNA-binding transcription factor, Znf217, binds CtBP utilizing not only the PXDLS motif, but also the RRT motif (*18, 45*). Interestingly, structural data demonstrate that the RRT binding site on one CtBP dimeric subunit rests on the same face as the PXDLS site of the second dimeric subunit, suggesting that proteins possessing both an RRT and PXDLS binding motifs bind both dimeric subunits

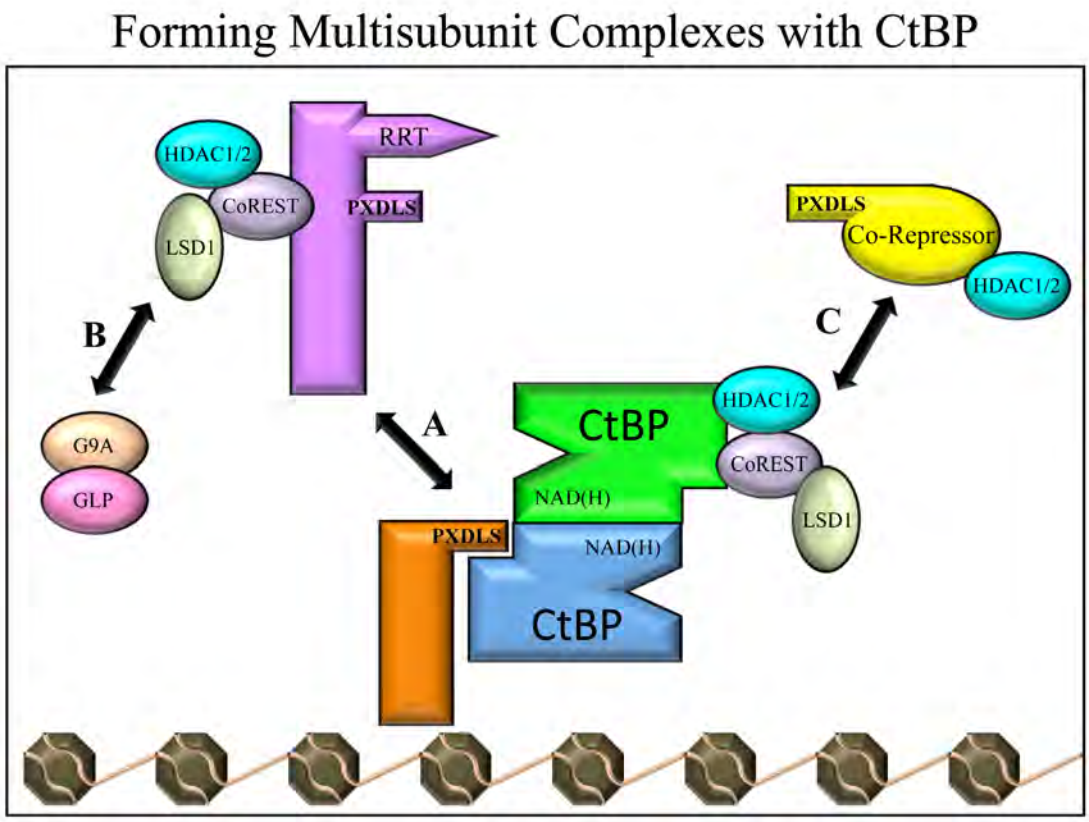
simultaneously (*44*), reinforcing the idea that oligomerization is critical for CtBP function.

Analyses of the CtBP transcriptional complexes revealed multiple types of chromatin modifying proteins. Class I histone deacetylases (HDAC 1/2) associate with the CtBP complex (*18, 46*) and yeast two-hybrid screens suggest Class II HDAC proteins may also interact with CtBP (*47*). Histone methyltransferases (HMT), G9A and GLP (EuHMTase I), and a lysine demethylase (LSD-1) also bind the CtBP complex (*18, 46*). The variation among chromatin modifying proteins is indicative of CtBP's role at different promoter regions, each requiring a specific set of transcriptional regulatory machinery.

Histone-modifying proteins bind to CtBP either directly or through adapter proteins. Members of the HDAC1/2 family have the ability to bind directly to CtBP, as does a corepressor, CoREST; these proteins utilize non-PXDLS interactions (*18, 46*). CoREST not only recruits HDAC1/2, but also demethylase LSD-1 to CtBP (*48, 49*). Unlike CoREST, corepressor LCoR exploits the CtBP PXDLS motif to tether HDAC1/2 to CtBP (*50*). DNA binding proteins that partner with CtBP can also recruit chromatin-modifying proteins, demonstrating yet another means of forming a chromatin-modifying complex that contains CtBP. In particular Znf217 binds CoREST, which in turn recruits HDAC1/2 and LSD1 (*45*). Another zinc finger protein, Wiz, binds directly to the G9A/GLP histone methyltransferases and CtBP when forming a regulatory complex (*51*). These observations again suggest that histone-modifying proteins are recruited through different DNA-binding proteins and adapters between various CtBP

Figure 1.2: CtBP binding proteins use common binding motifs to form histone-modifying complexes. (A) DNA-binding proteins (orange), such as Zeb1, TGIF (*52*), and Evi-1 (*53*), bind a CtBP dimeric subunit through a PXDLS motif. Other DNA-binding proteins (purple), such as ZNF217, utilize both the PXDLS and RRT motifs to interact with CtBP. (B) DNA-binding proteins can directly recruit histone deacetylation/demethylation complexes, such as the CoREST/HDAC/LSD1 complex, or methyltransferase complexes, such as G9A/GLP. (C) Some corepressors that recruit histone-modifying enzymes (yellow), such as LCoR1, utilize a PXDLS motif to bind CtBP. Alternatively CtBP can directly bind histone-modifying enzymes that utilize non-PXDLS interaction, such as HDAC1/2 and Co-REST/LSD1. CtBP binds a specific combination of DNA binding proteins, adapters, and histone modifying enzymes to target individual genes.

Figure 1.2


Forming Multisubunit Complexes with CtBP
HDAC1/2
CoREST
LSD1
RRT
PXDLS
B
G9A
GLP
A
CtBP
NAD(H)
PXDLS
NAD(H)
CtBP
HDAC1/2
CoREST
LSD1
C
PXDLS
Co-Repressor
HDAC1/2

regulatory complexes depending on the context, forming distinct chromatin-remodeling complexes to tailor transcriptional regulation (*15*) **(Figure 1.2)**.

Cytoplasmic Functions

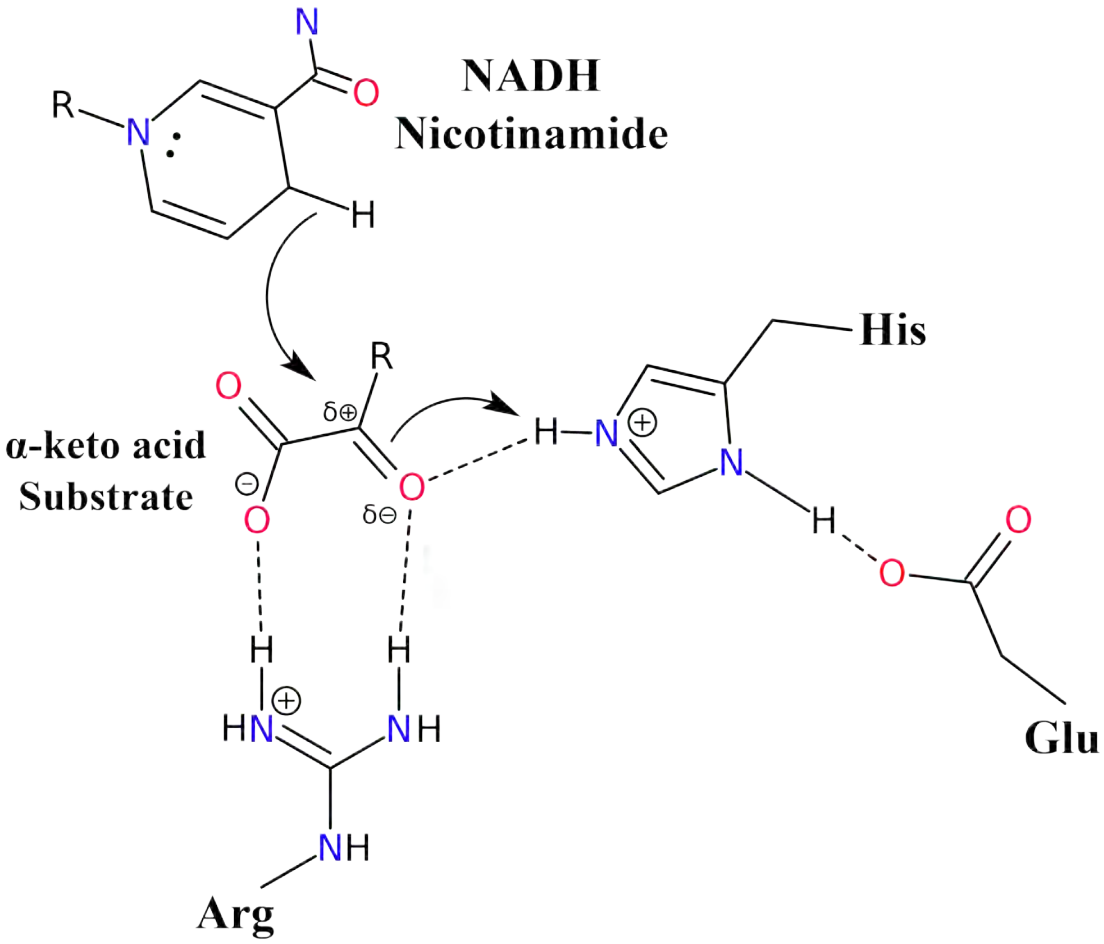
In addition to roles in the nucleus, CtBP also functions in the cytoplasm. CtBP1-S induces tubule constriction and fissioning at the Golgi (*54*), dependent upon CtBP binding long-chain acyl-coenzyme A (acyl-CoA) molecules. CtBP1-S also appears to possess acyl-transferase activity targeting lysophosphatidic acid (LPA), thereby altering the membrane properties to induce tubule constriction and fission (*55, 56*). However, the enzymatic role of CtBP1 remains controversial as other studies demonstrated CtBP1 either lacks acyl-transferase activity (*57*) or does not require acyl-transferase activity to induce fission (*58*). Complicating the issue, recent reports reveal that RIBEYE, or at least the CtBP2 domain of RIBEYE, also displays acyl-transferase activity (*59*) and functions in vesicle formation at ribbon synapses (*9, 14, 59*). Regardless of the exact mechanism, CtBP1 and CtBP2 have important roles in vesicle formation (*58*), Golgi partitioning during mitosis (*60*), and synapse signaling (*14, 59*).

CtBP Dehydrogenase Activity

Discovery of the sequence similarity of CtBP to other D2-HDHs resulted in studies attempting to determine if CtBP enzymatic activity existed and, if so, the identity of a potential substrate. D2-HDH catalysis involves hydride transfer to substrate during NADH oxidation, converting an α-keto acid 2-C carbonyl to the D-isomer hydroxyl

Figure 1.3: The D-2-hydroxyacid dehydrogenase enzymatic reaction. NADH and the α-keto acid substrate bind the active site. The substrate carbonyl and carboxylate oxygen are stabilized by a conserved arginine. An acidic residue, typically Glutamate, raises the pK_a of an adjacent histidine, stabilizing protonation. The histidine proton hydrogen bonds with and polarizes the substrate carbonyl. A hydride then transfers from the nicotinamide C4 carbon to the partial positive C2 substrate carbon. The carbonyl abstracts the histidine proton, and the D-isomer hydroxyacid product is released.

Figure 1.3



group (*21, 61, 62*) **(Figure 1.3)**. Evidence of CtBP enzymatic activity was first discovered for the ability to convert pyruvate to lactic acid through oxidation of NADH (*10, 27*). A more thorough search of potential substrates yielded the discovery that 4-methylthio-2-oxobutyric acid (MTOB) was an 80-fold better substrate for CtBP1 than pyruvate (*63*). Interestingly, substituting the MTOB sulfur for a methylene group, as in 2-oxohexanoic acid, reduced the enzymatic activity eight-fold relative to MTOB, suggesting the sulfur confers specificity (*63*). MTOB was identified as the final intermediate in the methionine salvage pathway, in which transamination of MTOB results in formation of methionine (*64*). Dietary supplements for livestock even employ MTOB to increase methionine levels (*65*). CtBP mediated transcriptional regulation does not appear to require enzymatic activity, suggesting that the NADH binding function has been conserved for the purpose of regulating CtBP transcriptional functions, not for catalytic turnover (*66*).

Despite the higher catalytic turnover relative to pyruvate, MTOB remains a fairly poor substrate when compared to the catalytic efficiencies of related D2-HDH family members. For example, D2-HDH family member rat PHGDH catalyzes substrate 400 times more efficiently, while *E. coli* DGDH processes substrate more than 2600 times more efficiently than MTOB is catalyzed by CtBP (*63, 67, 68*).

Roles in Cancer

Evidence continues to grow linking CtBP transcriptional regulatory function to development of cancer, especially in the context of the epithelial-to-mesenchymal

transition (EMT). EMT is a transdifferentiation process necessary in a multitude of developmental processes, in which non-motile epithelial cells lose polarity and cell adhesion, acquiring instead migratory and invasive characteristics (*69*). Although necessary in development and wound healing, EMT increases the metastatic potential of malignant tumors as well as resistance to apoptosis (*70*).

CtBP represses expression of certain cell adhesion molecules, cell-cell junction proteins, cytoskeletal elements, pro-apoptotic factors (*66*), and tumor suppressors (*71*), linking CtBP function to cell reprogramming associated with EMT. CtBP represses transcription of *E-cadherin*, a cell adhesion molecule whose loss is a key indicator of EMT (*66*). The apparent influence of CtBP across such varied targets is not surprising given recent data demonstrating CtBP binding throughout the mammalian genome at greater than 1,800 promoters (*72*). Given the pro-oncogenic nature of CtBP targets, research focuses on CtBP function in multiple types of cancer. Although limited evidence suggests roles for CtBP in progression of melanoma (*73*), pituitary tumors (*74*), and prostate cancer (*75*), larger studies implicate CtBP function in progression and severity of colon (*76-81*) and breast cancer (*8, 72, 82-84*).

Loss of the ability of tumor suppressors to target CtBP leads to formation and progression of colon cancer. *Adenomatous polyposis coli* (APC) possesses a tumor suppressor function that targets CtBP1 for degradation (*76*). Mutation of APC is a well established mechanism for development of colorectal tumors (*85*). The loss of APC tumor suppression results in elevated CtBP expression, an early event in cancer initiation (*76, 77*). A second tumor suppressor, $p14^{arf}$ (ARF), also targets CtBP for proteosomal

degradation. Direct ARF-CtBP interaction results in p53-dependent apoptosis (*78*). Further investigation revealed CtBP represses Bik, a pro-apoptotic factor, and that this was relieved through CtBP interaction with ARF (*79*). Bik serves to bind anti-apoptotic factors, antagonizing their function and sensitizing cells to apoptotic signals (*86*). CtBP therefore inhibits apoptosis directly through targeting Bik (*81*). Therefore, the loss of both APC and ARF tumor suppression increases CtBP activity and contributes to tumorigenesis.

Although CtBP is not known to directly target ARF, there is an inverse correlation between the expression levels of CtBP and ARF. Colon cancer tissue samples exhibited elevated levels of CtBP with simultaneous diminished expression of ARF. Adjacent healthy tissue exhibited inverse expression, displaying low expression of CtBP and normal expression of ARF (*81*). Additional evidence for the CtBP-ARF relationship in colon cancer arose by studying cell migration. Several studies demonstrate that the hypoxia-induced elevation in NADH levels increased cell migration in different colon cancer cell lines (*28, 80*), again suggesting NADH-mediated dimerization plays a role in CtBP function in cancer. Whereas loss of ARF resulted in higher rates of cell migration in a lung carcinoma cell line, introduction of exogenous ARF into an ARF-null colon adenocarcinoma cell line reduced cell migration under both normoxic and hypoxic conditions (*80*). Presumably, ARF antagonizes CtBP function, allowing Bik-mediated apoptosis, even under elevated NADH conditions. In colon adenocarcinomas, decreased ARF expression and increased CtBP levels result in cell survival and increased migration.

Deregulation of CtBP through loss of ARF, APC, or both proteins appears to significantly contribute to the development, survival, and progression of colon cancer.

The role of CtBP activity in breast cancer was examined across a representative group of various cell lines and tumor samples. Results indicate elevated CtBP activity allows cell survival and metastasis, whereas CtBP depletion arrests cell growth. CtBP1 displays elevated expression levels in invasive ductal carcinoma tissue, relative to normal breast tissue (*82*), and can even serve as a n indicator in median survival estimates in patients with basal-like triple negative breast cancer (*72*). Much like colon cancer, breast cancer tissue also exhibits the loss of *E-cadherin* expression. In addition, there is an inverse correlation between CtBP and the early-onset breast cancer gene, *BRCA1*, expression levels. Mutations in *BRCA1*, a DNA-repair protein, are typically associated with increased risk of developing breast cancer. *CtBP1* knockdown restores normal *E-cadherin* and *BRCA1* levels; further linking CtBP function to tumorigenesis (*82*). Additional evidence demonstrates *BRCA1* levels increase with a reduction of the NADH/NAD^+ ratio, causing removal of HDAC1 from the promoter (*83*). These results suggest a decrease in NADH levels negatively impacts CtBP ability to dimerize and form a competent transcriptional repressor complex, causing the loss in CtBP repressor activity. This observation potentially links elevated CtBP transcriptional activity with the Warburg effect (*87*), in which elevated glycolysis raises NADH levels in malignancies (*88*). Knockdown of *CtBP* appears to sensitize breast cancer cells to chemotherapeutics, such as cisplatin, as well as reduces cell proliferation (*8*). Overall, CtBP influences breast cancer through transcriptional repression of typical EMT related targets such as *E-*

cadherin, as well as cancer specific targets, such as *BRCA1*. The effects of CtBP repression may also explain why some breast cancer patients show a loss of *BRCA1* expression despite the absence of mutation in the *BRCA1* (*83*).

Targeting CtBP to Treat Cancer

Attempts to pharmacologically target CtBP in cancer models reveal that millimolar (mM) levels of the substrate MTOB shows anti-tumor effects in both colon and breast cancer models. Studies previously identified MTOB as a growth inhibitor and apoptosis-inducing agent for multiple transformed cell lines, although the targets of MTOB remained unclear (*89*). Later testing revealed MTOB displayed p53-independent cytotoxicity to colon adenocarcinoma HCT116 cells. However, the concentration required for high cytotoxicity were 4 mM or greater (*81*). MTOB treatment displaces CtBP from the Bik promoter, relieving *Bik* repression and leading to increased *Bik* expression relative to untreated cells. Either exogenous CtBP2 expression or silencing of *Bik* negated MTOB's effects, linking MTOB-induced apoptosis to CtBP repression of *Bik* (*81*). MTOB treatment of nude mice injected with p53-null HCT116 cells increased median survival slightly and dramatically decreased tumor burden relative to control mice (*81*). Finally, MTOB seems to target only malignant cells, as no cytotoxicity was observed in healthy tissue from mice injected with MTOB (*81*). Therefore, CtBP activity specific to colon cancer was safely inhibited by large quantities of MTOB, suggesting that CtBP is a viable target for cancer therapy.

MTOB's anti-cancer effect mediated through CtBP was corroborated in a second study examining the response in different breast cancer cell lines. Treatment of MCF-7 and MDA-MD231 cells with 10 mM MTOB evicted CtBP from a substantial number of promoters and attenuated repression of genes associated with genome stability, EMT, and stem cell pathways (*72*). Furthermore, markers for epithelial and mesenchymal phenotypes indicated a shift away from a mesenchymal phenotype towards an epithelial phenotype upon MTOB treatment, suggesting partial reversal of EMT (*72*). These data further support the idea of pharmacologically targeting CtBP in a number of different cancers. X-ray crystal structures of CtBP1 and CtBP2, solved prior to discovering the role of MTOB as a substrate and anti-cancer agent, suggested CtBP would be a good candidate for structure based drug design in the endeavor to develop new cancer therapeutics.

Scope of Thesis: CtBP and Structure Based Drug Design

The first portion of my research addresses the need for substrate bound CtBP structures to initiate development of rationally designed CtBP inhibitors. CtBP appeared to be a good candidate for structure based drug design as binary structures of both CtBP1 and CtBP2 in complex with NAD^+ had previously been solved (*10, 90*). In this endeavor I succeeded in co-crystallizing MTOB bound in the active site of both CtBP1 and CtBP2. The position and orientation of MTOB mimics other ligand bound D2-HDH family members. These structures provided data on the similarity of the architecture between

MTOB bound and previously published MTOB-free CtBP1 and CtBP2, as well as revealing unique aspects of MTOB binding.

Utilizing knowledge gained from the MTOB bound structures, the second part of this thesis focuses on identifying small molecules that tightly bind the CtBP active site. To this end, I identified multiple small molecules that inhibited CtBP enzymatic function. Co-crystallization was attempted with multiple compounds, the successes of which provided new details on ligand binding. At the suggestion of our collaborator, we explored the binding of 2-hydroxyimino-3-phenylpropionic acid, (HIPP) with CtBP1. My co-crystal structure of CtBP1 revealed that HIPP binding exploits the intended contacts with CtBP active site specificity determinants, while binding in an unexpected conformation. In addition, I found that HIPP shows a dramatic increase in affinity for CtBP of greater than 3 orders of magnitude compared with that of MTOB. With this novel co-crystal structure, new molecules can be synthesized to further enhance affinity and specificity. As new molecules are designed, insight into the mechanism behind small molecule inhibition of CtBP transcriptional regulatory activity can be deduced.

INTERFERON REGULATORY FACTORS

The IRF Family and Domain Arrangement

The Interferon Regulatory Factor (IRF) family consists of nine members (IRF1-IRF9) in humans and mice (*91, 92*). This family of proteins possesses diverse functions relating to immunity, immune cell development, and tumor suppression. The scope of research in this thesis focuses on two members of the IRF family: IRF3 and IRF5 (variant 4). Tamura *et al.* have reviewed the expression and function of the other IRF family members (*92*).

Each IRF member possesses a conserved N-terminal DNA binding domain (DBD; ~120 residues) that contains a tryptophan pentad and binds to a 5'-AAXXGAAA DNA sequence (*91, 93, 94*). Many interferon (IFN) stimulated genes, including IFN-α/β themselves, possess this sequence upstream of the transcriptional start site, often with multiple tandem repeats of the IRF binding consensus sequence above (*95*). IRF3 and IRF5 possess a DBD typical of the IRF family **(Figure 1.4)**.

The C-terminal domain of the IRF family members possesses much more diversity than the DBD. The C-terminal domain of most IRF family members, including IRF3 and IRF5, contains an IRF association domain (IAD) responsible for homo- and heteromeric associations (*96, 97*). Additionally, the IRF3 and IRF5 extreme C-terminus includes an autoinhibitory loop that prevents IAD associated activation (*98, 99*). The autoinhibitory region of IRF3 and IRF5 contains clusters of serine and threonine residues, which serve as phosphorylation targets. Phosphorylation activates the IRFs by triggering

oligomerization, localization to the nucleus, and transcriptional activation of target genes **(Figure 1.4)** (see *IRF Activation* below for details).

A linker region of ~75 amino acids separates the DBD from the IAD. The exact function of the linker region remains unknown, however, evidence is mounting that the linker has more functions than to simply tether the DBD and IAD to one another. For example, IRF3 and IRF5 each possess a nuclear export sequence (NES) in the linker region, important for proper localization (*100, 101*). Reports suggest that the IRF3 linker may not be unfolded as previously thought (*102*). Gaining insight into the function of the linker region is paramount for understanding IRF function.

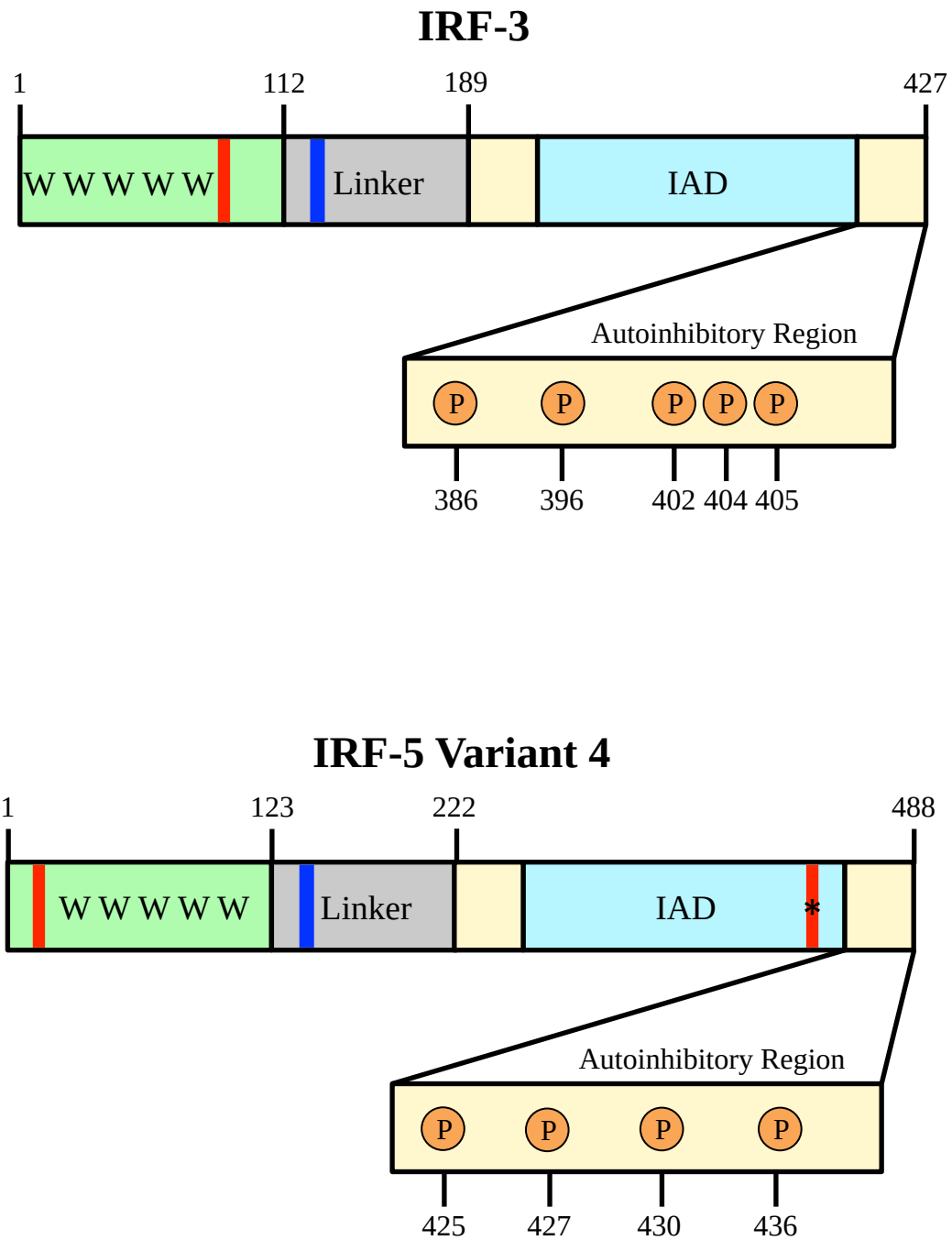
The Distinct Functions of IRF3 and IRF5

IRF3 and IRF5 have related but distinct roles in innate immunity. Concerning cellular expression, IRF3 is expressed constitutively in all cell types (*92, 103*), whereas IRF5, expressed in immune cells only, exhibits constitutive or inducible expression. Specifically, different variants of IRF5 were detected in B- and T-cells, dendritic cells, monocytes, natural killer cells, and peripheral blood mononuclear cells (*104*). The different expression patterns and higher variation of isoforms in IRF5 tailor the unique functions of each protein within and across different cell types. The IRF5 variant 4 sequence will be used for residue numbering throughout this dissertation.

Host innate immunity is critical for the initial response to pathogen invasion. Germline encoded receptors exist as soluble pattern recognition receptors (PRRs) or membrane bound Toll-like receptors (TLRs) and recognize pathogen associated

Figure 1.4: IRF domain arrangement. The IRF proteins consist of an N-terminal DNA Binding Domain (DBD; green), a linker region (grey), an IRF association domain (IAD; blue), and autoinhibitory region (yellow). The DBD contains a tryptophan pentad conserved throughout the IRF family, as well as NLS signals in both IRF3 and IRF5 (red bars). The linker region of both proteins also possesses an NES (blue bar). The C-terminal IAD is responsible for homo- and heteromeric interactions after IRF activation, including CBP binding. The IRF5 IAD contains an NLS / ubiquitination site (red bar with star) necessary for activation and nuclear localization. The autoinhibitory region prevents co-activator association with the IAD by shielding the binding site. This region also contains putative phosphorylation sites (orange spheres) necessary for IRF activation and dimerization.

Figure 1.4


IRF-3
1
112
189
427
W W W W W
Linker
IAD
Autoinhibitory Region
P
P
P
P
P
386
396
402
404
405
IRF-5 Variant 4
1
123
222
488
W W W W W
Linker
IAD
*
Autoinhibitory Region
P
P
P
P
425
427
430
436

molecular patterns (PAMPs). TLRs are expressed mainly in antigen presenting cells (APCs), while soluble PRRs are expressed more ubiquitously across different cell types. When activated, both receptor classes induce expression of Type I IFNs and pro-inflammatory cytokines (*92*). Human Type I IFNs consist of factors from 13 IFN-α genes and a single gene for IFN-β, -ε, -κ, and -ω each, which in turn generate multifaceted intra- and inter-cellular responses influencing hundreds of gene targets involved in the antiviral response, anti-proliferation, and immune signaling (*105*).

Although both IRF3 and IRF5 regulate Type I IFN expression (*92*, *106*, *107*), each protein becomes activated through a distinct set of pattern recognition receptors (PRRs). For example, both soluble PRRs (RIG1/MDA5) and membrane bound PRRs (TLR3, TLR4) each begin a signal cascade that ultimately leads to IRF3 activation (*108-110*). IRF5, however, becomes activated in response to signals from membrane bound TLR7, TLR8, and TLR9 (*111*, *112*). Other studies suggest the involvement of other TLRs in IRF5 activation, but results can be contradictory due to cell-specific expression of different IRF5 variants and activating pathways (*113*). IRF3 appears to have a larger role in the antiviral response, activating IFN expression across almost all cell types in response to viral infection. The IRF5 antiviral response is more limited as IRF5 is not ubiquitously expressed.

In addition to antiviral activity, IRF5 functions as a tumor suppressor. Studies indicate defects in both DNA damage and virus-induced apoptosis in IRF5-null models (*114*). Additionally IRF5 overexpression sensitizes tumor cells to apoptosis, an event

linked with IRF5 phosphorylation and activation (*115*). Thus IRF5 appears to possess dual roles in innate immunity and tumor suppression.

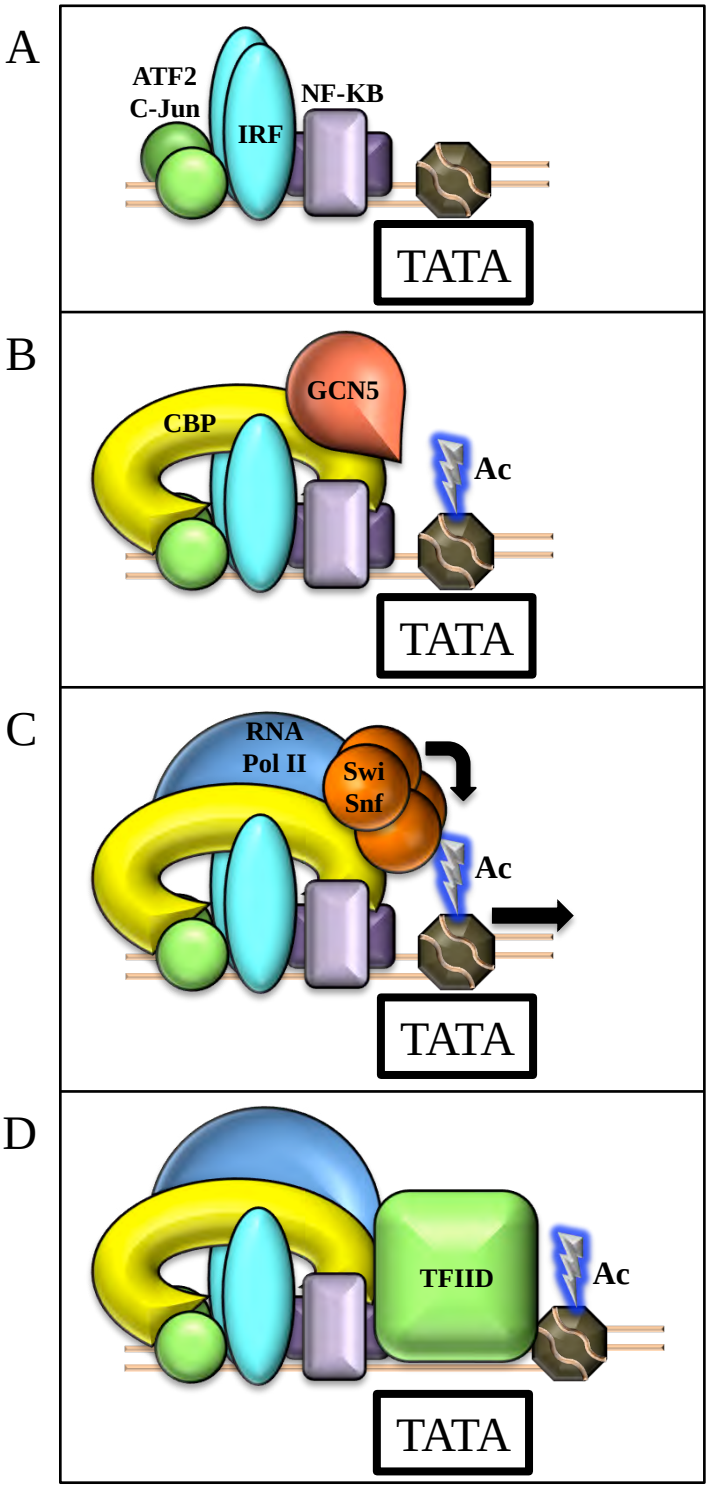
Innate Immunity: Activating Interferon Gene Transcription

The IFN-β enhancer has been well studied and serves as a model for transcriptional activation in innate immunity. IFN-β gene regulation involves formation of an activating transcriptional complex termed the enhanceosome. The enhanceosome binds to the Interferon Stimulated Response Element (ISRE) located about -50 bp from the transcriptional start site (*94*). The ISRE is broken up into four Positive Regulatory Domains (PRDs I-IV), with each PRD nucleotide sequence binding cofactors as they assemble on the ISRE. "Architectural" protein HMG-I (*116*) recruits the NF-κB p50/p65 heterodimer (PRD II) (*117*), ATF2/C-Jun heterodimer (PRD IV) (*118*), and IRF3 (*119*) or IRF-7 (*120*) (PRD III/I) to the ISRE, forming the enhanceosome.

The enhanceosome sequentially recruits histone acetyltransferases (HATs) and chromatin remodeling complexes to facilitate IFN-β gene transcription **(Figure 1.5)**. Although the ISRE remains nucleosome free, a nucleosome binds directly downstream of the TATA box, shielding the transcriptional start site of the IFN-β gene (*121*). In order to activate transcription, the enhanceosome assembly recruits two HATs: GCN5 and CBP, both of which are able to acetylate the nucleosome (*121*). Each of the ISRE binding factors contacts CREB Binding Protein (CBP) recruiting it to the enhancer region (*122-125*). After acetylation, GCN5 releases from the complex, while CBP remains to recruit the Pol II holoenzyme and Swi/Snf chromatin remodeling complex (*121*). Chromatin

Figure 1.5: The enhanceosome mediates transcriptional activation.
(A) A nucleosome blocks the transcriptional start site of a target gene. The elements comprising the enhanceosome, including ATF/C-jun heterodimer, IRFs, and NF-κB, assemble on the ISRE. (B) The enhanceosome recruits CBP and GCN5, which acetylate the histone. (C) CBP then recruits the Pol II holoenzyme and Swi/Snf complex. The Swi/Snf complex binds the acetylated nucleosome and repositions it, exposing the transcriptional start site. (D) Transcription factors are recruited and transcription begins.

Figure 1.5


A
ATF2
C-Jun
IRF
NF-KB
TATA
B
GCN5
CBP
Ac
TATA
C
RNA
Pol II
Swi
Snf
Ac
TATA
D
TFIID
Ac
TATA

remodeling shifts the nucleosome position, allowing transcription of IFN-β. Similar events occur at other IFN stimulated genes that also possess ISREs.

IRF3 Activation

Signals from a distinct set of PRRs result in IRF3 activation. Although these cascades may begin at different receptors, the separate pathways eventually converge, utilizing Tank Binding Kinase I (TBK-1) or IκB kinase ε (IKKε) to phosphorylate the C-terminus and activate IRF3 (*126, 127*). The presence of an NLS in the DBD and an NES in the linker region results in IRF3 shuttling between the nucleus and cytoplasm. Retention in the nucleus however, requires phosphorylated (activated) IRF3 binding to CBP, while inactive IRF3 localizes exclusively to the cytoplasm (*100*).

The IRF3 C-terminal autoinhibitory region possesses two clusters of Ser/Thr residues that may be targeted for phosphorylation. Site I includes S385 and S386 while Site II is composed of S396, S398, S402, T404, S405 (*92*). Residues S386 and S398 were identified as being critical to IRF3 dimerization and association with co-activator CBP (*128-130*). TBK1 appears to target Site II residues S396 and S402 for phosphorylation, followed by phosphorylation at Site I residue S386, which induces dimerization (*131, 132*). A study utilizing co-expression of the IRF3 IAD with IKKε revealed multiple distinct phosphorylated IRF3 species. Mass spectrometry identified phosphorylation patterns including mono-phosphorylation at S402, dual phosphorylation at S386/S402, and dual phosphorylation at S396/S402 (*133*). Only the dually phosphorylated S386/S402 species formed dimers and associated with CBP. Although multiple phosphorylation sites

have been identified as important, controversy still exists as to the exact mechanism, order, and residues targeted in IRF3 phosphorylation.

IRF5 Activation

IRF5 activation parallels the mechanism of activation of IRF3 despite responding to different PRRs and being expressed in only a subset of cells. Like IRF3, inactive IRF5 shuttles between compartments, but localizes to the cytoplasm under basal conditions. IRF5 possesses two NLSs, one in the N- and C-terminal domains each (*99*), and a single, dominant NES (*101*). Lysine residues in the C-terminal NLS are directly targeted for K-63 linked ubiquitination, a step necessary for nuclear translocation and IRF5 mediated transcriptional activation (*134*).

The exact pattern of phosphorylation for the IRF5 C-terminal serine cluster is less well understood than for IRF3. Initial experiments indicated that mutation of S425, S427, and S430 all impacted the ability of cells to respond to Newcastle Disease Virus (NDV) infection (*99*). Isothermal titration calorimetry (ITC) experiments tested phosphomimetic mutations (serine to aspartic acid) of the serine cluster for interaction with co-activator CBP. Results indicate S425D, S427D, S430D and S436D all induced tighter interactions with CBP, although the fold change in affinity between IRF5 and CBP was significantly lower compared to the parallel experiment in IRF3 (*130, 135*). The S430D mutation resulted not only in the strongest interaction with CBP, but also induced dimerization in the C-terminal IRF5 crystal structure (*135*). Recently, co-expression of IRF5 with different kinases or other activators resulted in phosphorylation of S425 and S436 (*136*).

Therefore, similar to IRF3, phosphorylation of a cluster of serine residues in the C-terminus of IRF5 appears critical to phosphorylation and activation of the molecule. However, the exact targets and pattern of phosphorylation remains to be determined.

IRFs and Disease

Pathogens frequently target IRF family members to evade the host immune response. In particular, many pathogens have evolved mechanisms to target IRF3 function. This is no surprise given that IRF3 functions in almost all cells, serving as a universal factor in immunity. As such, pathogens, particularly viruses, have evolved mechanisms to target the IRF3 pathway, or directly inactivate IRF3 to evade or suppress the immune response. For example, multiple human herpes viruses including HSV-1, varicella-zoster, Epstein-Barr, and human cytomegalovirus target IRF3 or its activation pathway (*137-141*). Similarly, members of the coronaviridae family, including SARS virus, and the filoviridae family, including Ebola and Marburg virus, evade immune response and establish infection in part by targeting IRF3 function (*142-145*). Understanding the mechanisms of IRF3 function may in turn lead to future prevention or treatment of various infections.

Despite its limited cell specific expression, some viruses still target IRF5 activity to suppress the immune response. The limited expression pattern of IRF5 curtails the role it plays in response to viral infection and as such, fewer examples exist of viral evolution targeting IRF5. One exception is human herpes virus 8, or Kaposi's Sarcoma-associated Herpes Virus (KSHV). KSHV infection can cause formation of small tumors, Kaposi's

sarcoma (*146*), and two lymphoproliferative disorders; primary effusion lymphoma (PEL) (*147, 148*) and multicentric Castleman's disease (*149*). During viral evolution KSHV acquired a viral homologue to the IRF family (vIRF), which produces four transcripts (vIRF1-4) (*150*). Of these vIRFs, vIRF3 blocks activation of both IRF5 and IRF7, blunting the IFN response (*151, 152*). Interestingly, PEL growth requires vIRF3 (*153*), possibly due to repression of IRF5 tumor suppressor functions (*152*). Thus, KSHV appears to target the IRFs to establish infection, but the resulting suppression of IRF5 function contributes to uncontrolled cell proliferation.

Although necessary for response to pathogens and tumor suppression, aberrant IRF5 activity is linked with the autoimmune disorder systemic lupus erythematosus (SLE). Although the cause of SLE is not completely understood, overactive immune cells, particularly B-cells, characterize the disease. This results in increased cytokines and autoantibody production (*154*). Various studies identified IRF5 polymorphisms as a marker for increased probability of developing SLE (*155-158*). The IRF5 polymorphisms are thought to upregulate expression (*159, 160*) or produce risk associated splice variants (*156, 161*). While other studies focused on the role of IFN production in SLE (*162*), two recent studies utilized SLE mouse models to demonstrate IRF5 expression is critical for SLE development independent of IFN expression (*163, 164*). These results suggest IRF5 regulation of other cytokines leads to development of SLE. Therefore, pharmacologically targeting IRF5 function may be a viable future treatment for SLE.

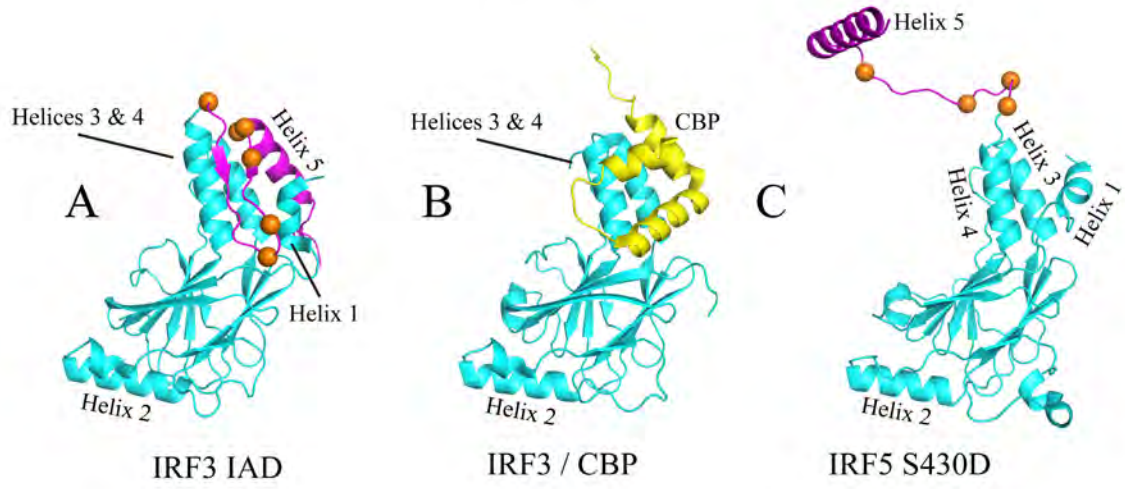
IRF Structures

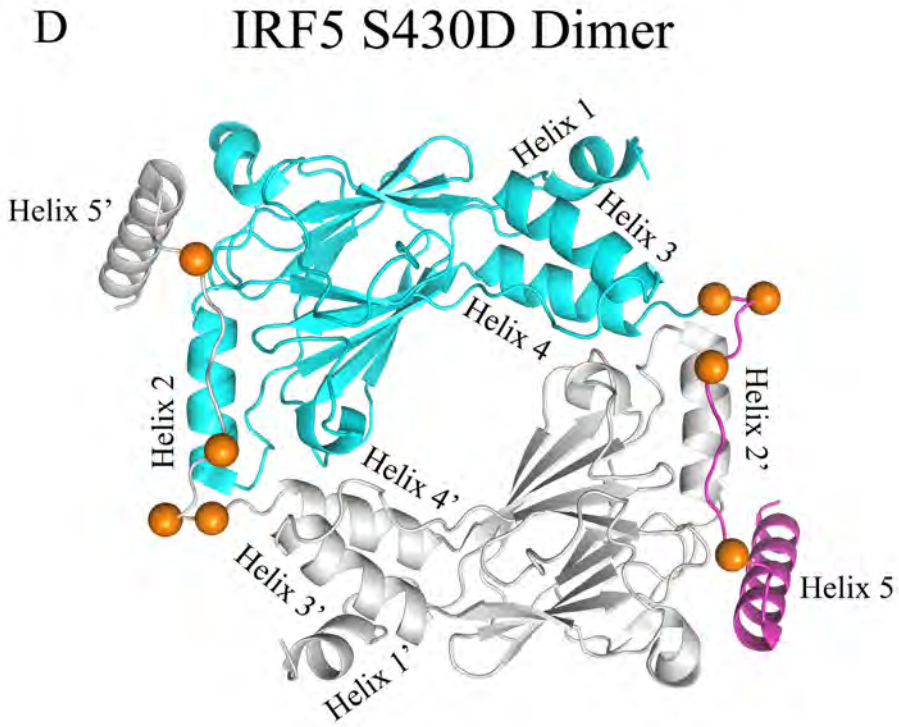
There are several IRF structures deposited in the Protein Data Bank that have advanced understanding of IRF function. The murine IRF1 and IRF2 DBD bound to DNA and both the free and DNA bound forms of the human IRF3 and IRF7 DBD, have been crystallized (*93, 94, 165-167*). No structures exist for an IRF containing the linker region between the DBD and IAD.

Although there are fewer structures published for the IAD, the existing structures have significantly impacted understanding of IRF regulation, including dimerization and co-activator binding during activation. The first C-terminal IRF structures consisted of an autoinhibited IRF3 IAD (*168, 169*). The structure revealed that residues responsible for autoinhibition pack against helices 3 and 4, shielding them from solvent exposure. This same region contains the Ser/Thr phosphorylation targets described previously (*168*). Based on these observations, the Kai Lin laboratory predicted that phosphorylation would unfold the autoinhibitory region, exposing the helices. To test this hypothesis, and IRF3 IAD lacking the autoinhibitory region was crystallized with the IRF binding domain of CBP, forming and IRF3 IAD/CBP complex (*170*). Removing the autoinhibitory region exposed the CBP binding site, allowing CBP to bind, suggesting that releasing autoinhibition through phosphorylation is critical for co-activator binding (*170*). Following the IRF3 IAD/CBP complex, a phosphomimetic IRF5 (S430D) IAD was also crystalized in an attempt to favor dimerization (*135*). Unlike the IRF3 IAD/CBP complex, the phosphomimetic IRF5 construct possessed its autoinhibitory region, which extended

Figure 1.6: Activation and dimerization in the IRF family. From the Kai Lin laboratory, adapted from Chen et al. (*135*) (A) The activated IRF-5 IAD S430D phosphomimetic mutant demonstrates that the autoinhibitory region (green) extends away from the globular domain (blue) exposing the cofactor binding site at helices 3 and 4. Putative phosphorylation sites are shown as orange spheres. (B) The IRF-3 IAD (blue) in the autoinhibited conformation. The C-terminal autoinhibitory region (green) shields helices 3 and 4. (C) Removal of the autoinhibitory region and incubation with co-activator CBP (yellow) reveals that activation through phosphorylation and subsequent displacement of the autoinhibitory loop exposes the CBP binding site. (D) The displaced autoinhibitory region, including helix 5, makes extensive contact along the IRF-5 IAD S430D dimer interface. The autoinhibitory region of one subunit (green) forms contacts near helix 2' of the second subunit (grey).

Figure 1.6


A
Helices 3 & 4
Helix 5
Helix 1
Helix 2
IRF3 IAD
B
Helices 3 & 4
CBP
Helix 2
IRF3 / CBP
C
Helix 5
Helix 3
Helix 4
Helix 1
Helix 2
IRF5 S430D


D
IRF5 S430D Dimer
Helix 1
Helix 5'
Helix 3
Helix 4
Helix 2
Helix 2'
Helix 4'
Helix 3'
Helix 5
Helix 1'

away from the globular domain to form the dimeric interface between IRF5 subunits (*135*). Mutational analysis supports the crystallographically observed dimer as representative of an active conformation for both IRF5 and IRF3 (*135*). Thus, a model was formed in which phosphorylation displaces the autoinhibitory region, simultaneously exposing the co-activator binding site and forming extensive contacts with a neighboring IRF5 subunit, resulting in a fully activated IRF dimer **(Figure 1.6)**. However, Takahasi et al. has challenged this model for IRF3 activation. They show a phosphorylated IRF3 structure that does not dimerize; instead it aligns very closely with the unphosphorylated IRF3 IAD (*171*). Flaws in their approach to crystallizing phosphorylated IRF-3 (see below) may have led to a similar conformation as seen in the unphosphorylated form.

Scope of The Thesis: Generating New IRF Structures to Understand IRF Function

Despite the large number of IRF structures, the model of IRF activation remains incomplete and unconfirmed. Takahasi et al. assert IRF3 does not dimerize in a manner similar to IRF5 based on their phosphorylated IRF3 structure (*171*). Instead of a model where the autoinhibitory region is displaced, they favor a model of direct interaction between the charged phosphate of one subunit with a charged surface of the second subunit. However, the structure of phosphorylated IRF3 does not demonstrate this direct interaction model, leaving their hypothesis unsupported. Despite this, they reason that IRF5 and IRF3 possess distinct activation mechanisms (*171*), ignoring the extensive sequence, structural and functional similarities between the two molecules.

Current IRF structures lack the linker region between DBD and IAD. The functional significance of the linker in all IRFs underscores the necessity of understanding the linker region's conformation. The lack of a structure that contains both a DBD and IAD also leaves a gap in the understanding of IRF function. The two domains may interact with one another through the influence of the linker. Recent data even suggests the domains may bind one another (*172*), adding the possibility of an additional layer of control over IRF activation. Similarly, a ~40 amino acid segment of a viral IRF homologue (vIRF3) binds portions of the N- and C-terminal domains of IRF5 and IRF7, despite a sequence gap of ~150 residues between the bound regions (*151, 152*). This suggests proximity of the DBD to the IAD, reinforcing the idea that IRF inter-domain interaction may have a role in protein function and inhibition. Through understanding the KSHV mediated inhibition of IRF5, future work could pharmacologically target IRF5 in specific autoimmune disorders.

In order to address the current gaps in understanding IRF activation, I sought to solve three different structures. The first structure, an IRF3 IAD/CBP construct that contains the autoinhibitory region, would address whether IRF3 dimerizes and, thus activates, in a mechanism analogous to IRF5. The IRF3 molecule would require phosphomimetic mutations to displace the autoinhibitory region analogous to the IRF5 structure. The second structure contains the missing linker region and inter-domain interactions. A full length IRF3 molecule would answer questions regarding conformations and inter-domain contacts in a full length IRF, and their implications on activation. The third structure was the complex of Kaposi's sarcoma associated Herpes

Virus (KSHV) protein vIRF3 bound to full length IRF5. Determining how vIRF3 achieves IRF5 inhibition by binding separate IRF5 domains would provide clues concerning not only IRF activation, but also methods to inhibit IRF5 in the context of autoimmune disease. Although my efforts to solve these structures were not successful, my work provides a platform for future attempts to complete the structural characterization of IRF activation and function.

CHAPTER II

Characterizing Interferon Regulatory Factor Activation

INTRODUCTION

The nine members of the Interferon Regulatory Factor (IRF) family of proteins possess diverse functions in immune response, immune cell development, and tumor suppression (*91, 92*). This chapter will focus on two members, IRF3 and IRF5, and their roles as transcription factors. All members of the IRF family share a highly conserved DNA binding domain (DBD) (*91, 93, 94*). A linker region of ~75 amino acids connects the DBD to the less well-conserved C-terminal domain. Many IRFs, including IRF3 and IRF5, possess a C-terminal IRF association domain (IAD) responsible for homo- and heteromeric interactions (*96, 97*). The C-terminus also contains an autoinhibitory region, restricting IRF association when the protein is in an inactive form (*98, 99*).

IRF3 and IRF5 possess similar functions, yet different expression patterns and unique activities set these two family members apart. Nearly ubiquitous expression of IRF3 across almost all cell types provides a generalized response to multiple types of infection (*92, 103, 107, 109, 110*). IRF5 however is expressed in a subset of immune cells, thereby limiting the role of IRF5 in antiviral responses (*104*). Despite its limited expression, IRF5 still participates in IFN activation upon infection (*106*). Unlike IRF3, IRF5 also functions in tumor suppression, with roles in both DNA-damage and virally induced apoptosis (*114, 115*). Understanding IRF5 activation is of particular interest, due to evidence that links IRF5 variants with autoimmune diseases, most strongly systemic lupus erythematosus (*155-161, 163, 164*).

IRF3 and IRF5 are activated upon signaling from germline encoded pattern recognition receptors (PRRs). These molecules recognize common pathogen associated

molecular patterns (PAMPs) such as LPS or viral nucleic acids. Cell soluble PRRs, such as RIG1 and MDA5, and membrane bound PRRs, including toll-like receptors (TLR) trigger signal transduction cascades upon PAMP binding. Although the participants in each PRR signaling differ, the cascades typically converge, triggering the IRFs to activate transcription of specific genes (*92*). Upon signal transduction, kinases phosphorylate IRF3 (*126, 127, 133*) and IRF5 (*136*) within serine/threonine rich regions of the autoinhibitory domain. Phosphorylation relieves autoinhibition; this leads to IRF oligomerization, nuclear translocation, and ultimately transcriptional activation (*96, 97*).

Upon localization to the nucleus, IRFs associate with other factors to bind DNA at enhancer regions and form a transcriptional activation complex termed the enhanceosome (*94, 117-119*). The interferon (IFN) β enhancer has been widely studied and is used as a model for IRF transcriptional activation. The IRFs and other factors assemble into the enhanceosome and recruit histone acetylation transferases (HAT), including CREB binding protein (CBP), which acetylate nucleosomes positioned to block transcription (*121-125*). Acetylation is followed by recruitment of chromatin remodeling complexes that move the nucleosome, allowing transcription to occur (*121*).

Multiple pathogens target IRF proteins to evade immune responses. Due to ubiquitous expression, viruses frequently target IRF3 activation pathways or IRF3 itself. These viruses include multiple herpes viruses, filoviruses (Ebola and Marburg viruses), as well as coronaviruses (including SARS) (*137-145*). Kaposi's Sarcoma-associated Herpes Virus (KSHV), which encodes four viral IRF homologues (vIRF), inhibits IRF5; vIRF3 targets the activity of both IRF5 and IRF7 (*150-152*). Understanding the activation

of IRF family members may provide insight into the mechanisms used by pathogens to inhibit IRF functions. Moreover, understanding how the IRFs become deactivated may eventually lead to new therapies for autoimmune disorders.

Structural studies have already made significant progress into the mechanisms underlying IRF activation. Multiple IRF DBD structures have been solved, both free and bound to DNA, providing information on specific DNA induced changes and contacts (*93, 94, 165-167*). More importantly for understanding regulation, structures have been solved for the IAD of IRF3 and IRF5 (*135, 168-171*). The initial IRF3 structures included the IAD complete with the autoinhibitory region (*168, 169*). The autoinhibitory region contains the putative phosphorylation sites and packs against a surface formed by helices 3 and 4, shielding them from solvent exposure. Based on this stereochemical arrangement, the laboratory of Kai Lin proposed that this surface could become exposed if phosphorylation triggers an unfolding of the autoinhibitory region. To test the hypothesis, the Lin laboratory determined a crystal structure of the IRF3 IAD without the C-terminal autoinhibitory region in complex with the IRF binding region of the co-activator CBP (*170*). This structure demonstrated that the surface obscured by the C-terminal autoinhibitory domain serves as a binding site for the co-activator CBP. In order for CBP to bind, helices 3 and 4 must be exposed, supporting the hypothesis that activation by phosphorylation causes the autoinhibitory domain to unfold. These results were followed by the first IRF5 crystal structure with an intact autoinhibitory domain, in which a phosphomimetic mutation (S430D) was introduced to favor dimerization. In this IAD structure the C-terminal autoinhibitory region was displaced from the CBP binding

site, forming extensive contacts with a second subunit to form a dimer (*135*). Thus a model was developed for IRF activation in which phosphorylation, dimerization, and CBP binding are coupled through the disposition of the autoinhibitory region. The applicability of this model to IRF3 has recently been challenged by another group that did not observe structural changes in the autoinhibitory region upon crystallization of a dually phosphorylated IRF3 C-terminus (*171*).

Although the model for IRF activation has begun to take shape, questions persist regarding fine details of IRF activation. In order to address the similarity between IRF3 and IRF5 activation, we attempted to solve the crystal structure of the dimeric IRF3 IAD, a construct that includes the autoinhibitory domain, in complex with CBP. Next, we endeavored to solve the structure of full-length IRF3 in an effort to determine if there are functionally important interactions between the DBD and IAD and to understand the role of the region linking these domains. This linker, absent in the existing IRF structures, may serve a role in IRF activation beyond simply tethering the domains together. The linker regions contain localization signals (*100, 101*) and may not be unfolded as previously thought (*102*). Finally, we sought to address how KSHV viral IRF homologue, vIRF3, achieves inhibition of IRF5. We addressed this issue by trying to solve the crystal structure of full-length IRF5 bound to vIRF3. Through structural studies, we intended to understand how vIRF3 inhibits IRF5 and IRF7 by simultaneously binding the N- and C-terminal regions despite a sequence gap of ~150 amino acids. The inhibition observed as a result of vIRF3 binding to both the DBD and IAD further suggests that the conformation of these domains relative to one another may have a role in IRF activation,

dimerization or the ability to bind DNA. A complex of IRF5 and vIRF3 would not only address the mechanism of KSHV mediated inhibition, but would also demonstrate possible means of pharmacologically targeting IRF5 for inhibition in autoimmune disorders.

MATERIALS AND METHODS

Constructs and Expression

IRF3 5D /CBP Complex:

The IRF3 phosphomimetic 5D construct (143-427) was cloned from the wild type C-terminal construct (*168*) in the pET28a vector (Novagen). The following phosphomimetic mutations were introduced through site directed mutagenesis: S386D, S396D, S402D, T404D, and S405D. A codon optimized CBP (2069-2113) was synthesized by Genescript and cloned into the pGEX6P-1 vector (GE). Both constructs were simultaneously transformed into HB101 cells and plated on agar containing Kanamycin (pET28a marker) and Ampicillin (pGEX6P-1 marker). Colonies were used to make starter cultures that were scaled up to six liters in tryptone-phosphate media (2% bacto-tryptone, 1.5% yeast extract, 0.2% dibasic sodium phosphate, 0.1% monopotassium phosphate, 0.8% sodium chloride). Cultures were grown at 25°C until induction with 0.05 mM IPTG at OD_{600}= 0.5. Temperature was lowered to 20°C and cells were harvested 18-24 hours after induction. Cells were pelleted and resuspended in buffer containing Complete EDTA-Free protease inhibitors (Roche), 50 mM Tris (pH 7.3), 100 mM sodium chloride, 0.2 mM EDTA and then flash frozen to be stored at -80°C.

Full Length IRF3 2D:

The full-length IRF3 2D (S386, 396D) phosphomimetic construct cloned into the pET28a vector was transformed into BL21-DE3 cells for expression. Culture growth, induction, and harvesting protocol were identical to above.

vIRF3:

Various vIRF3 single and co-expression constructs were created in pet28a, pGEX6P-1, and pMNO vectors. Michael Omartian modified a multi-cassette vector utilized in bacterial expression of hemoglobin (*173*) to create the pMNO co-expression vector. Expression of vIRF3 alone consisted mostly of vIRF3 (235-327) in pGEX6P-1 (N-terminal GST tag), with or without an additional C-terminal His_6 tag, transformed into BL21-DE3 cells. Growth, induction, and harvesting were identical to above with the exception of induction at 18-20°C with 0.2 mM IPTG. Co-expression constructs consisted of vIRF3 (235-327) or (240-277) in one cassette and His_6 tagged codon optimized synthetic (Genescript) IRF5 (1-467) or IRF5 (1-467) S430D in the second cassette of the pMNO vector. Constructs were transformed into *E.* coli Q cells (W3110*lacIq* L8) (*174*). Growth, induction and harvesting were identical to above with the exception of utilizing a higher IPTG (0.2 mM) concentration for culture induction.

Purification:

IRF3 5D /CBP Complex

Cells were lysed, cell debris pelleted, and lysate applied to nickel (NiNTA agarose) beads with the same protocol as purification of CtBP1 & CtBP2. After a 2-hour incubation, the nickel beads were applied to a column and washed with 100 mL buffer containing 100 mM sodium chloride, 50 mM Tris pH 7.3, 20 mM imidazole, and 1 mM DTT. 2-3 mL of wash was kept to cover the bed of beads for cleavage of the affinity tags. 15 µL of Thrombin (Novagen) diluted in 500 µL dilution buffer and 100 µL Prescission

protease (GE) were then added to the beads. After mixing the beads with a stir rod, the column and beads were incubated overnight at 14°C. The complex was eluted off the beads in eight 5 mL elutions with wash buffer. Complex elution was verified with samples of elutions and beads utilizing SDS-PAGE. Elutions containing the complex were pooled and dialyzed overnight into DEAE buffer (10 mM sodium chloride, 20 mM Tris pH 7.3, 1 mM DTT). The sample was loaded onto a pre-equilibrated DEAE sepharose column, and a 0-400 mM sodium chloride gradient was applied. The A_{280} was used to monitor elution. Elution samples were run on SDS-PAGE to determine purity and presence of both proteins. Elutions were concentrated to 28-66 mg/mL, and flash frozen for storage at -80°C.

Full Length IRF3 2D

Full length IRF3 2D purification utilized the same protocol as IRF3 5D / CBP with the following differences: The affinity (His_6) tag was not cleaved from the protein. Instead, protein was eluted from the beads with 300 mM imidazole in wash buffer. Protein was dialyzed and run on a DEAE sepharose column with a 0-300 mM sodium chloride gradient. Protein was concentrated to 4.5 mg/mL and flash frozen for storage at -80°C. Later crystallization trials demonstrated that in a lower salt buffer the protein could be concentrated as high as 12.7 mg/mL. (Note: Based on purification of full-length IRF5 constructs, size exclusion may be more efficient than ion exchange in removing contaminants and proteolytic degradation products from full-length IRF3 constructs.)

vIRF3 (235-327) pGEX6P-1

Cell lysis and cell debris removal were identical to above. Lysate was applied to 5 mL of water washed GST beads and incubated 2 hours at 4°C. Beads were washed with 150-200 mL of 100 mM sodium chloride, 50 mM Tris pH 7.3, 1 mM EDTA, and 1 mM DTT. 100 µL Prescission protease was added and protein was cleaved overnight at 4°C. Eight 5 mL elutions with the wash buffer were analyzed by SDS-PAGE. Elutions containing protein were pooled. Constructs with a C-terminal His_6 tag were incubated with nickel bead and eluted with imidazole as above. Constructs without an additional affinity tag were subjected to ion exchange as above.

vIRF3 (240-277) / Full Length IRF5 co-expression pMNO

Cell lysis and cell debris removal were identical to above. Lysate was applied to nickel beads (His_6 present on accompanying IRF5 constructs). Cleavage off of the beads with Thrombin was attempted, but proves consistently problematic with any construct containing the N-terminal domain of IRF5. As such, proteins were eluted with 300 mM imidazole, and dialyzed into 300 mM sodium chloride, 50 mM Tris pH 7.3, 1 mM EDTA, and 1 mM DTT. Eluted protein was concentrated and applied to a Superdex 200 size exclusion column. Fractions containing the protein were concentrated and flash frozen for storage at -80°C. (Note: This protocol, utilizing size exclusion, works well for purification of full-length IRF5 or IRF3 constructs expressed alone.)

Crystallization:

IRF3 5D /CBP Complex

Initial crystallization screening led to formation of small crystals in two conditions: 30% PEG 400 and 100 mM CAPS pH 10.5 (Wizard I crystal screen, Emerald Biosystems) and in 1.7 M 1,6 hexanediol, 100 mM magnesium chloride, and 50 mM Tris pH 8.5 (Hampton II crystal screen diluted 50%, Hampton Research). Protein (15 and 19 mg/mL in screens respectively) was mixed 1:1 with buffer. Crystals appeared in 2-4 weeks. Conditions were optimized and crystals used for diffraction were grown in hanging drops with 10-30 mg/mL protein grown in 1.8-2.4 M 1,6 hexanediol, 100 mM magnesium chloride, and 50 mM CHES pH 10.0 at a ratio of 3:2 buffer to protein. Optimized conditions resulted in crystal growth from 1-5 days.

Full Length IRF3 2D

Initial crystal screening resulted in crystallization of full-length IRF- 2D in 200 mM sodium or potassium citrate with 20% (w/v) PEG 3350 (PEG/Ion screen, Hampton Research). Conditions were optimized to grow crystals for diffraction in hanging drops with 12-16% PEG 3350, 100 mM potassium citrate, and 50 mM imidazole pH 7.8 at a 1:1 ratio with full-length IRF-3 2D (1.5-12.5 mg/mL).

RESULTS

Dimeric IRF-3 IAD / CBP complex

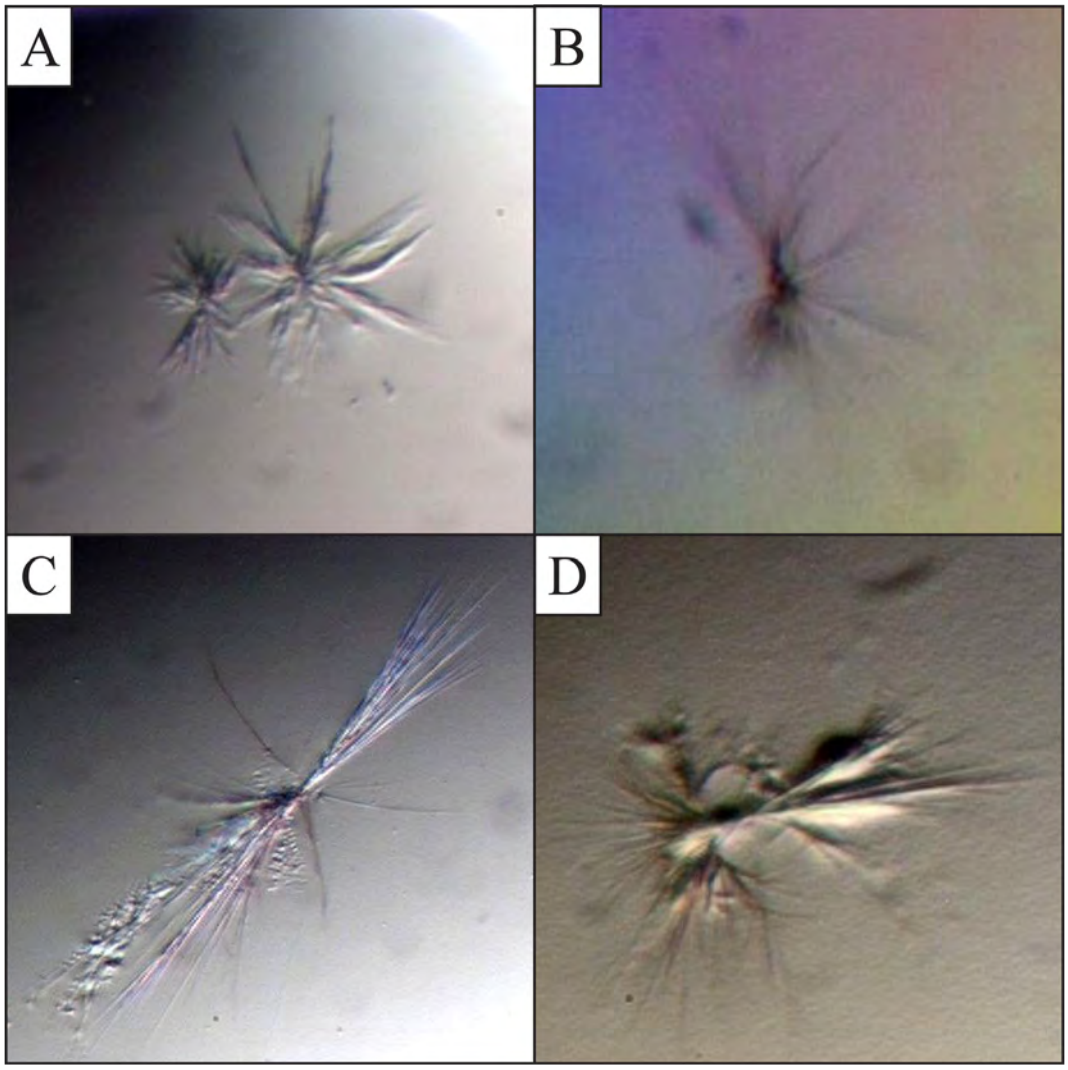
Solubility and crystallization

IRF3 5D (173-427) / CBP (2069-2113) co-expression produces a stable complex that can be readily purified. The IRF3 5D construct was designed to be similar to constitutively active, dimeric IRF3 phosphomimetics utilized in previous studies (*175, 176*). Initial IRF3 5D (173-427) purification resulted in a relatively low yield (0.5-0.75 mg/L culture) of protein with low maximum solubility (2 mg/mL). Co-expression of the CBP IRF binding domain (2069-2113) on a separate plasmid resulted in stable complex formation with vastly improved yield (5 mg/L) and solubility (66 mg/mL). The large increase in solubility and stability suggests the hydrophobic CBP binding site was solvent exposed, resulting in low solubility unless shielded by CBP. Interestingly, IRF5 S430D is stable as a dimer and was crystallized without CBP, highlighting stereochemical differences between IRF3 and IRF5 upon phosphomimetic activation. Although the exact cause of the difference remains unclear, partial covering of the exposed CBP binding site by the extreme IRF5 N-terminus, observed in the phosphomimetic dimer, shields a small portion of the hydrophobic surface from the solvent. IRF3 may be unable to achieve this partial coverage of the exposed CBP site as a result of sequence variation and construct differences, potentially explaining the loss of solubility and stability in the phosphomimetic IRF3 mutant.

The purified IRF3 5D / CBP complex was subjected to different crystallization screens, resulting in crystal growth in two different conditions. The crystals grew in a

Figure 2.1: IRF3 5D / CBP crystallization. (A) The initial IRF3 5D / CBP crystal screen hit in 30% PEG 400 and 0.1 M CAPS pH 10.5 (protein concentration 15 mg/mL). Crystals grew slowly over many weeks. (B) A second screening hit with similar morphology in 1.7 M 1,6 hexanediol, 100 mM $MgCl_2$, and 50 mM Tris pH 8.5 (protein concentration 19 mg/mL). (C) With the elevated pH level of condition in (A) and the salt and precipitant of the condition in (B), crystals grew in less than a week (1.8-2.4 M 1,6 hexanediol, 100 mM magnesium chloride, and 50 mM CHES pH 10.0 at a ratio of 3:2 buffer to protein). Protein concentrations ranged from 10-30 mg/mL. The length of the bunch of needles is roughly 600 μm. Individual needle widths were never larger than 15-20 μm. (D) At high protein concentrations a thick skin formed around the edge of the hanging drop. Crystals frequently grew at the edge and became completely embedded in the skin.

Figure 2.1



small “starbursts” of needles nucleating from a central point in both conditions, 3-4 weeks after hanging drops were set up **(Figure 2.1A-B)**. Focused screening resulted in optimized conditions combining the precipitant and salt of one initial condition (1,6 hexanediol and magnesium chloride) with the elevated pH of the second initial condition (pH 10.5). The combination resulted in more rapid formation and increased crystal size, although crystals still grew as needles from a central nucleation point **(Figure 2.1C)**. The number of needles ranged from half a dozen to hundreds. At higher protein concentrations, crystals formed embedded within a thick “skin” that formed on the edge of the droplets **(Figure 2.1D)**.

Diffraction

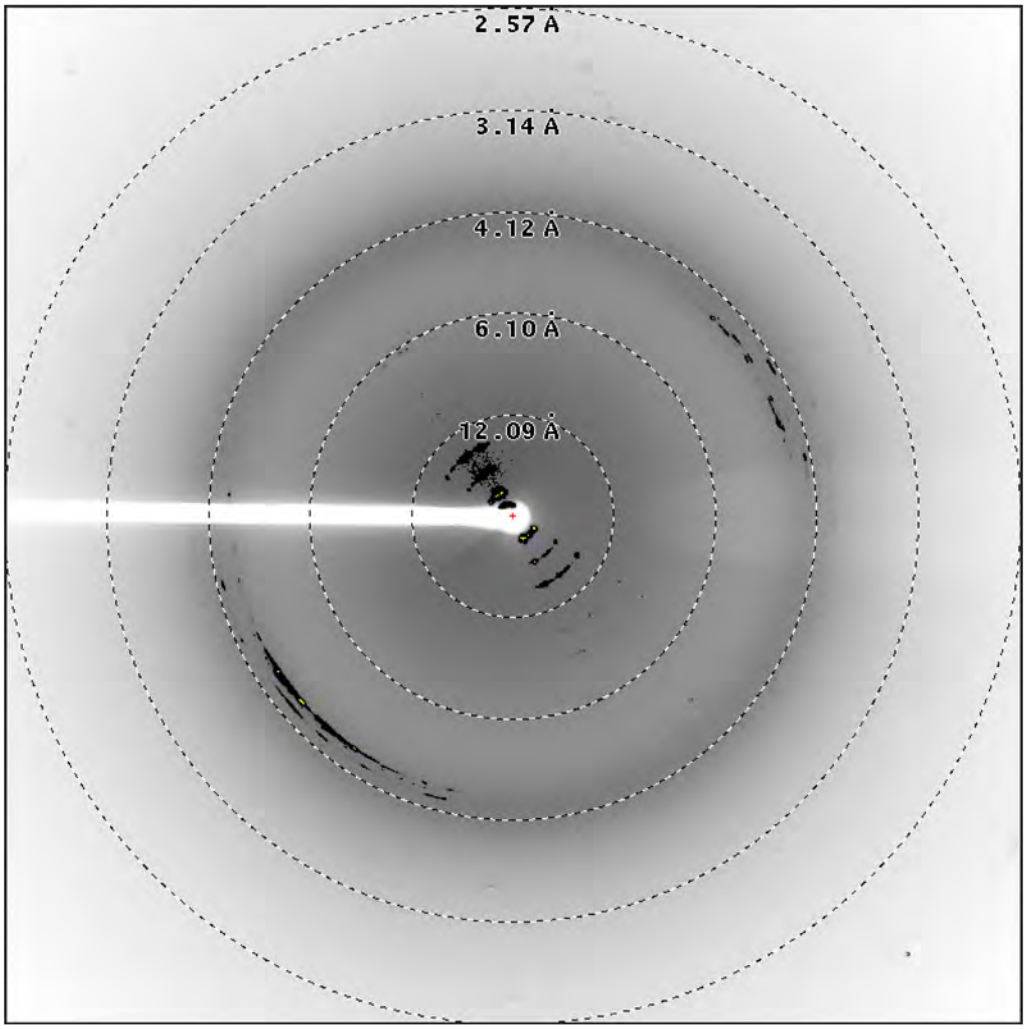
IRF3 5D / CBP crystals do not diffract well. Crystals were harvested for diffraction on the GM/CA CAT minibeam at APS. Large individual needles (100-200 μM length, 15-20 μM width) were collected as well as full clusters of needles. Crystals were frozen in various different cryogenic conditions or simply frozen in crystallization buffer. Despite the incredibly small width of the crystals, diffraction was observed **(Figure 2.2)**. However, the diffraction pattern was low resolution and resembled that of fiber diffraction. No dataset was successfully collected.

Attempts to modify the IRF3 construct to improve crystallization were met with severely reduced levels of expression. The IRF3 5D construct was truncated further on both the N- and C-termini. The N-terminus was modified to match the sequence of the phosphorylated IRF3 construct (*171*), which corresponds roughly to the first ordered

residue observed in all the IRF3 IAD structures. The C-terminal truncation was an attempt to remove residues not involved in secondary structures. However, it appears to destabilize the protein, causing the reduced expression and a noticeable increase in degradation. One option for optimization would be to leave the His-tag uncleaved on either the N- or C-terminus. Different expression protocols may also be able to address the current drawbacks of C-terminal truncation. Alternatively, surface entropy reduction, by means of mutation of large surface residues to alanine or serine, may facilitate new interface contacts and induce crystallization. In the meantime, work with an unmodified C-terminus is ongoing.

Figure 2.2: IRF3 5D / CBP fiber diffraction. The diffraction pattern from IRF3 5D / CBP crystals resembles a fiber pattern. Modification of the construct may induce formation of crystals instead of fibers, yielding the desired diffraction pattern.

Figure 2.2



Full-length IRF3

Crystallization

Full length IRF3 2D (S386, 396D) was purified and crystallized. Although the construct is not constitutively active, as is the IRF3 5D mutant, the construct was amenable to purification and thus utilized for crystal trials. Initial screening resulted in high nucleation and growth of small single crystal rods that reached a maximum size of 50 x 20 x 20 μm **(Figure 2.3A-B)**. Conditions were optimized and trials performed to minimize nucleation in an attempt to increase crystal size. Crystal growth increased, with maximum dimensions reaching 150 μm in length and 30-40 μm in width **(Figure 2.3 C-D)**.

Diffraction

Attempts to collect a full-length IRF3 2D dataset resulted in limited quality diffraction. Full length IRF3 2D crystals were harvested and flash frozen with glycerol or ethylene glycol as a cryoprotectant. Crystals were exposed to the same microbeam x-ray source as the IRF3 5D / CBP complex. Diffraction was rarely observed, however, two separate crystals showed medium-high resolution diffraction between 2.0 and 2.5 Å **(Figure 2.4 A,B)**. The diffraction spots were severely smeared, making data processing impossible. It remains unclear whether intrinsic disorder in the crystal or effects of the cryosolution were responsible for the smearing. Although no dataset was collected, the diffraction demonstrates the potential for solving a full-length IRF3 2D structure in the future.

Figure 2.3: Full-length IRF3 2D crystals. (A) and (B) Typical crystals grown before optimization. High nucleation and small crystal size (maximum 50 x 20 x 20 μm) plagued the first trials. The initial buffer conditions were 200 mM sodium or potassium citrate with 20% (w/v) PEG 3350. (C) and (D) After optimizing conditions there were fewer nucleation events, leading to larger crystal growth. Crystals eventually reached a maximum of 150 μm in length, with a maximum 30-40 μm in width.

Figure 2.3

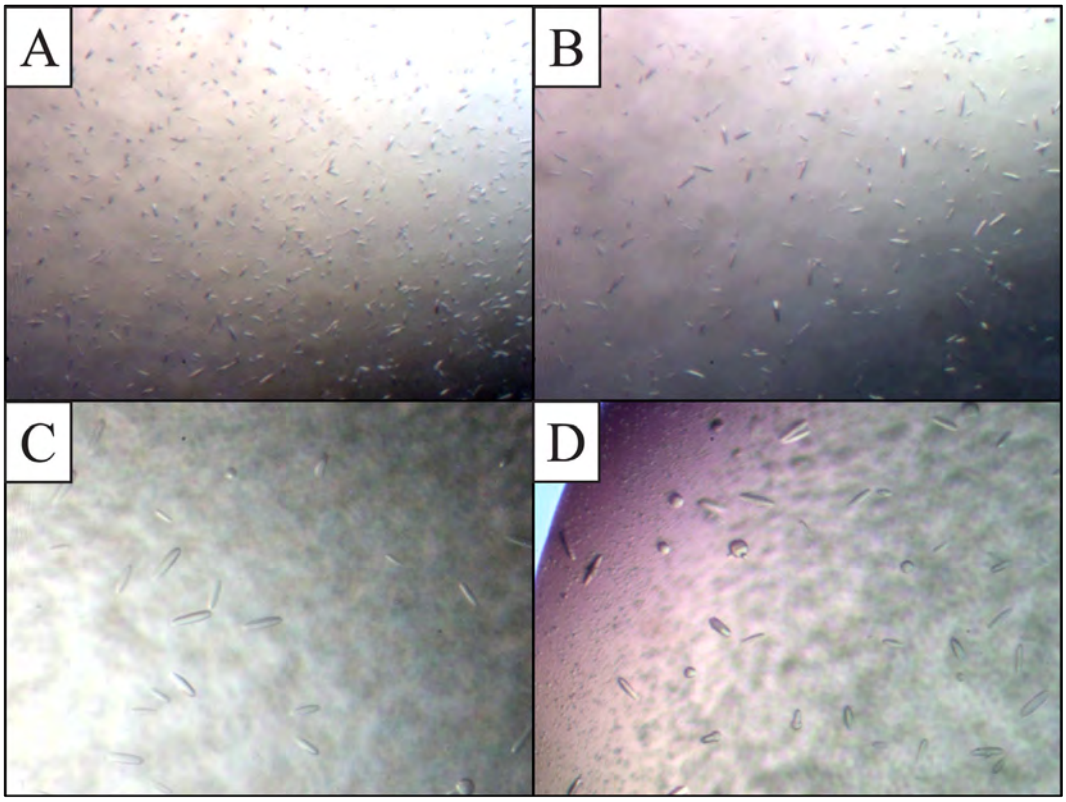
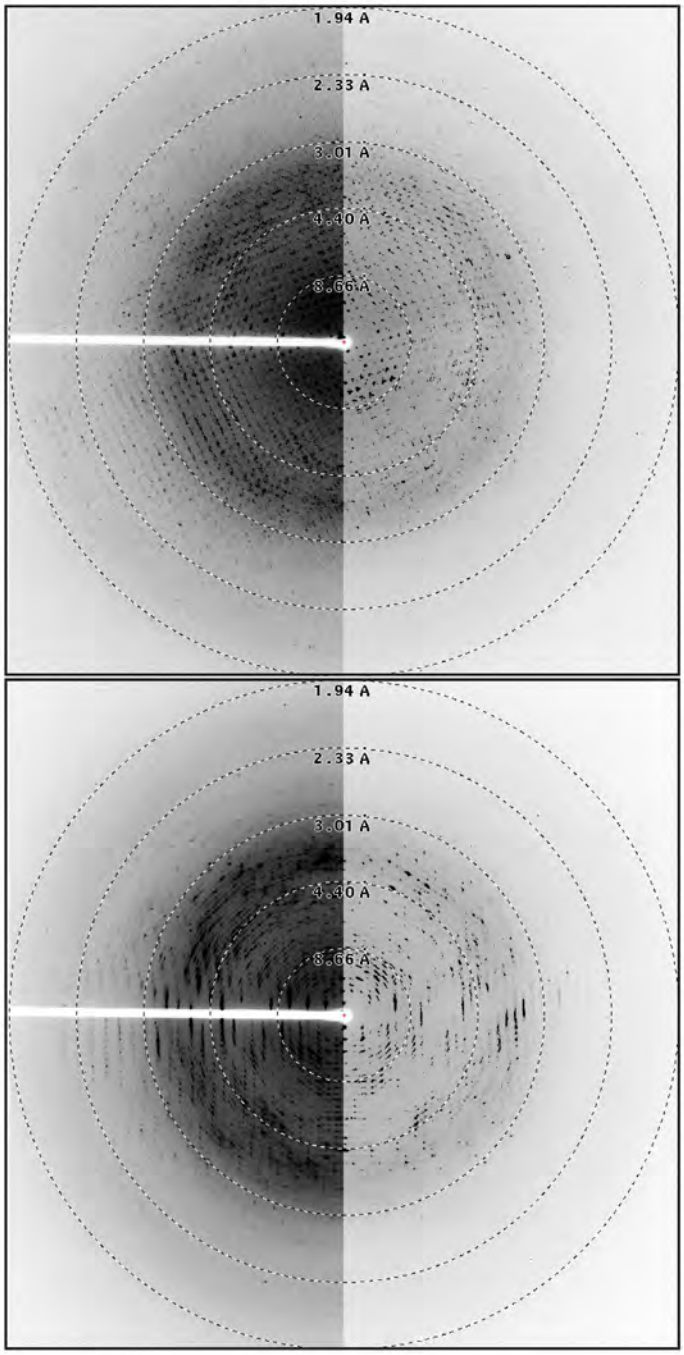


Figure 2.4: Full-length IRF3 2D crystal diffraction. High and low contrast composite picture of the diffraction pattern from full-length IRF3 2D crystals. The data were unable to be processed due to the smearing of the diffraction spots. Despite the poor quality data, diffraction was observed to about 2.0 Å, suggesting slight changes in crystallization might allow the full-length IRF3 2D structure to be solved.

Figure 2.4



vIRF3 and IRF5

Design and expression

The initial vIRF3 construct was designed to include the IRF5 / IRF7 binding domain (240-280) as well as adjacent regions predicted to be folded, for the purposes of increasing solubility and proper folding for crystallization. A protein folding prediction program, FoldIndex (*177*), indicated that in addition to the IRF binding residues, the ~50 amino acids C-terminal to the IRF binding residues also have a propensity to fold into secondary structures. Thus, the long vIRF3 construct was designed with residues 235-327 to include the IRF binding residues and the adjacent folded region or domain. However, expression of this region resulted in an unstable vIRF3 protein prone to proteolytic degradation.

For the purposes of co-expression with IRF5, we utilized both the long vIRF3 construct (235-327) and short vIRF3 construct possessing the minimum residues identified as responsible for IRF binding (240-277). vIRF3 residues 240-277 are predicted to form two helices. IRF5 expression was poor when utilizing the full-length protein (1-488), but greatly increased when unstructured residues were removed from the C-terminus, matching the previously crystallized IRF5 IAD construct. Co-expression of both constructs was attempted with IRF5 residues 1-467 hereon referred to as “full-length”. Unfortunately, neither vIRF3 construct was observed in the soluble fraction of co-expression trials with either the wild-type or the S430D phosphomimetic full-length IRF5.

Purification

Attempts to purify the long vIRF3 construct alone resulted in proteolytic degradation or tight binding to an *E. coli* protein contaminant **(Figure 2.5 A)**. Initial attempts to purify an N-terminal GST tagged 235-327 construct were met with inconsistent expression and continuous degradation throughout purification. Adding a C-terminal His_6 tag appeared to stabilize the protein, increasing the intensity of the largest band on Coomassie stained gels. The C-terminal tag also allowed selection of full-length construct over degraded vIRF3. However, expression remained inconsistent, often with no protein observed. When the protein was successfully expressed, a contaminating *E. coli* protein remained tightly bound to vIRF3. High salt and detergent washes, in addition to affinity tag and ion exchange purification steps, did not remove the contaminant. In the event that this protein was an *E. coli* chaperone, purified complex was incubated with ATP and magnesium. The contaminant, however, remained tightly bound. The low vIRF3 yield and contaminating protein led to abandoning this course of purification.

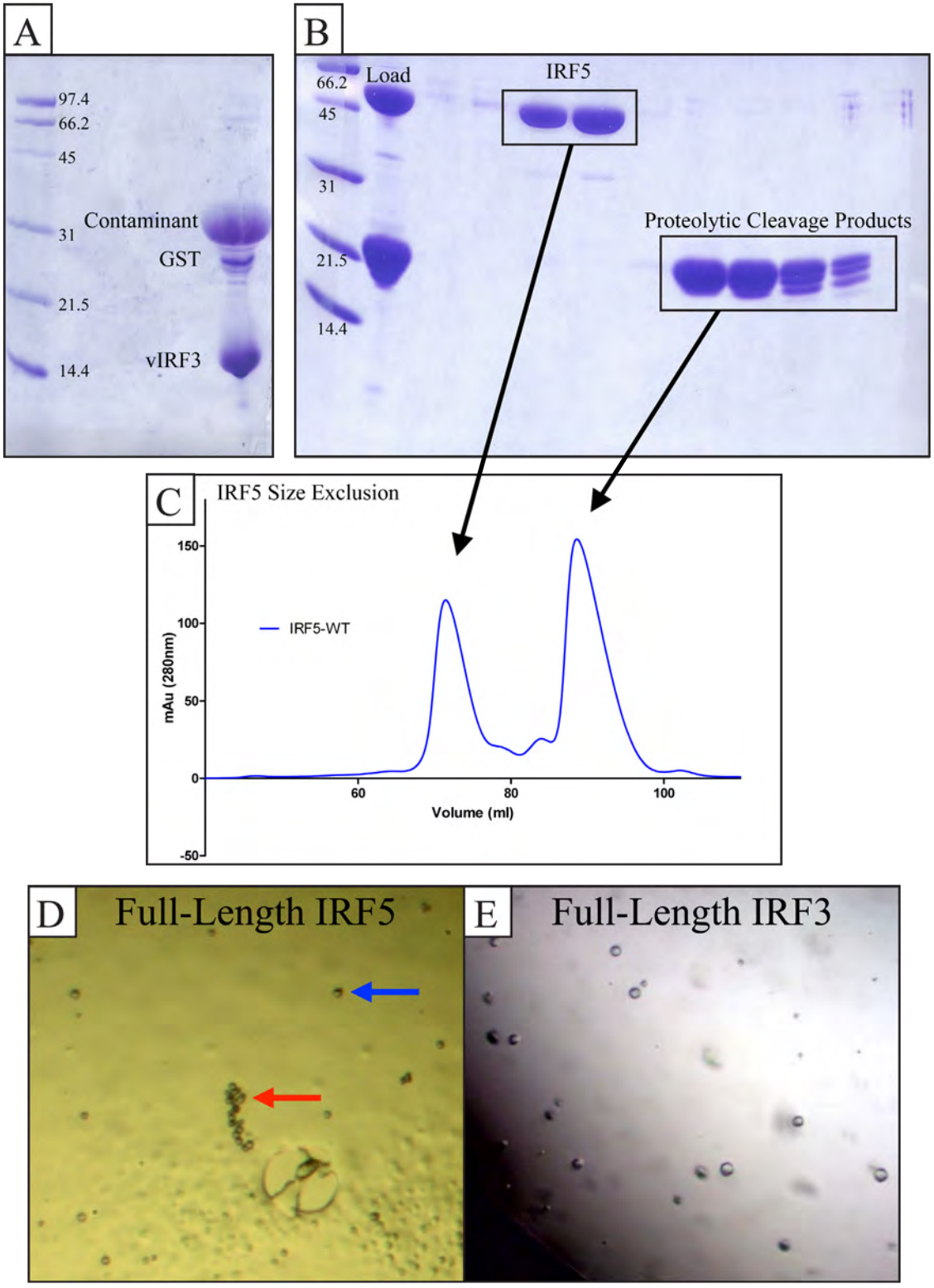
Co-expression of vIRF3 with IRF5 did not result in the desired effects of stabilizing vIRF3 expression, but did result in good expression of both the full-length IRF5 wild type and S430D mutant. Neither vIRF3 construct was observed at any point during or after purification on SDS-PAGE gels. The affinity tag purification followed by size exclusion however, purified large quantities of IRF5 **(Figure 2.5B)**. Wild type IRF5 was then expressed without any vIRF3 ensuring only pure IRF5 was recovered **(Figure 2.5C)**. The purified protein was concentrated to 8 mg/mL and utilized for crystal screening without success **(Figure 2.5D)**. However, the ordered precipitate observed in

the IRF5 crystal trials resembles that observed during full-length IRF3 2D crystallization **(Figure 2.5E)**. The similarities between ordered precipitates imply that like IRF3 2D, IRF5 may indeed be able to be crystallized. Overall, size exclusion appears to function better as a second step in full-length IRF purification than ion exchange.

The S430D mutant proved more difficult to purify than wild type IRF5. Size exclusion did not remove contaminating degradation products common to full-length IRF expression. Detergent washes followed by subsequent size exclusion with detergent present resulted in removal of contaminating degradation products. Unfortunately, the detergent (Triton-X 100) binds so tightly to IRF5 S430D that extensive dialysis and buffer exchanges failed to remove the detergent. Different detergents, more amenable to crystallization, will be utilized in subsequent IRF5 S430D 1-467 purification. The detergent may be binding the CBP binding site, suggesting co-expression with CBP could prevent detergent or proteolytic cleavage product binding to IRF5 S430D. The chromatogram indicated the S430D mutant might be in monomer-dimer equilibrium. At least three peaks containing IRF5 S430D existed between the expected monomer/dimer elution volumes. The unexpected number of peaks and the lack of resolution between peaks prevented drawing any firm conclusions about the oligomeric state.

Figure 2.5: vIRF3 and IRF5 expression. (A) The dual affinity tagged vIRF3 235-327 expressed inconsistently, and co-purified with an *E. coli* contaminant. The contaminant and residual free GST are present along with vIRF3 on a Coommassie stained SDS-PAGE gel. Detergent and high salt conditions did not remove the contaminant. (B) Co-expression of vIRF3 constructs and IRF5 1-467 showed expression of IRF5 only. Proteolytic cleavage products, which failed to separate from full-length protein during ion exchange, eluted separately using size exclusion. (C) Arrows from the Superdex 200 size exclusion elutions indicate the corresponding peaks on the chromatogram. vIRF3 was removed from constructs for expression of full-length IRF5 to be used in crystal trials. IRF5 expression was confirmed by western blot (data not shown). (D) Crystal screening for full-length IRF5 resulted in ordered precipitate but no crystallization. Drops commonly contained phase separation and what appeared to be translucent, spherical precipitate. Individual spheres (blue arrow) commonly bunched forming chains (red arrow). (E) Similar ordered precipitate was observed during full-length IRF3 2D screening.

Figure 2.5


A
97.4
66.2
45
31
21.5
14.4
Contaminant
GST
vIRF3
B
66.2
45
31
21.5
14.4
Load
IRF5
Proteolytic Cleavage Products
C
IRF5 Size Exclusion
150
100
50
0
-50
mAu (280nm)
IRF5-WT
60
80
100
Volume (ml)
D
Full-Length IRF5
E
Full-Length IRF3

DISCUSSION

Although no new IRF structures were determined, future crystallization and biochemistry will benefit from our current results. In the case of determining a dimeric IRF3 / CBP complex, important information was learned about the IRF3 – CBP interaction. Past ITC measurements demonstrated that CBP bound roughly 10 times more tightly to IRF3 phosphomimetic S386/396D than IRF5 phosphomimetic S430D (*130, 135*). The tighter interaction may help explain why the IRF3 5D construct solubility and yield increased roughly 30 and 10 fold, respectively, with CBP co-expression. The difference in binding again implies stereochemical differences between the C-terminal IRF3 and IRF5 phosphomimetic constructs. The increase in solubility and tight CBP binding suggest that the IRF3 autoinhibitory region may require CBP binding to adopt the extended conformation observed in IRF5 S430D. Additionally the IRF5 S430D structure has partial covering of the CBP binding site by the extreme N-terminal residues, a conformation IRF3 may not be able to achieve. CBP induced stability and solubility are perhaps why Takahasi et al. (*171*) failed to observe dimerization, either by our model or their model, in their phosphorylated IRF3 structure. The energetic costs of exposing the CBP binding surface may simply be greater than the destabilization introduced by phosphorylation, requiring CBP to lower the energy barrier, shifting equilibrium towards the extended autoinhibitory region conformation. Requiring CBP would also account for why Takahasi et al. did not detect evidence of their direct interaction model in their structure, but observed dimerization in the presence of CBP (*171*).

Ongoing work with the IRF3 5D IAD and CBP complex involves truncation of the N-terminus without the destabilizing truncation of the C-terminus. Expression levels remain elevated and degradation is minimal when purifying the modified construct complex. The newly purified complex will be subjected to more recently developed crystal screens (MCSG, etc.) with a wider variety of conditions. We are confident that the combination of subtle construct design changes with more varied crystallization conditions will result in diffraction quality crystals that do not exhibit the fiber pattern previously observed. More drastic construct alterations, such as surface entropy reduction, may be required if IRF3 5D / CBP continues to form fibers.

Suitable crystals for a full-length IRF3 structure may be obtained through modification of the full-length IRF3 2D construct or redesign of the phosphomimetic pattern. The IRF3 2D crystals grown to date were small and most did not diffract. Although the two crystals that did diffract displayed unacceptably high mosaicity, diffraction was observed near 2.0 Å. Minor truncations may be made at the N-terminus, but such truncations were not necessary for suitable crystallization of free and DNA bound IRF3 DBD. Truncation at the C-terminus will be avoided due to the problems observed in IRF3 5D construct degradation. Alternatively, the IRF3 2D construct may be swapped for a full-length 5D construct. The full-length IRF3 5D construct expressed very well, but remained contaminated with proteolytic cleavage products after two-step affinity tag and ion exchange purification. Utilizing size exclusion in place of the ion exchange proved vital to removing proteolytic cleavage products during full-length wild type IRF5 purification and may be applied here to full-length IRF3 5D, or any other full-

length IRF to remove degradation products. The 5D mutant should be more active than the 2D mutant, possibly inducing dimerization. The higher expression levels in the absence of CBP indicate the full-length IRF3 5D protein may not require CBP to dimerize.

Modification of full-length IRF5 constructs may also yield the elusive full-length IRF structure. IRF5 1-467 has yet to crystallize in any of the current screens. Modifying the IRF5 1-467 N-terminus may be more beneficial for crystallization than modification of the IRF3 N-terminus. The N-terminus of IRF5 possesses ~10 extra residues relative to IRF3. These residues do not align with the initial IRF3 residues, and are predicted not to be involved in any secondary structure. Removal of these residues may create a construct suitable for crystallization. An IRF5 structure with the DBD, linker, and IAD would not only provide the first IRF5 DBD structure, but also answer questions regarding linker conformation and inter-domain binding.

The issue of inter-domain binding may also be approached through small angle x-ray scattering (SAXS) techniques. SAXS analysis avoids the need to crystallize the protein, making it ideal for the difficult to crystallize full-length IRF constructs. SAXS typically results in low-resolution information on domain arrangements. However, existing DBD and IAD structures should make fitting domains into SAXS results easier. The analysis could be used to compare inactive versus active constructs, identifying any large-scale movement of domains. If successful, SAXS could be applied toward other structural questions involving co-activator binding, DNA binding, and dimerization mutants.

Finally, the KSHV vIRF3 / IRF5 complex may be solved by re-design of the constructs. In order to stably express the IRF binding region of vIRF3, constructs were designed by Dr. William Royer and Dr. Tara Kashav to tether vIRF3 residues 240-280 to the IRF5 N-terminus. Exact construct design and purification specifics are still being explored. At least one fusion construct has been purified and verified by western blot to be IRF5, presumably with a tethered N-terminal affinity tag followed by vIRF3 240-280. Continued work on this and the other IRF constructs will eventually yield our goals of understanding IRF structure and function.

CHAPTER III

Crystal Structures of Human CtBP in Complex with Substrate MTOB Reveal Active Site Features Useful for Inhibitor Design

INTRODUCTION

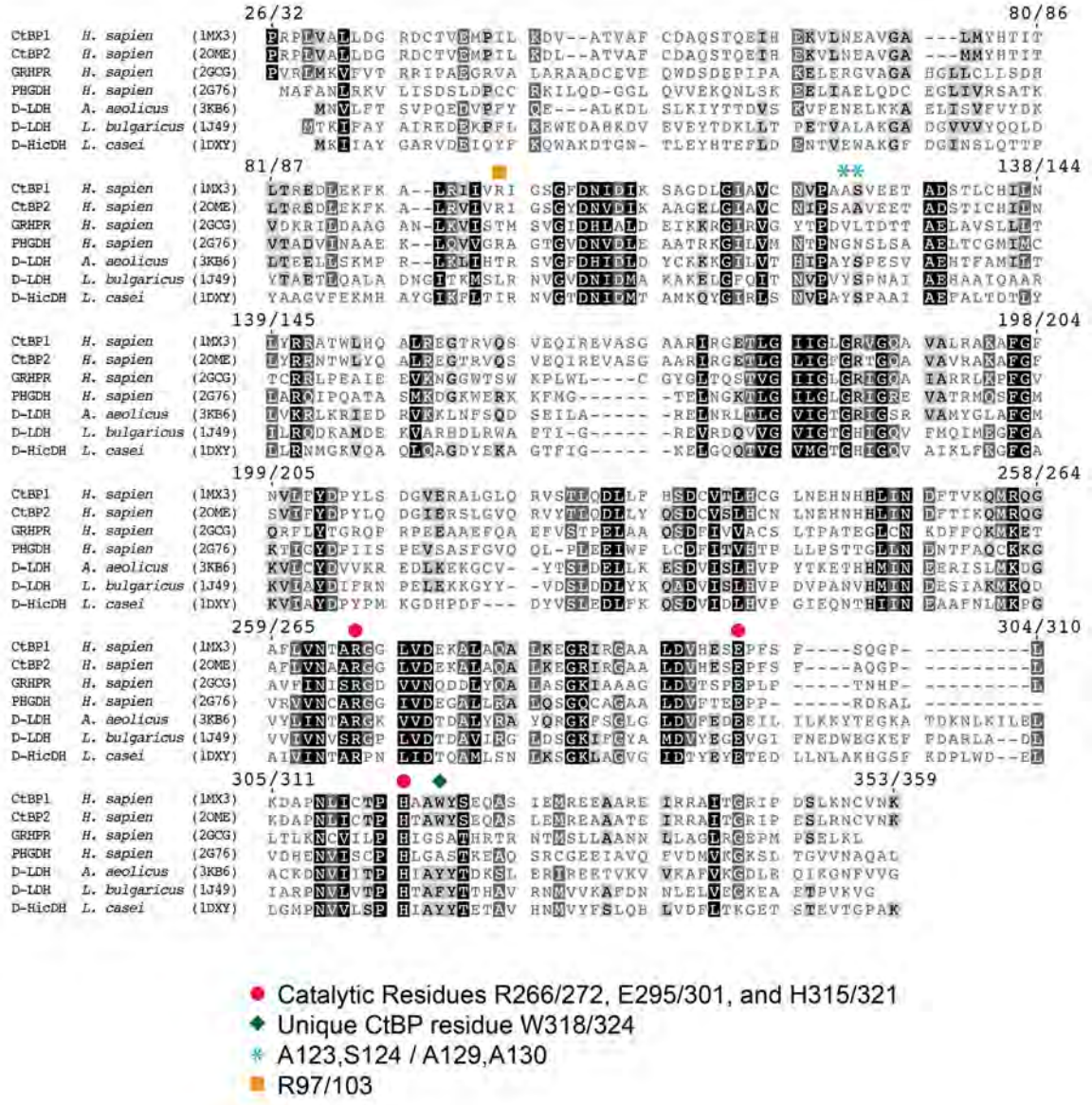
C-terminal Binding Protein (CtBP) 1 and 2 are critical modulators of multiple cellular processes. Discovered first for the ability of CtBP1 to bind adenovirus E1A protein and repress oncogenic transformation (*2, 3*) these two proteins have functional roles in the cytoplasm and nucleus. Although CtBP1 cytoplasmic functions such as trafficking and membrane fission (*9, 56*) are significant, both CtBP1 and CtBP2 (hereon referred to as CtBP) have been studied in the last 15 years primarily for their crucial role in transcriptional regulation.

CtBP functions as a transcriptional regulator by tethering chromatin remodeling proteins, such as histone deacetylases (HDAC), histone methyl transferases (HMT), and histone demethylases, to DNA bound transcription factors (*18, 71*). Intriguingly however, the amino acid sequence of CtBP clearly positions it as a member of a family of enzymes not known for transcriptional regulation **(Figure 3.1)**. Members of this enzyme family, D-isomer specific 2-hydroxyacid dehydrogenases (D2-HDH), reduce or oxidize substrates utilizing coenzyme NAD^+/NADH or NADP/NADPH (*2, 4*). This family includes human proteins Glyoxylate Reductase / Hydroxypyruvate Reducatse (GRHPR) and phosphoglycerate dehydrogenase (PHGDH), as well as bacterial specific members such as D-Lactate Dehydrogenase (D-LDH) and D-Hydroxyisocaproate Dehydrogenase (D-HicDH) (*19, 21, 24, 25*).

Dimerization, induced through binding NAD^+ or NADH (*27*), is critical for CtBP function. Dimerization modulates CtBP repressor activity by providing two available PXLDS binding motifs, one per subunit, to be utilized by multiple CtBP partners (*15*).

Figure 3.1: Sequence alignment of CtBP1 and 2 with Various D-isomer specific 2-hydroxyacid dehydrogenases (D2-HDH). Sequence alignment was carried out with the program Geneious (www.Geneious.com) (*178*). Residue numbers for CtBP1/CtBP2 are shown above the sequences. Red circles above the sequences designate conserved catalytic residues; the green diamond identifies Trp318/324, which provides a unique aromatic surface in the CtBP active site; blue asterisks designates small residues critical to formation of a hydrophilic cavity of CtBP; the orange square identifies Arg97/103, which exhibits conformational variability in CtBP1.

Figure 3.1


26/32
80/86
81/87
138/144
139/145
198/204
199/205
258/264
259/265
304/310
305/311
353/359
CtBP1 H. sapien (1MX3)
CtBP2 H. sapien (2OME)
GRHPR H. sapien (2GCG)
PHGDH H. sapien (2G76)
D-LDH A. aeolicus (3KB6)
D-LDH L. bulgaricus (1J49)
D-HicDH L. casei (1DXY)
Catalytic Residues R266/272, E295/301, and H315/321
Unique CtBP residue W318/324
A123,S124 / A129,A130
R97/103

Additional evidence shows dimerization is necessary to form a stable platform for non-PXDLS factor binding (*15*). The requirements of coenzyme binding and dimerization on transcriptional regulatory activity provide a rationale for evolutionary conservation of the enzymatic function of NADH binding. With a 100 fold higher affinity for NADH than NAD^+ (*26*), CtBP is able to sense the redox state of the cell through small changes in NADH concentration, increasing transcriptional regulatory function when stimuli, such as hypoxia and high extracellular glucose levels, increase the NADH/NAD^+ ratio (*28-30*).

Mounting evidence implicates CtBP repressor function in human cancer. CtBP co-repression targets pro-apoptotic factors (Bik, Noxa), cell adhesion molecules (Keratin-8, E-cadherin) (*66*), and tumor suppressors ($p16^{INK4a}$, $p15^{INK4b}$) (*71*), facilitating the Epithelial-to-Mesenchymal Transition (EMT), conferring resistance to apoptosis and promoting metastasis (*179*). Although repressing these genes is necessary during development, this activity of CtBP can also promote oncogenesis. CtBP itself is targeted by multiple tumor suppressors including HIPK2 (*180*), JNK1 (*181*), and ARF (*79*). Consistent with these cellular effects of CtBP are studies directly linking aberrant CtBP activity to colon, prostate, ovarian, and breast cancer (*8, 72, 75, 81, 182*).

A substrate for CtBP catalysis, 4-methylthio 2-oxobutyric acid (MTOB) antagonizes CtBP transcriptional regulation (*72, 81, 89*). MTOB, the final compound in methionine salvage, was established as a superior substrate for enzyme catalysis when compared to other common 2-keto acids, possibly linking methionine salvage and transcriptional modulation (*63*). MTOB induced apoptosis through displacement of CtBP2 from the *Bik* promoter in HCT116 colon cancer cells, a cell line displaying

elevated CtBP2 levels (*81*). Administering MTOB in mouse xenograft models resulted in decreased tumor burden and prolonged survival for MTOB treated mice versus untreated (*81*). Additionally, MTOB evicted CtBP from target promoters in breast cancer cell lines, shifting phenotypic indicators (E-cadherin/Vimentin) from mesenchymal to a more epithelial phenotype (*72*). These studies indicate that because CtBP enzyme activity is coupled to corepressor activity, modulating enzyme activity with a small molecule may inhibit corepressor function, inducing apoptosis. Although high MTOB concentrations (mM) are required for inhibition of CtBP, its clear effect on cancer cells provides proof of principle that small molecules could be developed to effectively treat cancers specifically regulated by CtBP activity.

Crystal structures of both CtBP1 and CtBP2 have been reported (*10, 39, 44, 90*), but none in complex with putative substrates. Given the ability of MTOB to inhibit the carcinogenic action of CtBP, we determined crystal structures of both CtBP1 and CtBP2 in complex with MTOB in order to provide a structural basis for the design of synthetic inhibitors. MTOB has similar but distinct binding modes in CtBP1 and CtBP2. The interactions of active site residues with the ligand suggest potential MTOB specificity determinants as well as stereochemical characteristics to exploit in future drug design. The structures reveal unique aspects of the CtBP active site size and shape that can be exploited to avoid cross reactivity of potential inhibitors with other members of the D2-HDH family. Thus, this work provides a starting point for the development of inhibitors that exclusively target CtBP1 and CtBP2 that could lead to a future CtBP targeted cancer treatment.

This chapter summarizes the crystal structures and binding of MTOB to both CtBP1 and CtBP2, as well as how the structures will provide a framework for future inhibitor development. The overall structures were examined to determine if a conformational change links enzymatic function to corepressor activity. The focus then moved to ligand binding through assessment of specific types of molecular interactions between MTOB and each protein. This was followed by identifying important individual residue contributions to MTOB binding. Finally, distinct CtBP active site characteristics were examined, providing insight into gaining specificity for future inhibitor design. The structural analysis of the complexes served as the starting point to design tighter binding, CtBP specific molecules. Future inhibitor designs will be tested for their ability to block CtBP transcriptional activity, helping elucidate the mechanism linking active site binding to transcriptional regulatory function.

MATERIALS AND METHODS

CtBP1 and 2 genes and protein expression

The full-length human CtBP1 and CtBP2 genes were generously provided by Dr. Steven Grossman. Residues 31-364 of CtBP2 (UniProtKB/Swiss-Prot accession: P56545.1) were cloned in the pET28a vector (Novagen) using the BamHI and XhoI restriction sites. Residues 28-353 of CtBP1 (UniProtKB/Swiss-Prot accession Q13363.2) were cloned into the pET28a vector using the NdeI and SacI restriction sites. Constructs were transformed into BL21-CodonPlus®(DE3)-RIL competent cells (Stratagene). 25 mL of starter cultures grown overnight at 37°C in LB were utilized to inoculate six 1 liter cultures, which in turn were grown at 37°C in tryptone-phosphate media (2% bacto-tryptone, 1.5% yeast extract, 0.2% dibasic sodium phosphate, 0.1% monopotassium phosphate, 0.8% sodium chloride). Cultures were induced at an OD_{600} of 0.5-0.6 by addition of 0.2 mM IPTG and the temperature was reduced to 30°C. Cells were harvested 4 hours after induction, pelleted and resuspended in 30 mL of buffer containing Complete EDTA-Free protease inhibitors (Roche), 50 mM Tris (pH 7.3), 100 mM sodium chloride, 0.2 mM EDTA. Total volume was divided equally into two tubes and cells were flash frozen in liquid nitrogen and stored at -80°C.

Recombinant CtBP2 purification

Cells were thawed and lysed in a cell disruptor at 80 psi. Lysate was supplemented with 10 mM magnesium chloride and 200 μM calcium chloride and 35 mg DNaseI (per 100 mL). The lysate was then stirred at 4°C for 30 min and the insoluble fraction was pelleted. The soluble fraction was collected and incubated with 5 mL of

nickel beads (Qiagen) with gentle stirring at 4°C for 90 minutes. The lysate and beads were applied to a column, and the lysate allowed to flow through. Beads were washed with 25 mL of wash buffer (50 mM Tris pH 7.3, 300 mM sodium chloride, 0.2 mM EDTA, 1 mM DTT, 40 mM imidazole) followed by 25 mL of wash buffer supplemented with 250 mM sodium pyruvate. Beads were washed with an additional 25 mL of wash buffer. Four 5 mL elutions were performed with wash buffer supplemented with 250 mM imidazole. Elutions 1 and 2 were diluted with an equal amount of dialysis buffer (50 mM Tris pH 7.3, 300 mM sodium chloride, 2 mM DTT, 1 mM EDTA) to prevent precipitation. All washes and elutions were carried out at 4°C. Samples were taken from each elution for analysis with SDS-PAGE. The His_6 affinity tag was cleaved with 15 µL Thrombin (Novagen) diluted in 500 µL Novagen Dilution Buffer that was added to the four pooled elution fractions. Fractions were dialyzed (6-8K MWCO) overnight into 2 liters of dialysis buffer at 4°C. Protein was either flash frozen and stored at -80°C, or concentrated for immediate loading on a Superdex 75 column. Size exclusion was performed in dialysis buffer. Elutions off of the column were analyzed by SDS-PAGE for purity before being pooled and concentrated. Protein was flash frozen and stored at -80°C.

Recombinant CtBP1 purification

The same protocol above was used for CtBP1 purification with the following differences: Protein bound to nickel beads was washed with 25 mL wash buffer, 50 mL wash buffer supplemented with 0.5% Triton-X 100, followed by 50 mL wash buffer supplemented with 2.0 M sodium chloride. 10 mL wash buffer supplemented with 250

mM sodium pyruvate was incubated for 5 minutes on the column at 4°C. Sodium pyruvate incubation was repeated 5 times. Beads were washed with final 25 mL of wash buffer. The His-tag was not cleaved off of the protein after elution. Eluted protein was dialyzed into dialysis buffer supplemented with 10% glycerol. Recombinant protein was then concentrated, flash frozen and stored at -80°C.

Crystallization of ternary CtBP2/NAD^+/MTOB complex

Protein (20-25 mg/mL) was incubated with a 25 or 50 molar excess of MTOB, and with or without an additional 10 molar excess of NADH. Protein, NADH, and MTOB were incubated overnight at 4°C. Concentrated protein was mixed in a 2:1 ratio with buffer and hanging drop vapor diffusion trays were setup in 24 well VDX plates. Crystal screening efforts resulted in diffraction quality crystals grown in buffers with combinations of different salts (potassium thiocyanate, potassium iodide, potassium nitrate) and PEGs (3350 and 8000). The highest quality diffraction was observed with crystals incubated with a 50 molar excess of MTOB and no supplemental NADH in a buffer containing 200 mM potassium nitrate, 15-20% PEG 3350, and 100 mM bis tris propane pH 7.0. Crystals typically grew as multiple joined plates. Seeding drops resulted in large single plates suitable for diffraction. Seed solutions were made by crushing crystals in 300 μL of mother liquor followed by further 3 fold dilution into mother liquor. Crystals were cryoprotected by submersion in mother liquor supplemented with 20% ethylene glycol for 5-10 s and then flash frozen in liquid nitrogen.

Crystallization of *CtBP1/NAD$^+$/MTOB complex*

Protein (~10 mg/mL) was mixed with a 25-50 molar excess of MTOB immediately before hanging drops were setup. Bipyramidal crystals grew overnight at room temperature in 200-300 mM magnesium chloride, 0-140 mM sodium formate, 100 mM Hepes pH 7.5, and 2.5 mM NAD$^+$. Crystals were cryoprotected by adding well buffer solution supplemented with increasing amounts of glycerol. Once the drop containing the crystal reached 20% v/v glycerol, the crystal was moved to 30% glycerol for 5 seconds and flash frozen in liquid nitrogen.

Data collection and structure solution

Diffraction data for the CtBP1/NAD+/MTOB and CtBP2/NAD+/MTOB complexes were collected on the BioCARS 14-BM-C beamline at the Advanced Photon Source of Argonne National Laboratory. The CCP4 suite was utilized for solving the CtBP1/NAD+/MTOB ternary complex (*183*). The initial model was solved by molecular replacement with PhaserMR (*184*). The CtBP1 binary complex (pdb code 1MX3) (*10*) and individual domains of a monomer in the CtBP2 binary complex (*90*) were used as search models for CtBP1 and CtBP2 respectively. Waters were automatically generated by Arp/Warp (*185*), and the structure was refined with RefMac5 (*186*). All other structures were refined through PHENIX (*187*). Model building between rounds of refinement was performed with Coot (*188*).

Hydrogen bond determination

Liganded CtBP structures were imported to the Schrödinger software suite. The Protein Prepwizard (*189, 190*) was utilized to add hydrogens, followed by minimization

with heavy atoms restrained. Hydrogen bonds were then determined using a maximum 2.5 Å hydrogen to acceptor (H-A) distance, 120° minimum donor-hydrogen-acceptor (D-H-A) angle, and 90° hydrogen-acceptor-acceptor antecedent (H-A-AA) angle. Hydrogen bond figures were generated in pymol (*191*).

Distance difference plot matrix

Distance difference plot matrices were generated as performed by Prabu-Jeyabalan et al. (*192*). Briefly, a distance matrix was created for the distance between every Cα pair with each structure. One matrix is then subtracted from the other generating a distance difference matrix. The resulting changes in Cα pair distances are displayed as a heat map. A crystallographic symmetry mate was generated for the CtBP1 asymmetric unit monomer to generate the physiological dimer. The CtBP2 asymmetric unit contains four dimers, thus generating symmetry mates was unnecessary.

van der Waals analysis

Van der Waals contact data were computed using a simplified Lennard-Jones potential as previously described (*193*). The Lennard-Jones potential, V_r, was calculated using the interatomic distance (r), depth of the well potential (ε), and the collision diameter (σ) with the following equation:

$$V_r = 4\varepsilon \left[(\sigma/r)^{12} - (\sigma/r)^{6} \right]$$

Energies were calculated for each ligand/protein atom pair with $r \leq 5$ Å, and potentials for non-bonded atom pairs separated by less than the distance at the minimum potential were equated to $-\varepsilon$.

RESULTS

CtBP1 and CtBP2 MTOB co-crystal structures

CtBP crystal structure determination

Both human CtBP1 (28-353) and CtBP2 (33-364) minimal dehydrogenase domains were each crystallized in ternary complexes with coenzyme NAD^+ and ligand MTOB. Molecular replacement utilizing the monomer from the CtBP1 binary structure (CtBP1/NAD^+, pdb code 1MX3) (*10*) as a search molecule provided the initial CtBP1 ternary complex phases. As with the CtBP1 binary complex, the ternary complex consists of a single monomer per asymmetric unit with the molecular dyad of the physiological dimer coincident with a two fold crystallographic axis in the $P6_422$ space group. The CtBP2 ternary complex molecular replacement employed separate substrate binding and coenzyme binding domains of a CtBP2 binary complex (CtBP2/NAD^+, pdb code 2OME) (*90*) as search molecules. The CtBP2 asymmetric unit contains eight monomers, arranged as four dimeric pairs **(Figure 3.2A)**. Molecular replacement located a solution for all eight coenzyme binding domains, but only six of eight substrate binding domains. The two remaining substrate binding domains were built manually during refinement. For full crystallographic statistics, see **Table 3.1**.

The CtBP binary and ternary complexes have similar conformations

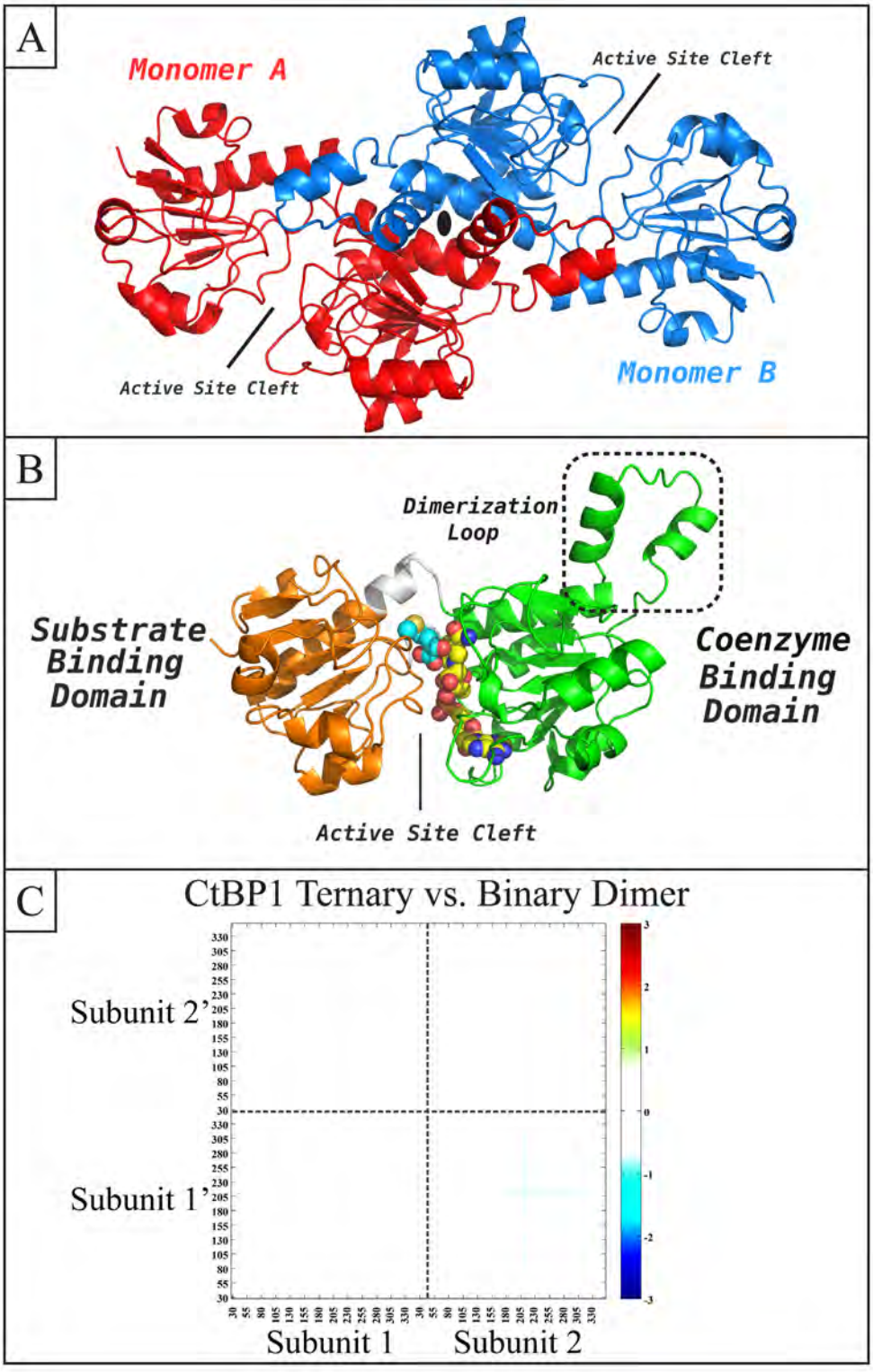
The overall CtBP/NAD^+/MTOB ternary structures closely resemble the previously published binary complexes with bound coenzyme. A discontinuous substrate binding domain (28-120 and 327-353) and coenzyme binding domain (125-319) form an active site cleft where both NAD^+ and MTOB bind **(Figure 3.2B)**. The effects of MTOB

Table 3.1

	CtBP1-MTOB	CtBP2-MTOB
Wavelength (Å)	0.99	0.99
Resolution range (Å)	29.92 - 2.38 (2.47 - 2.38)	32.30 - 2.86 (2.96 - 2.86)
Space group	$P6_422$	$P2_1$
Unit cell a b c (Å)	88.97 88.97 161.54	86.15 140.61 135.13
α β γ (°)	90 90 120	90 97.87 90
Total reflections	167249	298092
Unique reflections	15538	73354
Multiplicity	10.7 (11.6)	4.1 (4.0)
Completeness (%)	97.9 (99.9)	99.3 (96.9)
Mean I/sigma(I)	22.7 (7.1)	9.6 (3.3)
Wilson B-factor	52.8	54.3
R-sym (%)	6.5 (39.0)	9.3 (38.9)
R-factor (%)	20.0 (27.7)	21.4 (30.1)
R-free (%)	23.9 (38.9)	25.1 (33.3)
Number of atoms	2644	20058
Macromolecules	2463	19429
Ligands	97	424
Waters	84	205
Protein residues	327	2646
RMS (bonds)	0.009	0.009
RMS (angles)	1.27	1.15
Ramachandran favored (%)	96	96
Ramachandran outliers (%)	0	0
Clashscore	3.79	11.14
Average B-factor	55.5	52.7
Macromolecules	55.9	52.9
Solvent	53.5	44.3

Figure 3.2: Architecture of the CtBP ternary complexes. (A) Twofold symmetric CtBP1 dimer with one monomer in blue and one in red (black oval denotes crystallographic 2-fold operating relating to the subunits). Each monomer possesses an active site cleft where coenzyme and substrate bind. (B) CtBP1 monomer with the substrate (orange) and coenzyme binding domains (green) connected by a small hinge region (grey). MTOB (cyan) and NAD^+ (yellow) bind the CtBP1 active site cleft. A dimerization loop forms contacts with the other subunit of the dimeric complex. (C) A distance difference matrix plot (DDMP) compares the binary CtBP1 dimer to the ternary dimer. A distance matrix is created for each structure by calculating the distance between every Cα pair. The matrices are subtracted from one another, generating a distance difference matrix comparing the structures in which Cα pairwise differences are reported through a heat map (Å). For clarity, dashed lines mark the boundary between dimeric subunits. The lower left quadrant compares Subunit 1 and 1', and the upper right quadrant compares Subunit 2 and 2'. The upper left quadrant (Subunit 1' and 2) and lower right quadrant (Subunit 1 and 2') reveal any differences in dimeric subunit positions. The CtBP1 monomers show no tertiary or quaternary variations upon MTOB binding.

Figure 3.2


A
Monomer A
Active Site Cleft
Active Site Cleft
Monomer B
B
Dimerization Loop
Substrate Binding Domain
Coenzyme Binding Domain
Active Site Cleft
C
CtBP1 Ternary vs. Binary Dimer
Subunit 2'
Subunit 1'
Subunit 1
Subunit 2

binding on enzyme structure were investigated by a distance difference matrix plot comparing binary and ternary complex dimers **(Figure 3.2C)**. Only minor differences in Cα positions between dimeric CtBP1 binary (pdb code 1MX3) (*10*) and ternary complexes are apparent indicating no significant ligand-induced conformational changes in the overall tertiary or quaternary structure.

Analysis of the eight crystallographically unique CtBP2 monomers indicates that MTOB also does not induce any large conformational changes upon binding to CtBP2 **(Appendix, Figure A.1, A.3)**. As distance difference matrix plots for the eight crystallographically unique CtBP2 binary monomers (pdb 2OME) (*90*) display almost no Cα positional differences **(Appendix, Figure A.2)**, probably due to strong non-crystallographic (NCS) restraints, our comparison between the binary and ternary complex employed only a single CtBP2 binary complex monomer. Differences observed between binary and ternary dimers **(Appendix Figure A.3)** closely correlate with differences observed between the crystallographically unique subunits of the ternary complex, which occur primarily at lattice interactions (Appendix A.4). The subtle, inconsistent differences appear driven by asymmetry among the ternary complex monomers. The minor differences in MTOB bound versus MTOB free CtBP1 and CtBP2 are reflected in the rms deviations, 0.2-0.3 Å for monomers and 0.2-0.4 Å for dimers, suggesting the differences are purely crystallographic.

Consistent with the very high degree of homology (89% sequence identity between the CtBP1 and CtBP2 constructs, **Figure 3.1**), limited differences exist between MTOB bound CtBP1 and CtBP2 monomers. Not surprisingly, the small differences

observed occur at sequence variations, particularly those that impact lattice contacts. These include CtBP2 Gln214 (CtBP1 Ser208), CtBP2 Arg354 (CtBP1 Lys348), CtBP2 residues 48 and 54-55 **(Appendix Figure A.5A, B)**. An additional region of small differences between CtBP1 and CtBP2 results from the variability of the dimerization loop in CtBP2 monomers **(Appendix Figure A.5C)**. Overall, no large consistent differences occur between the CtBP1 and CtBP2 complexes, allowing for direct comparison of MTOB binding between the structures.

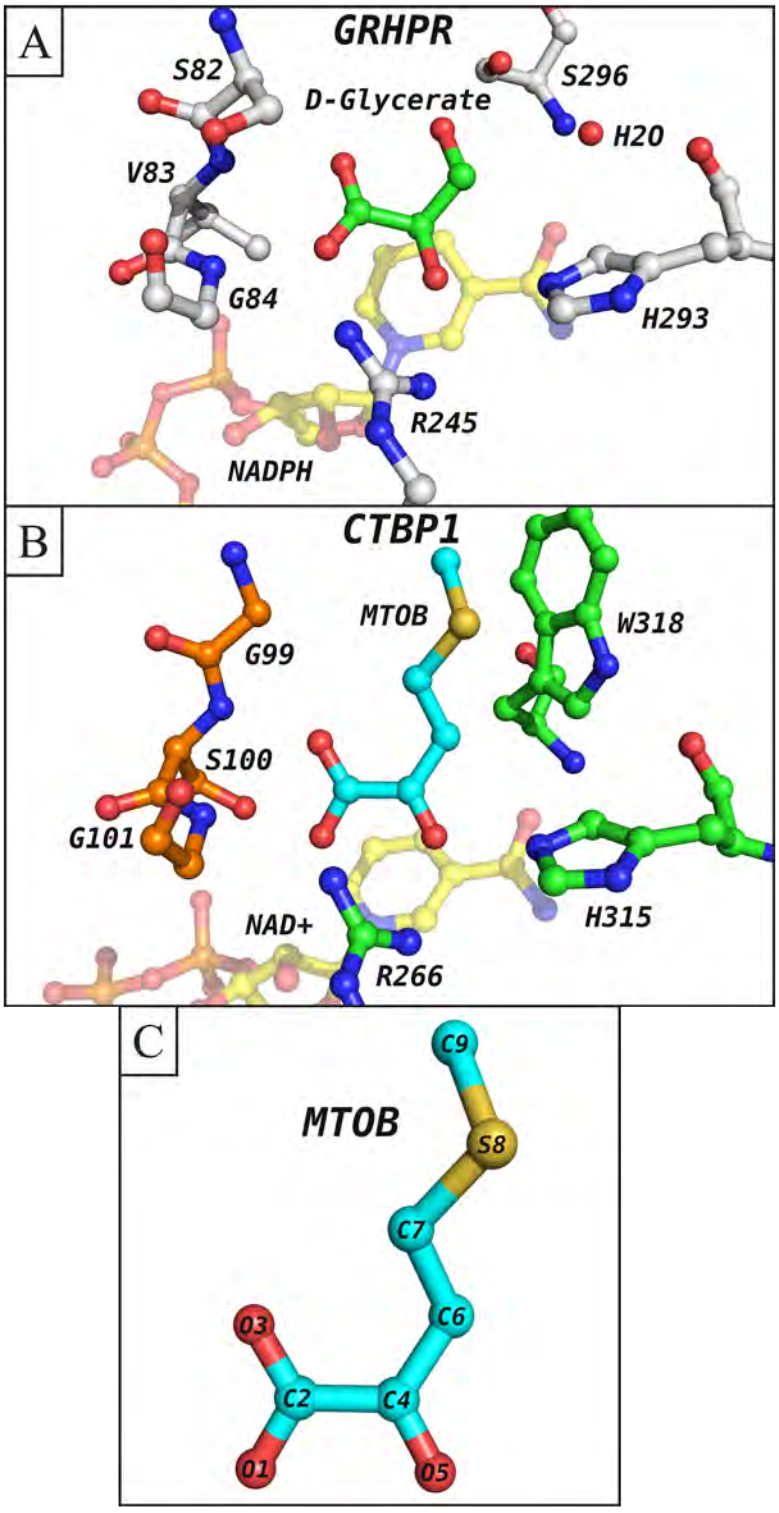
The above comparisons show that MTOB binding does not substantially alter the overall CtBP conformation. Additionally, the MTOB bound CtBP1 and CtBP2 structures are also nearly identical to one another, containing only small variations explained by crystal contacts, sequence variation, and backbone flexibility unrelated to MTOB binding. The lack of large structural changes upon binding argues against ligand-linked conformational changes driving transcriptional modulation by MTOB.

MTOB placement

MTOB is positioned in the active site similarly to ligands in other members of the D2-HDH protein family of proteins **(Figure 3.3A)** (*20, 21, 24*). Both the NAD^+ coenzyme and MTOB bind in the cleft between the substrate binding domain and coenzyme binding domain **(Figure 3.3B)**. MTOB binds between an active site loop from the substrate binding domain and catalytically conserved residues Arg and His of the coenzyme binding domain, while simultaneously positioned adjacent to the NAD^+ nicotinamide ring **(Figure 3.7B)**. The MTOB atom-labeling scheme utilized for discussion in this chapter derives from the PDB atom names **(Figure 3.3C)**.

Figure 3.3: The binding location of MTOB relative to other D2-HDH substrates. (A) Product D-glycerate (green) positioned in the active site of D2-HDH family member Glyoxylate Reductase / Hydroxypyruvate Reductase (GRHPR) (*24*). D-glycerate rests between catalytic residues Arg245 and His293 from the coenzyme binding domain and active site loop residues Val83 and Gly84 from the substrate binding domain. Two final contacts, from Ser296 and a crystallographic water (red), stabilize the ligand. (B) MTOB (cyan) positioned in the CtBP1 active site analogous to D-glycerate. MTOB sits in the cleft separating the coenzyme binding domain (green) and substrate binding domain (orange), also contacting the NAD^+ nicotinamide ring (yellow). Residues involved in MTOB binding from the coenzyme binding domain include catalytic residues Arg266 and His315. Active site loop residues Ser100 and Gly101 from the substrate binding domain stabilize the MTOB carboxylate. (C) MTOB atom labels, derived from the PDB, utilized in descriptions of MTOB atom positioning in both CtBP1 and CtBP2.

Figure 3.3


A
GRHPR
S82
S296
D-Glycerate
H20
V83
G84
H293
R245
NADPH
B
CTBP1
MTOB
W318
G99
S100
G101
H315
NAD+
R266
C
MTOB
C9
S8
C7
C6
O3
C2
C4
O1
O5

CtBP1 MTOB placement

Given the medium to low resolution of the crystallographic analysis (CtBP1 2.38 Å, CtBP2 2.86 Å), particular care was taken in modeling MTOB. To minimize bias, water molecules were first refined into the active site F_{obs}-F_{calc} (F_o-F_c) density for initial refinement. MTOB was then positioned into the active site utilizing the $2F_{obs}$-F_{calc} ($2F_o$-F_c) density and subjected to further rounds of refinement. The electron density is highest at the MTOB sulfur (S8) and along the C2-C4 bond in the α-keto acid moiety. The MTOB O3 and C9 atoms did not fit fully into the density, when contoured at 1σ **(Figure 3.4A)**. Positions of these atoms were chosen based on the F_o-F_c map and refined B-factor analysis. The F_o-F_c map indicated C9 likely occupies two different positions; one oriented away from the Arg97 guanidinium, and the other toward Arg97 **(Figure 3.4B-D)**. Water molecules in the binary CtBP1 structure occupy these positions and are displaced by MTOB **(Figure 3.5A)**. Given the moderate resolution of the analysis, only one C9 conformation of MTOB was chosen to satisfy the F_o-F_c map, retain optimal B-Factors, and avoid steric clashes, particularly with Arg97. The less favored position, which is predominant in CtBP2, appears to be disfavored by steric clashes in CtBP1. The MTOB O3 was positioned to minimize its B-factor and avoid negative F_o-F_c density, finalizing the ligand position in the active site.

The F_o-F_c map, contoured to 3σ, and refined B-factors indicate that MTOB does not fully occupy all the active sites in the CtBP1 crystal. MTOB occupancy was set at 67% based on the F_o-F_c map and local atomic B-factors. F_o-F_c density suggests another molecule may co-occupy a portion of the active site adjacent to the MTOB C7 atom

Figure 3.4: MTOB placement and conformations in the CtBP1 active site.
(A) MTOB positioned in the $2F_o$-F_c density (blue) contoured at 1σ (0.20 $e^-/Å^3$). MTOB atoms O3 and C9 protrude from the density. (B) The F_o-F_c density map (green), contoured to 3σ (0.10 $e^-/Å^3$), indicates that multiple MTOB conformations may exist around atoms C7 to C9. (C) A possible second conformation of MTOB is modeled (orange) in which the C9' atom (alternate C9 position) occupies the positive F_o-F_c density from refinement in the original C9 position. (D) The F_o-F_c map of the original MTOB conformation is replaced with a F_o-F_c map (0.10 $e^-/Å^3$) of MTOB refined in the alternate conformation. The $2F_o$-F_c map (not shown) is identical in the alternate conformation. Positive density now appears in the original C9 atom position, confirming MTOB can exist in both conformations. Poor B-factors and steric clashing with the Arg97 guanidinium prevent C9' modeling. Positive density still remains around the MTOB C7 atom, suggesting mutually exclusive occupancy of the active site by another molecule.

Figure 3.4

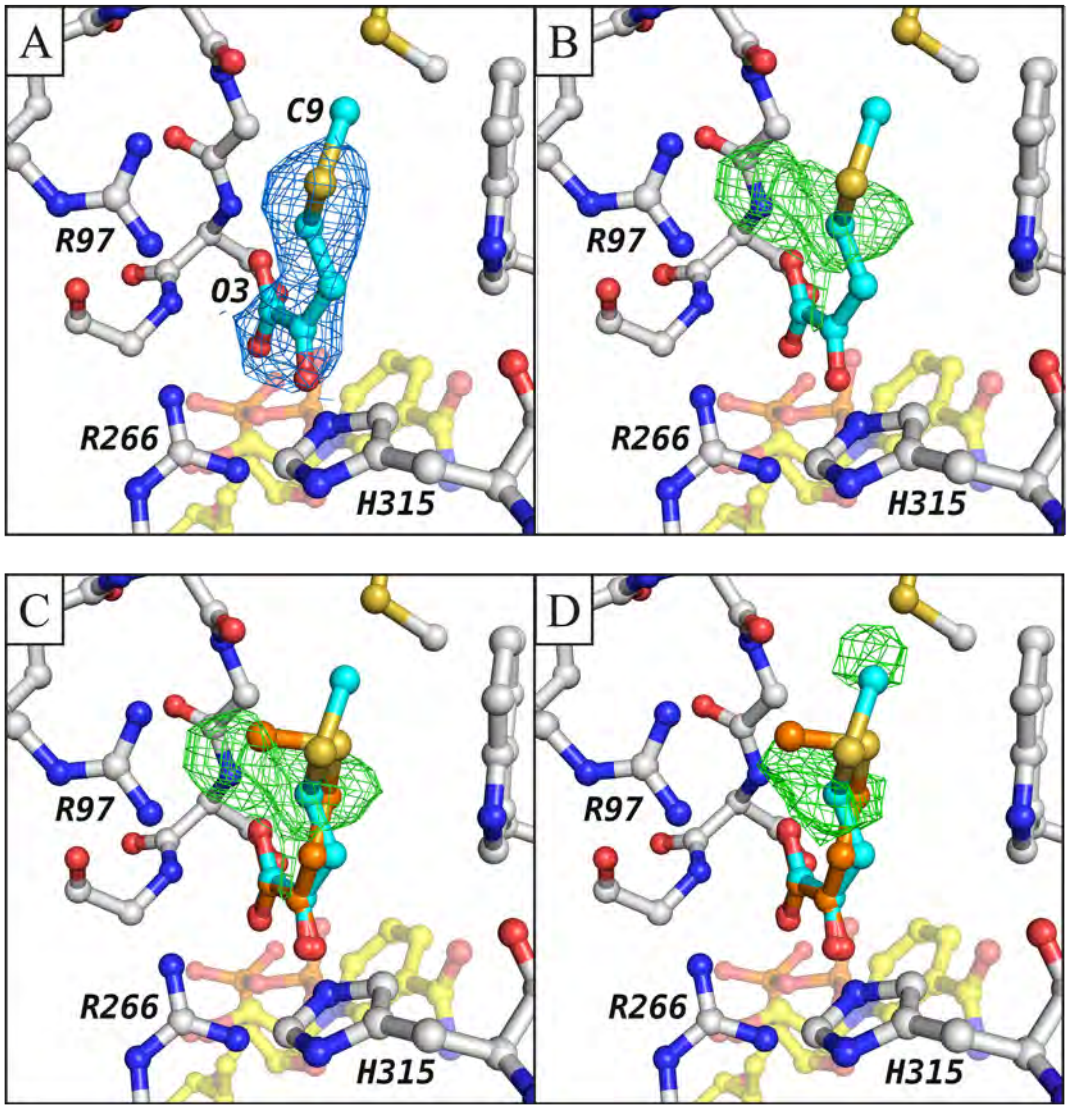
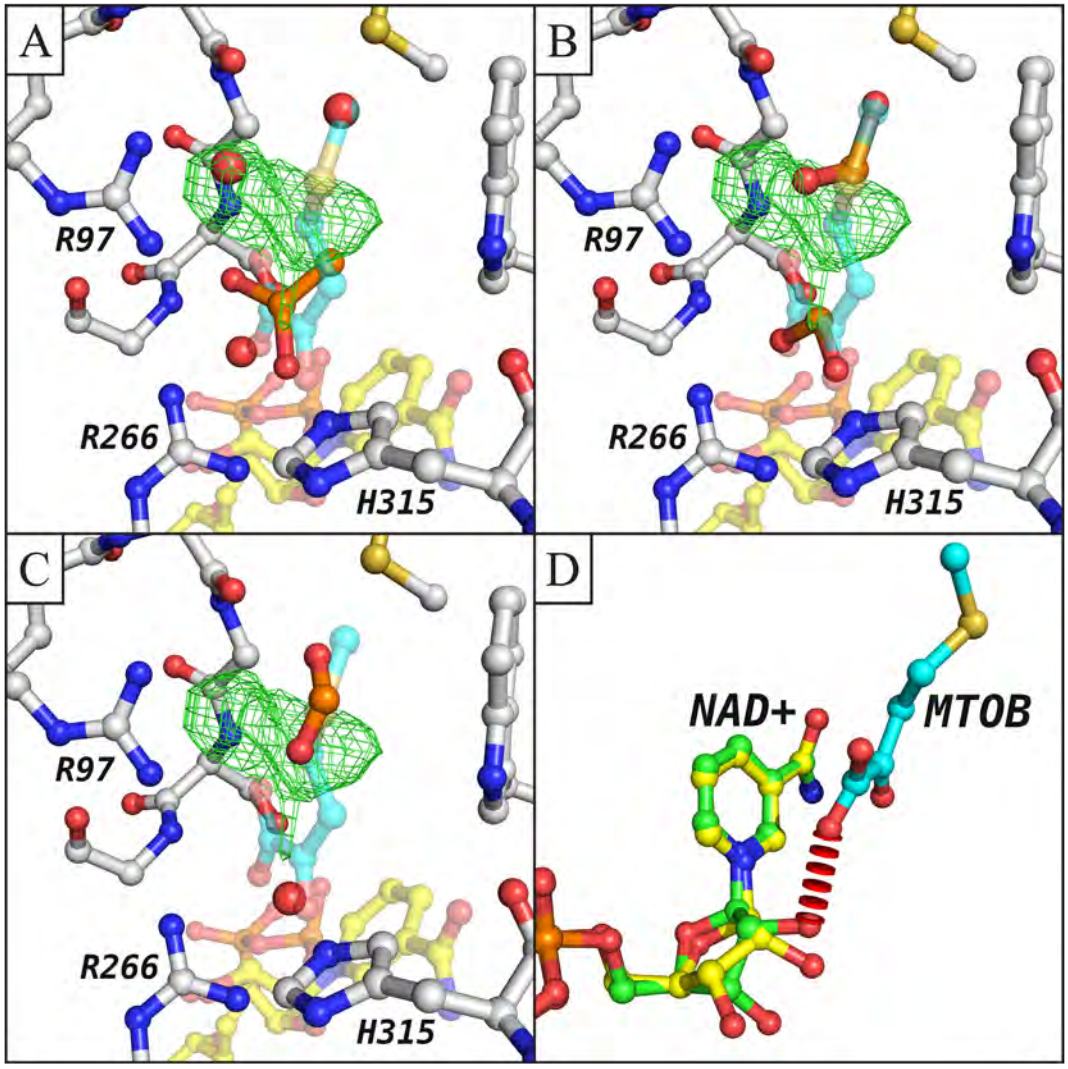

A
C9
R97
O3
R266
H315
B
R97
R266
H315
C
R97
R266
H315
D
R97
R266
H315

Figure 3.5: **Occupants of the CtBP1 active site.** The active sites of other CtBP1 structures were aligned with CtBP1 ternary complex. The MTOB structure F_o-F_c density (green) is included. (A) The CtBP1 binary complex active site contains an acetate ion (orange) and three water molecules (red). Two waters overlap with the MTOB C9 and C9' alternate positions. The third water rests adjacent to the MTOB carboxylate. Two rat CtBP1 structures, (B) (1HKU) (*39*) and (C) (2HU2) (*44*) contain formate ions (orange) that partially occupy the positive F_o-F_c density. Employed for CtBP1 ternary complex crystallization, formate may occupy the active site mutually exclusive to MTOB. Low occupancy or ambiguity caused by MTOB C9 motion resulted in failure of formate or water co-occupancy modeling. (D) NAD^+ modeled in two distinct conformations. The lower occupancy conformation (33%; green) has a large steric clash (red disks) with the MTOB carboxylate. The higher occupancy conformation (67%; yellow) does not have a severe clash with MTOB.

Figure 3.5



(Figure 3.4C,D). In addition to waters, binary structures of CtBP1/NAD(H) (pdb code 1MX3) (*10*) and rat CtB1/NAD(H) (pdb codes 1HKU and 2HU2) (*39, 44*) have acetate or formate ions in the active site (Figure 3.5A-C). The CtBP1 ternary structure F_o-F_c map suggests a formate ion, employed in crystallization, or alternatively a water molecule, exhibits mutually exclusive occupation of the active site adjacent to Arg97 and MTOB atoms C7-C9. Attempts to model these atoms failed due to low occupancy and ambiguity caused by the adjacent MTOB density and movement of the MTOB C9 atom. NAD^+ also exists in two distinct conformations **(Figure 3.5D)**. In one form, the nicotinamide ribose ring sterically clashes with the MTOB carboxylate. This conformer, whose occupancy was refined at 33%, is apparently predominant in the absence of MTOB. The second NAD^+ conformation, refined at 67%, does not clash with the carboxylate. Both NAD^+ conformations appear similar to existing NAD^+ conformations previously observed in other related structures (*10, 21, 25, 39*).

CTBP2 MTOB placement

MTOB was positioned in the CtBP2 active site similarly to CtBP1. Density from the $2F_o$-F_c map and an F_o-F_c omit map guided ligand placement. The quality of the $2F_o$-F_c density varies across the eight monomers, ranging from excellent to poor. Even in the poorest quality monomers, density was still present for the MTOB sulfur atom (S8) and at least some of the α-keto acid moiety. Omit F_o-F_c density maps confirm correct MTOB placement, yielding high confidence of MTOB conformation in monomers A, C and H. An NCS averaged $2F_o$-F_c map (average 2Fo-Fc density from all monomers in the asymmetric unit) however, indicates the catalytic Arg272 guanidinium group of

monomer H occupies a different position than the other monomers **(Figure 3.6A)**. Thus monomers A and C, where ligand and peptide fit best in all the considered maps, were utilized to ensure valid MTOB binding analysis **(Figure 3.6 B, C)**. See appendix for the fit of MTOB in the remaining six monomers **(Appendix Figure A.6, 7)**. As in the CtBP1 ternary structure, MTOB does not fully occupy the CtBP2 active sites. The occupancies of MTOB were refined in PHENIX to 79% in monomers A and C.

CtBP-MTOB interactions

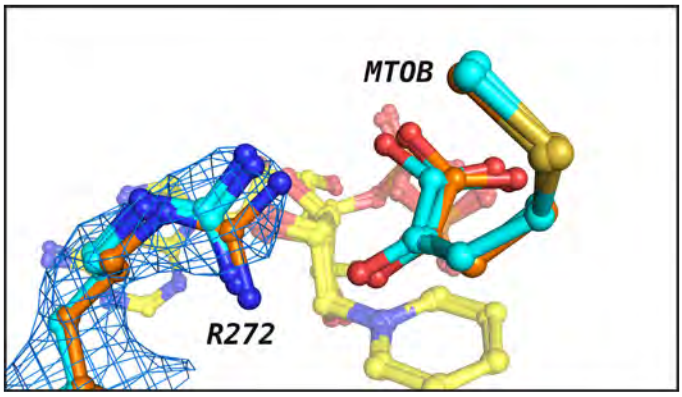
CtBP2 hydrogen bond network

As the network of hydrogen bonds between the MTOB α-keto acid moiety and CtBP2 closely resembles that of other D2-HDHs **(Figure 3.7A)**, we discuss it first, before considering the small differences observed with CtBP1 MTOB binding. Detailed discussion of the hydrogen bond network is restricted to monomers A and C, which show the clearest electron density maps – see above). The catalytically conserved Arg272 and His321 from the coenzyme binding domain provide half of the hydrogen donor atoms. The amide backbone and side chains of a flexible loop in the substrate binding domain provide the other three donors, helping bridge the active site cleft. Hydrogen bonds will be discussed with regards to distance and angle values shown in **Table 3.2**. The active site hydrogen bond network contains six bonds, two for each MTOB oxygen atom **(Figure 3.7B,C)**. Bonds from the catalytic His321 and Arg272 NH1 anchor the MTOB carbonyl. The Arg272 NH2 and the amide backbone of Gly107 provide the bonds for the first MTOB carboxylate oxygen (MTOB O1), while the Ser106 amide and side chain provide both of the bonds to the second MTOB carboxylate oxygen (MTOB O3). The

Figure 3.6: Positioning MTOB in the CtBP2 active site. Monomers A, C, and H have high quality maps for ligand positioning. (A) An NCS average $2F_o$-F_c (0.27 $e^-/Å^3$) map reveals that monomers A and C (cyan) fit well in the averaged density. The Arg272 guanidinium in monomer H (orange) situates outside the average density, leading to omission of monomer H from ligand binding analysis. (B) MTOB occupies the $2F_o$-F_c density at contour levels 1σ (blue, 0.27 $e^-/Å^3$) and 1.6σ (green, 0.43 $e^-/Å^3$), in monomers A and C. The sulfur is clearly visible at the higher contour level. (B) An omit map (F_o-F_c) generated by repeating the final round of refinement without MTOB or waters present, was contoured to 2.4σ (0.19 $e^-/Å^3$; blue mesh) and 3.5σ (0.27 $e^-/Å^3$; green surface). MTOB placements show good agreement with the omit map, including monomers A and C. The maps confirm the MTOB C9 in CtBP2 exists in the alternate conformation observed in CtBP1. The MTOB C9 atom does not clash with Arg97 in CtBP2 due to differences in side chain position.

Figure 3.6

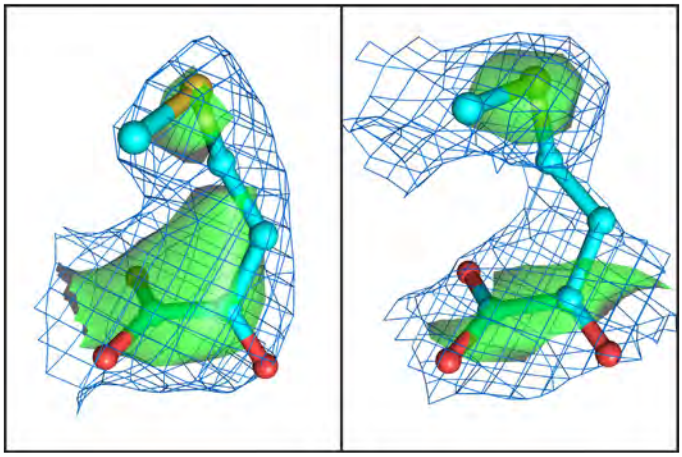
A



$2F_o$-F_c Map

Monomer A Monomer C

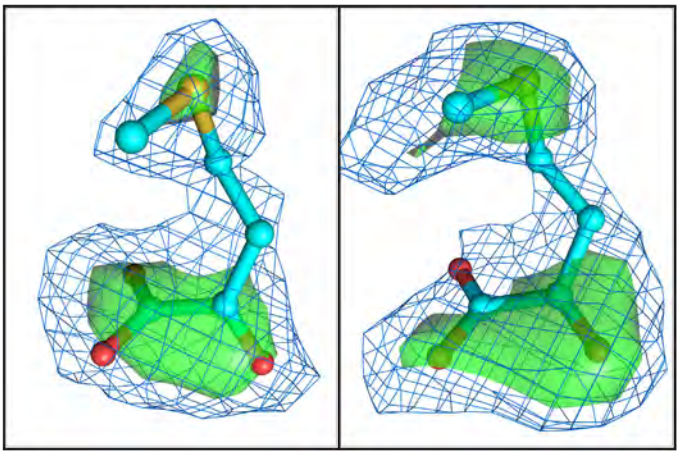
B



Omit F_o-F_c Map

Monomer A Monomer C

C



bond length from the Ser106 side chain, equivalent to the maximum allowable distance, suggests this bond is the weakest of the conserved bonds. The D2-HDH family lacks conservation of a side chain possessing a hydrogen bond donor at this position, suggesting a weak hydrogen bond to Ser106 Oγ may contribute only minimally.

In monomer A, Arg272 NH2 donates a non-conserved hydrogen bond to the MTOB carbonyl **(Figure 3.7B)**. The geometry of this bond indicates it is weak, providing a rationale for its inconsistent appearance in all CtBP2 subunits. The length and angles of the bond exist at the criteria limits in monomer A, and just beyond the limits in monomer C. Therefore, a network of six fully conserved hydrogen bonds, with a potential, weaker seventh bond, anchors MTOB in the active site in CtBP2.

The CtBP1 hydrogen bonding network is distinct from CtBP2

The network of hydrogen bonds in CtBP1 is distinct compared to CtBP2 **(Figure 3.7D)**. The differences arise from an unexpected twist in the MTOB carboxylate and the position of Arg97. The MTOB carboxylate group rotates relative to the plane of the O5 carbonyl, as evidenced by a shift in the MTOB O5 carbonyl – O1 carboxylate dihedral angle relative to CtBP2 monomers A and C by 72° and 44° respectively. The side chain of Arg97 situates further into the active site in the CtBP1 ternary complex than in either the CtBP1 binary complex or any of the MTOB bound CtBP2 monomers **(Figure 3.8)**. This new position allows for formation of a unique hydrogen bond and increased coulombic interaction to the MTOB carboxylate.

Only four of the six hydrogen bonds conserved in the CtBP2 complex are found in the CtBP1 structure. The two strong hydrogen bonds anchoring the MTOB carbonyl

Figure 3.7: **Substrate hydrogen bond network.** (A) The hydrogen bond network of D-glycerate in the GRHPR active site (*24*). Catalytic residues Arg245 and His293 from the coenzyme binding domain, and residues Val83 and Gly84 from the substrate binding domain hydrogen bond to the D-glycerate carboxylate and 2’ hydroxyl group (green dashes). Ser296 and a water molecule (red) donate two final hydrogen bonds to the 3’ hydroxyl. (B) The MTOB hydrogen bond network of CtBP2 Monomer A. Six conserved hydrogen bonds in Monomer A (orange dashes) and one Monomer A specific bond (yellow dashes) comprise the MTOB hydrogen bond network. (C) The hydrogen bond network of Monomer C contains only conserved bonds. (D) The network of hydrogen bonds in CtBP1 is distinct from CtBP2. MTOB’s conformation in CtBP1 allows for formation of only four of the six conserved bonds (orange dashes). Arg97 forms a unique hydrogen bond with MTOB in CtBP1 (blue dashes).

Figure 3.7

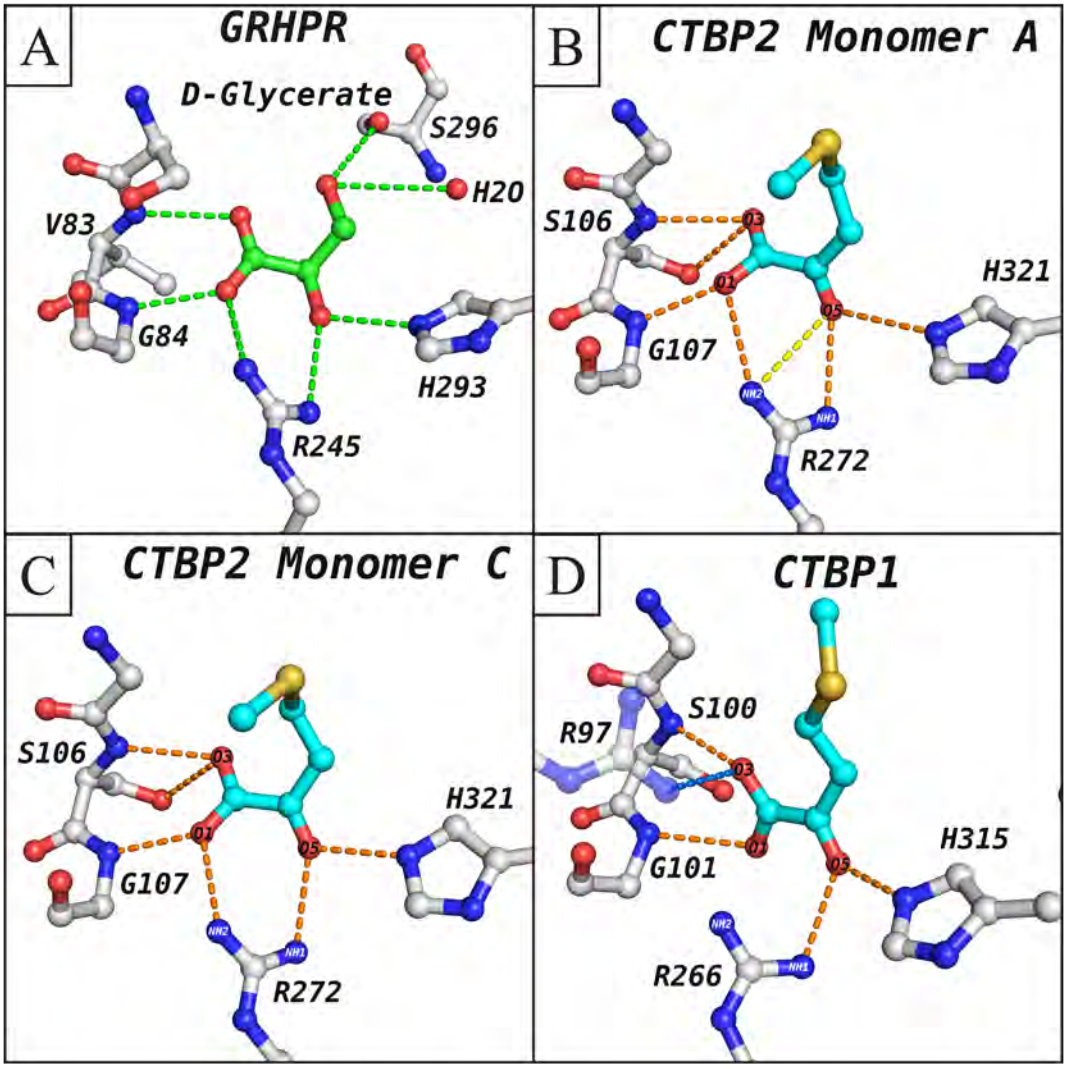

A
GRHPR
D-Glycerate
S296
H20
V83
G84
H293
R245
B
CTBP2 Monomer A
S106
O3
O1
O5
H321
G107
NH2
NH1
R272
C
CTBP2 Monomer C
S106
O3
O1
O5
H321
G107
NH2
NH1
R272
D
CTBP1
S100
R97
O3
O1
O5
H315
G101
NH2
NH1
R266

Table 3.2

CtBP - MTOB Hydrogen Bond Networks

Protein (Monomer)	Residue (Atom)	MTOB Atom	D-A Distance (Å)	H-A Distance (Å)	D-H-A Angle (°)	H-A-AA Angle (°)
CtBP2 (Monomer A)	H321 (NE1)	O5	3.0	2.0	170	112
	R272 (NH1)	O5	2.8	1.8	157	156
	R272 (NH2)	O1	3.0	2.1	151	108
	G107 (N)	O1	2.6	1.6	170	146
	S106 (N)	O3	2.8	1.8	169	97
	S106 (OG)	O3	3.2	2.3	160	106
	R272 (NH2)*	O5	3.2	2.4	132	93
CtBP2 (Monomer C)	H321 (NE1)	O5	2.7	1.7	158	108
	R272 (NH1)	O5	2.7	1.7	158	149
	R272 (NH2)	O1	2.7	1.8	158	118
	G107 (N)	O1	2.6	1.6	154	130
	S106 (N)	O3	2.8	1.8	160	119
	S106 (OG)	O3	3.3	2.5	145	110
CtBP1 (Monomer A)	H315 (NE1)	O5	2.6	1.6	177	122
	R266 (NH1)	O5	2.7	1.7	159	139
	G101 (N)	O1	2.8	2.0	143	93
	S100 (N)	O3	2.7	1.9	134	132
	R97 (NH1)**	O3	3.2	2.5	125	100

Hydrogen Bonds are defined as having a Hydrogen-Acceptor (H-A) distance < 2.5Å, Donor-Hydrogen-Acceptor (D-H-A) angle >120°, and Hydrogen-Acceptor-Acceptor Antecedent (H-A-AA) angle >90°.

*Bond is unique to CtBP2 Monomer A complex
**Bond is unique to CtBP1 complex

are present in the CtBP1 structure. The bonds from His315 and Arg266 NH1 are the most ideal hydrogen bonds of the CtBP1 network. Despite its rotation relative to the CtBP2 structure, the MTOB carboxylate still forms the conserved hydrogen bonds with the amide backbone of S100/106 and Gly101/107, albeit with less ideal bond geometry, suggesting these bonds are weaker in the CtBP1 complex. A slight improvement in geometry would form a CtBP1 unique hydrogen bond between Gly101 and alternate carboxylate oxygen, MTOB O3. However, as it is modeled, the geometry lies just outside the bond criteria. The CtBP1 complex lacks the conserved bonds donated from the CtBP2 Ser106 side chain and Arg272 NH2 to the MTOB O3 and O1 respectively.

The mobility of the Arg97 side chain evident from human and rat CtBP1 structures, permits formation of a hydrogen bond with CtBP1 MTOB not observed in the CtBP2 complex. In each CtBP1 structure, the Arg97 guanidinium group situates to interact with a molecule bound in the active site. This includes formate or acetate in the human and rat CtBP1 structures, and MTOB in the ternary structure **(Figure 3.8A,B)**. The same mobility is not observed in the CtBP2 structures **(Figure 3.8C,D)**. In the CtBP1 ternary structure, Arg97 protrusion into the active site allows the formation of a weak hydrogen bond between a guanidinium nitrogen and MTOB carboxylate oxygen O3. The bond distance and DHA angle exist at the limits of hydrogen bonding criteria, suggesting the hydrogen bond is weak.

The conformation of MTOB in CTBP1 results in a distinct, albeit weaker, hydrogen bond network. Only four of six hydrogen bonds conserved in CtBP2 also exist in CtBP1. Gaining the weak hydrogen bond between CtBP1 Arg97 and MTOB O3 does

not account for the loss of the other hydrogen bonds or for the reduction of quality in bond geometry between the MTOB carboxylate and the amide backbone. However, the movement of Arg97 into the active site cavity, coupled with the rotation of the MTOB carboxylate, suggests an increase in coulombic interaction may offset the energetic penalty incurred by the CtBP1 specific MTOB conformation.

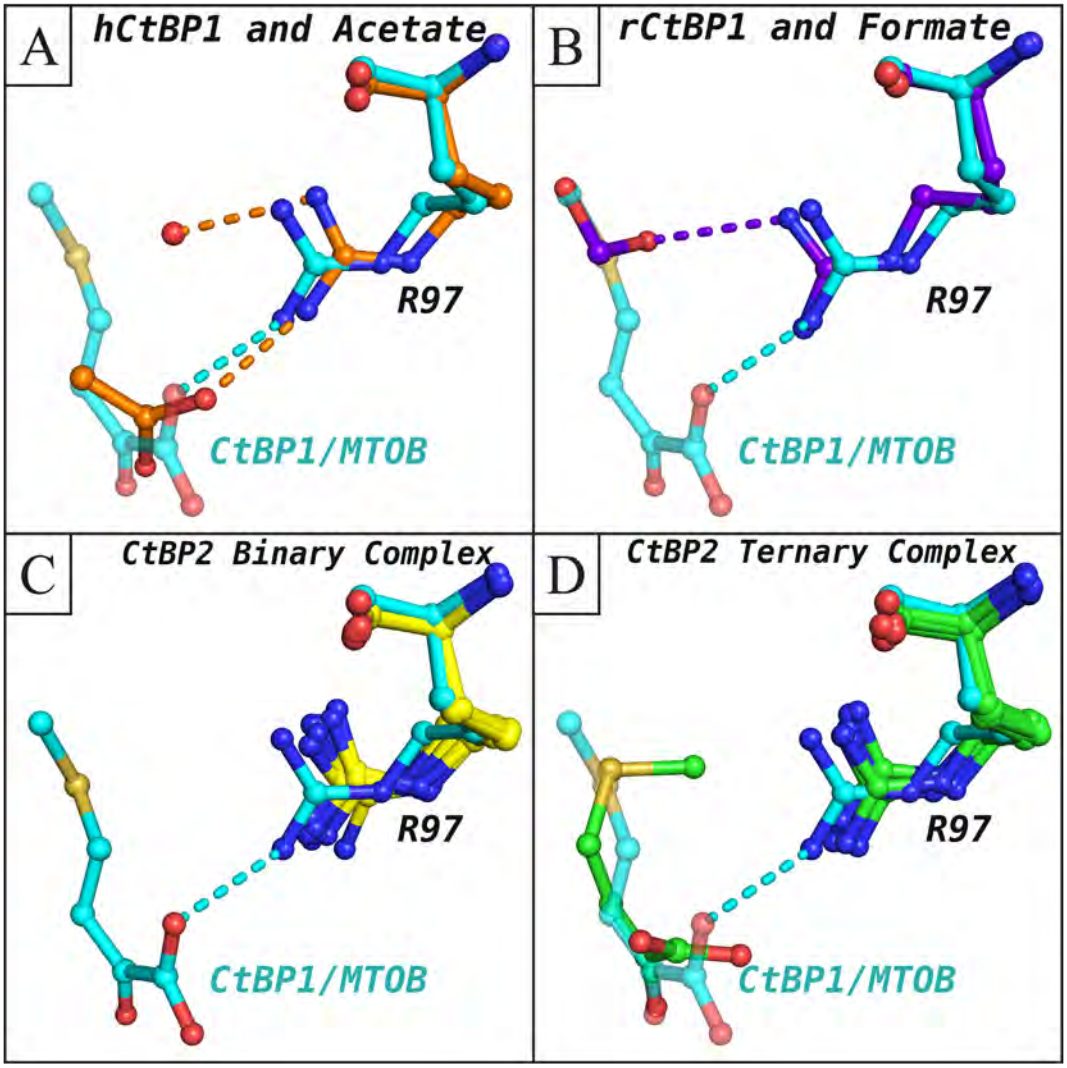
Unique Substrate Contacts and MTOB Specificity

Van der Waals analysis of the CtBP-MTOB complexes specifies details of contacts between MTOB and active site residues, providing insight into substrate specificity. Utilizing a program created by Ozen et al. (*193*), the energetics of each contact was calculated utilizing a Leonard-Jones potential. Only attractive forces were considered in the current analysis to avoid skewing of the binding profile by short hydrogen bonds **(Figure 3.9A)**. For details on the attractive and repulsive forces observed, see **Appendix Figure A.4**.

The MTOB α-keto group contacts CtBP similarly to other liganded D2-DHDs as discussed above. The major contacts include the coenzyme binding domain catalytic Arg266/272 (CtBP1/CtBP2) and His315/321 residues and the substrate binding domain active site loop, Gly99/105, Ser100/106, Gly101/107. CtBP however lacks a contact present in other D2-DHD, such as bacterial D-LDH, human GRHPR, and human PHGDH. The substrate α-keto acid moiety of other D2-HDHs contacts a large residue opposite from the catalytic arginine. The bulky residues originate from one of two equivalent CtBP positions, Ala123/129 and Ser124/Ala130 **(Figure 3.1)**, neither of which

Figure 3.8: The mobility of the CtBP1 Arg97 guanidinium group. In all panels, Arg97 and MTOB in the CtBP1 ternary complex are shown with cyan carbon atoms. (A) Arg97 of the human CtBP1 binary complex (1MX3, orange) (*10*) situates 2.8 Å and 3.1 Å (green dashes) from an acetate ion and water molecule (red sphere), respectively. Arg97 of the CtBP1 ternary complex (cyan) moves further into the active site, positioning 3.2 Å from the MTOB carboxylate. A new CtBP1 unique hydrogen bond is formed (cyan dashes) between the ternary complex Arg97 and MTOB. (B) The rat CtBP1 Arg97 (1HKU, purple) (*39*) pushes even further into the active site, creating a 3.2 Å contact (purple dashes) with a formate ion. The CtBP1 ternary complex R97 situates 4.0 Å from the same formate oxygen. (C, D) All eight CtBP2 binary complex (yellow) and ternary complex (green) Arg103 residues fail to shift into the active site, even when MTOB (monomer A) is present, suggesting the mobility of the arginine is unique to CtBP1.

Figure 3.8


A
hCtBP1 and Acetate
R97
CtBP1/MTOB
B
rCtBP1 and Formate
R97
CtBP1/MTOB
C
CtBP2 Binary Complex
R97
CtBP1/MTOB
D
CtBP2 Ternary Complex
R97
CtBP1/MTOB

contacts MTOB. In bacterial enzymes, this residue is typically a tyrosine, whereas in the human GRHPR and PHGDH, it is a leucine (Leu107) and asparagine (Asn106), respectively.

The MTOB atoms outside the α-keto acid moiety contact residues previously implicated in substrate specificity of other D2-DHDs. CtBP1 residues Tyr76/82, His77/83, and Met327/333, as well as Val159'/165' from the neighboring dimeric subunit, contact the terminal atoms on of MTOB, sterically preventing larger molecules from binding. Tyr76/82 and Met327/333 match residues implicated in substrate specificity and binding in bacterial D2-HDHs (*19*). CtBP unique residue His77/83 situates in an identical position as a leucine directly linked with substrate specificity in human protein GRHPR (*24*). These residues prevent α-keto acids larger than MTOB from binding. Variations exist in the level of contact between the MTOB terminal atoms across the CtBP complexes, due to the difference in C9 position between the CtBP complexes.

CtBP1 Trp318 (CtBP2 Trp324) dominates the active site and is within van der Waals distance of all MTOB atoms outside the α-keto acid moiety **(Figure 3.12A)**. Trp318/324 occupies a homologous position (Figure 1) to a phenylalanine in *L. bulgaricus* D-LDH, tyrosine of *A. aeolicus* D-LDH, and tyrosine of D-HicDH, all residues implicated in substrate specificity (*19-21*). Human PHGDH lacks an aromatic residue in this position, with an alanine instead permitting additional arginine residues that stabilize the substrate's negatively charged phosphate. Human GRHPR residues Ser296, Arg302, and a crystallographic water molecule occupy a position roughly equivalent to CtBP Trp318/324, serving as specificity determinants that contact the

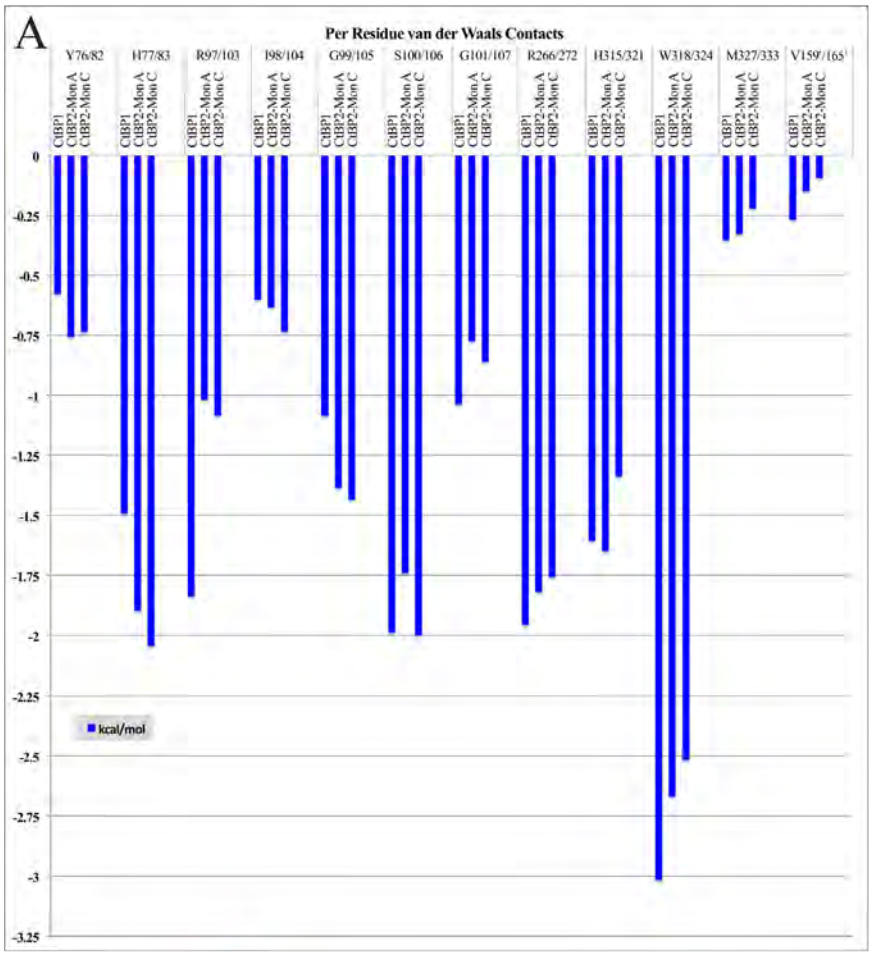
terminal hydroxyl group of the hydroxypyruvate substrate (*24*). The water molecule situates between the Ser296 and Trp141' donated by the neighboring dimeric GRHPR subunit. When aligned to CtBP1, this tryptophan positions between CtBP1 residues Trp318/324 and His77/83, helping to coordinate the GRHPR water, and making contact with the substrate. Alignment the CtBP structures with the GRHPR structure indicates MTOB, or any α-keto acids larger than hydroxypyruvate, would severely clash with the GRHPR Trp141'. Thus, CtBP Trp318/324 is a unique feature among human D2-HDH proteins that provides an extensive binding surface lining the active site.

Interestingly, the MTOB S8 atom centers over the CtBP Trp318/324 indole, rather than orienting its lone pair of electrons towards the positively charged Arg97/103 guanidinium **(Figure 3.9B)**. The positional preference suggests a sulfur-pi interaction commonly observed in the protein data bank between methionine and cysteine residues stacking over aromatic rings, plays a role in substrate specificity (*194-196*). The finding that replacing the MTOB sulfur with a methylene carbon decreased enzymatic efficiency 8 fold (*63*), suggests a sulfur-pi interaction with active site Trp318/324 contributes to substrate specificity of CtBP.

Only a single large magnitude difference exists in MTOB contacts between CtBP1 and CtBP2. CtBP1 Arg97 contacts MTOB more extensively in the CtBP1 complex than the homologous Arg103 in the CtBP2 complex due to the shift of the MTOB carboxylate, as discussed above in relation to hydrogen bonding. Not only does the CtBP1 carboxylate conformation create a unique hydrogen bond, but it also increases

Figure 3.9: MTOB van der Waals interactions. (A) The calculated van der Waals contacts with MTOB in CtBP1, CtBP2 monomer A, and CtBP2 monomer C. The largest magnitude difference between the CtBP1 and CtBP2 occurs with increased contacts of CtBP1 at Arg97/103. CtBP2 monomers possess greater contacts at His77/83, with offsetting lower contacts at Trp318/324 due to MTOB C9 conformational differences between structures. Trp318/324 forms the most extensive contacts with MTOB in all three structures. (B) The active site of CtBP1 (green) and CtBP2 Monomer A (cyan). The shift in CtBP1 MTOB carboxylate, and corresponding movement of Arg97 increase van der Waals contacts relative to CtBP2. Trp318/324 contacts all MTOB atoms outside the α-keto acid moiety in both structures.

Figure 3.9


A
Per Residue van der Waals Contacts
Y76/82
H77/83
R97/103
I98/104
G99/105
S100/106
G101/107
R266/272
H315/321
W318/324
M327/333
V159'/165'
CtBP1
CtBP2-Mon A
CtBP2-Mon C
0
-0.25
-0.5
-0.75
-1
-1.25
-1.5
-1.75
-2
-2.25
-2.5
-2.75
-3
-3.25
kcal/mol
B
W318/324
MTOB
R97/103
R266/272
H315/321

favorable van der Waals interactions and intensifies the coulombic interaction between MTOB and the Arg97 guanidinium **(Figure 3.9B)**. However, in terms of overall impact on binding, the shift of MTOB in CtBP1 does not greatly affect the total van der Waals interaction profile.

The unique hydrophilic cavity of the CtBP active site

A unique aspect of the CtBP active site, not shared by other members of the D2-HDH family, is a hydrophilic cavity between the MTOB binding site and NAD+. Four ordered water molecules fill this cavity in the CtBP1 ternary structure, linking MTOB with the nicotinamide phosphate group of NAD^+. Other D2-HDH family members possess residues with bulky side chains that interact with the substrate, sealing the binding site and preventing formation of the water network. CtBP lacks a bulky residue at this position (see Unique Contacts and Specificity above) allowing the existence of the water network. This unique hydrophilic cavity may be exploited for future small molecule affinity and specificity.

CtBP water network

The active site cavity in CtBP contains a network of waters observed in multiple structures. The network of waters, designated as waters A, B, C and D, fills a cavity that may provide useful, CtBP specific, contacts for future inhibitor development **(Figure 3.10A)**. Two other CtBP1 structures, the human CtBP1 binary complex (pdb code 1MX3) (*10*) and a rat CtBP1 binary complex (pdb code 1HKU) (*39*) display the full network of four water molecules, while a lower resolution rat CtBP1 (pdb code 2HU2) (*44*) retains only waters A and B **(Figure 3.10B)**. Individual monomers in the binary

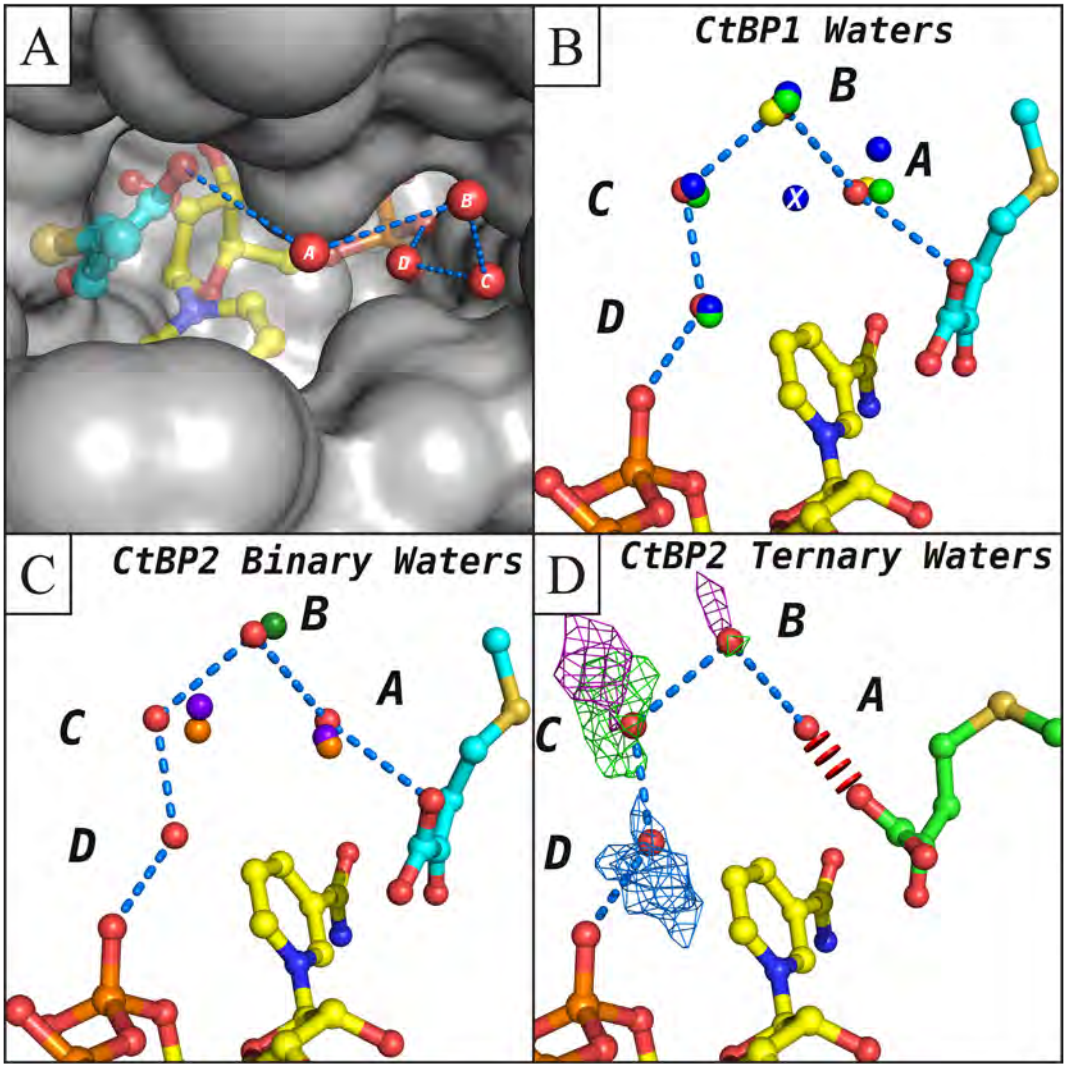
CtBP2 structure (pdb code 2OME) (*90*) contain some of the conserved waters despite the low resolution relative to the CtBP1 structures **(Figure 3.10C)**. Finally, the CtBP2 ternary complex contains no active site waters as attempts to model these waters resulted in low quality $2F_o$-F_c maps. However, the F_o-F_c omit map utilized for MTOB placement indicates positive density occurs at some positions of the water network providing evidence that water molecules bind at these positions **(Figure 3.10D)**.

The positions of water molecules that fill the active site cavity are very similar, with the exception of Water A **(Figure 3.10B,C)**. The precise binding location for water A depends on the occupants of the active site in various structures, with binding of ions (acetate or formate), MTOB, or other water molecules dictating the location of water A binding. Interestingly, the CtBP2 MTOB carboxylate sterically clashes with a water molecule at position A, which accounts for the lack of F_o-F_c density at this position **(Figure 3.10D)**. It appears therefore that the CtBP1 MTOB carboxylate shift towards Arg97 is required for water binding at position A. Water molecules in positions B-D show much less variation in position.

In addition to contacting NAD^+, the water network contacts protein residues that comprise the hinge region between the substrate binding domain and coenzyme binding domain, and a single active site loop residue. The CtBP1 residues involved in binding include 120-125 (VPAASV), and the equivalent CtBP2 residues 126-130 (IPSAAV). Despite the three amino acid variations between CtBP1 and CtBP2 in this loop, only the Ser124/Ala130 difference affects contacts within the water network. The side chains of the other amino acids orient outside of the cavity and do not affect the peptide backbone

Figure 3.10: Comparing other CtBP water networks to the CtBP1 ternary complex. (A) The water filled cavity in the CtBP1 ternary complex. Shown are the protein surface (grey), NAD (yellow), MTOB (CtBP1, cyan), and the water network (red spheres) labeled A to D. The water molecules form a network (blue dashes) between the MTOB carboxylate and an NAD^+ phosphate. (B) The CtBP1 ternary complex, network was compared to human CtBP1 binary complex (1MX3, blue spheres) (*10*), and two rat CtBP1 structures (1HKU, green spheres; 2HU2, yellow sphere) (*39*, *44*). 1MX3 has a fifth water (labeled with "X"), which is highly mobile, or incorrectly modeled (B factor = 97.1$Å^2$). MTOB, bound ions, or other waters dictate water A position across different structures. (C) CtBP2 binary complex (pdb 2OME) (*90*) monomers C (purple), F (orange), and G (dark green) contain partial water networks. (D) MTOB (green) from the CtBP2 ternary complex clashes with the CtBP1 water A position. CtBP2 ternary complex F_o-F_c map (0.24 $e^-/Å^3$) indicates positive density at water positions for monomers A (green mesh), G (purple mesh), and H (blue mesh). The low resolution of the CtBP2 ternary complex does not permit confident modeling of these water molecules.

Figure 3.10


A
B
CtBP1 Waters
B
A
C
X
D
C
CtBP2 Binary Waters
B
A
C
D
D
CtBP2 Ternary Waters
B
A
C
D

contacts. The substrate binding domain active site loop provides the final residue, CtBP Ser100/106, contributing to binding with waters at positions A and D, in addition to MTOB. An NAD^+ phosphate group oxygen hydrogen bonds with the final water in the network at position D.

The cavity exists because CtBP lacks a bulky residue at a key position

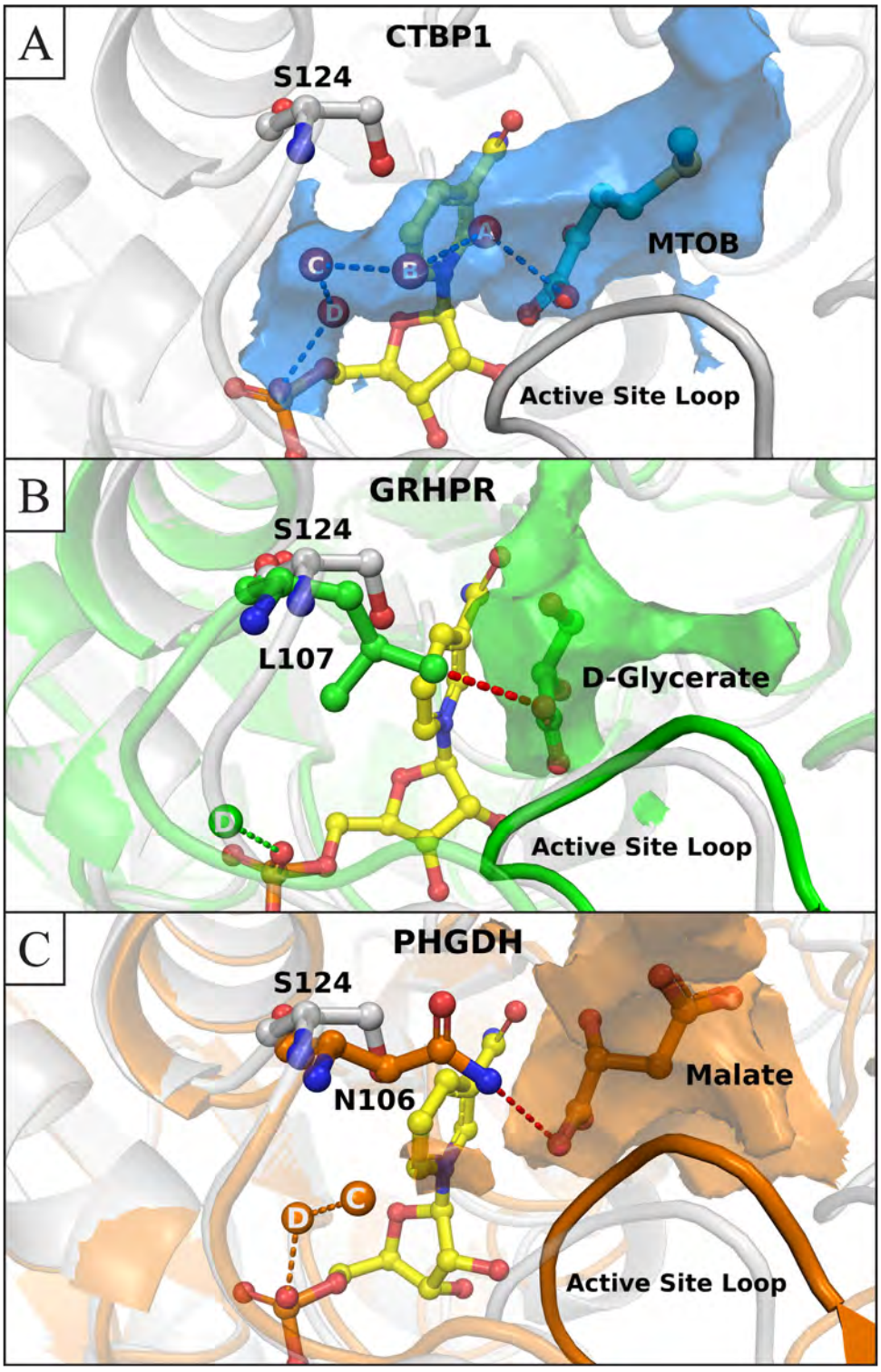
The observed active site cavity is present because CtBP lacks a residue typically involved in substrate binding in other D2-HDHs. In other family members a residue with a large side chain typically occupies the homologous position to CtBP Ala123/129 or Ser124/Ala130 in the hinge region just outside the coenzyme binding domain. These side chains interact with the substrate, as evidenced by the van der Waals analysis, but appear not to contribute to differentiation between different α–keto acids as the residues only contact the α-keto acid moiety (*19*). A tyrosine residue occupies the Ala123/129 position in three bacterial homologs, using its hydroxyl group for hydrogen bonding with the substrate carboxylate of *A. aeolicus* D-LDH and D-HicDH , or the sulfate ion oxygen in the case of *L. bulgaricus* D-LDH (*19-21*). In human D2-HDHs, PHGDH and GRHPR, similar bulky residue contact occurs with substrate. In GRHPR, Leu107 substitutes for CtBP Ser124/Ala130, providing van der Waals contacts to the carboxylate of the enzymatic product, D-Glycerate (*24*). Similarly, in PHGDH Asn106 substitutes for CtBP Ser124/Ala130, creating an interaction between the carboxamide nitrogen and a carboxylate belonging to a bound malate, an ion utilized in crystallization (*25*). The smaller CtBP residues in this position fail to contact MTOB in both the CtBP1 and CtBP2 ternary structures. The CtBP1 Ser124 hydroxyl oxygen and CtBP2 Ala130 methyl

group lay 5.7 and 5.6 Å from the MTOB carboxylate respectively. The lack of interaction between CtBP Ser124/Ala130 with MTOB may contribute to the $2x10^2$ slower k_{cat} of CtBP1 for MTOB compared to rat PHGDH with its endogenous substrate (*63, 67*).

The effect of the bulkier residues of GRHPR and PHGDH is to close the active site around the substrate, whereas the smaller CtBP residues leave a gap, allowing formation of the water network **(Figure 3.11A)**. Both PHGDH Asn106 and GRHPR Leu107 side chain approach the active site loop at the residue equivalent to CtBP Ser100/106, closing the cavity observed in CtBP. When aligning structures based on the well- conserved coenzyme binding domain, the GRHPR Leu107 side chain severely clashes with the water network, displacing waters in positions A-C **(Figure 3.11B)**. Similarly, the PHGDH Asn106 carboxamide nitrogen situates about 1.5 Å from the water A position in CtBP1 **(Figure 3.11C)**. These large residues do not completely destroy the water network; rather they restrict the substrate to the active site volume adjacent to the catalytic residues **(Figure 3.11B,C)**. Remaining waters are unable to form a complete network connecting coenzyme and substrate, instead situating equivalent to water positions C or D.

Figure 3.11: A small CtBP residue side chain allows formation of the water network. **(**A) Active site cavity surface maps of human D2-DHDs were generated in the absence of substrate atoms and aligned to CtBP1 (grey). The CtBP1 cavity surface map (blue surface) encapsulates MTOB (cyan) and the water network (red spheres) in one contiguous volume. CtBP1 hinge residue S124 fails to contact both MTOB and the active site loop, allowing water to fill the gap. (B) Leu107 of human GRHPR (green) closes the gap between the hinge and active site loop, closing off the active site volume, indicated by the termination of the green volume surface map. Leu107 contacts D-glycerate (red dashes) and displaces water positions A-C. One water molecule remains (green sphere) roughly equivalent to CtBP1 water position D. (C) PHGDH Asn101 (orange) also closes the gap between the hinge and active site loop, capping the active site volume (orange surface). The carboxamide nitrogen contacts a malate ion (red dashes), which situates approximately where substrate binds. Asn101 displaces waters at positions A and B, but leaves room for binding of waters at positions C and D adjacent.

Figure 3.11


A
CTBP1
S124
MTOB
Active Site Loop
B
GRHPR
S124
L107
D-Glycerate
Active Site Loop
C
PHGDH
S124
N106
Malate
Active Site Loop

DISCUSSION

Substantial evidence implicates CtBP in the pathogenesis of multiple forms of cancer (*8, 72, 75, 81, 83, 182*). A putative substrate for CtBP's dehydrogenase activity, MTOB, has been shown in two different cancer models to effectively interfere with CtBP transcriptional regulation (*72, 81*). Although high concentrations (mM) of MTOB are required, these studies validate CtBP as a therapeutic target. In the colon cancer model, MTOB evicted CtBP from its DNA-bound complex, reversing CtBP mediated repression of pro-apoptotic factors, thus inducing apoptosis in transformed cells (*81*). In the breast cancer model, CtBP mediated de-differentiation from epithelial to mesenchymal phenotype, a metabolic change observed specifically in transformed cells, was reversed in MTOB treated cells (*72*). These studies suggest that pharmacologically targeting CtBP would potentially have great benefits in treatment of various malignancies.

The effects of MTOB in these cancer models provided incentive to initiate the drug design process. MTOB was co-crystallized with CtBP1 and CtBP2 to ascertain how binding impacts the overall CtBP structure, and to determine the stereochemical features of the active site that could be exploited for drug design. Binding of MTOB failed to induce a significant overall tertiary or quaternary change in either CtBP1 or CtBP2, thus arguing against ligand-linked conformational changes as driving transcriptional modulation. Although the mechanism of inhibition remains undetermined, these structures provide the groundwork to begin rationally developing new small molecule inhibitors, which in turn should aid in understanding modulation of CtBP activity.

Our crystal structures of ternary complexes of MTOB and NAD^+ with CtBP1 and CtBP2 reveal minor differences in the active sites of the two enzymes that do not appear to preclude development of a single inhibitor for both enzymes. The observed differences in MTOB binding between the structures derive primarily from the flexibility of CtBP1 Arg97 relative to the homologous CtBP2 Arg103 residue. The MTOB α-keto acid group in CtBP2 forms hydrogen bonds and van der Waals contacts similar to other D2-DHDs, where as the CtBP1 network has distinctions. CtBP2 Arg103, whose conformation is similar in the binary and ternary complexes, remains outside of hydrogen bonding distance of the MTOB carboxylate, leaving space for the MTOB C9 atom to orient towards the Arg 103 guanidinium. In contrast, CtBP1 Arg97 attains distinct conformations in different crystal structures, exhibiting plasticity to interact with the specific ion or substrate occupying the active site. Movement of the Arg97 guanidinium group in the CtBP1 MTOB complex increases contact with the MTOB, apparently stabilizing the distinct carboxylate orientation relative to MTOB bound CtBP2. The change in carboxylate orientation disrupts some of the hydrogen bonds observed in CtBP2, but compensates through an additional hydrogen bond, van der Waals, and coulombic interactions with Arg97. Simultaneously, CtBP1 Arg97 movement causes minor clashes with MTOB C9, resulting in possible multiple conformations **(Figure 3.5C,D)**. Overall, the plasticity of Arg97 in CtBP1 relative to Arg103 in CtBP2 is not likely to preclude future inhibitors from binding both enzymes, as conformation of Arg97 in the binary complex closely resembles the CtBP2 position. Thus, small molecules

binding the CtBP2 active site may simply induce Arg97 in CtBP1 to assume the CtBP2 conformation.

Although the various CtBP structures fail to provide definitive reasons for the mobility of Arg97 in CtBP1, the plasticity is not expected to drastically affect drug design. The main difference between the CtBP1 and CtBP2 structures that may influence Arg97 mobility originates from sequence variation at a residue, CtBP1 Ile105 (CtBP2 Val111), which packs on the interior of the protein adjacent to the Arg97 (Arg103) side chain. The substitution of Val111 in CtBP2 diminishes contact with the Arg97/103 side chain. The change in packing with the larger CtBP1 isoleucine residue however, does not force Arg97 to shift relative to CtBP2, as evidenced by the ability of Arg97 in the CtBP1 binary complex to assume a similar position to the CtBP2 Arg103 **(Figure 3.8A, C)**. The similarity of the CtBP1 binary Arg97 and CtBP2 ternary structure, suggests other ligands and small molecules designed to bind the active site will simply induce CtBP1 Arg97 to adopt a conformation nearly identical to CtBP2 Arg103, thereby making binding conformation identical between the proteins.

CtBP1 Trp318 and CtBP2 Trp324, which line the active site, provide an extensive surface in each active site that is capable of accommodating a large moiety attached to an α-keto acid. In the ternary complexes, extensive interactions with Trp318/324 appear to contribute to the preference for MTOB over other substrates. These interactions include the packing of the MTOB sulfur atom (S8) against Trp318/324, which is reminiscent of sulfur-aromatic interactions identified through studies of protein and small molecule structures (*194-196*). The significant decrease in enzymatic efficiency when replacing the

sulfur with a methylene carbon (*63*) argues that the observed sulfur-aromatic interaction contributes to the specificity of CtBP for MTOB. The extensive aromatic surface formed by Trp318/324 represents an opportunity for inhibitor design. This surface is a characteristic not shared by other D2-HDH family members **(Figure 3.1)** and, thus, could be exploited for designing high affinity and specificity inhibitors to bind into the active site of CtBP.

A second important CtBP active site characteristic, a unique water filled cavity, would impart specificity on inhibitors designed fill the cavity. Other D2-HDHs have a bulky residue that forms stabilizing contacts with the substrate, isolating the active site. CtBP1 and CtBP2 however, possess a Ser (124) and Ala (130) at the homologous position, respectively, which do not contact the substrate and leave a gap in the active site wall that spills into another cavity. The water network that fills this extra cavity in CtBP1 (and presumably CtBP2) provides an opportunity to develop small molecules that bind with higher affinity and specificity. By displacing the water network with inhibitor atoms, additional contacts with CtBP will increase affinity while simultaneously reducing cross reactivity with other D2-HDHs that lack this cavity. Of course, increasing inhibitor affinity is contingent on the sum of new inhibitor interaction energies and entropic gains of water release being sufficient to offset the penalty of breaking the water network hydrogen bonds. If the water network can be displaced, the potential exists to eventually synthesize molecules that bind not only in the MTOB binding site and adjoining cavity, but also connects to moieties that take advantage at the NAD(H) binding site, greatly increasing inhibitor affinity, while maintaining CtBP specificity.

The goal of our ternary MTOB/NAD^+/CtBP structure determination was to obtain insights into the mechanism of inhibition and to determine detailed interactions that may facilitate inhibitor design. Given the lack of structural changes upon MTOB binding, mechanisms other than quaternary and tertiary rearrangements must underlie inhibition, perhaps involving substrate turnover and NAD^+ release, or altered coenzyme affinity. Any mechanism affecting coenzyme binding may effectively inhibit CtBP transcriptional modulation by disrupting dimerization. The detailed active site interactions described here provide a framework for inhibitor design. Two key features, the extensive interactions with Trp318/324 and a hydrophilic active site channel, may be particularly useful to exploit for inhibitor design, as they are unique to CtBP. Thus, the results presented here provide a first step guide design of highly specific inhibitors of CtBP that could become therapeutically valuable.

CHAPTER IV

Structure-Guided Design of CtBP Inhibitors

INTRODUCTION

CtBP represses transcription of a variety of genes involved in prevention of cancer formation. CtBP targeting of pro-apoptotic factors (*81*), cellular adhesion molecules (*66*), and tumor suppressors (*71*) facilitates progression from normal cellular function to malignancy. More broadly, CtBP appears to target thousands of genes for repression or activation, allowing epigenetic modulation of the metabolic state of a cell, including the epithelial to mesenchymal transition, a hallmark of cancer progression and invasiveness (*72*).

Two recent independent studies have demonstrated the potential, to pharmacologically target the role of CtBP in specific malignancies. These studies identified the role of CtBP in colon and breast cancer, and the ability of high levels of substrate, MTOB, to repress CtBP function (*72, 81*). However, MTOB treatment alone would be infeasible, due to the millimolar concentration required to abrogate CtBP function.

The impetus provided by the link between MTOB and CtBP inhibition, and the prospect of pharmacologically targeting CtBP to treat cancer, drove us to determine the structure of both CtBP1 and CtBP2 in complex with MTOB and cofactor NAD^+. These structures revealed unique characteristics of the CtBP active site, including a hydrated cavity not present in other related human proteins. By utilizing specific residue contacts, and the unique active site cavity, inhibitors that would improve on MTOB mediated CtBP inhibition may be identified or designed.

This chapter focuses on attempts to identify and characterize small molecule inhibitors of CtBP that will provide the groundwork for future structure based drug design. First, readily available α-keto acid compounds that would replicate and enhance the critical contacts observed in the CtBP structure were identified. These compounds were then screened by *in silico* docking to the CtBP1 active site to narrow down the number of compounds to test experimentally. Compounds that were predicted to bind similarly to MTOB were then subjected to co-crystallization trials with CtBP1 and binding affinity measurements with CtBP2. This approach allowed us to overcome the weak affinity of MTOB and identify substantially tighter binding compounds through structure-guided design combined with a bit of serendipity. Our best inhibitor, the tightest binder identified to date, will prove essential to the future progress in targeting CtBP for inhibition to treat cancer.

MATERIALS AND METHODS

CtBP constructs, expression, and purification

The constructs, expression, and purification were identical to those methods employed in Chapter III with the exception of the CtBP1 construct. A CtBP1 Q301A mutation was introduced to optimize lattice contacts.

Crystallization: *CtBP1*

Protein was thawed, and spin filtered in a Spin-X (Corning) column for 5 min at 4°C. Ligands were dissolved in stock solutions, neutralized by 20 mM Tris pH 7.5. Inhibitor was added at a 50 molar excess to protein. The solution was then mixed 1:1 with crystallization buffer. Previous CtBP1 crystals were frequently damaged during cryoprotection. New screening resulted in successful crystallization in buffers containing 75-300 mM magnesium chloride or calcium chloride, 100 mM Hepes pH 7.5, 2.5 mM NAD^+, and 10-40% PEG 400. Crystals grown in buffer containing greater than 20% PEG 400 were harvested and immediately flash frozen. Crystals grown in lower PEG 400 concentrations were transferred to higher PEG 400 concentrations for cryoprotection.

Data collection and structure solution

The CtBP1/NADH/Mandelic Acid complex diffraction data were collected on GM/CA CAT 23-ID-B beamline, at the Advanced Photon Source. Data for the CtBP1/NADH/2-hydroxyimino-3-phenyl propanoic acid complex were collected on a home source MicroMax007-HF/Saturn 944 CCD detector X-ray diffraction system. Model building and refinement were performed as in Chapter 2, utilizing Coot and the PHENIX software suite (*187, 188*).

Distance difference matrix plots

Distance difference matrix plots were constructed as described in Chapter III.

Docking

Docking experiments were performed using the Schrödinger Suite Glide program. Briefly, hydrogens were added, atom charges and protonation were calculated, and minimization was performed on the CtBP1-MTOB complex using the Protein Prepwizard (*189, 190*). MTOB was removed, and a docking grid was created around the CtBP1 active site. Ligand charges and stereochemistries were processed in LigPrep (*197*). Ligands were then docked into the rigid receptor using Glide (*198-201*). Docked ligands were minimized and scored.

Hydrogen bond determination

Hydrogen bonds were calculated as described in Chapter III.

Isothermal titration calorimetry

ITC experiments were performed on the ITC_{200} (GE). Purified CtBP2 was concentrated to 40-60 μM, and dialyzed overnight into 300 mM sodium chloride, 50 mM glycylglycine pH 8.5, 2 mM TCEP, with or without 1.5 mM NAD^+. Ligand stocks were diluted in dialysate. Experiments were performed at 20°C using twenty 2μL injections spaced by 180s. Data were processed in Origin.

RESULTS

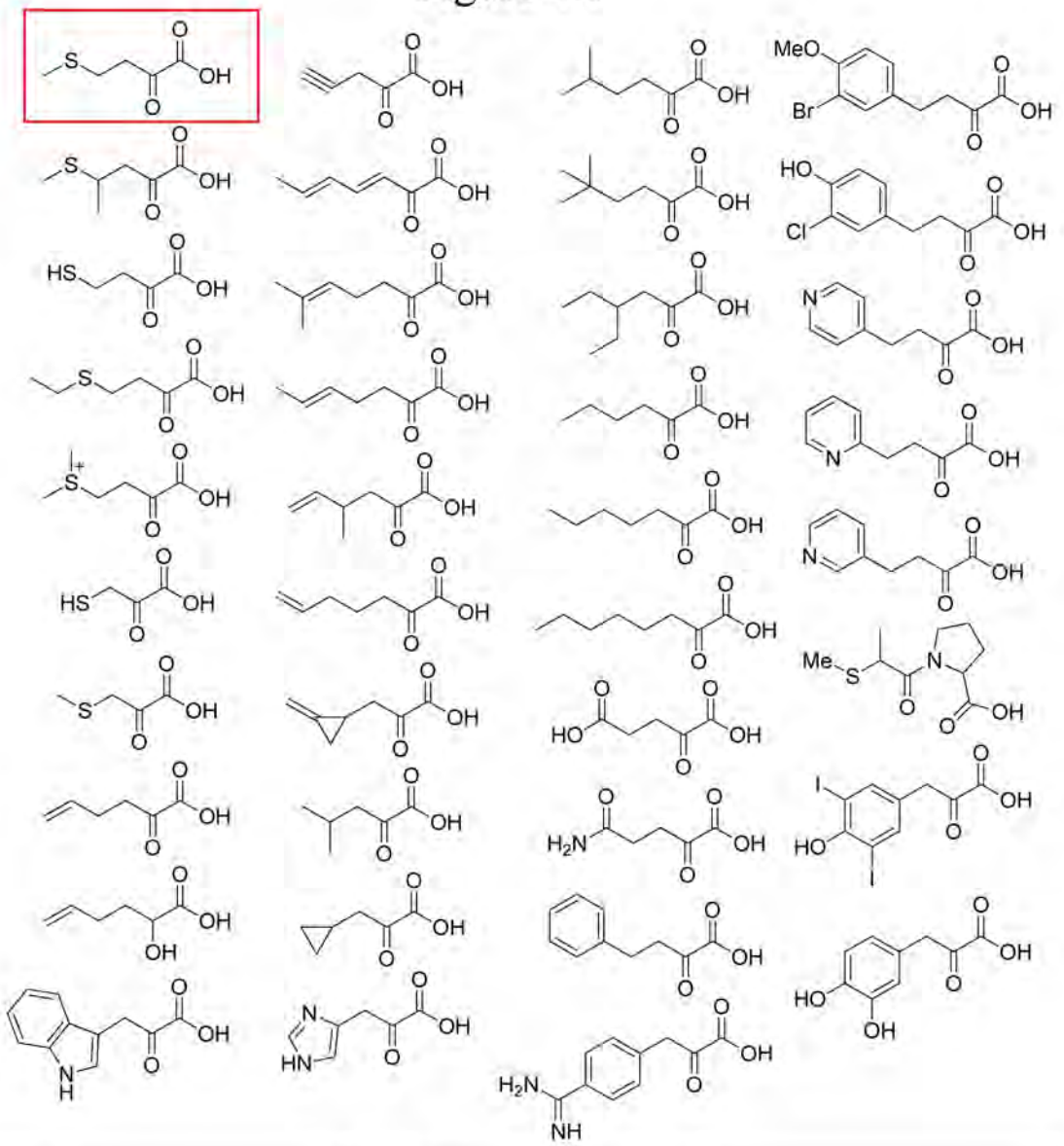
Exploiting CtBP - MTOB contacts to identify new inhibitors

The stereochemistry and contacts formed by MTOB in the co-crystal structures of CtBP1 and 2 aided in identifying new compounds to test for CtBP inhibition. Compounds from the Chemical Abstracts Service (CAS) registry chemically similar to MTOB were selected and tested *in silico* for their ability to dock into the CtBP1 active site. Initial compounds were selected to include an α-keto or α-hydroxy moiety to maintain favorable contacts observed with MTOB and other substrates across the D2-HDH family, with catalytic residues Arg266 and His315 and the backbone amides of the substrate binding domain active site loop **(Figure 3.7)**. Compounds varied in size and chemical properties outside the α-keto acid moiety **(Figure 4.1)**.

The variable region of the selected molecules contained different moieties designed to exploit the largest per residue interaction between MTOB and CtBP at Trp318, which dominates MTOB binding **(Figure 3.9A)**. Trp318 is unique among similar human D2-HDH family members, thereby conferring specificity for substrate and inhibitors. The MTOB sulfur atom (S8) contact with Trp318 appears to ascribe specificity over a methylene group, perhaps through a specific sulfur-pi interaction commonly observed in protein structures (*63, 194, 195*). Sulfur atoms were therefore incorporated into many of the initial docking candidates. Branched chain compounds were also tested for potential interactions and fit of various length and size compounds. Finally, compounds containing pi electron systems, including aromatic rings, were also tested in an attempt to induce pi interactions. Different length carbon linkers separated

Figure 4.1: Compounds docked into the CtBP1 active site. Compounds were chosen to mimic the conformation of the MTOB (red box) α-keto acid moiety while enhancing interactions with Trp318. Different chemical groups were tested to determine which compounds would assume the intended conformation. Sulfur containing compounds were chosen to exploit the MTOB sulfur-pi interaction. Branched chain compounds were selected for van der Waals/hydrophobic contacts. Alkene and alkyne compounds were chosen to induce pi-pi interactions. Finally, aromatic ring compounds were selected to maximize pi stacking with Trp318. Compounds with aromatic rings on the terminal carbon of two or three carbon chain length α-keto acids consistently had the highest docking scores (best predicted energies) and assumed a conformation similar to MTOB. Biochemical data indicated phenylpyruvate and a racemic mixture of ± vanillylmandelic acid bind the CtBP active site. Thus phenylpyruvate and D-mandelic acid (D-isomer product mimic) were selected for further biochemical and structural characterization.

Figure 4.1



Phenylpyruvate	± Vanillylmandelic Acid	D-Mandelic Acid

the α-keto moiety core and the pi systems. Although cation-pi interactions were considered, the CtBP active site contains two arginine residues and the protonated catalytic histidine, suggesting the repulsive forces of an additional cation would likely lower affinity of those compounds. See **Figure 4.1** for a full list of the initial compounds docked into the CtBP1 active site.

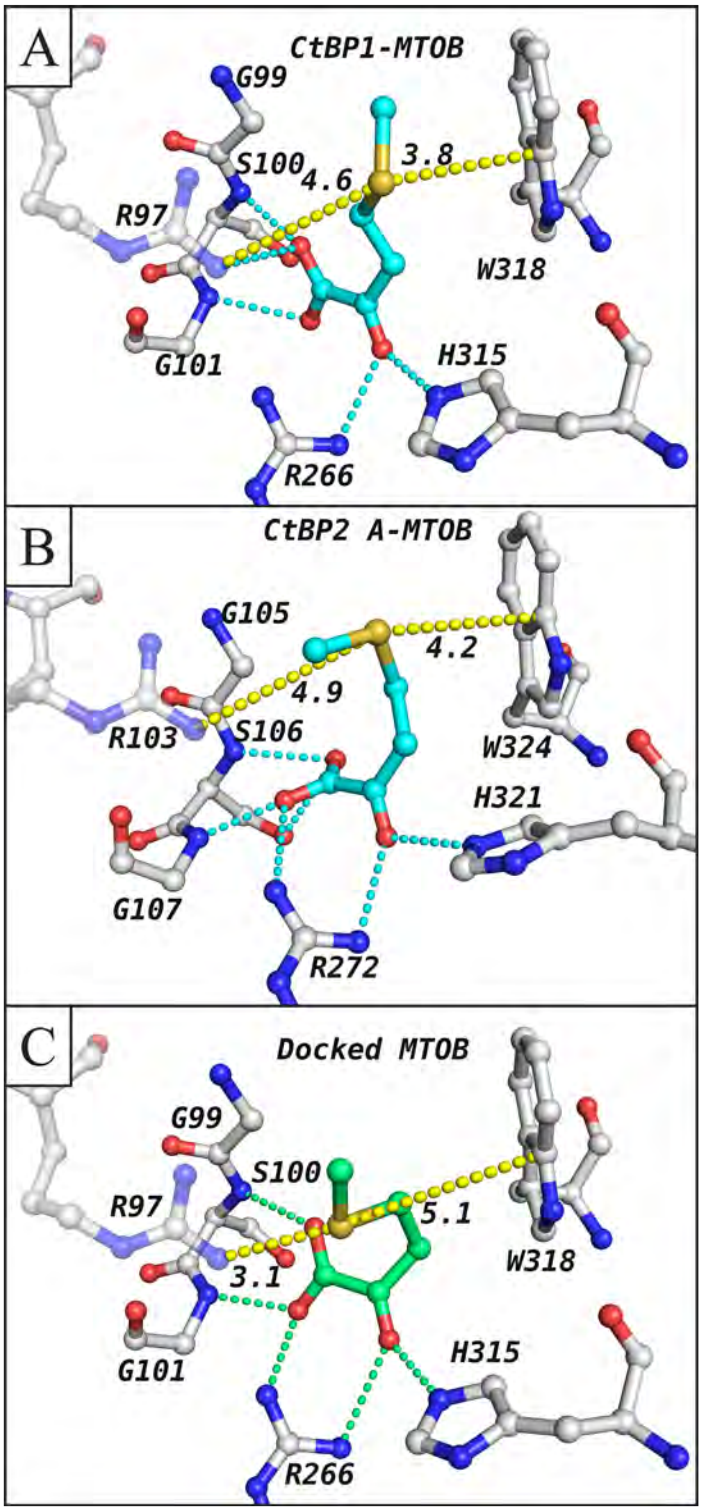
Interpreting docking results

The *in silico* docking of compounds into the CtBP1 active site resulted in selection of candidates for further biochemical and structural analysis. MTOB was docked into the CtBP1 active site to provide a control for identification of more tightly binding compounds. However, *in silico* docked MTOB failed to adopt the observed conformation present in the CtBP1 and 2 crystal structures. Rather, docking orients the sulfur atom towards Arg97 instead of positioning it adjacent to Trp318, potentially due to poor characterization of the energetic forces involved in sulfur-pi interactions. In this conformation, docking favors placement of the MTOB sulfur partial negative charge and hydrogen bond accepting lone pair in proximity to the positively charged, hydrogen bond donor guanidinium group **(Figure 4.2)**. Docking was able to accurately predict the conformation of the MTOB α-keto moiety, providing confidence in the ability to accurately dock other compounds despite the inability of the force fields to appropriately represent sulfur-pi interactions. Candidate molecules with strong predicted docking energies were aligned to the MTOB co-crystal structures to compare the position of the α-keto acid moieties. Sulfur containing candidate molecules, similarly to MTOB, were biased towards interaction with Arg97. These compounds therefore were allotted more

Figure 4.2: Docking software incorrectly docks MTOB. Panels (A) and (B) show the conformation and hydrogen bond network (cyan dashes) of MTOB in CtBP1 and CtBP2, respectively. The distance of the MTOB sulfur to the Trp318/324 Cε2 and Arg97/103 NH2 atoms is labeled (yellow dashes). (C) Docked MTOB has a similar hydrogen bond network as the crystal structures, but the sulfur atom is incorrectly positioned. The Glide docking software positions the MTOB sulfur atom within hydrogen bonding distance of Arg97/103. Additional docking suggests even when outside the maximum hydrogen bond distance, the partial negative character of the sulfur is still oriented towards Arg97/103. Because sulfur atoms were preferentially oriented towards Arg97/103 over Trp318/324, sulfur-containing compounds were allocated extra docking poses in an attempt capture conformations that interacted with Trp318/324.

Although docking of candidate compounds lacked an MTOB control with which to compare binding scores, the conformation of docked molecules still provided useful information regarding contacts with Trp318/324 and an overall fit within the active site.

Figure 4.2


A
CtBP1-MTOB
G99
S100
4.6
3.8
R97
W318
G101
H315
R266
B
CtBP2 A-MTOB
G105
4.2
4.9
R103
S106
W324
H321
G107
R272
C
Docked MTOB
G99
S100
R97
5.1
3.1
W318
G101
H315
R266

final docked conformations to capture those interacting with Trp318, in addition to the conformation interacting with Arg97. Molecules with both proper α-keto orientation and interaction with Trp318 were selected for co-crystallization and binding analysis.

Stacking aromatic rings with Trp318

Compounds with phenyl rings linked to an α-keto (2-oxoacid) moiety assumed a conformation in docking simulations similar to that in the MTOB co-crystal structures while also stacking the aromatic group with the Trp318 indole. Docking suggested two such compounds would have the highest probability of assuming a conformation similar to MTOB while stacking with Trp318. These compounds include 2-oxo-2-phenylethanoic acid (phenyl ring attached to the smallest possible α-keto acid) and 2-oxo-3-phenylpropanoic acid (phenylpyruvate; α-keto acid linked to phenyl ring with a single methylene carbon) **(Figure 4.1)**. Docking indicated the potential for clashes between active site residues and the ligand aromatic ring when two methylene carbons linked the phenyl ring to the α-keto acid moiety, as observed in 2-oxo-4-phenylbutanoic acid or larger compounds.

Biochemical data provided the final evidence necessary to choose compounds for co-crystallization and binding assays. Phenylpyruvate had previously been established as a CtBP1 substrate (*63*), indicating the compound could bind the active site. A co-crystal structure would determine if the ligand's aromatic ring stacked with the Trp318 indole as predicted by docking. Small molecule library screening for CtBP enzymatic inhibition by Grossman et al. (results not published) revealed D/L vanillylmandelic acid (VMA), a racemic mixture of an α-hydroxy acid similar to 2-oxo-2-phenylethanoic acid, inhibited

CtBP, presumably by binding the active site **(Figure 4.1)**. Based on the chemical screening results the VMA D-isomer and the D-2-hydroxy-2-phenylethanoic acid (D-mandelic acid) were docked with CtBP1 **(Figure 4.3)**. The results suggested that even with one less methylene carbon separating the α-keto/α-hydroxy group and the phenyl ring than phenylpyruvate, stacking may still occur with D-VMA or D-mandelic acid and the Trp318 indole. Although a racemic mixture of VMA is commercially available, the high cost and lack of enantiomer separation prohibited co-crystallization and binding assay development. A single enantiomer of an α-hydroxy acid enzymatic product mimic, D-mandelic acid, and phenylpyruvate were both reasonably priced, and therefore procured for co-crystallization and binding assays.

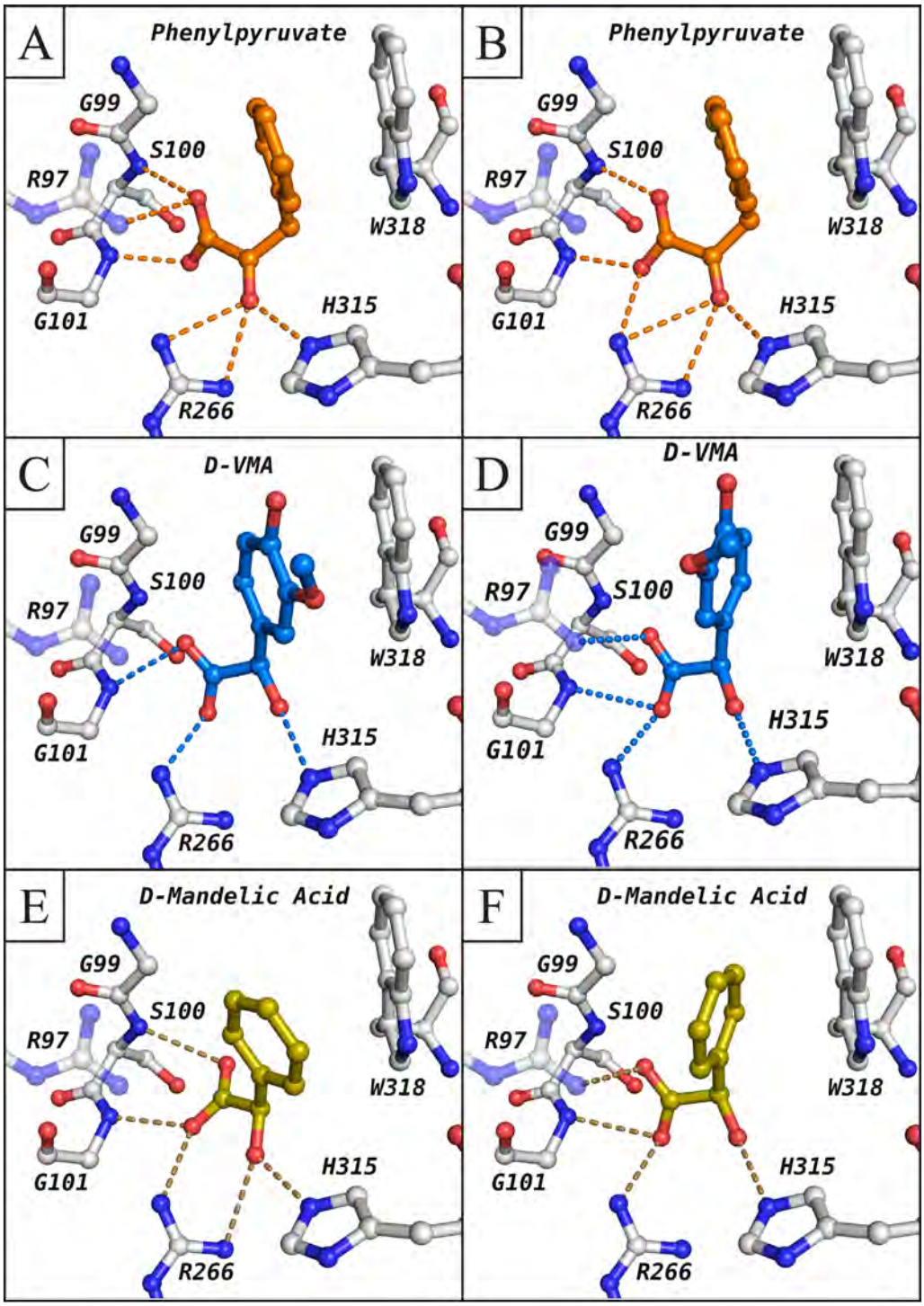
Small molecule CtBP1 co-crystallization and structure

Attempts were made to co-crystallize phenylpyruvate and D-mandelic acid with CtBP1 due to the typically higher resolution diffraction of this CtBP family member. The datasets collected from phenylpyruvate co-crystals failed to process beyond 3.5 Å resolution. The electron density in the CtBP1 active site was ambiguous and uninformative at this resolution. However, the D-mandelic acid co-crystals diffracted to 2.5 Å resolution. The active site was clearly bound by D-mandelic acid, with the phenyl ring positioned adjacent to Trp318 **(Figure 4.4)**.

D-mandelic acid was present in the active site, but the co-crystal structure contained serious flaws. Similarly to MTOB, D-mandelic acid failed to fully bind the CtBP1 active site, evidenced by the occupancy refining to 68%. The refined $2F_o$-F_c maps fail to completely encompass D-mandelic acid around a small portion of the phenyl ring

Figure 4.3: Candidates chosen for structural and biochemical characterization. Docking conformations and hydrogen bond networks (dashes) of candidate molecules are shown. (A) and (B) Separate docking conformations of phenylpyruvate (orange) into the CtBP1 active site. Docking predicted the α-keto acid moiety would assume a conformation similar to MTOB with the phenyl ring would stacking against Trp318. Very few conformational differences were observed under different docking conditions. (C) and (D) The α-hydroxy moiety of the D-isomer of VMA assumes the expected conformation. The aromatic ring interacts with Trp318, although different conformations are predicted. (E) and (F) D-mandelic acid (yellow) is predicted to bind similarly to D-VMA. Phenylpyruvate and D-mandelic acid were selected for structural and biochemical characterization.

Figure 4.3


A
Phenylpyruvate
G99
S100
R97
W318
G101
H315
R266
B
Phenylpyruvate
G99
S100
R97
W318
G101
H315
R266
C
D-VMA
G99
S100
R97
W318
G101
H315
R266
D
D-VMA
G99
S100
R97
W318
G101
H315
R266
E
D-Mandelic Acid
G99
S100
R97
W318
G101
H315
R266
F
D-Mandelic Acid
G99
S100
R97
W318
G101
H315
R266

and carboxylate group **(Figure 4.4A-C)**. The electron density suggests the phenyl ring position is correct. However, the same is not true for the accuracy of the carboxylate position. Similar to MTOB in the CtBP1 structure, the density around the carboxylate is incomplete. Unlike MTOB, every conformation tested contained steric clashes with Arg97, Arg266, or an active site water molecule when utilizing the existing density and refined B factors **(Figure 4.4B)**. The overall moderate to poor ligand density combined with the uncertainty of the carboxylate position prevents detailed analysis of the binding.

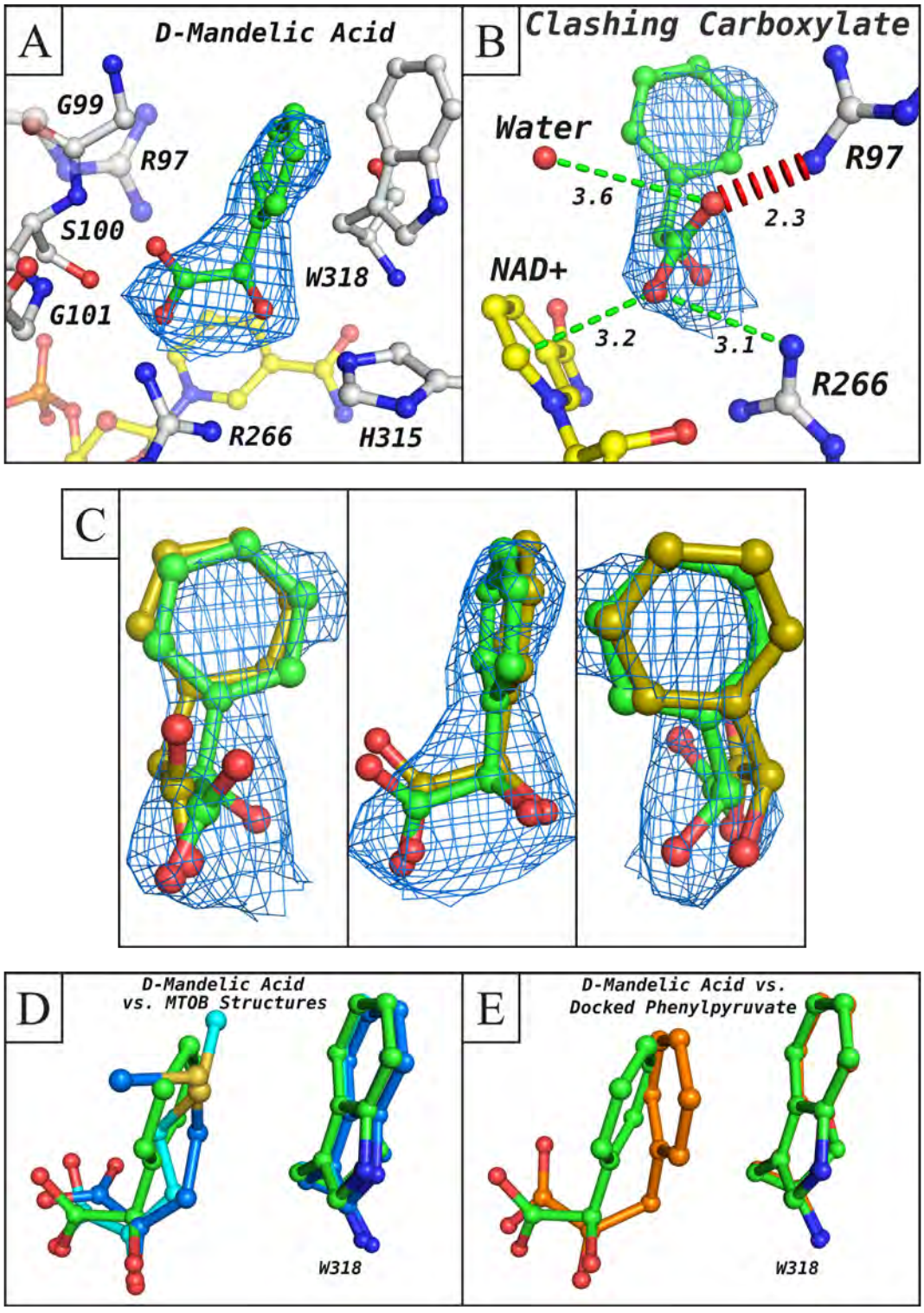
The position of the D-mandelic acid phenyl ring closely matches the *in silico* docking results, but D-mandelic acid does not appear to strongly bind the CtBP1 active site **(Figure 4.4C)**. The plane of the D-mandelic acid phenyl ring positions slightly further from the Trp318 indole than the MTOB sulfur atom **(Figure 4.4D)**, but well within distances for pi-pi interactions. Despite achieving the intended conformation and interaction with Trp318, the occupancy (68%) and poor density suggest D-Mandelic acid binding is weak. These results did not lead us to abandon characterization of D-Mandelic acid binding, but our priorities shifted to characterization of phenylpyruvate. Docking indicates that the additional methylene carbon in phenylpyruvate permits the aromatic ring to assume a distinct conformation relative to D-mandelic acid, altering and potentially strengthening contacts between the aromatic groups **(Figure 4.4E)**.

2-Hydroxyimino-3-phenyl propanoic acid

While attempting to measure phenylpyruvate binding, and continuing to generate new crystals for co-crystal structures, our collaborator Dr. Keith Ellis suggested utilizing 2-hydroxyimino-3-phenylpropanoic acid (HIPP) to mimic the transition state of

Figure 4.4: The 2.5 Å D-mandelic Acid / CtBP1 complex. (A) D-mandelic acid is positioned into the $2F_o$-F_c density contoured to 1 σ (0.17 e-/Å^3) in the CtBP1 active site. The density is not high quality, but it is readily apparent that the D-mandelic acid aromatic ring interacts with Trp318. (B) The ligand carboxylate cannot be accurately modeled. As modeled currently the carboxylate clashes with Arg97 (red disks). Rotation of the carboxylate away from Arg97 induces clashes with Arg266 or a crystallographic water. Occupancy refined to only 68%. (C) Comparing docked D-mandelic acid (yellow) with the crystal structure demonstrates their similarity. However, the docked conformation does not satisfy the electron density. (D) The co-structure reveals that the aromatic ring stacks with Trp318, similarly to the MTOB sulfur (dark blue, MTOB/CtBP2; cyan, MTOB/CtBP1). (E) Docking predicts the additional methylene carbon in phenylpyruvate causes a distinct, and potentially stronger interaction with Trp318 relative to D-mandelic acid.

Figure 4.4


A
D-Mandelic Acid
G99
R97
S100
G101
W318
R266
H315
B
Clashing Carboxylate
Water
R97
3.6
2.3
NAD+
3.2
3.1
R266
C
D
D-Mandelic Acid
vs. MTOB Structures
W318
E
D-Mandelic Acid vs.
Docked Phenylpyruvate
W318

phenylpyruvate conversion to D-phenyllactate. As previously mentioned, CtBP1 catalyzes the reduction of phenylpyruvate, but at a lower efficiency than MTOB (*63*). A compound mimicking the transition state could maintain the beneficial contacts between the chemical intermediate and CtBP without catalytic turnover and subsequent release of the product. HIPP substitutes the reactive α-keto acid carbonyl with a hydroxyimino group, preventing catalysis. Co-crystallization and binding affinity studies revealed HIPP binds CtBP with high affinity relative to MTOB, but assumes an unexpected conformation in the active site.

CtBP1 HIPP co-crystal structure

Structure determination

Human CtBP1 Q301A (28-353) was crystallized in complex with HIPP and NAD^+ **(Table 4.1)**. Unlike the active site electron densities for phenylpyruvate and D-mandelic acid, the HIPP electron density was immediately apparent despite the low resolution of the data. The MTOB bound CtBP1 complex served as a search model for molecular replacement, providing the initial phases. Identical to the other existing CtBP1 structures, the HIPP bound CtBP1 complex crystallized in space group $P6_422$, with a monomer in the asymmetric unit, with physiological dimers formed across a two fold crystallographic axis. The Q301A mutation, introduced to improve a crystal lattice contact, did not alter the crystallization or conformation of CtBP1.

HIPP Quaternary Structure

Distance difference matrix plot analysis indicates only small Cα variations exist in the HIPP-CtBP1 complex relative to both the binary CtBP1 complex (pdb 1MX3) (*10*)

Table 4.1

	CtBP1-Mandelic Acid	CtBP1-HIPP
Wavelength (Å)	0.99	1.54
Resolution range (Å)	34.97 - 2.45 (2.54 - 2.45)	24.56 - 3.06 (3.11 - 3.06)
Space group	$P6_422$	$P6_422$
Unit cell a b c (Å)	83.95 83.95 159.52	84.32 84.32 161.06
α β γ (°)	90 90 120	90 90 120
Total reflections	99465	68620
Unique reflections	12576	6857
Multiplicity	7.9 (8.7)	10.0 (6.6)
Completeness (%)	97.8 (97.3)	99.9 (99.9)
Mean I/sigma(I)	21.4 (4.5)	19.4 (4.9)
Wilson B-factor	49.9	56.5
R-sym (%)	9.9 (57.7)	7.5 (39.2)
R-factor (%)	20.4 (24.4)	23.3 (32.5)
R-free (%)	24.3 (27.3)	27.4 (36.8)
Number of atoms	2566	2500
Macromolecules	2463	2410
Ligands	55	57
Waters	48	33
Protein residues	327	325
RMS (bonds)	0.008	0.004
RMS (angles)	1.13	0.71
Ramachandran favored (%)	96	96
Ramachandran outliers (%)	0	0
Clashscore	5.65	10.55
Average B-factor	61.3	70.8
Macromolecules	61.6	71
Solvent	58.7	66.2

and MTOB bound CtBP1 **(Figure 4.5A,B)**. The small magnitude suggests these differences arise from artifacts of crystallization and the lower resolution of the HIPP structure, not from ligand induced changes. HIPP binding, similar to MTOB binding, does not induce any significant tertiary or quaternary changes in CtBP1.

HIPP placement and conformation

Difference maps guided the placement of HIPP into the CtBP1 active site, revealing an unexpected HIPP conformation. The active site of the initial model contained strong F_{obs}-F_{calc} (F_o-F_c) density. The prominence of the HIPP phenyl ring density made it immediately apparent that it stacks with Trp318 as intended. However, refining the HIPP carboxylate in a position analogous to the MTOB conformation resulted in positive and negative F_o-F_c density at different positions in the active site **(Figure 4.6A)**. In this conformation, the hydroxyimino group oriented towards His315, analogous to the carbonyl of MTOB. However, the HIPP hydroxyimino group (atom O6) and carboxylate oxygen (atom O1) sterically clashed with Arg266 and S100 respectively **(Figure 4.6B)**. Model building and subsequent refinement never relieved both the steric clashes and F_o-F_c peaks. The clashes and difference map peaks therefore suggested HIPP assumed a novel conformation, distinct from MTOB. Reorienting both the HIPP carboxylate oxygens towards the Arg266 guanidinium, and the hydroxyimino group towards the active site loop, satisfied the F_o-F_c peaks and relieved the steric clashes, while maintaining the aromatic stacking with Trp318 **(Figure 4.6C, D)**. Importantly, the maps and refinement indicate HIPP assumes full occupancy of the active site unlike all other liganded CtBP structures.

Figure 4.5: The Conformation of the HIPP structure matches the binary and MTOB bound CtBP1 structures. Distance difference plot matrix analysis indicates only small changes occur upon HIPP binding relative to the CtBP1 binary complex (pdb code 1MX3) (*10*). As before, generating symmetry mates created CtBP1dimers. Comparison of analogous monomers (subunit 1 and 1') reveals small changes in the substrate binding domain (discontinuous 28-120 and 319-353) relative to the dimerization loop (148-178) and a region (205-230) involved in lattice contacts. The same lattice contact showed variability when comparing binary and MTOB bound CtBP2 **(Figure A.3 purple boxes)** and when comparing MTOB bound CtBP1 and CtBP2 **(Figure A.5 Group 1)**. When viewing the position of subunit 1 to 2' (or 1' to 2) the highest positional variation occurs in the substrate binding, specifically residues 55-90, relative to one another (red boxes). (B) Analysis with the MTOB bound dimer shows the smaller magnitude movements in the same regions, suggesting ligand binding alone does not induce this shift. Likely, the differences in crystallization conditions combined with the lower resolution result in the movement of substrate binding domains relative to one another. A higher resolution HIPP structure would definitively demonstrate if a significant shift is occurring.

Figure 4.5

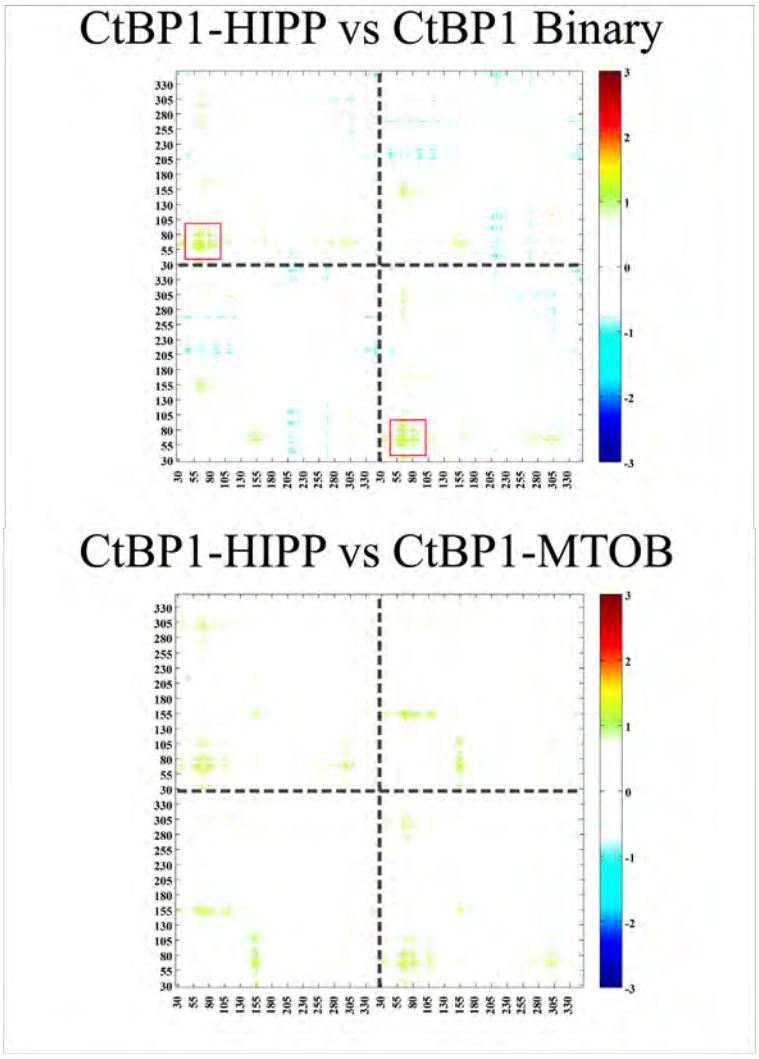
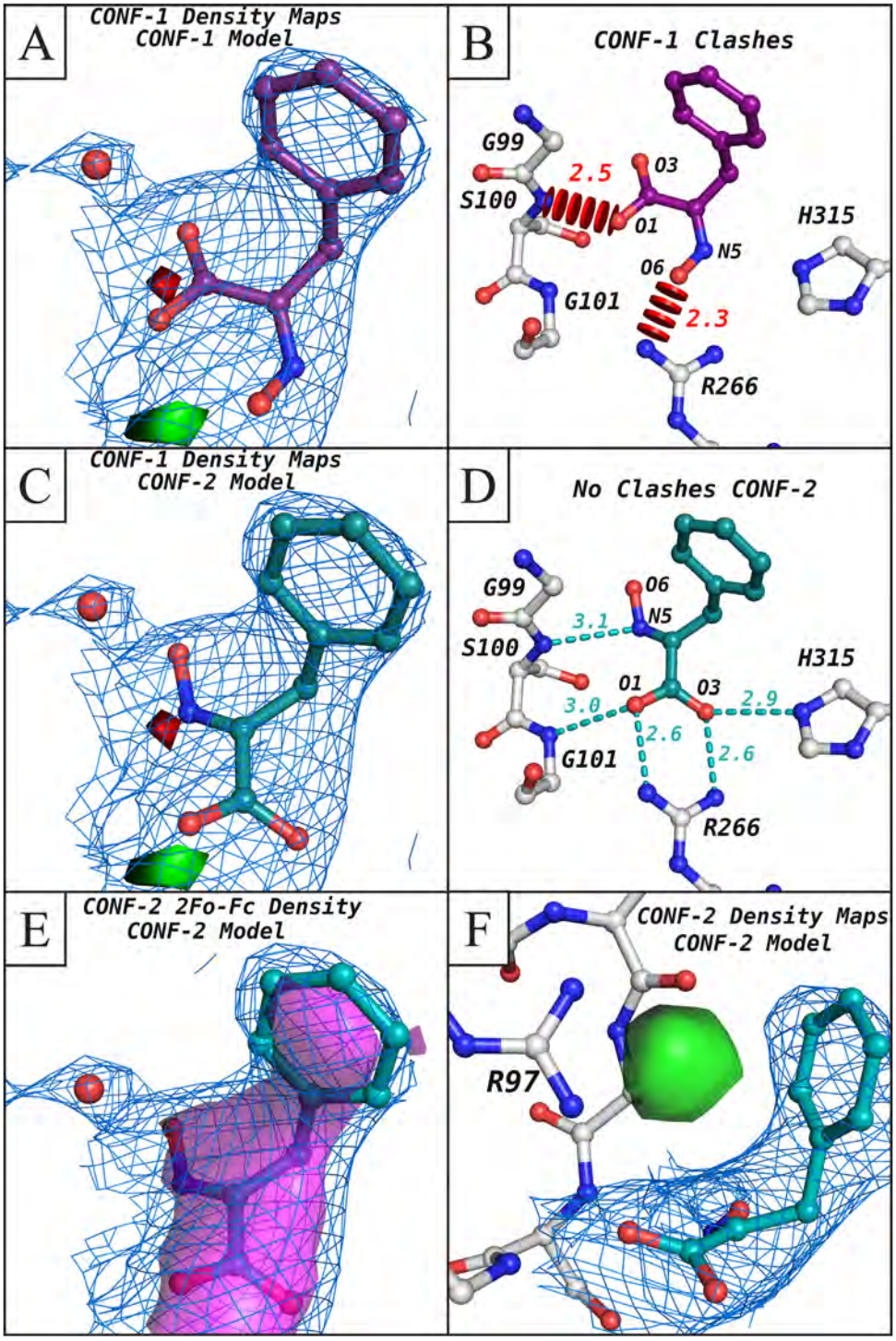


Figure 4.6 Positioning HIPP in the CtBP1 active site. (A) The initial HIPP conformation (Conf-1, purple) modeled orienting the hydroxyimino group analogous to the MTOB carbonyl. The phenyl ring position was immediately clear, as shown in the $2F_o$-F_c density (blue mesh) contoured to 1 σ (0.16 e-/Å^3). A water molecule (red sphere) was apparent in this active site density. The F_o-F_c map (3 σ; 0.21 e-/Å^3) showed both positive (green surface) and negative (red surface) density, indicating a need for more and less electron density, respectively. (B) Additionally, HIPP carboxylate oxygen O1 and hydroxyimino oxygen O6 sterically clash (red disks) with Ser100 and Arg266, respectively. (C, D) A new conformation of HIPP was chosen (Conf-2, teal) and placed into the Conf-1 active site maps. Conf-2 appeared to satisfy the difference map peaks and not possess clashes with active site residue. (E) Refinement of Conf-2 shows a good fit into the $2F_o$-F_c density, contoured to 0.16 e-/Å^3 (blue mesh) and 0.28 e-/Å^3 (purple surface). (F) As panel C suggests, the Conf-1 F_o-F_c peaks disappear when HIPP is refined in the Conf-2 orientation. The remaining peak (present in both structures but omitted from panel A) is positive density near Arg97. This density belongs to a water molecule that may partially occupy this position, as observed in the binary structure (1MX3) (*10*).

Figure 4.6


A
CONF-1 Density Maps
CONF-1 Model
B
CONF-1 Clashes
G99
S100
2.5
O3
O1
N5
O6
2.3
G101
H315
R266
C
CONF-1 Density Maps
CONF-2 Model
D
No Clashes CONF-2
G99
S100
3.1
O6
N5
H315
O1
O3
3.0
2.9
2.6
2.6
G101
R266
E
CONF-2 2Fo-Fc Density
CONF-2 Model
F
CONF-2 Density Maps
CONF-2 Model
R97

The only remaining active site F_o-F_c peak occurs away from HIPP, adjacent to the Arg97 guanidinium **(Figure 4.6E)**. The variability of Arg97 across CtBP1 structures suggests small movement of the guanidinium away from the active site, which would allow a water molecule or ion binding in the positive F_o-F_c peak, interacting with Arg97 and His77, but not clashing with HIPP. A water molecule is present at this position in the CtBP1 binary complex (pdb code 1MX3), where Arg97 has assumed a conformation further outside the active site. Movement of the CtBP1-HIPP complex Arg97 side-chain further into the active site to fill the F_o-F_c density would cause collisions with either His77 or HIPP.

After further refinement, an omit map confirmed the novel HIPP conformation. The novel HIPP conformation matches very well with the omit map F_o-F_c density, unlike the initial HIPP conformation based on MTOB binding **(Figure 4.7)**. Replacing the single oxygen atom of the MTOB carbonyl with the two atoms of the hydroxyimino group results in a compound that cannot physically situate in the active site analogous to D2-HDH substrate conformation, and instead adopts an unexpected conformation. HIPP *in silico* docking also results in this novel conformation instead of mimicking substrate or product binding, in agreement with the crystallographic finding **(Figure 4.8A)**.

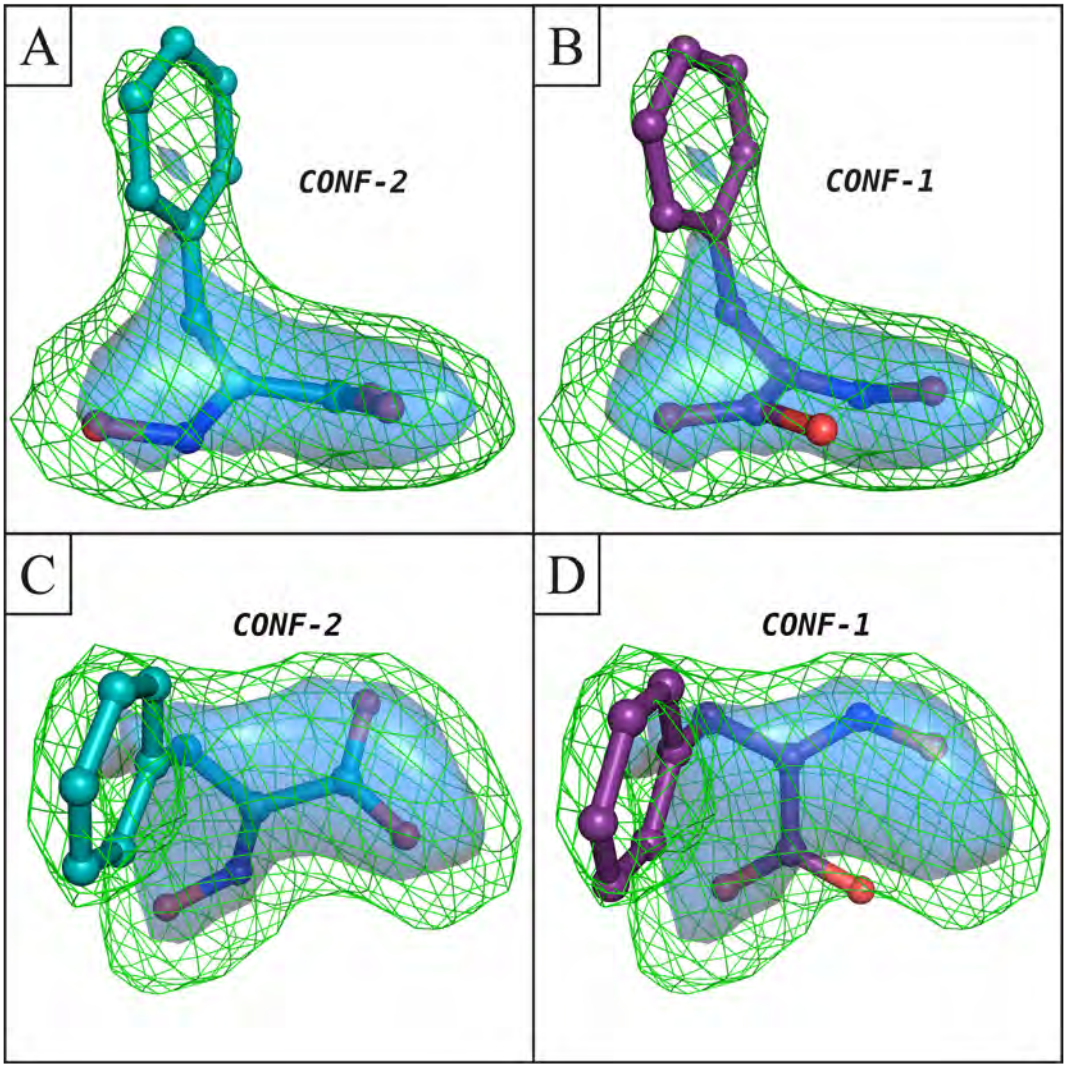
HIPP interactions and affinity

HIPP aromatic and coulombic interactions

After verifying that HIPP adopts a novel confirmation, computational analysis was performed to investigate the changes associated with the new confirmation. The HIPP carboxylate orients to maximize coulombic interactions **(Figure 4.8B, C)**.

Figure 4.7: The HIPP omit map confirms the novel conformation. An F_o-F_c omit map contoured to 0.21 e-/Å^3 (green mesh) and 0.30 e-/Å^3 (blue surface). (A) Conf-2 (teal) fits well within the difference map density. (B) Conf-1 (purple) does not fit as well, with a carboxylate oxygen protruding out of the density. This carboxylate position is the same location in Figure 4.6A where the Conf-1 F_o-F_c map indicated the need for less electron density. (C) Rotating the view 90° shows how well Conf-2 fits within the density, especially at the higher contour level (blue surface). (D) Again, Conf-1 fails to fit as well. The omit map helps confirms Conf-2, the novel conformation different from that of MTOB, is in fact the actual conformation of HIPP in the CtBP1 active site.

Figure 4.7



Both carboxylate oxygens align with the Arg266 guanidinium nitrogens. The carboxylate oxygens interact with two additional charged active site residues, positioning carboxylate oxygen O1 2.7Å from the Arg97 guanidinium and carboxylate oxygen O3 2.9Å from the protonated catalytic His315. Beyond the coulombic contacts, HIPP orients its phenyl ring to stack with Trp318, capturing the intended pi-stacking interactions despite the alternate conformation **(Figure 4.8B, D)**. Thus, the structure suggests very favorable interactions with CtBP1 across the entire HIPP molecule.

HIPP hydrogen bonding

The unique conformation of HIPP results in a hydrogen bond network distinct from MTOB that utilizes the same active site hydrogen bond donors **(Figure 4.9A)**. Catalytic residues His315 and Arg266, as well as substrate binding domain residue G101 form a network of four bonds with the HIPP carboxylate (O1 and O3). The distinctive HIPP conformation positions the imine nitrogen (N5) adjacent to S100, whereas the equivalent MTOB O5 carbonyl atom instead hydrogen bonds with His315. S100 can form two hydrogen bonds, either from the backbone amide or side chain Oγ, with the HIPP imino lone pair. A final hydrogen bond is formed between the HIPP hydroxyl group oxygen (O6) and a crystallographic water. Hydrogen bonds in an MTOB-like conformation would be weaker, supporting the observed novel confirmation **(Figure 4.9B)**.

Water network

Larger CtBP inhibitors may displace crystallographically observed active site water molecules when they bind CtBP. CtBP contains a unique network of four

Figure 4.8: HIPP interactions in the CtBP1 active site. (A) The HIPP co-crystal structure (grey protein, teal HIPP) compared to the Glide docking results (magenta). HIPP assumes a very similar conformation in both structures, verifying the unexpected conformation in the crystal structure. (B) The HIPP phenyl ring closely interacts with the Trp318 indole (teal dashes). The negatively charged HIPP carboxylate positions adjacent to Arg97, Arg266, and His315, maximizing coulombic interactions (yellow dashes). (C) A closer view of the coulombic interactions with the distances between atoms listed (Å). (D) A similar view for the aromatic ring interaction with Trp318. A portion of the HIPP phenyl ring positions at the minimum van der Waal distance from the Trp318 pyrrole group. The remaining portion of the ring positions slightly further away from the Trp318 benzene ring.

Figure 4.8

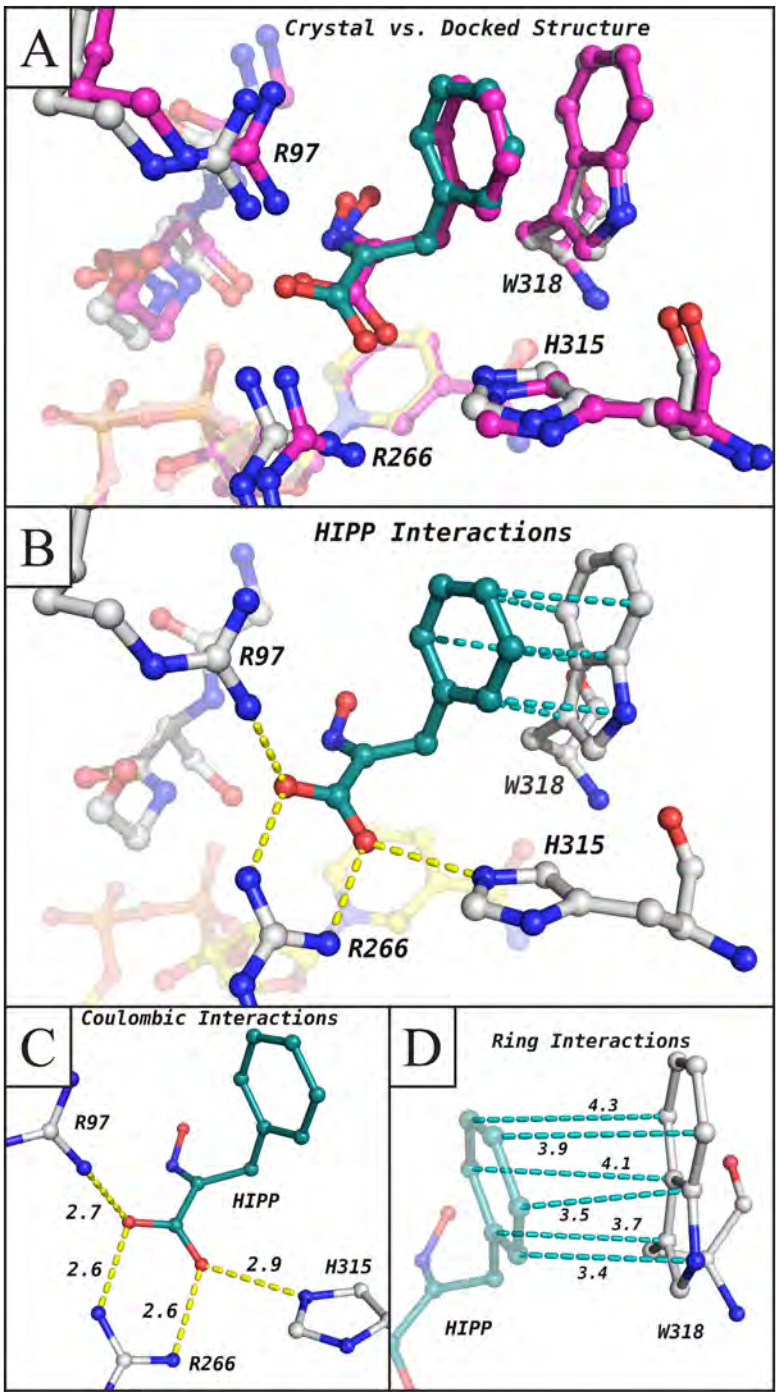
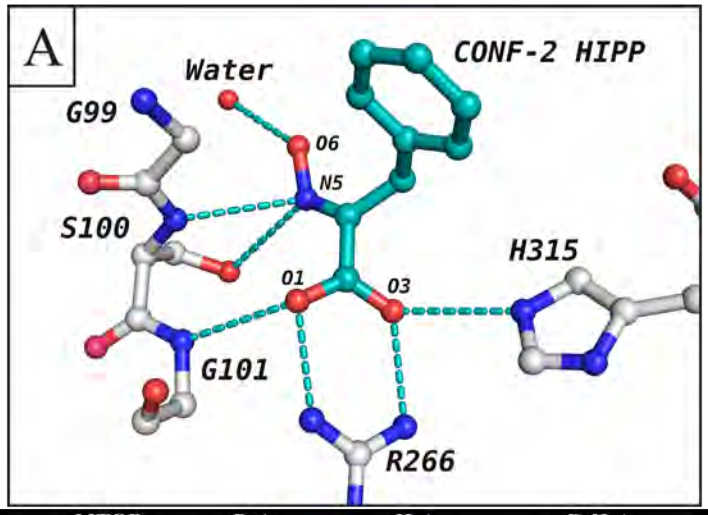
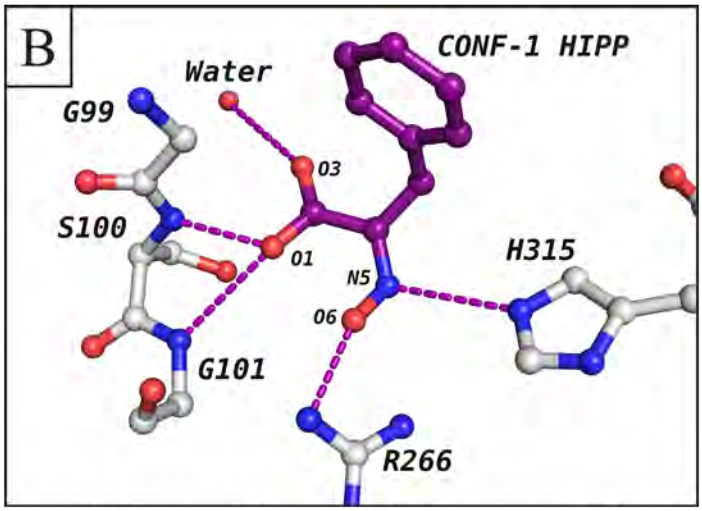

A
Crystal vs. Docked Structure
R97
W318
H315
R266
B
HIPP Interactions
R97
W318
H315
R266
C
Coulombic Interactions
R97
HIPP
2.7
2.6
2.9
H315
2.6
R266
D
Ring Interactions
4.3
3.9
4.1
3.5
3.7
3.4
HIPP
W318

Figure 4.9: The HIPP – CtBP1 hydrogen bond network. (A) The hydrogen bond network of HIPP Conf-2 (teal). Hydrogen bonds were calculated as described in the *Materials and Methods* (Chapter III). HIPP forms six hydrogen bonds with CtBP1 residues, and a seventh hydrogen bond with a crystallographic water. (B) HIPP Conf-1 (purple) has a less ideal hydrogen bond network. There are four hydrogen bonds to CtBP1, and a fifth hydrogen bond with the crystallographic water. The somewhat weaker hydrogen bond network for HIPP Conf-1 again suggests HIPP assumes the novel conformation (Conf-2).

Figure 4.9



Residue (Atom)	MTOB Atom	D-A Distance (Å)	H-A Distance (Å)	D-H-A Angle (°)	H-A-AA Angle (°)
H315	O3	2.9	2.0	149	136
R266 NH1	O3	2.6	1.6	165	135
R266 NH2	O1	2.6	1.6	167	106
G101	O1	2.7	2.1	152	132
S100 N	N5	3.1	2.1	168	97
S100 Og	N5	3.2	2.2	169	102
Water	O6	2.6	-	-	-



Residue (Atom)	MTOB Atom	D-A Distance (Å)	H-A Distance (Å)	D-H-A Angle (°)	H-A-AA Angle (°)
H315	N5	3.1	2.1	163	109
R266 NH2	O6	2.3	1.4	153	158
G101	O1	3.0	2.0	168	158
S100 N	O1	2.5	1.6	140	102
Water	O3	3.0	-	-	-

crystallographic water molecules (positions lettered A-D, as discussed in Chapter III). In the CtBP1-MTOB complex the waters form a chain connecting the MTOB carboxylate (water molecule position A) to the coenzyme nicotinamide phosphate oxygen (water molecule position D), threading through a cavity unique to CtBP. This cavity presents a unique opportunity for fitting larger, higher affinity, CtBP specific inhibitors at the energetic cost of displacing the water network. Previous studies with inhibitors of p38 alpha MAP kinase (*202*) and the fungal enzyme scytalone dehydratase (*203*) successfully displaced ordered water molecules to improve inhibitor binding and activity. However, other attempts to displace waters resulted in decreasing the inhibitor affinity, as observed with an EGFR kinase inhibitor (*204*), providing an important caveat with this approach.

HIPP binding alters the CtBP1 water network relative to the binary (*10*)and MTOB bound CtBP1 structures, suggesting water displacement may be a practical approach to future inhibitor design. Importantly, the HIPP hydroxyimino group completely displaces water from position A **(Figure 4.10C)**. Furthermore, alignment with the CtBP2 binary structure monomer G (pdb code 2OME), CtBP1 binary (pdb code 1MX3), rat CtBP1 (pdb codes 1HKU and 2HU2), and CtBP1-MTOB complexes indicates the position of water B has shifted about 1 Å (new position referred to as B'), moving the water molecule to a location 2.6 Å from the HIPP hydroxyimino oxygen (atom O6). The movement alters the contacts between water at position B' and the protein, relative to position B. Although the water B' position is unambiguous **(Figure 4.10A, B)**, even at the low resolution of the HIPP structure, water molecules are not observed at positions C and D, similar to the comparably low resolution rat CtBP1

structure (*205*) (pdb code 2HU2). Interestingly, water molecules in the B' position may interfere with water molecule binding at position C. Alignment of the HIPP structure to the CtBP1 structures containing full water networks indicates the distance between water positions B and C decreases from 2.9 Å to 2.4-2.6 Å upon shifting the water molecule to the B' position **(Figure 4.10C, D)**. Hence, HIPP binding displaced water at position A, and shifted the water molecule B into a position where it would displace or modify the position of water molecule C. These results imply the CtBP water network, or portions of the network, may be displaced by inhibitor moieties without an excessive energetic penalty. Successful displacement by newer inhibitors will require the sum of gaining new inhibitor contacts and entropically favored release of the waters offsetting the hydrogen bonds lost by the waters.

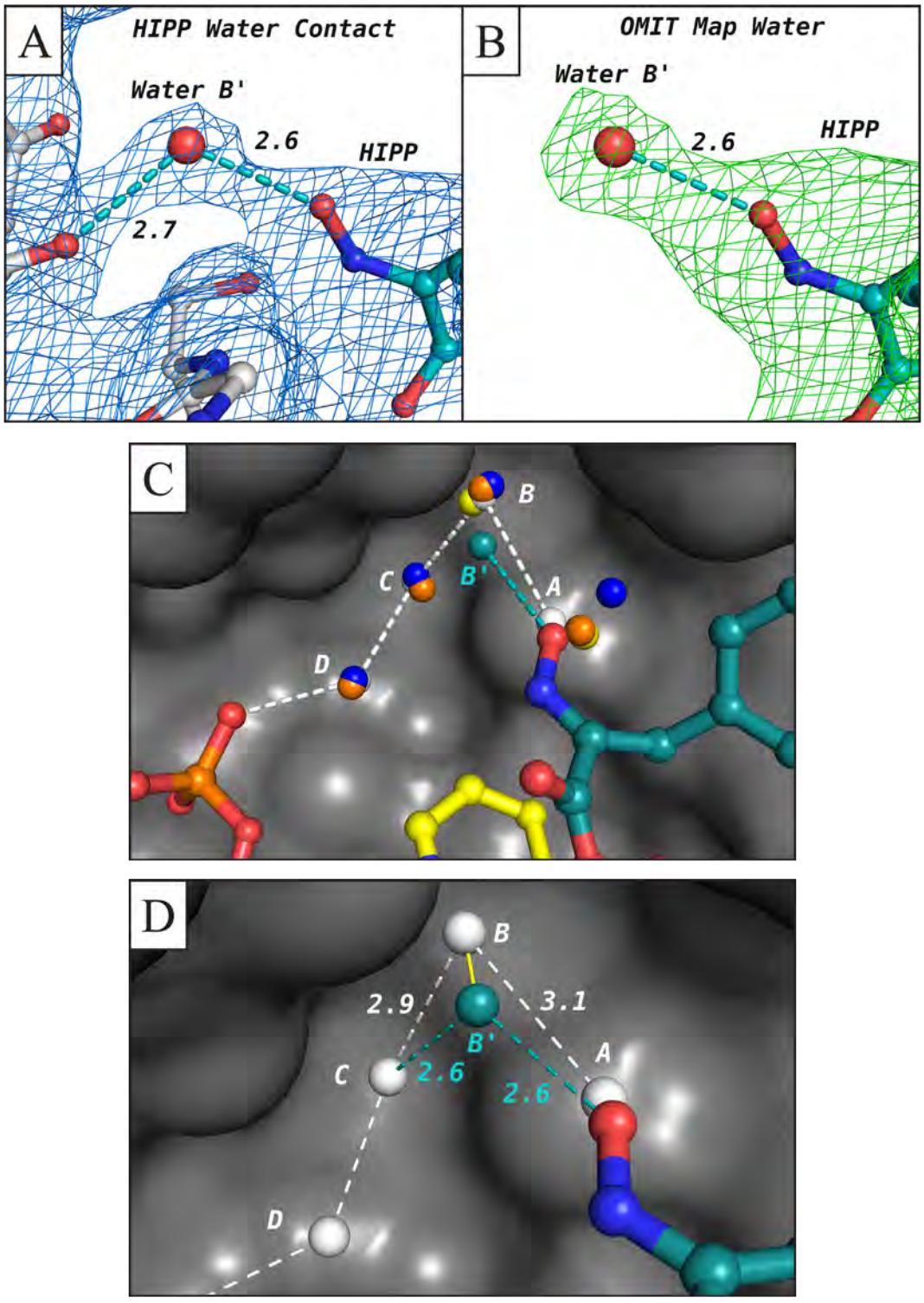
Experimental analysis of MTOB, phenylpyruvate, and HIPP binding affinity

Binding experiments were conducted first with MTOB and phenylpyruvate and finally with HIPP to determine their affinity for CtBP. The CtBP2 construct utilized for co-crystallization with MTOB possessed higher solubility than the CtBP1 construct and, thus, was used for isothermal titration calorimetry (ITC) experiments. An extensive survey of conditions for an ITC compatible buffer in which CtBP2 was stable eventually led to some success in determining the binding affinity for each small molecule tested.

The weak affinity of MTOB for CtBP2 caused difficulties in the experimental design. Typical successful ITC experiments require high affinity and solubility of reactants, which result in a sigmoidal binding isotherm when kcal/mol of injectant (ligand) is plotted versus the molar ratio of ligand to total protein concentration. The

Figure 4.10: The effect of HIPP on the CtBP water network. (A) The $2F_o$-F_c density contoured to 1 σ (0.16 e-/Å^3) shows a water molecule forming hydrogen bonds (teal dashes; distances in Å) with the backbone carbonyl of Val120, and the HIPP. (B) The F_o-F_c omit map contoured to 3 σ (0.21 e-/Å^3) confirms the presence of a water molecule. (C) Shown are full CtBP1 water networks (positions A through D) in MTOB – CtBP1 (white), binary CtBP1 (blue, pdb 1MX3) (*10*), and rat CtBP1 (orange, pdb 1HKU) (*39*) against the surface of CtBP1 (grey). The partial network (molecules A and B) is included from another rat CtBP1 structure (yellow, pdb 2HU2) (*44*). CtBP2 waters were omitted for clarity. The HIPP – CtBP1 hydroxyimino group displaces the water at position A, and shifts water molecule B (teal) into a new position, labeled B'. (D) A closer view of the MTOB – CtBP1 network (white) and the HIPP – CtBP1 water (distances indicated in Å). Position B' shifted about 1 Å from position B (yellow line), forming a short hydrogen bond with the HIPP. The proximity of position B' to C may shift or displace water molecule C. HIPP – CtBP1 water molecules C and D were not modeled due to their absence or the structure's low resolution.

Figure 4.10


A
HIPP Water Contact
Water B'
2.6
HIPP
2.7
B
OMIT Map Water
Water B'
2.6
HIPP
C
B
C
B'
A
D
D
B
2.9
3.1
B'
C
2.6
2.6
A
D

c value, defined as $c = n[M](1/K_d)$ where *n* is the number of binding sites or binding stoichiometry, [M] is the total protein concentration, and K_d is the dissociation constant, helps define the quality of the sigmoidal isotherm. Typically a value of $10 \leq c \leq 500$ results in a curve where one can accurately fit ΔH, K_d, and *n* (*206*). However, in the case of low affinity binding ($c < 1$), the three parameters cannot be fit as the curve appears nearly flat in the typical window of zero to two or three molar equivalents of ligand relative to protein. Extending the molar ratio window, by injecting higher concentrations of ligand (molar ratio window increased to ~30 equivalents), results in a hyperbolic curve where ΔH and the K_d can be fit, but only when *n* is fixed (*206*, *207*). However, if *n* is known, and the protein and ligand concentrations are accurate, the binding affinity can be confidently determined. MTOB affinity, due to the low signal upon binding, was measured through the low-*c* value approach, fixing $n = 1$ during curve fitting. The resulting K_d was calculated to be 2.38 ± 0.09 mM, confirming that MTOB only weakly binds CtBP **(Figure 4.11)**.

Binding affinity of phenylpyruvate with CtBP2 was stronger than MTOB, but still relatively weak. Typical experimental ITC setup and processing resulted in a K_d of approximately 80 μM **(Figure 4.11)**. Only an approximation is used here because this value was calculated from a single experiment with buffer conditions distinct from MTOB, and without having a proper ligand heat of dilution control. These experiments have not yet been successful due to low active protein concentration in the most recent attempts. The low c value ($c = 1.8$) indicates the curve is outside the optimal binding range. However, the obvious stronger binding signal relative to MTOB suggests phenyl

group resulted in a significant gain of affinity demonstrating our approach of aromatic ring stacking with Trp318 would be successful.

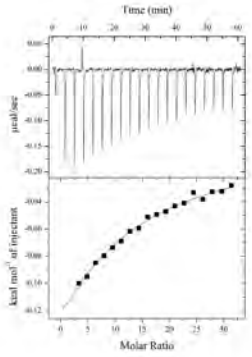
With full HIPP occupancy in the CtBP1 active site, not observed with either MTOB or D-mandelic acid, we expected a high HIPP affinity for CtBP2. ITC trials immediately yielded results showing HIPP indeed binds much more tightly to CtBP2 than MTOB or phenylpyruvate. In buffer containing 1.5 mM NAD^+, where CtBP2 is presumed to be dimeric, the HIPP K_d was measured at 1.12 ± 0.13 μM. In buffer without NAD^+ the K_d was determined to be 0.93 ± 0.05 μM **(Figure 4.11)**. These values denote an improvement in affinity of greater than three orders of magnitude over MTOB (identical conditions to the NAD^+ free binding experiment), and an approximately 80-fold increase in affinity relative to phenylpyruvate. MTOB binding appears enthalpically driven, with an unfavorable entropic contribution to the free energy. Favorable enthalpic and entropic interactions contribute to HIPP binding free energy, although entropic interactions dominate binding. The shift towards entropic binding suggests the release of ordered water molecules from the hydrophobic surfaces of HIPP, as well as release of solvent molecules from the CtBP active site.

Despite the tight binding of HIPP, the oligomeric state of CtBP2 in these conditions needs to be verified in future trials. The protein used in these experiments was either largely inactive, or substantial negative cooperativity is present, as indicated by a molar ratio close to $n = 0.5$ before adjusting it to $n = 1$ (1:1 binding). The low molar ratio was observed in every trial, even those without NAD^+ where CtBP2 should be monomeric and, therefore, unlikely to exhibit any cooperativity, negative or positive.

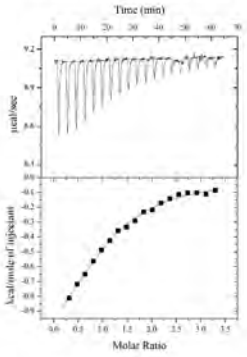
Figure 4.11: HIPP affinity for CtBP2 is higher than phenylpyruvate or MTOB. Isotherms from ITC experiments performed with MTOB, phenylpyruvate, and HIPP. The weak binding of MTOB (low *c* value) necessitated using a large molar excess of ligand. The molar ratio was fixed to $n = 1$ during curve fitting. The phenylpyruvate experiment was performed in different buffer conditions than the other experiments. The experiment lacked a heat of dilution control for the phenylpyruvate. Three separate experiments were performed with HIPP, without NAD^+ in the buffer. Protein and buffer were supplemented with NAD^+ during dialysis, and binding of HIPP was measured in four separate trials. The low molar ratio of the MTOB and HIPP experiments was treated as a result of inactive protein. Active protein concentrations were recalculated in the HIPP experiments so that $n = 1$.

Figure 4.11

MTOB



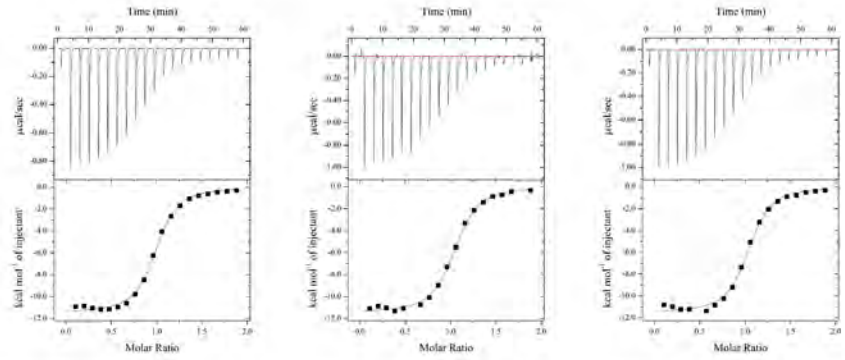
Phenylpyruvate



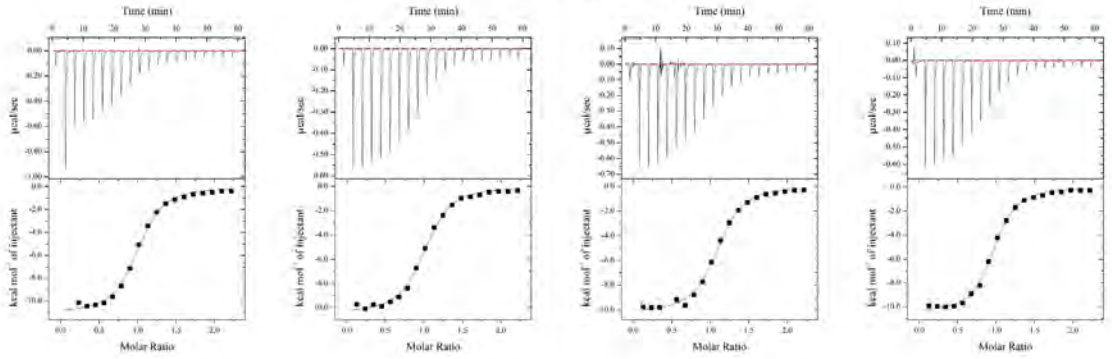
Molecule	Kd	ΔG (kcal)	ΔH (kcal)	-TΔS (kcal)
HIPP	0.93 ± 0.05 μM	-8.09 ± 0.03	-1.15 ± 0.004	-6.94
HIPP with NAD	1.12 ± 0.13 μM	-7.98 ± 0.06	-1.05 ± 0.05	-6.93
Phenylpyruvate*	79.37 ± 8.00 μM	-5.49 ± 0.06	-1.43 ± 0.11	-4.06
MTOB*	2.38 ± 0.09 mM	-3.51 ± 0.02	-4.51 ± 0.11	+0.99

*Measured from single experiment

HIPP



HIPP with NAD^+



Assuming negative cooperativity is not occurring, the low molar ratio would simply indicate the protein sample, which had previously been subjected to freeze thaw cycles and dialysis into multiple conditions, was roughly 50% inactive. This is the likely explanation given the following: (1) HIPP fully occupied the active site in the CtBP1 crystal structure, demonstrating both sites on a dimer can simultaneously be loaded with ligand; (2) A second binding event was never observed with different concentrations of HIPP; (3) The phenylpyruvate binding experiment, conducted with a different batch of purified protein, had a molar ratio of close to $n = 1$. Although a reduced activity of the protein is the likely explanation, the possibility exists that if a significant population of CtBP2 was able to dimerize in the absence of NAD^+, and negative cooperativity was occurring, then the molar ratio would be $n = 0.5$. Residual NADH from purification, or concentration dependent oligomerization could account for an unexpected oligomeric state. However, CtBP is incubated with pyruvate during purification to convert NADH to the weaker binding NAD^+. When done properly, this protocol removes NADH from CtBP (*26*). Regardless of the explanation for the low molar ratio, the measured affinity constant should remain virtually unchanged. The K_d does not change whether the curve is fit to $n = 0.5$ with fully active protein, or the protein concentration is adjusted so $n = 1$. Even if only half of the protein were active, this would be enough to accurately measure the K_d. Repeating binding will therefore only serve to rule out cooperativity.

DISCUSSION

The co-crystal structures of MTOB bound CtBP1 and CtBP2 proved instrumental for identifying higher affinity molecules that target the CtBP active site. We identified compounds predicted to interact with active site residue Trp318/324 while maintaining the contacts observed between the MTOB α-keto acid moiety and CtBP. The interaction with Trp318/324 would also confer specificity to any inhibitor minimizing cross reactivity with related human D2-HDH family members, GRHPR and PHGDH. Residues homologous to Trp318/324 in GRHPR and PHGDH are hydrophilic and positively charged, respectively, which contribute to substrate specificity. GRHPR hydrophilic residues would likely deter binding with hydrophobic/aromatic groups chosen to interact with Trp318/324. Although GRHPR residues could potentially form favorable cation-pi interactions, the multiple arginine residues designed to stabilize the natural substrate's phosphate group would likely cause some repulsion with an aromatic group if positioned near the edge of the ring.

The two molecules that seemed to best fit the desired criteria were phenylpyruvate and D-mandelic acid. Phenylpyruvate, a proven substrate (*63*), and D-mandelic acid, an enzymatic product mimic identified through a small molecule library screen, would potentially stack their aromatic rings with the Trp318/324 indole. The differing length between their α-keto/α-hydroxy acid "core" and the phenyl ring would help identify the ideal compound for further development. Co-crystallization with D-mandelic acid demonstrated the phenyl ring did indeed interact with Trp318/324 as predicted. However, D-mandelic acid failed to fully occupy the active site, and the electron density was too

poor to confidently model the carboxylate group, suggesting weaker than desired binding. The docked phenylpyruvate structure implied the aromatic ring would likely assume a distinct conformation relative to D-mandelic acid, possibly strengthening the interaction with Trp318/324.

While attempts continued to structurally and biochemically characterize phenylpyruvate binding, we began examining binding of 2-hydroxyimino-3-phenylpropionic acid (HIPP); a molecule that would mimic the transition state of catalytic conversion from phenylpyruvate to D-phenyllactate. Although this molecule was intended to mimic the conformation of the transition state within the CtBP active site, the additional volume of the HIPP hydroxyimino group, relative to the typical substrate carbonyl, forced HIPP to adopt an unexpected conformation distinct from any bound substrate in the D2-HDH family. Despite the unexpected conformation of the carboxylate and hydroxyimino group, the HIPP aromatic ring remained in a position to interact with Trp318 indole.

Experimental measures of the affinity of MTOB, phenylpyruvate, and HIPP proved that increasing interactions with Trp318/324 drastically increased ligand affinity. Phenylpyruvate affinity improved roughly 30-fold relative to the millimolar dissociation constant of MTOB. HIPP binding increased more considerably, with an affinity roughly 2000-fold stronger than MTOB. The K_d values correspond very well with unpublished enzymatic CtBP-HIPP K_I data measured by Grossman et al. Although the unique conformation of HIPP relative to phenylpyruvate most likely drives the difference in affinity, it is unclear why MTOB, D-mandelic acid, or any other substrate would not

assume the alternative confirmation as well. Utilizing ITC displacement assays (*208*), a more accurate affinity and thermodynamic profile will be measured for MTOB, phenylpyruvate, and D-mandelic acid.

HIPP binding also altered the CtBP water network, suggesting that displacing the water molecules with inhibitor moieties could improve future inhibitor binding. The HIPP hydroxyimino group completely displaces a water molecule (position A) found in other CtBP structures. The hydroxyimino group also shifts the position of the second water molecule in the network (position B), forming a hydrogen bond with HIPP. Movement of water molecule B away from the position observed in the other CtBP structures implies future inhibitor moieties may be able to displace it completely without an excessive energetic penalty. Water molecules at positions C and D were not observed in the HIPP structure, possibly due to the low resolution of the crystallographic analysis. The shift in water molecule B position places it in a position that would likely cause a shift or complete displacement of a water molecule at position C. If the water network can be displaced, new inhibitors can be produced that not only bind in the MTOB binding site and adjoining cavity, but also connect to moieties that take advantage at the NAD(H) binding site. This strategy would increase inhibitor affinity and maintain CtBP specificity, while competing for NADH binding, thus disrupting CtBP dimerization.

As the process of identifying and designing tighter binding, CtBP specific molecules continues, many questions still remain with regards to the mechanism behind MTOB mediated CtBP inhibition. Ligand binding to CtBP, including MTOB, fails to induce any large conformational changes, ruling it out as a mechanism of inhibition. A

full discussion regarding the possible mechanism of MTOB mediated inhibition and its impact on future inhibitor design is presented in Chapter V.

Chapter V

Discussion

Implications from CtBP co-crystal structures

The results reported in this dissertation provide the framework for the structure based design and development of potent inhibitors to pharmacologically target CtBP function in cancer. The first goal along the path to improving CtBP inhibition was to structurally characterize substrate MTOB binding to CtBP in order to gain insights into the mechanism of inhibition, as well as analyze specific contacts that may be exploited in inhibitor design. As described in Chapter III, we succeeded in co-crystallizing MTOB with both CtBP family members. Our original hypothesis of MTOB induced conformational change proved to be incorrect as there was no significant difference in conformation between the MTOB bound and unbound structures (Figure 3.2C and Figure A.3).

Our current hypothesis regarding the inhibition mechanism of MTOB is that enzymatic turnover of MTOB, which oxidizes NADH to NAD^+, drives inhibition of CtBP transcriptional regulation. Coenzyme binding induces dimerization (*27*), a change in oligomeric state necessary for most of CtBP's transcriptional regulation functions (*15, 29*). By preferentially binding NADH over NAD^+ (*26*) CtBP acts as a redox sensor, activating transcriptional regulatory functions when changes in a cell's metabolic state perturbs the NADH/NAD^+ ratio (*28, 30*). Catalysis oxidizes NADH, converting it to NAD^+, potentially leading to dissociation as coenzyme affinity decreases 100 fold upon oxidation (*26*). Upon coenzyme dissociation, the CtBP dimer would also dissociate, disrupting CtBP transcriptional regulatory function. MTOB may therefore sabotage CtBP redox sensing ability, inhibiting CtBP oligomerization and blocking complex formation.

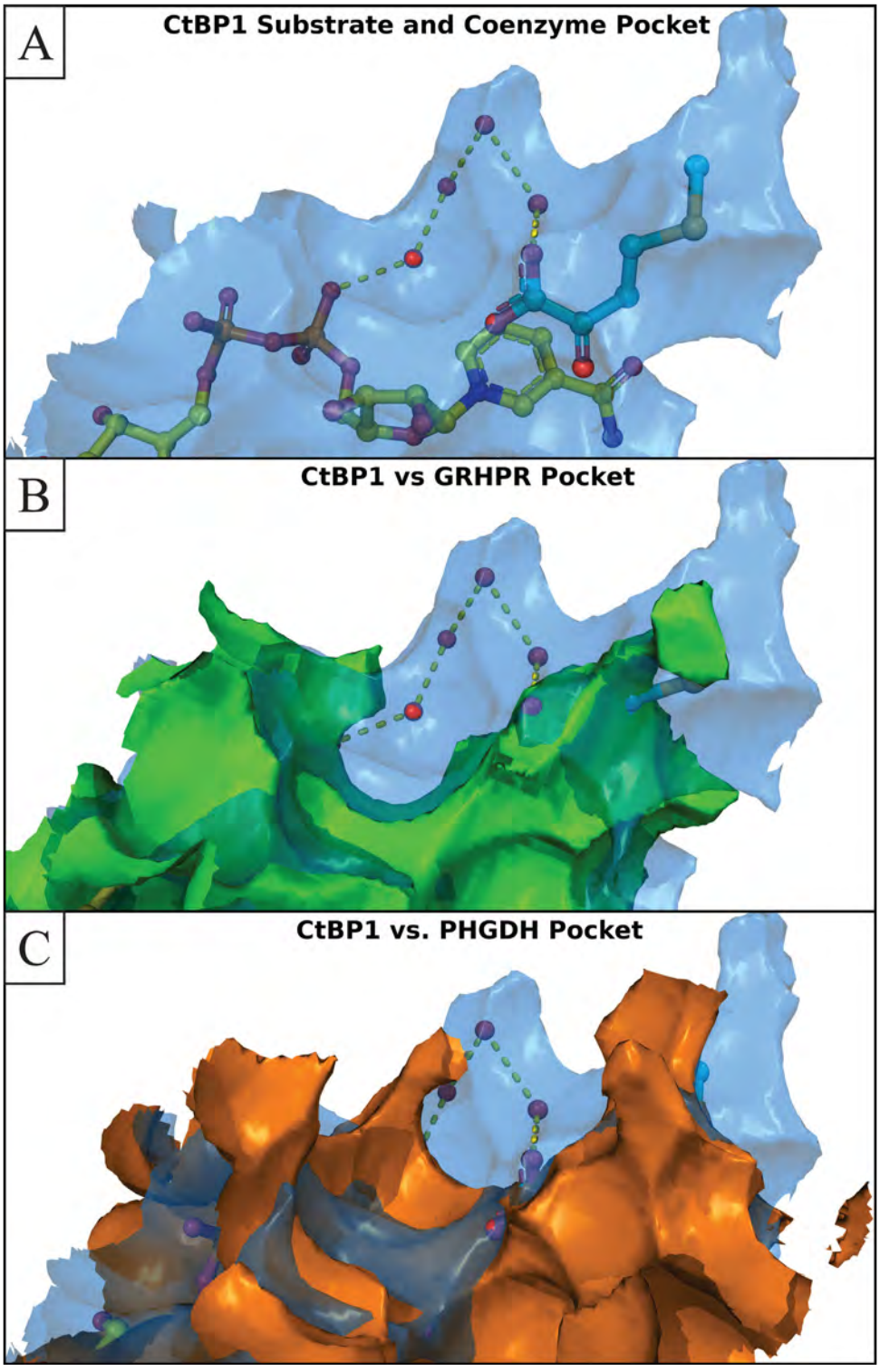
The interactions of MTOB with CtBP highlight important features of the active site. The α-keto acid moiety assumed a conformation similar to other D2-HDH family member substrates (*20, 21, 24*). The remaining MTOB atoms, which confer an 80-fold increase in substrate turnover relative to pyruvate (*63*), contact a number of active site residues, with the largest interaction occurring with Trp318/324. This side chain is unique among human D2-HDH family members, suggesting that targeting the active site tryptophan may confer specificity to CtBP.

A second key structural feature of CtBP is a water filled cavity adjoining the active site. Previous CtBP structures (*10, 39, 44, 90*) contain a portion or full network of four water molecules filling this cavity. It was not until co-crystallization with MTOB that we realized that this cavity and water network was exclusive to CtBP. Other D2-HDH family members possess a large amino acid that severs the continuity of the cavity volume by capping the active site and contacting the substrate **(Figure 5.1)**. In the MTOB – CtBP1 complex the four water molecules form a network that bridges MTOB with an NAD^+ phosphate. This observation spawned the idea of incorporating the cavity and, eventually, NAD(H) moieties into inhibitor design.

After achieving our first goal, we utilized the CtBP co-structures to identify and design new molecules that would more strongly bind the CtBP active site, forming the intended critical contacts observed in MTOB. As described in Chapter IV, our focus quickly led to α-keto/α-hydroxy acid compounds with aromatic rings, selected to maximize the interaction with Trp318/324. The critical variable in these initial compounds was the length of the carbon linker between the carbonyl/hydroxyl C2 carbon

Figure 5.1: The unique cavity in the CtBP active site will provide specificity for inhibitors. (A) The active site cavity and NAD(H) binding site represented by a surface map (blue) generated by the Sitemap program (*209-211*). MTOB (cyan), NAD^+ (yellow), and the water network (red spheres) are shown within the cavity. Panels (B) and (C) show alignment with the GRHPR and PHGDH active sites and coenzyme binding pockets (green and orange surface respectively). The CtBP water network cavity does not exist in the either protein (discussed in Chapter III), providing specificity for inhibitors designed to fill this volume.

Figure 5.1

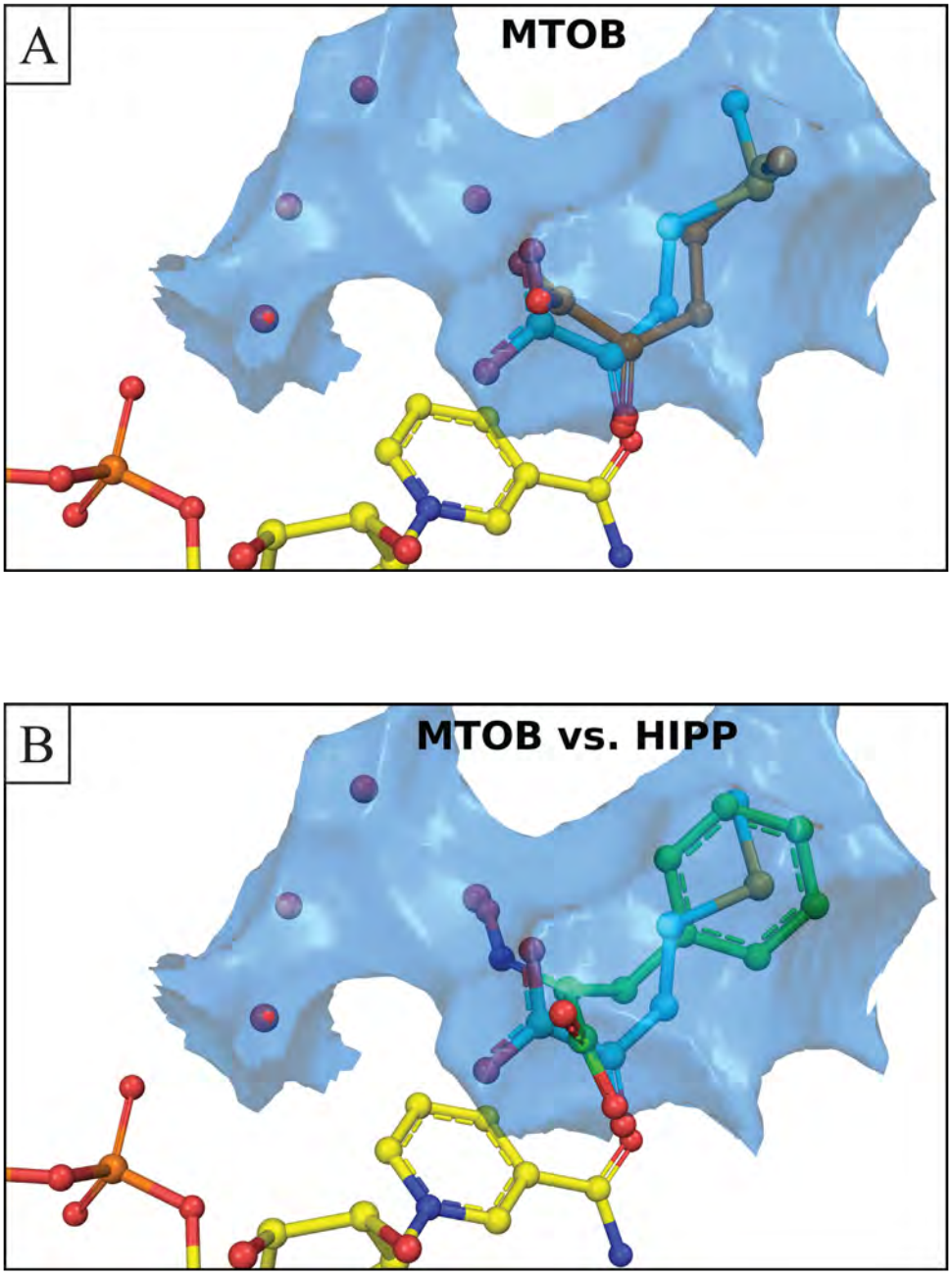

A
CtBP1 Substrate and Coenzyme Pocket
B
CtBP1 vs GRHPR Pocket
C
CtBP1 vs. PHGDH Pocket

and the aromatic group. Our first co-structure, with D-mandelic acid achieved the intended interaction with Trp318/324 but with less occupancy and less than ideal aromatic packing. Our focus continued with phenylpyruvate, a compound with an additional methylene carbon linking the aromatic group to the C2 carbon. Because phenylpyruvate is an established CtBP substrate (*63*), we tested a transition state mimic, 2-hydroxyimino-3-phenylpropionic acid (HIPP), which substitutes a hydroxyimino group for the carbonyl.

Co-crystallization with CtBP1 and binding affinity studies with CtBP2 revealed that HIPP not only binds in an unanticipated conformation **(Figure 5.2A, B)**, but also shows vastly increased affinity relative to MTOB **(Figure 4.11)**. The flexibility introduced by the extra methylene carbon, relative to D-mandelic acid, allowed the HIPP phenyl ring to maintain contact with Trp318 despite the new contacts with other active site residues. Initial ITC studies measured the HIPP dissociation constant at about 1 μM. The HIPP affinity represents an 80-fold increase over phenylpyruvate binding, and more than three orders of magnitude improvement relative to MTOB. Our findings reveal that the phenyl ring increases affinity in both phenylpyruvate and HIPP, and that the novel conformation of HIPP further strengthens interactions relative to phenylpyruvate. Additionally, HIPP appears to have altered the CtBP1 water network, displacing at least one water molecule, and shifting the position of a second. These results support our proposed strategy of water displacement by inhibitor moieties as a feasible approach to enhance potency and specificity in CtBP inhibitor design.

Figure 5.2: MTOB and HIPP fill only a small portion of the active site. MTOB and HIPP are shown in the CtBP1 active site and water network volume (blue surface) with waters (red) and NAD^+ (yellow). (A) MTOB from the CtBP1 structure (cyan) and CtBP2 structure monomer A (orange) positioned in the active site. (B) HIPP (green) compared with CtBP1 MTOB (cyan) shows the novel conformation of HIPP. The hydroxyimino group positions towards the water network, displacing the first water in the network. Future inhibitors will contain additional chemical groups, further occupying the water network cavity. Displacing the entire water network would position inhibitor atoms adjacent to an NAD^+ phosphate.

Figure 5.2



Future directions

Next steps in this project involve improving existing data on the HIPP –co-crystal structure and re-measuring affinity of weaker binding molecules via ITC. The relatively low resolution (3.06 Å) of the initial HIPP – CtBP1 crystal structure limits detailed analysis of the binding interactions. Although we have confidence in the placement of HIPP in the active site, a higher resolution structure would provide greater accuracy for specific contacts and determine the full extent of the changes observed in the HIPP water network. Understanding the full impact of HIPP binding on the network will help future inhibitor design. Since far fewer crystals were screened to obtain the HIPP-CtBP1 crystal structure than those used for the MTOB crystal structures, it is likely that a higher resolution diffraction data set can be obtained. Furthermore, despite CtBP2 structures typically crystallizing at lower resolution, an HIPP – CtBP2 molecule would determine if the HIPP conformation is identical between the CtBP family members.

The initial attempts to measure MTOB and phenylpyruvate affinity with ITC were deterred by the low affinity. However, the very recent success with HIPP binding provides a means to accurately measure binding affinity of MTOB, phenylpyruvate, D-mandelic acid, and any other low affinity compound. By measuring the change in HIPP affinity in the presence of the weaker binding compound (competitive inhibitor), we can accurately calculate the affinity and thermodynamics of binding (*208*). Performing the HIPP binding experiments again with fresh protein will have the added benefit of recalculating the molar ratio, which was low in the first round of experiments. Although

CtBP2 tends to be more soluble, future ITC experiments should be conducted with both CtBP1 and CtBP2 to determine any differences between the CtBP proteins.

Future inhibitor design

Our analysis of CtBP with substrates and inhibitors has provided a framework for the development of tighter binding inhibitors. In order for such inhibitors to become lead compounds for therapeutic agents, these compounds must also interfere with CtBP co-transcriptional activity. If the hypothesis regarding enzymatic turnover as the mechanism for MTOB mediated CtBP inhibition is correct, then only enzymatic substrates can utilize this mechanism to inhibit transcriptional regulation. Our inhibitors, which will likely not be substrates, would need to act through an inhibitory mechanism distinct from MTOB.

Given the importance of dimerization in CtBP activity, we surmise that targeting dimerization will effectively inhibit CtBP transcriptional activity. Dimerization, which is induced by NADH binding (*27*), is required for CtBP transcriptional regulatory function (*15, 29*). Mutational studies have shown that impeding CtBP dimerization through disruption of the dimer interface or NAD(H) binding motif effectively reduces or eliminates CtBP transcriptional regulation (*10, 18, 212*). Small molecule targeting NAD(H) binding also inhibited CtBP transcriptional function (*213*). Therefore, we intend to target dimerization with inhibitors designed to compete for NAD(H) binding, effectively replicating the mutagenesis and NAD(H) inhibitor studies.

Using the HIPP – CtBP1 complex as a guide, the next generation of inhibitors will extend into the unique CtBP active site cavity, displacing the water network in the hinge region between CtBP domains to gain affinity and specificity. The energy required

to displace the water network, and the chemical characteristics of the inhibitor moiety will determine if the thermodynamics of new inhibitor binding becomes more or less favorable. Currently, we are docking molecules with different chemistry at the imine position of HIPP. This includes substitutions with an amine, ether, β-amino acid, or β-hydroxy acid to maintain the hydrogen bond observed in the HIPP structure between the imine lone pair and Ser100. We will also explore removing the hydrogen bond donor from this site by replacing it with a methylene carbon (**Figure 5.3)**. All of the different chemistries at the HIPP imine position will be linked to various chemical groups designed to fill the active site cavity. Among different groups considered, five-atom aromatic heterocycles are particularly appealing as their large surface fits snuggly within cavity making a number of contacts without creating collisions. Additional groups added onto the heterocycles fill the entire water network cavity, positioning near the NAD(H) nicotinamide phosphate. Recent computational experiments tested the ability of molecules containing different combinations of the above groups to dock into CtBP1 **(Figure 5.3)**. Many of the resulting compounds had a higher docking score than HIPP while maintaining critical contacts at Arg266 and Trp318. The water network cavity appears to accommodate these large heterocycles, suggesting various sized chemical groups could be synthesized to fit the cavity **(Figure 5.4)**. After exploring the ability of these molecules to displace the water network, we will begin addition of chemical groups to compete for NAD(H) binding moieties. Before discussing NAD(H) moieties in inhibitor design, we consider the effects of NAD(H) binding on domain arrangement and CtBP activity. Movement of the CtBP substrate binding domain and coenzyme binding

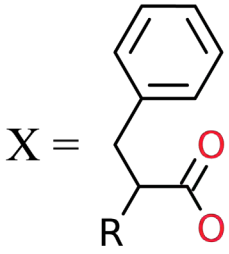
domain relative to one another may have a role in forming transcriptional regulatory complexes. The implications of domain movement in CtBP function must be factored into inhibitor design strategies to avoid causing activation rather than inhibition by small molecules designed to bind the active site and NAD(H) binding pocket.

The implications of domain movement on inhibitor design

Understanding the structural role of coenzyme binding to CtBP activation is paramount to the inhibitor design process, as this will impact the ability to develop inhibitors capable of disrupting dimerization and co-transcriptional activity. The structures of other D2-HDH family members exhibit movement of the substrate and coenzyme binding domains relative to one another, representing open and closed forms of the enzymes (*19, 21, 24*). To date, all structures of CtBP are in the closed form (*10, 39, 44, 90, 214*). Coenzyme binding has been linked to domain closure, a process that brings the substrate and coenzyme binding domains in proximity to form the active site (*19, 22*). Even ligands, or bound ions in the active site may contribute to domain closure and stability of the complex (*19, 22*). The energy barrier between open and closed conformation appears small as low energy crystal lattice contacts force some D2-HDH structure subunits into open conformations, reducing the ability of coenzyme and substrate to bind (*19, 21, 22, 24*). The current model of binding therefore begins with coenzyme and substrate binding to the coenzyme binding domain only of a D2-HDH protein in the open conformation. Equilibrium then favors domain closure, stabilized by weak electrostatic forces (*19, 21*). It remains unclear if domain closure is necessary for formation of active dimers since roughly 90 % of dimer interface forms between the

Figure 5.3: Recent docking with new compounds. In order to extend the inhibitors into the water network cavity, we have begun docking compounds with different chemistries into the CtBP active site. This figure serves as an example of the most recent docking experiment. Different R groups are attached to the C2 carbon of 3-phenylpropionic acid. Many of the new groups introduce chiral centers, allowing sampling of various stereochemistries during docking.

Figure 5.3



R Groups

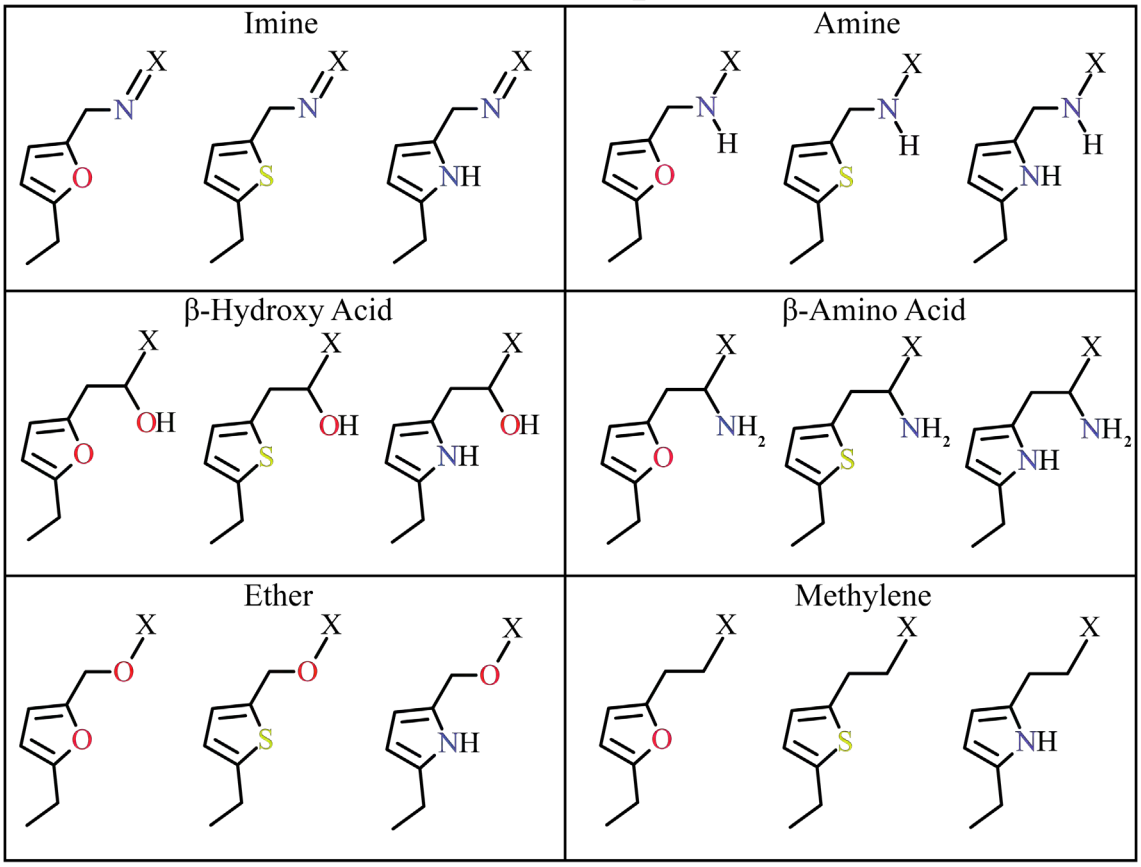
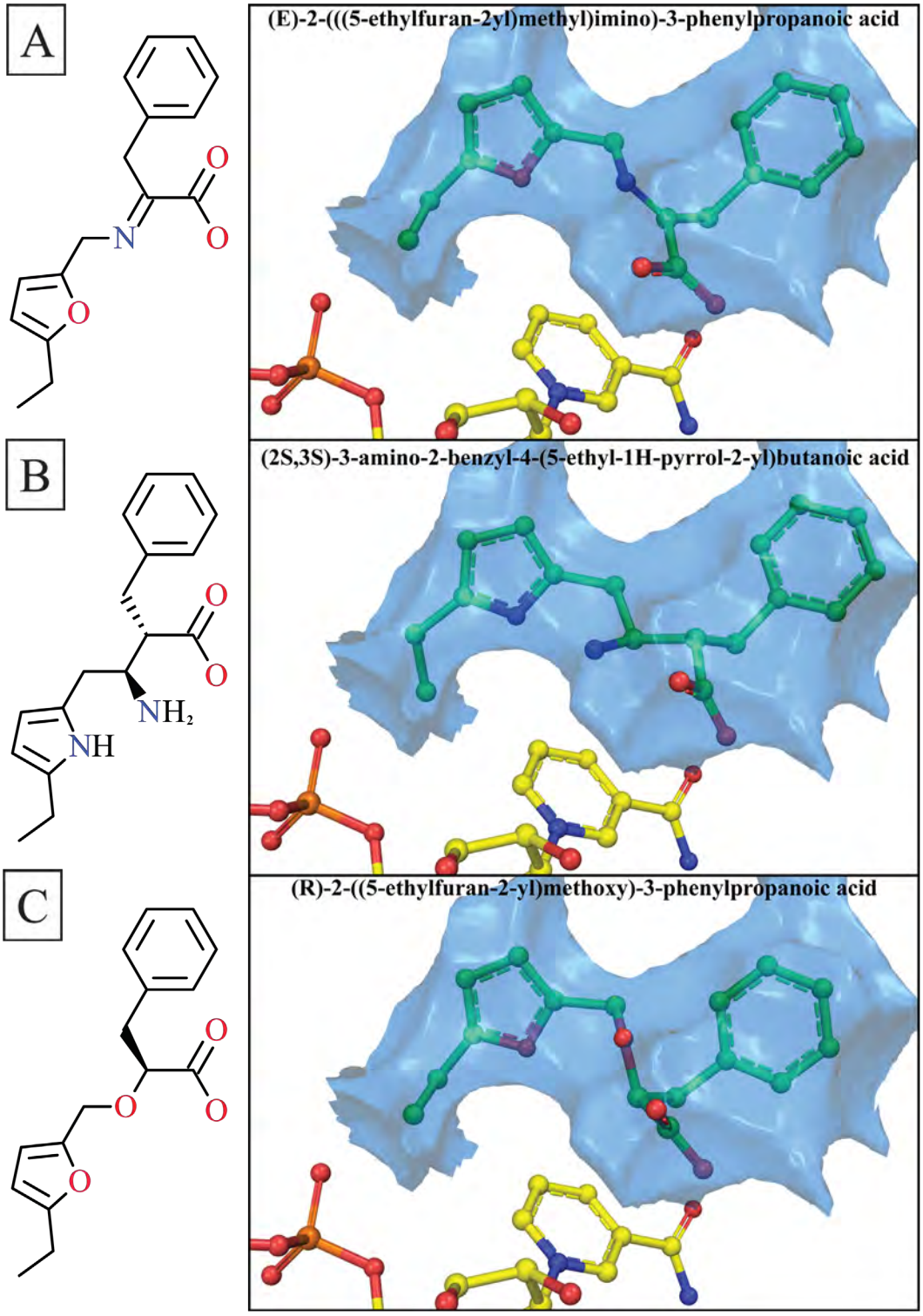


Figure 5.4: Docked compounds that would displace the water network. Three examples of compounds that dock well into the CtBP1 active site. The compounds fill the entire water network cavity. Future inhibitors could be expanded to include groups to compete with NAD(H) for binding.

Figure 5.4


A
(E)-2-(((5-ethylfuran-2yl)methyl)imino)-3-phenylpropanoic acid
B
(2S,3S)-3-amino-2-benzyl-4-(5-ethyl-1H-pyrrol-2-yl)butanoic acid
C
(R)-2-((5-ethylfuran-2-yl)methoxy)-3-phenylpropanoic acid

coenzyme binding domains of dimer subunits (*10, 21, 24, 39, 62*). However, the role of NAD(H) binding in dimerization and domain closure suggests a link between the two, such that domain closure may contribute to the stability of the dimer or otherwise impact transcriptional regulatory activity.

Although details of domain movement in CtBP are not yet well understood, information from other D2-HDH family members implies that domain movement must be factored into future inhibitor designs to avoid making favorable contacts that induce domain closure, as this could possibly increase CtBP transcriptional activity. One strategy would therefore be to design molecules that solely bind the coenzyme binding domain, avoiding formation of any inter-domain contacts. As an alternative strategy, inhibitors could be designed to force the domain equilibrium towards the open state, taking advantage of the low energy barrier between the open and closed states. The latter strategy would actively target known inter-domain contacts for disruption. Maintaining CtBP in an open conformation may turn out not to affect transcriptional regulatory functions, but we must be aware of the potential for complications caused by domain movement to modify inhibitor designs accordingly. Moreover, creating inhibitors that favor an open CtBP conformation will provide useful experimental tools for investigating the role of domain conformation in cell function and also possibly for structural studies with a stabilized open conformation of CtBP.

To investigate inhibitors that impact dimerization and domain orientation, we will add chemical groups to occupy the position of the NAD(H) phosphates. The NAD(H) atoms from the phosphates and the adenine moiety only contact the coenzyme binding

domain. Building inhibitor groups into the adenine-binding motif should significantly increase inhibitor affinity without inducing domain closure. Greater care must be taken when designing compounds that situate near the nicotinamide ribose ring, as this region contains the NAD(H) contacts bridging the substrate binding domain and coenzyme binding domain. Structures of open and closed D-LDH from *L. Bulgaricus* demonstrate a key inter-domain contact is made by coordination of a water by the NAD(H) nicotinamide ribose, a conserved substrate binding domain Asp residue, and the amide backbone of a coenzyme binding domain residue (*19*). A similar water coordination is present in CtBP. Inhibitors can be designed to avoid or disrupt this coordination, thus altering inter-domain contacts.

The other primary inter-domain contacts involve substrate binding in D2-HDH family members. MTOB binds the coenzyme binding domain of CtBP1/CtBP2 through catalytic Arg266/272 and His315/321, as well as specificity determinant Trp318/324. However, the substrate binding domain residues Gly99/105 and Ser100/106 form hydrogen bonds with the MTOB carboxylate, closing the active site **(Figure 3.7)**. S100/S106 also forms a single hydrogen bond with the HIPP imine lone pair. Thus the substrate bridges the two domains. Future inhibitor designs should minimize contacts with the active site loop to avoid facilitating domain closure, while exploiting contacts with the coenzyme binding domain residues in the active site. Fortunately, losing interaction with the active site loop does not appear to preclude ligand binding to coenzyme domain residues (*21*). Thus a strategy of focusing on coenzyme binding domain contacts while avoiding substrate binding domain contacts appears feasible.

Conformational change in the CtBP hinge region between domains may affect inhibitor binding in the open state. The water network in the unique CtBP active site cavity forms contacts with a section of the hinge region (CtBP1 residues 120-125). Our proposed strategy involves designing inhibitors to occupy this cavity in order to gain specificity for CtBP over other D2-HDH family members. A portion of the hinge region changes conformation in D-LDH proteins to create the movement of substrate binding domain relative to coenzyme binding domain (*19, 21, 24*). Inhibitors designed based on the closed conformation in the crystal structure may lose some contacts, particularly at CtBP 1/CtBP2 Val120/Ile126 and Pro121/127, upon domain opening. In D-LDH, movement of these residues affects water molecule coordination between the proline residue and an NAD(H) phosphate, equivalent to CtBP water position D (*19*). As the N-terminal portion of the hinge shifts, the CtBP cavity would become larger near water molecule positions C and D. Fortunately the conformational changes within the hinge are smaller in magnitude compared to the movements of the entire substrate binding domain. The hinge does not appear to have a large conformational change at key residues that convey specificity for CtBP (Ala123/129, Ser124/Ala130). Therefore inhibitors designed to bind CtBP would still be unable to bind other D2-HDH family members, even if the enzyme is in the open conformation. In summary, some inhibitor contacts may be lost in the hinge region upon opening, but moieties in this position should still convey specificity.

Determining the potency of new inhibitors

New inhibitors will be designed to enhance affinity by including binding in the NAD(H) binding motif, with the eventual goal of disrupting NAD(H) binding and dimerization. Improvements in newly designed inhibitors can be monitored *in vitro* by measuring dissociation constants with ITC, changes in CtBP oligomerization, and weakening of NAD(H) affinity. As with HIPP, ITC will measure the affinity of an new inhibitors for CtBP. Monitoring CtBP oligomerization through size exclusion chromatography (*27, 215*) will measure the efficacy of these newly designed inhibitors to block oligomerization in the presence of coenzyme. This assay will determine the ultimate effects of inhibitor binding regardless of whether domain movements have an effect on CtBP activity. Measuring changes in NAD(H) affinity will provide details on the competition between the inhibitor and coenzyme binding to CtBP. To date, we have been unable to measure NAD(H) affinity through ITC experiments or replicate NAD(H) binding data through Trp318 – NAD(H) FRET experiments (*26*). Although not critical, being able to measure affinity changes of NAD(H) for CtBP would help quantitate the effects of inhibitors on NAD(H) binding.

Concluding remarks

The results in this dissertation provide the foundation for developing inhibitors designed to pharmacologically target CtBP transcriptional regulatory functions in cancer. The benefit of targeting CtBP has become more apparent with each new study linking CtBP transcriptional regulatory activity to different types of cancer. Through structural studies we have elucidated the stereochemical basis for binding and important features of

the active site, facilitating the design of other potential inhibitors. New inhibitor designs will undoubtedly advance our understanding of the mechanism of CtBP inhibition on the pathway to creating an effective anti-cancer drug.

APPENDIX

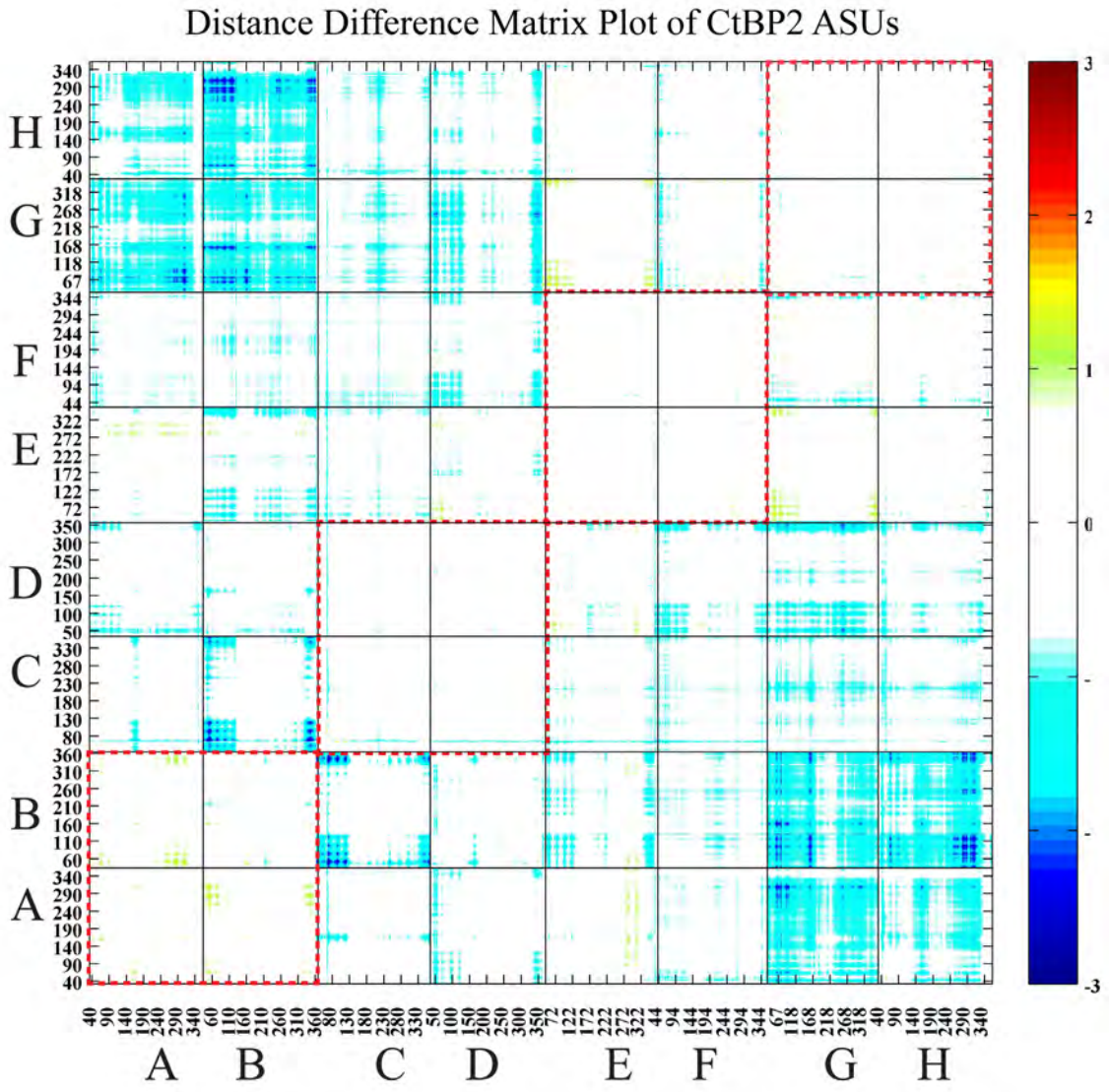
The effects of MTOB binding on CtBP conformation

As discussed in Chapter III, MTOB binding fails to induce a conformational change upon binding CtBP1 or CtBP2. Using distance difference matrix plots we carefully compared the MTOB bound CtBP1 and CtBP2 ternary complexes (CtBP/NAD^+/MTOB) to their respective binary (CtBP/NAD^+) structures. Our analysis concluded MTOB did not induce significant conformational changes. The first comparison between the MTOB bound CtBP1 dimer to the MTOB free (binary) dimer (*10*) revealed no significant differences **(Figure 3.2C)**. The equivalent CtBP2 analysis, discussed below, is more involved due to the eight monomers in the asymmetric unit.

In order to perform a similar analysis for CtBP2, we began by comparing all eight chains of the CtBP2 asymmetric unit between the MTOB bound (ternary) and MTOB free (binary) (*90*) complexes. This analysis provided a comparison between equivalent chains (A vs. A, etc.), equivalent dimers (AB vs. AB, etc.), and helped gauge the packing of the dimers relative to one another (AB position vs. GH position, etc.). The results show that the backbone of the individual monomers and dimer pairs does not exhibit any large changes **(Figure A.1**). However, the packing between the four dimer pairs slightly changes, resulting in the movement of the entire GH dimer 1-3 Å relative to AB dimer. These movements reflect a small reduction of the asymmetric unit volume in the MTOB bound CtBP2 structure relative to the binary structure. These packing changes and asymmetric unit dimensions may result from sequence differences in the remaining unstructured N-terminal tag sequence, or variation of the crystallization conditions between binary and ternary complexes.

Figure A.1: Comparing the asymmetric unit of the binary CtBP2 structure to the ternary structure. For clarity a grid separates individual monomers in the distance difference matrix plot. Equivalent monomer comparisons (Binary A vs. Ternary A, etc.) follow the diagonal from the origin. Equivalent dimers (AB vs. AB, etc.) are outlined by red dashes. Equivalent individual monomers and dimers do not exhibit large variations. However, comparing the positions of dimers relative to one another reveals changes in positions that culminate with a 1-3 Å shift of dimer GH relative to AB between the binary and ternary structures. The slight change in dimeric packing is reflected in a reduction of the unit cell dimensions differences of **a** = -1.6 Å, **b** = -1.0 Å, and **c** = -3.1 Å in the ternary structure. These packing differences may result from additional remaining N-terminal tag residues in the ternary complex relative to the binary complex.

Figure A.1


Distance Difference Matrix Plot of CtBP2 ASUs
A
B
C
D
E
F
G
H
3
2
1
0
-
-
-3

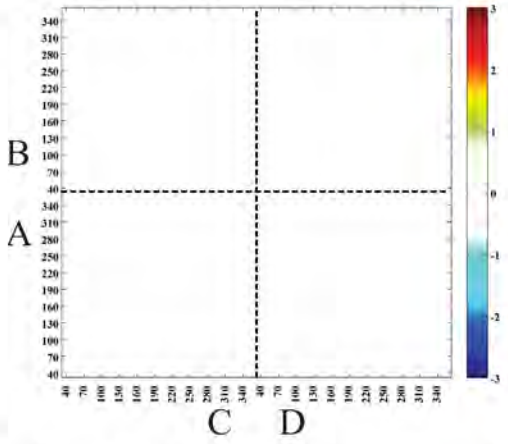
After comparing the entire CtBP2 asymmetric unit, we focused on ternary and binary CtBP2 dimer pairs to closely examine tertiary and quaternary structure differences. Analysis of the four dimers within the binary CtBP2 structure reveals almost no changes in Cα position, probably due to strong non-crystallographic symmetry (NCS) restraints used during refinement (**Figure A.2**). Since all binary complex monomers and dimer pairs were identical to one another, only the binary complex dimer composed of monomers A and B was used for comparison with each of the four MTOB-bound CtBP2 dimers. The binary to ternary dimer comparisons verify significant Cα variation does not occur upon MTOB binding in CtBP2, although Small variations are observed when comparing the binary CtBP2 dimer AB with the ternary AB dimer **(Figure A.3)**. The dimerization loop (148-178) of the ternary structure has shifted position in both monomers A and B relative to the binary structure. Monomer B displays additional small variations of coenzyme binding domain residues involved in an interface with monomer C (210-230) relative to a portion of substrate binding domain residues. These small variations sum to show a slight difference in domain packing of the binary CtBP2 AB dimer relative to the ternary AB dimer. MTOB binding however does not appear to be responsible for these differences as the remaining ternary dimers (CD, EF, GH) show only minute differences to the binary structure **(Figure A.3)**. As there are no consistent structural variations across dimers MTOB does not appear to induce changes upon binding CtBP2.

The individual monomers of the CtBP2 ternary complex were then compared against one another to examine conformational differences of the MTOB bound subunits.

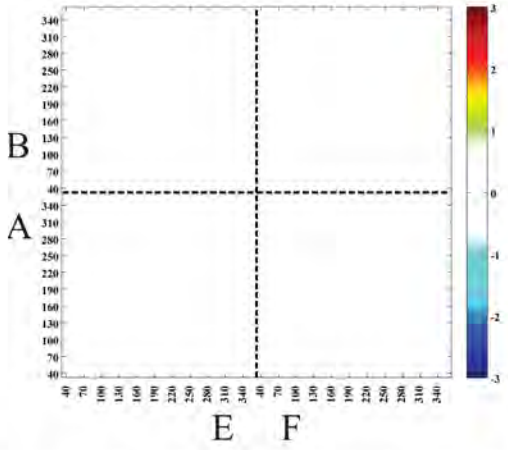
Figure A.2: The CtBP2 binary complex subunits are identical. Distance difference matrix plots show that there are no differences between the individual monomers or dimers of the CtBP2 structure. Therefore, CtBP2 binary complex dimer composed of monomers A and B was used for comparisons with CtBP2 ternary complex dimers.

Figure A.2

A. CtBP2 Binary Dimers: AB vs. CD



B. CtBP2 Binary Dimers: AB vs. EF



C. CtBP2 Binary Dimers: AB vs. GH

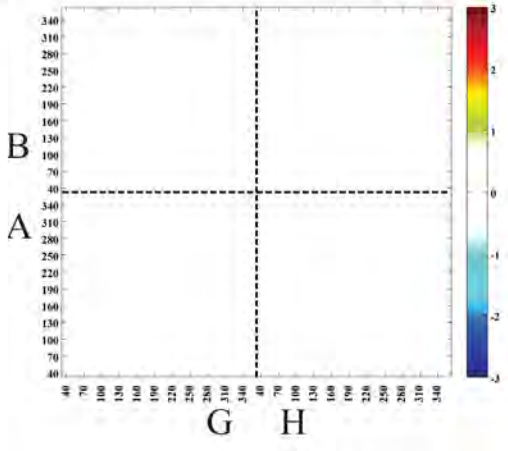
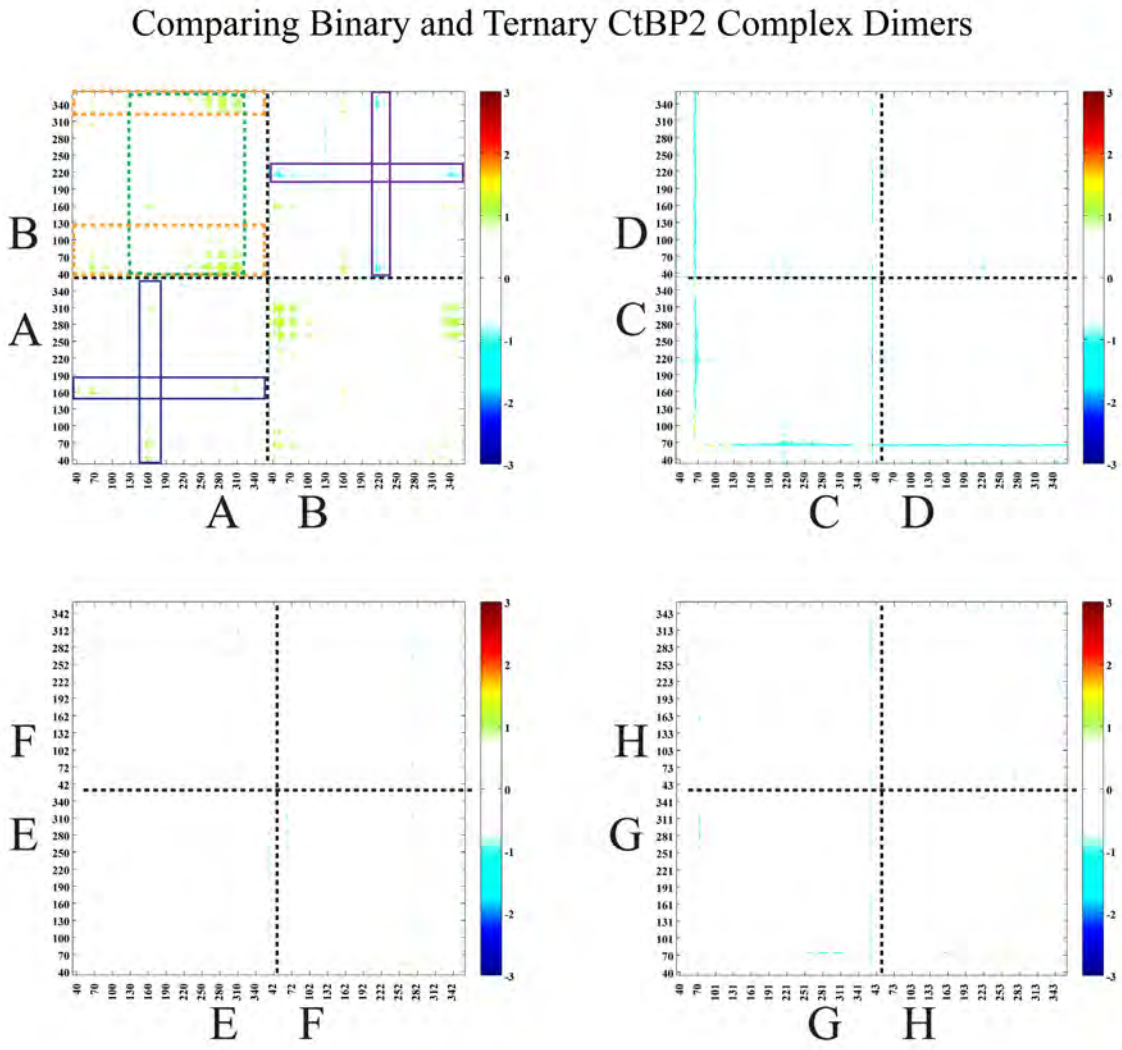


Figure A.3: MTOB does not induce conformational changes in CtBP2 dimers. CtBP2 binary and ternary dimer conformations were compared for Cα differences. The individual CtBP2 AB monomeric subunits and dimer display a small variation of Cα position. Monomers A and B show differences in the dimerization loop position (blue boxes). Monomer B displays additional variation at a crystal interface (purple boxes). Comparing Monomer A position relative to Monomer B, the quaternary structure exhibits a small shifts in portions of the Monomer B Substrate Binding Domain (orange dashes) and Monomer A Coenzyme Binding Domain (green dashes) between binary and ternary complexes. The differences in Cα position do not appear to be MTOB induced, as the remaining three dimer pairs do not exhibit any of the same shifts. The differences observed in monomer C are described in Figure A.4B.

Figure A.3


Comparing Binary and Ternary CtBP2 Complex Dimers
A
B
C
D
E
F
G
H

The differences observed between individual monomers within the ternary structure correspond to the small variation previously observed between binary and ternary CtBP2 dimers. Hence, there were no new significant changes within the ternary complex subunits to report. For the ternary complex analysis, monomer A Cα positions were compared directly to all other CtBP2 ternary complex monomers. Although most regions of the protein show little variation (**Figure A.4A**) two regions exhibit variances, with the most notable differences occurring at residues 63-66 in monomer C (**Figure A.4B**) and the dimerization loop in multiple monomers **(Figure A.4C)**. The backbone variation at residues 63-66 in monomer C, previously observed during comparisons of the binary and ternary dimers (dimer CD, Figure A.3), results from a unique ternary structure crystal lattice contact between monomer C and monomer H' of an adjacent asymmetric unit. The variability of the dimerization loop also mirrors a difference previously observed between binary and ternary (AB) dimers. The magnitude of the loop variation is not a consistent value between different monomers with the maximum difference observed between monomer A and E. These changes in magnitude suggest inherent dimerization loop flexibility causes it to assume a small array of conformations across the ternary structure asymmetric unit. Therefore, the observed changes in Cα positions are the result of altered crystal packing (monomer C), or flexibility of the dimerization loop conformation observed across different monomers. Other than these two regions, no single CtBP2 ternary complex monomer differs greatly from another, permitting accurate structural analysis across MTOB bound monomers.

Our final analysis of conformational differences was compared MTOB bound CtBP1 and MTOB bound CtBP2. Consistent with the very high degree of homology of the recombinant constructs, limited differences exist between MTOB bound CtBP1 and CtBP2 monomers. The first small differences observed occur at CtBP2 Gln214 (CtBP1 Ser208) where sequence variation in the dehydrogenase domain impacts lattice contacts **(Figure A.5A, B)**. Sequence variation at the C-terminus, CtBP2 Arg354 (CtBP1 Lys348), alters the side chain packing of the complementary surface of the Gln214/Ser208 lattice contact, facilitating the observed Cα shift. A second difference, observed at CtBP2 residues 48 and 54-55 results from similar sequence variation and crystal packing **(Figure A.5A, B)**. Although the sequence varies (CtBP2 Leu54; CtBP1 Val48), the greater impact on Cα position arises from a unique CtBP1 lattice contact involving these residues, for which there is no equivalent contact in the CtBP2 lattice. A third difference, limited to comparisons with monomer C, arises from a unique lattice contact conformation found only with ternary complex monomer C, and an adjacent asymmetric unit monomer H'. This is the same unique CtBP2 monomer C lattice contact responsible for the differences between CtBP2 structures in **Figure A.3 and A.4B**. Finally, because the dimerization loop changes position across different CtBP2 monomers (see **Figure A.3**), the same difference appears in comparisons of those CtBP2 monomers with CtBP1 **(Figure A.5C**).

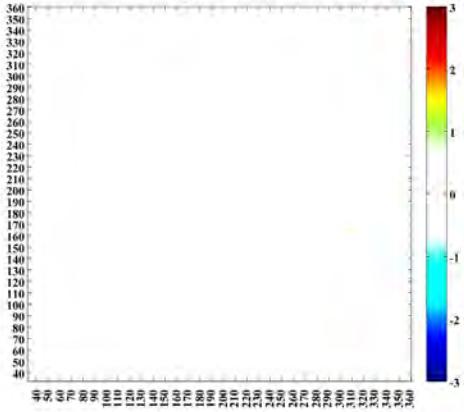
The distance difference matrix plot comparisons demonstrate that all the CtBP structures, regardless of the presence of MTOB, are essentially equivalent. The differences that are observed are the result of lattice contacts, sequence variation, and

backbone flexibility unrelated to MTOB binding. With no apparent conformational change upon MTOB binding, the mechanism of MTOB mediated CtBP inhibition remains unclear.

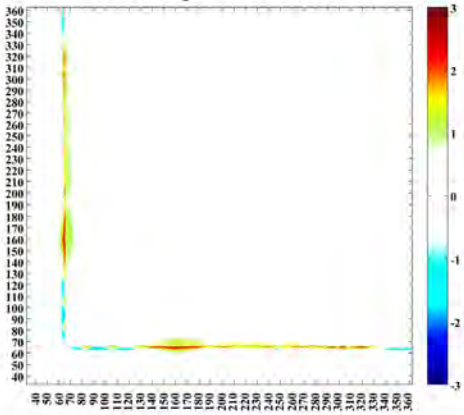
Figure A.4: The range of small differences between individual CtBP2 ternary structure monomers. (A) CtBP2 monomer A Cα position plotted against CtBP2 monomer H reveals no large variations occur between some of the subunits. (B) monomer A compared to monomer C reveals residues 63-66 show a sharp change in Cα position. A unique crystal lattice contact between monomer C and monomer H' of an adjacent asymmetric unit causes this variation compared to any other CtBP1 or CtBP2 monomer. (C) Variation between monomer A and monomer E shows the maximum differences observed in the "dimerization loop" (148-178) between subunits. Other CtBP2 monomers display varying patterns of dimerization loop movement relative to Monomers A and E, indicating the loop is flexible.

Figure A.4

A. CtBP2 Ternary Monomers: A vs. H



B. CtBP2 Ternary Monomers: A vs. C



C. CtBP2 Ternary Monomers: A vs. E

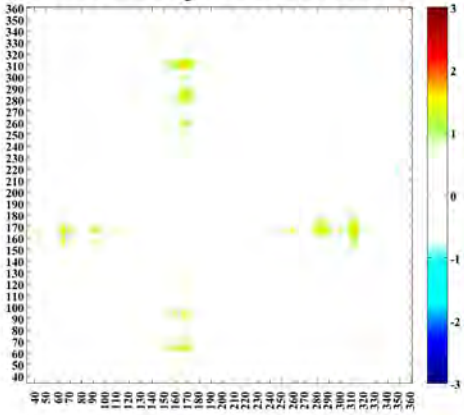
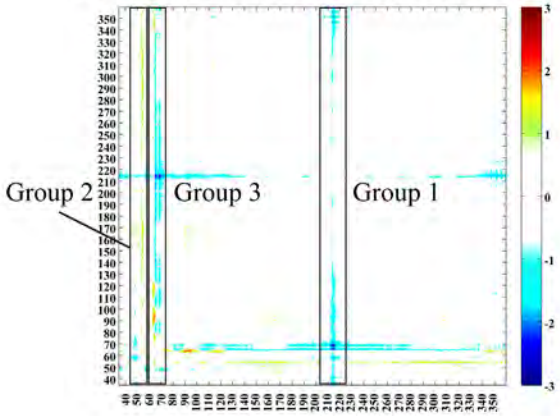


Figure A.5: Comparing MTOB bound CtBP1 and CtBP2 structures. For consistency, the same CtBP2 monomers in Figure A.4 are also used here. (A) The variance between CtBP2 Monomer C Cα position compared to CtBP1 (CtBP2 residue numbering) occurs at residue 214 and the extreme C-terminus (Group 1), residues 48 and 54-55 (Group 2), and residues involved a unique Monomer C lattice contact at residues 63-66 (Group 3). Sequence variation between CtBP1 and 2 (Q214S and R354K) results in Group 1 variation at a common lattice contact. Sequence variation (L54V) and a CtBP1 specific lattice contact causes Group 2 differences. See Figure 3.4 for an explanation of variation at Group 3. (B) The magnitude of Group 1 and 2 variation differs across CtBP2 monomers, seen here by a reduction in variation observed between CtBP1 and CtBP2 Monomer E. (C) In addition to Groups 1 and 2, CtBP 2 monomers show variation at the dimerization loop, residues 148-178 (Group 4), as observed here in Monomer A. MTOB bound CtBP1 and 2 exhibit only small differences, verifying the overall structures can be treated as equivalent.

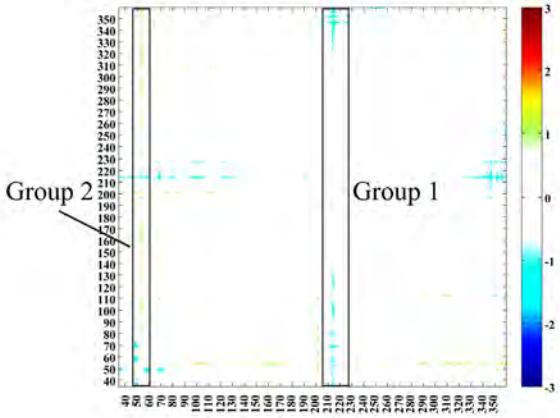
Figure A.5

Comparing MTOB Bound CtBP1 and CTBP2 Structures

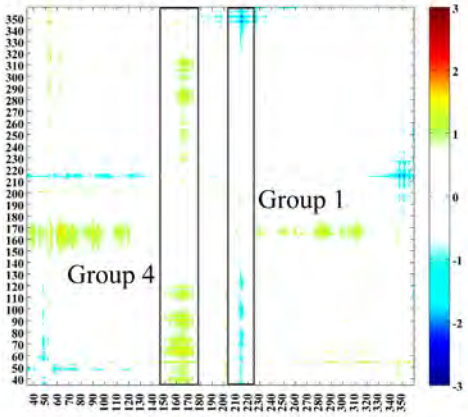
A. CtBP2 Monomer C vs. CtBP1



B. CtBP2 Monomer E vs. CtBP1



C. CtBP2 Monomer A vs. CtBP1



MTOB Placement in the CtBP2 Active Site

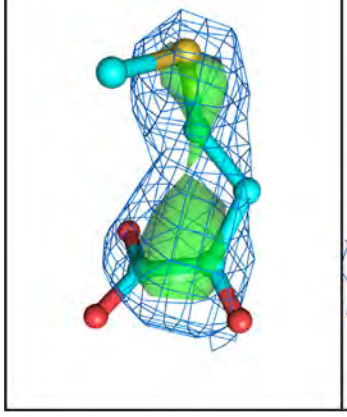
Chapter III details the methods used to place MTOB in the CtBP2 active site, describing the criteria used to select monomers A and C for evaluation of MTOB binding. This section displays the quality of the data of the other six subunits in the asymmetric unit. Specifically the $2F_o$-F_c density after refinement **(Figure A.6)** and the omit map F_o-F_c density **(Figure A.7)** are shown for monomers B, D, E, F, G, and H. Although the accuracy of placement in monomers A and C was superior, it is clear MTOB binds the active site of the other six monomers in a similar conformation. However, the accuracy of individual atoms in the most other monomers was questionable, and hence they were omitted from MTOB binding analysis. Monomer H is the exception, showing high quality density maps for MTOB equivalent to that observed in monomers A and C. Monomer H catalytic residue Arg272 assumes a conformation distinct from the other subunits **(Figure 3.6A)**, and was therefore omitted from binding analysis as well. Overall, monomers A and C the only two monomers that possessed high quality density maps for MTOB, and contained active site residues that fit within an NCS averaged map. As such, only these two monomers were selected for MTOB binding analysis.

Figure A.6: MTOB $2F_o$-F_c density in the CtBP2 active site. The $2F_o$-F_c density for the remaining six monomers not shown in Figure 3.6B. Maps are contoured to 1σ (blue mesh, 0.27 $e^-/Å^3$) and 1.6σ (green surface, 0.43 $e^-/Å^3$). All monomers display electron density around the sulfur atom. Monomer H has high quality density for MTOB, but active site residue Arg272 was not in accord with the non-crystallographic (NCS) averaged map (Figure 3.6A) leading to omission of monomer H from ligand analysis.

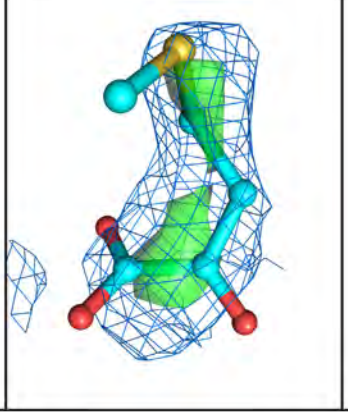
Appendix A.6

$2F_o$-F_c Map

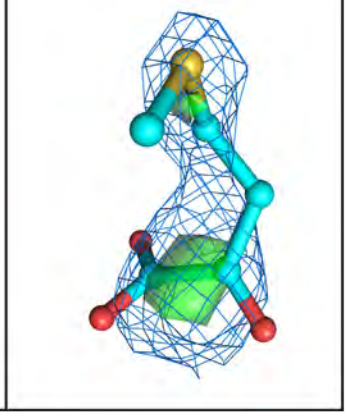
Monomer B



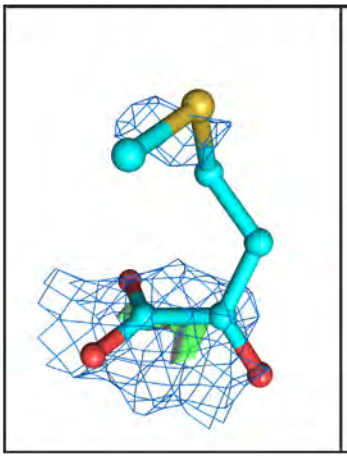
Monomer D



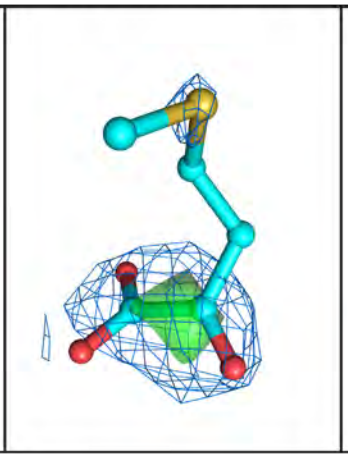
Monomer E



Monomer F



Monomer G



Monomer H

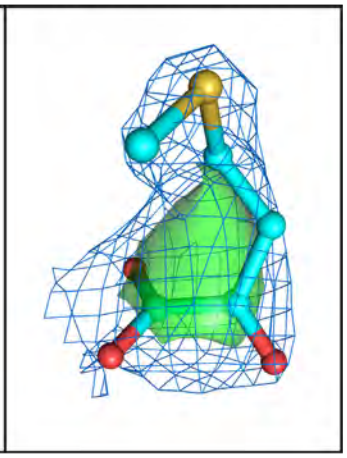
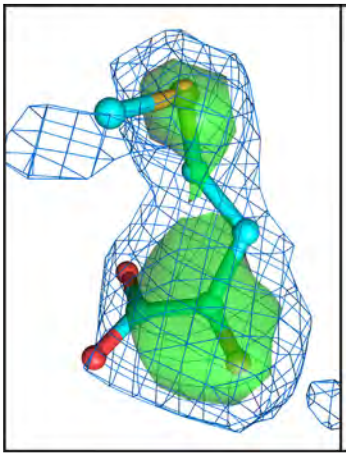


Figure A.7: MTOB omit map F_o-F_c density in the CtBP2 active site. The omit map F_o-F_c density for the remaining six monomers not shown in Figure 3.6C. Maps are contoured to 2.4σ (0.19 $e^-/Å^3$; blue mesh) and 3.5σ (0.27 $e^-/Å^3$; green surface). Ligand placement is generally in good agreement with the map density, with the poorest fit apparent in monomer G. Every other monomer displays density for the MTOB sulfur and α-keto acid group.

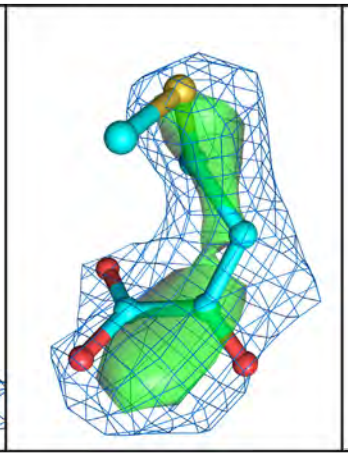
Appendix A.7

Omit F_o-F_c Map

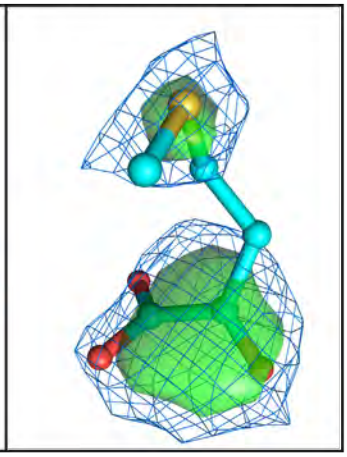
Monomer B



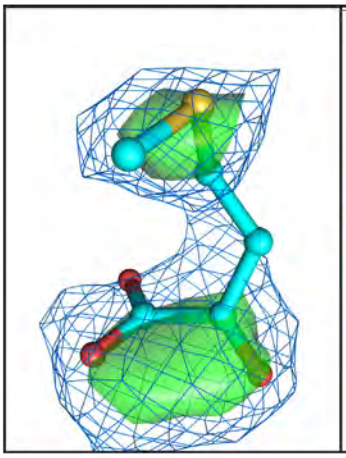
Monomer D



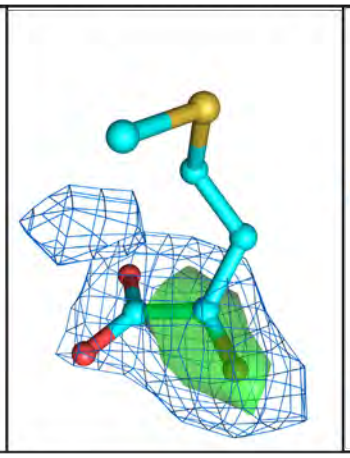
Monomer E



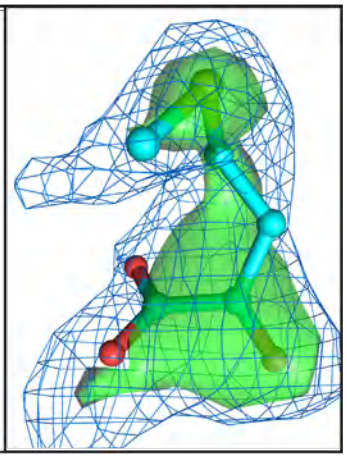
Monomer F



Monomer G



Monomer H



MTOB van der Waal analysis with repulsive terms

The van der Waal analysis of MTOB in the CtBP active site employed a Lennard-Jones potential to calculate the energy of interactions between individual atoms:

$$V_r = 4\varepsilon \left[(\sigma/r)^{12} - (\sigma/r)^{6} \right]$$

The Lennard-Jones potential, V_r, was calculated using the interatomic distance (r), depth of the well potential (ε), and the collision diameter (σ). The higher exponential term for the repulsive forces [$(\sigma/r)^{12}$] quickly dominates the lower exponential term for the attractive forces [$(\sigma/r)^{6}$] when the interatomic distance (r) becomes smaller than the collision diameter (σ). Hence, when atom pairs are equal to or just within the van der Waal radii, small variation across structures shows wide ranges of binding energies due to rapidly increase repulsive forces.

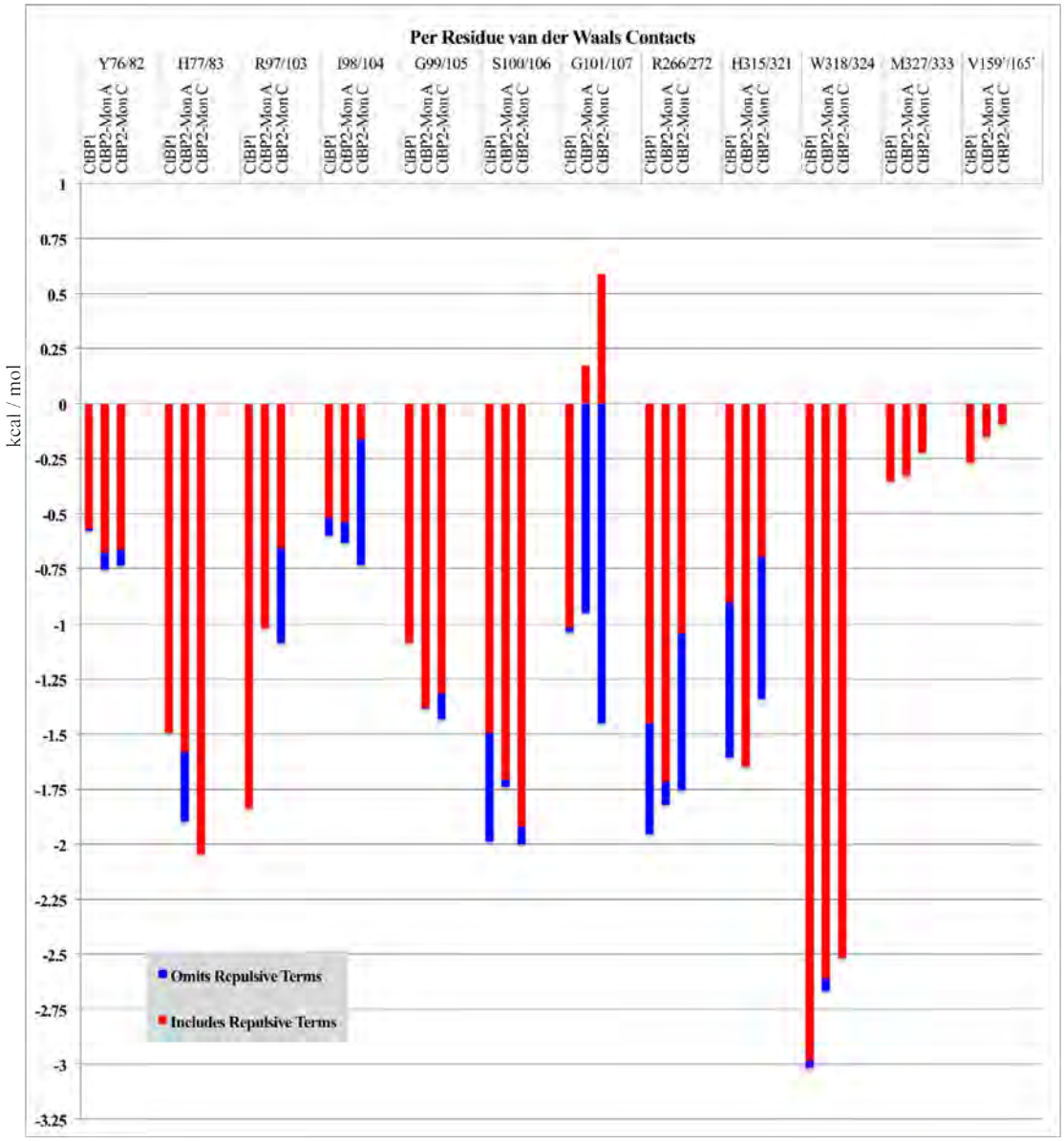
When applying the van der Waals analysis to MTOB in the CtBP1 and CtBP2 complexes most interactions are favorable (negative). Atoms involved in hydrogen bonds however, are typically closer together than the minimum van der Waal radii the program assigns for those atoms, thereby decreasing the favorable interactions. The van der Waal calculation is instead dominated by the repulsive term [$(\sigma/r)^{12}$]. When comparing hydrogen bonded atoms across different MTOB bound structures, slight atomic positional differences are amplified to large calculated differences in van der Waal potentials, making it appear that MTOB binding interactions vary greatly between structures. The positional variability inherent in moderate or low resolution structures exacerbates the problem, as the true positions of the atoms are less accurate. Very few clashes are

observed between non-hydrogen bonded atoms in the MTOB complexes suggesting the model is accurate.

We approached the problem of varying forces at hydrogen bonds by recording just the attractive forces, as reported in **Figure 3.9**, and also as the summation of attractive and repulsive forces here, in the appendix **(Figure A.8)**. By separating the attractive and repulsive values, we could identify where the repulsive terms were misleading due to tiny positional differences at hydrogen bonds. When repulsive terms were omitted, it became obvious that MTOB was in a similar position, with similar contacts in CtBP1 and CtBP2 monomers A and C. Including the repulsive terms, we see wildly varying values at the shortest hydrogen bonds, a trend observed when conducting similar analysis of other liganded D2-HDH family members (not shown). The repulsive terms at non-hydrogen bonded atom pairs however, was a good indicator of MTOB interactions, limited only by the accuracy of the crystal structure.

Figure A.8: Examining the repulsive van der Waals forces. The sum of attractive and repulsive terms is shown in red. Blue bars represent the final position of the binding energy calculated without the repulsive terms (values identical to Figure 3.9A). When no blue bars are apparent, there were no repulsive terms to remove. Repulsive (positive) values are occasionally obtained between atom pairs, usually involved in hydrogen bonds, which may place heavy atoms slightly within the van der Waals collision radius, such as the hydrogen bond at Gly101/107. Other smaller magnitude repulsions arise from model errors or dynamic atoms in these low to medium resolution CtBP complexes (I98/104). No large clashes (> 1 kcal) were observed, suggesting MTOB is properly situated in the active site.

Appendix A.8


Per Residue van der Waals Contacts
Y76/82
H77/83
R97/103
I98/104
G99/105
S100/106
G101/107
R266/272
H315/321
W318/324
M327/333
V159'/165'
CtBP1
CtBP2-Mon A
CtBP2-Mon C
kcal / mol
1
0.75
0.5
0.25
0
-0.25
-0.5
-0.75
-1
-1.25
-1.5
-1.75
-2
-2.25
-2.5
-2.75
-3
-3.25
Omits Repulsive Terms
Includes Repulsive Terms

BIBLIOGRAPHY

1. Berger, S. L. (2002) Histone modifications in transcriptional regulation, *Curr Opin Genet Dev 12*, 142-148.
2. Schaeper U, B. J., Uhlmann E, Subramanian T, Chinnadurai G. (1995) Molecular cloning and characterization of a cellular phosphoprotein that interacts with a conserved C-terminal domain of adenovirus E1A involved in negative modulation of oncogenic transformation, *PNAS 92*, 10467-10471.
3. Boyd, J. M., Subramanian, T., Schaeper, U., La Regina, M., Bayley, S., and Chinnadurai, G. (1993) A region in the C-terminus of adenovirus 2/5 E1a protein is required for association with a cellular phosphoprotein and important for the negative modulation of T24-ras mediated transformation, tumorigenesis and metastasis, *EMBO J 12*, 469-478.
4. Chinnadurai, G. (2002) CtBP, an unconventional transcriptional corepressor in development and oncogenesis, *Mol Cell 9*, 213-224.
5. Katsanis, N., and Fisher, E. M. (1998) A novel C-terminal binding protein (CTBP2) is closely related to CTBP1, an adenovirus E1A-binding protein, and maps to human chromosome 21q21.3, *Genomics 47*, 294-299.
6. Spano, S., Silletta, M. G., Colanzi, A., Alberti, S., Fiucci, G., Valente, C., Fusella, A., Salmona, M., Mironov, A., Luini, A., and Corda, D. (1999) Molecular cloning and functional characterization of brefeldin A-ADP-ribosylated substrate. A novel protein involved in the maintenance of the Golgi structure, *J Biol Chem 274*, 17705-17710.
7. Verger, A., Quinlan, K. G., Crofts, L. A., Spano, S., Corda, D., Kable, E. P., Braet, F., and Crossley, M. (2006) Mechanisms directing the nuclear localization of the CtBP family proteins, *Mol Cell Biol 26*, 4882-4894.
8. Birts, C. N., Harding, R., Soosaipillai, G., Halder, T., Azim-Araghi, A., Darley, M., Cutress, R. I., Bateman, A. C., and Blaydes, J. P. (2010) Expression of CtBP family protein isoforms in breast cancer and their role in chemoresistance, *Biol Cell 103*, 1-19.
9. Schmitz, F., Konigstorfer, A., and Sudhof, T. C. (2000) RIBEYE, a component of synaptic ribbons: a protein's journey through evolution provides insight into synaptic ribbon function, *Neuron 28*, 857-872.
10. Kumar, V., Carlson, J. E., Ohgi, K. A., Edwards, T. A., Rose, D. W., Escalante, C. R., Rosenfeld, M. G., and Aggarwal, A. K. (2002) Transcription corepressor CtBP is an NAD(+)-regulated dehydrogenase, *Mol Cell 10*, 857-869.
11. Rao, S. T., and Rossmann, M. G. (1973) Comparison of super-secondary structures in proteins, *J Mol Biol 76*, 241-256.
12. Nardini, M., Svergun, D., Konarev, P. V., Spano, S., Fasano, M., Bracco, C., Pesce, A., Donadini, A., Cericola, C., Secundo, F., Luini, A., Corda, D., and Bolognesi, M. (2006) The C-terminal domain of the transcriptional corepressor CtBP is intrinsically unstructured, *Protein Sci 15*, 1042-1050.
13. Bergman, L. M., Morris, L., Darley, M., Mirnezami, A. H., Gunatilake, S. C., and Blaydes, J. P. (2006) Role of the unique N-terminal domain of CtBP2 in determining the subcellular localisation of CtBP family proteins, *BMC Cell Biol 7*, 35.
14. Magupalli, V. G., Schwarz, K., Alpadi, K., Natarajan, S., Seigel, G. M., and Schmitz, F. (2008) Multiple RIBEYE-RIBEYE interactions create a dynamic scaffold for the formation of synaptic ribbons, *J Neurosci 28*, 7954-7967.
15. Zhao, L. J., Kuppuswamy, M., Vijayalingam, S., and Chinnadurai, G. (2009) Interaction of ZEB and histone deacetylase with the PLDLS-binding cleft region of monomeric C-terminal binding protein 2, *BMC Mol Biol 10*, 89.

16. Barnes, C. J., Vadlamudi, R. K., Mishra, S. K., Jacobson, R. H., Li, F., and Kumar, R. (2003) Functional inactivation of a transcriptional corepressor by a signaling kinase, *Nat Struct Biol 10*, 622-628.
17. Lin, X., Sun, B., Liang, M., Liang, Y. Y., Gast, A., Hildebrand, J., Brunicardi, F. C., Melchior, F., and Feng, X. H. (2003) Opposed regulation of corepressor CtBP by SUMOylation and PDZ binding, *Mol Cell 11*, 1389-1396.
18. Kuppuswamy, M., Vijayalingam, S., Zhao, L. J., Zhou, Y., Subramanian, T., Ryerse, J., and Chinnadurai, G. (2008) Role of the PLDLS-binding cleft region of CtBP1 in recruitment of core and auxiliary components of the corepressor complex, *Mol Cell Biol 28*, 269-281.
19. Razeto, A., Kochhar, S., Hottinger, H., Dauter, M., Wilson, K. S., and Lamzin, V. S. (2002) Domain closure, substrate specificity and catalysis of D-lactate dehydrogenase from Lactobacillus bulgaricus, *J Mol Biol 318*, 109-119.
20. Antonyuk, S. V., Strange, R. W., Ellis, M. J., Bessho, Y., Kuramitsu, S., Inoue, Y., Yokoyama, S., and Hasnain, S. S. (2009) Structure of D-lactate dehydrogenase from Aquifex aeolicus complexed with NAD(+) and lactic acid (or pyruvate), *Acta Crystallogr Sect F Struct Biol Cryst Commun 65*, 1209-1213.
21. Dengler, U., Niefind, K., Kiess, M., and Schomburg, D. (1997) Crystal structure of a ternary complex of D-2-hydroxyisocaproate dehydrogenase from Lactobacillus casei, NAD+ and 2-oxoisocaproate at 1.9 A resolution, *J Mol Biol 267*, 640-660.
22. Lamzin, V. S., Dauter, Z., Popov, V. O., Harutyunyan, E. H., and Wilson, K. S. (1994) High resolution structures of holo and apo formate dehydrogenase, *J Mol Biol 236*, 759-785.
23. Goldberg, J. D., Yoshida, T., and Brick, P. (1994) Crystal structure of a NAD-dependent D-glycerate dehydrogenase at 2.4 A resolution, *J Mol Biol 236*, 1123-1140.
24. Booth, M. P., Conners, R., Rumsby, G., and Brady, R. L. (2006) Structural basis of substrate specificity in human glyoxylate reductase/hydroxypyruvate reductase, *J Mol Biol 360*, 178-189.
25. Turnbull, A. P., Salah, E., Savitsky, P., Gileadi, O., von Delft, F., Edwards, A., Arrowsmith, C., Weigelt, J., Sundstrom, M., Oppermann, U. (2006) Crystal structure of human 3-phosphoglycerate dehydrogenase.
26. Fjeld, C. C., Birdsong, W. T., and Goodman, R. H. (2003) Differential binding of NAD+ and NADH allows the transcriptional corepressor carboxyl-terminal binding protein to serve as a metabolic sensor, *Proc Natl Acad Sci U S A 100*, 9202-9207.
27. Balasubramanian, P., Zhao, L. J., and Chinnadurai, G. (2003) Nicotinamide adenine dinucleotide stimulates oligomerization, interaction with adenovirus E1A and an intrinsic dehydrogenase activity of CtBP, *FEBS Lett 537*, 157-160.
28. Zhang, Q., Wang, S. Y., Nottke, A. C., Rocheleau, J. V., Piston, D. W., and Goodman, R. H. (2006) Redox sensor CtBP mediates hypoxia-induced tumor cell migration, *Proc Natl Acad Sci U S A 103*, 9029-9033.
29. Zhang, Q., Piston, D. W., and Goodman, R. H. (2002) Regulation of corepressor function by nuclear NADH, *Science 295*, 1895-1897.
30. Kim, J. H., and Youn, H. D. (2009) C-terminal binding protein maintains mitochondrial activities, *Cell Death Differ 16*, 584-592.
31. Bhambhani, C., Chang, J. L., Akey, D. L., and Cadigan, K. M. (2011) The oligomeric state of CtBP determines its role as a transcriptional co-activator and co-repressor of Wingless targets, *EMBO J 30*, 2031-2043.
32. Hildebrand, J. D., and Soriano, P. (2002) Overlapping and unique roles for C-terminal binding protein 1 (CtBP1) and CtBP2 during mouse development, *Mol Cell Biol 22*, 5296-5307.
33. Nagel, A. C., Krejci, A., Tenin, G., Bravo-Patino, A., Bray, S., Maier, D., and Preiss, A. (2005) Hairless-mediated repression of notch target genes requires the combined activity of Groucho and CtBP corepressors, *Mol Cell Biol 25*, 10433-10441.
34. Nagel, A. C., and Preiss, A. (2011) Fine tuning of Notch signaling by differential co-repressor recruitment during eye development of Drosophila, *Hereditas 148*, 77-84.

35. Hoang, C. Q., Burnett, M. E., and Curtiss, J. (2010) Drosophila CtBP regulates proliferation and differentiation of eye precursors and complexes with Eyeless, Dachshund, Dan, and Danr during eye and antennal development, *Dev Dyn 239*, 2367-2385.
36. Biryukova, I., and Heitzler, P. (2008) Drosophila C-terminal binding protein, dCtBP is required for sensory organ prepattern and sharpens proneural transcriptional activity of the GATA factor Pnr, *Dev Biol 323*, 64-75.
37. Nagel, A. C., Wech, I., Schwinkendorf, D., and Preiss, A. (2007) Involvement of co-repressors Groucho and CtBP in the regulation of single-minded in Drosophila, *Hereditas 144*, 195-205.
38. Zhang, Y. W., and Arnosti, D. N. (2011) Conserved catalytic and C-terminal regulatory domains of the C-terminal binding protein corepressor fine-tune the transcriptional response in development, *Mol Cell Biol 31*, 375-384.
39. Nardini, M., Spano, S., Cericola, C., Pesce, A., Massaro, A., Millo, E., Luini, A., Corda, D., and Bolognesi, M. (2003) CtBP/BARS: a dual-function protein involved in transcription co-repression and Golgi membrane fission, *EMBO J 22*, 3122-3130.
40. Postigo, A. A., and Dean, D. C. (1999) ZEB represses transcription through interaction with the corepressor CtBP, *Proc Natl Acad Sci U S A 96*, 6683-6688.
41. Wang, J., Lee, S., Teh, C. E., Bunting, K., Ma, L., and Shannon, M. F. (2009) The transcription repressor, ZEB1, cooperates with CtBP2 and HDAC1 to suppress IL-2 gene activation in T cells, *Int Immunol 21*, 227-235.
42. Nibu, Y., Zhang, H., Bajor, E., Barolo, S., Small, S., and Levine, M. (1998) dCtBP mediates transcriptional repression by Knirps, Kruppel and Snail in the Drosophila embryo, *EMBO J 17*, 7009-7020.
43. Turner, J., and Crossley, M. (2001) The CtBP family: enigmatic and enzymatic transcriptional co-repressors, *Bioessays 23*, 683-690.
44. Quinlan, K. G., Nardini, M., Verger, A., Francescato, P., Yaswen, P., Corda, D., Bolognesi, M., and Crossley, M. (2006) Specific recognition of ZNF217 and other zinc finger proteins at a surface groove of C-terminal binding proteins, *Mol Cell Biol 26*, 8159-8172.
45. Cowger, J. J., Zhao, Q., Isovic, M., and Torchia, J. (2007) Biochemical characterization of the zinc-finger protein 217 transcriptional repressor complex: identification of a ZNF217 consensus recognition sequence, *Oncogene 26*, 3378-3386.
46. Shi, Y., Sawada, J., Sui, G., Affar el, B., Whetstine, J. R., Lan, F., Ogawa, H., Luke, M. P., Nakatani, Y., and Shi, Y. (2003) Coordinated histone modifications mediated by a CtBP co-repressor complex, *Nature 422*, 735-738.
47. Zhang, C. L., McKinsey, T. A., Lu, J. R., and Olson, E. N. (2001) Association of COOH-terminal-binding protein (CtBP) and MEF2-interacting transcription repressor (MITR) contributes to transcriptional repression of the MEF2 transcription factor, *J Biol Chem 276*, 35-39.
48. Lee, M. G., Wynder, C., Cooch, N., and Shiekhattar, R. (2005) An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation, *Nature 437*, 432-435.
49. You, A., Tong, J. K., Grozinger, C. M., and Schreiber, S. L. (2001) CoREST is an integral component of the CoREST- human histone deacetylase complex, *Proc Natl Acad Sci U S A 98*, 1454-1458.
50. Palijan, A., Fernandes, I., Verway, M., Kourelis, M., Bastien, Y., Tavera-Mendoza, L. E., Sacheli, A., Bourdeau, V., Mader, S., and White, J. H. (2009) Ligand-dependent corepressor LCoR is an attenuator of progesterone-regulated gene expression, *J Biol Chem 284*, 30275-30287.
51. Ueda, J., Tachibana, M., Ikura, T., and Shinkai, Y. (2006) Zinc finger protein Wiz links G9a/GLP histone methyltransferases to the co-repressor molecule CtBP, *J Biol Chem 281*, 20120-20128.
52. Melhuish, T. A., and Wotton, D. (2000) The interaction of the carboxyl terminus-binding protein with the Smad corepressor TGIF is disrupted by a holoprosencephaly mutation in TGIF, *J Biol Chem 275*, 39762-39766.
53. Izutsu, K., Kurokawa, M., Imai, Y., Maki, K., Mitani, K., and Hirai, H. (2001) The corepressor CtBP interacts with Evi-1 to repress transforming growth factor beta signaling, *Blood 97*, 2815-2822.

54. Bonazzi, M., Spano, S., Turacchio, G., Cericola, C., Valente, C., Colanzi, A., Kweon, H. S., Hsu, V. W., Polishchuck, E. V., Polishchuck, R. S., Sallese, M., Pulvirenti, T., Corda, D., and Luini, A. (2005) CtBP3/BARS drives membrane fission in dynamin-independent transport pathways, *Nat Cell Biol 7*, 570-580.
55. Weigert, R., Silletta, M. G., Spano, S., Turacchio, G., Cericola, C., Colanzi, A., Senatore, S., Mancini, R., Polishchuk, E. V., Salmona, M., Facchiano, F., Burger, K. N., Mironov, A., Luini, A., and Corda, D. (1999) CtBP/BARS induces fission of Golgi membranes by acylating lysophosphatidic acid, *Nature 402*, 429-433.
56. Corda, D., Colanzi, A., and Luini, A. (2006) The multiple activities of CtBP/BARS proteins: the Golgi view, *Trends Cell Biol 16*, 167-173.
57. Gallop, J. L., Butler, P. J., and McMahon, H. T. (2005) Endophilin and CtBP/BARS are not acyl transferases in endocytosis or Golgi fission, *Nature 438*, 675-678.
58. Yang, J. S., Lee, S. Y., Spano, S., Gad, H., Zhang, L., Nie, Z., Bonazzi, M., Corda, D., Luini, A., and Hsu, V. W. (2005) A role for BARS at the fission step of COPI vesicle formation from Golgi membrane, *EMBO J 24*, 4133-4143.
59. Schwarz, K., Natarajan, S., Kassas, N., Vitale, N., and Schmitz, F. (2011) The synaptic ribbon is a site of phosphatidic acid generation in ribbon synapses, *J Neurosci 31*, 15996-16011.
60. Hidalgo Carcedo, C., Bonazzi, M., Spano, S., Turacchio, G., Colanzi, A., Luini, A., and Corda, D. (2004) Mitotic Golgi partitioning is driven by the membrane-fissioning protein CtBP3/BARS, *Science 305*, 93-96.
61. Vinals, C., De Bolle, X., Depiereux, E., and Feytmans, E. (1995) Knowledge-based modeling of the D-lactate dehydrogenase three-dimensional structure, *Proteins 21*, 307-318.
62. Stoll, V. S., Kimber, M. S., and Pai, E. F. (1996) Insights into substrate binding by D-2-ketoacid dehydrogenases from the structure of Lactobacillus pentosus D-lactate dehydrogenase, *Structure 4*, 437-447.
63. Achouri, Y., Noel, G., and Van Schaftingen, E. (2007) 2-Keto-4-methylthiobutyrate, an intermediate in the methionine salvage pathway, is a good substrate for CtBP1, *Biochem Biophys Res Commun 352*, 903-906.
64. Sekowska, A., Denervaud, V., Ashida, H., Michoud, K., Haas, D., Yokota, A., and Danchin, A. (2004) Bacterial variations on the methionine salvage pathway, *BMC Microbiol 4*, 9.
65. Dilger, R. N., Kobler, C., Weckbecker, C., Hoehler, D., and Baker, D. H. (2007) 2-keto-4-(methylthio)butyric acid (keto analog of methionine) is a safe and efficacious precursor of L-methionine in chicks, *J Nutr 137*, 1868-1873.
66. Grooteclaes, M., Deveraux, Q., Hildebrand, J., Zhang, Q., Goodman, R. H., and Frisch, S. M. (2003) C-terminal-binding protein corepresses epithelial and proapoptotic gene expression programs, *Proc Natl Acad Sci U S A 100*, 4568-4573.
67. Achouri, Y., Rider, M. H., Schaftingen, E. V., and Robbi, M. (1997) Cloning, sequencing and expression of rat liver 3-phosphoglycerate dehydrogenase, *Biochem J 323 (Pt 2)*, 365-370.
68. Nunez, M. F., Pellicer, M. T., Badia, J., Aguilar, J., and Baldoma, L. (2001) Biochemical characterization of the 2-ketoacid reductases encoded by ycdW and yiaE genes in Escherichia coli, *Biochem J 354*, 707-715.
69. Lim, J., and Thiery, J. P. (2012) Epithelial-mesenchymal transitions: insights from development, *Development 139*, 3471-3486.
70. Thiery, J. P., Acloque, H., Huang, R. Y., and Nieto, M. A. (2009) Epithelial-mesenchymal transitions in development and disease, *Cell 139*, 871-890.
71. Chinnadurai, G. (2009) The transcriptional corepressor CtBP: a foe of multiple tumor suppressors, *Cancer Res 69*, 731-734.
72. Di, L. J., Byun, J. S., Wong, M. M., Wakano, C., Taylor, T., Bilke, S., Baek, S., Hunter, K., Yang, H., Lee, M., Zvosec, C., Khramtsova, G., Cheng, F., Perou, C. M., Ryan Miller, C., Raab, R., Olopade, O. I., and Gardner, K. (2013) Genome-wide profiles of CtBP link metabolism with genome stability and epithelial reprogramming in breast cancer, *Nat Commun 4*, 1449.
73. Winklmeier, A., Poser, I., Hoek, K. S., and Bosserhoff, A. K. (2009) Loss of full length CtBP1 expression enhances the invasive potential of human melanoma, *BMC Cancer 9*, 52.

74. Dorman, K., Shen, Z., Yang, C., Ezzat, S., and Asa, S. L. (2012) CtBP1 interacts with Ikaros and modulates pituitary tumor cell survival and response to hypoxia, *Mol Endocrinol 26*, 447-457.
75. Wang, R., Asangani, I. A., Chakravarthi, B. V., Ateeq, B., Lonigro, R. J., Cao, Q., Mani, R. S., Camacho, D. F., McGregor, N., Schumann, T. E., Jing, X., Menawat, R., Tomlins, S. A., Zheng, H., Otte, A. P., Mehra, R., Siddiqui, J., Dhanasekaran, S. M., Nyati, M. K., Pienta, K. J., Palanisamy, N., Kunju, L. P., Rubin, M. A., Chinnaiyan, A. M., and Varambally, S. (2012) Role of transcriptional corepressor CtBP1 in prostate cancer progression, *Neoplasia 14*, 905-914.
76. Nadauld, L. D., Phelps, R., Moore, B. C., Eisinger, A., Sandoval, I. T., Chidester, S., Peterson, P. W., Manos, E. J., Sklow, B., Burt, R. W., and Jones, D. A. (2006) Adenomatous polyposis coli control of C-terminal binding protein-1 stability regulates expression of intestinal retinol dehydrogenases, *J Biol Chem 281*, 37828-37835.
77. Phelps, R. A., Chidester, S., Dehghanizadeh, S., Phelps, J., Sandoval, I. T., Rai, K., Broadbent, T., Sarkar, S., Burt, R. W., and Jones, D. A. (2009) A two-step model for colon adenoma initiation and progression caused by APC loss, *Cell 137*, 623-634.
78. Paliwal, S., Pande, S., Kovi, R. C., Sharpless, N. E., Bardeesy, N., and Grossman, S. R. (2006) Targeting of C-terminal binding protein (CtBP) by ARF results in p53-independent apoptosis, *Mol Cell Biol 26*, 2360-2372.
79. Kovi, R. C., Paliwal, S., Pande, S., and Grossman, S. R. (2010) An ARF/CtBP2 complex regulates BH3-only gene expression and p53-independent apoptosis, *Cell Death Differ 17*, 513-521.
80. Paliwal, S., Kovi, R. C., Nath, B., Chen, Y. W., Lewis, B. C., and Grossman, S. R. (2007) The alternative reading frame tumor suppressor antagonizes hypoxia-induced cancer cell migration via interaction with the COOH-terminal binding protein corepressor, *Cancer Res 67*, 9322-9329.
81. Straza, M. W., Paliwal, S., Kovi, R. C., Rajeshkumar, B., Trenh, P., Parker, D., Whalen, G. F., Lyle, S., Schiffer, C. A., and Grossman, S. R. (2010) Therapeutic targeting of C-terminal binding protein in human cancer, *Cell Cycle 9*, 3740-3750.
82. Deng, Y., Deng, H., Liu, J., Han, G., Malkoski, S., Liu, B., Zhao, R., Wang, X. J., and Zhang, Q. (2012) Transcriptional down-regulation of Brca1 and E-cadherin by CtBP1 in breast cancer, *Mol Carcinog 51*, 500-507.
83. Di, L. J., Fernandez, A. G., De Siervi, A., Longo, D. L., and Gardner, K. (2010) Transcriptional regulation of BRCA1 expression by a metabolic switch, *Nat Struct Mol Biol 17*, 1406-1413.
84. Bergman, L. M., Birts, C. N., Darley, M., Gabrielli, B., and Blaydes, J. P. (2009) CtBPs promote cell survival through the maintenance of mitotic fidelity, *Mol Cell Biol 29*, 4539-4551.
85. Fodde, R., Smits, R., and Clevers, H. (2001) APC, signal transduction and genetic instability in colorectal cancer, *Nat Rev Cancer 1*, 55-67.
86. Shimazu, T., Degenhardt, K., Nur, E. K. A., Zhang, J., Yoshida, T., Zhang, Y., Mathew, R., White, E., and Inouye, M. (2007) NBK/BIK antagonizes MCL-1 and BCL-XL and activates BAK-mediated apoptosis in response to protein synthesis inhibition, *Genes Dev 21*, 929-941.
87. Chiarugi, A., Dolle, C., Felici, R., and Ziegler, M. (2012) The NAD metabolome--a key determinant of cancer cell biology, *Nat Rev Cancer 12*, 741-752.
88. Uppal, A., and Gupta, P. K. (2003) Measurement of NADH concentration in normal and malignant human tissues from breast and oral cavity, *Biotechnol Appl Biochem 37*, 45-50.
89. Tang, B., Kadariya, Y., Murphy, M. E., and Kruger, W. D. (2006) The methionine salvage pathway compound 4-methylthio-2-oxobutanate causes apoptosis independent of down-regulation of ornithine decarboxylase, *Biochem Pharmacol 72*, 806-815.
90. Pilka, E. S., Guo, K., Rojkova, A., Debreczeni, J. E., Kavanagh, K. L., von Delft, F., Arrowsmith, C. H., Weigelt, J., Edwards, A., Sundstrom, M., and Oppermann, U. (2007) Crystal Structure of Human CtBP2 dehydrogenase complexed with NAD(H).
91. Taniguchi, T., Ogasawara, K., Takaoka, A., and Tanaka, N. (2001) IRF family of transcription factors as regulators of host defense, *Annu Rev Immunol 19*, 623-655.
92. Tamura, T., Yanai, H., Savitsky, D., and Taniguchi, T. (2008) The IRF family transcription factors in immunity and oncogenesis, *Annu Rev Immunol 26*, 535-584.
93. Escalante, C. R., Yie, J., Thanos, D., and Aggarwal, A. K. (1998) Structure of IRF-1 with bound DNA reveals determinants of interferon regulation, *Nature 391*, 103-106.

94. Fujii, Y., Shimizu, T., Kusumoto, M., Kyogoku, Y., Taniguchi, T., and Hakoshima, T. (1999) Crystal structure of an IRF-DNA complex reveals novel DNA recognition and cooperative binding to a tandem repeat of core sequences, *EMBO J 18*, 5028-5041.
95. Darnell, J. E., Jr., Kerr, I. M., and Stark, G. R. (1994) Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins, *Science 264*, 1415-1421.
96. Mamane, Y., Heylbroeck, C., Genin, P., Algarte, M., Servant, M. J., LePage, C., DeLuca, C., Kwon, H., Lin, R., and Hiscott, J. (1999) Interferon regulatory factors: the next generation, *Gene 237*, 1-14.
97. Lohoff, M., and Mak, T. W. (2005) Roles of interferon-regulatory factors in T-helper-cell differentiation, *Nat Rev Immunol 5*, 125-135.
98. Lin, R., Mamane, Y., and Hiscott, J. (1999) Structural and functional analysis of interferon regulatory factor 3: localization of the transactivation and autoinhibitory domains, *Mol Cell Biol 19*, 2465-2474.
99. Barnes, B. J., Kellum, M. J., Field, A. E., and Pitha, P. M. (2002) Multiple regulatory domains of IRF-5 control activation, cellular localization, and induction of chemokines that mediate recruitment of T lymphocytes, *Mol Cell Biol 22*, 5721-5740.
100. Kumar, K. P., McBride, K. M., Weaver, B. K., Dingwall, C., and Reich, N. C. (2000) Regulated nuclear-cytoplasmic localization of interferon regulatory factor 3, a subunit of double-stranded RNA-activated factor 1, *Mol Cell Biol 20*, 4159-4168.
101. Lin, R., Yang, L., Arguello, M., Penafuerte, C., and Hiscott, J. (2005) A CRM1-dependent nuclear export pathway is involved in the regulation of IRF-5 subcellular localization, *J Biol Chem 280*, 3088-3095.
102. Shukla, H., Vaitiekunas, P., Majumdar, A. K., Dragan, A. I., Dimitriadis, E. K., Kotova, S., Crane-Robinson, C., and Privalov, P. L. (2012) The linker of the interferon response factor 3 transcription factor is not unfolded, *Biochemistry 51*, 6320-6327.
103. Au, W. C., Moore, P. A., Lowther, W., Juang, Y. T., and Pitha, P. M. (1995) Identification of a member of the interferon regulatory factor family that binds to the interferon-stimulated response element and activates expression of interferon-induced genes, *Proc Natl Acad Sci U S A 92*, 11657-11661.
104. Mancl, M. E., Hu, G., Sangster-Guity, N., Olshalsky, S. L., Hoops, K., Fitzgerald-Bocarsly, P., Pitha, P. M., Pinder, K., and Barnes, B. J. (2005) Two discrete promoters regulate the alternatively spliced human interferon regulatory factor-5 isoforms. Multiple isoforms with distinct cell type-specific expression, localization, regulation, and function, *J Biol Chem 280*, 21078-21090.
105. Noppert, S. J., Fitzgerald, K. A., and Hertzog, P. J. (2007) The role of type I interferons in TLR responses, *Immunol Cell Biol 85*, 446-457.
106. Barnes, B. J., Moore, P. A., and Pitha, P. M. (2001) Virus-specific activation of a novel interferon regulatory factor, IRF-5, results in the induction of distinct interferon alpha genes, *J Biol Chem 276*, 23382-23390.
107. Sato, M., Tanaka, N., Hata, N., Oda, E., and Taniguchi, T. (1998) Involvement of the IRF family transcription factor IRF-3 in virus-induced activation of the IFN-beta gene, *FEBS Lett 425*, 112-116.
108. Matsumoto, M., Funami, K., Tanabe, M., Oshiumi, H., Shingai, M., Seto, Y., Yamamoto, A., and Seya, T. (2003) Subcellular localization of Toll-like receptor 3 in human dendritic cells, *J Immunol 171*, 3154-3162.
109. Kato, H., Takeuchi, O., Sato, S., Yoneyama, M., Yamamoto, M., Matsui, K., Uematsu, S., Jung, A., Kawai, T., Ishii, K. J., Yamaguchi, O., Otsu, K., Tsujimura, T., Koh, C. S., Reis e Sousa, C., Matsuura, Y., Fujita, T., and Akira, S. (2006) Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses, *Nature 441*, 101-105.
110. Fitzgerald, K. A., Rowe, D. C., Barnes, B. J., Caffrey, D. R., Visintin, A., Latz, E., Monks, B., Pitha, P. M., and Golenbock, D. T. (2003) LPS-TLR4 signaling to IRF-3/7 and NF-kappaB involves the toll adapters TRAM and TRIF, *J Exp Med 198*, 1043-1055.

111. Takaoka, A., Yanai, H., Kondo, S., Duncan, G., Negishi, H., Mizutani, T., Kano, S., Honda, K., Ohba, Y., Mak, T. W., and Taniguchi, T. (2005) Integral role of IRF-5 in the gene induction programme activated by Toll-like receptors, *Nature 434*, 243-249.
112. Schoenemeyer, A., Barnes, B. J., Mancl, M. E., Latz, E., Goutagny, N., Pitha, P. M., Fitzgerald, K. A., and Golenbock, D. T. (2005) The interferon regulatory factor, IRF5, is a central mediator of toll-like receptor 7 signaling, *J Biol Chem 280*, 17005-17012.
113. Paun, A., Reinert, J. T., Jiang, Z., Medin, C., Balkhi, M. Y., Fitzgerald, K. A., and Pitha, P. M. (2008) Functional characterization of murine interferon regulatory factor 5 (IRF-5) and its role in the innate antiviral response, *J Biol Chem 283*, 14295-14308.
114. Yanai, H., Chen, H. M., Inuzuka, T., Kondo, S., Mak, T. W., Takaoka, A., Honda, K., and Taniguchi, T. (2007) Role of IFN regulatory factor 5 transcription factor in antiviral immunity and tumor suppression, *Proc Natl Acad Sci U S A 104*, 3402-3407.
115. Hu, G., and Barnes, B. J. (2009) IRF-5 is a mediator of the death receptor-induced apoptotic signaling pathway, *J Biol Chem 284*, 2767-2777.
116. Ellwood, K. B., Yen, Y. M., Johnson, R. C., and Carey, M. (2000) Mechanism for specificity by HMG-1 in enhanceosome assembly, *Mol Cell Biol 20*, 4359-4370.
117. Thanos, D., and Maniatis, T. (1992) The high mobility group protein HMG I(Y) is required for NF-kappa B-dependent virus induction of the human IFN-beta gene, *Cell 71*, 777-789.
118. Falvo, J. V., Parekh, B. S., Lin, C. H., Fraenkel, E., and Maniatis, T. (2000) Assembly of a functional beta interferon enhanceosome is dependent on ATF-2-c-jun heterodimer orientation, *Mol Cell Biol 20*, 4814-4825.
119. Juang, Y. T., Lowther, W., Kellum, M., Au, W. C., Lin, R., Hiscott, J., and Pitha, P. M. (1998) Primary activation of interferon A and interferon B gene transcription by interferon regulatory factor 3, *Proc Natl Acad Sci U S A 95*, 9837-9842.
120. Sato, M., Hata, N., Asagiri, M., Nakaya, T., Taniguchi, T., and Tanaka, N. (1998) Positive feedback regulation of type I IFN genes by the IFN-inducible transcription factor IRF-7, *FEBS Lett 441*, 106-110.
121. Agalioti, T., Lomvardas, S., Parekh, B., Yie, J., Maniatis, T., and Thanos, D. (2000) Ordered recruitment of chromatin modifying and general transcription factors to the IFN-beta promoter, *Cell 103*, 667-678.
122. Merika, M., Williams, A. J., Chen, G., Collins, T., and Thanos, D. (1998) Recruitment of CBP/p300 by the IFN beta enhanceosome is required for synergistic activation of transcription, *Mol Cell 1*, 277-287.
123. Lin, R., Genin, P., Mamane, Y., and Hiscott, J. (2000) Selective DNA binding and association with the CREB binding protein coactivator contribute to differential activation of alpha/beta interferon genes by interferon regulatory factors 3 and 7, *Mol Cell Biol 20*, 6342-6353.
124. Bannister, A. J., Oehler, T., Wilhelm, D., Angel, P., and Kouzarides, T. (1995) Stimulation of c-Jun activity by CBP: c-Jun residues Ser63/73 are required for CBP induced stimulation in vivo and CBP binding in vitro, *Oncogene 11*, 2509-2514.
125. Kawasaki, H., Song, J., Eckner, R., Ugai, H., Chiu, R., Taira, K., Shi, Y., Jones, N., and Yokoyama, K. K. (1998) p300 and ATF-2 are components of the DRF complex, which regulates retinoic acid- and E1A-mediated transcription of the c-jun gene in F9 cells, *Genes Dev 12*, 233-245.
126. Fitzgerald, K. A., McWhirter, S. M., Faia, K. L., Rowe, D. C., Latz, E., Golenbock, D. T., Coyle, A. J., Liao, S. M., and Maniatis, T. (2003) IKKepsilon and TBK1 are essential components of the IRF3 signaling pathway, *Nat Immunol 4*, 491-496.
127. Hemmi, H., Takeuchi, O., Sato, S., Yamamoto, M., Kaisho, T., Sanjo, H., Kawai, T., Hoshino, K., Takeda, K., and Akira, S. (2004) The roles of two IkappaB kinase-related kinases in lipopolysaccharide and double stranded RNA signaling and viral infection, *J Exp Med 199*, 1641-1650.
128. Mori, M., Yoneyama, M., Ito, T., Takahashi, K., Inagaki, F., and Fujita, T. (2004) Identification of Ser-386 of interferon regulatory factor 3 as critical target for inducible phosphorylation that determines activation, *J Biol Chem 279*, 9698-9702.

129. Servant, M. J., Grandvaux, N., tenOever, B. R., Duguay, D., Lin, R., and Hiscott, J. (2003) Identification of the minimal phosphoacceptor site required for in vivo activation of interferon regulatory factor 3 in response to virus and double-stranded RNA, *J Biol Chem 278*, 9441-9447.
130. Chen, W., Srinath, H., Lam, S. S., Schiffer, C. A., Royer, W. E., Jr., and Lin, K. (2008) Contribution of Ser386 and Ser396 to activation of interferon regulatory factor 3, *J Mol Biol 379*, 251-260.
131. Panne, D., McWhirter, S. M., Maniatis, T., and Harrison, S. C. (2007) Interferon regulatory factor 3 is regulated by a dual phosphorylation-dependent switch, *J Biol Chem 282*, 22816-22822.
132. Clement, J. F., Bibeau-Poirier, A., Gravel, S. P., Grandvaux, N., Bonneil, E., Thibault, P., Meloche, S., and Servant, M. J. (2008) Phosphorylation of IRF-3 on Ser 339 generates a hyperactive form of IRF-3 through regulation of dimerization and CBP association, *J Virol 82*, 3984-3996.
133. Fujii, K., Nakamura, S., Takahashi, K., and Inagaki, F. (2010) Systematic characterization by mass spectrometric analysis of phosphorylation sites in IRF-3 regulatory domain activated by IKK-i, *J Proteomics 73*, 1196-1203.
134. Balkhi, M. Y., Fitzgerald, K. A., and Pitha, P. M. (2008) Functional regulation of MyD88-activated interferon regulatory factor 5 by K63-linked polyubiquitination, *Mol Cell Biol 28*, 7296-7308.
135. Chen, W., Lam, S. S., Srinath, H., Jiang, Z., Correia, J. J., Schiffer, C. A., Fitzgerald, K. A., Lin, K., and Royer, W. E., Jr. (2008) Insights into interferon regulatory factor activation from the crystal structure of dimeric IRF5, *Nat Struct Mol Biol 15*, 1213-1220.
136. Chang Foreman, H. C., Van Scoy, S., Cheng, T. F., and Reich, N. C. (2012) Activation of interferon regulatory factor 5 by site specific phosphorylation, *PLoS One 7*, e33098.
137. Ma, Y., Jin, H., Valyi-Nagy, T., Cao, Y., Yan, Z., and He, B. (2012) Inhibition of TANK binding kinase 1 by herpes simplex virus 1 facilitates productive infection, *J Virol 86*, 2188-2196.
138. Zhu, H., Zheng, C., Xing, J., Wang, S., Li, S., Lin, R., and Mossman, K. L. (2011) Varicella-zoster virus immediate-early protein ORF61 abrogates the IRF3-mediated innate immune response through degradation of activated IRF3, *J Virol 85*, 11079-11089.
139. Wang, J. T., Doong, S. L., Teng, S. C., Lee, C. P., Tsai, C. H., and Chen, M. R. (2009) Epstein-Barr virus BGLF4 kinase suppresses the interferon regulatory factor 3 signaling pathway, *J Virol 83*, 1856-1869.
140. Bentz, G. L., Liu, R., Hahn, A. M., Shackelford, J., and Pagano, J. S. (2010) Epstein-Barr virus BRLF1 inhibits transcription of IRF3 and IRF7 and suppresses induction of interferon-beta, *Virology 402*, 121-128.
141. Abate, D. A., Watanabe, S., and Mocarski, E. S. (2004) Major human cytomegalovirus structural protein pp65 (ppUL83) prevents interferon response factor 3 activation in the interferon response, *J Virol 78*, 10995-11006.
142. Sun, L., Xing, Y., Chen, X., Zheng, Y., Yang, Y., Nichols, D. B., Clementz, M. A., Banach, B. S., Li, K., Baker, S. C., and Chen, Z. (2012) Coronavirus papain-like proteases negatively regulate antiviral innate immune response through disruption of STING-mediated signaling, *PLoS One 7*, e30802.
143. Ramanan, P., Shabman, R. S., Brown, C. S., Amarasinghe, G. K., Basler, C. F., and Leung, D. W. (2011) Filoviral immune evasion mechanisms, *Viruses 3*, 1634-1649.
144. Prins, K. C., Cardenas, W. B., and Basler, C. F. (2009) Ebola virus protein VP35 impairs the function of interferon regulatory factor-activating kinases IKKepsilon and TBK-1, *J Virol 83*, 3069-3077.
145. Ramanan, P., Edwards, M. R., Shabman, R. S., Leung, D. W., Endlich-Frazier, A. C., Borek, D. M., Otwinowski, Z., Liu, G., Huh, J., Basler, C. F., and Amarasinghe, G. K. (2012) Structural basis for Marburg virus VP35-mediated immune evasion mechanisms, *Proc Natl Acad Sci U S A 109*, 20661-20666.
146. Radu, O., and Pantanowitz, L. (2013) Kaposi sarcoma, *Arch Pathol Lab Med 137*, 289-294.
147. Arvanitakis, L., Mesri, E. A., Nador, R. G., Said, J. W., Asch, A. S., Knowles, D. M., and Cesarman, E. (1996) Establishment and characterization of a primary effusion (body cavity-based)

lymphoma cell line (BC-3) harboring kaposi's sarcoma-associated herpesvirus (KSHV/HHV-8) in the absence of Epstein-Barr virus, *Blood 88*, 2648-2654.

148. Nador, R. G., Cesarman, E., Chadburn, A., Dawson, D. B., Ansari, M. Q., Sald, J., and Knowles, D. M. (1996) Primary effusion lymphoma: a distinct clinicopathologic entity associated with the Kaposi's sarcoma-associated herpes virus, *Blood 88*, 645-656.
149. Soulier, J., Grollet, L., Oksenhendler, E., Cacoub, P., Cazals-Hatem, D., Babinet, P., d'Agay, M. F., Clauvel, J. P., Raphael, M., Degos, L., and et al. (1995) Kaposi's sarcoma-associated herpesvirus-like DNA sequences in multicentric Castleman's disease, *Blood 86*, 1276-1280.
150. Offermann, M. K. (2007) Kaposi sarcoma herpesvirus-encoded interferon regulator factors, *Curr Top Microbiol Immunol 312*, 185-209.
151. Joo, C. H., Shin, Y. C., Gack, M., Wu, L., Levy, D., and Jung, J. U. (2007) Inhibition of interferon regulatory factor 7 (IRF7)-mediated interferon signal transduction by the Kaposi's sarcoma-associated herpesvirus viral IRF homolog vIRF3, *J Virol 81*, 8282-8292.
152. Wies, E., Hahn, A. S., Schmidt, K., Viebahn, C., Rohland, N., Lux, A., Schellhorn, T., Holzer, A., Jung, J. U., and Neipel, F. (2009) The Kaposi's Sarcoma-associated Herpesvirus-encoded vIRF-3 Inhibits Cellular IRF-5, *J Biol Chem 284*, 8525-8538.
153. Rivas, C., Thlick, A. E., Parravicini, C., Moore, P. S., and Chang, Y. (2001) Kaposi's sarcoma-associated herpesvirus LANA2 is a B-cell-specific latent viral protein that inhibits p53, *J Virol 75*, 429-438.
154. Frieri, M. (2013) Mechanisms of disease for the clinician: systemic lupus erythematosus, *Ann Allergy Asthma Immunol 110*, 228-232.
155. Sigurdsson, S., Nordmark, G., Goring, H. H., Lindroos, K., Wiman, A. C., Sturfelt, G., Jonsen, A., Rantapaa-Dahlqvist, S., Moller, B., Kere, J., Koskenmies, S., Widen, E., Eloranta, M. L., Julkunen, H., Kristjansdottir, H., Steinsson, K., Alm, G., Ronnblom, L., and Syvanen, A. C. (2005) Polymorphisms in the tyrosine kinase 2 and interferon regulatory factor 5 genes are associated with systemic lupus erythematosus, *Am J Hum Genet 76*, 528-537.
156. Graham, R. R., Kozyrev, S. V., Baechler, E. C., Reddy, M. V., Plenge, R. M., Bauer, J. W., Ortmann, W. A., Koeuth, T., Gonzalez Escribano, M. F., Pons-Estel, B., Petri, M., Daly, M., Gregersen, P. K., Martin, J., Altshuler, D., Behrens, T. W., and Alarcon-Riquelme, M. E. (2006) A common haplotype of interferon regulatory factor 5 (IRF5) regulates splicing and expression and is associated with increased risk of systemic lupus erythematosus, *Nat Genet 38*, 550-555.
157. Harley, J. B., Alarcon-Riquelme, M. E., Criswell, L. A., Jacob, C. O., Kimberly, R. P., Moser, K. L., Tsao, B. P., Vyse, T. J., Langefeld, C. D., Nath, S. K., Guthridge, J. M., Cobb, B. L., Mirel, D. B., Marion, M. C., Williams, A. H., Divers, J., Wang, W., Frank, S. G., Namjou, B., Gabriel, S. B., Lee, A. T., Gregersen, P. K., Behrens, T. W., Taylor, K. E., Fernando, M., Zidovetzki, R., Gaffney, P. M., Edberg, J. C., Rioux, J. D., Ojwang, J. O., James, J. A., Merrill, J. T., Gilkeson, G. S., Seldin, M. F., Yin, H., Baechler, E. C., Li, Q. Z., Wakeland, E. K., Bruner, G. R., Kaufman, K. M., and Kelly, J. A. (2008) Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci, *Nat Genet 40*, 204-210.
158. Sigurdsson, S., Goring, H. H., Kristjansdottir, G., Milani, L., Nordmark, G., Sandling, J. K., Eloranta, M. L., Feng, D., Sangster-Guity, N., Gunnarsson, I., Svenungsson, E., Sturfelt, G., Jonsen, A., Truedsson, L., Barnes, B. J., Alm, G., Ronnblom, L., and Syvanen, A. C. (2008) Comprehensive evaluation of the genetic variants of interferon regulatory factor 5 (IRF5) reveals a novel 5 bp length polymorphism as strong risk factor for systemic lupus erythematosus, *Hum Mol Genet 17*, 872-881.
159. Feng, D., Stone, R. C., Eloranta, M. L., Sangster-Guity, N., Nordmark, G., Sigurdsson, S., Wang, C., Alm, G., Syvanen, A. C., Ronnblom, L., and Barnes, B. J. (2010) Genetic variants and disease-associated factors contribute to enhanced interferon regulatory factor 5 expression in blood cells of patients with systemic lupus erythematosus, *Arthritis Rheum 62*, 562-573.
160. Ferreiro-Neira, I., Calaza, M., Alonso-Perez, E., Marchini, M., Scorza, R., Sebastiani, G. D., Blanco, F. J., Rego, I., Pullmann, R., Jr., Pullmann, R., Kallenberg, C. G., Bijl, M., Skopouli, F. N., Mavromati, M., Migliaresi, S., Barizzone, N., Ruzickova, S., Dostal, C., Schmidt, R. E., Witte,

T., Papasteriades, C., Kappou-Rigatou, I., Endreffy, E., Kovacs, A., Ordi-Ros, J., Balada, E., Carreira, P., Gomez-Reino, J. J., and Gonzalez, A. (2007) Opposed independent effects and epistasis in the complex association of IRF5 to SLE, *Genes Immun 8*, 429-438.

161. Wen, F., Ellingson, S. M., Kyogoku, C., Peterson, E. J., and Gaffney, P. M. (2011) Exon 6 variants carried on systemic lupus erythematosus (SLE) risk haplotypes modulate IRF5 function, *Autoimmunity 44*, 82-89.
162. Niewold, T. B., Kelly, J. A., Flesch, M. H., Espinoza, L. R., Harley, J. B., and Crow, M. K. (2008) Association of the IRF5 risk haplotype with high serum interferon-alpha activity in systemic lupus erythematosus patients, *Arthritis Rheum 58*, 2481-2487.
163. Richez, C., Yasuda, K., Bonegio, R. G., Watkins, A. A., Aprahamian, T., Busto, P., Richards, R. J., Liu, C. L., Cheung, R., Utz, P. J., Marshak-Rothstein, A., and Rifkin, I. R. (2010) IFN regulatory factor 5 is required for disease development in the FcgammaRIIB-/-Yaa and FcgammaRIIB-/- mouse models of systemic lupus erythematosus, *J Immunol 184*, 796-806.
164. Tada, Y., Kondo, S., Aoki, S., Koarada, S., Inoue, H., Suematsu, R., Ohta, A., Mak, T. W., and Nagasawa, K. (2011) Interferon regulatory factor 5 is critical for the development of lupus in MRL/lpr mice, *Arthritis Rheum 63*, 738-748.
165. Escalante, C. R., Nistal-Villan, E., Shen, L., Garcia-Sastre, A., and Aggarwal, A. K. (2007) Structure of IRF-3 bound to the PRDIII-I regulatory element of the human interferon-beta enhancer, *Mol Cell 26*, 703-716.
166. Panne, D., Maniatis, T., and Harrison, S. C. (2007) An atomic model of the interferon-beta enhanceosome, *Cell 129*, 1111-1123.
167. De Ioannes, P., Escalante, C. R., and Aggarwal, A. K. (2011) Structures of apo IRF-3 and IRF-7 DNA binding domains: effect of loop L1 on DNA binding, *Nucleic Acids Res 39*, 7300-7307.
168. Qin, B. Y., Liu, C., Lam, S. S., Srinath, H., Delston, R., Correia, J. J., Derynck, R., and Lin, K. (2003) Crystal structure of IRF-3 reveals mechanism of autoinhibition and virus-induced phosphoactivation, *Nat Struct Biol 10*, 913-921.
169. Takahasi, K., Suzuki, N. N., Horiuchi, M., Mori, M., Suhara, W., Okabe, Y., Fukuhara, Y., Terasawa, H., Akira, S., Fujita, T., and Inagaki, F. (2003) X-ray crystal structure of IRF-3 and its functional implications, *Nat Struct Biol 10*, 922-927.
170. Qin, B. Y., Liu, C., Srinath, H., Lam, S. S., Correia, J. J., Derynck, R., and Lin, K. (2005) Crystal structure of IRF-3 in complex with CBP, *Structure 13*, 1269-1277.
171. Takahasi, K., Horiuchi, M., Fujii, K., Nakamura, S., Noda, N. N., Yoneyama, M., Fujita, T., and Inagaki, F. (2010) Ser386 phosphorylation of transcription factor IRF-3 induces dimerization and association with CBP/p300 without overall conformational change, *Genes Cells 15*, 901-910.
172. Sathish, N., Zhu, F. X., Golub, E. E., Liang, Q., and Yuan, Y. (2011) Mechanisms of autoinhibition of IRF-7 and a probable model for inactivation of IRF-7 by Kaposi's sarcoma-associated herpesvirus protein ORF45, *J Biol Chem 286*, 746-756.
173. Summerford, C. M., Pardanani, A., Betts, A. H., Poteete, A. R., Colotti, G., and Royer, W. E., Jr. (1995) Bacterial expression of Scapharca dimeric hemoglobin: a simple model system for investigating protein cooperatively, *Protein Eng 8*, 593-599.
174. Brent, R., and Ptashne, M. (1981) Mechanism of action of the lexA gene product, *Proc Natl Acad Sci U S A 78*, 4204-4208.
175. Lin, R., Heylbroeck, C., Pitha, P. M., and Hiscott, J. (1998) Virus-dependent phosphorylation of the IRF-3 transcription factor regulates nuclear translocation, transactivation potential, and proteasome-mediated degradation, *Mol Cell Biol 18*, 2986-2996.
176. Dragan, A. I., Hargreaves, V. V., Makeyeva, E. N., and Privalov, P. L. (2007) Mechanisms of activation of interferon regulator factor 3: the role of C-terminal domain phosphorylation in IRF-3 dimerization and DNA binding, *Nucleic Acids Res 35*, 3525-3534.
177. Prilusky, J., Felder, C. E., Zeev-Ben-Mordehai, T., Rydberg, E. H., Man, O., Beckmann, J. S., Silman, I., and Sussman, J. L. (2005) FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded, *Bioinformatics 21*, 3435-3438.
178. Geneious Version 6.1, Biomatters.

179. Hanahan, D., and Weinberg, R. A. (2011) Hallmarks of cancer: the next generation, *Cell 144*, 646-674.
180. Zhang, Q., Yoshimatsu, Y., Hildebrand, J., Frisch, S. M., and Goodman, R. H. (2003) Homeodomain interacting protein kinase 2 promotes apoptosis by downregulating the transcriptional corepressor CtBP, *Cell 115*, 177-186.
181. Wang, S. Y., Iordanov, M., and Zhang, Q. (2006) c-Jun NH2-terminal kinase promotes apoptosis by down-regulating the transcriptional co-repressor CtBP, *J Biol Chem 281*, 34810-34815.
182. Barroilhet, L., Yang, J., Hasselblatt, K., Paranal, R. M., Ng, S. K., Rauh-Hain, J. A., Welch, W. R., Bradner, J. E., Berkowitz, R. S., and Ng, S. W. (2012) C-terminal binding protein-2 regulates response of epithelial ovarian cancer cells to histone deacetylase inhibitors, *Oncogene*.
183. (1994) The CCP4 suite: programs for protein crystallography, *Acta Crystallogr D Biol Crystallogr 50*, 760-763.
184. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software, *J Appl Crystallogr 40*, 658-674.
185. Lamzin, V. S., and Wilson, K. S. (1993) Automated refinement of protein models, *Acta Crystallogr D Biol Crystallogr 49*, 129-147.
186. Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures, *Acta Crystallogr D Biol Crystallogr 67*, 355-367.
187. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr D Biol Crystallogr 66*, 213-221.
188. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallogr D Biol Crystallogr 60*, 2126-2132.
189. (2012) Suite 2012: Schrödinger Suite 2011 Protein Preparation Wizard; Epik version 2.3; Impact version 5.8; Prime version 3.1 Schrödinger, LLC, New York, NY, 2012.
190. Madhavi Sastry, G., Adzhigirey, M., Day, T., Annabhimoju, R., and Sherman, W. (2013) Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments, *J Comput Aided Mol Des 27*, 221-234.
191. The PyMOL Molecular Graphics System, Version 1.5.0.4, Schrödinger, LLC.
192. Prabu-Jeyabalan, M., Nalivaika, E. A., King, N. M., and Schiffer, C. A. (2004) Structural basis for coevolution of a human immunodeficiency virus type 1 nucleocapsid-p1 cleavage site with a V82A drug-resistant mutation in viral protease, *J Virol 78*, 12446-12454.
193. Nalam, M. N., Ali, A., Altman, M. D., Reddy, G. S., Chellappan, S., Kairys, V., Ozen, A., Cao, H., Gilson, M. K., Tidor, B., Rana, T. M., and Schiffer, C. A. (2010) Evaluating the substrate-envelope hypothesis: structural analysis of novel HIV-1 protease inhibitors designed to be robust against drug resistance, *J Virol 84*, 5368-5378.
194. Meyer, E. A., Castellano, R. K., and Diederich, F. (2003) Interactions with aromatic rings in chemical and biological recognition, *Angew Chem Int Ed Engl 42*, 1210-1250.
195. Zauhar, R. J., Colbert, C. L., Morgan, R. S., and Welsh, W. J. (2000) Evidence for a strong sulfur-aromatic interaction derived from crystallographic data, *Biopolymers 53*, 233-248.
196. Morgan, R. S., Tatsch, C. E., Gushard, R. H., McAdon, J., and Warme, P. K. (1978) Chains of alternating sulfur and pi-bonded atoms in eight small proteins, *Int J Pept Protein Res 11*, 209-217.
197. (2012) Suite 2012: LigPrep, version 2.5, Schrödinger, LLC, New York, NY, 2012.
198. (2012) Suite 2012: Glide, version 5.8, Schrödinger, LLC, New York, NY, 2012.
199. Friesner, R. A., Murphy, R. B., Repasky, M. P., Frye, L. L., Greenwood, J. R., Halgren, T. A., Sanschagrin, P. C., and Mainz, D. T. (2006) Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes, *J Med Chem 49*, 6177-6196.

200. Halgren, T. A., Murphy, R. B., Friesner, R. A., Beard, H. S., Frye, L. L., Pollard, W. T., and Banks, J. L. (2004) Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening, *J Med Chem 47*, 1750-1759.
201. Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., Repasky, M. P., Knoll, E. H., Shelley, M., Perry, J. K., Shaw, D. E., Francis, P., and Shenkin, P. S. (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy, *J Med Chem 47*, 1739-1749.
202. Liu, C., Wrobleski, S. T., Lin, J., Ahmed, G., Metzger, A., Wityak, J., Gillooly, K. M., Shuster, D. J., McIntyre, K. W., Pitt, S., Shen, D. R., Zhang, R. F., Zhang, H., Doweyko, A. M., Diller, D., Henderson, I., Barrish, J. C., Dodd, J. H., Schieven, G. L., and Leftheris, K. (2005) 5-Cyanopyrimidine derivatives as a novel class of potent, selective, and orally active inhibitors of p38alpha MAP kinase, *J Med Chem 48*, 6261-6270.
203. Chen, J. M., Xu, S. L., Wawrzak, Z., Basarab, G. S., and Jordan, D. B. (1998) Structure-based design of potent inhibitors of scytalone dehydratase: displacement of a water molecule from the active site, *Biochemistry 37*, 17735-17744.
204. Wissner, A., Berger, D. M., Boschelli, D. H., Floyd, M. B., Jr., Greenberger, L. M., Gruber, B. C., Johnson, B. D., Mamuya, N., Nilakantan, R., Reich, M. F., Shen, R., Tsou, H. R., Upeslacis, E., Wang, Y. F., Wu, B., Ye, F., and Zhang, N. (2000) 4-Anilino-6,7-dialkoxyquinoline-3-carbonitrile inhibitors of epidermal growth factor receptor kinase and their bioisosteric relationship to the 4-anilino-6,7-dialkoxyquinazoline inhibitors, *J Med Chem 43*, 3244-3256.
205. Quinlan, K. G., Verger, A., Kwok, A., Lee, S. H., Perdomo, J., Nardini, M., Bolognesi, M., and Crossley, M. (2006) Role of the C-terminal binding protein PXDLS motif binding cleft in protein interactions and transcriptional repression, *Mol Cell Biol 26*, 8202-8213.
206. Turnbull, W. B., and Daranas, A. H. (2003) On the value of c: can low affinity systems be studied by isothermal titration calorimetry?, *J Am Chem Soc 125*, 14859-14866.
207. Tellinghuisen, J. (2008) Isothermal titration calorimetry at very low c, *Anal Biochem 373*, 395-397.
208. Sigurskjold, B. W. (2000) Exact analysis of competition ligand binding by displacement isothermal titration calorimetry, *Anal Biochem 277*, 260-266.
209. Halgren, T. A. (2009) Identifying and characterizing binding sites and assessing druggability, *J Chem Inf Model 49*, 377-389.
210. Halgren, T. (2007) New method for fast and accurate binding-site identification and analysis, *Chem Biol Drug Des 69*, 146-148.
211. (2012) Suite 2012: Sitemap, version 2.6, Schrodinger, LLC, New York, New York 2012.
212. Thio, S. S., Bonventre, J. V., and Hsu, S. I. (2004) The CtBP2 co-repressor is regulated by NADH-dependent dimerization and possesses a novel N-terminal repression domain, *Nucleic Acids Res 32*, 1836-1847.
213. Birts, C. N., Nijjar, S.K., Mardle, C.A., Hoakwie, F., Duriez, P.J., Blaydes, J.P., Tavassoli, A. (2013) A cyclic peptide inhibitor of C-terminal binding protein dimerization links metabolism with mitotic fidelity in breast cancer cells, *Chemical Science 4*, 3046-3057.
214. Nardini, M., Valente, C., Ricagno, S., Luini, A., Corda, D., and Bolognesi, M. (2009) CtBP1/BARS Gly172-->Glu mutant structure: impairing NAD(H)-binding and dimerization, *Biochem Biophys Res Commun 381*, 70-74.
215. Colanzi, A., Grimaldi, G., Catara, G., Valente, C., Cericola, C., Liberali, P., Ronci, M., Lalioti, V. S., Bruno, A., Beccari, A. R., Urbani, A., De Flora, A., Nardini, M., Bolognesi, M., Luini, A., and Corda, D. (2013) Molecular mechanism and functional role of brefeldin A-mediated ADP-ribosylation of CtBP1/BARS, *Proc Natl Acad Sci U S A 110*, 9794-9799.