STRUCTURAL STUDIES OF THE VON HIPPEL-LINDAU TUMOR SUPPRESSOR AND THE ONCOGENE CHAPERONE HSP90

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Charles E. Stebbins, Ph.D.

May 1999

Thesis Advisor: Nikola Pavletich

STRUCTURAL STUDIES OF THE VON HIPPEL-LINDAU TUMOR SUPPRESSOR AND THE ONCOGENE CHAPERONE HSP90

Charles E. Stebbins, Ph.D.

Cornell University 1999

Tumor development is associated with two broad types of genetic alterations: gain of function mutations in proto-oncogenes, and loss of function mutations in tumor suppressor genes. This thesis describes the x-ray crystal structures of two important proteins associated with each category.

Mutation of the von Hippel-Lindau (VHL) tumor suppressor is associated with the inherited VHL cancer syndrome and the majority of kidney cancers. VHL binds the ElonginC-ElonginB complex and regulates levels of hypoxia-inducible proteins. The structure of the ternary complex makes four important contributions toward understanding VHL function: (i) it shows two separate binary interfaces, one between VHL and ElonginC, and another between ElonginC and ElonginB, (ii) it shows that VHL uses a 35 residue domain to bind to ElonginC, and that this domain is a frequent target of tumorigenic mutations, (iii) it reveals a mutational patch on a separate domain of VHL likely corresponding to a second macromolecular binding site important for VHL function, and (iv) it extends the similarities between the

ElonginC-VHL complex and the Skp1–F-box protein complexes that target substrates for degradation, supporting the hypothesis that VHL may function as an adaptor in an analogous combinatorial regulatory pathway.

The heat shock protein 90 (Hsp90) chaperone is required for the activation of several families of eukaryotic protein kinases and nuclear hormone receptors, many of which are proto-oncogenic and play a prominent role in cancer. The geldanamycin antibiotic has antiproliferative and antitumor effects as it binds to Hsp90, inhibits Hsp90 activity, and results in the degradation of Hsp90 substrates. The structure of the geldanamycin-binding domain of Hsp90 reveals a conserved ATP binding motif harbored by a diverse range of proteins. This domain contains a pronounced pocket, 15 Å deep, that is highly conserved across species and in which geldanamycin binds. The structure of this complex reveals how this drug inhibits the binding of ATP to Hsp90, and provides a foundation for the rational design of improved geldanamycin derivatives.

BIOGRAPHICAL SKETCH

Charles Erec Stebbins was born on December 5, 1969 in Lincoln, Nebraska. His father is Charles Ernest Stebbins, a professor of romance philology at the University of Nebraska at Lincoln. His mother is Patricia Francis Patterson, a clinical therapist specializing in child psychology and the treatment of the chronically mentally ill. In 1988 he graduated from Huntsville High School in Huntsville, Alabama, shortly after which he attended Oberlin College, in Oberlin, Ohio. There he pursued studies of the physical sciences and mathematics, while spending several summers as a research assistant in the laboratory of astrophysicist and infrared astronomer Charles Telesco at NASA's Marshall Space Flight Center. He received his B.A. in physics from Oberlin in 1992. From September 1992 to May 1993, he studied as a graduate research fellow in the department of physics at the State University of New York at Stony Brook. Following a one year leave of absence from Stony Brook during which he worked as a research assistant in the laboratory of x-ray crystallographer Seth Darst, in 1994 he enrolled as a graduate student at the Joan and Stanford I. Weill Graduate School of Medical Sciences of Cornell University in New York City. There he completed the work described in this thesis under the guidance of crystallographer Nikola P. Pavletich.

PUBLICATIONS

1. S.A. Darst, C.E. Stebbins, S. Borukhov, M. Orlova, G. Feng, R. Landick, and A. Goldfard. (1994). "Crystallization of GreA, a Transcript Cleavage Factor from *Escherichia coli*." **Journal of Molecular Biology** 242, 582-585.

2. C.E. Stebbins, S. Borukhov, M. Orlova, A. Polyakov, A. Goldfard, and S.A. Darst. (1995). "Crystal Structure of the GreA Transcript Cleavage Factor from *Escherichia coli*." **Nature** 373, 636-640.

3. C.E. Stebbins, A.A. Russo, C. Schneider, N. Rosen, F.U. Hartl, and N.P. Pavletich. (1997). "Crystal Structure of an Hsp90-Geldanamycin Complex: Targeting of a Protein Chaperone by an Antitumor Agent." **Cell** 89, 239-250.

4. C.E. Stebbins, W.G.Kaelin, and N.P. Pavletich. (1999). "Structure of the VHL-ElonginC-ElonginB complex: implications for VHL tumor suppressor function." **Science** 284, 455-461.

To my mother, for revealing the discipline in compassion.

To my father, for demonstrating the compassion in discipline.

ACKNOWLEDGMENTS

The work herein would not have been possible were it not for the resources, stimulation, and motivation derived from the special research environment created by my thesis advisor, Nikola Pavletich. I wish to thank him for continually pursuing the highest levels of achievement and encouraging others to do the same.

Sincere thanks are also due to my Special Committee members Dinshaw Patel, Jerry Hurwitz, and David Hajjar for donating their time and advice to further my education and career in science.

I wish to thank Alicia Russo for sharing with me her vast knowledge of biochemistry and molecular biology, and likewise Phil Jeffrey for dispensing consistently fruitful advice in crystallographic theory and practice. Many thanks to my colleagues Paul Kussie, Hoang Nguyen, Lilly Tong, Alona Cohen, Michael Finnin, Hijuan Yang, Jie-Oh Lee, Ning Zhang, Brenda Shulman, and Lan Huang.

Special thanks are due to Seth Darst for taking the risk to allow a wayward physicist to undertake the crystal structure of GreA, thereby opening many doors for me in the biological sciences.

I am deeply indebted to my mother and father for their constant support and encouragement. Thanks to my brother for his timely sense of humor and perspective. And much gratitude to my mother and father-in-law for embracing and supporting me as one of their own.

Finally, for the sharp and critical clarity of her mind, her unwavering encouragement and love, I wish to reserve a most special appreciation for my wife, Nina.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

LIST OF ABBREVIATIONS

ATP	adenosine triphosphate
Bis	bisine
CH_3CN	acetonitrile
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
DTT	dithiothreitol
EDTA	ethylenediaminotetraacetic acid
FPLC	fast performance liquid chromatography
GBD	geldanamycin binding domain
GDM	geldanamycin
GST	glutathione-S-transferase
H_2O	water
HPLC	High pressure liquid chromatography
IPTG	isopropylthiogalactopyranoside
kDa	kilodalton
L	liter
LB	Luria broth
M	molar
mg	milligram
MIR	multiple isomorphous replacement
ml	milliliter
μl	microliter
mM	millimolar
m.w.	molecular weight
NaCl	sodium chloride
NCI	National Cancer Institute
NP40	nonionic detergent P-40
OD	optical density
PCR	polymerase chain reaction
pfu	thermal stable DNA polymerase from *Pyrococcus furiosus*
PMSF	phenylmethylsulfonyl fluoride
pmol	picomol
PAGE	polyacrylamide gel electrophoresis
rbs	ribosomal binding site
SCF	Skp-Cullin-F-box complex
SDS	sodium dodecyl sulfate
SIR	single isomorphous replacement
SOCS	suppressors of cytokine signaling
TFA	triflouroacetic acid
Tris	Tris-(hydroxymethyl)-aminomethane
VCB	VHL-ElonginC-ElonginB complex
wt	weight

PREFACE

It was necessary to make a choice regarding the inclusion of material describing the theory and practice underlying the methodology of macromolecular x-ray crystallography. Such material would include an analysis of the theory and experimental methods involved with both the crystallization of proteins as well as diffraction, phasing of diffraction data, and refinement of the molecular model. For two reasons, I have decided to omit any detailed discussion of these topics.

The first reason is a practical matter of space. This thesis, describing the crystal structures of two different macromolecular complexes, contains a substantial amount of information without the addition of a section on theory. It was thought to be in the best interests of the coherency of the dissertation to limit extraneous discussion and focus on the novel biological results revealed by these structures.

The second reason reflects the fact that the basic theory underlying modern crystallography has been understood for almost fifty years, and a plethora of texts by distinguished experts in the field discuss these issues to great depth. For these reasons, I feel it more appropriate to simply reference such material rather than endeavor to reproduce it here in what will surely be the less elegant presentation of an experimentalist.

I have therefore limited my discussion of crystallographic methods to those techniques which were directly relevant to the work described in this dissertation. In order to provide some reference to statistical quantities used to assess crystallographic data, I have also included in Appendix B a brief description of the reliability indices mentioned in the text.

CHAPTER ONE:

INTRODUCTION

1.1. Tumorigenesis

Over one hundred different diseases, involving pathologies in diverse tissues and with wildly variable patient prognosis, fall under the classification of cancer. Until the advent of modern molecular biology, the only common denominator in these myriad afflictions was the clear deregulation of tissue growth, leading to the formation of tumors. It is tumor formation in sensitive organs, and its subsequent spread to other tissues, which leads to most cancer-related mortality (Vogelstein and Kinzler, 1993; Bishop, 1995; Weinberg, 1996).

Cell growth in cancer involves not only the rapid division of tumor cells relative to healthy cells, but also the deregulation of growth in the context of the tissue in which the cell is embedded. Tumor cells ignore the signals from the extracellular environment and divide to form rebellious masses in the midst of an otherwise well-functioning organ. Tumors will also stimulate the formation of a supporting architecture which will allow them to survive independently of the tissue from which they developed. These components include the stimulation of a vascular network around and into the tumor which connects it to the nutritional and waste removal systems of the body (Vogelstein and Kinzler, 1993; Bishop, 1995; Weinberg, 1996).

It is now understood that tumor formation proceeds from the deregulated growth of a single progeny cell. In other words, tumors are

clonal, or at least would be if their tendency to diverge genetically, due to their high DNA mutation rate and propensity to produce gross chromosomal alterations, is disregarded. In fact, the diversity of cancer is distilled when the causes of cellular growth deregulation are examined, as it is found that a common set of genetic alterations underlie almost all known cancers. These alterations involve two broad classes of genes. One set involves genes which code for molecules which participate in driving the cell towards growth. The other set involves genes which code for proteins that inhibit cell growth. Although the development of tumors is dependent upon a cell accruing mutations in many different genes, the alteration of these two types of growth regulatory genes is a necessary and central event in tumorigenesis (Vogelstein and Kinzler, 1993; Bishop, 1995; Weinberg, 1996).

Tumor suppressors actively inhibit the growth of cells. Some, like the SMAD4 protein (Heldin et al., 1997; Kretzschmar and Massagué, 1998) do so by sensing signals from the extracellular environment, while others, like the p53 protein, seem to detect a variety of signals which require growth arrest, such as damage to the chromosomal DNA or the deregulation of growth stimulatory pathways in the cell (Vogelstein and Kinzler, 1993). Loss of function mutations in tumor suppressors therefore remove these inhibitory activities, and a cell with a deactivated tumor suppressor is primed for deregulated growth.

Proto-oncogenes function in pathways which promote cell growth, and do so in a regulated manner until gain of function mutations (or the overexpression of the proto-oncogene) cause their growth stimulatory activity to become abnormally high. Examples of proto-oncogenes include the well studied Ras molecule (involved in more than 25% of all cancers) whose

deactivation is often observed in colon, lung, and pancreatic cancers, Myc (involved in many leukemias), and Mdm2 (a protein which inactivates the p53 tumor suppressor) (Vogelstein and Kinzler, 1993; Bishop, 1995; Weinberg, 1996)..

Because alterations in tumor suppressors and oncogenes is so central an event in tumorigenesis, they have been actively examined by the biomedical research community. Each of the two molecules studied in the experiments described in this dissertation are members of one or the other of these two classes. The von Hippel-Lindau (VHL) tumor suppressor is a classic growth arrest factor whose inactivation can predispose may tissues toward malignancy (Kaelin and Maher, 1998). This property of VHL, and the fact that so little is currently known regarding its biochemistry in the cell, made it an attractive target for structural studies. The heat shock protein 90 (Hsp90) chaperone maintains many proto-oncogenes (and oncogenes) in an active state by modulating their folded structure (Xu and Lindquist, 1993; Aligue et al., 1994; Nathan and Lindquist, 1995; Chavany et al., 1996; Stepanova et al., 1996). It is thus likely to be a key player in tumorigenesis, and its activity to be necessary for the deregulation of many growth stimulatory pathways in the cell. The existence of the Hsp90 inhibitor, geldanamycin, established as an antitumor compound (DeBoer et al., 1970; Omura et al., 1979; Ono et al., 1982), made a structural analysis in the context of therapeutics very appealing.

1.2. The VHL Tumor Suppressor

VHL syndrome is characterized by the dominantly inherited predisposition to develop tumors of the central nervous system, kidney,

retina, pancreas, and adrenal gland (Melmon and Rosen, 1964). Affected individuals inherit an inactivated form of the gene in one allele, and formation of tumors depends upon a second inactivating mutation in the remaining wild-type VHL allele. VHL syndrome has a birth incidence of at least 1 in 36,000, almost total penetrance by age 65, and an average life span of 49 years (Melmon and Rosen, 1964).

Physicians began to make associations between various VHL manifestations, as well as between these symptoms and the inherited nature of the disease, as far back as the mid 19th century. At this time, scattered reports from ophthalmologists noted a correlation between the occurrence of retinal hemangioblastomas and similar lesions in the cerebellum. It was not until 1904, however, that the German ophthalmologist von Hippel recognized the inherited nature of the retinal tumors. In 1926, the Swedish ophthalmologist Avid Lindau connected the retinal and cerebellar hemangioblastomas, and classified them together as belonging to a single disorder (Choyke et al., 1995). In 1964, Melmon and Rosen published a foundational work in which they associated the tumors studied by von Hippel and Lindau with a much broader set of inherited cancers in diverse tissues such as the pancreas, adrenal gland, and spine (Melmon and Rosen, 1964). It was in this paper that the term "von Hippel-Lindau" began to be applied to this inherited syndrome. The gene altered in VHL syndrome was cloned by a collaborative effort between several groups at the National Cancer Institute (NCI) in 1993 (Latif et al., 1993).

Individuals who inherit an inactivated allele of the VHL gene possess a 70% probability of developing tumors of the kidney and central nervous system (Kaelin and Maher, 1998). The kidneys of VHL patients can be especially ravaged by the disease, containing on the order of 500 or more solid

tumors and over 1000 cysts (Pause et al., 1997). Over 80% of the mortality related to VHL syndrome is due to metastatic clear cell renal carcinoma or from cerebellar hemangioblastoma (Choyke et al., 1995). The tendency of individuals with VHL syndrome to continually develop new tumors in a variety of sensitive tissues implies the continual need for surgical and chemotherapeutic intervention. The financial and emotional costs of the disease to the individual are usually enormous.

VHL syndrome is caused by germline mutations in the VHL tumor suppressor, a 213 amino acid protein expressed in a wide variety of tissues (Latif et al., 1993 ; Whaley et al., 1994 ; Crossey et al., 1994 ; Linehan et al., 1995). In addition to the familial syndrome, VHL is also inactivated in about 80% of sporadic clear cell renal carcinomas (RCC), the predominant form of kidney cancer, as well as in the majority of sporadic hemangioblastomas of the cerebellum (Gnarra et al., 1994 ; Shuin et al., 1994 ; Foster et al., 1994 ; Herman et al., 1994). VHL related cancers affect upwards of 30,000 people a year (Linehan et al., 1995).

The importance of the VHL tumor suppressor in RCC is underscored by the fact that reintroduction of the wild-type VHL gene into certain normally tumorigenic VHL (-/-) renal carcinoma cells prevents the formation of tumors when such cells are grafted into nude mice (Iliopoulos et al., 1995 ; Gnarra et al., 1996). Growth suppression was also observed in cell culture, provided that the cultures were grown as multicellular spheroids (Lieubeau-Teillet et al., 1998) rather than as monolayers, for which inconsistent growth suppressive results have been obtained (Iliopoulos et al., 1995).

VHL-/- tumors are highly vascularized, and VHL null mice die during embryogenesis with defects in placental vasculogenesis (Gnarra et al., 1997).

In accordance with these observations, recent studies have shown that VHL negatively regulates the production of hypoxia-inducible factors such as the angiogenic vascular endothelial growth factor (VEGF) and the GLUT1 glucose transporter (Gnarra et al., 1996 ; Iliopoulos et al., 1996 ; Siemeister et al., 1996). VHL-/- cells produce high levels of these factors, and reintroduction of VHL downregulates them under normoxic conditions (Gnarra et al., 1996; Iliopoulos et al., 1996; Siemeister et al., 1996). The mechanism of this VHL effect is not yet understood, but appears to be mediated, in part, by the destabilization of hypoxia-regulated mRNA transcripts (Gnarra et al., 1996; Iliopoulos et al., 1996; Siemeister et al., 1996); it also requires the trafficking of VHL between the nuclear and cytoplasmic compartments (Lee et al., 1999). In addition, VHL-/- cells show high levels of other growth factors, including platelet-derived growth factor B (PDGF B), and transforming growth factor α (TGF-α), which are likewise down regulated upon reintroduction of wild-type VHL (Iliopoulos et al., 1996 ; Knebelmann et al., 1998).

VHL has also been linked to the regulation of the cell cycle. In normal cells, the withdrawal of serum from the culture media induces cell cycle arrest, and this process is linked to the stabilization of the G1/S cyclin-dependent kinase inhibitor $p27^{Kip1}$ (Pause et al., 1998). Cultured RCC cells lacking functional VHL not only fail to exit the cell cycle upon serum withdrawal, but also fail to stabilize p27. Stabilization of p27 and cell cycle arrest occur upon the reintroduction of wild type VHL in these tumor derived cell lines (Pause et al., 1998).

VHL is not homologous to any known protein. Initial biochemical studies revealed that VHL binds the ElonginC and ElonginB proteins, forming a ternary complex with a central role in VHL function (Duan et al., 1995; Kibel

et al., 1995; Kishida et al., 1995). Mutagenesis and peptide mapping studies showed that a 12 amino acid VHL region that contains nearly a quarter of the tumor-derived mutations makes key interactions with ElonginC (Kibel et al., 1995), and most, though not all, of the tumor derived mutations tested disrupted the VHL-ElonginC-ElonginB complex (Duan et al., 1995; Kibel et al., 1995; Lonergan et al., 1998; Kishida et al., 1995).

Recently, the ternary VHL-ElonginC-ElonginB complex (henceforth called the VCB complex) was shown to associate with Cul2, a member of the cullin family, in part through ElonginC-Cul2 binding (Pause et al., 1997; Lonergan et al., 1998). Cullins are thought to be involved in protein degradation because Cul1 is part of the multiprotein SCF ubiquitin ligase system that targets many cell cycle regulatory proteins for ubiquitin-mediated proteolysis (Kipreos et al., 1996; Lyapina et al., 1998; Feldman et al., 1997; Skowyra et al., 1997; Lisztwan et al., 1998). In addition to Cul1, the SCF complex contains (i) Skp1, which shares sequence homology with ElonginC, and (ii) a protein from the F-box family, which confers specificity for the target to be degraded (Bai et al., 1996; Zhang et al., 1995). The role of Cul2 in the VHL tumor suppressor pathway is reinforced by the requirement of the VCB-Cul2 complex for the VHL-dependent destabilization of VEGF mRNA, and by the inability of tumor-derived VHL mutants to form this quaternary complex (Pause et al., 1997; Lonergan et al., 1998).

ElonginC, in association with ElonginB, also binds to the suppressor of cytokine signaling (SOCS) superfamily (Starr et al., 1997; Hilton et al., 1998) of proteins that share the 40 residue SOCS-box motif (Kamura et al., 1998), and to the ElonginA transcriptional elongation factor (Aso et al., 1995). The SOCS-

box and ElonginA have homology to the 12 residue region of VHL that is critical for ElonginC-ElonginB binding.

ElonginB and ElonginC associate with ElonginA to form the SIII (Elongin) transcriptional elongation complex which suppresses transient pausing of RNA polymerase II (Aso et al., 1995). The binding of ElonginB and ElonginC to ElonginA results in a 40-fold stimulation of Elongin activity in transcription-runoff assays (Aso et al., 1995). It was demonstrated that wild type VHL could inhibit such *in vitro* transcriptional elongation assays (Duan et al., 1995), presumably by titrating the elongins B and C away from ElonginA. However, the vast excess (Kamura et al., 1998) of ElonginB and ElonginC in the cell (100 to 1000-fold higher than that of ElonginA and VHL), the lack of demonstrable specificity in the elongation inhibition assays, and the dearth of evidence indicating that this inhibition is significant *in vivo*, have raised doubts as to whether the primary function of VHL in the cell is to inhibit transcriptional elongation.

The search for VHL-interacting proteins revealed the possibility of an *in vivo* association with the extracellular matrix protein fibronectin (Ohh et al.,1998). VHL (-/-) RCC cells, like most transformed cells in general, exhibit defects in the fibronectin matrix assembly. Reintroduction of wild type VHL partially restores proper fibronectin matrix formation (Ohh et al.,1998).

Finally, some experiments suggest that VHL regulates the transcription of VEGF (Mukhopadhyay et al., 1997), potentially through inhibiting protein kinase C phosphorylation of the transcription factor Sp1 (Pal et al., 1997 ; Pal et al., 1998). These reports show evidence for the formation of a complex containing VHL, PKC, and Sp1 (Mukhopadhyay et al., 1997; Pal et al., 1997 ; Pal et al., 1998).

1.3. The Hsp90 Chaperone

The heat shock response consists of the activation or enhanced expression of genes which allow the cell to respond to harmful environmental conditions such as elevated temperature. The response is conserved across all biological kingdoms, and many heat shock proteins themselves are highly conserved across widely divergent species. Most, if not all, of the heat shock proteins are molecular chaperones, and thus function to prevent stress-induced aggregation of misfolded polypeptides or to assist them to obtain the native state (Hartl, 1995; Hartl and Martin, 1995; Scheibel and Buchner, 1998).

Although many proteins can be refolded *in vitro*, it is not clear that the partially folded intermediates along the path to the native conformation, some of which expose hydrophobic regions destined for the core of the molecule, can efficiently fold unassisted within the crowded conditions in the cell, where protein concentrations are on the order of those found in macromolecular crystals. Molecular chaperones evolved to serve two primary functions (Hartl, 1996): 1) to prevent the aggregation of unfolded or partially folded proteins (accomplished typically through protecting exposed hydrophobic residues), and, 2) to actively assist in the folding of a molecule toward the native conformation. Examples of the first class include DnaJ and DnaK, while the classic example from the second class is the GroEL-GroES chaperone (Hartl, 1995; Hartl and Martin, 1995). Recently, a third class of chaperones, defined by the discovery of heat shock protein 33 (Hsp33), have been uncovered. Hsp33 appears to function during oxidative stress, harboring highly reactive cysteines which, in the presence of oxidizing conditions, form disulfide bonds

activating the chaperone which then protects other macromolecules from oxidation (Jakob et al., 1999).

It has recently become clear that under normal conditions some chaperones also serve other functions in the cell, namely as regulators of activity through the modulation of a molecule's folded state. In this role, chaperones shepherd several classes of proteins which are unable to achieve or maintain activity unassisted, requiring either the presence of the chaperone or a co-factor to achieve conformational maturation and/or stability (Rutherford, 1994).

Heat shock protein 90 (Hsp90) is one of the more abundant proteins in eukaryotic cells, and its protein levels account for approximately 2% of the total cellular protein even under non-stress conditions (Buchner, 1996). Although not an oncogene, Hsp90 may be classified as a molecule which can lead to the development of cancer, and one which if inactivated in tumors could impair tumor growth. Hsp90 enjoys the curious and important status of serving as a chaperone for several different oncogenes, and thus represents a member of a higher order class. It is this position of Hsp90 in regulating a large class of proto-oncogenes which makes it an attractive target for therapeutic compounds, as its inactivation would render many possible causes of tumor growth inactive (Scheibel and Buchner, 1998).

Geldanamycin, and its closely related analogues herbimycin and macbecin are naturally occurring antitumor antibiotics (DeBoer et al., 1970; Omura et al., 1979; Ono et al., 1982). In the NCI's *in vitro* screen for antitumor agents, geldanamycin has shown potent activity, achieving 50 % growth inhibition (GI_{50}) at concentrations as low as 13 nM against the most responsive cell lines, with a mean GI_{50} of 180 nM against all 60 of the tumor cell lines

(Supko et al., 1995). Geldanamycin is active in mouse tumor models as well (Sasaki et al., 1979; Ono et al., 1982), and it has been selected for preclinical development as an antitumor agent by the NCI (Supko et al., 1995).

The antitumor effects of geldanamycin likely result from its ability to deplete cells of two broad classes of growth-regulatory signaling proteins: (i) proto-oncogenic protein kinases, including the erbB2 (Miller et al., 1994; Chavany et al., 1996) and EGF (Murakami et al., 1994) receptor tyrosine kinases, the v-src family of non-receptor tyrosine kinases (June et al., 1990; Xu and Lindquist, 1993; Whitesell et al., 1994; Hartson and Matts, 1994), and the Raf-1 (Schulte et al., 1995; Schneider et al., 1996) and CDK4 Ser/Thr kinases (Stepanova et al., 1996), whose overexpression, or otherwise deregulation, has been observed in diverse human cancers (Bouchard et al., 1989; Hunter and Pines, 1994; Tronick and Aaronson, 1995); and (ii) the nuclear hormone receptor family, including the estrogen and androgen hormone receptors (Smith et al., 1995; Whitesell and Cook, 1996; Nair et al., 1996), which can drive the growth of hormone dependent cancers of the breast (Osborne et al., 1980) and prostate (Isaacs and Coffey, 1979), respectively.

Initially thought to be a nonspecific kinase inhibitor, geldanamycin's target has only recently been identified as the heat shock protein Hsp90 (Whitesell et al., 1994), and its endoplasmic reticulum homologue GP96 (Chavany et al., 1996). In eukaryotes, Hsp90 has dual chaperone functions participating in the conformational maturation of nuclear hormone receptors and the aforementioned protein kinases, as well as in the cellular stress response (Bohen and Yamamoto, 1994; Jakob and Buchner, 1994; Pratt and Welsh, 1994). These two processes are likely to have in common the ability of Hsp90, in cooperation with Hsp70 and other factors, to prevent protein

aggregation and mediate the ATP-dependent refolding of heat-denatured proteins *in vitro* and *in vivo* (Freeman and Morimoto, 1996; Schneider et al., 1996). [*]

The best studied Hsp90-mediated conformational maturation is that of the nuclear hormone receptors, which require the Hsp90 system in order to acquire or maintain a state competent to bind hormone (Bresnick et al., 1989; Picard et al., 1990; Nathan and Lindquist, 1995). In this ATP-dependent reaction, an initial hormone receptor complex that contains Hsp90, Hsp70 and at least two co-chaperones, p60 and Hip (p48), is in equilibrium with a complex in which the receptor is in a metastable, nearly mature state competent to bind hormone (Smith et al., 1995; Dittmar et al., 1996). This nearly mature complex lacks Hsp70, p60 and Hip, but contains two new proteins, p23 and one of the three large immunophilins, FKBP52, FKBP54, or CyP40 (Smith et al., 1995; Dittmar et al., 1996). Upon hormone binding, the receptor is released as an active transcription factor (Smith et al., 1995; Dittmar et al., 1996). Treatment with geldanamycin appears to block the conversion to the nearly mature complex, preventing hormone binding and activation, and results in the degradation of the hormone receptor (Whitesell and Cook, 1996).

The requirement for Hsp90 in the activation of the aforementioned protein kinases has been demonstrated both genetically (Xu and Lindquist, 1993; Aligue et al., 1994; Nathan and Lindquist, 1995) and biochemically (Hartson and Matts, 1994; Chavany et al., 1996; Stepanova et al., 1996). This reaction is not as well understood, but it appears to involve a subset of the

---

[*] Other Hsp90 family members have been found bound to mutants of p53 (Sepehrnia et al., 1996) as well as to the retinoblastoma protein during mitosis (Chen et al., 1996). For example, GP96 (GRP94/endoplasmin) is localized to the endoplasmic reticulum and is involved in the assembly of immunoglobulin proteins (Melnick et al., 1992) and the antigenic peptide-induced maturation of MHC class I molecules (Li and Srivastava, 1993).

molecules found in the complexes with hormone receptors, as well as a protein kinase-specific cofactor, $p50^{Cdc37}$ (Whitelaw et al., 1991; Stancato et al., 1993; Cutforth and Rubin, 1994; Stepanova et al., 1996). For some of these kinases, such as Raf-1, Hsp90 association is a prerequisite for their trafficking to the plasma membrane (Schulte et al., 1995); for others, the mechanism of Hsp90-mediated activation is not yet understood. Again, geldanamycin interferes with Hsp90 function and induces the rapid degradation of these kinases by the proteosome system (Murakami et al., 1994; Schulte et al., 1995; Schneider et al., 1996; Stepanova et al., 1996).

In addition to its role in the conformational maturation of signal transduction molecules, the eukaryotic Hsp90 system participates in the refolding of certain thermally denatured polypeptides during the recovery of cells from heat stress (Borkovich et al., 1989). Using firefly luciferase as a model substrate, it was demonstrated that Hsp90 cooperates in this process with Hsp70, Hsp40 (a DnaJ homologue), p60, Hip and p23 (Schumacher et al., 1994; Schneider et al., 1996). Thus, the Hsp90 complexes formed with the thermally denatured polypeptide resemble those of hormone receptors and protein kinases. Geldanamycin inhibits the Hsp90-mediated refolding of luciferase, both in vitro and in vivo, and as a result luciferase is retained in the Hsp90 complex in an unfolded, degradation-sensitive state (Schneider et al., 1996); this shifts the balance from refolding to the proteolytic degradation of the protein (Schneider et al., 1996). Such a mechanism might also explain

how, under non-stress conditions, geldanamycin and its homologues cause the degradation of the signaling molecules. [*]

[*] In addition to geldanamycin and other benzoquinone ansamycins, the antifungal compound radicicol, a macrocyclic drug, also suppresses cellular transformation by binding to and inhibiting the Hsp90 chaperone activity toward signal transduction kinases (Sharma et al., 1998; Schulte et al., 1998; and Roe et al., 1999).

CHAPTER TWO:

MATERIALS AND METHODS

2.1. Reagents and equipment

Unless noted, all reagents used in the cloning and biochemistry of these projects were purchased from commercial distributors. In particular: restriction enzymes and their buffers were from New England Biolabs; salts, buffering agents, glycerol, DTT (dithiothreitol), thrombin (purchased as a semi-pure powder from bovine sources and purified in the laboratory on a heparin-sulfate resin), agarose, protease inhibitors, and detergents were obtained from either Sigma-Aldrich or Fisher Scientific, with no appreciable alteration of experimental results between the reagents from either company; polyethylene glycol compounds were obtained from Fluka (now part of Sigma-Aldrich); oligonucleotides were synthesized at Genelink; SDS-PAGE reagents were from BioRad; DNA markers for agarose gels were obtained from Promega and Gibo-BRL, as was geldanamycin; FPLC columns and resins were from Amersham-Pharmacia, except for the heparin-sulfate resin from TosoHaus; and crystallization screens and supplies were purchased from Hampton.

Equipment utilized consisted of standard SDS-PAGE and agarose gel casters and boxes from BioRad, an FPLC system with a model LCC-501 Plus Liquid Chromatography Controller running two P-500 pumps from Amersham-Pharmacia, an HPLC system with a Waters 486 Tunable

absorption detector and a Waters 600E system controlle running two Waters 600 pumps (Millipore), a model 8452A diode array uv/visible spectrophotometer from Hewlett-Packard, and a Nikon SMZ-U optical microscope. Diffraction data were collected on an R-AXISIIC imaging plate detector (Molecular Structure Corporation) mounted on a Rigaku 200HB x-ray generator for the Hsp90 project. For VHL, data for the MIR solution were collected at the National Synchrotron Light Source (NSLS) at Brookhaven National Laboratories on a prototype CCD detector manufactured by Mar-Research. The final native data set used in the refinement was collected at the Cornell High Energy Synchrotron Source (CHESS) on a 2x2 mosaic ADSC Quantum-4 CDD detector. Crystallographic programs were executed on unix-based Silicon Graphics workstations harboring processors of the R8000 or R10000 generation.

2.2. Biochemistry and Crystallography

The materials and methods utilized in the cloning, expression, purification, crystallization, and structural determination of both the VCB ternary complex and the Hsp90 apo and Hsp90-GDM complexes are involved and lengthy. Therefore, in order to preserve some continuity in the discussions of these two different crystallographic projects, separate chapters have been allocated summarizing the materials, methods, and results specific to each. Even this division was not quite sufficient, and the cloning, expression, and purification protocols associated with the VHL project were summarized in Chapter 3, but the details of these topics are withheld until Appendix A. Generally, the cloning of various genes was undertaken by

methods which have become standard in the current molecular biological techniques, involving amplification of constructs by PCR and site specific ligation of the inserts using restriction endonuclease digested ends. Other biochemical aspects of the projects were more involved and are described in later chapters.

Both projects involved structural solutions using the multiple isomorphous replacement (MIR) method originally developed by Max Peurtz in elucidating the structure of hemoglobin (Bragg and Perutz, 1954). The modern variants are well reviewed in several textbooks (Blundel and Johnson, 1976; Stout and Jenson, 1989; Drenth, 1994). Alternate Hsp90 crystal forms were solved by the method of molecular replacement (Rossmann and Blow, 1962) with the addition of Patterson correlation (PC) refinement (Brünger, 1990). The use of modern implementations of density modification as originally developed by Wang (1995) was important in the structural determination.

2.3. Testing of Cross-reactivity of SCF and VCB Components

These studies, discussed in Chapter Five, were conducted utilizing the co-expression scheme which produced the soluble VCB ternary complex (discussed in Chapter Three). ElonginB, Skp1 and GST-VHL were co-expressed in bacteria, and GST-VHL was isolated with glutathione beads, and the elution (10mM glutathione) was analyzed by SDS-PAGE. No Skp1 was found to co-purifying with GST-VHL; this suggested that VHL has no any affinity for Skp1. Conversely, ElonginB, ElonginC and GST-Skp2 were co-

expressed in bacteria and similarly purified on glutathione beads. Although a small amount of the ElonginC-ElonginB complex associated with GST-Skp2; upon further purification, this complex did not persist like the VCB complex would through anion-exchange chromatography, indicating that any ElonginC-Skp2 interaction would be very weak relative to the ElonginC-VHL or the Skp1-Skp2 interaction. ElonginC (17-112), the fragment used in these experiments, serves as a useful marker in bacterial lysates, as it runs anomalously small at approximately 6kD at a location devoid of bacterial proteins in BL21(DE3) strains. On the anion exchange gradient, the fractions were monitored for the presence of ElonginC. In the elution with a gradient in NaCl, a small peak was observed harboring ElonginC and what appeared to be stoichiometric amounts of ElonginB. Western blots of these fractions analyzed by SDS-PAGE (using a polyclonal antibody to an N-terminal epitope of Skp2), indicated that there was no corresponding peak of Skp2 co-eluting with ElonginC. Although there was a Skp2 fragment staining in these fractions, this band was of equal intensity over a range of salt concentrations, not tracking ElonginC.

2.4. Sequence threading

Sequence threading was performed using the program THREADER2 (Miller et al., 1996) with equal weights for threading and sequence alignment, and was performed against library of 1925 folds augmented with the folds of the VHL α and β domains and of ElonginC. The sequences of Skp1 (1-112), the SOCS1 SOCS box (170-211), and the Skp2 F box (107-147) yielded Z scores

of 3.7 (ElonginC structure ranked first), 2.5 (VHL α domain ranked first), and 1.6 (VHL α domain ranked third), respectively. In threading analyses of additional F box sequences we also utilized a randomization test based on sequence shuffling (Miller et al., 1996). The best alignment was obtained with the CyclinF F box which ranked the VHL α domain first with a score of 1.6. Most F boxes produced consistent alignments with the VHL α domain, with the exception that some lacked the one residue insertion between the H2 and H3 helices. ElonginA sequences encompassing both the region of homology with the VHL H1 helix and the proposed F-box sequences (residues 545-610) were threaded as above, and yielding a Z score of 0.85 (704^{th}).

2.5. Molecular graphics

All molecular illustrations in this thesis were generated by computer programs. Most figures were designed and produced on a Silicon Graphics unix workstation and modified on an Apple Macintosh using the program Adobe Photoshop. The exceptions are sequence alignments which were generated by programs on the Apple Macintosh which are part of the DNASTAR suite (DNASTAR Corporation), using sequences from Genebank downloaded from the internet using the web browser Netscape (Netscape Corporation). These alignments were then imported into the program Canvas (Deneba Systems) and annotated with symbols and line drawings. Ribbon diagrams of proteins and atomic models of side chains were produced with Molscript (Kraulis, 1991). Rendered images were produced by converting Molscript files to a format compatible with Raster3D (Merrit and Murphy, 1994), and rending them with this latter program. Illustrations of solvent

accessible surface area were produced using GRASP (Nicholls et al., 1991). Ramachandran plots and surface accessibility graphics used with sequences were generated by PROCHECK (Laskowski et al., 1993). Finally, images of electron density were produced by converting plot files from O (Jones et al., 1991) to Protein Databank format, and generating the figure with Molscript.

CHAPTER THREE:

VHL-ElonginC-ElonginB EXPERIMENTAL RESULTS

3.1. Purification and crystallization of the VCB ternary complex

The production of the ternary complex was complicated by the insolubility of VHL and ElonginC, and to a lesser extent, ElonginB, in bacteria (*E. coli*). Therefore, initial biochemical work was performed with recombinant human VHL, ElonginB, and ElonginC which were individually expressed in bacteria, isolated from the cell pellet by solubilization in 6.4 M guanidine-HCl or 8 M urea, and purified by reverse-phase HPLC or a nickel-chelating affinity resin. Initial work with full-length VHL utilized the affinity method with hexahistidine tags on all three proteins. Material purified by this method served for the domain delineating proteolytic digestion experiments (described below). Subsequent work with the truncated VHL fragment (also described below) followed the HPLC protocol and utilized untagged Elongin proteins and a hexahistidine tag for VHL (present only to insure adequate expression levels of the polypeptide). The lyophilized peaks for each polypeptide from HPLC were solubilized in 8M urea, mixed in a near 1:1:1 molar ratio, and reconstituted by dialysis into a native buffer. Productively refolded material was separated from that which was misfolded and partially aggregated by size-exclusion chromatography and ion-exchange

chromatography. Appendix A provides the details of the expression and purification schemes for each stage of the project.

It should be noted that the gel-filtration and ion-exchange (Source Q resin from Pharmacia) steps were critical not for separating the ternary complex from other polypeptides, as the HPLC purification achieved remarkably good separation, but for separating misfolded aggregates of the ternary complex from those which had productively refolded in the dialysis. The dialyzed material was loaded on a Superdex 200 gel-filtration column (Pharmacia). Following a broad peak at high molecular weight which contains non-stoichiometric quantities of the three proteins, there is a sharp peak at a molecular weight of approximately 45kD, the correct mass for the ternary complex. The dialyzed material was also placed on an anion-exchange column as described in Appendix A. The elution profile revealed that several different species of the ternary complex could be found. In particular, there is a central peak in which most of the ternary complex material is located, followed by a broad shoulder out to high salt. When the peak and shoulder are separately pooled and loaded on a gel-filtration column, only the peak elutes at the appropriate molecular weight, whereas the shoulder elutes as an aggregate. In addition, two separate binary complexes can be seen. At low salt, the VHL-ElonginC complex, and at high salt the ElonginC-ElonginB complex.

In order to identify compact structural domains amenable to crystallization, the ternary complex was subjected to limited proteolysis with the non-specific protease subtilisin. The digestion produced a fragment of

VHL (comprising residues 54-213), identified by sequence analysis and mass spectroscopy (Memorial Sloan-Kettering Cancer Center Microchemistry Laboratory), which co-migrated with ElonginB and ElonginC on size-exclusion chromatography. This domain (Figure 3.1), which contains both the core region of homology with rodent VHL as well as the sites for more than 99% of tumor derived mutations, was subsequently shown to have equivalent biological activity to the full length protein, and to constitute a significant fraction of the VHL expressed in the cell (Schoenfeld et al., 1998; Iliopoulos et al., 1998). In the remaining studies, this alternative endogenous VHL polypeptide was used in the place of full length VHL.

In order to remove the possibility of structural heterogeneity potentially incurred during the *in vitro* refolding of the VCB complex, a scheme was devised by which the three proteins were co-expressed in *E. coli* (Figure 3.2), using BL21(DE3) cells. This strain encodes T7 RNA polymerase within its genomic DNA which is under the control of the *lac* operator. VHL and ElonginB were engineered into a standard expression plasmid (pGEX-4T-3 from Pharmacia) as a dicistronic message (see Appendix A), and this vector was co-transformed with a second plasmid harboring a compatible origin of replication and which contained the gene for ElonginC (pBB75 – Appendix A). Each plasmid contained a different antibiotic resistance gene, and the presence of both plasmids was maintained by dual antibiotic selection.

Although the individual proteins were insoluble in bacterial expression systems, a soluble ternary complex could be isolated from cells expressing all

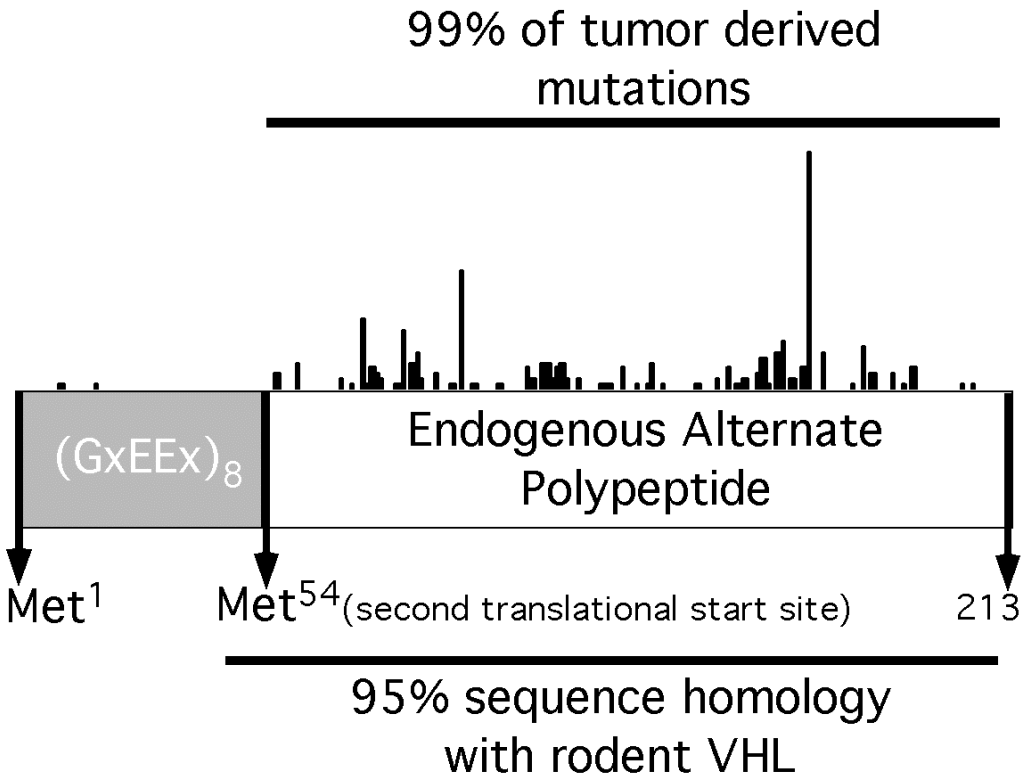


Figure 3.1. VHL construct in crystal structure

Illustration of the alternate translational protein product of VHL (white) in the context of the full length protein. Shown are the acidic repeat at the N-terminus (grey), the homologous region with rat and mouse VHL, and a histogram of missense mutational frequency.

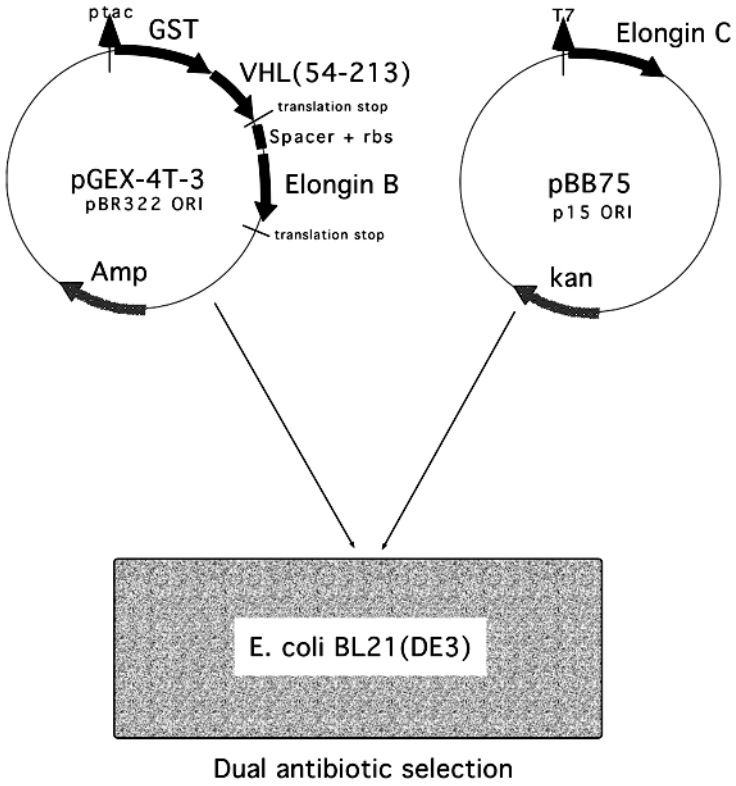


Figure 3.2. Bacterial co-expression scheme.

Circles represent the plasmid DNA with the approximate relative locations of the promotors (ptac and T7), the genes expressed (GST-VHL fusion, Elongin B, and Elongin C, and the **amp**icillin and **kan**amycin resistance genes), and the non-expressed DNA elements important for the construct (the spacer between the dicistronic VHL-ElonginB transcript and the second ribosomal binding site before ElonginB, noted "rbs"). ORI referes to origin of replication, the number preceding ORI denoting the family of ORI to which the plasmid belongs.

three genes. VHL(54-213) was expressed as a glutathione-S-transferase (GST) fusion protein, and the ternary complex was isolated on a glutathione affinity resin (Pharmacia) and separated from GST by site specific proteolysis with thrombin at a consensus peptide separating the two proteins (Figure 3.3). Digestion of the affinity purified ternary complex with the thrombin protease also produced a fortuitous, non-consensus but specific cleavage at residue 17 in ElonginC (as identified by N-terminal sequencing and mass spectroscopy). This fragment was not a product in the subtilisin digestion experiments. These first seventeen amino acids are not contained in the drosophila or yeast homologs of ElonginC, and the ternary complex formed with this fragment was as stable as full length. The removal of this N-terminal segment radically altered the chromatographic behavior of the ternary complex, causing its elution from anion exchange (Source Q resin from Pharmacia) to occur at an NaCl concentration of 100mM instead of 160mM as observed for the ternary complex with full length ElonginC.

This fragment of ElonginC (17-112) was cloned into pBB75 and co-transformed with the dicistronic GST-VHL(54-213) and full length ElonginB. While crystals could be obtained from the V(54-213)C(1-112)B complex, diffraction quality crystals only formed with the truncated ElonginC construct (see below). Following the cleavage of VHL from GST, the ternary complex was further purified by a second pass over the glutathione-sepharose resin, anion-exchange, and finally gel-filtration chromatography. The complex was concentrated by ultrafiltration to 25mg ml^{-1} in a buffer of 5mM Bis-Tris-Propane, 200mM NaCl, 2mM DTT, pH 7.0.

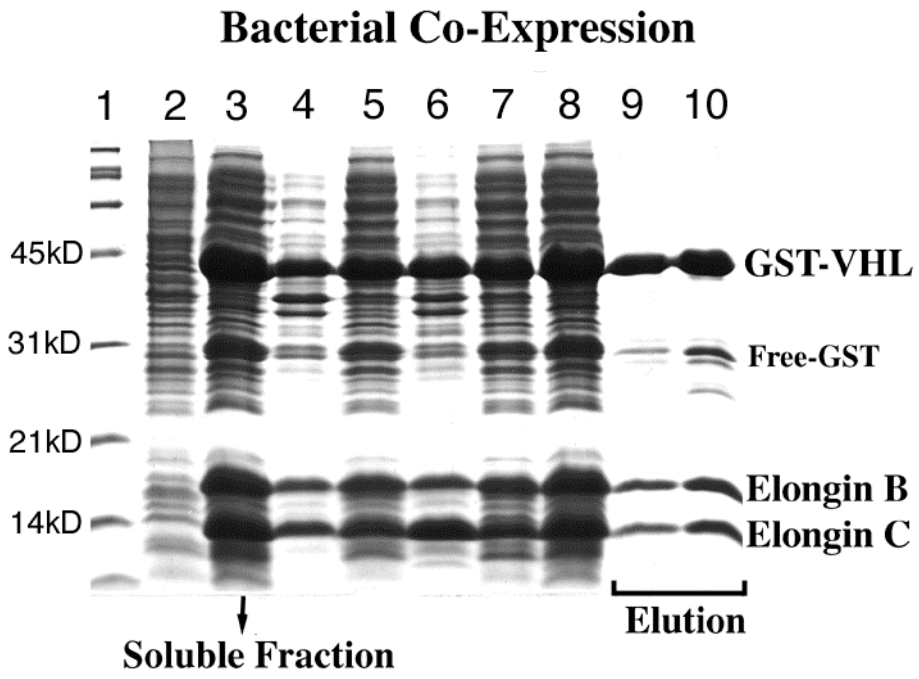


Figure 3.3. Soluble VCB complex visualized on SDS-PAGE

Description of gel lanes: (1) Broad-range molecular weight markers, (2) uninduced bacterial cultures, (3) 25°C induced, lysed, centrifuged bacterial fraction, (4) insoluble pellet from (3), (5) 37°C induced, lysed, centrifuged bacterial fraction, (6) insoluble pellet from (5), (7) flow through from initial pass of soluble lysate over 50ml glutathione-sepharose gravity column, (8) as (7) with second half of total lysate volume, (9) 5mM glutathione elution of protein from resin from column corresponding to lane (7) flow through, (10) as (9) but corresponding to lane (8).

Extensive screening with the refolded V(54-213)C(1-112)B ternary complex revealed that small, needle-like crystals could be obtained by the hanging-drop vapor-diffusion method (Figure 3.4) from 12-15% 2-propanol and 25mM Bis-Tris-Propane pH 7.0, as well as in identical buffering conditions with 10-14% polyethylene glycol monomethyl-ether 5000. Despite prolonged efforts to improve the crystal quality by systematically varying precipitant concentrations, solution pH, growth conditions, and the concentration of numerous additives, the crystals obtained from these VCB constructs were no larger than 50 μm in the largest dimension, and rarely larger than 10-20 μm in the smallest dimension, and were too small for effective crystallographic analysis.

Upon the development of the co-expression scheme and the obtaining of native ternary complex from bacteria, various combinations of polypeptides in the ternary VCB complex were examined. Following screening of various precipitating conditions, crystals of the V(54-213)C(17-112)B ternary complex were grown at 4°C by the hanging-drop vapor-diffusion method, from 12 - 17% polyethylene glycol 8000 (PEG8000) supplemented with 0.2 M magnesium acetate, 5 mM DTT, and 0.1 M sodium cacodylate, pH 5.7 (Native I, Table 3.1). The crystals grew as thin plates measuring about 50 μm in the smallest dimension. They formed in space group $P4_122$, with a=b=93.7Å and c=363.5 Å, and contained four VCB ternary complexes in the asymmetric

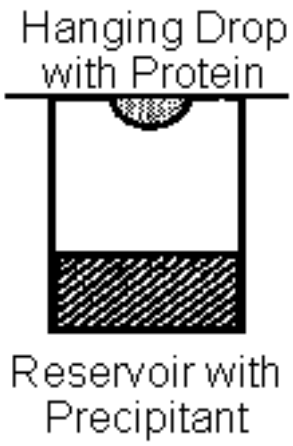


Figure 3.4. The Hanging-drop, vapor diffusion method.

The protein is mixed with the precipitating solution from the reservoir (well-buffer) and sealed over the well with vacuum grease to prevent the diffusion of vapor from the experiment.

unit (total of 170kD), with a solvent content of 43%.

Subsequent to the MIR solution, it was discovered that crystals of identical symmetry and unit cell (Native II, Table 3.1) could be grown from 10-15% polyethylene glycol m.w. 2000 Da (PEG2000), 0.2M magnesium acetate, and 0.1M sodium cacodylate, pH 5.7, but which were large in all dimensions, commonly obtaining edges in excess of 0.2mm (Figure 3.5). Data from this morphology (Native II) were used for the final model refinement.

The growth of well diffracting crystals was extremely sensitive to the pH of the precipitating solution. Crystals grown in the range of pH 5.0 to 5.8 have sharp edged crystal morphology and optimal diffraction. As the pH increased toward neutrality, the crystals quickly lost their sharp edges and produce much more mosaic diffraction to substantially lower resolution. By pH 7.0, crystals no longer grew reproducibly. The concentration of magnesium acetate also had a profound effect on crystal quality, with either 100mM above or below the optimal concentration of 200mM producing less ordered crystals plagued by twinning.

Even in the best conditions obtained, the crystals still display a tendency toward twinning, most often in the Native I (PEG8000) morphology. Typically, however, one could discern two independent crystal lattices in the diffraction pattern which could be integrated separately by the processing software, although the autoindexing algorithm used by DENZO would often

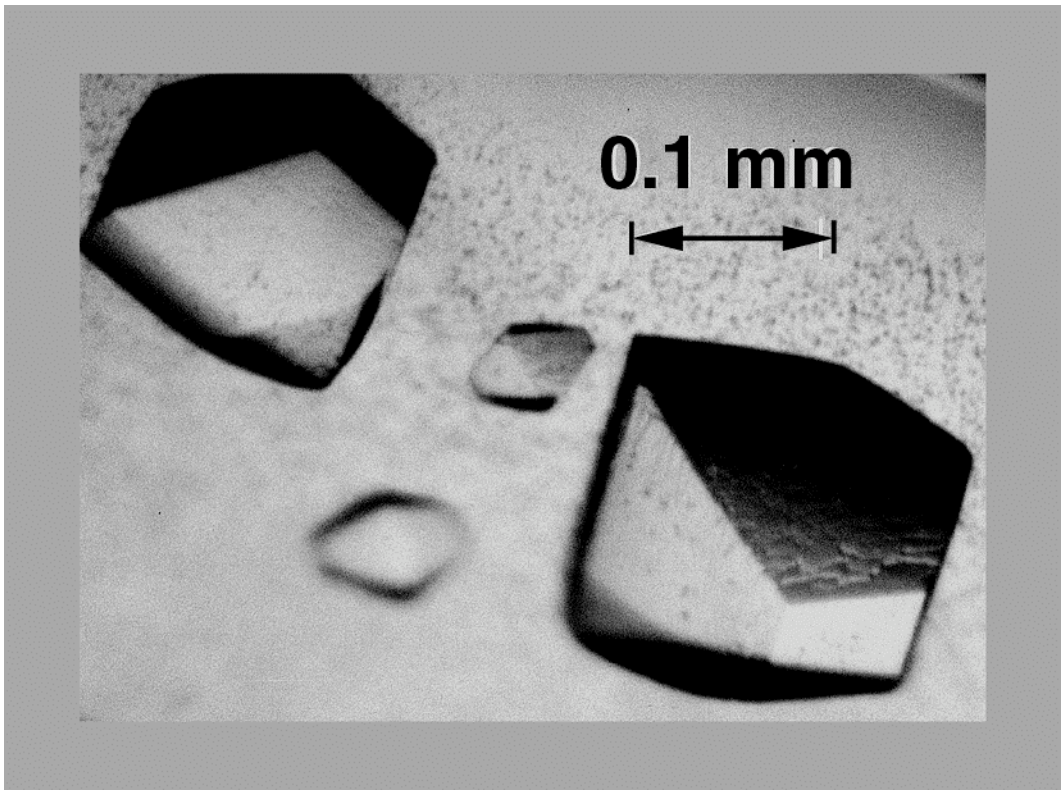


Figure 3.5. The Native II (PEG2000) VCB crystal form.

fail without extensive manual intervention, typically consisting of directly selecting the appropriate lattice for the set of reflections utilized. The integration and refinement of unit cell and orientation parameters was prone to instability, however, and extensively modified integration schemes (varying from crystal to crystal) were required to optimize the final data processing.

3.2. Crystallography of the VCB ternary complex

3.2.1. Data collection and the MIR Solution

Because of the large unit cell, the collection of data on a laboratory apparatus was impractical. In particular, the close spacing of reflections along the c* axis (c=363.5 Å) proved too difficult to separate in the laboratory. Therefore, all data sets used in the structural determination were collected at synchrotron radiation sources, where the low beam divergence and higher intensity allowed the data collection strategy to be engineered such that the spot overlap along the long axis was minimized and the final resolution of the data was maximized.

The data sets used in the multiple isomorphous replacement method were collected at Brookhaven National Laboratories at the X9B beamline. Data for both natives and derivatives were collected with crystals flash-cooled to -160°C in stabilization buffer supplemented with 15% glycerol. Heavy-atom soaks were performed in a stabilization buffer lacking DTT and containing 25% PEG8000, 75mM magnesium acetate, and 0.1M sodium cacodylate, pH 5.7, with 5mM thimerosal for 6 hours, 0.3mM K_2PtBr_6 for 1 hour, or 1mM $KAu(CN)_2$ for 4 hours. A summary of the diffraction data can

be found in Table 3.1. The diffraction data was processed with the programs DENZO and SCALEPACK (Otwinowski and Minor, 1997)

It should be noted that the term "stabilization" buffer is perhaps a misnomer for this solution. While the solution described above served for the short duration soaks performed, even non-derivatized crystals incubated in this buffer for more than a day were compromised. The transfer to the cryo-buffer was found to be optimized when performed as rapidly as possible. Therefore, crystals were flash-cooled in the following manner: 1) the drop containing the crystal was opened and 10-20 λ of stabilization buffer was added, 2) the crystal of interest was removed from the drop by a pipetman and transferred to 200 λ solution of stabilization buffer. This solution was then mixed with a pipet tip. Immediately after mixing was complete, the crystal was transferred to the cryo-buffer and mixed similarly. As soon as the solution was thoroughly mixed (no more than 20 seconds), the crystal was mounted in a nylon cryo-loop and flash-cooled in the gaseous stream. The incubation of the PEG8000 crystals for longer periods of time in the cryo-buffer resulted in a decrease in resolution. In contrast, the larger PEG2000 crystals required a more gradual exchange into the cryo-buffer, as the immediate transfer which proved best for the smaller crystals osmotically shocked the Native II form, causing their surface to erupt in a roadmap of fissures. Therefore, the larger crystals were transferred in three linear, discrete increments of glycerol and PEG2000 (to final 16% and 25%, respectively).

Table 3.1. Summary of VCB crystallographic data

		Reflections		Data Statistics		MIR Stats	
Crystal	Resolution (Å)	Unique	Measured	Coverage (%)	R_{sym} (%)	R_{iso} (%)	PhP
$KAu(CN)_2$	3.40	22,385	90,176	94.0	5.0	-	-
Thimerosal	3.50	19,433	70,784	91.4	7.4	32	1.20
K_2PtBr_6	4.20	10,067	32,616	80.9	7.4	13	1.11
Native I	3.80	15,729	48,137	93.3	9.1	14	0.81
Native II	2.70	41,219	123,622	90.9	7.0	-	-

Refinement and Analysis of VCB Molecular Model

Crystal	Resolution (Å)	Reflections ($\|F\|>2\sigma$)	Atoms modeled (protein/H_2O)	R_{cryst}/R_{free} (%, %)	r.m.s deviations bonds, angles, B-factors
Native II	20 - 2.70	38,609	10,874/ 454	23.9, 28.6	0.014Å,1.8°,3.4Å^2

$R_{sym}= \sum_h \sum_i \; | \; I_{h,i} - I_h \; | \; / \; \sum_h \sum_i I_{h,i}$, for the intensity ($I$) of i observations of reflection h. $R_{iso} = \sum \; | \; F_{PH} - F_P \; | \; / \; \sum \; | \; F_P \; |$; PhP = Phasing Power = $\sum [F_H(calc)^2 \; / \; (F_{PH}(obs) - F_{PH}(calc))^2]^{1/2}$ F_P = native structure factor amplitude; F_{PH} = derivative structure factor amplitude. $F_{PH}(obs)$, $F_{PH}(calc)$ = observed, calculated derivative structure factors, respectively. $R = \sum \; | \; F_P - F_{calc} \; | \; / \; \sum F_P$. F_{calc} = model structure factor; 5% data omitted for R_{free}. Bond and angle deviations are from ideal values; B-factor deviations are between bonded atoms. In the MIR analysis, the gold derivative was used as a native, and the other derivatives and native were modeled with negative occupancy sites.

The structure of the VCB complex was determined by the multiple isomorphous replacement (MIR) method combined with four-fold non-crystallographic symmetry (ncs) averaging (Table 3.1). Initial heavy atom positions for the gold derivative were determined in an automated fashion using SHELX-90 (Sheldrick, 1991), and were used together with strong native Patterson peaks to determine the three ncs operators, which were primarily translations (Figure 3.6). The subsequent derivative sites were obtained using difference Fourier maps. The gold derivative was used as a native due to its high quality data, and the remaining derivatives and Native I were modeled with negative occupancy gold "holes."

Initial phases (mean figure of merit 0.47) were calculated to 3.5 Å resolution with the program MLPHARE (CCP4, 1994). These phases produced maps which were not interpretable. The phases were improved by using density modification (solvent flattening and histogram matching) and multi-domain four-fold ncs averaging with the program DM (CCP4, 1994), eventually using individual masks for VHL, ElonginC, and ElonginB. A single mask (initially determined automatically by DM) resulted in good density for some portions of the ternary complex, but uninterpretable density for other regions. A polyalanine model was built into the those regions of clear electron density, and was used to define approximate locations for protein domains. As the model was extended in subsequent cycles of model building, the use of individual masks for each protein (and each domain of VHL) resulted in continuous and interpretable density across the complex (Figure 3.7).

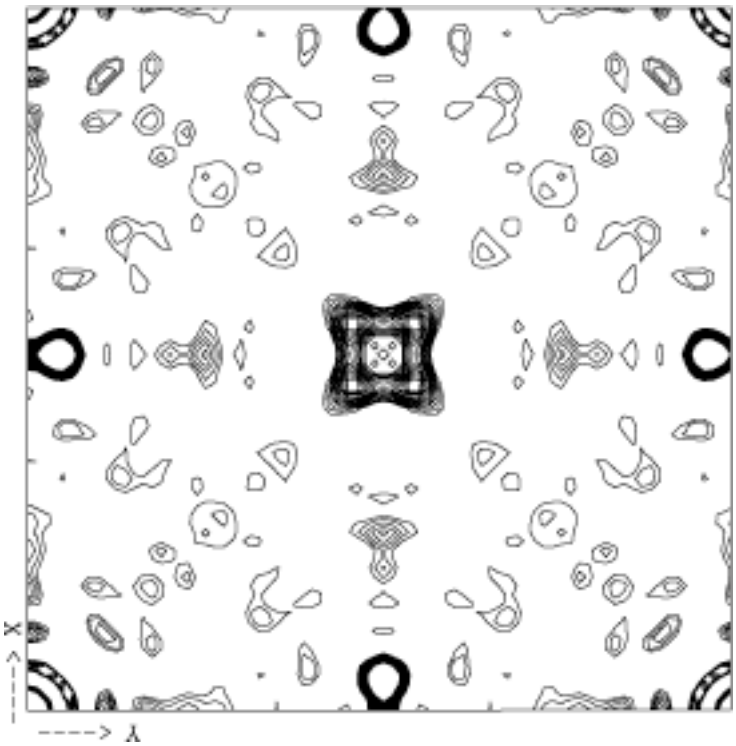


Figure 3.6. VCB Native Patterson

Z=0 section from self-Patterson function illustrating the peaks produced by the translational non-crystallographic symmetry. Peaks have been truncated at a cut-off value by the software and do not therefore indicate the full height in the contour map.

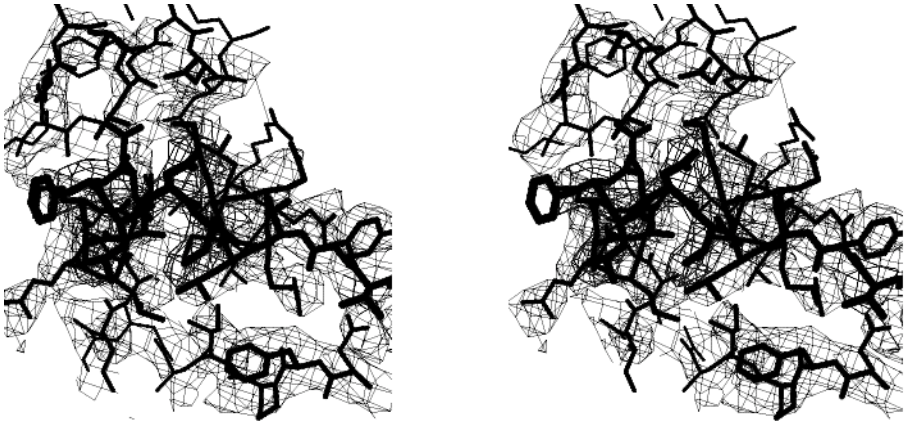


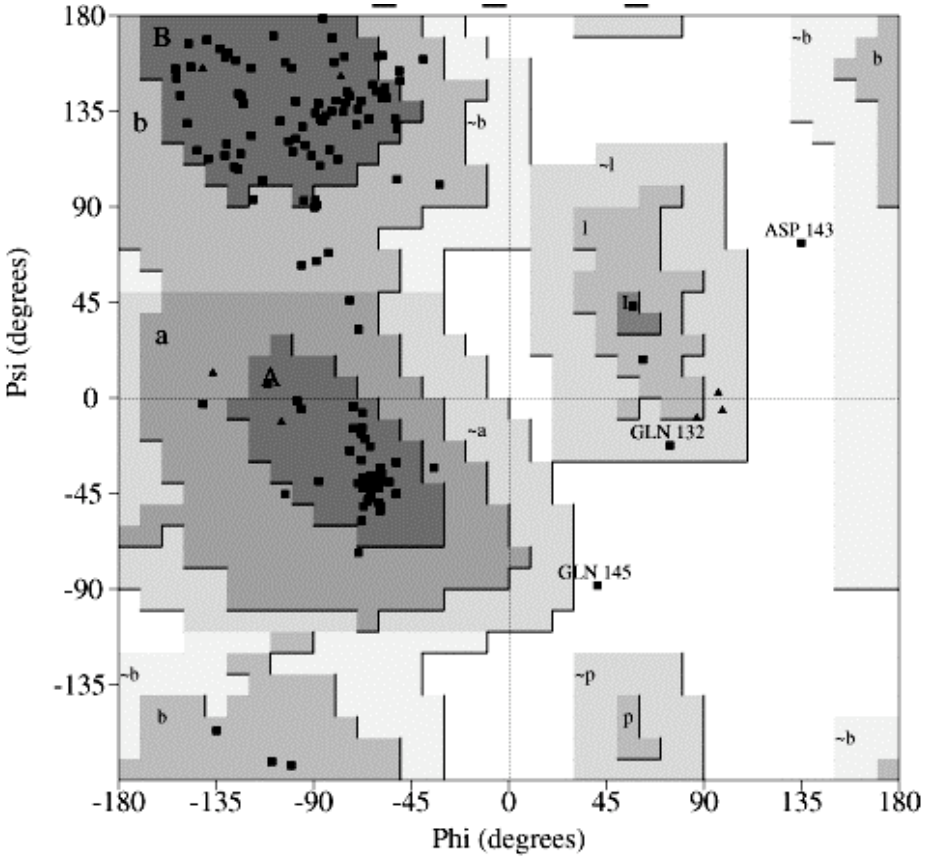
Figure 3.7. Multi-domain, 4-fold averaged, 2Fo-Fc map

Shown here is a stereo image of the four-fold ncs averaged electron density 2Fo-Fc map phased with the polyalanine model. Superimposed on the density is the final VHL model. The region shown is the α domain interaction with ElonginC (see Chapter Four).

3.2.2. Model Building and Refinement

An initial polyalanine model was built into the resultant maps with O (Jones et al., 1991) and refined using a maximum likelihood target function with strong ncs restraints (force constant = 500 kcal mol^{-1} $Å^{-2}$) with the program CNS (Brünger, et al., 1998). This initial model comprised most of ElonginB (excluding the extended tail), ElonginC, and the VHL β domain along with the H1 and H4 helices along with the L8 loop linking the two VHL domains. Model-phased maps calculated at 2.7 Å resolution were multi-domain averaged with the program RAVE (Jones, et al., 1991) and were used to assign the amino acid sequence and complete the model (Figure 3.7). Initially, the hydrophobic cores of each protein were interpreted, due to the presence of numerous bulky hydrophobic side chains which were more readily identifiable in the electron density maps. This allowed most of the connectivity to be established. It was noticed at this time that the ncs masks utilized were part of two distinct ternary complexes. In particular the α domain (see Chapter Four) of VHL needed to be associated with a different ncs group. This was accomplished, leading to the interpretation and building of the helical domain. Iterative cycles of model building and refinement were performed until the R_{free} no longer improved. The final model consists of over 10,000 atoms and has a crystallographic R_{free} of 28.6% to 2.7 Å resolution.. The following residues have poor electron density and are presumed to be disordered in the crystals: residues 54-62 and 205-213 from the N- and C-termini of VHL, respectively, residues 139-144 of VHL, residues 50-57 of ElonginC, and residues 80-87 and 99-120 of ElonginB. Outside these segments, there are no residues in the disallowed region of the Ramachandran

plot (Ramachandran et al., 1963; Ramachandran et al., 1968) and more than 85% of the residues are in the most favored regions (Figures 3.8, 3.9, 3.10 for VHL, ElonginC, and ElonginB, respectively). Table 3.1 summarizes the refinement statistics.

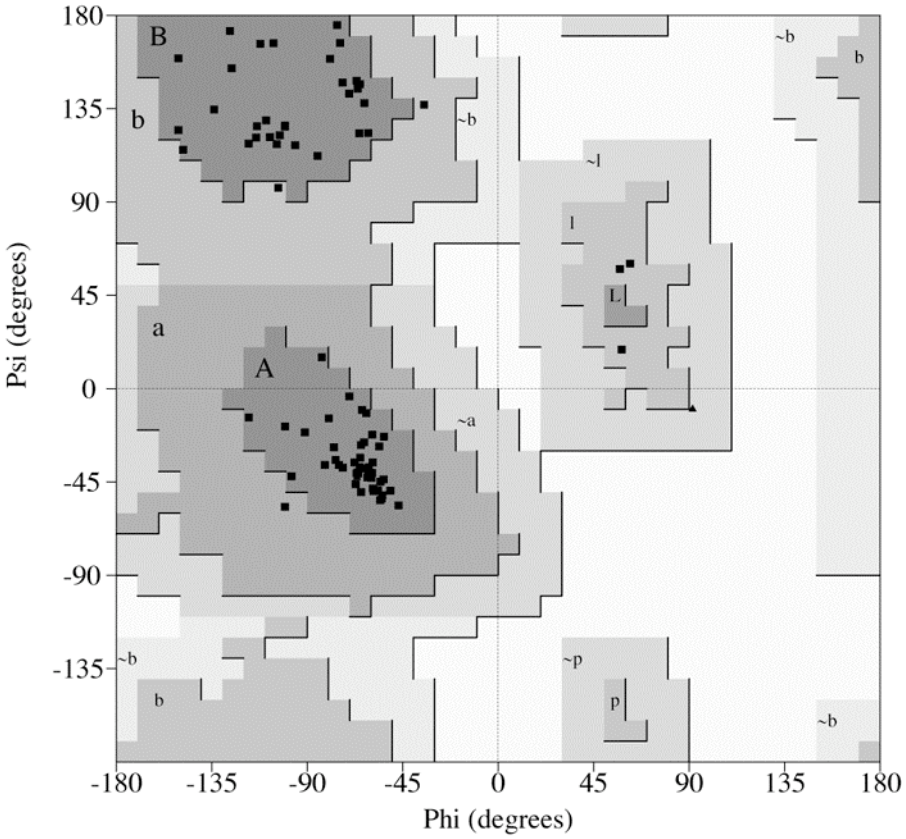


Plot statistics

Residues in most favoured regions [A,B,L]	96	80.0%
Residues in additional allowed regions [a,b,l,p]	21	17.5%
Residues in generously allowed regions [~a,~b,~l,~p]	1	0.8%
Residues in disallowed regions	2	1.7%
Number of non-glycine and non-proline residues	120	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	7	
Number of proline residues	13	
Total number of residues	142	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

Figure 3.8. VHL Ramachandran Plot

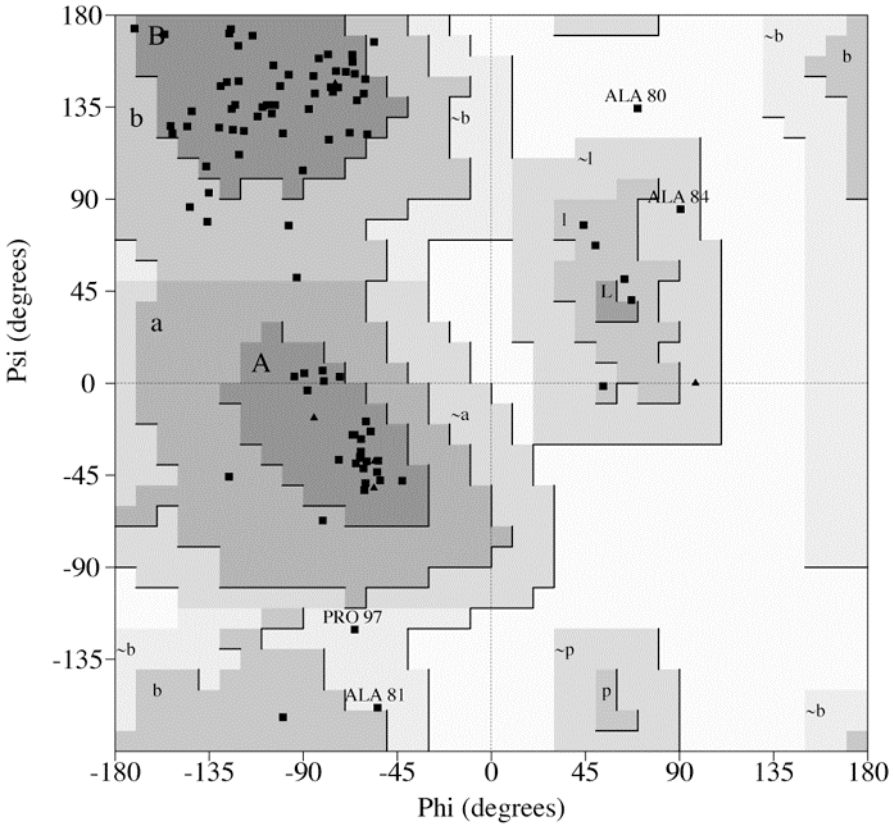


Plot statistics

Residues in most favoured regions [A,B,L]	72	92.3%
Residues in additional allowed regions [a,b,l,p]	6	7.7%
Residues in generously allowed regions [~a,~b,~l,~p]	0	0.0%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	78	100.0%
Number of end-residues (excl. Gly and Pro)	4	
Number of glycine residues (shown as triangles)	2	
Number of proline residues	4	
Total number of residues	88	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

Figure 3.9. ElonginC Ramachandran Plot



Plot statistics

Residues in most favoured regions [A,B,L]	66	78.6%
Residues in additional allowed regions [a,b,l,p]	15	17.9%
Residues in generously allowed regions [~a,~b,~l,~p]	2	2.4%
Residues in disallowed regions	1	1.2%
Number of non-glycine and non-proline residues	84	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	5	
Number of proline residues	7	
Total number of residues	98	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

Figure 3.10. ElonginB Ramachandran Plot

CHAPTER FOUR:

THE VHL-ElonginC-ElonginB STRUCTURE

4.1. Overview of the VCB ternary complex

In the crystal structure of this ternary complex, ElonginC binds ElonginB and VHL across two distinct interfaces, whereas VHL and ElonginB do not interact. The proteins are not co-linear, however, but instead form an overall shape reminiscent of a boomerang, with ElonginC at the center and ElonginB and VHL at its ends (Figure 4.1). The two interfaces are comparable in extent, each burying a total surface area of about 2000 $Å^2$.

VHL has two domains: an N-terminal β sandwich of roughly 100 amino acids (β domain) and a smaller α-helical domain (α domain), held together by two linkers and a polar interface. A large portion of the α domain surface, and a small portion of the β domain, interact with ElonginC. About half of the tumorigenic mutations target the α domain and its residues that contact ElonginC. The remaining mutations target the β domain, and significantly a β domain surface patch uninvolved in ElonginC binding. This indicates that two intact protein binding sites may be required for the tumor suppressor effects of VHL.

ElonginC folds into an α/β roll with three β strands packing against four α helices, and binds VHL through helices and loops at its carboxy-

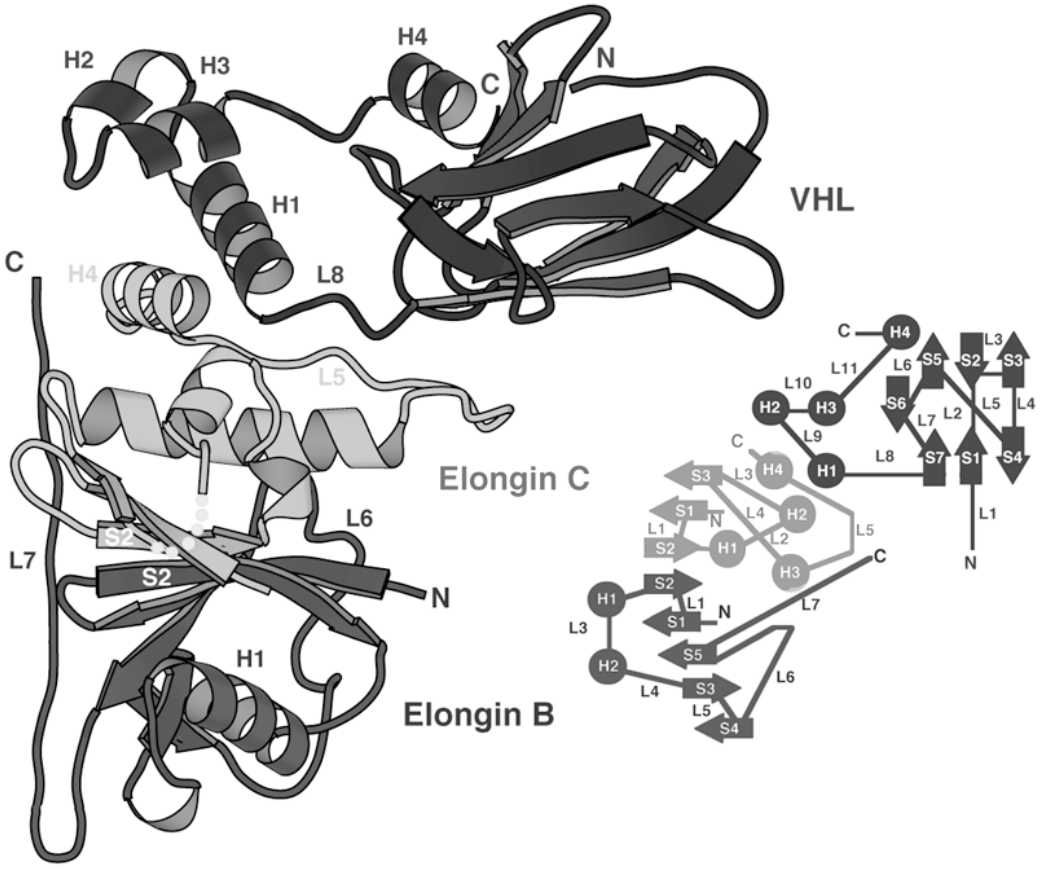


Figure 4.1. Overall view of the VCB ternary complex.

The VCB ternary complex consists of two binary interfaces: one between VHL and ElonginC, and the other between ElonginC and ElonginB. On the left is a ribbon diagram illustrating the secondary structure of the VCB complex. On the right is a topology diagram in which circles indicate helices and wide arrows indicate strands. Only the structural elements discussed in the text are labeled in the ribbon diagram.

terminus (Figure 4.1). ElonginB possesses an α/β roll structure very similar to that of ubiquitin, as expected from sequence homology (Garrett, 1995; Vijay-Kumar, 1985). The ElonginB-ElonginC interface is dominated by the juxtaposition of two β sheets to form a continuous intermolecular sheet. Although VHL and ElonginB do not interact, the carboxy-terminus of ElonginB runs along side ElonginC and reaches near VHL (Figure 4.1).

4.1.1. The Structure of the von Hippel-Lindau Tumor Suppressor

The amino-terminal β domain of VHL (residues 63-154) consists of a seven-stranded β sandwich and an α helix (Figure 4.1, 4.3A). One sheet is comprised of strands S1, S4, S6, and S7, while the other half of the sandwich contains a sheet with S2, S3, and S5. The 4+3 β sandwich has a Greek key topology similar to the transthyretin family in the SCOP database (Murzin, 1995), and the N-terminal domain aligns with one member of this class, the starch binding domain of cyclodextrin glucanotransferase (Penninga et al., 1996), with a r.m.s. deviation of 2.1Å over a span of 63 residues (Figure 4.2). The amphipathic α helix (H4, residues 193-204) comes back from the carboxy terminus of the protein and packs against the outer face of the four stranded β sheet through hydrophobic interactions.

The α domain of VHL (residues 155-192) consists of three α helices (H1, H2, and H3; Figure 4.1, 4.3A). These pack in an arrangement reminiscent of a four-helix cluster ('folded leaf' classification (Murzin, 1995)), except for the absence of a fourth VHL helix. A helix from ElonginC (H4) fits into this gap and completes the four helix cluster arrangement, giving two pairs of helices

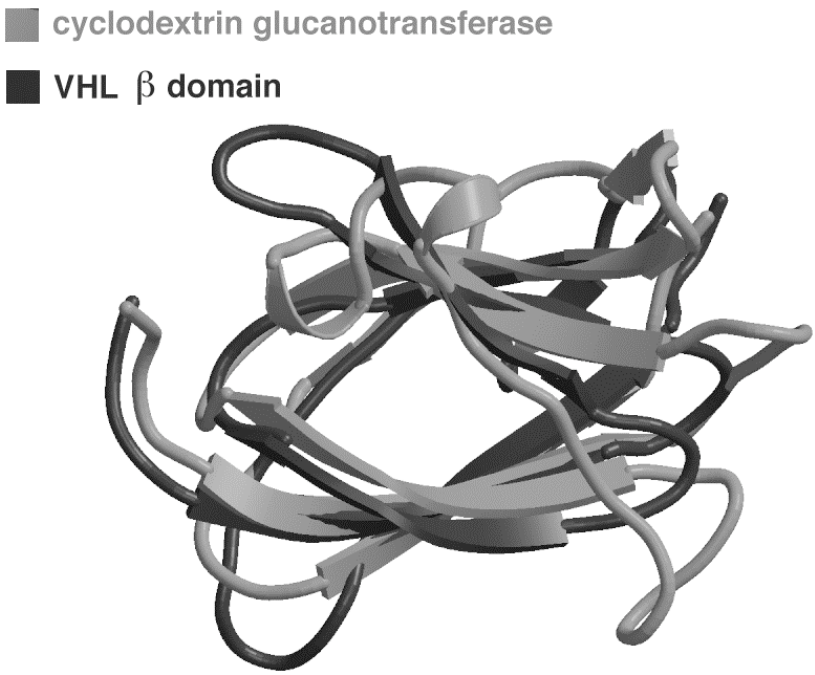


Figure 4.2. VHL β domain and cyclodextrin glucanotransferase

Structural alignment based on Cα positions of the VHL β domain and residues 581-686 of cyclodextrin glucanotransferase (A chain, pdb entry code 1pam).

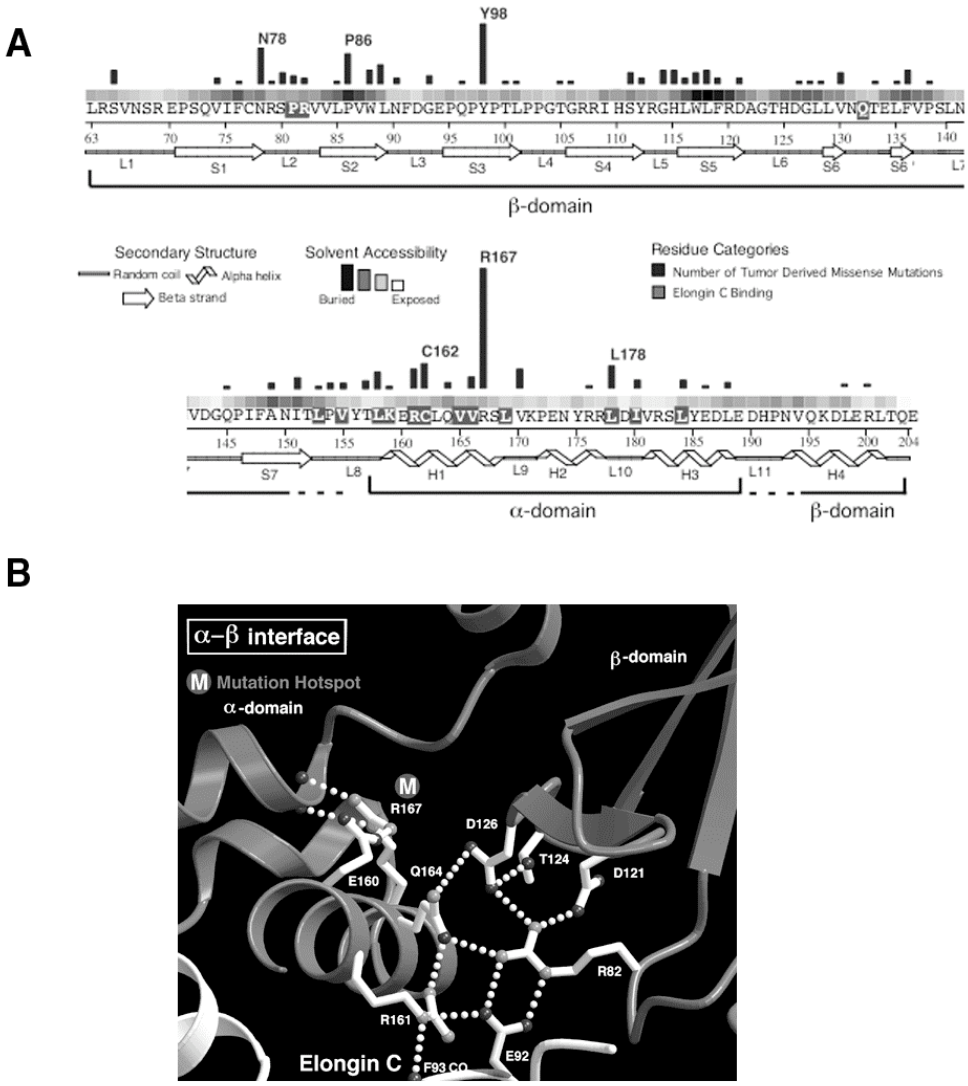


Figure 4.3. Sequence conservation and the VHL α-β interface

(**A**) A histogram of 279 missense mutations in the database (labeled are the six most frequently mutated residues). Shaded squares describe the relative solvent exposure of a residue in a hypothetical VHL monomer. White on black lettering indicates residues contact ElonginC. (**B**) α–β interface: VHL is grey and ElonginC is white. White dashed lines indicate hydrogen bonds, dark atoms indicate oxygen and light nitrogen. A grey circle with the letter "M" indicates one of the six most frequently mutated residues in tumors.

packing at a perpendicular angle (Figure 4.1). The H1 helix of VHL coincides with the 12 amino acid segment (amino acids 158 to 169) shown to play a critical role in ElonginC binding and which is one of the primary targets of tumorigenic mutations.

The α and β domains are connected by two short polypeptide linkers (residues 154-156 and 189-194). The interface between the two domains is well packed, but contains no hydrophobic residues. Instead, it is stabilized by hydrogen-bond networks from the H1 helix, the β sandwich, and ElonginC (Figure 4.3B). These involve eight charged and two polar residues, as well as several polypeptide backbone groups. Many of these residues are frequently mutated, and in fact one of them, Arg^{167}, is the most frequently mutated VHL residue (Figure 4.3,A and B).

The β sandwich forms an extensive and enclosed hydrophobic core containing many residues often found mutated in cancer. The helical α domain, however, in the absence of ElonginC, would not form a self-contained hydrophobic core, but instead, barring conformational alterations, would possess an exposed hydrophobic interface between H1, H2, H3 and the loops L9 and L10 which connect them. The three-helix cluster would thus appear unstable in the absence of ElonginC, as it would have its hydrophobic core partially exposed. It is conceivable that in monomeric VHL there is a conformational change that transfers the amphipathic H4 helix from the β domain to the gap in the α domain and thus sequesters the exposed hydrophobic residues.

4.1.2. The Structure of ElonginC

ElonginC consists of an α/β fold with a three-stranded β sheet packing with one face against four α helices, the last of which are separated by a long, well-ordered loop in an extended conformation (Figure 4.1). A hydrophobic core is included between the β sheet and the helices. The H2, H3, and H4 helices along with the extended L5 loop form a concave surface with a central pocket in which VHL binds (Figure 4.1). Mutagenesis of these elements disrupts VHL binding (Takagi, 1997).

The overall surface of ElonginC is unusual in that the portion that is hydrophobic is significantly higher than that for a small globular protein. The most significant hydrophobic patch is the VHL-binding surface, and in particular the H4 helix which protrudes out of the side of this surface. A smaller hydrophobic patch occurs at the opposite side, at the site of ElonginB binding. A third patch, which is solvent exposed, exists in between the VHL and ElonginB binding sites. This pattern of surface hydrophobicity would be conserved in the drosophila and yeast ElonginC homologues (Figure 4.3A).

Unexpectedly, ElonginC shares significant structural similarity (1.5 Å rmsd for 60 residues) with the tetramerization (T1) domain of the *Shaker* potassium channel (Holm, 1998; Kreusch, 1998), although the regions of ElonginC important for its protein-binding functions, such as the L5 loop, are not similar, however (Figure 4.3B). The proteins evince no obvious sequence similarity. All secondary structural elements, except for the fourth β strand in the T1 domain (corresponding to the disordered residues 50-57 in ElonginC) are conserved with an rmsd of 1.5Å (60 residues).

A

B

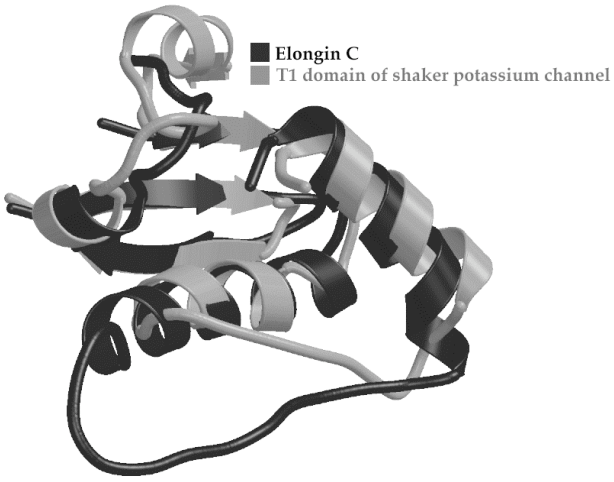


Figure 4.4. Sequence and structural conservation of ElonginC

(A) ElonginC homologues aligned with THREADER2. Residues contacting VHL are white on grey, and those contacting ElonginB are white on black. A disordered segment in ElonginC is indicated with a dashed line. **(B)** Structural alignment of ElonginC with the *Shaker* potassium channel.

4.1.3. The Structure of ElonginB

The ElonginB and ubiquitin structures are nearly identical (1.25 Å rmsd over 71 residues), consistent with the 24% sequence identity (Garrett et al., 1995; Vijay-Kumar et al., 1985). The ElonginB structure consists of two helices and several loops packing against a five stranded β sheet (Figure 4.1, and 4.4A). ElonginB differs from ubiquitin in two respects. It contains an insertion of nine amino acids between the S4 and S5 strands relative to ubiquitin (residues 63-71; Figure 4.4A,B), and residues from this insertion make contacts to ElonginC. It also contains an extension of 30 amino acids at the C-terminus. In the crystal structure, this extension adopts an extended conformation running along one side of the molecule, interacting with ElonginC, and nearly reaching VHL before it becomes disordered. Twenty additional residues, partially conserved (Figure 4.4A), remain in this disordered segment.

The ubiquitin superfold is present in several protein families involved in diverse processes including protein degradation, nuclear transport, and signal transduction (Hochstrasser, 1998). The ElonginB structure is closer to ubiquitin (1.25 Å rmsd) than to the ubiquitin folds of non-conjugated proteins such as RalGDS and RafRBD (1.7 to 2.0 Å rmsd), but it is not as close as Rub1 (0.7 Å rmsd from ubiquitin), which is conjugated like ubiquitin (Nassar et al., 1995; Huang et al., 1998; Rao-Naik et al., 1998).

A

B

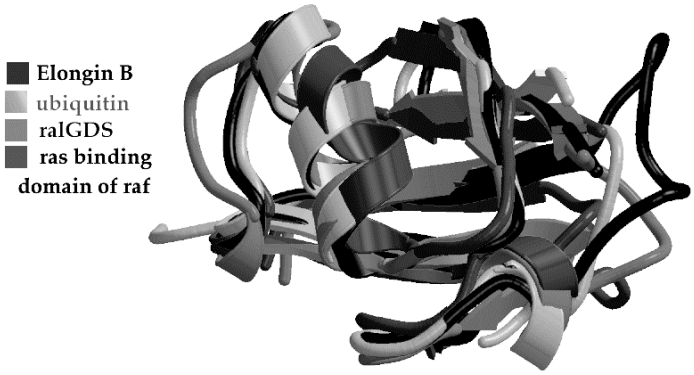


Figure 4.5. Sequence and Structural Conservation of ElonginB

(A) Sequence identity between human ElonginB and ElonginB homologues as generated by a structure based alignment with ubiquitin Residues which make hydrogen bonds or van der Waals contacts with ElonginC are indicated with white on black lettering. **(B)** Alignment of ElonginB and ubiquitin-fold proteins.

4.2. The VHL-ElonginC Interface

The formation of the VHL-ElonginC complex involves the joining of two complementary hydrophobic architectures that bury a total surface area of 2050Å^2. The three helices in the α-domain of VHL, and the concave surface of ElonginC which donates its C-terminal helix to the α-domain, create an inter-molecular four-helix cluster which buries the extensive hydrophobic surfaces exposed on both proteins. Significant but less extensive contacts are made with ElonginC by the β-domain and the first linker between the two domains of VHL (Figure 4.6).

The VHL-ElonginC interface is almost completely hydrophobic with only a handful of significant hydrogen bonds (Figure 4.6). The two interacting surfaces have highly complementary undulations. The H1 helix of VHL fits into the concave surface of ElonginC, inserting hydrophobic side chains into pockets along the surface (Figure 4.6). Reciprocally, the H4 helix of ElonginC, which bulges out from the side of the concave surface, fits into an extended groove formed by the H1, H2, and H3 helices of VHL and completes the 4-helix folded leaf arrangement (Figure 4.6). Two graphs illustrating the per residue surface area buried upon complex formation are shown in Figure 4.7A and B.

The H1 helix in the α domain of VHL has a central role in the VHL-ElonginC interface. This amphipathic helix extends for three turns with hydrophillic residues facing back toward the β-domain of VHL and the non-polar residues facing the hydrophobic face of ElonginC (Figure 4.8, A and B).

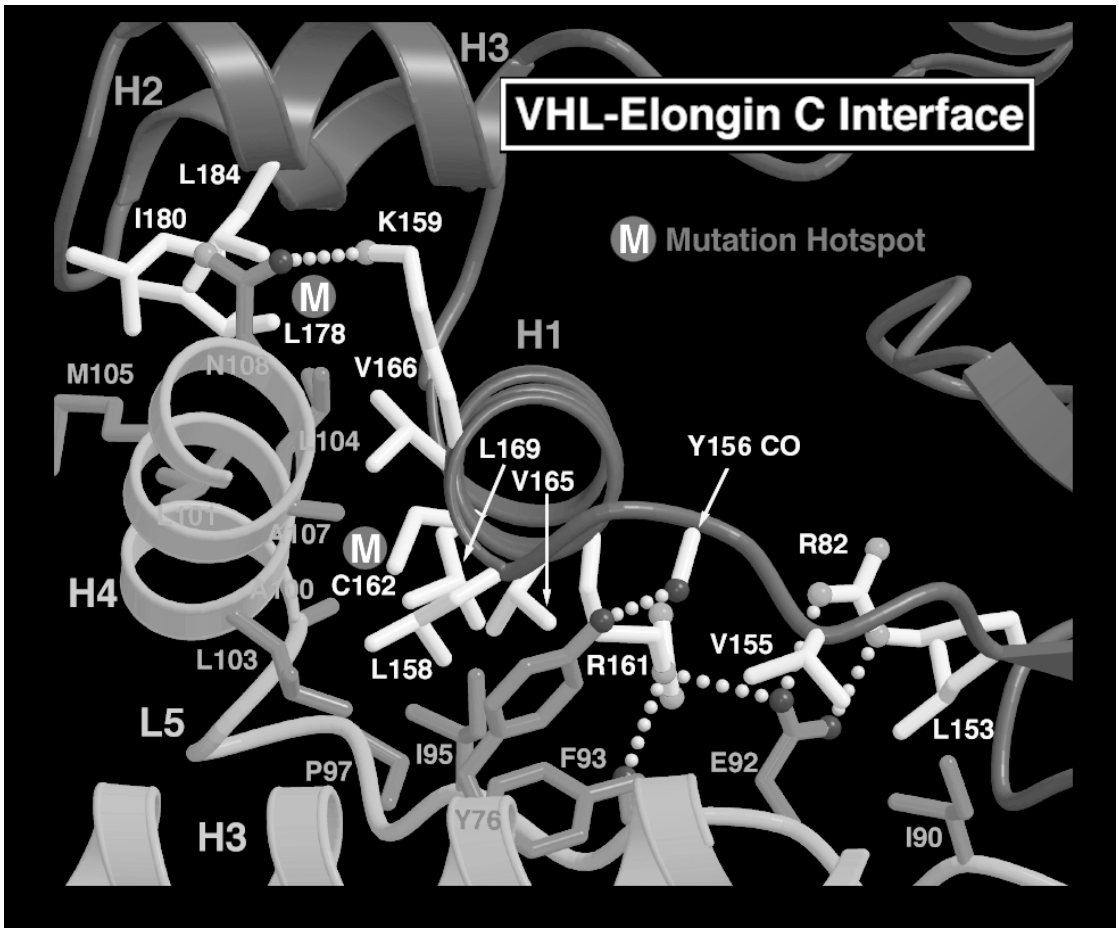
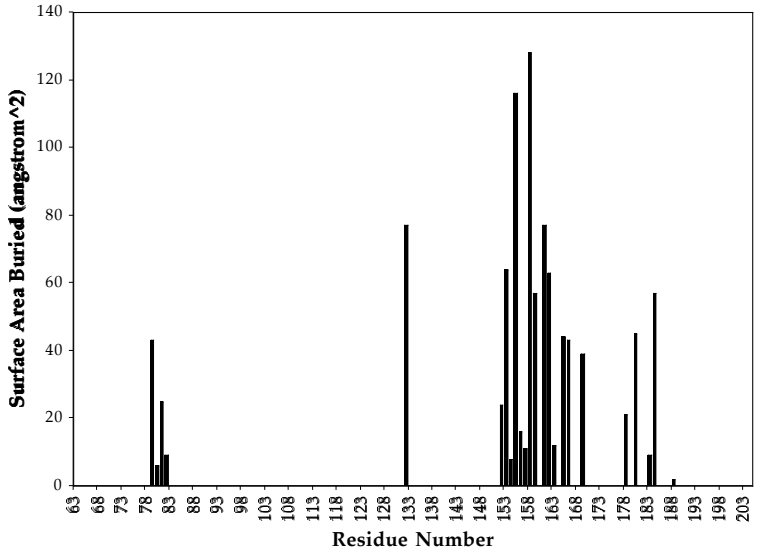


Figure 4.6. The VHL-ElonginC interface

The VHL-ElonginC interface consists of an intermolecular hydrophobic four-helix cluster augmented by additional contacts. VHL and ElonginC secondary structural elements are shown in dark and light grey, respectively. VHL amino acids are in white while those of ElonginC are in grey. Hydrogen bonds are indicated by white dashed lines. A grey circle with the letter "M" indicates a residue which is one of the six most frequently mutated in cancer.

A



B

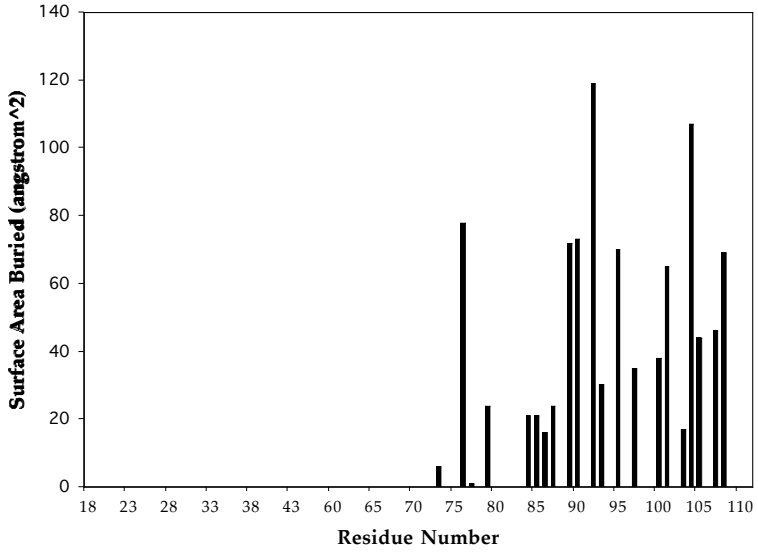
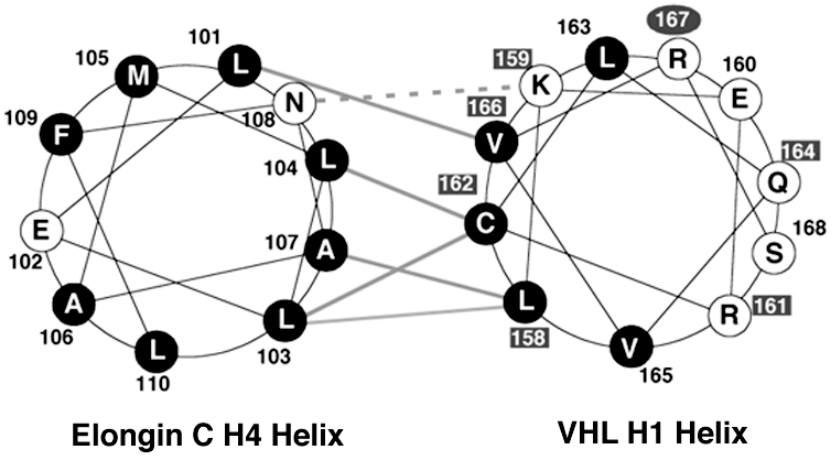


Figure 4.7. Buried surface area upon VHL-ElonginC complex formation

Histogram showing the surface area buried for each residue in going from a hypothetical monomer to the complex. **(A)** VHL **(B)** ElonginC.

A



B

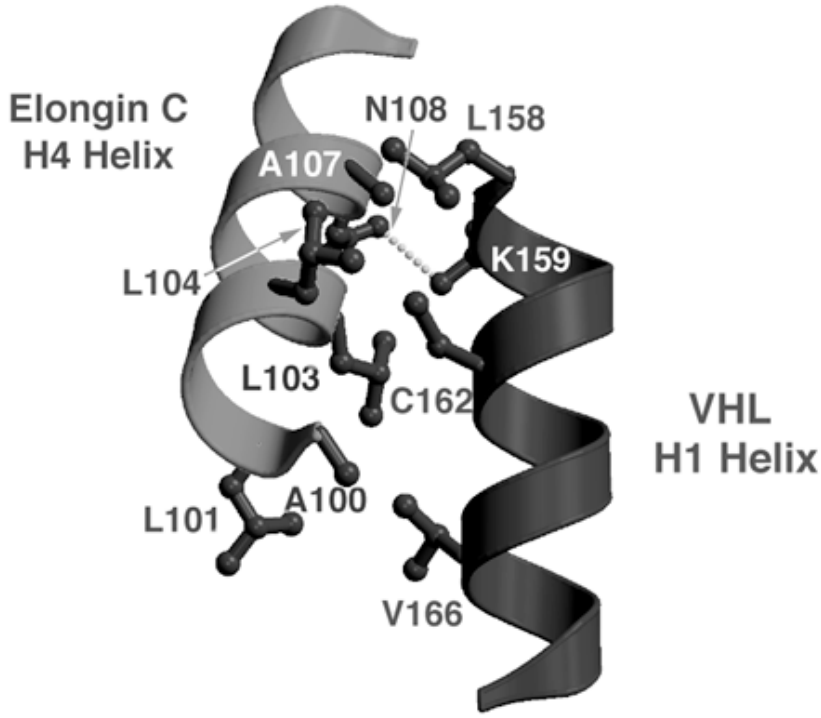


Figure 4.8. Hydrophobic zipper

VHL H1 and ElonginC H3 run antiparallel at an angle of 26.4˚, the first two turns of H1 contacting the first two turns of H3. The interaction between these two helices serves to bury a surface area of close to 600 Å^2. The most extensive contacts here are made by Leu^{158} of VHL, which protrudes out of the α domain and fits into an ElonginC pocket made up of Tyr^{76}, Phe^{93}, Ile^{95} Leu^{103}, Ala^{107}, and Cys^{112} (Figures 4.6-4.8). Other key contacts from the VHL H1 helix are made by Cys^{162}, which packs exclusively along the C-terminal helix (H3) of ElonginC, and Arg^{161}, which forms hydrogen bonds with Glu^{92} of ElonginC (Figure 4.3B and 4.6). These are augmented by lower-density contacts from the Lys^{159}, Val^{165}, Val^{166}, and Leu^{169} side chains of VHL (Figure 4.5).

The central interface role of the H1 helix explains how a peptide replica of this region can compete with full length VHL for binding to ElonginC when at adequate excess (Kibel et al., 1995). The other two helices of the α domain also contribute contacts (Leu^{178}, Ile^{180}, and Leu^{184}; Figure 4.6), with Leu^{184} making the most extensive ones in this region. Additional contacts are made by residues in the first α-β linker (Leu^{153} and Val^{155}), and by Arg^{82} from the β domain (Figure 4.6). The hydrogen bonds made by Arg^{82}, together with those made by Arg^{161} from the H1 helix, represent the few significant hydrogen-bond contacts made at the VHL-ElonginC interface. These arginine side chains are also anchored in the hydrogen bond networks of the VHL α-β interdomain interface (Figure 4.3B). A summary of VHL-ElonginC contacts can be found in Table 4.1.

Table 4.1. Summary of VHL-ElonginC contacts

VHL	SS	ElonginC	SS	Region (VHL ǀ ElonginC)	Type
P81	L2	E92	L5	side ǀ side	VDW
R82 NE	**L2**	**E92 OE1**	**L5**	**side ǀ side**	**hyd(2.9Å)**
R82 NH2	**L2**	**E92 OE2**	**L5**	**side ǀ side**	**hyd(3.2Å)**
Q132	L7	S86	L5	side ǀ main	VDW
L153	L8	I90	L5	main ǀ side	VDW
V155	L8	Y83	H2	side ǀ main	VDW
V155	L8	T84	L5	side ǀ main	VDW
V155	L8	I90	L5	side ǀ side	VDW
Y156	L8	Y76	H2	main ǀ side	VDW
Y156 O	L8	Y76 OH	H2	main ǀ side	**hyd(3.0Å)**
T157	L8	C112	H3	side ǀ main	VDW
L158	H1	C112	H3	main ǀ main	VDW
L158	H1	A107	H3	side ǀ side	VDW
L158	H1	F93	L5	side ǀ side	VDW
L158	H1	L103	H3	side ǀ side	VDW
K159	H1	A107	H3	main ǀ side	VDW
K159	H1	N108	H3	side ǀ side	VDW
K159 NZ	**H1**	**N108 OD1**	**H3**	**side ǀ side**	**hyd(3.0Å)**
R161	H1	I95	L5	side ǀ side	VDW
R161 NE	**H1**	**E92 OE2**	**L5**	**side ǀ side**	**hyd(3.2Å)**
R161	H1	E92	L5	side ǀ side	VDW
R161	H1	F93	L5	side ǀ main	VDW
R161 NH2	**H1**	**E92 OE2**	**L5**	**side ǀ side**	**hyd(3.3Å)**
R161 NH2	**H1**	**F93 O**	**L5**	**side ǀ side**	**hyd(3.3Å)**
C162 Cβ	H1	L104	H3	side ǀ s/m	VDW
C162 S	H1	L104	H3	side ǀ s/m	VDW
C162	H1	I95	L5	side ǀ side	VDW
C162	H1	L103	H3	side ǀ s/m	VDW
V165	H1	I95	L5	side ǀ side	VDW
V166	H1	L101	H3	side ǀ side	VDW
L169	H1	P97	L5	side ǀ side	VDW
L184	H3	M105	H3	side ǀ side	VDW
L184	H3	L104	H3	side ǀ side	VDW

Van der Waals interactions (VDW) are defined as atoms approaching between 3.5Å and 4.0 Å. Hydrogen bonds (hyd) are indicated when atoms are between 2.2Å and 3.4Å with appropriate donor and receptors. "Side" and "main" indicate contacts to side chain and main chain atoms ("s/m" to both).

Nearly one-half of the ElonginC side chains that contact VHL, including Glu^{92}, are identical in yeast ElonginC (Figure 4.3A). Although an ortholog of VHL is not yet apparent in yeast, the presence of a functional homologue would be likely.

4.3. The ElonginC-ElonginB interface

The ElonginC-ElonginB interface is similar in extent to the ElonginC-VHL interface, but differs in character by being less hydrophobic and by lacking the complementary pockets and protrusions characterizing the ElonginC-VHL interface. The interaction between the two molecules is dominated by the edge-to-edge interaction of the ElonginC and ElonginB β sheets (Figure 4.1, 4.9). The S2 β strands from each protein (12-19 in ElonginB and 28-32 in ElonginC) align in parallel and make four canonical parallel β sheet hydrogen bonds, involving the backbone amides of ElonginB residues Thr^{13} and Phe^{15} and ElonginC Glu^{28} and Ile^{30}, which bond with the carbonyl oxygens of ElonginC Glu^{28} and Ile^{30} and ElonginB Lys^{11} and Thr^{13}, respectively.

The sheet junction is reinforced by van der Waals contacts from hydrophobic residues residing on the S2 strands of both Elongins, and by interactions from additional structural elements on both faces of the sheet (Figure 4.9 and 4.10). On one side are the interactions between Thr^{13} and Phe^{15} (from S2), Pro^{69} from L6, and Phe^{93} (in L7) from ElonginB with the amino acids Phe^{29} and Val^{31} (S2), His^{27} (L1), and Thr^{78} (from H3) of ElonginC. On the other face of the sheet, there are hydrophobic interactions between

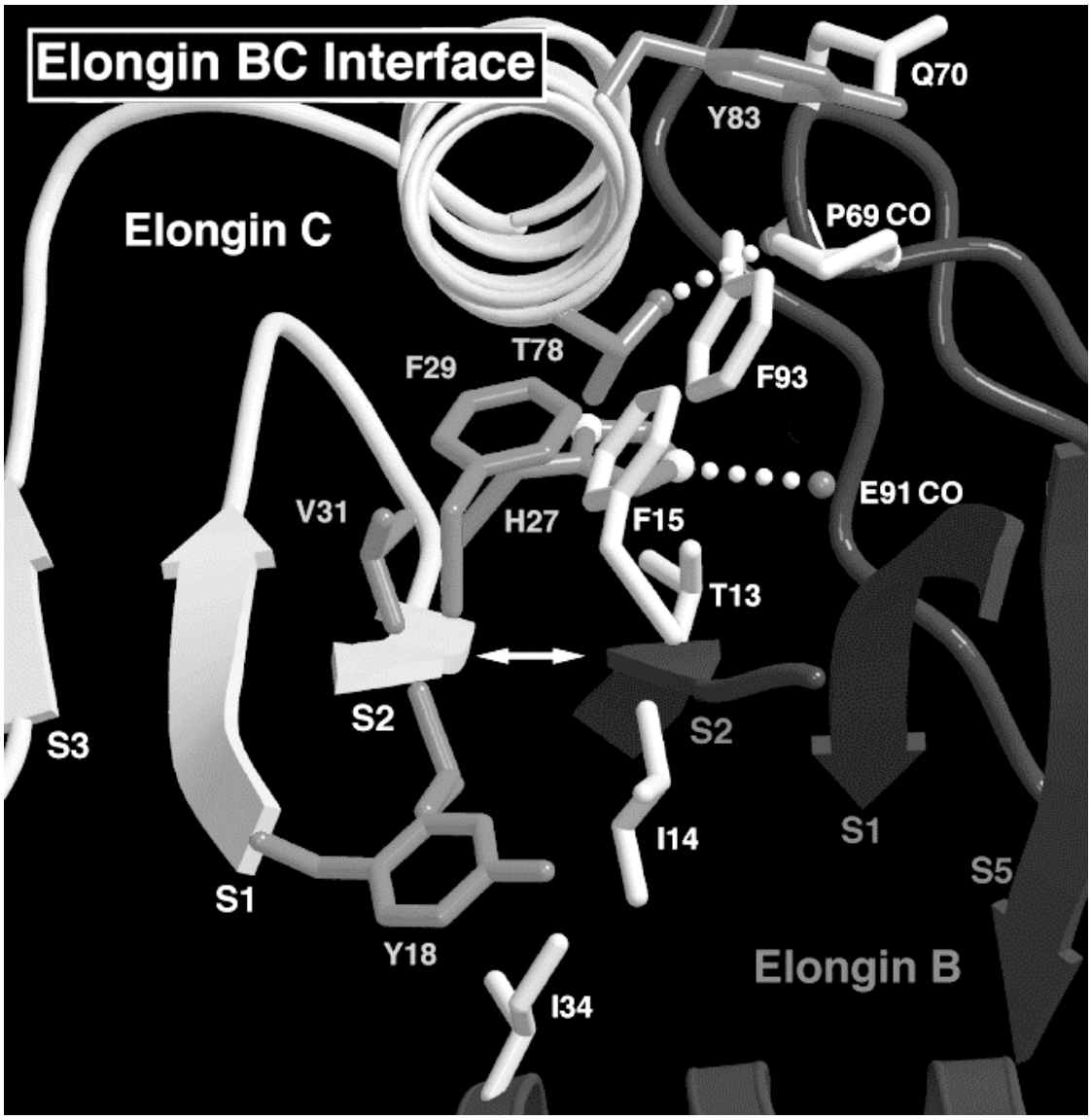
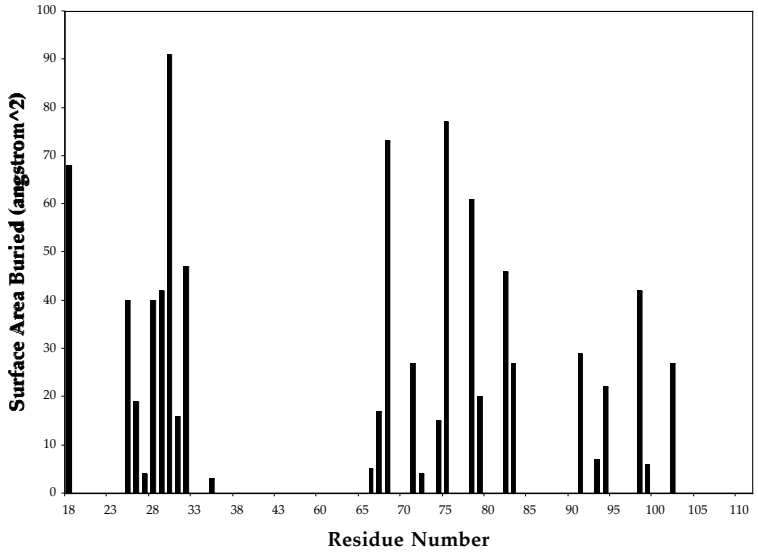


Figure 4.9. The ElonginC- ElonginB interface

The parallel β strand - β strand interaction that dominates the ElonginC-ElonginB interface is indicated by a two-headed arrow (four hydrogen bonds). ElonginB and ElonginC secondary structural elements are shown in grey and white, respectively. ElonginB amino acids are in white while those of ElonginC are in grey.

A



B

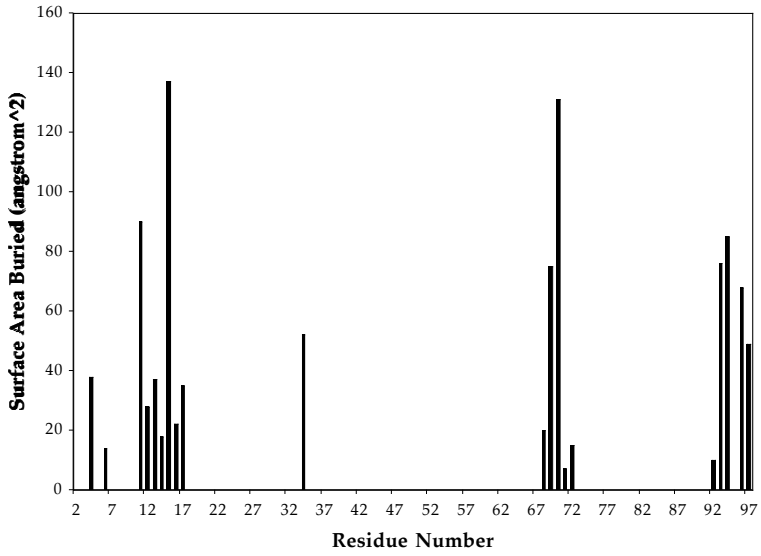


Figure 4.10. Buried surface area upon ElonginBC complex formation

Histogram showing the surface area buried for each residue in going from a hypothetical monomer to the complex. **(A)** ElonginC **(B)** ElonginB.

Ile^{14} and Ile^{34} of ElonginB and Tyr^{18} and Ile^{30} of ElonginC. A summary of ElonginC-ElonginB contacts can be found in Table 4.2.

The ElonginC residues whose mutation abolishes ElonginB binding (Takagi et al., 1996; Lonergan et al., 1998) are included in these interacting regions. The carboxy-terminal tail of ElonginB (Figure 4.1), not present in the structures of ubiquitin or ubiquitin-like molecules, packs through loose van der Waals contacts in an extended conformation alongside ElonginC (Figure 4.1). In general, the ElonginC residues that make side chain contacts to ElonginB are less conserved than those that contact VHL (Figure 4.4). The ElonginC N-terminal residues that make backbone β-sheet hydrogen bonds to ElonginB are an exception, as they are conserved because of their role in the hydrophobic core.

An intermolecular strand-strand interaction forming a continuous β sheet is seen in the binding of the ubiquitin-like domains of ralGDS and rafRBD to Ras (Nassar et al., 1995; Huang et al., 1998). The ElonginB-ElonginC interface differs from these regulatory interfaces in that i) it is nearly twice the size (2000 $Å^2$ compared to about 1200 $Å^2$ in the binding of rafRBD and ralGDS to the switch one region of the proto-oncogene ras), ii) it has a significant hydrophobic component compared to the other two complexes that rely primarily on hydrogen bond contacts, and iii) it has a parallel strand pairing in contrast to the antiparallel strand-strand interactions displayed in these other intermolecular sheet forming complexes.

Table 4.2. Summary of ElonginB-ElonginC contacts

ElonginB	SS	ElonginC	SS	Region	Type
				ElonginB | ElonginC	
R8 NH2	S1	H27 NE2	L1	side | side	VDW
K11 O	**L1**	**E28 N**	**S2**	**main | main**	**hyd(3.0Å)**
T12	S2	E28	S2	main | main	VDW
T12	S2	I30	S2	side | side	VDW
T13 N	**S2**	**E28 O**	**S2**	**main | main**	**hyd(2.9Å)**
T13	S2	E28	S2	main | main	VDW
T13	S2	F29	S2	side | side	VDW
T13	S2	E28	S2	side | main	VDW
T13	S2	F29	S2	main | main	VDW
T13 O	**S2**	**I30 N**	**S2**	**main | main**	**hyd(3.0Å)**
I14	S2	I30	S2	main | main	VDW
I14	S2	I30	S2	side | side	VDW
F15 N	**S2**	**I30 O**	**S2**	**main | main**	**hyd(3.0Å)**
F15	S2	I30	S2	s/m | main	VDW
F15	S2	V31	S2	side | side	VDW
F15	S2	F29	S2	side | side	VDW
F15	S2	C74	H2	side | side	VDW
F15	S2	Y18	S1	main | side	VDW
T16	S2	Y18	S1	side | side	VDW
D17	S2	L32	S2	side | side	VDW
H27 NE2	**L5**	**E91 O**	**H2**	**main | side**	**hyd(2.8Å)**
I34	H1	I30	S2	side | side	VDW
I34	H1	Y18	S1	side | side	VDW
P69	L5	M75	H2	main | main	VDW
P69 CO	**L5**	**T78 Oγ1**	**H2**	**main | side**	**hyd(3.4Å)**
P69	L5	R82	H2	Side | side	VDW
Q70	L5	Y83	H2	side | side	VDW
F93	L6	H27	L1	side | side	VDW
F93	L6	F29	S2	side | side	VDW
F93	L6	H68	H2	side | main	VDW

Van der Waals interactions (VDW) are defined as atoms approaching between 3.5Å and 4.0 Å. Hydrogen bonds (hyd) are indicated when atoms are between 2.2Å and 3.4Å with appropriate donor and receptors. "Side" and "main" indicate contacts to side chain and main chain atoms ("s/m" to both). The intermolecular strand-strand hydrogen bonds are boxed.

4.4. Comparison with mutagenesis data

Takagi *et al.* (1997) report the generation and evaluation of clustered alanine mutations in the C-terminus of ElonginC . Mutants spanning 92-94, 104-105, and 110-112 were defective in binding VHL, consistent with the interaction of residues Glu^{92}, Phe^{93}, Leu^{104}, and Cys^{112} with VHL as observed in the crystal structure. In addition, mutations in residues 89-91 and 98-99, possessing no demonstrable contacts with VHL, had no effect on complex formation. Clustered mutations spanning 95-97 and 101-103 in ElonginC, while not abrogating VHL binding in their assays, did severely impair activation of ElonginA, which shares a region of homology with the H1 helix of VHL (H1 alignment Figure), and thus likely interacts with ElonginC in a manner very similar to that of VHL. One may understand these results as indicating that these mutations cause a destabilization in the complex which is only partially detectable with their assay. Due to the limited contacts Leu^{101} makes with Val^{166} of VHL, it is not unreasonable to suppose that the disruption of these interactions would have milder effects of complex formation than the alteration of residues more central to the interface. In ElonginA, however, there is a leucine in the place of Val^{166}, and the larger hydrophobic residue could make more extensive contacts with Leu^{101} and perhaps other residues in H3 of ElonginC, leading to a greater loss in binding energy upon disruption of the interaction. It is more difficult to understand the substitution of Leu^{103} with alanine, as Leu^{158} from VHL and Leu^{103} make extensive contacts between their terminal methyl groups. However, both the Cβ atom and the backbone carbon atom of Leu^{103} in ElonginC also make close van der Waals contacts with the sulfur of Cys^{162} in VHL. Thus, while

removing the leucine-leucine interaction, the alanine mutant still preserves two of the four atomic contacts between VHL and ElonginC created by Leu^{103}. It is not as clear why mutations in Pro^{97} and Ile^{95} destabilize ElonginA complex formation more than VHL.

In addition, Takagi *et al*. (1996) also define regions of ElonginC important for its interaction with ElonginB. Deletion of 28 amino acids from the C-terminus, corresponding to removal of the L5 loop and the H3 helix, did not prevent the isolation of a complex between ElonginB and ElonginC. Consistent with this data are the facts that neither L5 nor H3 form significant components of the ElonginC hydrophobic core or the interaction region with ElonginB. Deletion of the first 17 amino acids did not interfere with binding, consistent with the proteolytic and size exclusion data on the ElonginC (17-112) fragment. However, deletion of the first 22 residues, or the internal deletion of amino acids 21-30, prevented complex formation. These results are entirely consistent with the crystallographic data, from which it is observed that residues 21-31 form the two β strands of ElonginC closest to ElonginB, with the amino acids 28-32 (S2) making the direct backbone contacts with S2 of ElonginB. In addition to the deletion analysis, four ElonginC clustered alanine scanning mutants were reported: 19-21 and 22-24, which did not prevent the isolation of ElonginBC complexes, and two sets which abrogated binding, 25-27 and 28-30. No contacts are observed between amino acids 19-24 in ElonginC with residues in ElonginB (Table 4.1), whereas the side chains from residues Asp^{25}, His^{27}, Phe^{29}, and Ile^{30} all make contact to atoms in ElonginB.

CHAPTER FIVE:

ANALYSIS OF THE VHL-ElonginC-ElonginB STRUCTURE

5.1. Analysis of tumor derived missense mutations in VHL

About half of the mutations in the VHL database (total of 579 entries (Beroud, 1998)) are missense mutations, and they map comparably to the α and β domains (Figure 4.3A, and 5.1). As with the p53 tumor suppressor (Cho et al., 1994), a subset of amino acids in VHL are disproportionately represented in the database of tumor derived mutations. Of the more than 60 residues found altered in cancer, the six most frequently mutated amino acids - Asn^{78}, Pro^{86}, and Tyr^{98} from the β domain, along with Cys^{162}, Arg^{167}, and Leu^{178} from the α domain - account for over 40 % of all mutations found in tumors. Arg^{167} (17 % of missense mutations) and Tyr^{98} (9 % of mutations) are obvious mutational "hotspots," contributing to one fourth of all mutations. Residues Asn^{78}, Pro^{86}, Arg^{167}, and Leu^{178} belong to an architectural class of mutants which are involved in stabilizing the three dimensional structure of VHL, whereas Tyr^{98}, and Cys^{162} are non-structural and are located on the surface of the molecule.

In the α domain, the H1 helix is the primary target of tumorigenic mutations (Figure 4.3A and 5.1). It contains the most frequently mutated residue, Arg^{167}, which has a structural role stabilizing the H1 helix and the α-β

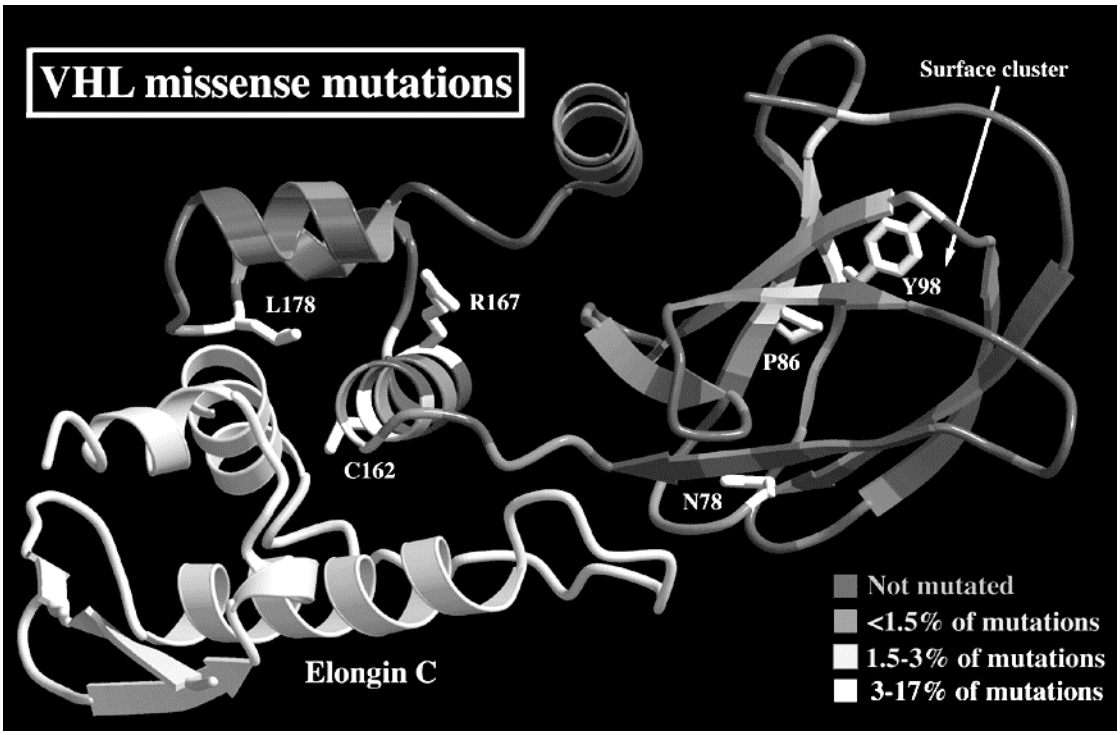


Figure 5.1. Frequency and location of tumorigenic missense mutations

The VHL ribbon diagram is colored by mutational frequency using a gradient in half-tones. The white residues correspond to the six most frequently mutated residues and are labeled. ElonginC is shown in a light grey-white.

interdomain interface (Fig 4.3B). The function Arg^{167} is purely structural, serving as the center of a chemical bonding network of linked hydrogen bonds and van der Waals contacts connecting the two domains of VHL and all major structural components of the α-domain. The guanidinium group of Arg^{167} forms a salt bridge with the carboxylate of Glu^{160} (NH1), and forms hydrogen bonds the backbone oxygens of Asp^{187} (NH2) and Leu^{188} (NE) of H3. Glu^{160} makes van der Waals contacts with Gln^{164} in H1, which itself forms hydrogen bonds with Arg^{161} and Arg^{82} (from the β domain) while stacking with Asp^{126}. Asp^{126}, also located in the β sandwich, makes van der Waals contacts with Pro^{192} (in L11) from the α domain and hydrogen bonds with Arg^{82} in the L2 loop of the β domain. The NE nitrogen of Arg^{167} forms additional van der Waals contacts with Leu^{163} (in H1) and Tyr^{175} (in H2).

In tumors, Arg^{167} is mutated either to tryptophan or glutamine. The tryptophan mutation would likely induce the greatest alteration in VHL. Given the tight bonding network described above, it is hard to imagine that this bulky hydrophobic amino acid could insert into the space occupied by Arg^{167} in wild type VHL without inducing significant structural rearrangements, potentially disrupting both the interface between the two domains of VHL as well as destabilizing the packing of the three helices in the α domain. Although the arginine to glutamine mutation seems much more conservative, this substitution deprives the surrounding amino acids of several important interactions. The shorter and uncharged glutamine would be unable to form the hydrogen bonds and van der Waals interactions made by the head group of arginine. This would remove important links between the H1 and H3 helices, and L11 loop, as well as disrupting the chain of interactions which links Arg^{167} to the β domain in wild type VHL. Thus, while not producing the steric clash of the Arg^{167}Trp mutation,

the $Arg^{167}Gln$ mutant would cause a destabilization of the α domain and the link between the two domains of VHL.

The hypothesis that $Arg^{167}Trp$ and $Arg^{167}Gln$ would destabilize the α domain is supported by *in vitro* binding studies of these mutants in the presence and absence of detergent. In the absence of detergent, both tumorigenic arginine mutants still bind ElonginC, albeit with reduced ability. The addition of detergent to these mutants, which does not effect the binding of wild-type VHL to ElonginC, completely abolishes the formation of isolable ternary complexes (Duan et al., 1995; Lonergan et al., 1998). Since Arg^{167} is not directly involved in binding ElonginC, its mutation would not be expected to alter the protein-protein interactions, provided that the fold of the α domain is maintained. The addition of detergent, a destabilizing act in and of itself, would combine with the structural destabilization incurred by mutation to partially or completely unfold the entire α domain.

The other two frequently mutated residues in the α domain, Cys^{162} and Leu^{178}, are on the H1 and H2 helices, respectively. Leu^{178} has dual roles, interacting with ElonginC and also stabilizing the H2-H3 packing of the α domain (Figure 4.6). The primary role of Cys^{162} is to bind to ElonginC (Figure 4.6). In cancer Cys^{162} is typically mutated to phenylalanine, tryptophan, tyrosine, or arginine. Given the tight packing between Cys^{162} and the three hydrophobic residues Ile^{95}, Leu^{103}, and Leu^{104} of ElonginC, the mutation to any of the four large amino acids listed would be likely to severely disrupt the interaction of the VHL H1 helix with the H3 helix from ElonginC. As this interface is central to the formation of the binary complex, these mutations in Cys^{162} are likely to substantially impair the binding of these two proteins. Support for this interpretation is found in the experiments (Kibel et al., 1995) in which a peptide corresponding to the H1 helix loses the ability to compete with

full length protein upon the alteration of Cys^{162} to phenylalanine. Mutations in Leu^{158} (to proline or valine) and Arg^{161} (to proline or glutamine) result in the loss of contacts due to the substitution of shorter amino acids.

The remaining α domain mutations target residues involved in the packing of the helices (such as Val^{170}, Ile^{180}, Leu^{184}, Leu^{188}), or in stabilizing the α-β interdomain interface (Arg^{161}, Gln^{164}; Fig 4.3B). Several of these residues also make important contacts to ElonginC (Arg^{161}, Ile^{180} and Leu^{184}; Figure 4.5). This pattern of mutations in the α domain solidifies the role ElonginC binding has in the tumor suppressor activity of VHL.

In the β domain, most, but not all of the mutations map to residues important for the structural integrity of the β sandwich. Some map to the hydrophobic core of the protein, where nineteen of the twenty-two residues are mutated (such as Phe^{76}, Pro^{86}, Phe^{119}, Trp^{117}, and Val^{130}); others map to buried polar residues that hold loops together (Asn^{78}; Figure 4.3A, and 5.1). These observations indicate that the structural integrity of the β-domain is critical to the tumor suppressing activity of VHL.

Tyr^{98}, which is the second most frequently mutated residue in VHL (Figure 4.3A), is a significant exception to this trend of structural β domain mutations. Tyr^{98} occurs on the surface of the β sandwich, opposite from where ElonginC binds, and has no apparent structural role (Figure 5.1, 5.2, and 5.3). That Tyr^{98} neither functions to maintain the structural integrity of VHL nor to bind ElonginC is supported by the ability of a Tyr^{98} tumor-derived mutant to bind ElonginC *in vivo* (Duan et al., 1995, Kishida et al., 1995, Ohh et al., 1998)). Several additional solvent-exposed residues surrounding Tyr^{98} are also mutated (Figure 5.2). These include Asn^{90}, Gln^{96}, and Tyr^{112}, none of which appear to have any significant structural role (Figure 5.2). In addition, the partially buried Trp^{88}, and important component of the hydrophobic core of VHL, extends to the

surface and contributes through the aromatic portion of its indol ring to the mutational surface patch (Figure 5.2). The amino acids comprising the patch account for roughly one fifth of mutations found in cancer. These observations suggest that this mutational patch of solvent exposed residues may correspond to another protein binding site of VHL, and that the use of this site is important for the tumor suppressor activity of VHL.

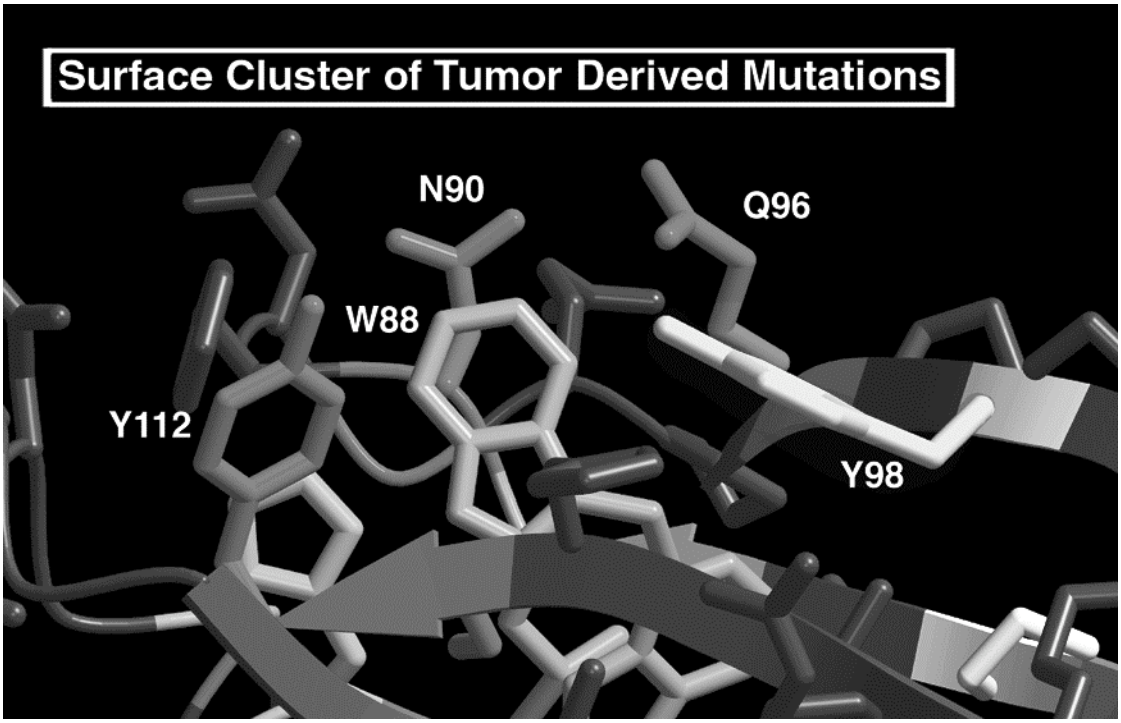


Figure 5.2. Mutational surface patch near Tyr^{98}

Close-up view of Tyr^{98} and the aromatic and polar residues in its vicinity, colored as in 5.1. To make these residues easier to see, view is rotated about the horizontal axis relative to 5.1.

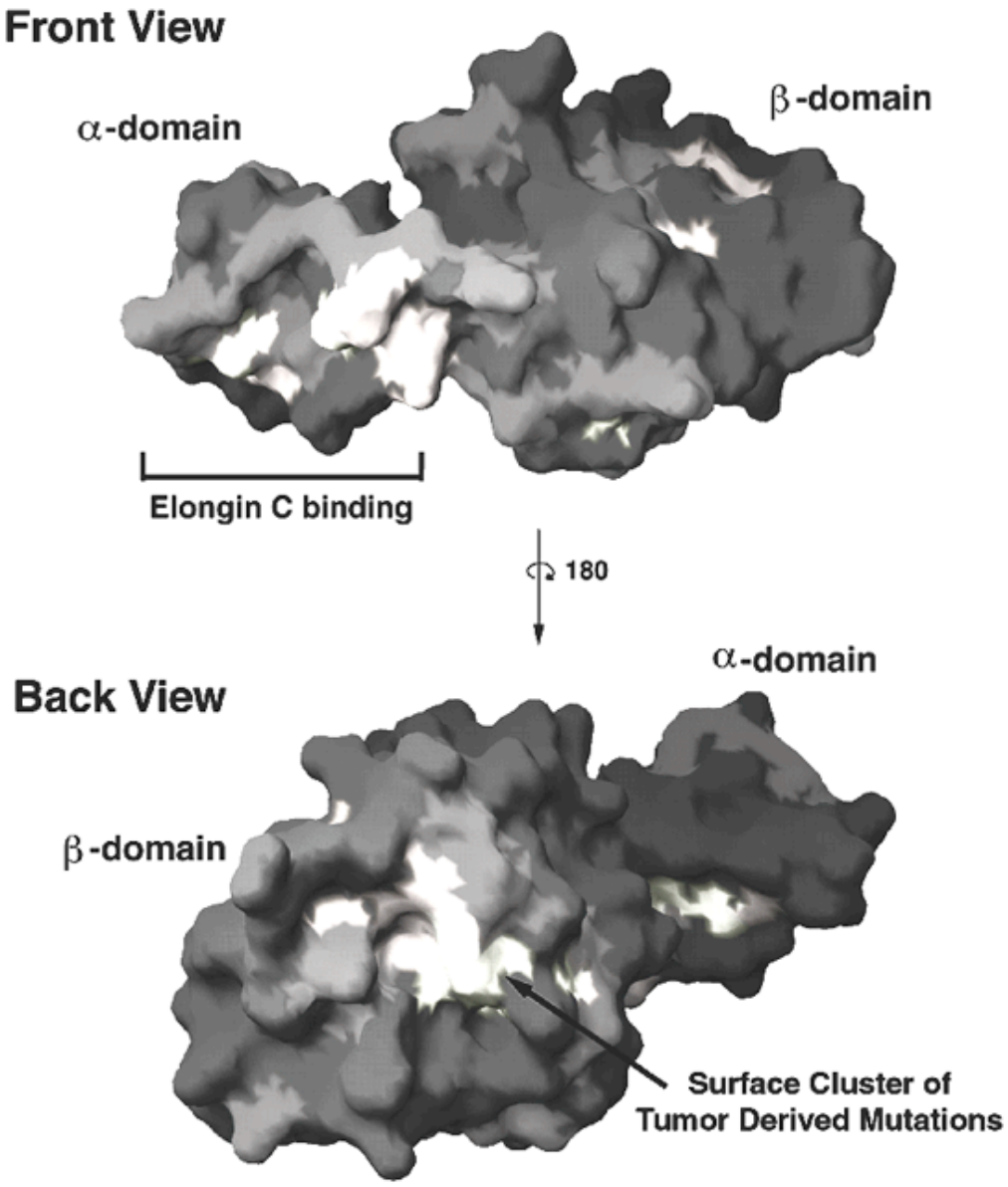


Figure 5.3. Molecular surface representation

Molecular surface representation of VHL in two orientations. Coloring of surface residues is according to mutational frequency as in 5.1. The 'front view' is as in 5.1; the 'back view' is rotated 180° about the vertical axis.

5.2. Structure basis for type II VHL syndrome

VHL families have clinically heterogeneous symptoms, and are divided into two types differentiated by the presence (type II) or absence (type I) of pheochromocytoma (Chen et al., 1995). In type II syndrome, the alteration is almost always (99% of cases) a missense mutation, whereas in type I the mutation can be either a missense mutation or deletion/insertion/frameshift mutation (Chen et al., 1995).

An analysis of missense mutations associated with type II VHL syndrome (Chen et al., 1995) reveals a striking bias against mutations that would be predicted to cause total loss of function. In contrast to type I missense mutations that usually target the β domain hydrophobic core and which would be predicted to cause a denaturation of the folded state, none of the type II mutations map to the hydrophobic core of the β sandwich. Most map either to residues that contact ElonginC (Leu^{158}, Arg^{161}), or to exposed residues that make up the putative second binding site (Tyr^{98}, Tyr^{112}). Even when they map to structural residues they would likely cause mostly local defects. For example, Arg^{167} mutants, which are very frequent in the type II syndrome, retain a residual ability to bind ElonginC under mild buffer conditions, such as the absence of detergent (Duan et al., 1995; Ohh et al., 1998). This class of missense mutations is consistent with the total lack of deletion/insertion/frameshift mutations in type II cases.

One mechanism that could explain these observations is if the occurrence of pheochromocytoma requires gain of function mutations, where the mutant VHL interferes with the wild type allele in a dominant fashion. Alternatively, the total loss of VHL function may be lethal in some tissues. The gain of function

model is, in principle, plausible as VHL has two protein binding sites, and a mutant having only one of the two defective may sequester key components of the pathway. Experimentally, it is supported by the observation of a dominant negative effect with an overexpressed mutant (frameshift at residue 187). This mutant is predicted to have lost its α domain structure but have retained an intact β domain and second protein binding site (Gnarra et al., 1996).

5.3. VHL α domain homology to SOCS-box

To investigate whether the 40-amino acid SOCS-box sequence motif (Starr et al., 1997; Hilton et al., 1998),which possesses sequence homology to the H1 helix of VHL (Figure 5.4), and which was recently shown to bind ElonginC (Kamura et al., 1998), is structurally related to the VHL α domain, a sequence threading analysis was performed with the program THREADER2 (Miller et al., 1996). This analysis gives a measure of how likely a particular amino acid sequence is to adopt a given structure. When the sequence of the SOCS-1 SOCS box was threaded against a library of nearly 2000 structures, the VHL α domain ranked first. Inspection of the THREADER2 alignment (Figure 5.5A) shows that both the hydrophobic residues that are structural determinants of the α domain and the exposed hydrophobic residues that bind ElonginC are conserved in character (Figure 5.5, A and B). These findings suggest that the SOCS box and VHL α domain represent a common structural motif, and in conjunction with the reported SOCS box-ElonginC binding data indicate that the SOCS box and VHL α domain are functional homologues.

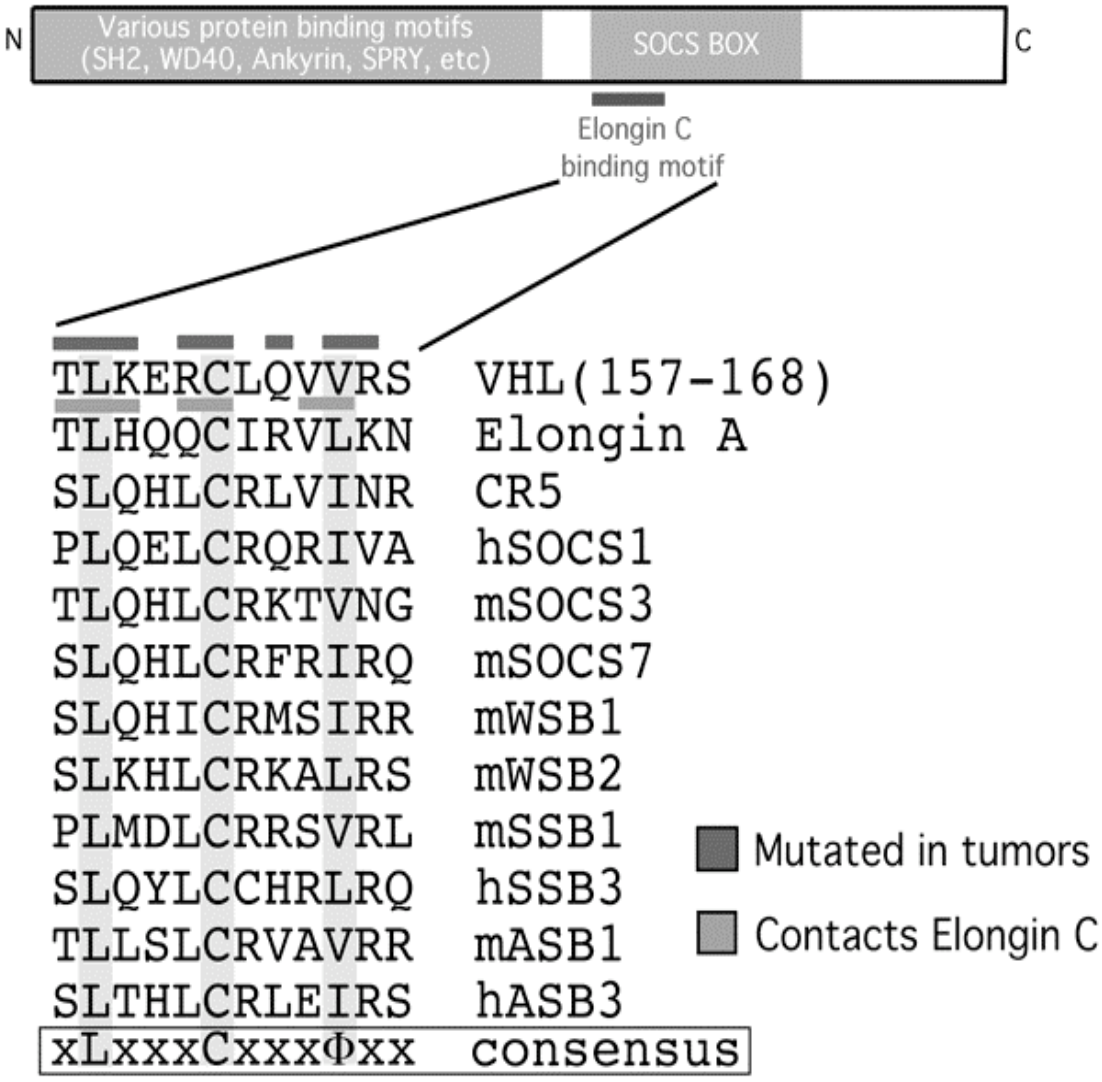
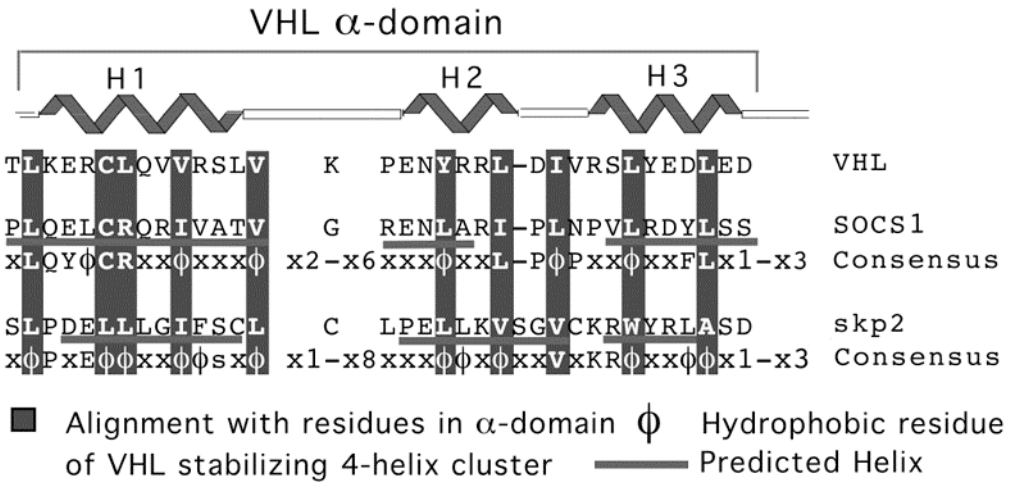


Figure 5.4. Proteins with homology to the H1 helix of VHL

A



B

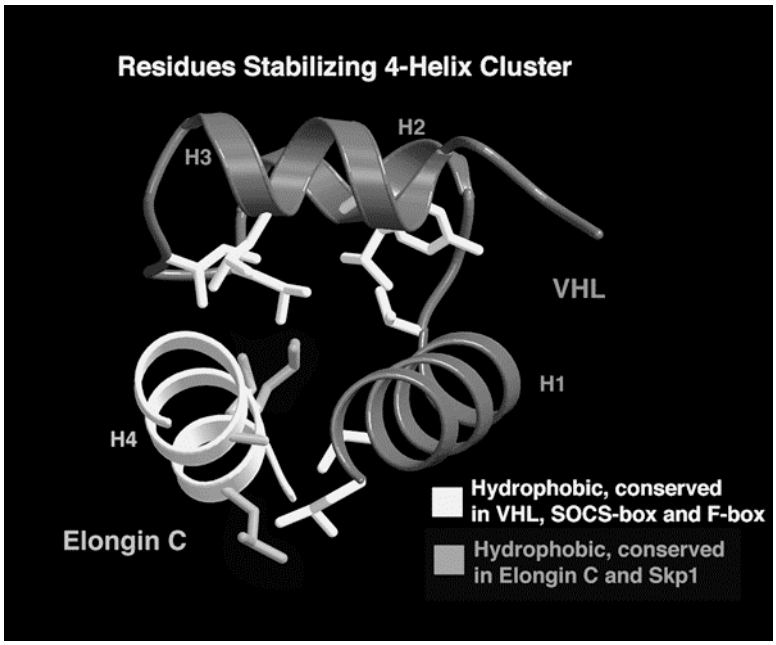


Figure 5.5. Homology modeling of VHL α domain

(A) Alignment of the VHL α domain, the SOCS1 and the SOCS-box consensus, and the Skp2 and F-box consensus sequences. (B) Residues that contribute significantly to the stability of the four-helix cluster.

ElonginA also has sequence homology with the 12-res VHL H1 helix (Figure 5.4). In the threading analysis, the VHL α domain structure ranked 704th, however, suggesting that the ElonginC binding activity of ElonginA may be divergent.

5.4. Analogy between VCB and SCF complexes

Threading analysis was used to further investigate the homology between ElonginC and the amino-terminal two-thirds of Skp1 (31 % sequence identity). The threading analysis of the Skp1 sequence ranked the ElonginC structure first, and this was the only structure to score in a range indicative of a high probability for the same structure (Z score of 3.7). In the threader alignment, many of the hydrophobic core residues of ElonginC are either conserved in Skp1, or are conservatively substituted (Figure 4.3A). These observations strongly indicate that Skp1 will contain a structure very similar to that of ElonginC.

It should then be asked whether the similarity in structure could also support a similarity in protein-binding function. The ElonginC-VHL interface relies on several characteristic structural and sequence features of ElonginC: the concave surface with its protruding H4 helix and its H3-L5 loop forming pockets into which VHL side chains insert, and the highly hydrophobic nature of this surface. The threading-based alignment (Figure 4.3A) indicates that these ElonginC structural features and their characteristic hydrophobicity will be conserved in Skp1 (Figure 5.6). The ElonginC surface residues that hydrogen bond to VHL are also conserved in Skp1 (Glu^{92} and Asn^{108}; Figure 4.3A and 4.5).

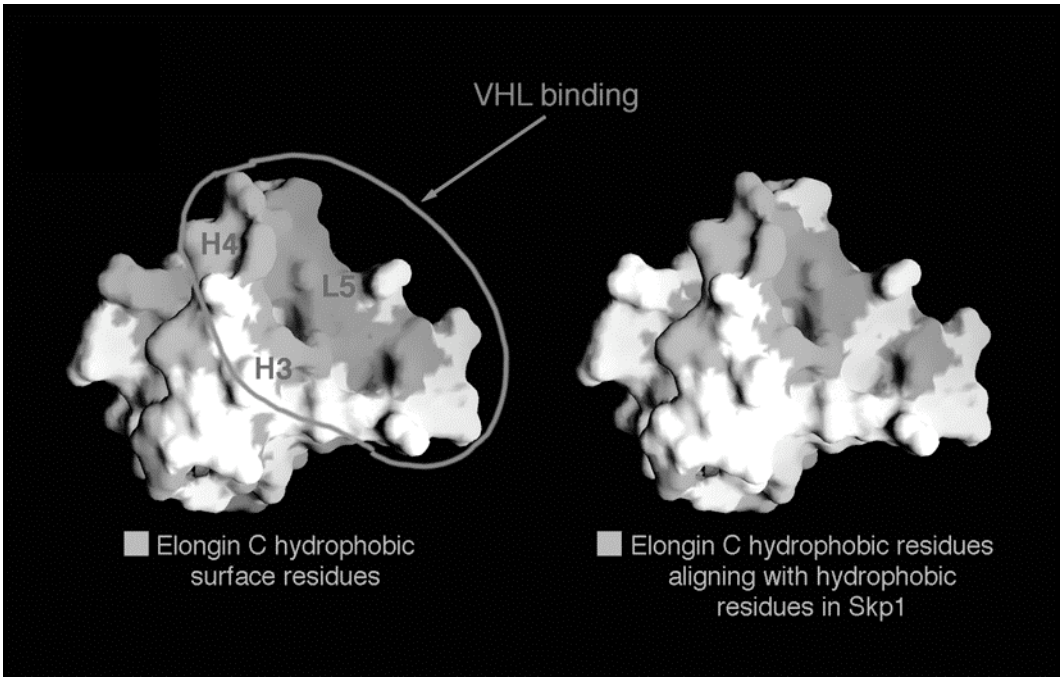


Figure 5.6. Surface property conservation between ElonginC and Skp1

ElonginC and Skp1 likely share a similar tertiary structure and a similar protein binding site. The leftmost molecular surface illustrates hydrophobic surface residues in ElonginC (grey) along with the approximate locations of the secondary structural elements (H3, L5, and H4) underlying the VHL-binding surface. The rightmost figure displays, again in grey, only those hydrophobic surface residues of ElonginC which maintain their hydrophobic character in Skp1 according to the alignment shown in Fig. 4.4A (A, C, F, I, L, M, P, V, W, and Y are considered hydrophobic in this analysis).

These observations suggest that Skp1 may use this portion of its structure to bind other proteins with α domain analogous to the α domain of VHL.

Because Skp1 binds proteins that contain the 45-amino acid F-box sequence motif, we asked whether there is any structural similarity between the F-box and the VHL α domain. In the threading analysis of the Skp2 F-box sequence, the VHL α domain ranked third and its score did not indicate a close structural conservation. The VHL α domain did not consistently rank high with other F-box sequences, although it did rank first with the CyclinF F-box. On the other hand, most F box sequences gave the same alignment with VHL α domain helices, and this alignment showed a good match between the hydrophobic VHL residues involved in α domain stabilization and hydrophobic F box residues (Figure 5.5, A and B). These observations suggest that there may be a loose structural similarity between the VHL α domain and the F box.

In the threading-based alignment, hydrophobic VHL residues that bind ElonginC are also conserved in the F-box motif, suggesting that the F box may bind Skp1 in a manner analogous to the VHL α domain-ElonginC complex. This raises the possibility of cross-reactivity between the two systems, but, as described in Chapter 2, we found little evidence for this in experiments in which the proteins were co-expressed in various combinations in *E. Coli*. This may indicate that no such cross-reactivity exists, although the negative result may be an consequence of the co-expression assay which was utilized.

Tyr^{98}, the second most frequently mutated residue found in VHL-related tumors, does not play any structural roles in VHL and makes no interactions with ElonginB or ElonginC. Its frequent mutation and location in the mutated patch on the surface of VHL argues for a functional role for this residue and the patch as a whole in the interaction of VHL with other macromolecules, most likely proteins. Ample space exists for such interactions to occur without

disturbing the VCB ternary complex. Thus one may envision a quaternary complex involving VCB and another protein, even one of considerable size (Figures 5.1, 5.2, and 5.3).

Several lines of genetic and biochemical evidence support such a role for VHL. Truncation of amino acids 1-93 results in a VHL protein which binds the ElonginBC complex as well as CUL2, but which is deficient in the regulation of hypoxia-inducible mRNAs (Lonergan et al., 1998). In addition, the mutations $Tyr^{98}His$ and $Trp^{117}Arg$, both of which are residues exposed or partially exposed in the surface patch, prevent the observed interactions with fibronectin while maintaining VCB-CUL2 binding (Ohh et al., 1998). Thus, there are classes of mutations which interfere with reported VHL functions but which do not abrogate VHL interactions with the elongins or with CUL2. Structurally, this corresponds to a disruption in the β domain which leaves the α domain at least partially functional.

These results raise the possibility that the primary function of VHL is not to inhibit transcriptional elongation by competing with ElonginBC for ElonginA. In fact, several problems have surfaced with this initial model for VHL function. To begin, reported levels of the ElonginBC complex in the cell are 100 to 1000-fold higher than those of ElonginA and VHL (Kamura et al., 1998), which is inconsistent with a simple competition model. Secondly, the inhibition of elongation by VHL demonstrated by *in vitro* experiments has never been shown to be specific, and the effects of mass action could explain the majority of these biochemical experiments. Finally, there exists at present no data demonstrating that the inhibition of SIII activity observed *in vitro* is significant *in vivo*.

If VHL functions in a pathway different or in addition to transcriptional elongation, the model most consistent with the majority of published data involves an analogy with the SCF complex. It is certainly difficult to ignore the

extensive sequence homology between the various components of a functional SCF complex (Cdc53, Skp1, and ubiquitin) and that of the VCB-CUL2 complex. The simplest extension of the analogy would be to suppose that the VCB-CUL2 complex, like the SCF complexes, functioned to regulate the degradation of specific polypeptides. In such a system VHL would be employed as an adaptor molecule targeting proteins to the CUL2-BC complex for ubiquitin or ElonginB conjugation and eventual proteolysis. Such a role would be consistent with the conclusions reached above that VHL function involves both its binding to CUL2-BC via the α domain and its binding to other molecules through the β domain mutational patch.

This process could occur through the ubiquitination of substrates, or, given the presence of ElonginB in the complex, through the conjugation of this ubquitin-like (Ubl) molecule to other proteins. The presence of the ubiquitin-like molecule ElonginB in the ElonginC-VHL complexes has also been suggested to support a role in protein degradation (Lonergan et al., 1998), although it is not yet known whether ElonginB is processed for conjugation. The possibility is intriguing for several reasons. First, all known ElonginB homologs contain a carboxy-terminal glycine residue as do ubiquitin and conjugated-Ubl molecules by which these latter proteins are conjugated to other polypeptides. The carboxy-terminus of ElonginB extends out to within less than 10 Å of VHL before the density in the crystal becomes disordered. Twenty amino acids remain, however, and this long extended tail of ElonginB could allow its conjugation to proteins bound to the distal face of VHL. In the three-dimensional structure of ElonginB, lysines are located in positions similar to those in ubiquitin (Lys^{29} and Lys^{48}) which are implicated as sites for further ubiquitination in the production of poly-ubiquitin moieties. It is conceivable, therefore, that an analogous process occurs in the VCB-CUL2 complex in which multiple molecules of ElonginB or

ubiquitin are linked by these lysines and serve, just as poly-ubiquitination does, as a signal for degradation by the 26S proteasome. It should also be recalled that the ElonginB structure is closer to the ubiquitin structure than non-conjugated ubiquitin-like proteins, although it is farther than those that are known to be conjugated. It is possible, however, that ElonginB may serve a structural role helping to stabilize the ElonginC structure. This would be consistent with the extensive ElonginB-ElonginC interface observed in the structure (Figure 4.9). It may even serve a ubiquitin-like targeting role through non-covalent attachment.

Regardless of the exact role of ElonginB, a model in which VCB-CUL2 targets proteins for degradation could explain several experimental observations. For example, the accumulation in VHL (-/-) cells of mRNAs for several proteins, including VEGF, appears to be largely due to the stabilization of their mRNA transcripts (Levy et al., 1996; Iliopoulos et al., 1996; Gnarra, 1996). Recently, three proteins were identified which form a complex which binds to adenylate-uridylate-rich elements (ARE) in the 3'-untranslated region of the mRNA for VEGF and another hypoxia-induced gene GLUT-1 (Levy et al., 1996). ARE sequences are often found to be destabilizing to mRNA transcripts, and the modulation of transcript stability is often achieved by protein-RNA interactions along these regions (Levy et al., 1996). These proteins which bind ARE sequences for VEGF and GLUT-1 are hypoxia-inducible in normal cells, but are constitutively elevated in tumor cells lines lacking functional VHL (RCC 780-O cells). Upon the reintroduction of wildtype VHL, this complex regains normal regulation by hypoxia, and under normal oxygen levels cannot be detected in the cell (Levy et al., 1996). Thus, these ARE binding proteins are affected in a similar manner to VEGF mRNA levels. If they do indeed stabilize these mRNAs, then VHL-related targeting of such genes for proteolysis would diminish the stability of the message transcript, whereas in cells lacking functional VHL, these proteins

would persist, stabilizing the mRNA, and potentially leading to the abnormal levels of the associated growth factors found in tumor cells.

A second example involves the protein fibronectin. Certain misfolded polypeptides in the ER have been shown to undergo retrograde transport to the cytoplasmic surface, where they are targeted for proteolysis by ubiquitination (Ohh et al., 1996). The coimmunoprecipitation experiments which indicated that VHL and fibronectin form a complex *in vivo* may be explained by an analogous process by which compromised fibronectin is specifically targeted by VHL to the CUL2-BC complex to be marked for degradation. The abnormal fibronectin matrix found in tumors lacking functional VHL may be attributable, at least in part, to the failure to degrade compromised fibronectin and its subsequent deposition in the extracellular matrix.

Results indicating that VHL inhibits PKC phosphorylation of Sp1 could be consistent with a VHL induced proteolysis of this and potentially other kinases, although no evidence currently supports such a conclusion. It is less obvious how the stabilization of p27 could be associated with the reintroduction of functional VHL into tumor cells. One possibility is that p27 lies far down in pathways regulated by VHL, and that its stabilization is due to indirect effects from other molecules.

The logical culmination of this model is also in analogy to the SCF, and would encompass the combinatorial nature of the F-box hypothesis (Patton et al., 1998a) as illustrated in Figure 5.7. In the ubiquitin proteolytic pathway, E3 enzymes, which are often multiprotein complexes such as the SCF, serve as determinants of substrate specificity for ubiquitination, and the wide variety of proteins degraded by this pathway is matched by a similar variety in E3 complexes. Cdc53 and Skp1 are invariant members of the SCF in yeast, whereas the F-box protein changes depending on the substrate. Thus the F-box protein

functions as a molecular adaptor, allowing for the targeting of many proteins to the same complex for ubiquitination. In yeast, at least three different F-box proteins - Cdc4, Grr1, and Met30 - target other molecules for ubiquitination by the SCF. Cdc4 and Met30 bind the target proteins via WD40 motifs, and Grr1 does so through a leucine rich repeat. Interestingly, the different families of SOCS-box proteins, all likely sharing structural homology to the VHL α domain and several of which are now known to bind the Skp1 homolog ElonginC, all possess different N-terminal motifs known to mediate protein-protein interactions, including WD40 repeats, ankyrin repeats, and SH2 domains (Patton et al., 1998a). Thus, the CUL2-BC complex may also possess a repertoire of adaptor molecules which can interact with both ElonginC as well as other classes of proteins, and this complex may target these molecules for degradation. The high excess of the ElonginBC complex to VHL in the cell would be explained by such a system, as its role would be far more extensive than interactions with any one class of molecules.

Such a model has several implications. The most obvious is that the CUL2-VCB complex may have ubiquitin, or ubiquitin-like, ligase activity, and that conjugated proteins (whether with Ub or ElonginB) are targeted for proteolysis (or potentially a different fate such intracellular trafficking). The second implication is that VHL and other potential adaptors bind proteins which are regulated by specific proteolysis, and, in the case of VHL, this binding (and subsequent targeting) is dependent upon the mutational patch in the β domain. This model is consistent with results in which the mutation of residues on the distal face of the β sandwich of VHL lead to impairment of the regulation of hypoxia-inducible mRNAs and the assembly of the extracellular fibronectin matrix, and that this impairment occurs despite the interaction of VHL with the CUL2-BC complex.

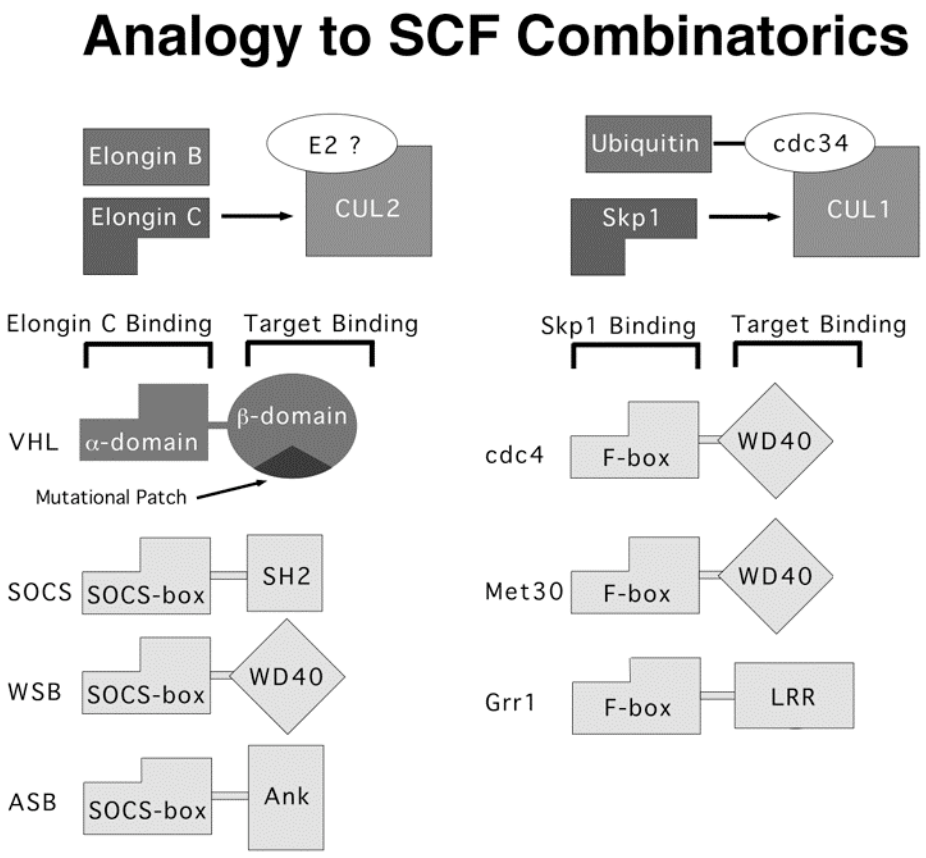


Figure 5.7. VCB analogy to the SCF complex

The structure of the VHL-ElonginC-ElonginB complex supports the hypothesis that ElonginC may function in a combinatorial regulatory system analogous to the SCF complex. It is not yet clear whether ElonginB would have a ubiquitin-like role in the ElonginC system.

5.5. Summary

A function for the VHL-ElonginC-ElonginB complex in protein degradation has been proposed based on the sequence homology between ElonginC and Skp1 as well as on the binding of ElonginC and Skp1 to the cullin proteins Cul2 and Cul1, respectively (Lonergan et al., 1998).

The structure of the ElonginB-ElonginC-VHL complex extends these similarities: (i) it indicates that Skp1 would not only have the same overall tertiary structure as ElonginC, but would also contain a protein binding site similar to that used by ElonginC to bind VHL; (ii) it suggests that the F box may contain a three-helix structure similar to that of the VHL α domain and that the F box-Skp1 complex may form in an overall analogous way to the VHL-ElonginC complex; (iii) it reveals that VHL has two domains, each containing a protein binding site, much like the modular F box proteins that contain multiple protein-binding motifs, such as WD-40 and leucine-rich repeats; and (iv) it indicates that VHL is part of the SOCS-box superfamily, based on the prediction of structural conservation between the SOCS box and the VHL α domain. Members of this family contain additional protein-binding modules such as SH2 (SOCS proteins), ankyrin (ASB proteins), or WD-40 (WSB proteins) domains, paralleling the F box proteins which form a diverse family of protein adaptors (Patton et al., 1998a; Patton et al., 1998b).

The ElonginC system probably regulates a cellular process in a combinatorial manner by binding different adaptor proteins. That this process may be protein degradation is further supported by the additional analogies between the ElonginC-VHL complex and the SCF complex revealed in this study.

It is conceivable, however, that this combinatorial system regulates a different process. Identifying the proteins that bind the β domain of VHL could help us understand what process is regulated by this complex and how VHL exerts its tumor suppressor effects.

VHL thus appears to function in the context of an intricate biological pathway which potentially regulates a diverse set of proteins and possibly mirrors similar pathways regulating different classes of molecules from yeast to humans. Viewing VHL from such a central perspective is in sharp contrast to the years when the associated syndrome was viewed as a rare medical condition of lesser significance in understanding biological function than disorders which received more fanfare. This revolution in the perception of von Hippel-Lindau disease is aptly summarized by a pioneer in the field of tumorigenesis, Dr. Alfred Knudson, when he spoke at the keynote address for the First International Symposium on von Hippel-Lindau which took place in Freiburg, Germany in 1994:

> "In the last ten years, von Hippel-Lindau disease has gone from an obscure medical curiosity to a condition with far-reaching implications in oncology."

These implications are still being worked out even as this thesis is written, but it is hoped that the work herein will substantially contribute to this growing field.

CHAPTER SIX:

HSP90 EXPERIMENTAL RESULTS

6.1. Purification and crystallization of Hsp90 and Hsp90 complexes

The cloning, biochemistry, and initial crystallization for this project were performed by Alicia Russo, as has been described (Stebbins et al., 1997). Briefly, bovine Hsp90, prepared from brains as described (Koyasu et al., 1986), was digested by subtilisin and the products were delineated by N-terminal sequencing and mass spectroscopy. The C-terminus of the middle domain was estimated based on its mobility in SDS-PAGE because this fragment was not amenable to mass spectroscopic analysis. The corresponding fragments of human Hsp90 were overexpressed in E. coli using the pET3d vector (Novagen), and the recombinant domains were purified by ion exchange and gel filtration chromatography.

Proteolytic digestion determined that the bovine Hsp90 contains three structural domains (Figure 2.4): an N-terminal domain of approximately 25 kDa, a middle domain of approximately 35 kDa, and a C-terminal domain of approximately 10 kDa. Each of the three corresponding recombinant domains of human Hsp90 were then assayed for geldanamycin binding using gel filtration chromatography to remove excess drug, and reversed phased HPLC to detect the bound drug. It was determined that the N-terminal domain, which has the highest sequence conservation among the three domains, binds geldanamycin at an approximate one molar ratio.

For crystallization, the recombinant human Hsp90-GBD (residues 9-236) was concentrated by ultrafiltration to 30 mg/ml, and was further purified by gel filtration chromatography. The geldanamycin–Hsp90-GBD complex was prepared by mixing the purified protein with a 2-fold molar excess of geldanamycin (Gibco-BRL), fractionation by gel filtration, and concentration by ultrafiltration to 17 mg/ml. Two crystal forms were grown at 7° C using the method of hanging-drop vapor diffusion (Ollis and White, 1990). One form (Native1 and Native2 in Table 2.2) grew from a buffer of 0.1 M sodium cacodylate, 0.2 M ammonium sulfate, and 30 % w/v polyethylene glycol 8000 (PEG8000), pH 6.5, in spacegroup I222 with a = 66.2 Å, b = 90.1 Å, c = 101.2 Å. Another form (Native3) grew from 0.1 M Tris-HCl, 0.2 M magnesium chloride, and 30 % PEG4000, pH 8.5, in spacegroup $P2_1$ with a = 53.4 Å, b = 44.1 Å, c = 54.0 Å, $\beta = 115.7°$. Crystals of the geldanamycin–Hsp90-GBD complex (GDM-complex in Table 2.2) grew from a buffer of 0.1 M Tris-HCl, 0.2 M sodium acetate, and 33 % w/v PEG4000, pH 8.5 in spacegroup $P2_1$, with a = 53.7Å, b = 44.3Å, c = 54.6Å, $\beta = 116.1°$ (same space group as Native3), and had a violet coloring. In addition to magnesium and geldanamycin, the growth of the $P2_1$ form can be induced by acidic pH as well.

6.2. Crystallography of Hsp90 and the Hsp90-geldanamycin complex

6.2.1. Data collection and the MIR Solution

Data were collected on an R-AXISIIC imaging plate detector mounted on a Rigaku 200HB x-ray generator, at 10 °C for Native1 and the derivatized crystals. Crystals were mounted in capillary tubes (Charles Supper Company) along with a small amount of mother liquor at one end of the tube. The tube

was then sealed at both ends with wax. For the other data sets, crystals were flash-cooled at -160.0°C with gaseous nitrogen either in 30 % PEG8000 and 20 % glycerol (Native2), or in 35 % PEG4000 (Native3 and GDM-complex). The flash-cooled crystals were mounted in thin nylon loops.

Derivatives were obtained by transferring individual protein crystals into stabilization solutions containing various heavy atom compounds at different concentrations. The process required that first a protein-free stabilization solution, i.e., a solution in which the protein crystals remained unharmed and retained their high resolution diffraction, be obtained. Through trial and error it was determined that a solution containing 75mM sodium cacodylate (pH 6.5), 0.15M ammonium sulfate, and 40% w/v PEG8000 possessed these properties. Soaking of the crystals proceeded in two steps. The first consisted of a washing of the protein crystal in 200μL of stabilization buffer in order to remove any residual non-crystalline protein. The drop containing the crystal was first diluted with stabilization buffer (typically adding approximately 50 μL to the 6 μL drop volume). The crystal was transferred using a standard hand-held pipetman in which only a small volume of drop solution (on the order of 5-10 μL) was carried along with the crystal. Following this initial washing, the crystal was similarly transferred into stabilization buffer plus heavy atom.

Initially, heavy atoms were screened for reactivity by soaking crystals in very high concentrations of various compounds. Examples included 10mM (or saturated) of mercury chloride, gold chloride, uranyl acetate, and lead acetate, among many others. Conditions which lead to crystal cracking and/or discoloration were pursued more extensively at more reasonable concentrations in a trial and error fashion. Such soaks were screened by

analyzing diffraction for significant mean isomorphous differences in reflections at an early stage in data collection when completeness was low (on the order of 10-20%). Soaks which demonstrated significant mean isomorphous differences (between 10 and 30%) were exposed to the beam until data sets were relatively complete, after which were constructed isomorphous difference Patterson maps. These maps were examined for significant peaks along possible Harkker sections both for I222 and $I2_12_12_1$.

The initial search for heavy atom derivatives was performed on crystals held at 10°C. Since the crystals were grown at 7°C, this represented little change. Although frozen data of high quality was used after the 10°C structure was determined, its use was avoided in the initial phasing in order to prevent non-isomorphism problems often associated with flash freezing of protein crystals. The non-frozen crystals were mounted in quartz capillary tubes and cooled through the use of gaseous nitrogen

The heavy atom derivatives used in the multiple isomorphous replacement (MIR) solution were prepared by soaking I222 form crystals in a buffer of 75 mM sodium cacodylate, 0.15 ammonium sulfate, and 40 % PEG8000, pH 6.5 containing one of the following heavy atom solutions: saturated K_2AuCl_4 for 52 h, 20mM K_2PtCl_4 for 24 h, and 20mM K_2PtCl_6 for 60 h. An additional derivative was obtained in 5 mM sodium cacodylate, 250 mM NaCl, 33 % PEG8000 with 0.5 mM $Sm(OAc)_3$, pH 7.0 for 14 h.

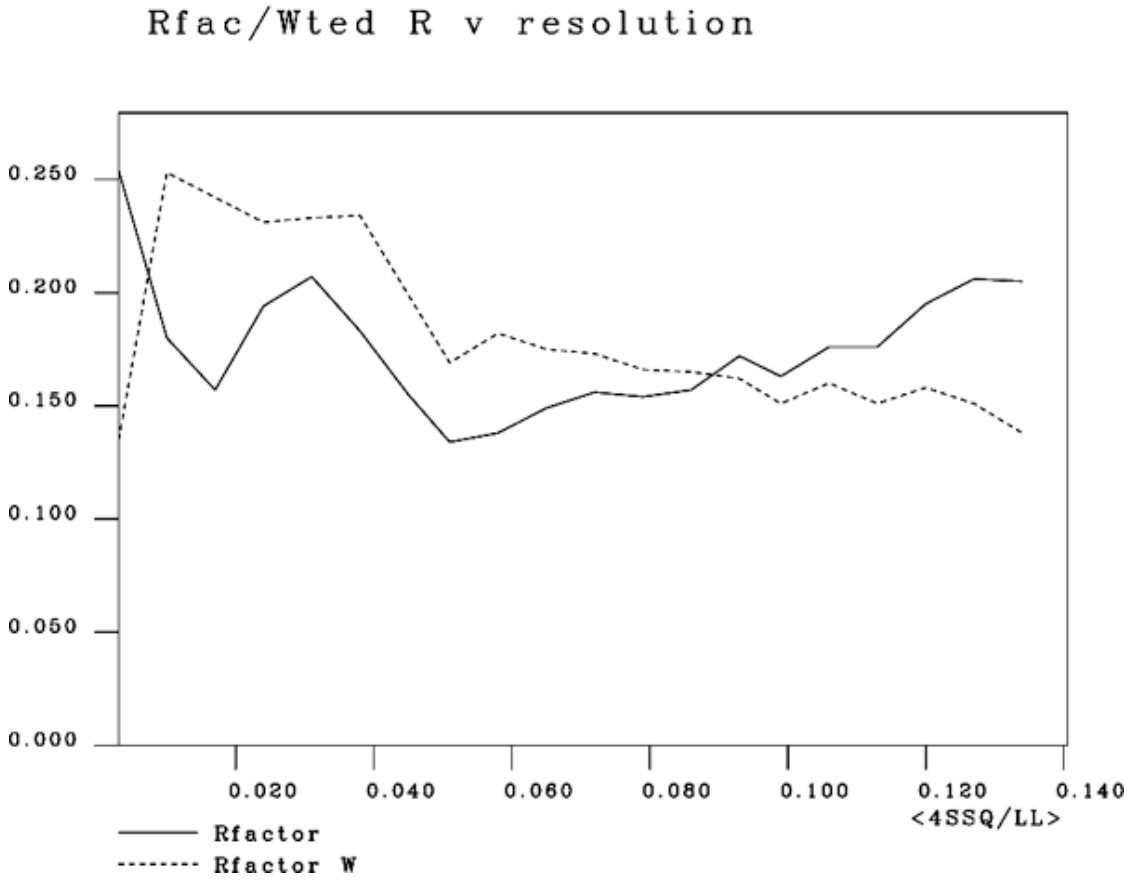
The obtaining of an isomorphous heavy atom derivative with a strong signal and data of high quality serves as the focal point for subsequent phasing and identification of other heavy atom sites through difference Fourier synthesis. Although the first derivative with an identifiable signal was obtained with a platinum compound (Table 6.1), it proved weak in the

calculation of phases due to non-isomorphism between it and the native. A strong first derivative was found in the compound gold chloride, which while insoluble beyond 1mM in the stabilization buffer used, non-the-less succeeded in substituting very well in the crystal, in fact coloring the crystal a yellow-gold. In addition, this derivative was very isomorphous with the native, and phase information extended to 2.5Å resolution (Figure 6.1A). This derivative was obtained by soaking a crystal in saturated gold chloride for 52 hours. A summary of the data associated with this soak and others used in the MIR solution can be found in Table 6.1.

The Patterson calculated from the gold derivative data revealed a strong and obvious peak (Figure 6.1B) on the Harkker sections associated with constant integer values, as consistent with the I222 spacegroup, and not on the one-half integer sections, as would be expected with a spacegroup containing screw axes such as $I2_12_12_1$. This fact determined that the spacegroup was I222. From this peak (on the three Harkker sections) the real space coordinates of the heavy atom site (or their enantiamorphs) were determined. SIR phases were calculated using the program MLPHARE (CCP4, 1994) and improved with solvent flattening with DM (CCP4, 1994). The solvent flattened phases were used in difference Fourier synthesis to search for additional, weaker sites in the gold chloride derivative, and for sites in the data sets from different soaks.

A total of four different compounds were used in five different data sets for the final calculation of the phases. The gold chloride compound was used twice: in one data set, the soak time was 52 hours with a saturated solution, whereas in the other it was 20 hours with 0.5mM gold chloride. These different soaking conditions led to different occupancies for the heavy

A



B

Figure 6.1. The gold chloride MFID and Patterson plots

Table 6.1. Summary of Hsp90 crystallographic data

		Diffraction Data				MIR Analysis		
Crystal	Resolution (Å)	Unique Reflections	Measured Reflections	Coverage (%)	R_{sym} (%)	R_{iso} (%)	PhP	R_c
Native1	2.30	13532	43812	98.0	4.4	-	-	-
Native2	1.65	34516	361958	98.2	3.8	-	-	-
Native3	2.20	11381	51794	97.7	3.5	-	-	-
Complex	1.90	16906	86335	92.2	2.5	-	-	-
K_2AuCl_4	2.50	10254	49642	94.6	5.7	17	2.43	0.54
K_2PtCl_4	3.20	4713	20129	91.6	5.8	28	1.07	0.83
K_2PtBr_6	3.20	4831	25025	93.1	4.7	22	0.75	0.90
$Sm(OAc)_3$	3.20	4442	16281	86.0	7.1	30	0.65	0.92

Refinement and Analysis of Hsp90 Molecular Models

Crystal	Resolution (Å)	Reflections ($\|F\| > 2\sigma$)	Atoms Modeled (protein, H_20)	R_{cryst} / R_{free} (%, %)	r.m.s deviations bonds(Å), angles(°), B-factors(Å^2)
Native1	7.0 - 2.30	12445	1626, 96	16.9, 25.2	
Native2	6.0 - 1.65	32524	1626, 335	19.4, 24.4	0.007, 1.6, 3.3
Native3	7.0 - 2.20	10845	1679, 246	18.9, -	0.010, 1.7, 3.2
Complex	6.0 - 1.90	16080	1679, 288	18.9, -	0.009, 1.8, 3.2

$R_{sym} = \sum_h \sum_i \; | \; I_{h,i} - I_h \; | \; / \; \sum_h \sum_i I_{h,i}$, for the intensity ($I$) of i observations of reflection h. $R_{iso} = \sum \; | \; F_{PH} - F_P \; | \; / \; \sum \; | \; F_P \; |$; PhP = Phasing Power = $\sum [F_H(calc)^2 \; / \; (F_{PH}(obs) - F_{PH}(calc))^2]^{1/2}$ F_P = native structure factor amplitude; F_{PH} = derivative structure factor amplitude. $F_{PH}(obs)$, $F_{PH}(calc)$ = observed, calculated derivative structure factors, respectively. $R = \sum \; | \; F_P - F_{calc} \; | \; / \; \sum F_P$. F_{calc} = model structure factor; 10% data omitted for R_{free}. Bond and angle deviations are from ideal values; B-factor deviations are between bonded atoms.

atom sites, providing slightly different phasing information. Multiple isomorphous replacement phases were calculated and heavy atom positions and occupancies refined (although never simultaneously) using the maximum likelihood method as implemented in the program MLPHARE..

Once the refinement had converged, the phases were used in a real space density modification algorithm as implemented in the program DM. Solvent flattening and histogram matching were used in conjunction with phase extension to the limit of the gold chloride data (2.5Å). Throughout the use of density modification, a value somewhat analogous to the Brünger free R value (Brünger, 1992) was used to judge the progress and to modify the parameters used in DM. It was found that those protocols which minimized the DM free R value corresponded to the best electron density maps, and following the first few such evaluations, a protocol was established the consisted of using only solvent flattening and histogram matching with phase extension to 2.5Å (Sayre's equation was not used nor skeletonization, as both of these algorithms resulted in diminished quality to the electron density maps and higher DM free R values). In addition, it was found that using phase extension with many very small steps gave superior results to larger but fewer steps (in the end, I settled on 100 steps, extending the resolution from the 3.2 Å of the initial MIR solution to 2.5Å).

Following density modification, the resultant phases were used as a starting point in MLPHARE for further refinement of the heavy atom parameters, and then a subsequent recalculation of the phases. This process of cycling between phase refinement of the heavy atom positions and density modification was carried out until the FOM, as calculated by MLPHARE, was optimized (four such cycles). This stage corresponded to the best electron

density maps (Figure 6.2), and the phase extension allowed for the use of higher resolution maps in later stages of model building in which the sequence was interpreted. Final values for the MIR solution are shown in Table 2.3. Analysis of these statistics suggests that the resultant phases were the product of the extremely strong gold chloride derivative weakly supplemented with poor phase information from the three other compounds. Thus, the electron density maps could be viewed as very good SIR maps plus weak additional phasing from other derivatives, with the final phases improved extensively by solvent flattening and histogram matching.

6.2.2. Model Building and Refinement of apo-I222 Form

The initial electron density maps revealed several unambiguous and extended regions of secondary structure. Two α helices and several β strands were built initially with a protein chain consisting only of alanine side chains. The electron density maps were visualized and the atomic model built using the program O. Although residues were placed conservatively, i.e., only where there was unambiguous secondary structure, it was possible from the initial maps to trace 80% of the amino acids in the Hsp90-GBD (corresponding to approximately 60% of the total atoms) with the poly-alanine chain. Several cycles of positional , least-squares refinement followed by simulated annealing with the program X-PLOR (Brünger, 1988), weighted combination between model phases and MIR phases using SIGMAA (CCP4 , 1994), and model building with O, allowed for a continuous trace and sequence interpretation of amino acids Val^{17}-Glu^{223}. Amino acids 9-16 and 223-236 of the fragment used in these studies were not visible in the electron density maps and presumed to

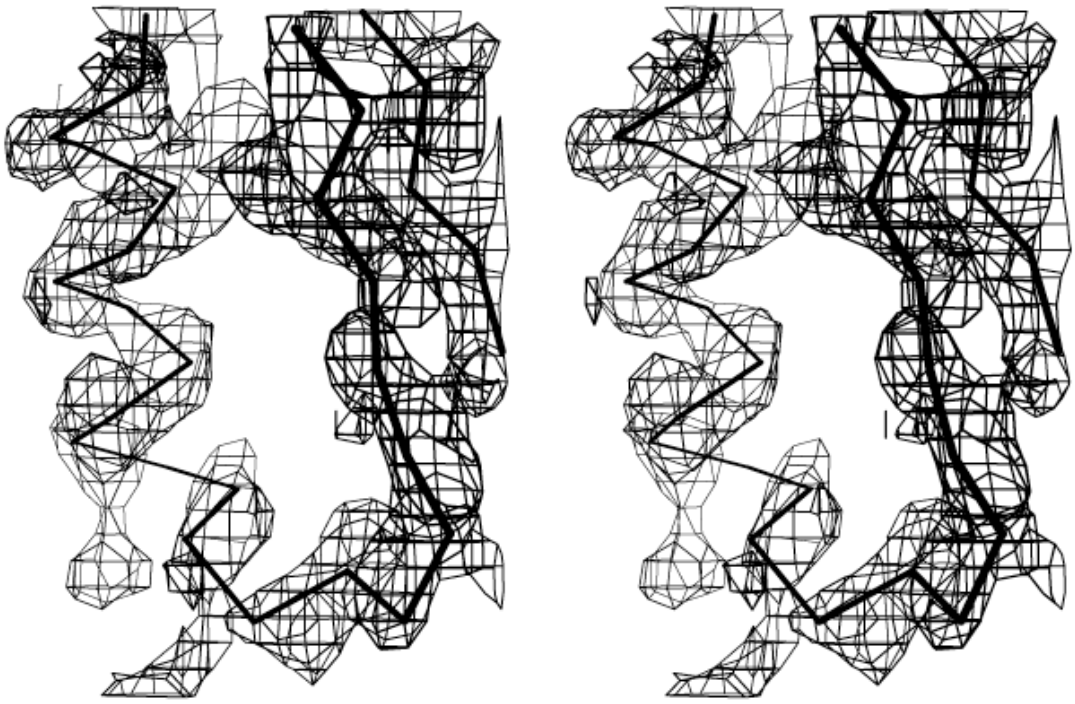


Figure 6.2 MIR density

Stereo image of the solvent flattened MIR density at 3.8Å resolution with Cα trace from the final model .

be disordered. In the end, 16 partial models were constructed, and a total of 22 models were constructed before the structure was finalized. Throughout all steps in the refinement, both the standard R value (R_{cryst}) and the Brünger free R (R_{free}) were monitored carefully and used to judge the outcome of each stage of model building. In addition, as the model became more complete, each stage was also judged by the Ramachandran plot of backbone torsion angles.

The use of weighted phase combination between the phases calculated from the partial model and those from the MIR/solvent flattened calculations proved very important to the connection of polypeptide chain segments for which the electron density was discontinuous. As mentioned, at the initial stages it was simple to trace nearly 80% of the polypeptide backbone, but what remained consisted of stretches of secondary structure unconnected. Phase combined maps had the tendency to show continuous density, or near continuous density, in regions where the MIR maps were unclear. Thus, after a few significant stages of rebuilding and refinement, some new stretch of hitherto discontinuous density would become clear, and allow for the connection of two regions of secondary structure. This in turn would improve the model phases so that the next stage would produce phase combined maps perhaps allowing another connection to be made. This process continued until model number 16, when all connections were made, and the trace from amino-acid 17 through 223 was complete. In the later stages of model building, when the sequence had been extensively assigned, severe constraints were put on which connections were possible, and this was a great aid in the establishment of connectivity.

The sequence assignment began with the modeling of obvious aromatic side chains. Fortuitously, there were five sequences containing multiple aromatics which served well to assign large stretches of secondary structure. However, these stretches were clustered near the beginning of the protein, or near its end, and a large stretch, consisting of amino-acids 45-133 (roughly 40% of the total residues) remained without such clear signposts. This remainder was filled by extending the known regions as far as possible before encountering a discontinuity. In addition, this stretch contained three methionine residues, and data collected from frozen crystals was used to detect the anomalous signal in the sulfur atoms, and difference Fourier maps revealed the presence of several peaks, confirming the sequence assignment.

Once R_{cryst} was below 30%, individual (and one overall) isotropic temperature factors were allowed to vary in the refinements, although never simultaneously with the spatial coordinates. This produced an immediate drop in both the R_{cryst} and R_{free} of about 4%. This and model adjustment soon brought the R_{cryst} to near 25%, at which time ordered solvent molecules were modeled. Initially, very stringent constraints were placed on the modeling of solvent in order to isolate the best ordered molecules. Such a protocol would improve the model phases most quickly with the least risk of modeling noise. Once the well ordered solvent molecules were placed, the constraints were relaxed and more molecules added until the R_{free} was no longer beneficially affected, or the individual temperature factors for the added solvent molecules were too high compared to the rms B value for the atoms. For the Hsp90-GBD, all solvent molecules were modeled as water. Finally, once the addition of ordered solvent had minimized the R values, an overall anisotropic temperature factor matrix was refined.

By model number 22, the model refined against the 10 ˚C data had been optimized, with final R_{cryst} and R_{free} of 16.9% and 25.2%, respectively (Table 6.1, Native 1). This model, without water molecules, was used as a starting point to refine against a high resolution, frozen data set of the same crystal form. The frozen data (Table 6.1) extended beyond 1.5Å resolution, with high quality data to at least 1.65Å resolution, although due to detector geometry the completeness was limited much beyond this resolution. Initial refinement was limited to rigid body to allow for displacements incurred from the freezing process. This was followed by positional least-squares refinement and simulated annealing until the R_{cryst} was 30.8% and the R_{free} 34.3%. Water molecules were added and individual temperature factors refined until the R_{cryst} and R_{free} values approached 21% and 25% respectively, to a resolution of 1.65Å. The use of anisotropic temperature factor refinement was used as well, and the final refinement statistics may be found in Table 6.1 (Native 2).

The model quality analysis will focus on the R_{free} and Ramachandran stereochemical plot in order to achieve the best assessment. As shown in Table 6.1 and Figure 6.3 (along with more conventional, although less reliable, statistics such as the R_{cryst} and rms deviations from small molecule structures of the amino acids in bond lengths and angles) the R_{free} steadily decreased during the model building process and reached excellent final values of 25.2% and 24.4% in the 10 ˚C and frozen data sets, respectively. Finally, as illustrated in Figure 6.5, in both structures, all residues were found in the allowed regions of the Ramachandran plot, indicating that all backbone torsion angles were stereochemically feasible. In addition, both structures are characterized by near 90% of the residues falling into the most favored regions on the plot.

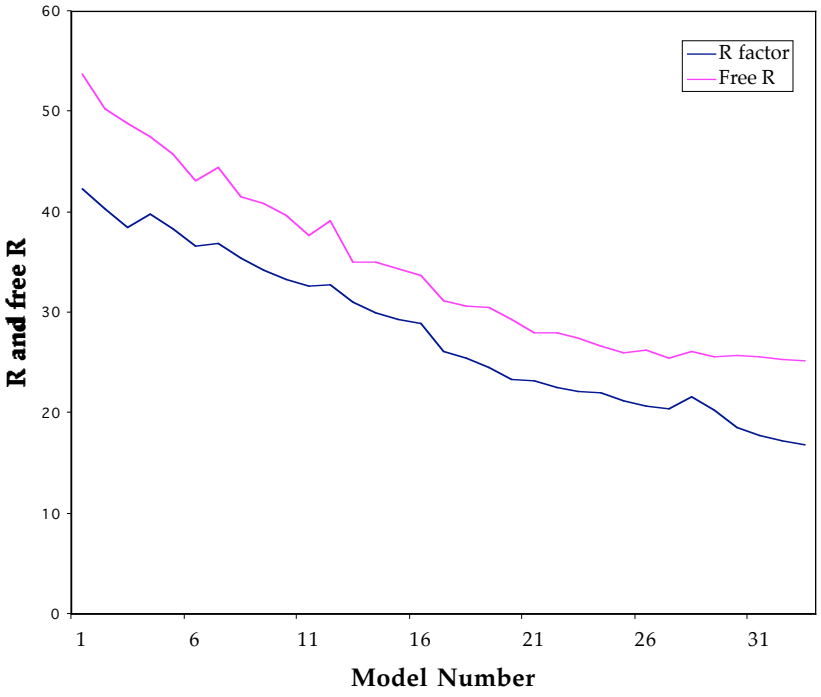
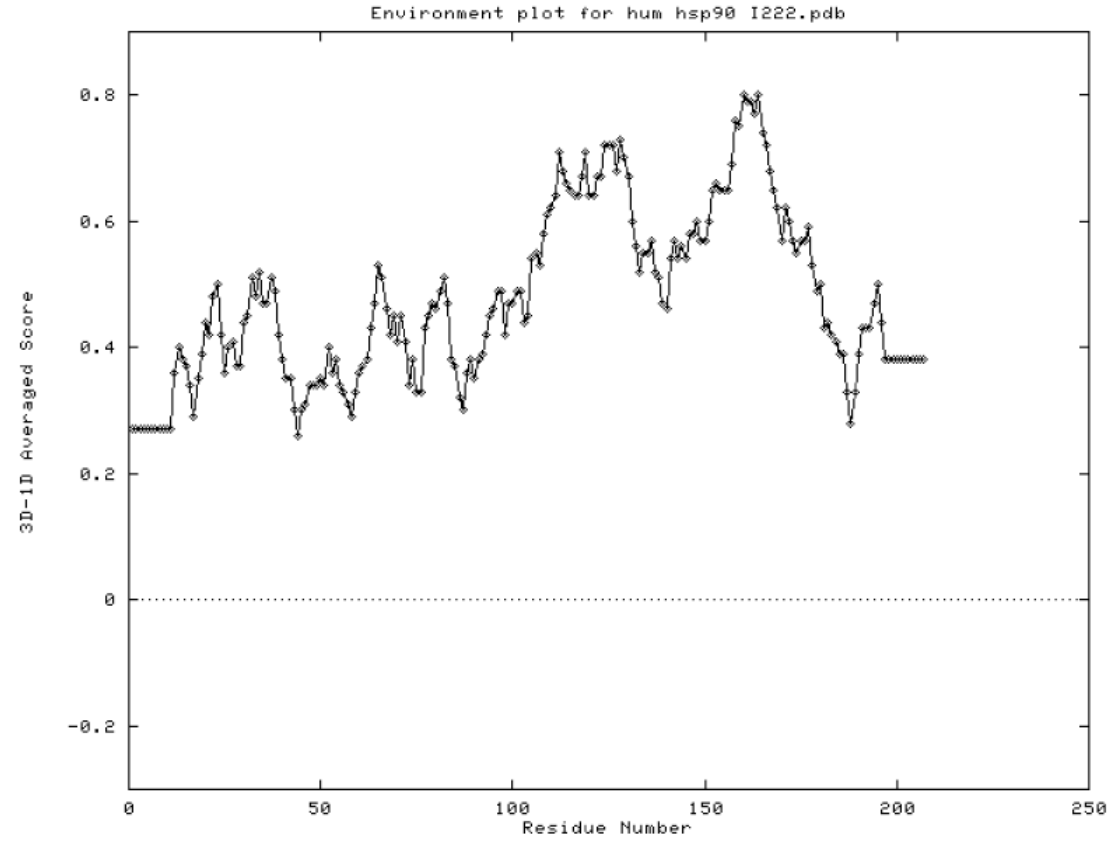


Figure 6.3: apo-Hsp90 free R graph

A



B

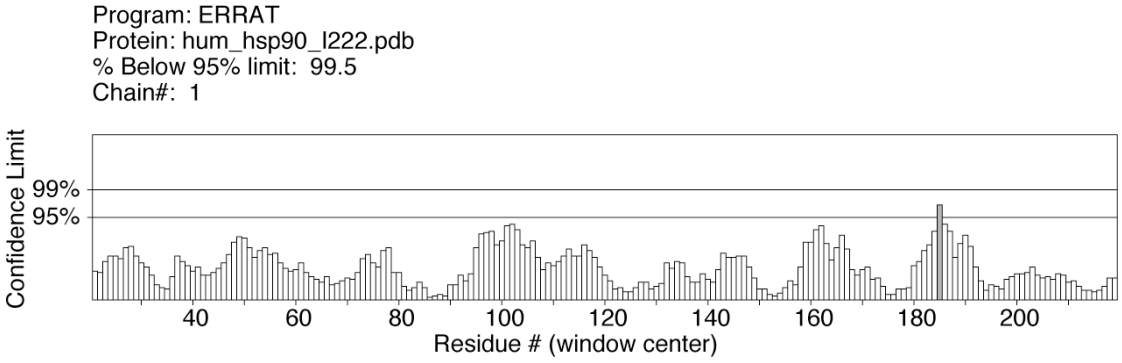
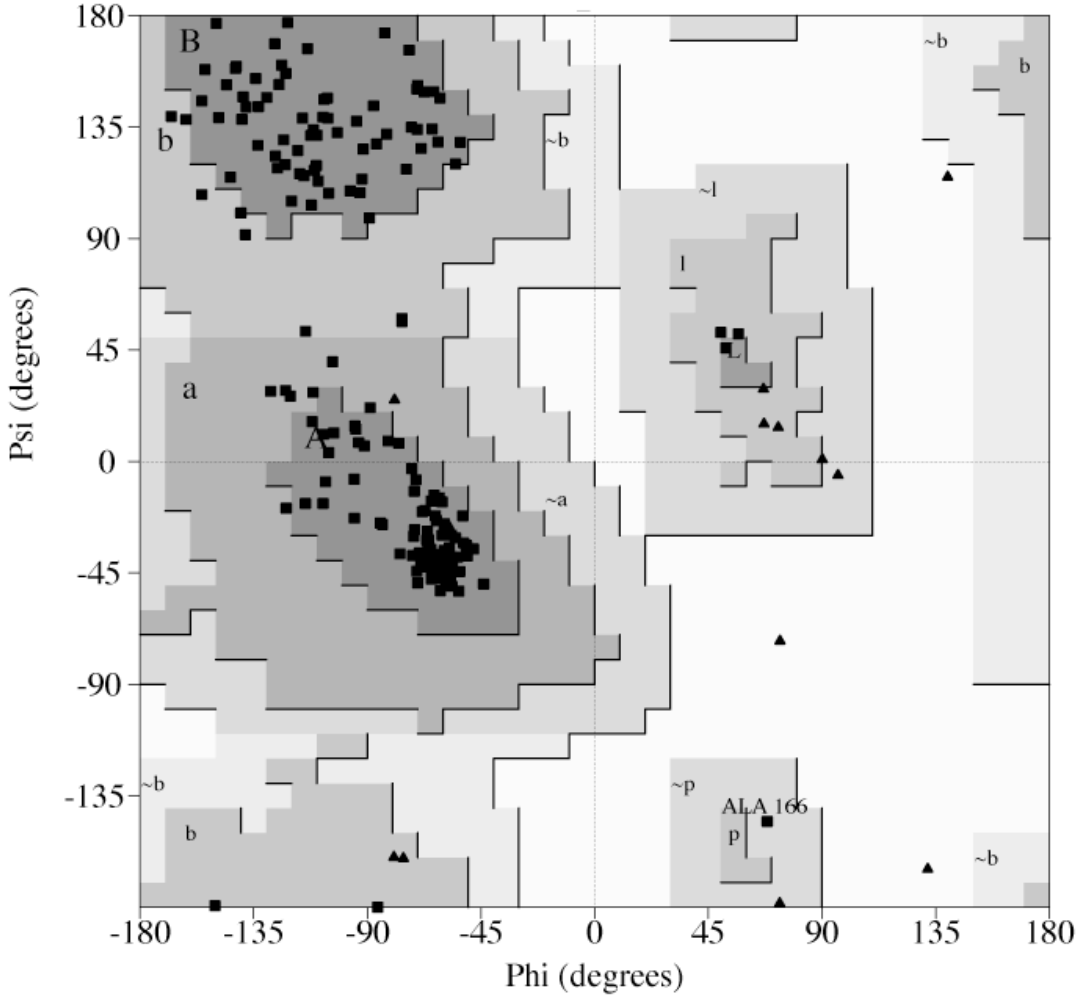


Figure 6.4. -160°C I222 model profile analysis

(A) Eisenberg profile **(B)** ERRAT profile indicating the number of residues which should be lower than a given value.



Plot statistics

Residues in most favoured regions [A,B,L]	165	88.7%
Residues in additional allowed regions [a,b,l,p]	20	10.8%
Residues in generously allowed regions [~a,~b,~l,~p]	1	0.5%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	186	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	15	
Number of proline residues	4	
Total number of residues	207	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

Figure 6.5. -160°C I222 apo-Hsp90 Ramachandran plot

6.2.3. $P2_1$ Crystal Form

As described above, crystallization experiments with the apo-Hsp90-GBD produced crystals in several different conditions. The I222 form crystals were obtained in 0.1 M sodium cacodylate, 0.2M ammonium sulfate, and 30% w/v PEG8000, pH 6.5 as well as in the condition 0.1M Tris-HCl, 0.2 M magnesium chloride, and 30% PEG4000, pH 8.5. This latter condition also yielded a different crystal form which was difficult to reproduce , generating crystals belonging to space group $P2_1$, with unit cell parameters a = 53.4 Å, b = 44.1 Å, c = 54.0 Å, $\beta = 115.7^{\circ}$. This rare result (no more than two or three such crystals were obtained) proved to be quite fortunate, as an examination of this crystal form revealed an important conformational change with implications on the function and regulation of this domain.

Having determined the structure of the Hsp90-GBD in the I222 crystal form, and assuming that no large scale conformational changes had occurred between crystal forms (i.e., that in the $P2_1$ form crystals, the protein was characterized by essentially the same structure as in the I222 crystals), it was logical to use the method of molecular replacement to place the I222 model into the correct orientation in the $P2_1$ crystals. This was performed using the program X-PLOR.

The initial rotational search after PC refinement revealed a clear correlation peak 3.5σ above the remaining possible solutions. This peak was used in a translational search which gave another unambiguous solution. Using these rotational and translation parameters, the I222 model (without water molecules) was positioned into the $P2_1$ unit cell, and rigid body

refinement, used in order to correct for gross errors in the rotational and translational variables, was performed with X-PLOR. The R_{cryst} after rigid-body refinement had a value of 36% to 4Å.

Upon examining the $2F_o$-F_c electron density maps (phased with the rigid-body refined I222 model phases), it was discovered that the loop L2, and the portions of the H4 and H5 helices immediately adjacent to the L2 loop, did not have any clear electron density (Figure 6.6). The rest of the model was characterized by electron density that matched the I222 model quite well. Although it was first supposed that perhaps this loop region (corresponding to amino acids 105-132) was disordered in the $P2_1$ crystal form, after carefully searching the surrounding space, continuous density for a differing conformation was discovered. The residues were fit into this new density, and the resultant structure revealed a significant and perhaps functionally relevant conformation change, although the structural change was certainly a small enough fraction of the total molecule for the molecular replacement solution to succeed.

To improve the density in the vicinity of the L2 loop, in order that it might be modeled as accurately as possible, water molecules were added with the stipulation that and any water molecules near this region were removed before electron density maps were calculated. This led to an improvement in the density in this region, but not sufficiently for trivial model building. Since the omission of correct model, as much as the incorrect placement of atoms, will lead to inaccuracies in the phases, I reasoned that perhaps by allowing water molecules to be placed where difference density was quite high might lead to maps which were more interpretable. This was indeed found to be the case, with the added advantage that the inserted water

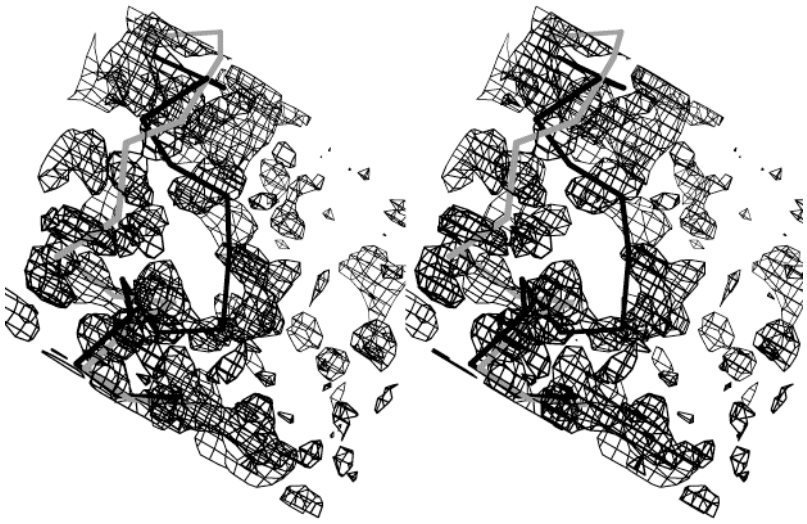


Figure 6.6. Comparison of 2Fo-Fc loop density for I222 and $P2_1$ models.

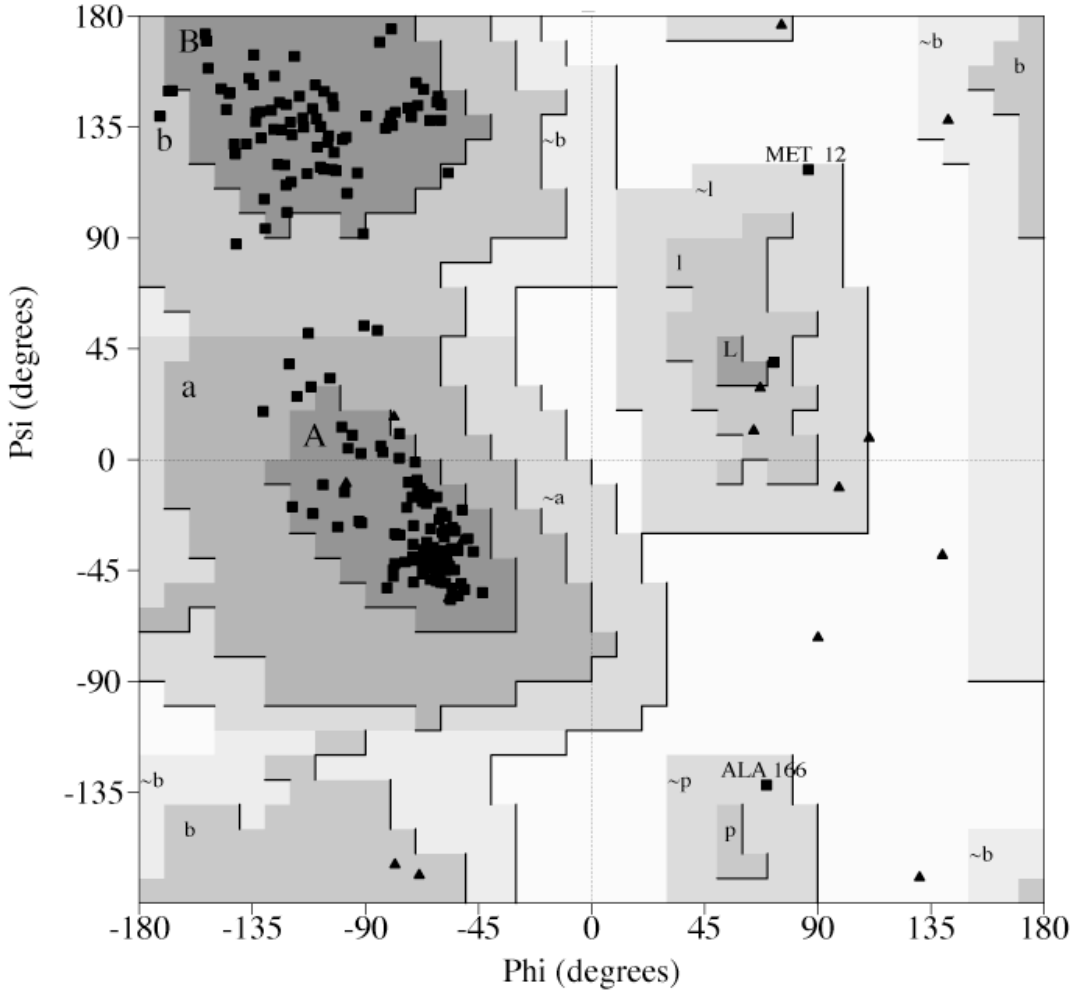
Stereo image of the 2fo-fc density for the $P2_1$ crystal form showing the models for the flexible loop in the I222 model (dark) and the final $P2_1$ model (light).

molecules traced the polypeptide chain quite well (often with waters occurring at C_α positions), and thus aided in model building. Examples of the $P2_1$ form electron density around the I222 model L2 loop, and around the final $P2_1$ form loop are shown in Figure 6.6. Final refinement statistics for the $P2_1$ model are in Table 6.1 (Native 3), and Figure 6.7 presents a Ramachandran plot of this model.

The rebuilding of the H4-L2-H5-H6 secondary structural elements reveals that a segment of approximately 35 amino acids, spanning residues 100-134, has two different conformations in the two different apo-Hsp90-GBD crystal forms. These may be termed the "open" and "closed" conformations, because they result in different pocket size and accessibility, with only the open conformation being compatible with the binding of the geldanamycin antitumor drug.

6.2.4. The Hsp90-Geldanamycin Complex

Attempts to soak geldanamycin into existing crystals of the I222 form (there were not enough $P2_1$ form crystals generated for this purpose) were not successful. In retrospect, with the knowledge that the closed I222 form was incompatible with drug binding, it is clear why these experiments failed. In is interesting to note the effect of such soaks of the I222 crystals. Short term soaks resulted in slightly etched crystals which still diffracted well, but which showed no sign of drug binding in the difference Fourier maps. Long term soaks resulted in some dissolving of the crystals, and more dramatically, a



Plot statistics

Residues in most favoured regions [A,B,L]	175	91.1%
Residues in additional allowed regions [a,b,l,p]	15	7.8%
Residues in generously allowed regions [~a,~b,~l,~p]	2	1.0%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	192	100.0%
Number of end-residues (excl. Gly and Pro)	1	
Number of glycine residues (shown as triangles)	15	
Number of proline residues	5	
Total number of residues	213	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

Figure 6.7. Ramachandran plot for $P2_1$ model.

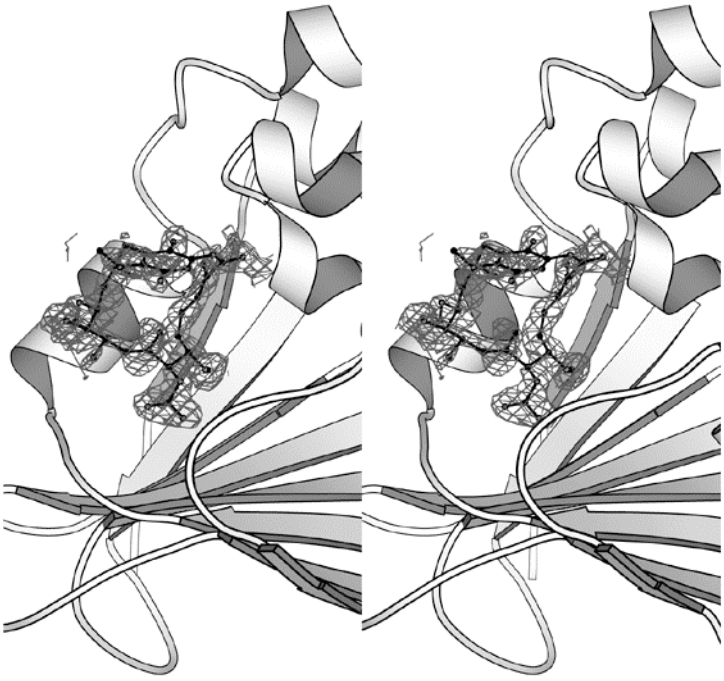
change in spacegroup. Diffraction from such soaks, of much poorer quality than unsoaked crystals (diffracting only to about 2.8Å), is characterized by loss of centering, with symmetry of the $P2_12_12$ class (also orthorhombic), but with essentially the same unit cell. This results in halving the number of asymmetric units and thus creating a dimer in the asymmetric unit. Molecular replacement was used to position the model in this crystal form. At the time this was done, the apo $P2_1$ form had not yet been solved, and thus when it was discovered that no density could be found for the L2 loop, it was assumed that it had become disordered in the new crystal packing conditions, and no further effort was invested in examining the density once it was clear that the drug was not present. Still, it is intriguing that the soaking of the drug caused a change in the protein structure, likely toward that of the $P2_1$ open form, the conformation consistent with drug binding. Further studies might perhaps reveal the presence of drug, with low occupancy due to the degradation of the data and the difficult of the drug entering the pocket through the solvent channels.

It was therefore necessary to co-crystallize the drug and protein. Since the complex was stable over a gel filtration column (more than 30 minutes), there was much confidence that co-crystallization would succeed for the appropriate conditions of crystallization. The Hsp90-geldanamycin complex was prepared by mixing the purified protein with a 2-fold molar excess of geldanamycin (Gibco-BRL), fractionation by gel filtration, and concentration by ultrafiltration to 17 mg/ml. Crystals of the complex grew from a buffer of 0.1 M Tris-HCl, 0.2 M sodium acetate, and 33% w/v PEG4000, pH 8.5 in spacegroup $P2_1$, with a = 53.7Å, b = 44.3Å, c = 54.6Å, b = 116.1° (the same space group as the apo-Hsp90-GBD second crystal).

The crystals were difficult to obtain, and required several weeks of growth to reach diffraction size. In fact, the crystal used in the structure determination was at the time the only such grown, and found by chance in one of many conditions. It appeared intriguing under the light microscope as it had a slight tan or brown color (although several soaks in geldanamycin of native I222 form crystals had similar color, but, as mentioned, did not contain any bound drug). When mounted and frozen for data collection, it was clear under better lighting conditions that the crystal was in fact a striking violet. This was even more interesting as it is known that the reduction of the benzoquinone in the drug causes geldanamycin solutions to be colored violet, and it was hypothesized that a similar electron distribution upon binding the protein may have been induced in the ring. Diffraction was measured to 1.9Å, and the statistics of the data can be found in Table 6.1 (Complex).

Since the crystals of the complex were of the $P2_1$ form of nearly identical unit cell dimensions as the apo-$P2_1$ crystals, the open, apo structure (omitting water molecules) was used to refine against the complex data. Rigid-body refinement was performed, after which a F_o - F_c difference Fourier map was then examined. This map (Figure 6.8) revealed unambiguous density for the drug in the pocket of the protein. To improve the density, water molecules were added outside the region and refinements of positional parameters and temperature factors undertaken (both least squares and simulated annealing).

A



B

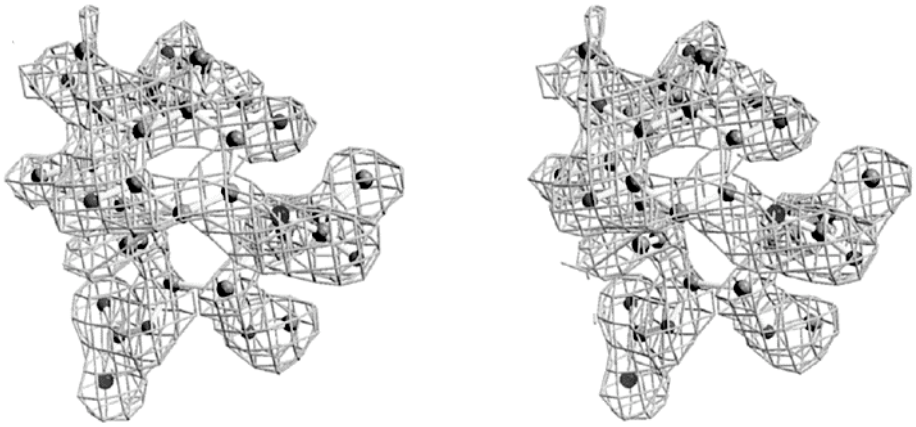
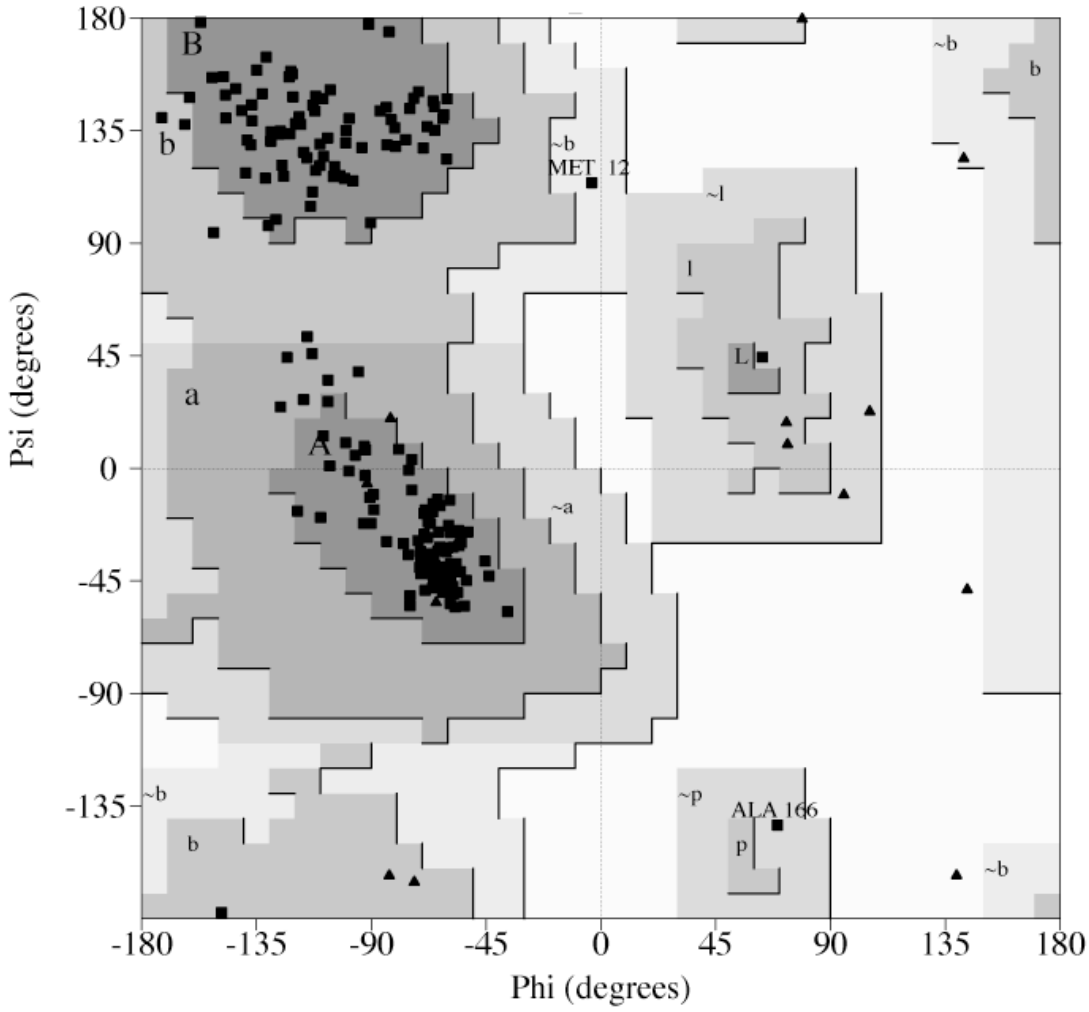


Figure 6.8. Fo-Fc map of Hsp90-GDM phased with $P2_1$ model.
Two views of geldanamycin density before refinement with the drug. **(A)** With the final model shown alongside as a ribbon diagram, and **(B)** with the final drug model.

In order to refine the complex structure, it was necessary to build a chemical description of the drug for use by X-PLOR (databases for protein structural elements are well developed and contained within the program distributions). Since it was possible only to obtain the coordinates for the related drug herbimycin A (Furusaki et al., 1980), this structure was used as a starting point and modified using the model building program O to be structurally identical to geldanamycin. These coordinates were then submitted to the program XPLOR2D (Jones et al., 1991), which generates appropriate chemical parameters for the given structure in the format used by X-PLOR. Although automated, the results required manual alteration as not all the chemical parameters were correct. The resultant topology and parameter files were used in conjunction with those for proteins in the X-PLOR refinements.

Geldanamycin was built into the improved difference density maps, and the structure refined to a final R_{cryst} of 18.9%. The final refinement statistics are in Table 6.1 (Complex), and a Ramachandran plot of the refined structure in Figure 6.9. There are no other structural alterations (apart from the L2 loop movement) in the Hsp90-GBD induced upon the binding of geldanamycin.



Plot statistics

Residues in most favoured regions [A,B,L]	176	91.7%
Residues in additional allowed regions [a,b,l,p]	14	7.3%
Residues in generously allowed regions [~a,~b,~l,~p]	2	1.0%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	192	100.0%
Number of end-residues (excl. Gly and Pro)	1	
Number of glycine residues (shown as triangles)	15	
Number of proline residues	5	
Total number of residues	213	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

Figure 6.9. Ramachandran plot of Hsp90-GDM protein model.

CHAPTER SEVEN:

THE HSP90 AND HSP90-GELDANAMYCIN STRUCTURES

7.1. Overall structure of the geldanamycin–Hsp90-GBD complex

The Hsp90-GBD has nine helices and an antiparallel β sheet of eight strands which together fold into an α+β sandwich (Figure 7.1, A-D). The β sheet is slightly twisted, and contains eight strands all running anti-parallel with the exception of S8, which runs parallel with S2 near the carboxy-terminus of the molecule. One face of the β sheet is hydrophobic and packs against a layer of five helices; four of these helices (H1, H2, H4, and H9) pack flatly against the β sheet, with their axes parallel to the β strands, while H7 packs against the β sheet at a steeper, almost perpendicular angle. The structure has a second layer of helices that pack on the first layer (H5 and H6), and two smaller helices (H3 and H8) at the periphery of the sandwich (Figure 7.1, A-D). It is noteworthy that four of the nine helices are of the 3_{10} type and comprise 11.2 % of the amino acids, which is significantly higher than the 3.4 % average in the protein database (Barlow and Thornton, 1988). At its center, the helical face of the sandwich has a wide opening that extends into the hydrophobic core of the structure and results in a pronounced pocket, about

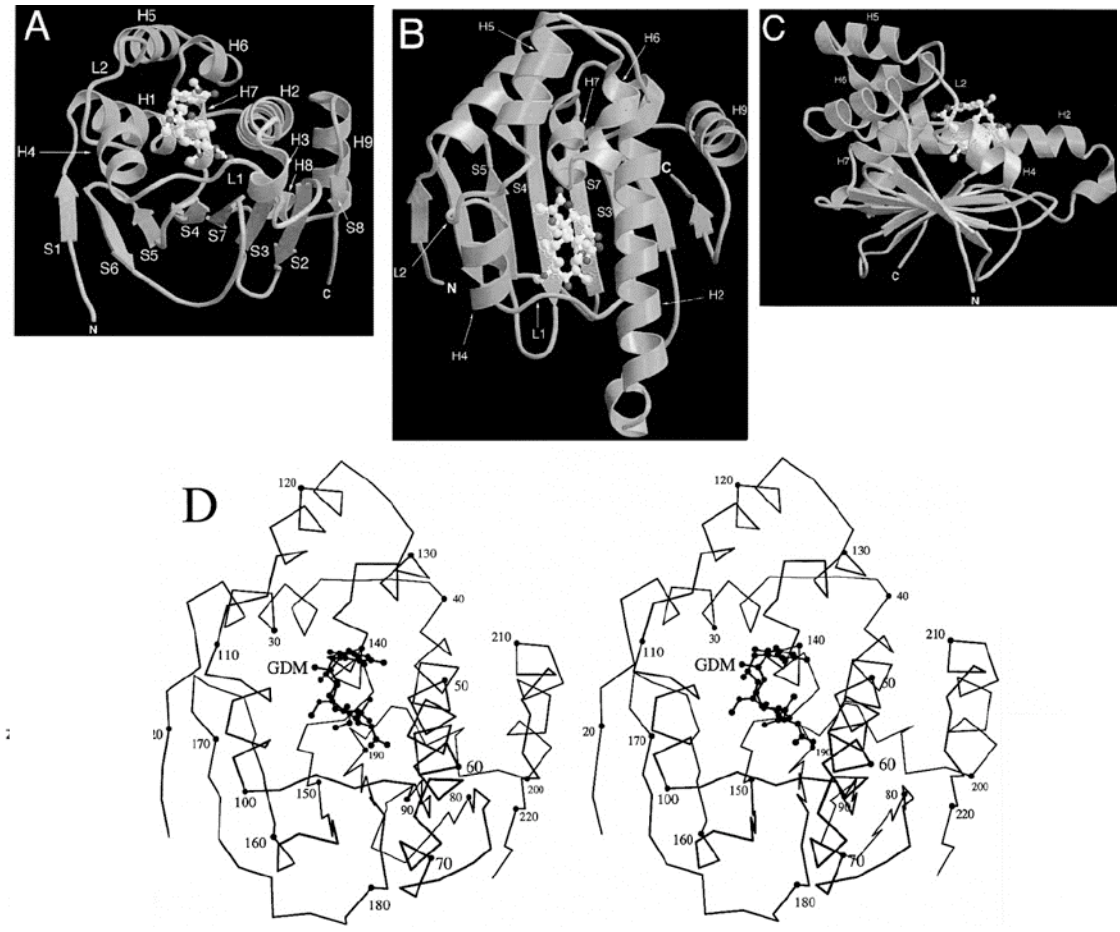


Figure 7.1. The Hsp90-GBD with bound geldanamycin.

(A-C) Three orthogonal views of the protein-drug complex. Carbon atoms are white, and oxygen and nitrogen atoms are in light and dark grey, respectively. **(D)** Stereo view of Cα trace similar to the orientation in (B).

15 Å deep (Figure 7.1, A-D). The pocket has the β sheet as its base, three helices and a loop as its walls, and contains residues highly conserved across species (7.2B). The helical face of the sandwich also has a surface groove that leads into the pocket.

In the complex, the 560 Da geldanamycin adopts a compact conformation and binds inside the pocket, filling all but its deepest portion (Figure 7.1, A-D). The benzoquinone group of geldanamycin binds near the entrance of the pocket, whereas the ansamycin ring inserts into the pocket. There is extensive, though not complete, surface complementarity between geldanamycin and the pocket, and this allows for a high density of van der Waals contacts. Although there are few hydrogen bond contacts (total of five), one pair from the geldanamycin carbamate group (Figure 7.3) can be reasonably described as one of the most important intermolecular interactions in the complex.

The entrance and width of the pocket are likely regulated by a conformational change, as evidenced by structural and positional alterations in three helices and a loop in the two different apo-Hsp90-GBD crystal forms (Chapter Six). One of the conformations, which is essentially identical to that observed in the crystals of the geldanamycin–Hsp90-GBD complex, has a wide enough entrance to allow geldanamycin binding. The other conformation has a narrower entrance, and it would be incompatible with geldanamycin binding.

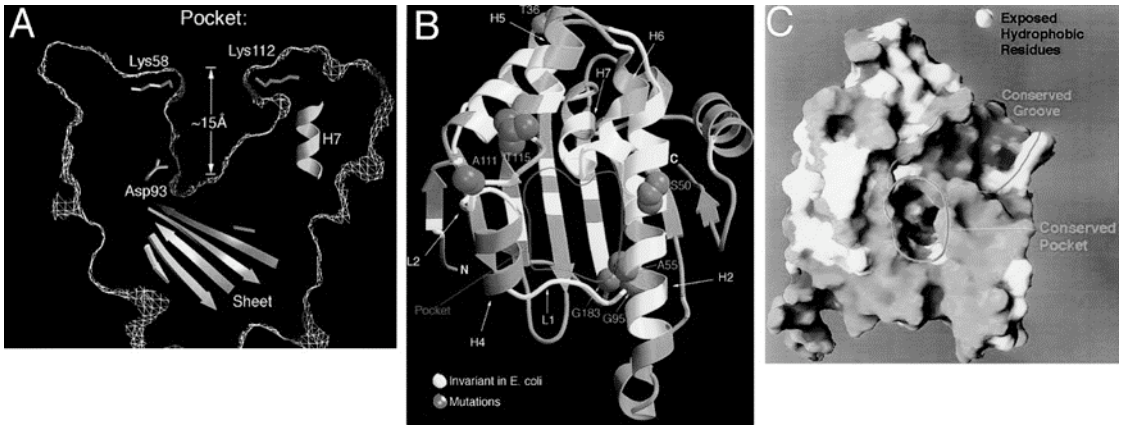


Figure 7.2. The Hsp90-GBD pocket

(A) A thin slice of the Hsp90-GBD molecular surface, depicted as a white net, stressing the pocket dimensions. The approximate locations of the β sheet and of the H7 helix are indicated; the H2 helix would approximately be above the plane of the figure, whereas the H4 helix would be bellow. **(B)** White highlights the positions of residues invariant in the E. coli Hsp90 homologue; dark grey highlights residues, shown in space filling representation, where either temperature sensitive or inactivating mutations map. **(C)** Surface representation indicating the pocket and groove. Also indicated, in white, are several patches of exposed hydrophobic amino acids in the vicinity of the pocket and groove.

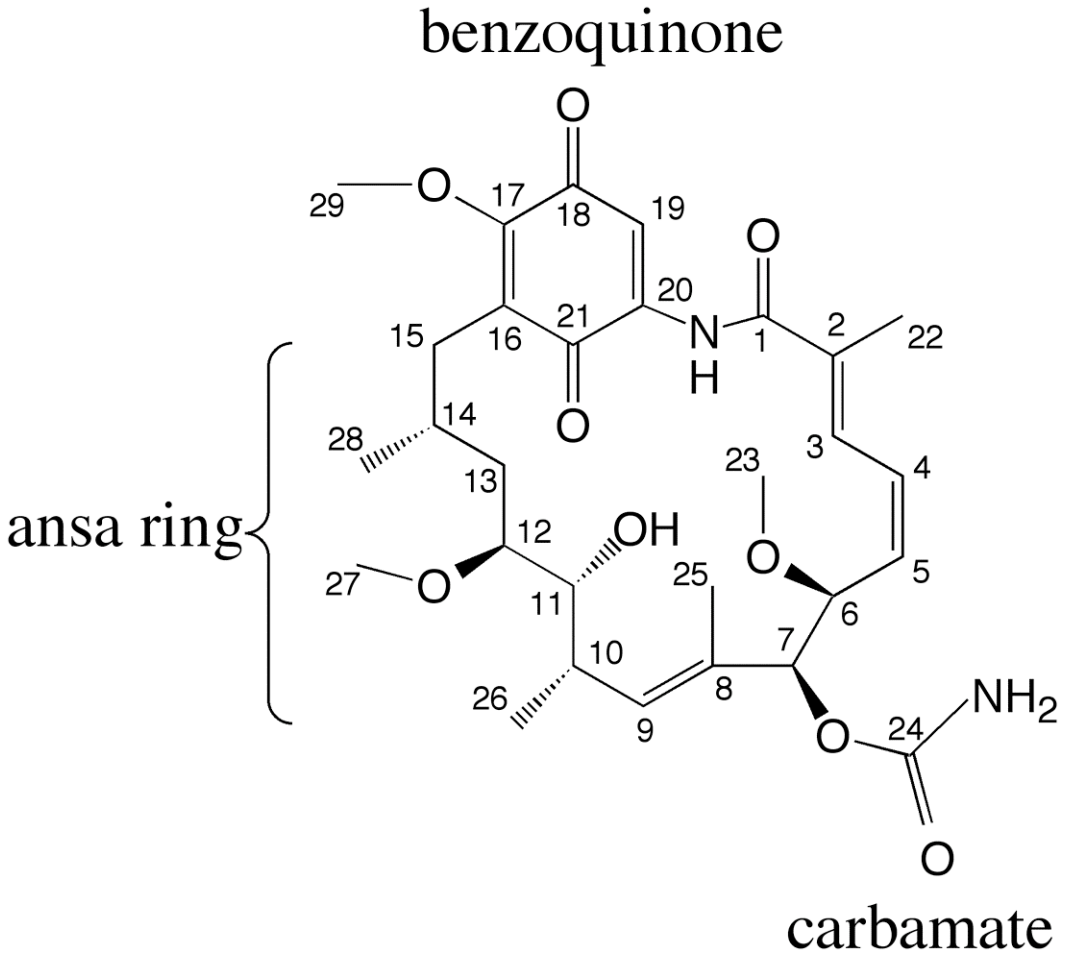


Figure 7.3. The chemical structure of geldanamycin

The numbers indicate the location of carbon atoms. The numbering scheme is chosen to reflect the most common ordering used in the literature.

7.2. Architecture and properties of the drug-binding pocket

The pocket can be described as a flat-bottomed cone: it is about 15 Å deep, 12 Å in diameter near its entrance, 8 Å mid-way down, and wide enough at the bottom to hold three water molecules (Figure 7.2, A-C). The bottom of the pocket is formed by the central portion of the antiparallel β sheet (strands S3, S4, and S7) and the walls by the H2, H4, and H7 helices and the L1 loop. The pocket is deepest by the H7 helix, which packs at an almost perpendicular angle to the β-sheet bottom, and is shallowest by the L1 loop. The pocket is of mixed hydrophobic and polar character, with roughly half of the 17 amino acids lining its interior being hydrophobic, a quarter polar, and a quarter charged (these residues are Leu^{48}, Asn^{51}, Asp^{54}, Ala^{55}, Lys^{58}, Ile^{91}, Asp^{93}, Ile^{96}, Gly^{97}, Met^{98}, Asn^{106}, Leu^{107}, Lys^{112}, Gly^{135}, Phe^{138}, Val^{150}, Thr^{184}, and Val^{186}). Although the pocket becomes increasingly more hydrophobic towards the bottom, it retains one charged and one polar residue at its deepest portion (Asp^{93} and Thr^{184} from the β sheet).

The Hsp90-GBD domain is highly conserved across species, with 43 % of its residues identical in the *Escherichia coli* Hsp90 homologue. The structure reveals that the conserved residues are not distributed homogeneously on the structure, but instead show a striking tendency to cluster in and around the pocket, with 82 % of the residues lining the interior of the pocket being invariant from *E. coli* to humans (Figures 7.2B and 7.4). This identifies the pocket as holding the key to understanding the function of this domain. Of particular interest is Asp^{93}, because it is at the pocket bottom in an otherwise

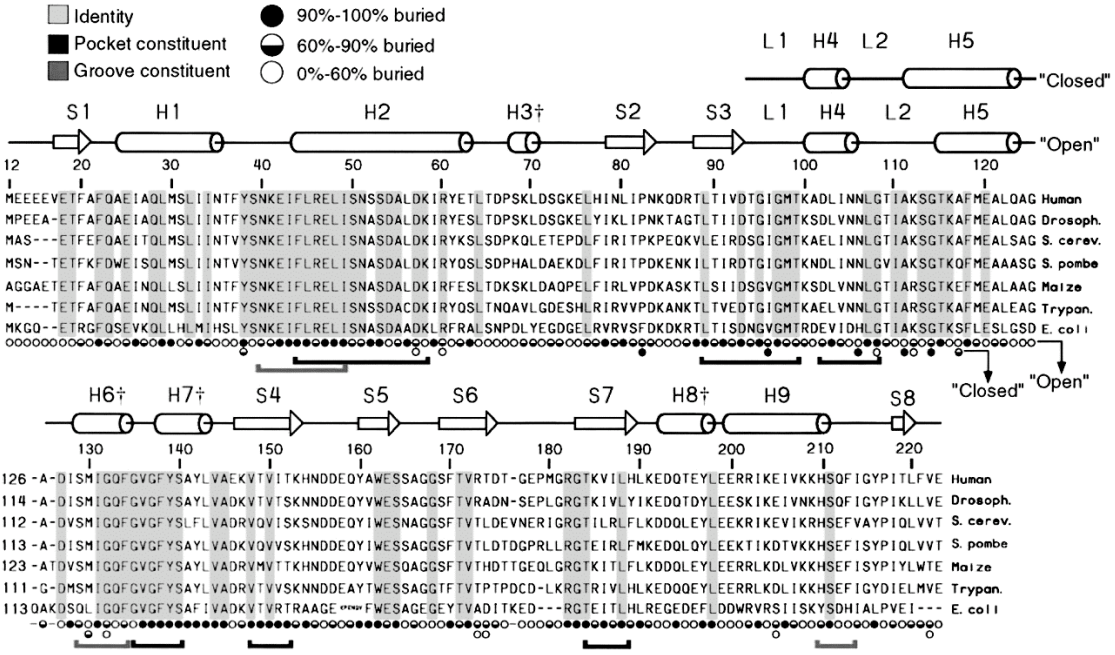


Figure 7.4. Alignment of Hsp90-GBD sequences from several species

Secondary structural elements for the $P2_1$ (open) form are indicated above the sequence together with the changes observed in the I222 (closed) form. Invariant residues are with a grey background, residues which make up the pocket are underlined in black, and those in the groove in dark grey. Residue solvent accessibilities are indicated with circles. The dagger symbol indicates a 3_{10} helix.

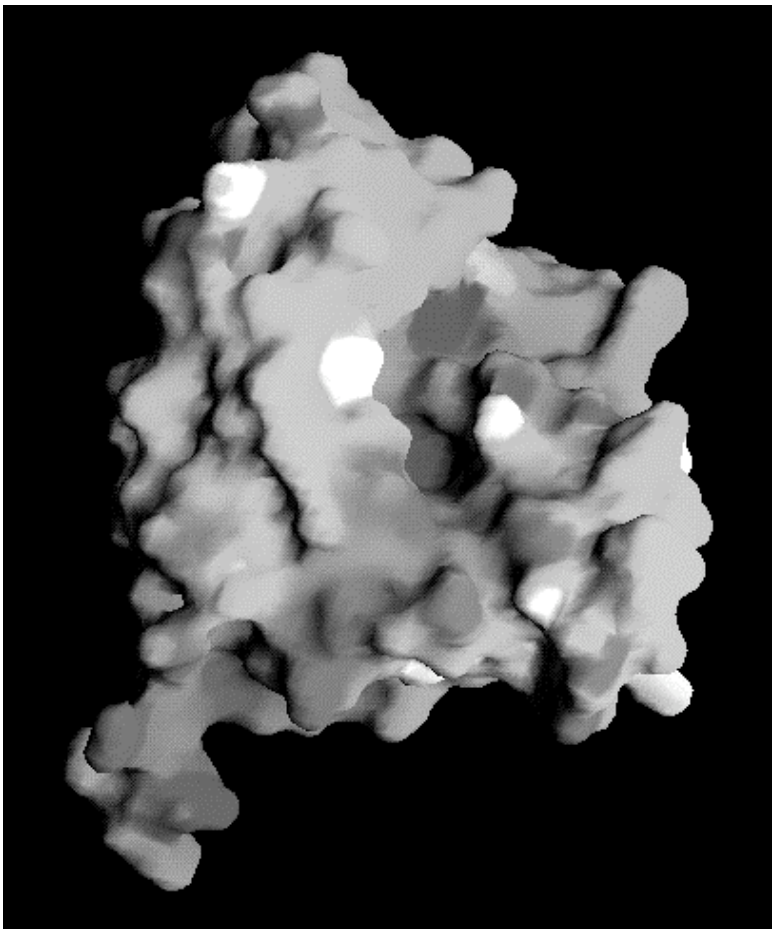


Figure 7.5. Surface electrostatic potential of the Hsp90-GDM

A gradient from dark grey (acidic or negatively charged) to white (basic or positively charged) colors the solvent accessible surface of the Hsp90-GBD centered on the pocket.

mostly hydrophobic environment, and because it is conserved in all known Hsp90 homologues from 35 species.

Supporting a functional significance for this pocket are mutations in the yeast homologue of Hsp90 (Hsp82) that result in either temperature sensitive or reduced activity phenotypes (Sullivan and Toft, 1993; Bohen and Yamamoto, 1993; Kimura et al., 1994; Nathan and Lindquist, 1995), as well as a mutation in the Drosophila Hsp90 (Hsp83) that impairs signaling by the sevenless receptor tyrosine kinase (Cutforth and Rubin, 1994). These mutations map either in the pocket ($Ala^{55}Val$, $Gly^{95}Ser$, and $Gly^{183}Asp$), or in the immediate vicinity of the pocket ($Thr^{36}Asn$, $Ser^{50}Leu$, $Ala^{111}Thr$, and $Thr^{115}Ile$), and the structure suggests that most of these mutations would disrupt the structural integrity of the pocket (Figures 7.2B and 7.4).

The valley, including the pocket, is characterized by a negative surface electrostatic potential (acidic in nature) relative to the surrounding regions (Figure 7.5). This negative potential is due primarily to the amino acids Glu^{47}, Asp^{54}, Asp^{57}, Asp^{93} (deep within the pocket), and Asp^{102}, all absolutely conserved in the alignment shown (Figure 7.4) with the exception of Asp^{102} which becomes glutamic acid in yeast (*pombe*), *trypanosome cruszi*, and *E. coli*. The buried Asp^{93} side chain is located near the bottom of the pocket, and the otherwise hydrophobic/uncharged composition of the rest of the pocket causes the negative potential to dissipate as one moves away from the pocket bottom. The pocket is partially obscured by the positively charged Lys^{58} in the I222 crystal form. By its displacement in the $P2_1$ crystal form, caused by the alternate conformation of the L2 loop, a more acidic and much more extensive opening to the hydrophobic core of the apo-Hsp90-GBD is created.

Leading to the pocket, there is a surface groove formed by the H6 and H9 helices on the sides, and the H2 helix at its bottom (Figure 7.2C). Its shape is not as pronounced as the pocket, being shallower (about 7-8 Å), broader (about 15 Å), and mostly polar. This groove is conserved, although not as well as the pocket, with 73 % of the residues lining its concave surface identical in the *E. coli* homologue (these residues are Asn^{40}, Glu^{42}, Ile^{43}, Arg^{46}, Glu^{47}, Ser^{50}, Ser^{129}, Gln^{133}, His^{210}, Ser^{211} and Ile^{214}; Figure 3). Since many of these residues are polar/charged, without any significant roles in the structural integrity of this domain, their conservation suggests that the groove may participate in intermolecular interactions important for Hsp90 function.

As described in Chapter 2, the rebuilding of the H4-L2-H5-H6 secondary structural elements in the apo-$P2_1$ crystal form reveals that a segment of approximately 35 amino acids, spanning residues 100-134, has two different conformations in the two different apo-Hsp90-GBD crystal forms. These may be termed the "open" and "closed" conformations, because they result in different pocket size and accessibility, with only the open conformation being compatible with the binding of the geldanamycin antitumor drug. Figure 7.4 illustrates the change in secondary structure above a sequence alignment, and Figure 7.6 displays an alignment of the two apo-Hsp90-GBD structures, with geldanamycin as bound in the complex superimposed over the two structures.

At the center of the open-to-closed conformational change (or *from* the $P2_1$ structure *to* I222 structure), is a displacement of the L2 loop into the pocket by more than 9 Å (C_α-C_α distance), whereupon the L2 loop replaces the H4 helix as one of the pocket walls. The motion of the L2 loop is facilitated by portions of the H4 and H5 helices undergoing helix-to-coil and coil-to-helix

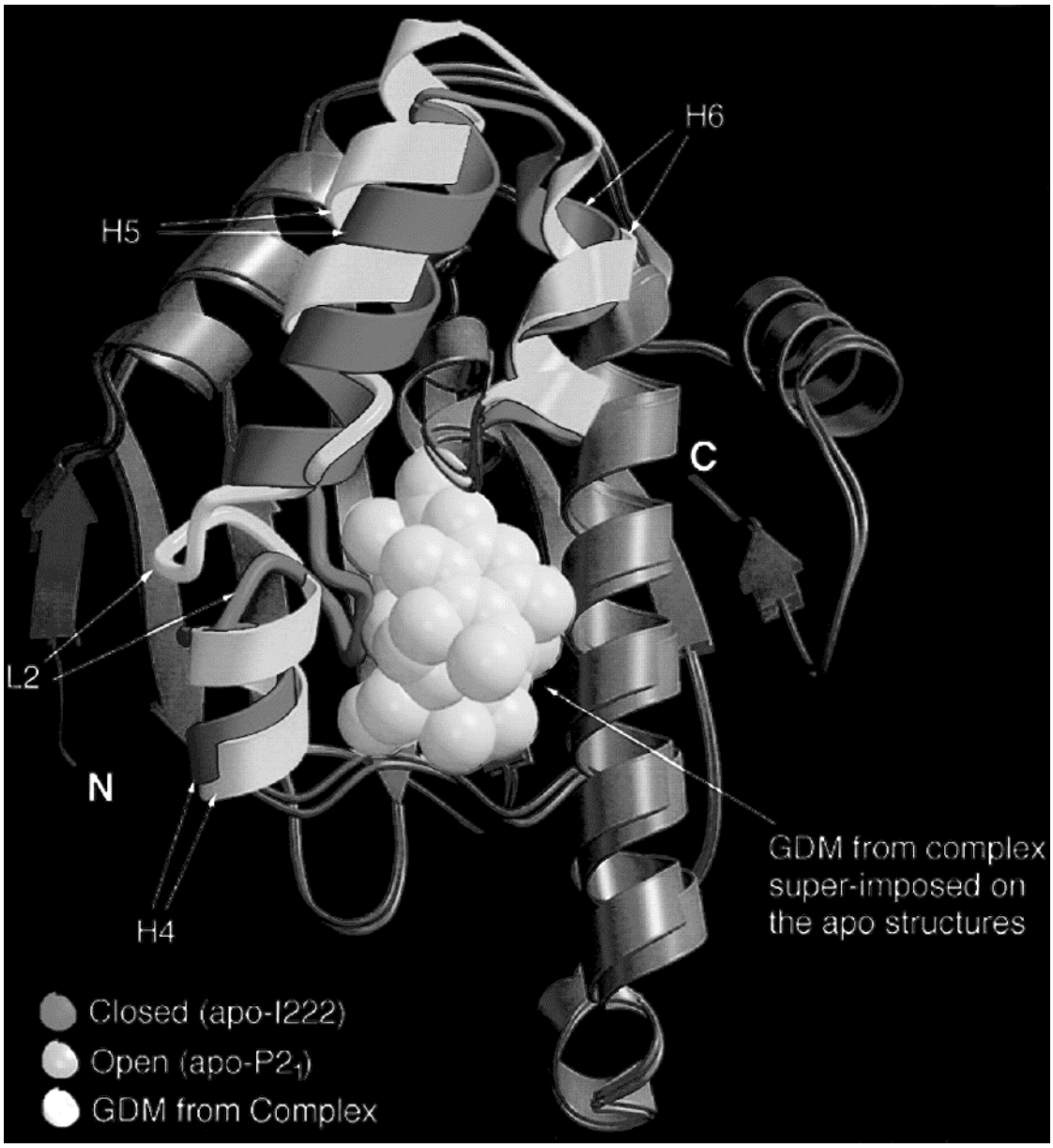


Figure 7.6. Alignment of I222 and P2$_1$ structures
The secondary structural elements that undergo positional and conformational changes are shown in grey for the I222 (closed) and in white for the P2$_1$ (open) apo-Hsp90-GBD crystal forms. Geldanamycin, in space filling representation, is superimposed on the figure to stress the effects the conformational change has on the pocket size.

transitions, respectively (Figures 7.4 and 7.6). The H4 and H5 helices also are displaced, by up to 6 Å. The functional consequence of this conformational change is that the L2 loop acts as a gate that constricts the pocket entrance from a width of 12 Å to 8 Å (Figure 7.6).

7.3. Structure of Hsp90-bound geldanamycin

Geldanamycin consists of a closed ansa ring with a planar benzoquinone embedded in it (Figure 7.3). The ansa ring is sterically hindered because (i) its backbone consists of a planar amide and three carbon-carbon double bonds (two of them arranged in a 1,3 diene), and (ii) of its sixteen backbone atoms, nine carry non-hydrogen substituents: a carbonyl, a carbamate (-OC(O)NH2), a hydroxyl, two methoxy, and four methyl groups (Figure 7.3).

Hsp90-bound geldanamycin is highly compact and internally well packed (overall dimensions of about 9 x 9 x 9 Å). Its ansa ring is folded over the benzoquinone forming a structure which from the side has the shape of a letter “C”, with the benzoquinone forming the top of the "C" and the ansa ring forming the stem and bottom of the "C" (Figure 7.7A). Two of the methyl groups from the ansa ring are centrally positioned to maximize intramolecular van der Waals contacts: the C25 methyl group from the tip of the ansa ring packs with the benzoquinone, thus bridging the top and bottom halves of the "C", and the C28 methyl group packs with the diene carbon atoms, bridging the two sides of the ansa ring (Figure 7.7A).

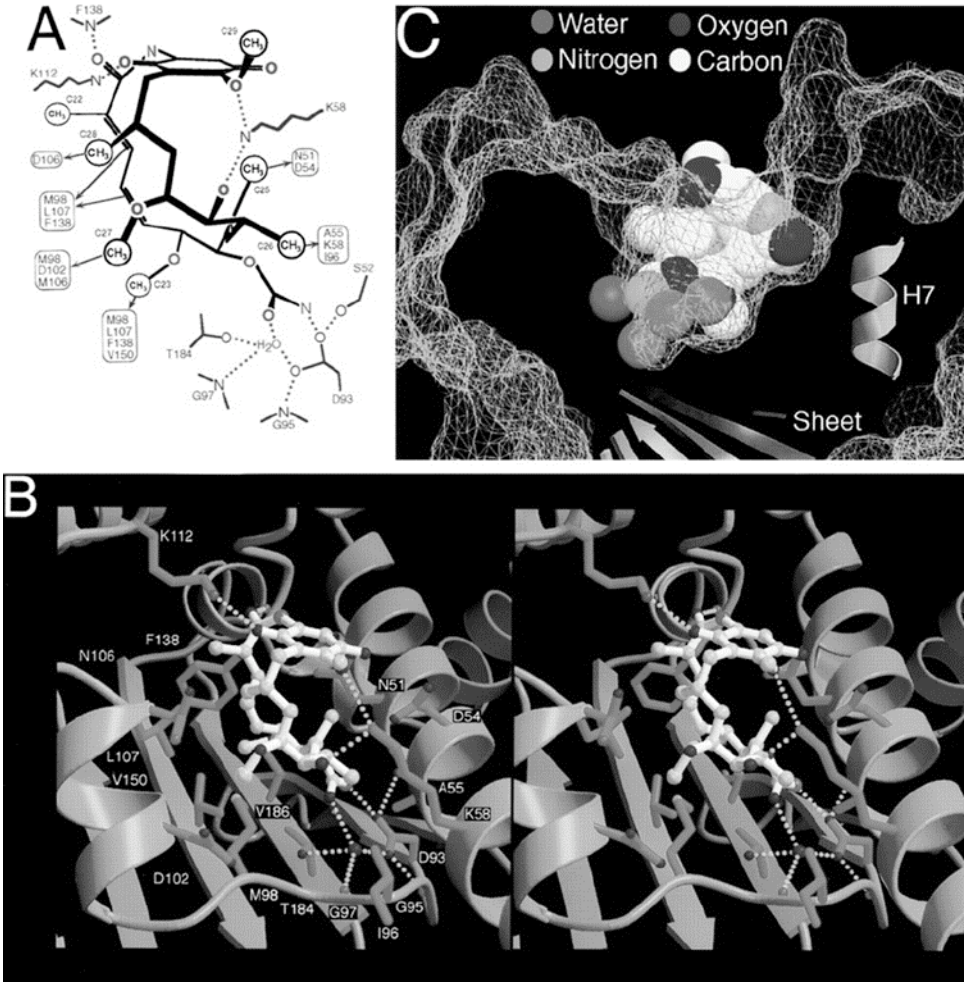


Figure 7.7. Hsp90-geldanamycin interactions

(A) Bound geldanamycin conformation summarizing the hydrogen bond (dotted lines) and van der Waals (arrows) contacts it makes to Hsp90. **(B)** Stereo view of the geldanamycin–Hsp90-GBD interactions. Hydrogen bonds are indicated by dotted lines, and the water molecule that bridges the geldanamycin carbamate group with Asp93 is shown as a dark sphere. **(C)** Geldanamycin, in space filling representation, adopts a structure that is overall complementary to the pocket, shown as a molecular surface net.

This compact conformation differs from those observed with free geldanamycin (Rinehart and Shield, 1976) or herbimycin (Furusaki et al., 1980), which crystallize from organic solvents in significantly more open, extended conformations. The differences are likely due, in part, to the loss of co-planarity between the amide group and the benzoquinone in Hsp90-bound geldanamycin, as well as to the differences in the crystallization solvents.

7.4. Geldanamycin–Hsp90 contacts

GDM sits in the pocket of Hsp90 in a very compact form; the benzoquinone group is relatively solvent exposed at the pocket entrance, whereas the bottom of the ansa ring is buried deep in the pocket (Figure 7.7). Most of the Hsp90-GDM interactions are van der Waals in nature, occurring along the sides of the ansa ring, extending from the benzoquinone down toward the carbamate. The tip of the geldanamycin ansa ring (bottom of "C"), which has the carbamate, C23 methoxy, and C25 and C26 methyl group substituents binds near the bottom of the Hsp90 pocket and makes a high density of van der Waals contacts, whereas the benzoquinone (top of "C") is positioned near the entrance of the pocket and makes only a few contacts (Figure 7.7). There is extensive surface complementarity between the compact geldanamycin structure and the pocket, as attested by the fact that a remarkable 85 % of the surface area of geldanamycin is buried in the complex (609 $Å^2$ out of 719 $Å^2$).

The most critical portion of the interface is likely to be a hydrogen bond network between the geldanamycin carbamate group and the Asp^{93} side chain from Hsp90, which are buried at the pocket bottom in an otherwise

mostly hydrophobic environment. The carbamate amino group makes a hydrogen bond to one of the Asp^{93} side chain oxygen atoms, and its carbonyl oxygen atom makes a water-bridged hydrogen bond with the other Asp^{93} side chain oxygen atom (Figure 7.7). The Asp^{93} side chain, which has an identical conformation in the apo-Hsp90 structures, also hydrogen bonds to the Ser^{52} OH and the backbone amide of residue 95. This hydrogen bond network is buried about 10 Å away from the nearest solvent surface, and the low dielectric constant of this environment would significantly increase the coulombic attraction due to the partial charges. The importance of this hydrogen bond network is supported by structure-activity studies demonstrating that the removal of the carbamate group, or the attachment of additional atoms to it, completely abolished geldanamycin activity (Schnur et al., 1995b).

At the tip of the ansa ring, the carbamate group contact is flanked by the C23 methoxy and C25 and C26 methyl groups making high density van der Waals contacts with the bottom and sides of the pocket (Figures 6A and 6B). Although the surface complementarity at the pocket bottom is extensive, there remains a buried cavity filled with three water molecules near the C23 methoxy and the carbamate groups (Figure 7.7C). It is conceivable that the cellular substrates of Hsp90 can fill this cavity as well. Modifications to this portion of geldanamycin, for example by adding to the C23 position a group of 4-6 non-hydrogen atoms could increase the steric and hydrogen bond complementarity and may yield derivatives with increased affinity for Hsp90.

Halfway between the pocket bottom and the surface, the widening of the pocket causes a decrease in the complementarity and contact density, although the fit is still adequate to exclude bulk solvent (Figure 7.7C). In this

region, the diene and the C27 methoxy groups make van der Waals contacts, and the O5 hydroxyl group makes a hydrogen bond contact with Lys^{58} (Figures 7.7A and 7.7B). As we get closer to the surface, the contact density decreases further. In this region, the C28 methyl group appears to be in an overall unfavorable environment, being within van der Waals contact distance to a backbone carbonyl oxygen atom and a partially buried water molecule (Figures 7.7A and 7.7B). The loose fit in this region, both sterically and electrostatically, points to a potential for a significant improvement of the affinity of geldanamycin for Hsp90 (Figure 7.7C).

At the pocket entrance, the carbonyl oxygen atom of the ansa ring makes a hydrogen bond to a backbone amide (Phe^{138}) at the N-terminal portion of the H7 helix (Figures 7.7A and 7.7B). The benzoquinone, which is at the pocket entrance with one face solvent-exposed, makes only a few contacts, and these are primarily solvent-exposed hydrogen bonds made to long lysine chains: one of the benzoquinone oxygen atoms hydrogen bonds with Lys^{112}, and the C29 methoxy group oxygen atom makes a long-distance (3.5 Å) hydrogen bond with Lys^{58}.

CHAPTER EIGHT:

ANALYSIS OF THE HSP90-GELDANAMYCIN STRUCTURES

8.1. Structures of Hsp90 homologs

Shortly following the publication of our Hsp90 structures (Stebbins et al., 1997), the laboratory of Dr. Lawrence Pearl at the University College of London published the crystal structure of the yeast homolog of the Hsp90-GBD (Prodromou et al., 1997a; Prodromou et al., 1997b). This structure proved nearly identical to the human protein (as is to be expected from the high degree of sequence conservation) with an identical topology and a root mean square deviation of 4Å over 6 strands and 5 helices. The published structural results contained no mention of the pocket structure, but instead focused on the hydrophillic side of the β sheet, which in their crystals, packs against the same side of the sheet from another yeast Hsp90 molecule, creating a dimer in the asymmetric unit. It is hypothesized by the authors that this dimer interface, which in the crystals forms an extensive 'tunnel' , is a site of peptide binding from the client protein. Arguments based on the dimerization of full-length Hsp90 are used to support this position (Prodromou et al., 1997a). Finally, two additional crystal structures revealed this domain to be an ATP-binding module (Prodromou et al., 1997b and Obermann et al., 1998).

In 1991 the structure of DNA gyrase B, a bacterial topoisomerase II, was reported (Wigley et al., 1991). The coordinates for this structure had not been

deposited in the Brookhaven Protein Data Bank, so that the computational methods commonly used to identify structural homology could not detect it. However, Dr. Don Wigley informed us subsequent to our publication of the Hsp90-geldanamycin structure of this homology.

In the two years following the publication of the first Hsp90 structure, two additional structures of Hsp90 homologs were reported. The first was the structure of MutL, a protein found in species ranging from bacteria to mammals which is essential for DNA mismatch repair (Modrich and Lahue, 1996). Mutation of the two human MutL homologs (MLH1 and MSH2) account for 90% of reported hereditary nonpolyposis colorectal cancer (Peltomäki and Vasen, 1997), and mutation of MLH1 has been identified in human leukemia and lymphoma cell lines (Hangaishi et al., 1997). All MutL family members possess a conserved, approximately 300 amino acid N-terminal domain which is required for their activity. The majority of inactivating mutations in MutL map to this domain (Aronshtam and Marinus, 1996), and the crystal structure of this domain reveals that it shares strong homology with the N-terminal geldanamycin binding domain of Hsp90, including extensive conservation of the pocket architecture (Ban and Yang, 1998). The second structure was that of CheA, a histidine kinase which functions in signal transduction pathways in bacteria, plants, and fungi (Parkinson and Kofoid, 1992; Appleby et al., 1996). The structure of CheA reveals that the kinase domain is structurally similar to Hsp90, aligning well with the four β strands and three α helices which are the primary components of the pocket structure (Bilwes et al., 1999).

It is interesting to note here, and will be discussed in the next section, that all of the Hsp90 homologs - DNA gyrase B, MutL, and CheA – three

classes of very distinct function (a topoisomerase, a DNA repair enzyme, and a histidine kinase), all bind and hydrolyze ATP using the domain homologous with Hsp90, and in fact bind the nucleoside triphosphate in the equivalent of the Hsp90 pocket (which is the most conserved region between these different proteins). This strongly confirms the ATP binding function of Hsp90, and indicates that this domain is an ancient ATP binding motif which has become adapted as a modular unit conferring hydrolysis activity to a disparate class of molecules.

8.2. The pocket as an ATP binding site

The pocket of the Hsp90-GBD is has classical features of a substrate binding site, for there are extensive similarities between the pocket and a typical enzyme active site, as noted in the following: (1) The pocket has adequate space and hydrophobic content, and is adequately shielded from bulk solvent, to tightly bind substrates of a few hundred daltons, (2) there is a non-uniform electrostatic distribution resulting from an aspartic acid at the otherwise hydrophobic bottom of the pocket, (3) the shape and size of the pocket entrance may be regulated by a conformational change, (4) the geldanamycin-binding domain of Hsp90 is the best conserved among its three structural domains, and within this domain, the conserved residues cluster in and around the pocket, (5) mutations that inactivate Hsp90 map in and around the pocket, and, finally, (6) the geldanamycin inhibitor of Hsp90 function binds inside this pocket.

The size and composition of the pocket immediately suggested to us when we determined the apo-Hsp90-GBD crystal structure that it could

function to bind small molecule compounds, and given the role of ATP in many other refolding reactions catalyzed by molecular chaperones, the binding of a nucleotide triphosphate was an intriguing possibility. However, at the time of the structural determination, the biochemical data concerning ATP binding with regard to Hsp90 was inconsistent. Many studies concluded that Hsp90 was an ATP independent chaperone. These include: (1) that Hsp90 purified from different sources all lacked ATPase activity (Wiech et al., 1993; Scheibel et al., 1997), (2) that the ATPase activity detected in some experiments was attributable to contaminating kinases (Shi et al., 1994; Wearsch and Nicchitta, 1996), (3) that Hsp90 possessed chaperone activity *in vitro* in the absence of ATP (Wiech et al., 1992), and, in what many considered to be the definitive experiments, (4) that Hsp90 was unable to be affinity labeled with azido-ATP, could not be retained on an ATP agarose resin, and could not enhance the fluroescence of an ADP derivative, whereas the chaperone Hsp70, a known ATPase, could accomplish each of these feats (Jakob et al., 1996). On the affirmative side of ATP involvement, studies include the following: (1) that Hsp90 or Hsp90 homologs were highly active ATPases (Nadeau et al., 1992, 1993) or in some studies, GTPases (Nardai et al., 1996) with autophosphorylating activity (Csermely and Kahn, 1991 ; Csermely et al., 1995), (2) that Hsp90 undergoes a conformational change in the presence of ATP (Csermely et al., 1993; Sullivan et al., 1997), and, (3) that the interaction of Hsp90 with co-chaperones is ATP dependent (Johnson and Toft, 1994). Despite the contradictory data, it had become the consensus in the chaperone literature that Hsp90 was indeed an ATP-independent chaperone (Buchner, 1996; Schneider et al., 1996; Panaretou et al., 1998).

In line with such reasoning, it was proposed that the pocket bound a portion of the polypeptide substrate to be refolded (Stebbins et al., 1997). In this model, the pocket-polypeptide association is likely to occur as a late step in the reaction pathway because Hsp70, which is present in all the early Hsp90 complexes (Smith, 1993; Smith et al., 1995; Dittmar et al., 1996; Johnson et al., 1996), can bind unfolded proteins (Flynn et al., 1991; Zhu et al., 1996) and is thus the likely initial recruiter of the unfolded polypeptide to the Hsp90 complex. This model was not completely satisfactory, however, as it was hard to envision the mechanism by which such a binding could be incorporated into the refolding process. Most importantly, this model went against our structural intuition to some extent, which would pigeonhole the pocket as characteristic of an enzyme acting on a small molecule. Thus, although unsatisfactory, these conclusions seemed to best reflect the available biochemical data.

Shortly after, several reports predicted that the Hsp90-GBD shared a similar fold and crucial binding residues with an ATP binding motif present in several topoisomerase type II proteins (Bergerat et al., 1997; Gerloff, et al., 1997; Dunbrack, et al., 1997; Mushegian et al., 1997), including the bacterial DNA gyrase B (Wigley et al., 1991). The residues of the motif align well with some of the most conserved residues in the Hsp90-GBD, and when the structure is examined, are revealed to line the pocket interior and surround the inhibitor of Hsp90 function, geldanamycin. Supporting these predictions, crystal structures of the yeast Hsp90 homolog (Prodromou et al., 1997b) and human Hsp90 (Obermann et al., 1998), both with bound Mg^{+2}-ADP, were published, revealing a strong structural homology in the binding of the nucleotide to DNA gyrase B.

In these structures of Hsp90, the L2 loop framing the pocket is in the open conformation, with a slight displacement averaging about 1Å over the length of this segment (residues 106-124). In these Hsp90-ATP complexes, the nucleotide triphosphate is found inserted into the pocket in which geldanamycin binds. ATP inserts such that the adenine ring is at the bottom of the pocket and the phosphates are at the pocket entrance, with the γ-phosphate not visible in any of the high resolution (beyond 2 Å) crystal structures, either because it is inherently flexible or because it was hydrolyzed. The purine occupies the pocket bottom, in a manner analogous to the bottom of the ansa ring of geldanamycin, with the N6 group forming direct hydrogen bonds with Asp^{93} (Asp^{79} in yeast Hsp83).and water-mediated hydrogen bonds with the Ser^{52} and Leu^{48} backbone carbonyl oxygens (Prodromou et al., 1997b ; Obermann et al., 1998). In a manner extremely similar to geldanamycin, the N7 group of the purine forms water-mediated hydrogen-bonds with the Asn^{51} side chain, and the N1 group with the Gly^{97} backbone and Asp^{93} and Thr^{81} side chain groups (Prodromou et al., 1997b ; Obermann et al., 1998). These results strongly suggest that geldanamycin is acting as an ATP mimic.

Support for this hypothesis is found in structure-based mutagenesis experiments in which it is found that: (1) Asp^{93} is required for ATP binding and hydrolysis in Hsp90, (2) Glu^{47} (located at the entrance of the pocket in a position where the γ-phosphate would be, were it visible in the density, proximal to the bound Mg^{+2} ion) is required for the hydrolysis of ATP, and, (3) neither mutations in Asp^{93} nor Glu^{47} are viable in yeast which have been depleted of redundant Hsp90 family members (Obermann et al., 1998). This same study found in addition the interesting result that the hydrolysis of ATP

does not occur in the N-terminal GBD fragment alone, but is dependent on the presence of full length Hsp90. This raises the possibility that the ATP hydrolysis activity of Hsp90, like that of DNA gyrase B (Wigley et al., 1991), is dependent on the interactions of residues from a second domain of the ATPase with the ATP binding domain, creating, in effect, a catalytic structure which is composed of two domains. Moreover, the interaction of the p23 co-factor to Hsp90 follows an ATP hydrolysis-dependent reaction cycle (Obermann et al., 1998).

Current models for Hsp90 function typically incorporate a multi-step reaction mechanism in which: (1) the nucleotide free or ADP bound form of Hsp90 binds partially unfolded substrates (through two binding sites at the N and C-terminal domains (Scheibel et al., 1997)) (2) this substrate bound form then binds ATP in the N-terminal domain pocket, inducing a conformational change in Hsp90, (3) this creates a short lived transition state which has much decreased affinity for the substrate, and finally, (4) the transition state releases the substrate almost immediately, after which Hsp90 hydrolyzes ATP. Geldanamycin may interfere with this process by blocking the binding of ATP and preventing the release of substrate, or by displacing ATP after substrate is bound, which may lead to retention or release of substrate depending on the nature of the transition state. Consistent with this model, it has been observed that the binding of geldanamycin does in some cases stabilize Hsp90-client complexes (Chen et al., 1997; Dasgupta and Momand, 1997). Since geldanamycin binds Hsp90 in an manner closely mimicking ATP, it is almost certain that the cellular effects of this compound can be understood by its inhibition of the ATPase cycle of Hsp90, and that this cycle is an integral

component of the maturation/refolding reaction catalyzed by this heat shock protein.

Prodromou *et al.* (1997b) provide a structural explanation for the discrepancies in the literature concerning the binding of ATP to Hsp90, the ideas of which were also articulated by Grenert *et al.* (1997). The primary experiments which had convinced many in the field that Hsp90 was ATP-independent were those of Ursula Jakob (1996) in which it was demonstrated, for the first time with highly purified recombinant Hsp90, that this chaperone, in contrast to Hsp70, could not be retained on an ATP agarose resin, and could not enhance the fluroescence of an ADP derivative. However, these experiments were founded upon a subtle assumption, namely that the binding of ATP to Hsp90 did not occur primarily through its adenine moiety. The ATP in the experiments of Jakob and others which failed to retain Hsp90 was conjugated to the agarose resin by the ribose or adenine constituents (Jakob et al., 1996). Since these elements are deeply buried in the Hsp90-GBD pocket, it is clear in retrospect why these experiments failed. This binding geometry of ATP also would explain why Hsp90 could not be detected with fluorescent or photoaffinity labeling analogs. In contrast, the crystal structure of Hsp70 bound to ATP (Flaherty et al., 1994), the ATP is positioned in a substantially more exposed orientation. The 8 position of the adenine (buried in the pocket of Hsp90) is unhindered, allowing the 8-substituted analogs to retain this molecule. This set of "idiosyncratic" (Prodromou et al., 1997b) interactions between ATP and Hsp90 explain why geldanamycin targets this ATP binding protein specifically over the plethora of ATP binding proteins in the cell.

8.3. Implications of conformational change

As mentioned in Chapter Six, there is a conformation change observed between two different apo-Hsp90 crystal forms. It is not clear what causes the conformational change *in vitro*; the L2 loop, which changes the most, is not involved in any significant crystal packing contacts, though the H5 and H6 helices are in both crystal forms. In addition, the $P2_1$ crystal form can be induced by the presence of geldanamycin and by acidic (pH 4.6) growth conditions (although the crystals in the latter conditions suffer from mild to severe twinning problems). Both of the alternate structures are associated with well-packed side chain arrangements. Because the conformational change is relatively free of crystal packing effects, and is associated with internally well packed side chain orientations, one may propose with some confidence that this conformational change is not an artifact of crystallization, but is a *bona fide in vivo* occurrence fortuitously captured in crystallization. Supporting an *in vivo* role for the closed conformation is the $Ala^{111}Thr$ mutation on the L2 loop that results in a temperature sensitive phenotype in yeast (Kimura et al., 1994). This mutation would disrupt only the closed conformation structure where Ala^{111} is fully buried and participates in the hydrophobic packing of the loop, but not the open conformation structure where Ala^{111} is fully solvent-exposed with no apparent structure-stabilizing role.

The similarities of the pocket to a substrate binding site, discussed above, invite similar speculation, as many enzymes are regulated by such means. In particular, the conformational change observed in our crystals may be part of the ATPase cycle, regulating the binding and release of client

polypeptides and/or co-chaperones in the larger multi-chaperone complex of which Hsp90 is only one member.

8.4. The prospects for rational drug design

Among the naturally occurring homologues, geldanamycin is the most potent. Compared to herbimycin, for example, geldanamycin is more potent by a factor of about four in achieving a 50 % reduction (IC_{50}) of erbB-2 kinase activity in human SKBr3 breast cancer cells (Miller et al., 1994). The structure suggests that this is due to the hydrogen bonds that the O5 hydroxyl and C29 methoxy oxygen atoms of geldanamycin make with Lys58, because herbimycin does not have the C29 methoxy group.

Among the published synthetic derivatives (Schnur et al., 1995; Schnur et al., 1995b), the highest potency, reflecting an improvement in the IC_{50} of a factor of 4-5, was achieved by an amino substitution at the C17 position of the benzoquinone (Schnur et al., 1995). It was proposed that this and several related 17-amino substitutions improved cellular activity indirectly, by stabilizing the quinone form over the reduced hydroquinone (Schnur et al., 1995). The structure supports this hypothesis as this position is highly solvent exposed in the complex and is a poor candidate for additional Hsp90 contacts.

Overall, however, most modifications decreased or eliminated activity (Schnur et al., 1995; Schnur et al., 1995b). The maintenance of the carbamate group and of the closed, cyclic nature of the ansa ring proved necessary for activity (Schnur et al., 1995b), consistent with the extended hydrogen bond network that the carbamate group makes, and with the importance of a closed cyclic ring in limiting conformational flexibility, respectively. Modifications at

other positions either had little effect or reduced the activity of geldanamycin in line with space considerations revealed by the crystal structure. For example, additions of small groups to the amide nitrogen and to the C19 position of the benzoquinone, which are juxtaposed in a partially solvent-exposed region of the complex, were tolerated (Schnur et al., 1995; Schnur et al., 1995b), whereas bulky substituents at these positions reduced or eliminated activity (Schnur et al., 1995b).

The crystal structure of the Hsp90-GBD complex with geldanamycin suggests modifications which could improve the steric and hydrogen bond complementarity between geldanamycin and Hsp90 and could thus provide analogues with increased Hsp90 affinity (a total synthesis of geldanamycin has recently been reported (Miller, 1995)). An example is the buried cavity filled with three water molecules near the C23 methoxy and the carbamate groups discussed in Chapter 11. Modifications to this portion of geldanamycin, for example by adding to the C23 position a group of 4-6 non-hydrogen atoms could increase the steric and hydrogen bond complementarity and may yield derivatives with increased affinity for Hsp90. Another example is the unfavorable environment about the C28 methyl group. The loose fit in this region, both sterically and electrostatically, points to a potential for a significant improvement of the affinity of geldanamycin for Hsp90. In addition, the work of others identifying the pocket as a site of ATP binding will allow drug design to proceed in the context of geldanamycin's role as an ATP mimic, opening a new avenue for the development of novel anti-chaperone compounds.

APPENDIX A:

PROTOCOLS

A.1. Initial Cloning of VHL, ElonginB, and ElonginC

The genes for VHL, ElonginB, and ElonginC were kindly provided in the SP2 vector (Strategene) by Dr. William Kaelin at the Dana Farber Cancer Institute of Harvard Medical School. For the initial refolding and subtilisin digestion experiments, these proteins were cloned by PCR into pET28b (Novagen, T7 promotor) with a hexahistidine tag fused to their N-terminus. The oligonucleotides for the three genes are listed below.

VHL(1-213) 5′ oligonucleotide

5 ′GCCGTGAC TCA GCTAGC ATG CCC CGG AGG GCG GAG AAC TGG **3′**
CLAMP Nhe I **M P R R A E N W**

VHL(1-213) 3′ oligonucleotide

5 ′ GCTAGGCATGTC GGATCC CTA ATC TCC CAT CCG TTG ATG TGC **3′**
CLAMP BamHI STOP **D G M R Q H A**

ElonginB 5′ oligonucleotide

5 ′ GCCGTGAC TCA CATATG GAC GTG TTT CTC ATG ATC CGG **3′**
CLAMP Nde I **D V F L M I R**

ElonginB 3′ oligonucleotide

5 ′ GCTAGGCATGTC GGATCC CTA CTG CAC AGC TTG TTC ATT GGC **3′**
CLAMP BamHI STOP **Q V A Q E N A**

ElonginC 5′ oligonucleotide

5′ GCCGTGAC TCA CAT**ATG GAT** GGA GAG GAG AAA ACC TAT **3′**
CLAMP Nde I **D G E E K T Y**

ElonginC 3′ oligonucleotide

5′ GCTAGGCATGTCGGATCC CTA ACA ATC TAA GAA GTT CGC AGC **3′**
CLAMP BamHI STOP **C D L F N A A**

The PCR protocol (100μl reactions) was the following:

- 20-150 pmoles of primer
- 0.1 mM dNTP's
- 1xPFU polymerase buffer (from 10X stock)
- 40 ng plasmid containing template
- 1μL PFU polymerase
- H_2O until 100 μL

A typical amplification protocol would consist of the following:

A-cycle: 94°C 8 minutes
58°C 1 minutes **x1**
72°C 2 minutes

B-cycle: 94°C 30 seconds
50°C 30 seconds **x3**
72°C 1 minutes

C-cycle: 92°C 1 minutes
62°C 1 minutes **x21**
72°C 2 minutes

D-hold: 72°C 5 minutes

Annealing temperatures would vary depending on the template, although the optimal temperature rarely corresponded to the calculated melting temperature and were determined by trial and error.

After the VHL structural domain had been identified by limited proteolysis, VHL(54-213 was cloned by PCR into pET28b as a hexahistidine

fusion (the alternate VHL construct did not express well without a tag), and the two Elongin proteins were cloned by PCR into a non-tagged construct by ligation into pET-3d (Novagen, T7 promotor). The PCR protocol was similar to that used in the initial cloning. The primers are described below.

VHL(54-213) 5′ oligonucleotide

5 ′GCCGTGAC TCA GCTAGC ATG CCC CGG AGG GCG GAG AAC TGG **3′**
CLAMP Nhe I **M P R R A E N W**

VHL(1-213) 3′ oligonucleotide

5 ′ GCTAGGCATGTC GGATCC CTA ATC TCC CAT CCG TTG ATG TGC **3′**
CLAMP BamHI STOP **D G M R Q H A**

ElonginB 5′ oligonucleotide

5 ′ GCCGTGAC TCA CC**ATGG** AC GTG TTT CTC ATG ATC CGG **3′**
CLAMP Nco I **D V F L M I R**

ElonginB 3′ oligonucleotide

5 ′ GCTAGGCATGTC GGATCC CTA CTG CAC AGC TTG TTC ATT GGC **3′**
CLAMP BamHI STOP **Q V A Q E N A**

ElonginC 5′ oligonucleotide

5 ′ GCCGTGAC TCA CC**ATGG AT** GGA GAG GAG AAA ACC TAT **3′**
CLAMP Nco I **D G E E K T Y**

ElonginC 3′ oligonucleotide

5 ′ GCTAGGCATGTCGGATCC CTA ACA ATC TAA GAA GTT CGC AGC **3′**
CLAMP BamHI STOP **C D L F N A A**

Following the PCR, the DNA was purified by standard phenol-chloroform precipitation. Briefly, a 1:1 ratio of phenol/$CHCl_3$ was added in a

1:1 volume ratio to the PCR reaction. The mixture was vortexed, centrifuged on at 14,000rpm in an Eppendorf microfuge for five minutes at room temperature and the top layer (the aqueous phase) saved. This was washed with $CHCl_3$ to remove residual phenol by a 1:1 volume addition, vortexing, centrifugation at 14,000rpm in an eppendorf microfuge, and the top layer saved. This solution was brought to 0.3M sodium acetate (add to that 1/10 volume 3M sodium acetate pH 5.5), after which 2 volumes of 100% ethanol was added and the solution transferred to ice for 15 minutes. This was then vortexed at 14,000rpm in an eppendorf microfuge for 15 minutes at 4°C, the liquid removed, the pellet dried and resuspended in water.

Restriction enzyme digests were routinely performed as double digests in a buffer compatible with both nucleases, and overcutting (enzyme units to μg of DNA on the order of 20:1) was typically employed with the limit that the solution glycerol concentration not exceed 5%. Digests were purified on 0.8-1.2% agarose gels. The digest would be run out on the gel and the band excised and purified by a protocol from Qiagen (Qiaquick Gel Extraction) involving the dissolving of the agarose and binding of the DNA to a proprietary resin contained in small spin columns. Ligation reactions were typically performed as follows:

- 50ng vector
- 3-100 fold molar excess insert
- Heat to 55°C for 2', slow cool on bench, then on ice
- 1x ligase buffer (typically New England Biolabs)
- 1μL ligase
- H_2O until 20μL

The ligations were incubated for 1-24 hours at 16°C. Variations to improve yield were at times carried out, but in most cases the cloning was

without complications. Bacteria were transformed by the following protocol: 1) 5 μL of competent BL21(DE3) cells prepared by standard protocols (except without the addition of $CaCl_2$ before freezing) were added to a pre-chilled tube containing 50μL of 0.1M $CaCl_2$, 2) no more than 50ng of DNA from the ligation reaction are added to this mixture and left on ice for 30 minutes, 3) bacteria were then subjected to a 42°C heat shock for approximately one minute, and then cooled one ice for one minute, 4) 200μL of LB (no antibiotic) is added and the transformation placed at 37 °C for 30 minutes, 5) the transformation is plated on LB/agar plates with the appropriate antibiotic and screen for colony growth.

Colonies were examined for the presence of the insert by restriction digests of purified plasmid DNA. Briefly, 5ml LB cultures were grown overnight and pure plasmid DNA was extracted using a proprietary protocol by Qiagen using spin columns. Restriction digests were performed with appropriate enzymes and the reaction examined on an agarose gel. Positive colonies (those which dropped appropriately sized inserts) were then screened for protein expression.

A.2. Affinty purification of the VCB ternary complex from inclusion bodies

Full length VHL, ElonginB, and ElonginC were grown in LB media supplemented with 60mg/L kanamycin (pET28b) at 37°C until the OD at 600 nm was between 1.0 and 2.0. The cultures were then induced with 1mM IPTG for 12 hours at same temperature. Cells were harvested by centrifugation (4500rpm in a Sorvall RC 3C swinging bucket centrifuge) and resuspended in 20ml per *E. coli* liter with Buffer A: 8M Urea, 01.M NaPhosphate, 10mM Tris pH 8.0.

The followiing procedure was iterated for each polypeptide. The solubilized cells were loaded onto TosoHaas AF-chelating resin which had been first pre-equilibrated in $NiSO_4$ overnight, then equilibrated in buffer A. The mixture was then nutated for one hour. The resin was then centrifuged in the swinging bucket Sorvall RC 3C at 2500rpm and washed with 200ml buffer A ("washing" consists of removing the supernatant after a spin, adding the wash buffer, and mixing with a long metallic spatula until the resin in resuspended). The resin was then washed with 200ml buffer: 8M Urea, 0.1M NaPhosphate, 30mM Imidazol, 10mM Tris pH 6.3, in order to remove some high molecular weight proteins which bind to the resin.. Finally, the protein was eluted with 200ml buffer: 8M Urea, 0.1M NaPhosphate, 0.5M Imidazol, 10mM Tris pH 4.5.

The three polypeptides were refolded in the following manner. They were first mixed in an approximate 1:1:1 molar ratio at a concentration of 0.2mg/ml for VHL and 0.1 mg/ml for ElonginB and ElonginC (diluted by adding buffer A). The final volume was about 1L. The solution was then brought by dialysis (dialysis membranes from Spectrapor with a 1000 m.w. cut off) to native buffer: 40mM HEPES pH 8.0, 0.1M KCl, 10% Glycerol, using four buffer exchanges of volume 5L. The dialysate was harvested by centrifugation at 13,000rpm in Sorvall RC 5C fixed angle centrifuge. From 500mgs of total protein in the refolding reaction, roughly 100mgs were contained in the soluble fraction.

The soluble fraction was loaded on a Q-sepharose FF column (Pharmacia) in 50mM Tris pH 7.8, 75mM NaCl, 10mM DTT, washed with the same buffer for 10 column volumes, and eluted with a 0.5M NaCl bump. The

eluted protein was concentrated to ~10-15mg/ml using a ultrafiltration device (Amicon) employing a filter with a 3,500 m.w. cutoff.

The concentrated protein was loaded on a Superdex 200 gel-filtration column that had been equilibrated in 25mM BTP 6.8, 200mM NaCl, 5mM DTT. The peak fraction eluting at 15mls (corresponding to 45kDa) was collected, and the shoulders from 13ml to 17 mls we pooled (60mgs of protein). The protein was flash frozen in liquid nitrogen in 50-100 μl aliquots.

A.3. Proteolytic Digestion with Subtilisin

The material which was submitted for mass spectroscopic analysis and N-terminal sequencing was produced by proteolytic digestion in the following manner. Purified, refolded ternary complex produced as described above was diluted to 1mg/ml in buffer E supplemented with $CaCl_2$ to final concentration of 2.5 mM. Subtilisin was added to 1% wt:wt ratio and the reaction was allowed to proceed for 10 hours on ice. After 10 hours, PMSF was added to a final concentration of 0.5mM. The digest was loaded on a SOURCE-Q column (Pharmacia) in 50mM Tris pH 7.8, 75mM NaCl, 10mM DTT, and bumped off in 0.5M NaCl in order to concentrate the protein for gel filtration.. The elution was run on Superdex 200 in 25mM BTP 6.8, 200mM NaCl, 5mM DTT. The peak fraction was concentrated by ultrafiltration to 7-7.5mg/ml and flash-frozen as 40μL aliquots. 1-2 mgs of this material was loaded on HPLC (see below for general protocol) and the various fragments were separated and submitted to the Memorial Sloan-Kettering Cancer Center Microchemistry Facility for N-terminal sequencing and mass spectroscopic analysis.

A.4. HPLC purification of the VCB ternary complex from inclusion bodies

Using freshly plated colonies, a small culture in LB would be inoculated overnight. VHL(54-213) cultures (pET28b) were grown using 60mg/ml kanamycin as before for the full length protein, and ElonginB and ElonginC (pET3d) were grown using 200mg/ml. The next morning, 12-24 liters of LB would be inoculated with several mls of the overnight culture and grown to OD at 600nm equal to 1.0. Cultures would then be induced with 1mM IPTG for 3-4 hours, harvested in 1L bottles at 4500rpm using swinging bucket centrifuge.

Cells were resuspended in about 20ml per *E. coli* liter with a buffer containing 50mM Tris pH 7.5, 200mM NaCl, DTT, and 0.05% NP-40. To the resuspended cells, 0.3mg/ml lysozyme and 0.5mM PMSF were added, and the mixture allowed to nutate for one hour at room temperature. Another 0.5mM PMSF was then added again. Cells were thoroughly lysed using a cell disruptor (Emusifelx-C5 from Avestin) and the lysate spun down at 11,000rpm by centrifugation in a Sorvall RC 5C.

The pellet was solubilized in 6.4M Guanidine-HCl, 150mM DTT, 0.5M NaCl, 50mM Tris pH 8.0 (entire solution should be pH 8.0), and 150mM DTT. The solution was then heated to 55˚C and stirred vigorously with stir bar for one hour. After this, the mixture was cooled on ice and brought to a TFA concentration of 0.1% (15l/L pellet) and stir a little more. The remaining insoluble material was spun down at 13,000 rpm and the supernatant placed on ice.

C4 resin (The Nerst Group) was placed in a 250ml tissue culture conical tube (Falcon), and wetted with 100% CH_3CN, such that the total wet C4

volume corresponded to about 100-150ml per 12 liters of *E.coli*. Each polypeptide was purified on a separate resin. The resin was washed to remove CH_3CN with H2O+0.1% TFA (three washes), after which the resuspended pellet was applied and mixed thoroughly for several minutes. The flow through was saved (if it was not milky white in appearance but instead yellow or even brownish in color, then the resin had been overloaded with protein, and the flow through would be re-applied after the purification to extract additional material). The resin was then washed with H_2O+0.1% TFA twice, washed twice 20% CH_3CN+0.1% TFA, and the protein eluted with 70% CH_3CN+0.1% TFA.. The eluted material was lyophilized using a model 12XL Freezemobile (Virtis).

The lyophilized material was resuspended in 6.4M Guanidine-HCl, 50mM DTT, 0.5mM NaCl, and 50mM Tris pH 8.0 (one ml per liter of *E. coli* pellet). The resuspension could take as long as 30min to 1 hour before most everything was dissolved (and at times required doubling the resuspension volume). The resuspended material was filtered through a 0.45 micron syringe filter (Millipore). These would clog quickly, requiring the use of several for large protein preparations.

The resuspended material was then purified by HPLC. As much as 100mgs protein were loaded at one time on a 80ml C4 resin HPLC column was pre-equilibrated in H_2O+0.1% TFA . A gradient in CH_3CN was used:

0% => 35% CH_3CN to 10 minutes
35% => 50% CH_3CN to 55 minutes all at 8 ml/min
50% => 100% CH_3CN to 65 minutes

VHL(54-213) elutes at about CH_3CN 42-46%, ElonginB at about 40% CH_3CN, and ElonginC at about 48-50% CH_3CN. The peaks were then pooled and lyophilized.

The powder was resuspended in 8M Urea, 0.1 Na-Phosphate, 10mM Tris 8.0, and10mM DTT. The three proteins were then mixed together in an approximate 1:1:1 molar ratio. The molecular weights for the three proteins are:

VHL(54-213): 18.5 kDa
ElonginB: 13kDa
ElonginC: 12kDa

This would lead to an approximate VHL: ElonginC: ElonginB mass ratio of 2:1:1. To insure that all the VHL is refolded, I typically would add excess ElonginC and ElonginB such that the final ration was closer to a 3:2:2 mass ratio (or 1.5:1:1). The mixture was diluted in 8M Urea, 0.1 Na-Phosphate, 10mM Tris 8.0, and10mM DTT for a final total protein concentration of 0.1mg/ml. Using dialysis membranes (Spectrapor) with a 1000 m.w. cut off, and a dialysate to buffer (40mM HEPES pH 8.0, 200mM KCl, 10% Glycerol, 10mM DTT) ratio of 1:5 the refolding reaction was carried out overnight with 5 buffer exchanges spaced apart by about 5 hours each. The dialysate was harvested by centrifugation at 13,000rpm in Sorvall RC 5C.

The soluble fraction was loaded on a Q-sepharose column (Pharmacia) in 50mM Tris pH 7.8, 75mM NaCl, 10mM DTT, washed with the same buffer for 10 column volumes, and eluted with a 0.5M NaCl bump. The eluted protein was concentrated to ~10mg/ml using a ultrafiltration device (Amicon) employing a filter with a 3,500 m.w. cutoff.

The concentrated protein was loaded on a Superdex 200 gel-filtration column that had been equilibrated in 25mM BTP 6.8, 200mM NaCl, 5mM DTT. The peak fraction eluting at 15mls (corresponding to 45kDa) was collected, and the shoulders from 13ml to 17 mls we pooled (60mgs of protein). The protein was flash frozen in liquid nitrogen in 50-100 μl aliquots.

The hexahistidine tag on VHL was removed by thrombin digestion. Thrombin cleavage buffer (10X from Novagen) was added to the protein solution, and then thrombin (Novagen) in a thrombin:VCB mass:mass ratio of 1:2000 (1 unit of Novagen thrombin per mg of VCB complex) was added the reaction was allowed to proceed at 4˚C for four hours. Thrombin was removed by ion-exchange chromatography by using a Heparin-sulfate column (TosoHaus). The column was pre-equilibrated in 25mM Tris 7.5, 200mM NaCl, and 5 mM DTT. The VCB ternary complex does not bind the resin at this salt concentration, whereas thrombin does. The thrombin digest would then be passed over the column and the flow through collected.

Following the removal of thrombin, the VCB ternary complex was further purified by anion-exchange chromatography. The heparin sulfate flow through was diluted to 50mM NaCl with a buffer containing 50mM Tris 7.5 and 10mM DTT. This material was loaded on a clean SOURCE-Q column (Pharmacia) equilibrated in the same buffer. A gradient from 0 mM to 250mM NaCl was run, and a peak containing the ternary complex was eluted at approximately 180mM NaCl with an extended tail possessing secondary, small peaks. The peak fractions were pooled for further purification.

The peak from SOURCE-Q was concentrated to 30-35mg/ml and purified by gel filtration on a Superdex 200 column as described above, with the exception that the final DTT concentration was reduced to 2mM. The tip

of the peak was hand-collected, typically yielding material in the range 20-25mg/ml. This was flash frozen in liquid nitrogen in 40λ aliquots.

A.5. The bacterial co-expression scheme

In the co-expression scheme adopted for producing protein used in the crystallographic analysis, two plasmids with compatible origins of replication were co-transformed into *E. coli*. One plasmid (pGEX-4T-3, Pharmacia) contained a dicistronic transcript harboring the VHL(54-213) and full length ElonginB genes and conferred resistance to ampicillin. The second plasmid (pBB75, a plasmid designed by Adrian Batchler and kindly supplied to our laboratory) contained the ElonginC gene (full length or residues 17-112) and conferred resistance to kanamycin.

In creating the dicistronic plasmid, VHL(54-213) was first cloned into pGEX-4T-3. The VHL gene was amplified by PCR as described in part A1 above. The 5′ oligonucleotide contained a 21 base pairs of sequence complimentary to the region of VHL beginning at residue 54. Preceding this segment to anneal were a clamp and restriction enzyme site, a sequence coding for 7 amino acids from the flexible spacer of the p107 protein (a homolog of the retinoblastoma protein), and a sequence coding for a consensus thrombin cleavage site. The design was used to engineer a flexible region between GST and the proteolytically identified structural domain of VHL, as cloning without a spacer into the BamH I site immediately following the thrombin site engineered at the end of GST in pGEX-4T-3 resulted in a fusion construct which poorly cleaved with thrombin. The 3′ oligonucleotide was designed to anneal at the end of the VHL gene.

VHL(54-213) 5′ oligonucleotide

5′ GCCGTGACGTCA GGATCC
CLAMP BamH I

p107 spacer Thrombin Site
AAT GTG CAG GGA CAT CTT CCC CTC GTA CCA CGA GGT TCA
N V Q G H L P L V P R G S

VHL (Met^{54})
ATG GAG GCC GGG CGG CCG CGG CCC **3′**
M E A G R P R P

VHL(54-213) 3′ oligonucleotide

5′ GCTAGGCATGTC CTA GAATTC ATC TCC CAT CCG TTG ATG TGC **3′**
CLAMP STOP EcoR I **D G M R Q H A**

ElonginB was inserted after the VHL stop codon by amplifying by PCR a large region just after the T7 promoter of the expression cassette of pET-3d (Novagen) into which ElonginB had been cloned (see above) such that approximately 30 base pairs (including a second ribosomal binding site, or rbs) separated the stop for VHL(54-213) and the ATG start codon for ElonginB. The 5′ primer was therefore complementary to a region of pET-3d, and the 3′ primer was complementary to the C-terminus of ElonginB. The oligonucleotides were designed as shown below.

ElonginB 5′ oligonucleotide

5′
GCCGTGACGTCA CCCGGG CTCTAGAAATAATTTTGTTTAACTTTAAG**3′**
CLAMP Xma I pre-rbs as in pET-3d

ElonginB 3′ oligonucleotide

5′ GCTAGGCATGTC CTCGAG AACCTT CCC CTA
CLAMP Xho I HindIII **G** STOP

CTG CAC AGC TTG TTC ATT GGC **3′**
Q V A Q E N A

ElonginC was cloned into pBB75 using the restriction endonucleases Nde I and BamH I (this was an altered version of pBB75 in which the EcoR I site was replaced with BamH I). The oligonucleotides for this PCR are described here. For both the full length and 17-112 ElonginC fragment, the cloning protocols were essentially the same.

ElonginC (full length) 5′ oligonucleotide

M
GCCGTGACGTCA CAT**ATG** GAT GGA GAG GAG AAA ACC TAT
CLAMP Nde I **D G E E K T Y**

ElonginC (17-112) 5′ oligonucleotide

M
GCCGTGACGTCA CAT**ATG** ATG TAT GTC AAA TTG ATA TCA
CLAMP Nde I **M Y V K L I S**

ElonginC 3′ oligonucleotide

GCTAGGCATGTC GGATCC CTA ACA ATC TAA GAA GTT CGC AGC
CLAMP BamHI STOP **C D L F N A A**

A.6. Purification of the VCB ternary complex from the native state

From freshly plated colonies (dual transformants on kanamyacin/ampicillin plates at the usual concentrations), 50ml cultures were grown overnight in LB containing both antibiotics. The next morning this

culture was used to inoculate 6L (with 8ml each) of LB (dual antibiotic), and these cultures were grown at 37°C until they reached an OD at 600nm of 1.0. The cultures were then equilibrated to 25°C and induced with fresh IPTG to 1mM. At this time an addition 200mg/L of ampicillin was added to the cultures. After six hours, the cultures were brought to 2mM IPTG with a thread dose of 200mg/ml ampicillin. The expression was allowed to continue for 16-24 hours.

Cells were then harvested in 20 ml per *E. coli* liter with 50mM Tris, 200mM NaCl, and 10mM DTT, pH 8.0. Cells ere lysed by a single pass through the cell disruptor, after which the lysate was immediately brought to 1mM PMSF. At this point, the solution was brought to 5mM $MgCl_2$, and Dnase I was added to reduce solution viscosity. Finally, the solution was brought to 0.5mM in EDTA.

The lysate was then spun at 17,000 rpm in a Sorvall RC 5C and the supernatant was passed over 20-30ml of Q-sepahrose resin in large gravity column. The flow through was applied to a 50ml glutathione-sepharose resin (Pharmacia) which had been equilibrated in 50mM Tris 8.0, 200mM NaCl, and 10mM DTT. After application of the Q-sepharose flow through, the column was washed with 10 column volumes with 50mM Tris 8.0, 500mM NaCl, and 10mM DTT. Finally, the ternary complex was eluted with 5 column volumes of 50mM Tris 8.0, 200mM NaCl, 10mM DTT, and 5mM reduced glutathione (Boheringer Mannheim).

In order both to concentrate the elution as well as to remove glutathione (which impairs thrombin proteolysis), the protein was applied to a

Q-sepharose column equilibrated in 50mM Tris pH 8.0, 50mM NaCl, and 10mM DTT. After 10 column volumes wash in the same buffer, the protein was eluted with 300mM NaCl. This would typically produce 50mls of fairly concentrated (5-10mg/ml) protein.

This material would then be digested overnight with a 1:25 wt:wt ratio of thrombin in a solution brought to 2.5 mM $CaCl_2$ (the high concentration was necessary to completely cut the second thrombin site adjacent to VHL). The digested material was then reapplied to a regenerated glutathione column. Thrombin was removed by ion-exchange chromatography by using a heparin-sulfate column (TosoHaus). The column was pre-equilibrated in 25mM Tris 7.5, 200mM NaCl, and 5 mM DTT. The VCB ternary complex does not bind the resin at this salt concentration, whereas thrombin does. The thrombin digest would then be passed over the column and the flow through collected.

The ternary complex was then diluted to 50mM NaCl with a buffer containing 50mM Tris pH 8.0 and10mM DTT. This diluted material was applied to a SOURCE-Q column. A gradient from 0 to 200mM NaCl eluted sharp peaks at 160mM for a ternary complex containing full length ElonginC or at 100mM for complexes which contained ElonginC (17-112). The peak fractions were pooled and concentrated to 35mg/ml, after which the material was run over a Superdex 200 sizing column pre-equilibrated in 5mM BTP pH 7.0, 200mM NaCl, and 2mM DTT. The tip of the peak was collected (typically around 25mg/ml) and the material flash frozen in liquid nitrogen as 80 μL aliquots.

APPENDIX B

CRYSTALLOGRAPHIC RELIABILITY INDICES

B.1. R_{sym}: agreement of symmetry related reflections

This reliability index compares reflections which are related by crystallographic and Friedel symmetry, and which, therefore, should ideally have identical intensities (ignoring contributions from anomalous scattering). The definition is

$$R_{sym} = \sum_h \sum_i \mid I_{h,i} - I_h \mid / \sum_h \sum_i I_{h,I}$$

for the intensity (I) of i observations of reflection h. Values under 10% overall for a dataset are usually considered adequate, and below 5% indicates data of excellent quality. The diffraction limit is usually set to the highest resolution bin in which the R_{sym} is less than or equal to 20-30%.

B.2. Indices for the MIR analysis

B.2.1. R_{iso}: the mean fractional isomorphous difference

To assess the quantity of heavy atom substitution, and to ascertain whether the derivative suffers from non-isomorphism, the behavior of the mean fractional isomorphous difference (R_{iso}) with respect to resolution is plotted. The overall R_{iso} is defined as

$$R_{iso} = \sum \mid F_{PH} - F_P \mid / \sum \mid F_P \mid,$$

where F_P is the native structure factor amplitude, and F_{PH} is the derivative structure factor amplitude.

An overall value in the range 10% to 30% is indicative of substitution at a level which will prove useful for phasing (assuming isomorphism). To estimate how isomorphous the native and derivative crystals are, R_{iso} is plotted as a function of resolution. As the average intensity change for a perfectly isomorphous crystal should be relatively constant at a resolution beyond 5Å, non-isomorphism will be manifest as a sharp rise in the R_{iso} with respect to resolution. These evaluations are integral to the screening of heavy atom derivatives.

B.2.2. Rc: The cullis R-factor

The Cullis R-factor is defined as

$$R_c = \sum \| F_{PH} \pm F_P | - F_{H(calc)} | \; / \; \sum | F_{PH} \pm F_P |,$$

where $F_{H(calc)}$ is the calculated heavy atom structure factor, and where F_P and F_{PH} are the native and derivative structure factor amplitudes, respectively. The combination of F_{PH} and F_P is determined by addition or subtraction depending on whether F_{PH} and F_P have opposite signs (additive case), or the have the same sign (in which case the difference is computed).

B.2.3.. PhP: the phasing power

This is the ratio of the root mean square value for the calculated heavy atom structure factor to the root mean square lack of closure error (called this because only in the case of zero error do the protein and heavy atom structure factors sum to give the derivatized structure factor - if error exists, then the vector sum triangle does not "close"). The latter term measures the difference between observed and calculated derivative structure factors, and thus gives some sense of quality of the heavy atom model. This phasing power, PhP, therefore scales the heavy atom contribution to the accuracy of the heavy atom model. The phasing power is defined as

$$\text{PhP} = (F_{\text{H(calc)}}^2 \;/\; (F_{\text{PH(obs)}} - F_{\text{PH(calc)}})^2)^{1/2},$$

where $F_{\text{PH(obs)}}$ and $F_{\text{PH(calc)}}$ are the observed and calculated derivative structure factors, respectively. In practice with the program MLPHARE in the CCP4 software suite, a phasing power above 1.0 indicates a useful derivative of low quality, 1.5 an average derivative, and above 2.0 an excellent derivative.

B.2.4. FOM: the figure of merit

Once the phases have been estimated by the MIR method, a probability distribution is constructed based on a Gaussian distribution in the error of the phase. The electron density calculated from the measured structure factors

$$\rho(x,y,z) = \frac{1}{V}\sum_{hkl} \left|\mathbf{F}_{hkl}\right| e^{i\alpha} e^{-2\pi i(hx+ky+lz)}$$

where ρ is the value of the electron density at the point(x,y,z), V is the volume of the volume of the unit cell, F_{hkl} is the structure factor of the hkl^{th} reflection, and α is the phase.

To reflect the fact that the data and heavy atom calculations are imperfect, some statistical weight indicating the confidence in the MIR determined phase should modulate this summation. In determining a statistical weight for each structure factor in this sum, the best weight is likely to be the one which minimizes the overall error in the electron density map, since this is the data (the map) which ultimately has impact on the model. The weight which accomplishes this is called the figure of merit.

The magnitude of figure of merit for a given reflection has a value between 0 and 1, and increases from 0 toward 1 as the phase probability distribution becomes more unimodular. It is thus a measure of the quality of the phase for a typical reflection, and the average value over the data gives a rough indication of the quality of the phases. This average is calculated with

$$FOM = <\sum P(\alpha) \exp(i\alpha) \; / \; \sum P(\alpha)>,$$

where P(α) is the probability distribution for the phase α. With current density modification algorithms, an FOM in the range 0.5-0.6 will likely yield density sufficiently interpretable to determine a crystal structure. With the addition of ncs averaging, the overall FOM can be even lower.

B.3. Indices for the assessment of model quality

These are the reliability indices which compare the model to the diffraction data. Both the standard and cross-validated residuals involve simulating a diffraction pattern based on the molecular model and comparing the calculated reflections with the observed reflections. These quantities may be summarized by

$$R = \sum \mid F_P - F_{calc} \mid / \sum F_P,$$

where F_{calc} is the calculated protein structure factor from the atomic model; R_{free} is the R-factor calculated using 10% of the reflection data chosen randomly and omitted from the refinement process, whereas R_{cryst} is calculated with the remaining data used in the refinement.

B.3.1. R_{cryst}: the standard model quality index

Until recently, R_{cryst} values below 25% were considered indicative of a correct structure (Laskowski et al., 1998). However, the R_{cryst} value on its own is an extremely poor indicator of model quality. This is due to the fact that modern refinement algorithms, especially simulated annealing, are extremely powerful minimizers, so much so that incorrect protein structure can be manipulated so as to bring the R_{cryst} to values once considered to indicate a finalized, high quality model. This is because these programs incorporate as part of the energy to be minimized a term related to this quantity. Kleywegt and Jones (1995) demonstrated the uselessness of an unaccompanied R_{cryst} by tracing the chain of a known structure backward and then refining the R_{cryst} to near 20% for a structure at 3Å resolution.

B.3.2. R_{free}: the cross-validated model quality index

The problems with potential bias in the R_{cryst} were addressed by Axel Brünger, who proposed a new assessment of model quality based on the cross-validation principle in mathematical statistics (Brünger, 1992). In calculating this new R-factor, called the R_{free}, a subset of the reflections (5-10%) are set aside before refinement such that only remaining data is used in the refinement process. The R_{free} calculated from this subset not used in the refinement should, in principle, give an unbiased assessment of the model quality. R_{free} values which correspond to an essentially correct model are not yet delineated, although a growing consensus is that this value should be at least below 35% (Laskowski et al., 1998). Even low R_{free} values are not a perfect quality assurance, and the R_{free} seems to be less sensitive to small local errors than it is to more large scale problems (as in sequence register errors, strand-strand misconnections, etc.).

BIBLIOGRAPHY

Aligue, R., Akhavan-Niak, H., and Russell, P. (1994). A role for Hsp90 in cell cycle control: Wee1 tyrosine kinase activity requires interaction with Hsp90. EMBO J. *13*, 6099-6106.

Appleby, J.L., Parkinson, J.S., and bourret, R.B. (1996). Signal transduction via the multistep phosphorelay: not necessarily a road less traveled. Cell *86*, 845-848.

Aronshtam, A., and Marinus, M.G. (1996). Dominant negative mutator mutations in the mutL gene of *Escherichia coli*. Nucleic Acids Res. *24*, 2498-2504.

Aso, T., Lane, W. S., Conaway, J. W., and Conaway, R. C. (1995). Elongin (SIII): a multisubunit regulator of elongation by RNA polymerase II. Science *269*, 1439-43.

Bai, C., Sen, P., Hofmann, K., Ma, L., Goebl, M., Harper, J. W., and Elledge, S. J. (1996). SKP1 connects cell cycle regulators to the ubiquitin proteolysis machinery through a novel motif, the F-box. Cell *86*, 263-74.

Ban, C., and Yang, W. (1998). Crystal structure and ATPase activity of MutL:Implications for DNA repair and mutagenesis. Cell *95*, 541-552.

Barlow, D. J., and Thornton, J. M. (1988). Helix Geometry in Proteins. J. Mol. Biol. *201*, 601-619.

Bayer, P., Arndt, A., Metzger, S., Mahajan, R., Melchior, F., Jaenicke, R., and Becker, J. (1998). Structure determination of the small ubiquitin-related modifier SUMO-1. J Mol Biol *280*, 275-86.

Bergerat, A., Massy, B., Gadekke, D., Variytas, P.-C., Nicolas, A., and Forterre. (1997). An atypical topoisomerase II from archaea with implications for meiotic recombination. Nature *386*, 414-417.

Beroud, C., Joly, D., Gallou, C., Staroz, F., Orfanelli, M. T., and Junien, C. (1998). Software and database for the analysis of mutations in the VHL gene. Nucleic Acids Res *26*, 256-8.

Bishop, J.M. (1995). Cancer: The rise of the genetic paradigm. Genes and Development, *9*(11), 1309-1315.

Bilwes, A.M., Alex, L.A., Crane, B.R., and Simon, M.I. (1999). Structure of CheA, a signal-transducing histidine kinase. Cell *96*, 131-141.

Blow, D.M. and Crick, F.H.C. (1959) The treatment of errors in the isomorphous replacement method. Acta Cryst. *12*, 794-802.

Blundell, T.L. and Johnson, L.N. (1976) Protein Crystallography. Academic Press, New York, New York.

Bohen, S. P., and Yamamoto, K. R. (1994). The Biology of Heat Shock Proteins and Molecular Chaperones, R. I. Morimoto, A. Tissieres and C. Georgopoulos, eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

Bohen, S. P., and Yamamoto, K. R. (1993). Isolation of Hsp90 mutants by screening for decreased steroid receptor function. Proc. Natl. Acad. Sci. U S A *90*, 11424-11428.

Borkovich, K. A., Farrelly, F. W., Finkelstein, D. B., Taulien, J., and Lindquist, S. (1989). hsp82 is an essential protein that is required in higher concentrations for growth of cells at higher temperatures. Mol. Cell. Biol. *9*, 3919-3930.

Bose, S., Weikl, T., Bugl, H., and Buchner, J. (1996). Chaperone function of hsp90-associated proteins. Science *274*, 1715-1717.

Bouchard, L., Lamarre, L., and Tremblay, P. J. (1989). Stochastic appearance of mammary tumors in transgenic mice carrying the MMTV/c-neu oncogene. Cell *57*, 931-940.

Bragg, W.L., and Perutz, M.F. (1954). Proc. Roy. Soc. A *225*, 315.

Bresnick, E. H., Dalman, F. C., Sanchez, E. R., and Pratt, W. B. (1989). Evidence that the 90-kDa heat shock protein is necessary for the steroid binding conformation of the L cell glucocorticoid receptor. J Biol Chem *264*, 4992-7.

Buchner, J. (1996) Supervising the fold: functional principles of molecular chaperones. FASEB J. *10*, 10-19.

Brünger, A. T. (1988) Crystallographic refinement by simulated annealing. Application to a 2.8Å resolution structure of aspartate aminotransferase. J. Mol. Biol. *203*, 803-816.

Brünger, A. T. (1990). Extension of molecular replacement: a new search strategy based on Patterson correlation refinement. Acta Cryst. A *46*, 46-47.

Brünger, A. T. (1991). X-PLOR, a system for X-ray crystallography and NMR, Version 3.1, Yale University. (New Haven, CT: Yale Univ. Press).

Brünger, A. T. (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature *355*, 472-475.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstieve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998). Crystallography & NMR system: A new software asuite for macromolecular structure determination. Acta Cryst. D *54*, 905-921.

CCP4 (1994). Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. Acta Cryst. D *50*, 760-763.

Chavany, C., Mimnaugh, E., Miller, P., Bitton, R., Nguyen, P., Trepel, J., Whitesell, L., Schnur, R., Moyer, J. D., and Neckers, L. (1996). p185(erbB2) binds to GRP94 in vivo - Dissociation of the p185(erbB2)/GRP94 heterocomplex by benzoauinone ansamycins precedes depletion of p185(erbB2). J. Biol. Chem. *271*, 4974-4977.

Chen, F., Kishida, T., Yao, M., Hustad, T., Glavac, D., Dean, M., Gnarra, J. R., Orcutt, M. L., Duh, F. M., Glenn, G., and et al. (1995). Germline mutations in the von Hippel-Lindau disease tumor suppressor gene: correlations with phenotype. Hum Mutat *5*, 66-75.

Chen, C.F., Chen, Y.M., Dai, K., Chen, P.l., Riley, D.J., and Lee, W.H. (1996). A new member of the Hsp90 family of molecular chaperones interacts with the retinoblastoma protein during mitosis and after heat shock. Mol. Cell. Biol. *16*, 4691-4699.

Chen, H.S., Singh, S.S., and Perdew, G.H. (1997). The Ah receptor is a sensitive target of geldanamycin-induced protein turnover. Arch. Biochemi. Biophys. *348*, 190-198.

Cho, Y., Gorina, S., Jeffrey, P.D., and Pavletich, N.P. (1994) Crystal structure of a p53 tumor suppressor-DNA complex: Understanding tumorigenic mutations. Science *265*, 346-355.

Choyke, P.L., Glenn, G.M., Walther, M.M., Patronas, N.J., Linehan, W.M., Zbar, B. (1995) von Hippel-Lindau disease: Genetic, Clinical, and Imaging Features. Radiology, *146*, 629-642.

Collaborative Computational Project, Number 4. (1994). The CCP4 Suite: Programs for Protein Crystallography. Acta Cryst. *D50*, 760-763.

Crossey, P. A., Richards, F. M., Foster, K., Green, J. S., Prowse, A., Latif, F., Lerman, M. I., Zbar, B., Affara, N. A., Ferguson-Smith, M. A., and et al. (1994). Identification of intragenic mutations in the von Hippel-Lindau disease tumour suppressor gene and correlation with disease phenotype. Hum Mol Genet *3*, 1303-8.

Csermely, P., and Kahn, C.R., (1991) the 90 kDa heat-shock protein (Hsp90) possesses an ATP binding-site and autophosphorylation activity. J. Biol. Chem. *266*, 4943-4950.

Csermely, P., *et al*. (1993). ATP induces a confomrational change in the 90-kDa heat shock protein (hsp90). J. Biol. Chem. *268*, 1901-1907.

Csermely, P., Miyata, Y., Schaider, T., and Miyata, I. (1995). Autophosphorylation of GRP94 (endoplasmin). J. Biol. Chem. *270*, 6381-6388.

Cutforth, T., and Rubin, G. M. (1994). Mutations in Hsp83 and cdc37 impair signaling by the sevenless receptor tyrosine kinase in Drosophila. Cell *77*, 1027-1036.

Dasgupta, G., and Momand, J. (1997). Geldanamycin prevents nuclear translocation of mutant p53. Exp. Cell Res. *237*, 29-37.

DeBoer, C., Meulman, P. A., Wnuk, R. J., and Peterson, D. H. (1970). Geldanamycin, a new antibiotic. J. Antibiot. (Tokyo) *23*, 442-447.

Dittmar, K. D., Hutchison, K. A., Owens-Grillo, J. K., and Pratt, W. B. (1996). Reconstitution of the steroid receptor-hsp90 heterocomplex assembly system of rabbit reticulocyte lysate. J. Biol. Chem. *271*, 12833-12839.

Drenth, J. (1994). Principles of protein x-ray crystallography. Springer-Verlag Press, New York.

Duan, D. R., Pause, A., Burgess, W. H., Aso, T., Chen, D. Y., Garrett, K. P., Conaway, R. C., Conaway, J. W., Linehan, W. M., and Klausner, R. D. (1995). Inhibition of transcription elongation by the VHL tumor suppressor protein. Science *269*, 1402-6.

Duina, A. A., Chang, H. J., Marsh, J. A., Lindquist, S., and Gaber, R. F. (1996). A cyclophilin function in hsp90-dependent signal transduction. Science *274*, 1713-1715.

Dunbrack, R.L., Jr., Gerloff, D.L., Bower, M., Chen, X., Lichtarge, O., and Cohen, F.E. (1997). Meeting review: the second meeting on the critical assessment of techniques for protein structure prediction (CASP2), Asilomar, California, December 13-16, 1996. Folding and Design *1*, R27-R42.

Flaherty, K.M., Wilbanks, S.M., De Luca-Flaherty, C., and McKay, D.B. (1994). Structural basis of the 70kDa heat shock cognate protein ATP hydrolytic activity. II. Structure of the active site with ATP or ADP bound to wild-type and mutant ATPase fragment. J. Biol. Chem. *269*, 12899-12907.

Flynn, G. C., Pohl, J., Flocco, M. T., and Rothman, J. E. (1991). Peptide-binding specificity of the molecular chaperone BiP. Nature *353*, 726-730.

Foster, K., Prowse, A., van den Berg, A., Fleming, S., Hulsbeek, M. M., Crossey, P. A., Richards, F. M., Cairns, P., Affara, N. A., Ferguson-Smith, M. A., and et al. (1994). Somatic mutations of the von Hippel-Lindau disease tumour suppressor gene in non-familial clear cell renal carcinoma. Hum Mol Genet *3*, 2169-73.

Freeman, B. C., and Morimoto, R. I. (1996). The human cytosolic molecular chaperones hsp90, hsp70 (hsc70) and hdj-1 have distinct roles in recognition of a non-native protein and protein refolding. EMBO J. *15*, 2969-2979.

Freeman, B. C., Toft, D. O., and Morimoto, R. I. (1996). Molecular chaperone machines: chaperone activities of the cyclophilin cyp-40 and the steroid aporeceptor-associated protein p23. Science *274*, 1718-1720.

Furusaki, A., Matsumoto, T., Nakagawa, A., and Omura, S. (1980). Herbimycin A: an ansamycin antibiotic; X-ray crystal structure. J. Antibiot. (Tokyo) *33*, 781-782.

Garrett, K. P., Aso, T., Bradsher, J. N., Foundling, S. I., Lane, W. S., Conaway, R. C., and Conaway, J. W. (1995). Positive regulation of general transcription factor SIII by a tailed ubiquitin homolog. Proc Natl Acad Sci U S A *92*, 7172-6.

Gerloff, D.L., Cohen, F.E., Korostensky, C., Turcotte, M., Gonnet, G.H., Benner, S.A. (1997). A predicted consensus structure for the N-terminal fragment of the heat shock protein HSP90 family. Proteins *27*, 450-458.

Gnarra, J. R., Tory, K., Weng, Y., Schmidt, L., Wei, M. H., Li, H., Latif, F., Liu, S., Chen, F., Duh, F. M., and et al. (1994). Mutations of the VHL tumour suppressor gene in renal carcinoma. Nat Genet *7*, 85-90.

Gnarra, J. R., Ward, J. M., Porter, F. D., Wagner, J. R., Devor, D. E., Grinberg, A., Emmert-Buck, M. R., Westphal, H., Klausner, R. D., and Linehan, W. M. (1997). Defective placental vasculogenesis causes embryonic lethality in VHL-deficient mice. Proc Natl Acad Sci U S A *94*, 9102-7.

Gnarra, J. R., Zhou, S., Merrill, M. J., Wagner, J. R., Krumm, A., Papavassiliou, E., Oldfield, E. H., Klausner, R. D., and Linehan, W. M. (1996). Post-transcriptional regulation of vascular endothelial growth factor mRNA by the product of the VHL tumor suppressor gene. Proc Natl Acad Sci U S A *93*, 10589-94.

Grenert J.P., Sullivan W.P., Fadden P., Haystead T.A.J., Clark J., Mimnaugh E., Krutzsch H., Ochel H.J., Schulte T.W., Sausville E., Neckers L.M., Toft D.O. (1997). The amino-terminal domain of heat shock protein 90 (hsp90) that binds geldanamycin is an ATP/ADP switch domain that regulates hsp90 conformation. J Biol Chem *272*, 23843-50

Hangaishi, A., Ogawa, S., Mini, K., Hosoya, N., Chiba, S., Yazaki, Y., and Hirai, H. (1997). Mutations and loss of expression of a mismatch repair gene., hMHL1, in leukemia and lymphoma cell lines. Blood *89*, 1740-1747.

Hartl, F. U. (1995). Principles of chaperone-mediated protein folding. Philos Trans R Soc Lond B Biol Sci *348*, 107-12.

Hartl, F. U., and Martin, J. (1995). Molecular chaperones in cellular protein folding. Curr Opin Struct Biol *5*, 92-102.

Hartson, S. D., and Matts, R. L. (1994). Association of Hsp90 with cellular Src-family kinases in a cell-free system correlates with altered kinase structure and function. Biochemistry *33*, 8912-8920.

Heldin, C.-H., Miyazomo, K., and ten Dijke, P. (1997). TGF-β signaling from cell membrane to nucleus through SMAD proteins. Nature *390*, 465-471.

Hendrick, J. P., and Hartl, F. U. (1995). The role of molecular chaperones in protein folding. Faseb J *9*, 1559-69.

Herman, J. G., Latif, F., Weng, Y., Lerman, M. I., Zbar, B., Liu, S., Samid, D., Duan, D. S., Gnarra, J. R., Linehan, W. M., and et al. (1994). Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma. Proc Natl Acad Sci U S A *91*, 9700-4.

Hilton, D. J., Richardson, R. T., Alexander, W. S., Viney, E. M., Willson, T. A., Sprigg, N. S., Starr, R., Nicholson, S. E., Metcalf, D., and Nicola, N. A. (1998). Twenty proteins containing a C-terminal SOCS box form five structural classes. Proc Natl Acad Sci U S A *95*, 114-9.

Hochstrasser, M. (1998). There's the Rub: a novel ubiquitin-like modification linked to cell cycle regulation. Genes and Development *12*, 901-907.

Holm, L., and Sander, C. (1998). Touring protein fold space with Dali/FSSP. Nucl. Acids Res. *26*, 316-319.

Huang, L., Hofer, F., Martin, G. S., and Kim, S.-H. (1998). Structural basis for the interaction of Ras with RalGDS. Nature Struct. Biol. *5*, 422-426.

Hunter, T., and Pines, J. (1994). Cyclins and cancer II: cyclin D and CDK inhibitors come of age. Cell *79*, 573-582.

Iliopoulos, O., Kibel, A., Gray, S., and Kaelin, W. G., Jr. (1995). Tumour suppression by the human von Hippel-Lindau gene product. Nat Med 1, 822-6.

Iliopoulos, O., Levy, A. P., Jiang, C., Kaelin, W. G., Jr., and Goldberg, M. A. (1996). Negative regulation of hypoxia-inducible genes by the von Hippel-Lindau protein. Proc Natl Acad Sci U S A 93, 10595-9.

Iliopoulos, O., Ohh, M., and Kaelin, W. G., Jr. (1998). pVHL19 is a biologically active product of the von Hippel-Lindau gene arising from internal translation initiation. Proc Natl Acad Sci U S A 95, 11661-6.

Isaacs, J. T., and Coffey, D. S. (1979). Androgenic control of prostatic growth: Regulation of steroid levels. UICC Monograph (Prostatic Cancer) 48, 112-122.

Jakob, U., and Buchner, J. (1994). Assisting spontaneity: the role of Hsp90 and small Hsps as molecular chaperones. Trends Biochem. Sci. 19, 205-211.

Jakob, U., Scheibel;T., Bose, S., Reinstein, J., and Buchner, J. (1996). Assessment of the ATP binding properties of Hsp90. J. Biol. Chem. 271, 10035-10041.

Jakob, U., Muse, W., Eser, M., and Bardwell, J.C.A. (1999). Chaperone activity with a redox switch. Cell 96, 341-352.

Johnson, J.L., and Toft, D.O. (1994). A novel chaperone complex for steroid-receptors involving heat-shock proteins, immunophilins, and p23. J. Biol. Chem. 269, 24989-24993.

Johnson, J., Corbisier, R., Stensgard, B., and Toft, D. (1996). The involvement of p23, hsp90, and immunophilins in the assembly of progesterone receptor complexes. J. Steroid Biochem. Mol. Biol. 56, 31-37.

Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Cryst. A 47.

June, C. H., Fletcher, M. C., Ledbetter, J. A., Schieven, G. L., Siegel, J. N., Phillips, A. F., and Samelson, L. E. (1990). Inhibition of tyrosine phosphorylation prevents T-cell receptor-mediated signal transduction. Proc. Natl. Acad. Sci. USA 87, 7722-7726.

Kamura, T., Sato, S., Haque, D., Liu, L., Kaelin, W. G., Conaway, R. C., and Conaway, J. W. (1998). The ElonginBC complex interacts with the conserved SOCS-box motif present in members of the SOCS, ras, WD-40 repeat, and ankyrin repeat families. Genes Dev. 12, 3872-3881.

Kanstan, M.B. (1992). Cell 71, 587.

Kibel, A., Iliopoulos, O., DeCaprio, J. A., and Kaelin, W. G., Jr. (1995). Binding of the von Hippel-Lindau tumor suppressor protein to ElonginB and C. Science 269, 1444-6.

Kimura, Y., Matsumoto, S., and Yahara, I. (1994). Temperature-sensitive mutants of hsp82 of the budding yeast Saccharomyces cerevisiae. Mol. Gen. Genet. 242, 517-527.

Kipreos, E. T., Lander, J. P., Wing, W. W. H., and E.M., H. (1996). cul-1 is required for cell cycle exit in C. elegans and identifies a novel gene family. Cell 85, 829-839.

Kishida, T., Stackhouse, T. M., Chen, F., Lerman, M. I., and Zbar, B. (1995). Cellular proteins that bind the von Hippel-Lindau disease gene product: mapping of binding domains and the effect of missense mutations. Cancer Res 55, 4544-8.

Kleywegt, G.J. and Jones T.A. (1995). Where freedom is given, liberties are taken. Structure 3, 535-540.

Kleywegt, G.J. and Jones T.A. (1996) Good model-building and refinement practice. Meth. Enzymol. in press.

Kleywegt, G.J. and Brunger, A.T. (1996). Checking your imagination: applications of the free R value. Structure 4, 897-904.

Knebelmann, B., Ananth, S., Cohen, H. T., and Sukhatme, V. P. (1998). Transforming growth factor alpha is a target for the von Hippel-Lindau tumor suppressor. Cancer Res 58, 226-31.

Koyasu, S., Nishida, E., Kadowaki, T., Matsuzaki, F., Iida, K., Harada, F., Kasuga, M., Sakai, H., and Yahara, I. (1986). Two mammalian heat shock proteins, HSP90 and HSP100, are actin-binding proteins. Proc. Natl. Acad. Sci. U S A 83, 8054-8058.

Kraulis, P. J. (1991). Molscript: A program to produce both detailed and schematic plots of protein structures. J. Appl. Cryst. 24, 946-950.

Kretzschmar, M., and Massagué,J. (1998). SMADS: mediators and regulators of TGF-b signaling. Curr. Opin. Genet. Dev. 8, 103-111

Kreusch, A., Pfaffinger, P. J., Stevens, C. F., and Choe, S. (1998). Crystal structure of the tetramerization domain of the Shaker potassium channel [see comments]. Nature 392, 945-8.

Lamiell, J. M., Salazar, F. G., and Hsia, Y. E. (1989). von Hippel-Lindau disease affecting 43 members of a single kindred. Medicine (Baltimore) 68, 1-29.

Laskowski, R.A., MacArthur, M.W., and Thornton, J.M. (1998). Validation of protein models derived from experiment. Curr. Opin. Struct. Biol. 8, 631-639.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 26, 283-291.

Latif, F., Tory, K., Gnarra, J., Yao, M., Duh, F. M., Orcutt, M. L., Stackhouse, T., Kuzmin, I., Modi, W., Geil, L., and et al. (1993). Identification of the von Hippel-Lindau disease tumor suppressor gene [see comments]. Science 260, 1317-20.

Lee, S., Neumann, M., Stearman, R., Stauber, R., Pause, A., Pavlakis, G., and Klausner, R. (1999). Transcription-Dependent Nuclear-Cytoplasmic Trafficking Is Required for the Function of the von Hippel-Lindau Tumor Suppressor Protein. Mol Cell Biol 19, 1486-1497.

Levy, A. P., Levy, N. S., and Goldberg, M. A. (1996). Hypoxia-inducible protein binding to vascular endothelial growth factor mRNA and its modulation by the von Hippel-Lindau protein. J Biol Chem 271, 25492-7.

Li, Z., and Srivastava, P.K. (1993). Tumor rejection antigen GP96/GRP94 is an ATPase: implication for antigen presentation and protein folding. EMBO J., 12 3143-3151.

Linehan, W. M., Lerman, M. I., and Zbar, B. (1995). Identification of the von Hippel-Lindau (VHL) gene. Its role in renal cancer. Jama 273, 564-70.

Lonergan, K. M., Iliopoulos, O., Ohh, M., Kamura, T., Conaway, R. C., Conaway, J. W., and Kaelin, W. G., Jr. (1998). Regulation of hypoxia-inducible mRNAs by the von Hippel-Lindau tumor suppressor protein requires binding to complexes containing elongins B/C and Cul2. Mol Cell Biol 18, 732-41.

Lyapina, S. A., Correll, C. C., Kipreos, E. T., and Deshaies, R. J. (1998). Human CUL1 forms an evolutionarily conserved ubiquitin ligase complex (SCF) with SKP1 and an F-box protein. Proc Natl Acad Sci U S A 95, 7451-6.

Maddock, I. R., Moran, A., Maher, E. R., Teare, M. D., Norman, A., Payne, S. J., Whitehouse, R., Dodd, C., Lavin, M., Hartley, N., Super, M., and Evans, D. G. (1996). A genetic register for von Hippel-Lindau disease. J Med Genet 33, 120-7.

Melmon, K., and Rosen, S. (1964). Lindau's disease: Review of the literature and study of a large kindred. Am J Med 36, 595-617.

Melnick, J., Aviel, S., and Argon, Y. (1992). The endoplasmic-retulum stress protein-GRP94, in addition to BiP associates with unassembled immunoglobulin-chains. J. Biol. Chem. 267, 21303-21306.

Merrit, E. A., and Murphy, M. E. (1994). Raster3D Version 2.0: A Program for Photorealistic Molecular Graphics. Acta Cryst. D50, 869-873.

Miller, R. T., Jones, D. T., and Thornton, J. M. (1996). Protein fold recognition by sequence threading - tools and assessment techniques. FASEB J. 10, 171-178.

Miller, P., DiOrio, C., Moyer, M., Schnur, R. C., Bruskin, A., Cullen, W., and Moyer, J. D. (1994). Depletion of the erbB-2 Gene Product p185 by Benzoquinoid Ansamycins. Cancer Res. 54, 2724-2730.

Miller, S. J. (1995). I. The asymmetric synthesis of the antitumor antibiotic macbecin. Diss. Abstr. Int. [B] 55, 3313.

Modrich, P., and Lahue, R. (1996). Mismatch repair in replication fidelity, genetic recombination, and cancer biology. Annu, Rev. Biochem. 65, 101-133.

Mukhopadhyay, D., Knebelmann, B., Cohen, H. T., Ananth, S., and Sukhatme, V. P. (1997). The von Hippel-Lindau tumor suppressor gene product interacts with Sp1 to repress vascular endothelial growth factor promoter activity. Mol Cell Biol *17*, 5629-39.

Murakami, Y., Mizuno, S., and Uehara, Y. (1994). Accelerated degradation of 160 kDa epidermal growth factor (EGF) receptor precursor by the tyrosine kinase inhibitor herbimycin A in the endoplasmic reticulum of A431 human epidermoid carcinoma cells. Biochem. J. *301*, 63-68.

Murzin, A. G., Brenner, S. E., Hubbard, T., and Chothia, C. (1995). **scop**:a structural classification of proteins database fo the investigation of sequences and structures. J Mol Biol *247*, 536-540.

Mushegian, A.R., Bassett, Jr., D.E., Bogushki, M.S., Bork, P., and Koonin, E.V. (1997). Positionally cloned human disease genes: patterns of evolutional conservation and functional motifs. Proc. Natl. Acad. Sci. USA *94*, 5831-5836.

Nadeau, K., Sullivan, M.A., Bradley, M., Engman, D.M., and Walsh, C.T. (1992). 83-kilodalton heat-shock proteins of trypanasomes are potent peptide-stimulated ATPases. Protein Sci. *1*, 970-979.

Nadeau, K., Das, A., and Walsh, C.T. (1993). Hsp90 chaperonins possess ATPase activity and bind heat-shock transcription factors and peptidyl proyl isomerases. J Biol. Chem. *268*, 1479-1487.

Nair, S. C., Toran, E. J., Rimerman, R. A., Hjermstad, S., Smithgall, T. E., and Smith, D. F. (1996). A pathway of multi-chaperone interactions common to diverse regulatory proteins: estrogen receptor, Fes tyrosine kinase, heat shock transcription factor Hsf1, and the aryl hydrocarbon receptor. Cell Stress & Chaperones *1*, 237-250.

Nardai, G., Schnaider, T., Söti, C., Ryan, M.T., Hoj, P.B., Somogy, J., and Csermely, P. (1996). Characterization of the 90 kDa heat shock protein (Hsp90)-associated ATP/GTPase. J. Biosci. *21*, 179-190.

Nassar, N., Horn, G., Herrmann, C., Scherer, A., McCormick, F., and Wittinghofer, A. (1995). The 2.2 A crystal structure of the Ras-binding domain of the serine/threonine kinase c-Raf1 in complex with Rap1A and a GTP analogue. Nature *375*, 554-60.

Nathan, D. F., and Lindquist, S. (1995). Mutational analysis of Hsp90 function: interactions with a steroid receptor and a protein kinase. Mol. Cell. Biol. *15*, 3917-3925.

Nicholls, A., Sharp, K. A., and Honig, B. (1991). Protein Folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins Struct. Funct. Genet.*11*, 281-296.

Nieland, T. J. F., Tan, M. C. A. A., Monne-van Muijen, M., Koning, F., Kruisbeek, A. M., and Vanbleek, G. M. (1996). Isolation of an immunodominant viral peptide that is endogenously bound to the stress protein GP96/GRP94. Proc. Natl. Acad. Sci. USA *93*, 6135-6139.

Obermann, W.M.J., Sondermann, H., Russo, A.A., Pavletich, N.P., and Hartl, F.U. (1998). In vivo function of Hsp90 is dependent on ATP binding and ATP hydrolysis. J. Cell Biol. *143*, 901-910.

Ohh, M., Yauch, R. L., Lonergan, K. M., Whaley, J. M., Stemmer-Rachamimov, A. O., Louis, D. N., Gavin, B. J., Kley, N., Kaelin, W. G., Jr., and Iliopoulos, O. (1998). The von Hippel-Lindau tumor suppressor protein is required for proper assembly of an extracellular fibronectin matrix. Mol Cell *1*, 959-68.

Ollis, D. and White, S. (1990) Protein crystallization. Meth. Enzymol. *182*, 646-659.

Omura, S., Iwai, Y., Takahashi, Y., Sadakane, N., Nakagawa, A., Oiwa, H., Hasegawa, Y., and Ikai, T. (1979). Herbimycin, a new antibiotic produced by a strain of Streptomyces. J. Antibiot. (Tokyo) *32*, 255-261.

Ono, Y., Kozai, Y., and Ootsu, K. (1982). Antitumor and cytocidal activities of a newly isolated benzenoid ansamycin, macbecin I. Gann. *73*, 938-944.

Osborne, C. K., Yochmowitz, M. G., Knight, W. A., and McGuire, W. L. (1980). The value of estrogen and progesterone receptors in the treatment of breast cancer. Cancer *46*, 2884-2888.

Otwinowski, Z., and Minor, W. (1997). Processing of x-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307-326.

Pal, S., Claffey, K. P., Cohen, H. T., and Mukhopadhyay, D. (1998). Activation of Sp1-mediated vascular permeability factor/vascular endothelial growth factor transcription requires specific interaction with protein kinase C zeta. J Biol Chem *273*, 26277-80.

Pal, S., Claffey, K. P., Dvorak, H. F., and Mukhopadhyay, D. (1997). The von Hippel-Lindau gene product inhibits vascular permeability factor/vascular endothelial growth factor expression in renal cell carcinoma by blocking protein kinase C pathways. J Biol Chem *272*, 27509-12.

Panaretou, B., Prodromou, C., Roe, S.M., O'Brien, R., Ladbury, J.E., Piper, P.W. and Pearl, L.H. (1998) ATP binding and hydrolysis are essential to the function of the Hsp90 molecular chaperone *in vivo*. EMBO J. *17*, 4829-4836.

Parkinson, J.S., and Kofoid, E.C. (1992). Communication modules in bacterial signaling proteins. Annu. Rev. Genet. *26*, 71-112.

Patton, E. E., Willems, A. R., Tyers, M. (1998a). Combinatorial control in ubiquitin-dependent proteolysis: don't Skp the F-box hypothesis. Trends Genet *14*, 236-243.

Patton, E. E., Willems, A. R., Sa, D., Kuras, L., Thomas, D., Craig, K. L., Tyers, M. (1998b). Cdc53 is a scaffold protein for multiple Cdc34/Skp1/F-box proteincomplexes that regulate cell division and methionine biosynthesis in yeast Genes Dev *12*, 692-705.

Pause, A., Lee, S., Lonergan, K. M., and Klausner, R. D. (1998). The von Hippel-Lindau tumor suppressor gene is required for cell cycle exit upon serum withdrawal. Proc Natl Acad Sci U S A *95*, 993-8.

Pause, A., Lee, S., Worrell, R. A., Chen, D. Y., Burgess, W. H., Linehan, W. M., and Klausner, R. D. (1997). The von Hippel-Lindau tumor-suppressor gene product forms a stable complex with human CUL-2, a member of the Cdc53 family of proteins. Proc Natl Acad Sci U S A *94*, 2156-61.

Peltomaki, P., and Vasen, H.F. (1997). Mutations predisposing to hereditary nonpolyposis colorectal cancer: database and results of a collaborative study. Gastroenterology *113*, 1146-1158.

Penninga, D., van der Veen, B.A., Knegtel, R.M.A., van Hijum, S.A.F.T., Rozeboom, H.J., Kalk, K.H., Dijkstra, B.W., and Dijkhuizen, L. (1996). The raw starch binding domain of cyclodextrin glycosyltransferase from *Bacillus circulans* strain 251. J Biol Chem *271*, 32777-32784.

Picard, D., Khursheed, B., Garabedian, B., Fortin, M. G., Lindquist, S., and Yamamoto, K. R. (1990). Reduced levels of hsp90 compromise steroid receptor action *in vivo*. Nature *348*, 166-168.

Prodromou, C., Roe, S.M., Piper, P.W., and Pearl, L.H..(1997a). A molecular clamp in the crystal structure of the N-terminal domain of the yeast Hsp90 chaperone. Nat. Struct. Biol. *4*, 477-482.

Prodromou, C., Roe, S.M., O'Brien, R., Ladbury, J.E.Piper, P.W., and Pearl, L.H..(1997b). Identification and structural characterization of the ATP/ADP-binding site in the Hsp90 molecular chaperone. Cell *90*, 65-75.

Pratt, W. B., and Welsh, M. J. (1994). Chaperone functions of the heat shock proteins associated with steroid receptors. Semin. Cell Biol. *5*, 83-93.

Ramachandran, G.N., Ramakrishnan, C, .and Sasisekharan, V. (1963). Stereochemistry of Polypeptide Chain Congfiguration. J. Mol. Biol., 7, 95-99.

Ramachandran, G.N., and Sasisekharan, V. (1968). Conformation of polypeptides and proteins. Adv. Protein Chem., 23, 283-437.

Rao-Naik, C., delaCruz, W., Laplaza, J., Tan, S., Callis, J., and Fisher, A. (1998). The Rub Family of Ubiquitin-like Proteins. J Biol Chem *273*, 34976-34982.

Rinehart, K. L., Jr., and Shield, L. S. (1976). Chemistry of the ansamycin antibiotics. Fortschr. Chem. Org. Naturst. *33*, 231-307.

Roe, S.M., Prodromou, C., O'Brien, R., Ladbury, J.E., Piper, P.W., and Pearl, L.H. (1999). Structural basis for inhibition of the hsp90 molecular chaperone by the antitumor antibiotics radiciol and geldanamycin. J. Med. Chem. *42*, 260-266.

Rossmann, M.G., and Blow, D.M. (1962) The detection of sub-units within the crystallographic asymmetric unit. Acta Cryst. *15*, 24-31.

Rost, B., and Sander, C. (1993). Prediction of protein structure at better than 70% accuracy. J. Mol. Biol. *232*, 584-599.

Sasaki, K., Yasuda, H., and Onodera, K. (1979). Growth inhibition of virus transformed cells in vitro and antitumor activity in vivo of geldanamycin and its derivatives. J. Antibiot. (Tokyo) *32*, 849-851.

Scheibel, T., Neuhofen, S., Weikl, T., mayr, C, Reinstein, J., Vogel, P.D., and., Buchner, J. (1997). ATP-binding properties of human Hsp90. J. Biol. Chem. *272*, 18608-18613.

Scheibel, T., and Buchner, J. (1998). The Hsp90 complex – a super-chaperone machine as a novel drug target. Biochem. Pharmacol. *56*, 675-682.

Scheibel, T., Weikl, T., and Buchner, J. (1998). Two chaperone sites in Hsp90 differing in substrate specificity and ATP dependence. Proc. Natl. Acad. Sci. USA *95*, 1495-1499.

Schneider, C., Sepp-Lorenzino, L., Nimmesgern, E., Ouerfelli, O., Danishefsky, S., Rosen, N., and Hartl, F. U. (1996). Pharmacologic shifting of a balance between protein refolding and degradation mediated by Hsp90. Proc. Natl. Acad. Sci. U S A *93*, 14536-14541.

Schnur, R. C., Corman, R. J., Gallaschum, R. J., Cooper, B. A., Dee, J. L., Doty, J. L., Muzzi, M. L., DiOrio, C. I., Barbacci, E. G., Miller, P. E., Pollack, V. A., Savage, D. M., Sloan, D. E., Pustilnik, L. R., Moyer, J. D., and Moyer, M. P. (1995). erbB-2 Oncogene Inhibition by Geldanamycin Derivatives: Synthesis, Mechanism of Action, and Structure-Activity Relationships. J. Med. Chem. *38*, 3813-3820.

Schnur, R. C., Corman, R. J., Gallaschum, R. J., Cooper, B. A., Dee, J. L., Doty, J. L., Muzzi, M. L., Moyer, J. D., DiOrio, C. I., Barbacci, E. G., Miller, P. E., O'Brien, M. J., Morin, B. A., Pollack, V. A., Savage, D. M., Sloan, D. E., Pustilnik, L. R., and Moyer, M. P. (1995). Inhibition of the Oncogene Product p185(erbB-2) in Vitro and in Vivo by Geldanamycin and dihydrogeldanamycin Derivatives. J. Med. Chem. *38*, 3806-3812.

Schoenfeld, A., Davidowitz, E. J., and Burk, R. D. (1998). A second major native von Hippel-Lindau gene product, initiated from an internal translation start site, functions as a tumor suppressor. Proc Natl Acad Sci U S A *95*, 8817-22.

Schulte, T. W., Blagosklonny, M. V., Ingui, C., and Neckers, L.M. (1995). Disruption of the Raf-1-Hsp90 molecular complex results in destabilization of Raf-1 and loss of Raf-1-Ras association. J. Biol. Chem. *270*, 24585-24588.

Schulte, T.W., Akinaga, S., Soga, S., Sullivan, W., Stensgard, B., Toft, D., and Neckers, L.M. (1998). Antibiotic radicicol binds to the N-terminal domain of Hsp90 and shares important biologic activities with geldanamycin. Cell Stress Chaperones *3*, 100-108.

Schumacher, R. J., Hurst, R., Sullivan, W. P., McMahon, N. J., Toft, D. O., and Matts, R. L. (1994). ATP-dependent chaperoning activity of reticulocyte lysate. J. Biol. Chem. *269*, 9493-9499.

Sepehrnia, B. Paz, I.B., Dasgupta, G., and Momand, J. (1996). Heat-shock protein 84 forms a complex with mutant p53 protein predominantly within a cytoplasmic compartment of the cell. J. Biol. Chem. *271*, 15084-15090.

Sharma, S.V., Agatsuma, T., and Nakano, H. (1998). Targeting of the protein chaperone, HSP90, by the transformation suppressing agent, radicicol. Oncogene *16*, 2639-2645.

Sheldrick, G. (1991). Patterson Interpretation and the Use of Macromolecular Delta-F Data. Daresbury, UK.

Shi, Y., Brown, E.D., and Walsh, C.T. (1994). Expression of recombinant human casein kinase-II and recombinant heat-shock protein-90 in *Escherichia coli* and characteriztion of their interactions. Proc. Natl. Acad. Sci. USA *91*, 2767-2771.

Shuin, T., Kondo, K., Torigoe, S., Kishida, T., Kubota, Y., Hosaka, M., Nagashima, Y., Kitamura, H., Latif, F., Zbar, B., and et al. (1994). Frequent somatic mutations and loss of heterozygosity of the von Hippel-Lindau tumor suppressor gene in primary human renal cell carcinomas. Cancer Res *54*, 2852-5.

Siemeister, G., Weindel, K., Mohrs, K., Barleon, B., Martiny-Baron, G., and Marme, D. (1996). Reversion of deregulated expression of vascular endothelial growth factor in human renal carcinoma cells by von Hippel-Lindau tumor suppressor protein. Cancer Res *56*, 2299-301.

Smith, D. F. (1993). Dynamics of heat shock protein 90-progesterone receptor binding and the disactivation loop model for steroid receptor complexes. Mol. Endocrinol. *7*, 1418-1429.

Smith, D. F., Whitesell, L., Nair, S. C., Chen, S. Y., Prapapanich, V., and Rimerman, R. A. (1995). Progesterone receptor structure and function altered by geldanamycin, an hsp90-binding agent. Mol. Cell Biol. *15*, 6804-6812.

Stancato, L. F., Chow, Y. H., Hutchison, K. A., Perdew, G. H., Jove, R., and Pratt, W. B. (1993). Raf exists in a native heterocomplex with hsp90 and p50 that can be reconstituted in a cell-free system. J. Biol. Chem. *268*, 21711-21716.

Starr, R., Willson, T. A., Viney, E. M., Murray, L. J., Rayner, J. R., Jenkins, B. J., Gonda, T. J., Alexander, W. S., Metcalf, D., Nicola, N. A., and Hilton, D. J. (1997). A family of cytokine-inducible inhibitors of signalling. Nature *387*, 917-21.

Stebbins, C.E., Russo, A.A., Schneider, C., Rosen, N., Hartl, F.U., and Pavletich, N.P. (1997). Crystal Structure of an Hsp90-Geldanamycin Complex: Targeting of a Protein Chaperone by an Antitumor Agent." Cell *89*, 239-250.

Stebbins, C.E., Kaelin, W.G., and Pavletich, N.P. (1999). "Structure of the VHL-ElonginC-ElonginB complex: implications for VHL tumor suppressor function." Science, *284*, 455-461.

Stepanova, L., Leng, X., Parker, S. B., and Harper, J. W. (1996). Mammalian p50Cdc37 is a protein kinase-targeting subunit of Hsp90 that binds and stabilizes Cdk4. Genes Dev. *10*, 1491-1502.

Stout, G.H., and Jensen, L.H. (1989). X-ray structure determination: a practical guide (second ed.). New York: John Wiley and Sons.

Sullivan, W. P., and Toft, D. O. (1993). Mutational analysis of hsp90 binding to the progesterone receptor. J. Biol. Chem. *268*, 20373-20379.

Sullivan, W., Stensgard, B., Caucott, G., Bartha, B., McMahon, N., Alnemri, E.S., Litwack, G., and Toft, D.O. (1997). Nucleotides and two functional states of Hsp90. J. biol. Chem. *272*, 8007-8012.

Supko, J. G., Hickman, R. L., Grever, M. R., and Malspeis, L. (1995). Preclinical pharmacologic evaluation of geldanamycin as an antitumor agent. Cancer Chemother. Pharmacol. *36*, 305-315.

Takagi, Y., Conaway, R. C., and Conaway, J. W. (1996). Characterization of ElonginC functional domains required for interaction with ElonginB and activation of elongin A. J Biol Chem *271*, 25562-8.

Takagi, Y., Pause, A., Conaway, R. C., and Conaway, J. W. (1997). Identification of ElonginC sequences required for interaction with the von Hippel-Lindau tumor suppressor protein. J Biol Chem *272*, 27444-9.

Tronick, S. R., and Aaronson, S. A. (1995). Growth factors and signal transduction. In: The molecular basis of cancer, J. Mendelsohn, P. Howley, M. Israel and L. Liotta, eds., pp. 117-140.

Vijay-Kumar, S., Bugg, C. E., Wilkinson, K. D., and Cook, W. J. (1985). Three-dimensional structure of ubiquitin at 2.8 A resolution. Proc Natl Acad Sci U S A *82*, 3582-5.

Vogelstein, B. and Kinzler, K.W. (1993). The multistep nature of cancer. Trends in Genetics, *9*(4), 138-141.

Wang, B.-C. (1985) Resolution of phase ambiguity in macromolecular crystallography. Meth. in Enzymol. *115*, 90-112.

Wearsch, P. A., and Nicchitta, C. V. (1996). Endoplasmic reticulum chaperone GRP94 subunit assembly is regulated through a defined oligomerization domain. Biochemistry *35*, 16760-16769.

Weinberg, R.A. (1996) How Cancer Arises. Scientific American, 62-70.

Whaley, J. M., Naglich, J., Gelbert, L., Hsia, Y. E., Lamiell, J. M., Green, J. S., Collins, D., Neumann, H. P., Laidlaw, J., Li, F. P., and et al. (1994). Germ-line mutations in the von Hippel-Lindau tumor-suppressor gene are similar to somatic von Hippel-Lindau aberrations in sporadic renal cell carcinoma [published erratum appears in Am J Hum Genet 1995 Jan;56(1):356]. Am J Hum Genet *55*, 1092-102.

Wiech, H. Buchner, J., Zimmermann, R., and Jakob, U. (1992). Hsp90 chaperones protein folding *in vitro*. Nature *358*, 169-170.

Wiech, H. Buchner, J., Zimmermann, M.,Zimmermann, R., and Jakob, U. (1993). Hsc70, immunoglobin heavy-chain binding-protein, and Hsp90 differ in their ability to stimulate transport of precursor proteins into mammalian microsomes. J. Biol. Chem. *268*, 7414-74221.

Whitby, F., Gang, X., Pickart, C., and Hill, C. (1998). Crystal Structure of the Human Ubiqtuin-like Protein NEDD8 and Interactions with Ubiquitin Pathway Enzymes. J Biol Chem *273*, 34983-34991.

Whitelaw, M. L., Hutchison, K., and Perdew, G. H. (1991). A 50-kDa cytosolic protein complexed with the 90-kDa heat shock protein (hsp90) is the same protein complexed with pp60v-src hsp90 in cells transformed by the Rous sarcoma virus. J. Biol. Chem. *266*, 16436-16440.

Whitesell, L., and Cook, P. (1996). Stable and specific binding of heat shock protein 90 by geldanamycin disrupts glucocorticoid receptor function in intact cells. Mol. Endocrinol. *10*, 705-712.

Whitesell, L., Mimnaugh, E. G., De Costa, B., Myers, C. E., and Neckers, L. M. (1994). Inhibition of heat shock protein HSP90-pp60$^{v\text{-}src}$ heteroprotein complex formation by benzoquinone ansamycins: essential role for stress proteins in oncogenic transformation. Proc. Natl. Acad. Sci. U S A *91*, 8324-8328.

Wigley, D.B., Davies, G.J., Dodson, E.J., Maxwell, A., and Dodson, G. (1991). Crystal structure of an N-terminal fragment of the DNA gyrase B protein. Nature *351*, 624-629.

Xu, Y., and Lindquist, S. (1993). Heat-shock protein hsp90 governs the activity of pp60v-src kinase. Proc. Natl. Acad. Sci. U S A *90*, 7074-7078.

Zhu, X., Zhao, X., Burkholder, W. F., Gragerov, A., Ogata, C. M., Gottesman, M. E., and Hendrickson, W. A. (1996). Structural analysis of substrate binding by the molecular chaperone DnaK. Science *272*, 1606-1614.