Heterologous expression of laminin peptide 11 on a virus particle surface for use in malignant tumor cell targeting
by Thomas Darmody Arnold

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Biochemistry
Montana State University
© Copyright by Thomas Darmody Arnold (2000)

Abstract:
This thesis details the construction of a chimeric tumor-targeted bio-imaging agent. There are two major components to this system: 1) the tumor-targeting ligand, which directs a host scaffold (macromolecule) to tumor tissues, and 2) the "host" scaffold, which carries imaging agents and displays, on its surface, the tumor-targeting ligand. Peptide 11 was chosen as the tumor-targeting ligand of the chimera. Peptide 11 (CDPGYIGSR), is a nine amino acid sequence derived from the β1 chain of the basement membrane molecule, laminin-1. This peptide is the major ligand binding domain for the 67 kDa laminin binding protein (LBP). The expression level of LBP shows a positive correlation with progression in many solid tumors. Synthetic peptide 11 blocks tumor lung colonization and tumor angiogenesis. Radio-labeled derivatives of peptide 11 have target and image tumor cells that overexpress LBP. We have chosen the protein capsid of Cowpea chlorotic mottle virus (CCMV) to serve as a macromolecular scaffold from which to present peptide 11. The CCMV scaffold has been chosen for its resilient and dynamic structural attributes and for its ability to bind large numbers of the magnetic resonance imaging contrast agent, gadolinium (Gd), on its surface. In theory, the CCMV-peptide 11 chimera will target multiple solvent exposed Gd ions to metastatic tumor cells expressing LBP on or near their surface. The specific aims of this thesis project are: 1) To design. an appropriate system for the construction of a protein engineered tumor-targeted bio-imaging agent; and 2) To genetically modify the coat protein subunit of CCMV to express exogenous peptide 11 at one of five possible exterior loop positions. The peptide 11 sequence has been inserted into four of the five surface loops of the viral coat protein and we have demonstrated stable virus assembly under various conditions. Systemic infection and efficient production of chimeras from plants appears to be compromised in some viral constructs. We are therefore using a novel yeast expression system to provide high concentration of empty capsid particles which can be loaded with contrast/therapeutic agents. Chimeras purified from yeast look and behave like native CCMV virions. The ability of these chimeras to target LBP, and eventually, to image metastatic tumor cells, will be investigated by others in this research group.

HETEROLOGOUS EXPRESSION OF LAMININ PEPTIDE 11 ON A VIRUS PARTICLE SURFACE FOR USE IN MALIGNANT TUMOR CELL TARGETING

by

Thomas Darmody Arnold

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Biochemistry

MONTANA STATE UNIVERSITY
Bozeman, Montana

July 2000

N378
Ar661

APPROVAL

of a thesis submitted by

Thomas Darmody Arnold

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

Dr. Martin Teintze _____ (Signature) 7/24/00 Date

Approved for the Department of Biochemistry

Dr. Paul Grieco _____ (Signature) 7-24-00 Date

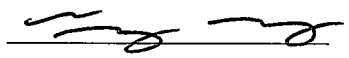
Approved for the College of Graduate Studies

Dr. Bruce McLeod _____ (Signature) 7-24-00 Date

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library.

If I have indicated my intention to copyright this thesis by including a copyright notice page, copying is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for permission for extended quotation from or reproduction of this thesis in whole or in parts may be granted only by the copyright holder.

Signature 

Date 7-17-00

ACKNOWLEDGMENTS

I would like to express my thanks and appreciation to a number of people:

To my advisor, Dr. Mark Young, for broadening my scientific perspective and for supporting me in my academic, professional, and personal aspirations. Also, to the other members of my committee: Drs. Jean Starkey, Martin Tientze, and Martin Lawrence, for their guidance and support during my graduate studies these past two years.

To the Mark Young "labbies" past and present: Debbie, Sue, Lina, George, Ken, Jamie, Chris, Zach, Blake, Linda, Carlie, Ellie, Annie, Ann, and Wan, for all your help, and all the laughs.

To the Microbiology Department at MSU, Bozeman, for "taking me under your wing" this past year and a half.

And to my family: Bill, for your support and friendship and for getting me to come to Bozeman in the first place; Randy and Sandy, for your love and for giving me a great place to relax; Dorian and Mark, for your caring and your humor; Mom and Dad, for supporting all my wacky ideas and for your unconditional love.

TABLE OF CONTENTS

LIST OF FIGURES VII

ABSTRACT IX

1. INTRODUCTION 1

2. BACKGROUND AND PROJECT RATIONAL 3

INTRODUCTION 3
A TWO COMPONENT CHIMERA: PEPTIDE 11 PLUS CCMV 4
Why peptide 11? 4
The 67-kDa Laminin Binding Protein (LBP) 5
Peptide 11 9
Uses for Peptide 11 10
Why CCMV? 12
THE DESIGN CYCLE: DESIGN, CREATE AND TEST 23
SUMMARY 27

3. METHODS 29

DESIGN OF CCMV-PEPTIDE 11 CHIMERAS 29
Structure and Sequence Analysis 29
Selection of Insertion Sites 29
B-Factor Mapping 29
Surface Accessibility Mapping 30
ENGINEERING CCMV-PEPTIDE 11 CHIMERAS (PLANTS) 31
Construction of Loop-Specific BamH1 Cloning Sites 31
Insertion of Peptide 11 Sequence into Loop Positions 31
In Vitro Transcription for Plant Inoculation 35
Plant Inoculation and Virus Propagation 35
Virus Purification from Plants 36
Virus Characterization from Plants 37
ELISA 37
RT-PCR Amplification and Restriction Digestion 38
SDS-PAGE and Western Blot Analysis 39
ENGINEERING CCMV-PEPTIDE 11 CHIMERAS (*P. PASTORIS*) 40
Construction of pPicZA/pep11 Plasmids 40
Transformation into *P. Pastoris* 42
Small Scale Time Course Expression Screening 43
Intermediate-Scale Expression of CCMV-Peptide 11 Coat Protein in *P. pastoris* 44
Electron Microscopy 45
Expression and Purification of CCMV-Peptide 11 Chimeras from High-Density Fermenter Cultures 46
Glycerol Batch Phase 46

Glycerol-Fed Batch Phase 47
Methanol-Fed Batch Phase 47
Lysis of Cells and Sample Preparation 48
Sample Purification and Virion Characterization 49

4. RESULTS AND DISCUSSION 50

DESIGN OF CCMV-PEPTIDE 11 CHIMERAS 50
Structure and Sequence Analysis 50
Selection of Insertion Sites 54
ENGINEERING CCMV-PEPTIDE 11 CHIMERAS (PLANTS) 59
Insertion of Peptide 11 Sequence into Loop Positions 59
Expression of CCMV-Peptide 11 Chimeras in Plants 65
ENGINEERING CCMV-PEPTIDE 11 CHIMERAS (*P. PASTORIS*) 72
Construction of pPicZA/pep11 Plasmids and *P. pastoris* Transformation 73
Small Scale Time Course Expression Screening 77
Intermediate Scale Expression of CCMV-Peptide 11 Coat Protein in *P. pastoris* ... 81
Expression, Purification and Characterization of CCMV-Peptide 11 Chimeras from High-Density Fermenter Cultures 83

5. CONCLUSIONS 93

SUMMARY OF RESULTS 94
FURTHER STUDIES 101
CONCLUSIONS 106

REFERENCES CITED 108

APPENDICES 115

APPENDIX A - 1CWP: RESIDUE WISE INTERACTION ENERGIES 116
APPENDIX B - SEQUENCING DATA 132
APPENDIX C - THE CANCER PROBLEM 134

LIST OF FIGURES

Figure Page

1. Cartoon depicting the basement membrane molecule, laminin-1 and the peptide 11 sequence found within the β1 chain....8

2. A modeled structure of peptide 11....10

3. The 3D crystal structure of the adenovirus (Ad5) fiber knob protein....15

4. Shapes and structures common to viruses....18

5. Structural representations of CCMV. 22

6. Structural comparison of the CCMV capsid subunit to the Polyoma virus VP1 capsid subunit....52

7. Structural comparison between C subunits of Cucumber mosaic virus (CMV) and CCMV....53

8. B-factor mapping of CCMV coat protein capsomers....55

9. Accessibility mapping of CCMV coat protein capsomers....57

10. CCMV coat protein with loop insert sites (amino acid positions) colored....58

11. Schematic representation of plasmid-based vectors used as templates for making infectious RNA....61-62

12. Coupled *in vitro* transcription/translation from pCC3/pep11 constructs....64

13. Examples of RNA used to inoculate plants 68

14. ELISA results from plants 69

15. Western blot of coat protein products from plants....70

16. RT-PCR amplification and diagnostic restriction digestion....71

17. Schematic representation of pPicZA/pep11 construction....74

LIST OF FIGURES (CONTINUED)

Figure Page

18. Screening for coat protein plus peptide 11 insertion into pPicZA.....75

19. Lyticase PCR-based screening (integration into *P. pastoris* genome).....76

20. Small scale time course expression screening.....79

21. Small scale time course expression screening (Western blots).....80

22. Intermediate scale expression culture of CCMV-Peptide 11 Coat Protein in *P. Pastoris*.....82

23. Isolation of chimeras by PEG precipitation.....86

24. Isolation of chimeras by ultracentrifugation.....89

25. CsCl purification of chimeric capsids (loops 129 and 63).....92

26. (Appendix C) Cancer tissue has a distinctive appearance under the microscope.....140

27. (Appendix C) Some prefixes used in naming cancers.....141

ABSTRACT

This thesis details the construction of a chimeric tumor-targeted bio-imaging agent. There are two major components to this system: 1) the tumor-targeting ligand, which directs a host scaffold (macromolecule) to tumor tissues, and 2) the "host" scaffold, which carries imaging agents and displays, on its surface, the tumor-targeting ligand. Peptide 11 was chosen as the tumor-targeting ligand of the chimera. Peptide 11 (CDPGYIGSR), is a nine amino acid sequence derived from the β1 chain of the basement membrane molecule, laminin-1. This peptide is the major ligand binding domain for the 67 kDa laminin binding protein (LBP). The expression level of LBP shows a positive correlation with progression in many solid tumors. Synthetic peptide 11 blocks tumor lung colonization and tumor angiogenesis. Radio-labeled derivatives of peptide 11 have target and image tumor cells that overexpress LBP. We have chosen the protein capsid of Cowpea chlorotic mottle virus (CCMV) to serve as a macromolecular scaffold from which to present peptide 11. The CCMV scaffold has been chosen for its resilient and dynamic structural attributes and for its ability to bind large numbers of the magnetic resonance imaging contrast agent, gadolinium (Gd), on its surface. In theory, the CCMV-peptide 11 chimera will target multiple solvent exposed Gd ions to metastatic tumor cells expressing LBP on or near their surface. The specific aims of this thesis project are: 1) *To design an appropriate system for the construction of a protein engineered tumor-targeted bio-imaging agent; and 2) To genetically modify the coat protein subunit of CCMV to express exogenous peptide 11 at one of five possible exterior loop positions.* The peptide 11 sequence has been inserted into four of the five surface loops of the viral coat protein and we have demonstrated stable virus assembly under various conditions. Systemic infection and efficient production of chimeras from plants appears to be compromised in some viral constructs. We are therefore using a novel yeast expression system to provide high concentration of empty capsid particles which can be loaded with contrast/therapeutic agents. Chimeras purified from yeast look and behave like native CCMV virions. The ability of these chimeras to target LBP, and eventually, to image metastatic tumor cells, will be investigated by others in this research group.

CHAPTER 1

INTRODUCTION

The ***overall goal*** of this thesis project is to begin development of a tumor-targeted bio-imaging system based on a plant virus scaffold (Cowpea chlorotic mottle virus - CCMV) – peptide 11 chimera. The CCMV capsid scaffold has been chosen for its resilient and dynamic structural attributes and for its unique ability to bind large numbers of gadolinium (Gd) ions (up to 180 ions per capsid). Peptide 11 was chosen as the tumor-targeting ligand of the chimera because it has been shown to specifically bind LBP shed from and/or on the surface of certain metastatic tumor cells. Although appropriate collaborations have already been arranged to evaluate the effectiveness of our constructs for bio-imaging applications, this thesis is limited to the their design, construction, and preliminary characterization.

The specific aims of this thesis project are:

1) *To design an appropriate system for the construction of a protein engineered tumor-targeted bio-imaging agent.* Structural elements (CCMV template scaffold and peptide 11) will be evaluated in a "rational approach" to designing chimeric CCMV-peptide11 capsid constructs.

2) *To genetically modify the coat protein subunit of Cowpea chlorotic mottle virus (CCMV) to express exogenous peptide 11 at one of five possible exterior loop positions.* Standard cloning techniques will be used to insert a peptide 11-based sequence into CCMV genomic cDNA. A variety of heterologous and homologous expression systems will be evaluated for their ability to generate the modified virion in a stable, assembled form. Assembled virions will be biochemically characterized.

CHAPTER 2

BACKGROUND AND PROJECT RATIONAL

Introduction

This thesis represents our *initial efforts* to design, engineer and test the potential for viral protein cages to serve as scaffolds for targeted delivery of cancer diagnostic agents. My thesis research details the construction of such an agent. The goal of this thesis background section is to introduce the two fundamental components that we used to create a protein scaffold-tumor ligand chimera and to briefly discuss why we chose these components. This is important since the goals of my thesis research (see Chapter 1, Introduction) originated from such a conceptual foundation. In the final section of this review, I will summarize the strategies that we used to create our tumor-targeted bio-imaging agent, and the criteria that we will use in evaluating its first forms. For a more in-depth look at "the cancer problem" that underlies my motivations for attempting this thesis, please see Appendix C.

A Two Component Chimera: Peptide 11 Plus CCMV

There are at least two fundamental components that make up a tumor-targeted macromolecular bio-imaging agent: 1) the tumor-targeting ligand, which directs the host scaffold to tumor tissues, and 2) a "host" scaffold, which carries imaging agents and displays, on its surface, the tumor-targeting ligand. We have chosen the protein capsid of Cowpea chlorotic mottle virus (CCMV) to serve as a macromolecular host and the LBP-seeking peptide 11, as its tumor targeting ligand. In theory, the CCMV capsid will target multiple solvent exposed Gd ions (MRI contrast agent) to metastatic tumor cells expressing LBP on or near their surface. Here, I will briefly review the two major components of our agent.

Why Peptide 11?:

Peptide 11 is the major ligand for the tumor cell surface receptor, laminin binding protein (LBP). In order to adequately discuss peptide 11 as a tumor-targeting ligand, its cognate receptor will first be discussed.

The 67-kDa Laminin Binding Protein (LBP). The 67 kDa LBP is a major nonintegrin cell surface receptor that mediates high-affinity interactions between cells and the basement membrane molecule, laminin-1. In 1983, two research groups independently reported the discovery of LBP, separated from extracellular matrix on laminin-Sepharose columns[1,2]. This protein was found to be expressed by a variety of mammalian cells and to be *particularly abundant on the surface of cancer cells*[3]. Several studies to elucidate this receptor's role in tumor progression have clearly demonstrated an increase in LBP expression in tumors compared with normal tissues[4,5], a correlation between LBP expression, invasive phenotype of the tumor[6,7], and poor clinical prognosis[8,9]. High expression of LBP is also associated with the invasive phenotype of normal embryonic trophoblastic tissues[10], migratory inflammatory cells[11], and embryonic cells during gastrulation[12]. The strong correlation between LBP overexpression and the invasiveness of certain cells in non-cancer situations and the metastatic potential of tumor cells suggests that this receptor plays an important role in the acquisition of a metastatic phenotype by cancer cells. (For a more thorough discussion on metastasis, please see Appendix C.) High expression of LBP is therefore a good biological marker of metastatic cells.

In an effort to characterize LBP, Rao *et al*, in 1989, isolated full-length cDNA clones from a human expression library using monoclonal antibodies against the purified 67 kDa protein[13]. Predicted molecular mass and *in vitro* translation of the selectively hybridized mRNA identified a protein of 32 kDa. Previous pulse-chase experiments had predicted this apparent LBP precursor protein[14], later termed laminin receptor precursor

(LRP). Posttranslational modification of this 37 kDa LRP including fatty acid acylation by palmitate, oleate and stearate may lead to dimerization of LRP, and provide a mechanism for the formation and plasma membrane association of functional LBP[15]. Interestingly, LRP is, by itself, associated with polysomes (and no cell surface localization) in oocytes and sea urchin embryos[16]. A homologous mouse protein, p40, is specifically associated with the 40S ribosomal subunit[17]. Moreover, duplicate protein homologues in yeast (YST1 and YST2) have been identified as essential protein components of the 40S ribosomal subunit. The assembly of these proteins into immature 40S subunits and the subsequent processing of 20S rRNA to 18s rRNA (through protein or catalytic RNA modifications) represent critical steps in defining the translational capacity of yeast cells[18]. Taken together, these results suggest a dual role for LRP in carcinogenesis: It may be involved in maintaining the translational capacity of cells (overexpression of the precursor protein in tumors could function to assure an adequate supply of active 40S subunits to meet the demands for protein synthesis during tumor cell growth and proliferation). As LRP dimerizes and is exported to the cell surface it then acts as a laminin binding protein (and extracellular matrix modulator). Recent phylogenetic analysis indicates that the laminin-binding function (specifically, the LMWWML palindrome from peptide G - a purported active site for laminin binding[19,20]) was acquired during evolution, around the same time that laminin-related molecules came to exist in nature[17].

Although the exact mechanism of LBP's action during tumor cell invasion and metastasis is not clear, recent data from the Starkey laboratory (Montana State

University, Bozeman, MT) may shed some light on this issue. In a recent study[21], chinese hamster ovary (CHO) cells were stably co-transfected with expression vectors for dihydrofolate reductase (dhfr), a full-length human β1 integrin chain, and a full-length α6 integrin chain. Growth in methotrexate (MTX)-containing medium resulted in amplification of both α6 and β1 expression products, which dimerized and then localized at the cell surface. The α6β1 integrin is a laminin-specific receptor that also appears to promote tumor cell migration, basement membrane matrix invasion, and metastasis[22] (and references therein). Upon MTX amplification and co-overexpression, subsequent surface expression of α6β1 integrin was increased. Endogenous LBP, under autonomous regulation in the cell (LBP expression appears to be regulated at the level of translation[22]), was concomitantly co-overexpressed with the laminin-specific integrin. Moreover, a significant amount of substrate-associated LBP (in addition to the LBP that was localized to the cell surface near focal adhesion plaques) was observed. This matrix-associated LBP may have been shed into the culture medium from tumor cells in response to the cells' interaction with matrix (or matrix components) via integrin mediated signaling. LBP shedding occurs in a variety of tumor cell lines and may be proportional to the invasive ability of the cells. It therefore seems likely that upregulation, and then induced shedding, of LBP promotes aggressive malignant behavior in solid tumors. The shed-LBP retains its ability to bind laminin[23] and is active in its ability to modify the basement membrane. Surface expressed and shed-LBP both appear to have protein sulfhydryl oxidase (SOX) catalytic activity[24]. Recent analysis demonstrating high similarity between the N-terminal region of LRP and various members of the chaperonin-

like HSP70 family lend support to the hypothesis that LBP could actively rearrange basement membrane laminin-1[19]. A possible model for LBP's role in invasion and metastasis might therefore include both tumor cell attachment via surface LBP-laminin interactions as well as an active reorganization of the basement membrane to facilitate cellular invasion.

The major ligand for LBP is the noncollagenous basement membrane component, laminin-1 (Fig 1). The laminin glycoprotein is composed of three chains, α_1 (400 kDa), β_1 (230 kDa), and γ_1 (220 kDa), arranged in a cruciform-like structure (as imaged by electron microscopy), with one long arm and three short arms each containing globular end regions [25,26]. The carbohydrate composition of laminin-1 is heterogeneous and specific sugar moieties are enriched on the globular end regions[23]. LBP exhibits a relatively high affinity for laminin-1 with a published K_d on the order of 2 nM[27].

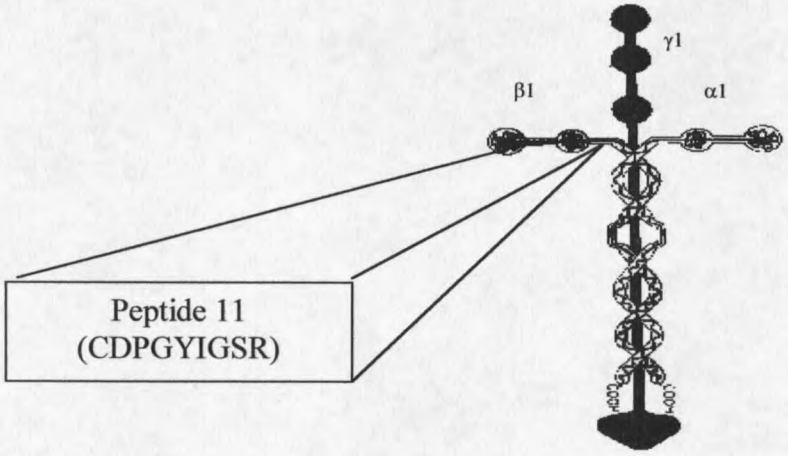


Fig 1. Cartoon depicting the basement membrane molecule, laminin-1 and the peptide 11 sequence found within the β1 chain.

Peptide 11. The nine amino acid sequence, CDPGYIGSR, known as peptide 11 (Fig. 1), is the primary ligand binding domain for the 67 kDa LBP. Peptide 11 is located within epidermal growth factor (EGF)-like repeats (LE repeats) in domain III of the laminin β_1 chain (residues 925-933)[23,26-28]. Free peptide 11, within purified protease-derived fragments of laminin or from direct synthesis of the small molecule, effectively blocks invasion of basement membranes by tumor cells[26,28,29], reduces experimental tumor lung colonization (metastasis)[28,29], and inhibits tumor angiogenesis[30]. Peptide 11 specifically elutes phage populations displaying mimotopes for three different regions of LBP (including the peptide G palindrome mentioned above) from a nine-mer random phage display library[20]. Also, a synthetic UV-light-activatable crosslinker based on the peptide 11 sequence (called the peptide 11 photoprobe) binds to and cross-links with surface and shed LBP[21]. This research demonstrates peptide 11's specificity for LBP, a molecule heavily expressed on the surface of tumor cells.

Alanine scan substitution experiments indicate that Tyr^5, Ile^6 and Arg^9 in peptide 11 contribute significantly to the anti-invasive activity of synthetic peptide 11. Based on structure:function experiments with D/L alanine for glycine analogues, free active peptide 11 tends to favor an "S"-shaped backbone conformation in solution[31]. Cyclic disulfide peptide11 (*-CDPGYIGSRC-*) is thought to possibly mimic the structure of the peptide 11 sequence found in laminin-1, since the peptide 11 sequence is derived from a disulfide bonded c-loop of an LE repeat. This cyclic isoform appears to adopt an S-shaped conformation[29] similar to the shape of non-cyclic isoforms, and exhibits good anti-invasive and anti-metastatic activity. Structure/function studies indicate that the

active conformation for peptide 11 is one which projects Tyr^5, Ile^6 and Arg^9 sidechains on the same face of the peptide, and extends Tyr^5 away from the compact core of the peptide where it can functionally interact with its LBP receptor (Fig. 2)[25,29,30]. The cyclic form of peptide 11 appears to satisfy these requirements[29]. The tyrosine analogue, 4-benzoyl-L-phenylalanine, in the peptide 11-based photoprobe induces receptor-specific crosslinking, further implicating an active role for Tyr^5 in binding LBP[21].

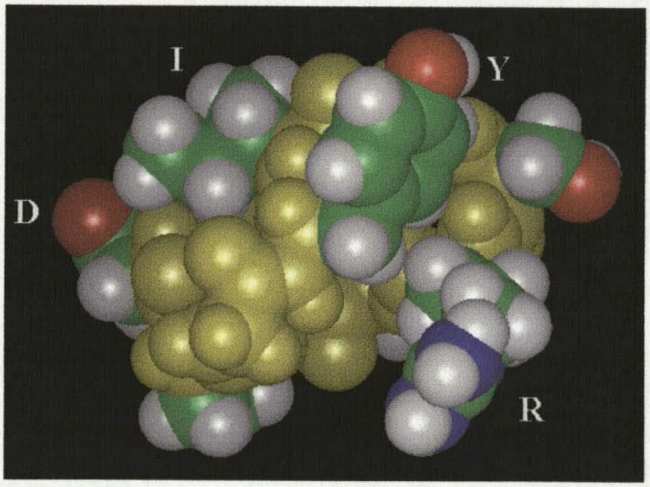


Fig.2. A modeled structure of peptide 11. Notice the compact core and Tyr^5 (Y) projecting from this core (out of the page).

Uses for Peptide 11. Recent advances in elucidating the structure, function, extracellular matrix binding, and intracellular signaling involved in metastatic/angiogenic pathways have led to new and exciting modalities for cancer therapeutics and diagnosis (see Appendix C). Broad-based pharmaceutical efforts have focused on developing inhibitors of matrix metalloproteinases (MMPs) in an effort to interfere with angiogenesis

and metastasis (see Appendix C). The adhesion polypeptides, YIGSR and RGD (representing the minimal sites of cell binding domain of basement membrane components like fibronectin, laminin-1, and collagen) have been found to adhere to integrins ($\alpha1\beta1$, $\alpha2\beta1$, $\alpha3\beta1$, $\alpha5\beta1$, $\alpha6\beta1$ and $\alpha7\beta1$) and the 67kDa laminin binding protein on the surface of cancer cells, and block cell functions such as spreading and motility. These peptides can also inhibit tumor angiogenesis by interfering with the functions of similar receptors on endothelial cells. These peptides show promise in inhibiting the binding of cancer cells to basement membrane, thereby preventing the dissemination of certain cancers[32].

Because of its biological activity, peptide 11 is an obvious candidate for small-molecule-based therapy of solid tumors, specifically to inhibit metastasis. However, because of its susceptibility to proteolytic degradation in physiological environments, free synthetic peptide 11 is unlikely to perform effectively as a therapeutic. The cyclic form of peptide 11 is likely to be more stable in physiological environments. Still, this peptide as well as other peptide 11 analogues (i.e. partial D-AA conformers, complete D-AA conformers, specific substitutional variants) are either inactive or also quite susceptible to degradation. Small therapeutic peptides similar to peptide 11 have been injected intraperitoneally to prevent the dissemination of gastric cancer. Apparently, these peptides have an increased half-life (relative to serum half-life) in the peritoneal cavity due to their high molecular weight. Alternatively, to increase peptide 11's stability while retaining its biological attributes, various synthetic peptide 11 mimetics are currently being explored. Another alternative is to stabilize peptide 11 on a robust

platform. The core of this thesis is the creation of such a peptide 11 chimera system. We have chosen to use a plant virus, CCMV, as a protein engineered scaffold to display and stabilize peptide 11 for *bio-imaging* purposes. There are also potential therapeutic applications for such a construct, but these will not be discussed here.

Why CCMV?

Many small (9–40 amino acids in length) synthetic peptides, have a broad spectrum of uses ranging from antigens for epitope-based vaccines to specific inhibitors of cell surface interactions to targeted imaging. The efficacy of such peptides is generally increased by attaching them to a carrier protein in order to extend their lifetime in physiological environments and to present them in a manner in which they will be better recognized by their biological target[41-49]. An effective way to "attach" a polypeptide to a protein is to genetically insert the appropriate oligonucleotide into a protein-encoding gene and then to express the altered gene as a fusion protein[41,46,49]. It is possible to engineer the fusion between a particular peptide and a viral capsid protein, which then can assemble into a macromolecular scaffolding structure[46]. The coat/capsid proteins from a variety of both plant and animal viruses have been used in this way[44-49]. Plant virus-based scaffolds have been mostly limited to foreign antigen display for use as vaccines. These epitope-display chimeras can be conceptually extended to include the presentation of receptor ligands (discussed below).

There are numerous examples of mammalian viruses that have been modified to carry an inserted peptide sequence on their surface, targeting their cargo (usually therapeutic genes) to tissues of interest. Tissue-targeted adenovirus-based vectors are a case in point. While these vectors would probably not be used as scaffold to present peptide 11 for targeted bio-imaging purposes (the long-term focus of this thesis report), the methods used to create them can be applied to the research presented here. Two general approaches have been taken to create targeted adenoviral vectors [33].

> In the 'two component' approach a bi-specific molecule [usually an antibody-chimera with two distinct specificities] is complexed with the adenovirus. The bispecific component simultaneously blocks native receptor binding and redirects virus binding to a tissue-specific receptor. In the 'one-component' approach the adenovirus is genetically modified to ablate native receptor interactions and a novel ligand is genetically incorporated into one of the adenovirus coat proteins[51].

Two component systems offer greater flexibility in rapidly validating the feasibility of targeting via a particular receptor. Drawbacks to this method include the additional production and characterization of bispecific antibodies, the potential clearance of virus by Fc receptors, and potential activation of the complement system. One-component systems offer the best advantages in producing a manufacturable therapeutic and may offer the most direct way (without the need for extra components) to target an adenoviral vector and, by analogy, other virus-based vectors, to specific ligands.

There have been several successful attempts at creating one-component, targeted adenoviral vectors. First generation vectors were created by incorporating different peptide motifs into the C-terminus of the viral fiber protein. The fiber protein, composed of three copies of a single protein, projects from the surface of the icosahedral adenovirus

capsid particle. The N-termini of the three fiber proteins are embedded in the "penton base". The C-termini contribute to the formation of a terminal fiber "knob". A potential advantage to making C-terminal fiber fusions (as opposed to internal insertions) is that there may be less likelihood of the insert placing structural constraints on macromolecular assembly. Terminal fusions may also represent a choice of convenience when high resolution three dimensional structures are not available to rationally consider internal insertions. In a study by Wichkham *et al*[34], several peptides, including the peptide 11-based motif, $(GYIGSR)_3$, were evaluated in terms of their ability to be expressed as C-terminal fusion proteins that are generally nondisruptive to viral functioning, their ability to target the virus to specific receptors. Two peptide insertions constructs were considered successful by these criteria, the integrin receptor targeting sequence, ACDCRGDCFGG, and the heparan sulfate targeting sequence, KKKKKKK. Other less successful insertion peptides, including the peptide 11-based sequence, were considered "incompatible with the correct folding of the fiber protein". Unsuccessful constructs encoded C-terminal additions that were five or six amino acids *shorter* than the two longest constructs that were successfully made. The authors suggest that the "targeting peptide composition, rather than absolute peptide length, is most critical for making short peptide additions to the [coat protein C-terminus]."

Recently, the 3D crystal structure of an adenovirus fiber knob (Fig. 3) was solved[35]. This structure enabled researchers to design chimeras using a more rational approach than was previously possible[36]. The fiber protein, like the coat proteins from many viruses (see below), has a core β-barrel fold with surface exposed loops. It

assembles into a trimer with three fold symmetry. Structural analysis of the fiber knob reveals one exterior surface loop, the HI loop, with a number of features predicting its utility for ligand incorporation.

> Specifically, the HI loop does not contribute to intramolecular interactions in the knob [fiber trimerization]. The loop consists of mostly hydrophilic amino acid residues, [with a] high degree of flexibility, [that are] exposed outside the knob. Lengths of HI loops vary significantly in knobs of different adenovirus serotypes, [suggesting] that alterations of the original structure of the loop, such as insertions and deletions, should be compatible with the correct folding of the entire knob domain[36].



Fig 3. The 3D crystal structure of the adenovirus (Ad5) fiber knob protein. The HI loop, exposed outside the knob, connects the β-strands H and I, which are involved in the formation of the cell-binding site. (Rendered using atomic coordinates obtained from the Brookhaven Protein Data Base (PDB ID 1KNB) constructed with SwissPDBViewer (v3.63b.)

A variety of cell-targeting peptides have been successfully inserted into the HI loop of the fiber protein. The first generation of these chimeras were composed of HA (AYPYDVPDYA)[37] or FLAG (DYKDDDDK)[36] epitope insertions. These epitopes do not target biologically important ligands. However, high-affinity monoclonal antibodies to these epitopes are readily available since these epitopes are commonly used to "tag" proteins to which they are heterologously fused. The HA and FLAG sequences were therefore chosen to establish a first stage proof of concept: Can a short peptide be expressed on the surface of the viral protein that will target the virus to cellular receptors? These epitopes were proven successful in this regard. The fiber protein correctly folded with these sequences inserted. Moreover, the HA epitope specifically targeted viable virus to cells expressing an HA "pseudo-receptor" (a membrane-anchored single-chain antibody to HA)[37]. Since the production of these first-generation chimeras, several chimeras have been made with more biologically important ligands, including a tumor-targeting "RGD" motif [38], inserted in the HI loop of the fiber protein. For a review of these, see reference 33. Research investigating the use of adenoviruses to present cell-targeting ligands is important to this thesis, not because this vector provides a potential scaffold to present peptide 11, but because it represents successful "rational design" approaches used to create cell-targeted virus-based chimeras. This work is also a good example of how a researcher might exploit viruses, agents that are commonly associated with the diseases they can cause, in an effort to achieve novel health-improving functionalities that might otherwise be difficult to create *de novo*. These systems generally take advantage of the innate attributes of virus structure and function.

Virions are nucleoprotein particles designed to transfer their genomes between cells of a susceptible host and between hosts. The structure of the protein capsid that contains the viral genome is intimately linked with these core functions. In particular, capsid protein interactions must be sufficiently stable to protect the packaged nucleic acid genome yet be adjustable to first allow its encapsidation and later its release. Evidence for this "capsid structure = viral function" premise comes from the observation that almost all viruses seem to come in two presumably convergent shapes: roughly spherical (usually based on icosahedral symmetry) and rod shaped (usually based on helical symmetry) (Fig. 4). Furthermore, a core motif, the 8-stranded antiparallel β-barrel fold (Fig. 4), seems to be quite common among capsid subunits (coat proteins); this fold is structurally conserved in many icosahedral viruses regardless of whether they infect animals, plants, insects, fungi or bacteria[39]. The dynamic stability of viral protein capsids is also evident in the inherent ability of certain coat proteins to self assemble into the highly symmetrical capsid structures that define them. In all icosahedral viruses, the coat protein subunits assemble into hexameric and pentameric capsomer units. The ratio of hexamers to pentamers determines the curvature, which in turn defines the overall structure (size and shape) of the final virion[40,41]; the higher the hexamer to pentamer ratio, the larger the diameter of the structure. Spherical viruses range in size from 18 to 500 nm, while rod shaped viruses can be as large as 900 nm.

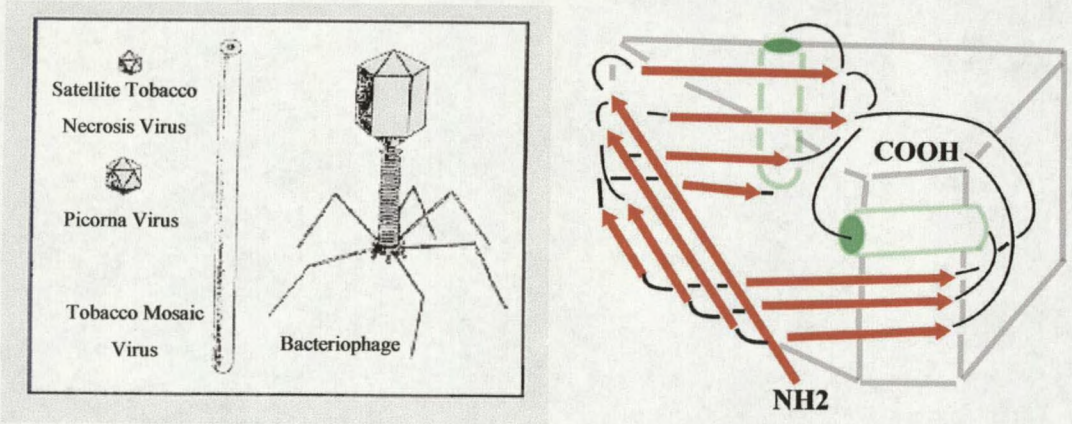


Fig. 4. Shapes and structures common to viruses. Left image: Small genomes and structural constraints force most viruses into convergent shapes: simple spheres, helical rods, and complex structures made of multiple symmetric subunits. Right image: A generalized β-barrel protein fold common to many viral capsid proteins.

Considering the seemingly canonical capsid structure = viral function link, it is surprising that some capsids are remarkably tolerant to sequence/chemical alterations in their composite subunits. Resilient core folds (the β-barrel) and robust capsid architecture can probably be thanked for this apparent flexibility. *By exploiting the dynamic stability and inherent mutability of an icosahedral virus, we aim to create a tumor-targeted bio-imaging protein scaffold.*

The small spherical virus, Cowpea chlorotic mottle virus (CCMV), appears to be an ideal system for developing such a scaffold. CCMV is a member of the bromovirus group of the Bromoviridae (a member of the alpha family supergroup)[42,43]. Bromoviruses are 25-28 nm icosahedral viruses with a four component positive sense

single stranded RNA genome. RNAs 1 and 2 code for replicase enzymes. RNA 3 codes for a protein that is involved in viral movement within its plant host. RNA 4 (a subgenomic RNA derived from RNA 3) codes for the 20-kDa viral coat protein (see Chapter 4, Results and Discussion, Figs. 1 and 2).

Previous studies have demonstrated that the CCMV coat protein can be genetically modified without interfering with its ability to form particles. The coat protein can be modified by making specific insertions/deletions/mutations in the coat protein gene region of cDNA cloned from CCMV genomic RNA3 isolates. *In vitro* transcription of this modified cDNA creates infectious RNA chimeras. Expressed coat protein can self assemble in the absence of nucleic acid under varying pH and ionic conditions[44]. CCMV chimeras can readily be produced in both heterologous (*E. coli* and *Pichia pastoris* yeast) and homologous (plant protoplasts and various plant hosts) systems (see Chapter 3, Methods, and Chapter 4, Results and Discussion). CCMV assembly is remarkably tolerant of alterations, including loop insertions (see below), often maintaining replication at or near wild type levels in plants. Mutant CCMV accumulates to high levels and can easily be purified from these hosts. Assembly of mutant capsids from heterologous systems is also quite reliable. The crystal structure of CCMV has been determined to 3.2 Å resolution[45]. This structure provides a clearer picture of the coat protein interactions that appear to be critical to capsid stability and dynamics and provides an appropriate framework from which to design protein engineering projects (see below). *These attributes make CCMV an ideal system for the virion surface expression of exogenous peptide 11.*

Any attempt at designing novelties into a preexisting protein template requires some understanding of the template (and, of course, an understanding of the new characteristics being incorporated). The 3.2 Å resolution structure of CCMV (Fig. 5)[44] provides a wealth of information helpful in considering our prospective template and is helpful in guiding rational design and experimentation. The virion is made up of 180 quasi-equivalent copies of the coat protein subunit organized in 20 hexameric and 12 pentameric capsomers which make up a T=3 quasi-symmetric capsid (Fig. 5a). Each coat protein subunit (Fig. 5e) is composed of an eight-stranded, anti-parallel β-barrel core motif (formed by amino acids 52-176). Extending in opposite directions from this core are long N-terminal (residues 1-51; residues 1-27 are not ordered in the crystal structure) and C-terminal arms (residues 176-190). These arms appear to provide an intricate framework that solidifies subunit interactions (see ring-like structure in the center of the Hexamer, Fig. 5c). The orientation of the coat protein β-barrel is nearly parallel to the five-fold and quasi six-fold axes resulting in the outward projection of five loops (βB-βC, βD-βE, βF-βG, βE-αEF, and βH-βI; Fig. 5). These loops appear to be especially permissive to sequence alteration/additions. The observation that there is significant sequence and structural variability in similar external loops in viruses closely related to CCMV supports this contention (see Chapter 4, Results and Discussion, Figs. 6-9). Structural analysis, homology modeling and previous protein engineering experiments suggest that certain loops may be more amenable to certain insertions (see Chapter 4, Results and Discussion). Surrounding each of the 60 quasi three-fold axes located on the interface between hexamer and pentamer capsomers are predicted Ca^{2+} binding sites (Fig.

5). There are 180 such Ca^{2+} binding sites per virion, each consisting of five metal coordinating residues (Glu81, Gln85, Glu148 from one subunit; Gln149 and Asp153 from an adjacent subunit). A long-term goal of a complimentary project in the Young laboratory (Montana State University) involves modifying these Ca^{2+} binding sites to enhance Gd binding for tumor-targeted magnetic resonance imaging (MRI)-based cancer imaging applications. (Gd is a widely used MRI contrast T_1-relaxation agent. A review of MRI and other imaging techniques can be found in Appendix C).

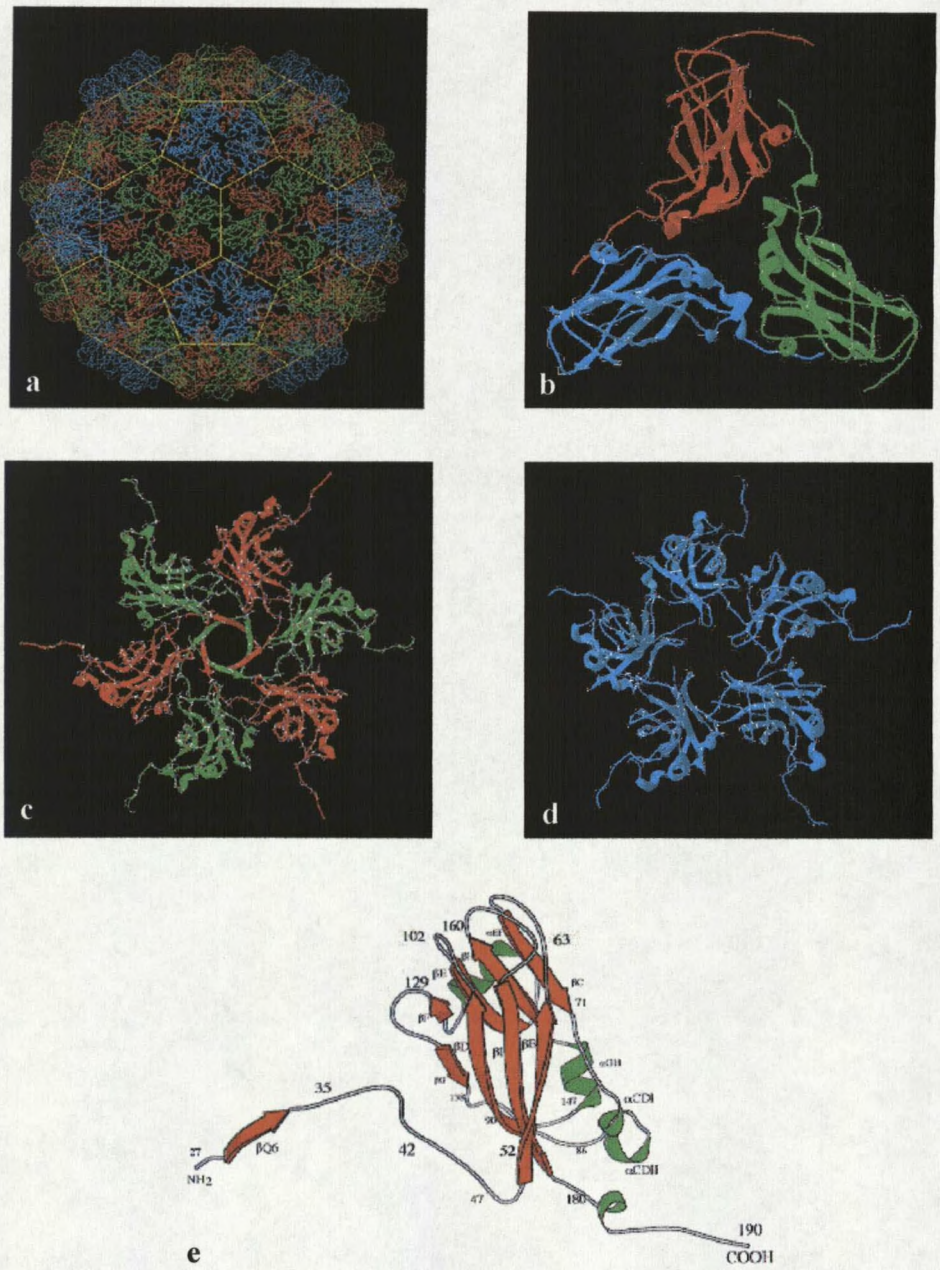


Fig.5. Structural representations of CCMV. a. CCMV virion α-carbon tracing showing 180 coat protein subunits forming the T=3 capsid structure. Pentameric capsomers (c) are in blue, hexameric capsomers (d) are in red and green. b. CCMV virion quasi-three fold axis, where 3 calcium ions are predicted to bind. e. Ribbon diagram of a single CCMV coat protein (see Fig. 6 for larger view).

CCMV-based protein scaffolds are quite stable. Native CCMV virions are stable over broad ranges of pH (pH 2-8) and temperature (-80°C to 72°C)[44,46]. The coat protein substitutions, R26C or K42R, are two important mutations that have been found to further stabilize CCMV capsids. Presumably, one or both of these mutations can be introduced into CCMV capsids that have been modified in other ways (i.e. exterior loop presentation) in order to further stabilize these chimeras. This approach is currently being explored with the chimeras that were created for this thesis project (see Chapter 5, Conclusions).

The Design Cycle: Design, Create and Test

Up to this point, I have reviewed the two major constituents that make up our envisioned tumor-targeting chimera. Now, I will discuss the approach that we took to put these components together.

Basically, we are taking an intuitive approach to designing a scaffold with novel tumor-targeting and bio-imaging capabilities. So far, we have expressed the protein and put it into the "design cycle" (including target validation). Steps in this design approach include:

i) Finding a natural protein(s) that can be easily modified to perform novel functions: For reasons outlined above, CCMV makes a natural scaffold template for our project. Importantly, the capsid is stable, easy to mutate, mutations are well tolerated, 180

identical copies of an identical coat protein provides for multiple site presentation in preferred exterior positions, and it can be heavily loaded with MRI contrast agents. Other commonly used scaffolding molecules[47] and references therein (e.g. ribosomes (ribosome display), protein toxins[48,49], phage tails (phage display), polymers (starburst dendrimers), cell surface display, and liposome display) do not share all of these attributes with CCMV. Examples of other plant virus-based scaffolds include tobacco mosaic virus (TMV)[50,51,52], tomato bushy stunt virus (TBSV)[51,53], plum pox potyvirus[54], and cowpea mosaic virus (CPMV)[51,55]. These plant virus scaffolds have generally been used to express peptides on the virion surface (as either terminal extensions or spliced into surface exposed loops) to aid in peptide purification or, more commonly, for use as epitope presentation systems (vaccines). To date, there are no reports describing the use of plant virus scaffolds for displaying a cell-targeting receptor ligand, making it difficult to evaluate these systems for the approach that we are taking in this thesis. The only examples of virus-based cell-targeting chimeras that were found at the time this thesis was written detailed the use of mammalian virus-based scaffolds for targeted gene delivery applications (discussed above). These scaffolds are not useful for the novelties which we aim to engineer, namely the ability to carry bio-imaging contrast agents.

We have chosen peptide 11 as our tumor targeting ligand to be spliced into the exterior of CCMV. The rational behind choosing peptide 11 include low reported toxicity of the free peptide, and a short, compact, and likely unique structure (among the known sequences for animal proteins, peptide 11 is specific to the laminin-1 molecule).

In addition, synthetic peptide 11 is effective (selective and specific) in targeting tumors of a metastatic phenotype, specifically those with increased surface expression and shedding of LBP. Examples of cancers that could possible be targeted by peptide 11-based agents include: breast carcinoma, prostatic carcinoma, colorectal carcinoma, lung carcinoma, gastric carcinoma, pancreatic carcinoma, ovarian carcinoma, and melanoma. Together, these cancers account for approximately 500,000 new cancer patients every year in the United States[56,57].

ii) Create structural models for proteins/peptides: As outlined above, the known (or estimated) structures of both CCMV (our scaffold template) and peptide 11 (our tumor-targeting ligand) are readily available. As discussed below, these models played an important part in conceiving this project.

iii) Analyze structures to identify residues that are critical to scaffold stability and residues that can be modified to achieve novel function(s): The five exterior loops of the CCMV coat protein appear to be excellent choices for peptide 11 insertion sites (see above and Chapter 4, Results and Discussion). We speculate that peptide 11 might be more readily recognized by its target receptor (LBP) when it is constrained by a larger scaffolding structure (here, the CCMV host scaffold). This phenomena was observed with cell-targeted adenovirus-based coat protein chimeras (see above). 3D structure analysis (surface accessibility mapping) and preliminary investigation suggest that certain

loops of the CCMV coat protein might be more amenable to the incorporation of peptide 11 than others (see Chapter 4, Results and Discussion).

iv) Experimental approaches – measuring success: An important step in design is the biophysical characterization of the final engineered product. The first measure of success used in this project was the successful expression of the chimera in cellular systems (homologous and heterologous; see Chapters 3, 4, and 5 concerning expression of our chimeras). Unfolded/unassembled proteins are generally cleared very efficiently by cellular proteases and cannot be expressed (or cannot accumulate) to significant levels. A stably folded native state also implies solubility while a tendency to aggregate indicates that elements of a protein's hydrophobic core may be accessible, leading to nonspecific interactions (aggregation).

Productive assembly of intact mutant virions was our second measure of success. Assuming sufficient quantities of soluble protein are available in appropriate concentrations, and as long as the loop insertions do not interfere with capsid assembly, virions should form (see Chapter 4, Results and Discussion, and reference 58). Biophysical characterization of assembled particles can discriminate robust (useful) virions from labile (less useful) virion constructs, and provides important clues for improving future rounds of protein design and engineering.

A final measure of success in creating our macromolecular bio-imaging agent is target validation: do our constructs find their LBP-expressing tumor targets? A target can

be validated in the laboratory using functional *in vitro* and in animal assays (including specific targeting inhibition assays using competing ligands and/or monospecific antibodies raised against the targeted receptor). Since we have just begun "target validation" of our constructs and have no significant data concerning their performance, this step in the design cycle will not be fully discussed in this thesis.

Once a construct meets these requirements (suitable expression, stable assembly, and target validation) it will be evaluated for its ability to target Gd to LBP and metastatic tumor cells, and for its usefulness in the early detection of micrometastases. These future experiments will not be discussed since they are beyond the scope of this thesis.

Summary

The long-term goal of this project is to create a tumor-targeted virus-based bio-imaging agent for use in the early detection of metastases. We have chosen two basic components for its construction: the tumor targeting ligand, peptide 11 and its eventual host scaffold, CCMV. The independent attributes of peptide 11 and CCMV suggest that a marriage between these two structures would be an ideal system for creating a tumor-targeted bio-imaging agent. An agent of this sort may be helpful in estimating the metastatic potential of tumors, thereby aiding in staging and prognostic-related decisions on cancer treatment. We consider this to be a novel approach to cancer imaging and

believe that this work may ultimately lead to new approaches to reduce patients' suffering from cancer and therefore generally improve human health.

CHAPTER 3

METHODS

Design of CCMV-Peptide 11 Chimeras

Structure and Sequence Analysis

Structures of the CCMV coat protein based on atomic coordinates obtained from the Brookhaven Protein Data Base (PDB ID 1CWP) were rendered using Swiss-PDBViewer (version (v) 3.6b3). Coat protein oligomers (pentamers and hexamers) were generated using the internet-based "oligomer generator" program on the VIPER website[59] or were obtained (with atomic coordinates in PDB form) from the Scripps University, CA website[60]. Peptide 11 structures (atomic positions estimated from molecular dynamics experiments based on populations of NMR-derived structures) were obtained from the Starkey laboratory.

Selection of Insertion Sites

B-Factor Mapping. The B-factor (or temperature factor) is a measure of the disorder in a crystallized protein resolved by X-ray crystallography. "Dynamic disorder" ($<ud^2>$) is caused by the temperature-dependant vibration of atoms in the structure. Also,

molecules, or parts of molecules, in different unit cells of the resolved structure do not occupy exactly the same position and do not have exactly the same orientation. This is called "static disorder" ($\langle us^2 \rangle$). The B-factor combines dynamic and static disorder in the equation: $B = 8\ \pi 2\ (\langle ud^2 \rangle + \langle us^2 \rangle)$. B-factors tend to be high for exterior (surface exposed) atoms and low for interior (buried) atoms. Using the program, Swiss-PDBViewer (v3.6b3)[61,62], individual residues within CCMV subunits, capsomers (pentamer and hexamer), and capsids (the entire CCMV structure) were colored according to their associated B-factor, from dark blue for low B-factor to red for high B-factor. The highest B-factor of any backbone atom is attributed to all backbone atoms, the same is true for sidechains.

Surface Accessibility Mapping. Each amino acid in the coat protein was colored by its relative accessibility using Swiss-PDBViewer (v3.6b3)[60,61]. Maximum accessibility is defined as being the accessible surface of an amino-acid X in a pentapeptide GGXGG in extended conformation. This is only an approximative scale, but adequate to differentiate core amino-acids from surface ones. Dark blue color is attributed to completely buried amino-acids, whereas red color is attributed to amino-acids with at least 75% of their relative surface "totally" accessible.

Inter-subunit association energies were calculated based on the atomic buried surface areas multiplied by solvation parameters, as previously described[63]. At each interface, contribution of individual residues were estimated and plotted as a function of residue numbers. In addition, the residue-wise contributions (for each side of the

interface) were sorted in decreasing order and the top 10 residues and their contributions were listed. These calculations were generated by the VIPER staff and are reported in the VIPER website[64]. They have been reprinted (with permission) in Appendix A.

Engineering CCMV-Peptide 11 Chimeras (Plants)

Construction of Loop-Specific BamH1 Cloning Sites

All DNA, RNA and protein manipulations were carried out with previously described methods[65] unless stated otherwise. A GeneAmp PCR System 9700 thermocycler (Perkin-Elmer – Applied Biosystems, Foster City, CA) was used for all PCR amplifications.

Insertion of Peptide 11 Sequence into Loop Positions

Loop-specific BamH1 cloning sites (amino acid positions 62-63, 101-102, 114-115, 129-130, and 160-161 in the wild-type coat protein; between residue sets) were identified during the design stage (see above and Chapter 4, Results and Discussion). As a first attempt to express exogenous peptide 11 in each of these sites, we inserted the peptide 11-based sequence (GICDPGYIGSRCGIL) *within* the loop sequence. To do this, plasmid-based vectors (called pCC3 with BamH1 site, see Fig. 11) were constructed for cloning DNA sequences encoding for heterologous fusion proteins into the surface

exposed loops of CCMV. (This work was done prior to my arrival at Montana State University.)

Described briefly, full length cDNA copies of genomic RNA1, RNA2, and RNA3 of CCMV were previously cloned 1 base downstream from a T7 RNA polymerase promoter in a pUC118 based vector background[66]. These constructs are referred to as pCC1, pCC2, and pCC3 (see Fig. 11). Capped transcripts made by *in vitro* transcription (see below) from these plasmid templates (after linearizing with Xba1) were previously shown to infect plants when inoculated simultaneously.

Using PCR "QuikChange" oligonucleotide site-directed mutagenesis (as described[67]) a unique BamH1 restriction site was inserted into each of the five pre-selected regions of the CCMV coat protein cDNA using the pCC3 plasmid as template (see Fig. 3a). Deoxyoligonucleotides used for mutagenesis were: BamH1-loop 63 (5'TCT TGT GCG GCT GCC GGG ATC CTG GAA GCT AAA GTA ACC); BamH1-loop 102 (5'GGG TTG CTT CCC AGT GTT GGG ATC CTG AGT GGC ACA GTG AAA TCC); BamH1-loop 114 (5'ACA GAG ACG CAG ACT GGG ATC CTG ACT GCT GCT GCC TCC); BamH1-loop 129 (5'GCT GTG GCC GAC AAC GGG ATC CTG TCG AAA GAT GTT GTC GCT); BamH1-loop 160 (5'ATC TAC TTG TAC AGC GGG ATC CTG AGT GCG GCT CTC ACT). 3' primers were complimentary to the 5' primers. Nucleotides in primer that are designed to introduce nucleotides in coat protein gene are colored in red. BamH1 restriction sites are underlined (see Fig. 11).

The resulting pCC3 with BamH1 constructs were digested with BamH1 and free plasmid ends were dephosphorylated with 1/10 volume calf intestinal alkaline

phosphatase (CIP). Synthetic deoxyoligonucleotide duplexes encoding peptide 11 (with complimentary BamHI ends underlined in the following sequence: 5'G ATC TGC GAT **CCC GGG** TAT ATA GGC TCG CGA *TGT* GG; 3'GAT CCC *ACA* TCG CGA GCC TAT ATA **CCC GGG** ATC GCA) were ligated into the "compatible ends" of restriction digested BamH1 sites. The ligation reaction included a ratio of 1:3 insert:plasmid with 1μl bacteriophage T4 DNA ligase (Promega, Madison, WI) and 1μl 10x ligation buffer (containing 100mM $MgCl_2$ and 10mM ATP, Promega) in 10 μl total volume. This cocktail was slowly cooled from 27°C to 4°C overnight to facilitate ligation. The ligation reaction was transformed into XL2-Blue Ultracompetent *E. coli* cells (Stratagene, Cedar Creek, TX) and cells were grown at 37°C overnight on LB agar plates under ampicillin (Amp) selection. Single colonies were picked and grown overnight in liquid LB under Amp selection. Plasmids were prepared using Perfectprep plasmid mini kits (Eppendorf, Westbury, NY) as directed. The resulting constructs, called pCC3/pep11-63, -102, -114, -129, and -160 (see Fig. 11), encoded for a C-terminal (to peptide 11) cysteine (italicized in above sequence) and contained a unique SmaI restriction site (bold in above sequence) that was used for screening by diagnostic restriction digestion. 750 μl of *E. coli* culture from clones found to be positive for the peptide 11 insertion was stored in 15% (v/v) glycerol at –80°C. Prepared DNA from positive clones was tested to verify constructive cloning (below).

Two independent methods were used to verify constructive cloning (in frame and forward orientation of peptide 11 sequence; no detrimental mutations):

1.) DNA sequence analysis: ABI Prism BigDye (Perkin-Elmer – Applied Biosystems, Foster City, CA) PCR-based terminator cycle sequencing of constructs (as previously described[68]) was performed. "Nested set" deoxynucleotide primers (5'GCG TAT CGA AGT GGA CCG CC; 3'GTT GTT CAA GGG TTA TAC CCC TAA AGG CCT CGG G, or 3'GTA ACG GTC GAC AGC GGG C) complimentary to regions of the coat protein gene that flank the peptide 11 insert sequences were used to generate labeled PCR fragments from pCC3/pep11 templates. These fragments were PCR cycle sequenced.

2.) Coupled *in vitro* transcription/translation assay: This second assay has previously been described[69]. Briefly, the chimeric CCMV-peptide 11 coat protein gene was PCR amplified from pCC3/pep11 constructs using these "nested set" deoxynucleotide primers: 5'TAA TAC GAC TCA CTA TAG GGT AAT TTA TC*A TG*T CTA CAG TCG GAA CAG GG; and 3'GTA ACG GTC GAC AGC GGG C. The 5' primer introduces a bacteriophage T7 RNA polymerase promoter element (underlined in above sequence; the ribosomal start codon (ATG) is italicized) to the 5' end of the modified coat protein gene (at wild type RNA 3 nucleotide position 1350). The 3' primer is complimentary to sequence immediately 3' of the coat protein stop codon, TAG. DNA quantitation and integrity was determined by UV spectrophotometry and visualization on agarose gels. Chimeric CCMV-peptide 11 RNAs were synthesized from PCR amplified coat protein genes using T7 mMessage mMachine (Ambion, Austin, TX). RNA quantitation and integrity was determined by UV spectrophotometry and visualization on denaturing agarose gels[56]. Resulting 5'capped mRNAs were translated using cell free

wheat germ extracts as described in the translation kit protocol (Promega, Madison, WI). Wheat germ extracts contain the cellular components necessary for protein synthesis (tRNA, ribosomes, initiation, elongation and termination factors). Radioactive [^{35}S]methionine was utilized so that translation products could be visualized (after separation of products by SDS-PAGE, see below) using autoradiography.

In Vitro Transcription for Plant Inoculation

Plasmids (pCC1, pCC2, pCC3 (wild type), and pCC3/pep11 plasmid constructs; see Fig. 11) were prepared for *in vitro* transcription by linearization with Xba1 followed by Proteinase K (Sigma, St. Louis, MO) digestion for 30 min at 65°, phenol-chloroform extraction and ethanol precipitation. Chimeric CCMV-peptide 11 RNAs were synthesized by *in vitro* transcription from approximately 1.0 μg of linearized DNA with T7 Megascript (Ambion, Austin, TX) in a 20 μl reaction using this company's protocols. RNA quantitation and integrity were determined by UV spectrophotometry and visualization on denaturing agarose gels[56].

Plant Inoculation and Virus Propagation

Carborundum dusted leaves of cow pea plants (*Vigna unguiculata* (L.) var. California Blackeye) and tobacco plants (*Nicotiana benthamiana*) at the primary (two leaf) stage were manually inoculated with 20μl/leaf of a mixture containing 4 μl of each

in vitro transcription reaction (for approximately 1μg of pCC1-, pCC2-, and pCC3 (wild type)- or pCC3/pep11 (chimeric)-based RNA/leaf) in 28 μl phosphate buffer (0.01M KPO_4, pH 6.0). This). For plant infection to occur, all three RNAs (here, generated *in vitro* from pCC1, pCC2, and pCC3 or pCC3/pep11) need to be inoculated concomitantly (see Fig. 3a). Alternatively, solution (homogenization) can be prepared using leaves that have been productively infected from transcript inoculated plants. 50 μg/ml of this solution was used to "passage" virus from plant-to-plant.

Virus Purification from Plants

Following inoculation and 2-4 weeks of growth, virus was purified from systemically infected leaves using previously described methods[70]. Briefly, infected primary and secondary leaves (kept separately) were frozen at –20°C or were frozen by immersing in liquid nitrogen, crushed, and placed into a Waring blender. The contents were blended for approximately 3 minutes in homogenization buffer (0.2 M sodium acetate, 0.2 M acetic acid, 0.01 M ascorbic acid, 0.01 M disodium EDTA, pH 4.8). The homogenate was expressed through 2 layers of cheesecloth, and the foam was removed. The emulsion was allowed to separate on ice for one hour before subjecting to low speed centrifugation (20 min at 20,000 x g) to pellet leaf tissue. Virus in the partially cleared supernatant fluid was precipitated by addition of 10% (w/v) polyethylene glycol (PEG) MW 8000, followed by slow overnight stirring. The precipitate was pelleted by low speed centrifugation and then resuspended in virus buffer (0.1 M sodium acetate/acetic

acid, 1 mM sodium azide, 1 mM disodium EDTA, pH 4.8, 0.22 μm filtered). The resuspension was cleared of undissolved material using low speed centrifugation. Virus was precipitated a second time with 15% (w/v) PEG MW 8000 and resuspended in virus buffer. This second virus resuspension was subjected to isopycnic ultracentrifugation. Virus buffer plus 38% CsCl (w/w) was placed in Beckman SW41 centrifuge tubes (10-11 ml/tube). Each tube was top loaded with ≤ 15 mg of virus in 1-2 ml virus buffer and then centrifuged at 38,000 r.p.m. for 18-24 hrs. The virus band was dialyzed against three volume changes of virus buffer and then concentrated in Centricon™ 100 microconcentrators and stored at 4°C in virus buffer. Typical yields of wild type CCMV were 200 – 300 mg of virus per kg of leaf tissue ($E_{260}^{1mg/ml} = 5.87$); yields of chimeric complexes from plants varied, but were generally 100 fold lower than wild type virus yields.

Virus Characterization from Plants.

ELISA. ELISA-based detection ("double-sandwich" method) of virus particles isolated from plants (or *P. pastoris* yeast; see below) was preformed as previously described[71]. Briefly, 100 μl of 1:5000 dilution of 1 mg/ml purified CCMV polyclonal antibody (Ab) serum (rabbit anti-CCMV 1°) in coating buffer (1.59 g/L Na_2CO_3, 2.93 g/L $NaHCO_3$, 0.2 g/L NaN_3, ph 9.6, brought to 1L volume with dH_2O) was adsorbed onto the surface of wells of Costar EIA/RIA 96 well flat bottom high binding microtiter plates (Corning Incorporated, Corning, NY) for at least 12 hrs at 4°C. Plates were washed 3

times with 200 µl of PBS (8.0 g/L NaCl, 0.2 g/L KH_2PO_4, 2.9 g/L $Na_2HPO_4.12H_2O$, 0.2 g KCl, pH 7.4, brought to 1L volume with dH_2O) containing 0.05% Tween 20 (washing buffer). Wells were coated with 100 µl of antigen (CCMV (wild-type) or CCMV-peptide 11 chimeras). Antigen was prepared by grinding 100 mg leaf tissue (in dH_2O or antigen buffer (0.1% K_2PO_4, 2% polyvinylpyrrolidone (40,000 MW; Sigma, St. Louis, MO), pH7)). Antigen coated plates were incubated overnight at 4°C and then washed three times with washing buffer. A 1:500 dilution (in antigen buffer) of alkaline phosphatases (AP)-conjugated rabbit anti-CCMV polyclonal antibody was added with 2 mg/ml chicken albumin and incubated for 2 hrs at 37°C. After incubation, wells were washed 3 times with washing buffer. Then, 100 µl of p-nitrophenyl phosphate, disodium AP substrate was added (Sigma, St. Louis, MO; prepared at 1 mg/ml by weight in substrate buffer: 9.7% Diethanolamine, 0.02% NaN_3, pH9.8). Color reaction was measured in a Molecular Devices kinetic microplate reader (Molecular Devices, Sunnyvale, CA) at OD 405 nm. In parallel experiments, serial dilutions of plant purified wild type CCMV were used as standards. Background signal was subtracted out from reported ODs. This ELISA procedure was used to detect and quantify samples isolated from *P. pastoris* yeast (see below) as well as those isolated from plant tissue.

RT-PCR Amplification and Restriction Digestion. Viral RNA was isolated from plant tissue using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA) as directed. cDNA synthesis from this RNA was performed using "nested set" complimentary deoxyoligonucleotide primers that flank the coat protein gene (5'GAT AGT AAG TCG

ACA TGT CTA CAG TCG G; 3'GTA ACG GTC GAC AGC GGG C). The Promega Access RT-PCR system and protocols (Promega, Madison, WI) were used. Amplified products were phenol-chloroform extracted and ethanol precipitated prior to sequencing (as described above) or restriction enzyme digestion (SmaI restriction digestion was used to confirm presence of peptide 11 sequence in clones and in cDNA isolates).

SDS-PAGE and Western Blot Analysis. Purified virus, virus samples isolated from plants (by grinding in liquid nitrogen), or *in vitro* translation products were boiled and electrophoresed on pre-cast 12.5% or 20% SDS/polyacryamide gels according manufacturer guidelines (PhastSystem, Pharmacia Biotech, Uppsala, Sweden). The gel-loading buffer contained 2.5% β-mercaptoethanol reducing agent. Proteins were visualized by autoradiography ([^{35}S]-labeled *in vitro* translation products) or by western blot as described in PhastaSystem manual. Briefly, proteins were electroblotted onto nitrocellulose membrane (PhastaSystem manual) and the membrane was blocked overnight at 4°C in milk solution (10 ml TBS-tween (8 g/L NaCl, 0.2 g/L KCl, 3 g/L Tris base, pH 7.4, brought to 1L volume with dH_2O, add 1 ml/L Tween-20) plus 0.5 g nonfat dried milk). The membrane was washed (three quick rinses with TBS-tween followed by three 5 min washes in TBS-tween). Protein was probed with 1:5000 dilution of rabbit anti-CCMV polyclonal Ab in 10 ml TBS-tween. Horseraddish peroxidase (HP)-conjugated goat anti-rabbit IgG was the secondary antibody used for enhanced chemiluminescence detection (Amersham, Buckinghamshire, England) or, alternativly, for Opti-4CN cholorimetric detection (Bio-Rad Laboratories, Hercules, CA). This SDS-

PAGE/western blot protocol was also used to characterize proteins isolated from *P. pastoris* (see below).

Engineering CCMV-Peptide 11 Chimeras (*P. pastoris*)

In order to meet the demand for large amounts of easily to purify capsid constructs, we chose to transfer the core of the pCC3/pep11 expression cassettes into the *P. pastoris* heterologous expression system. This system has gained widespread attention because of its ability to express high levels of heterologous proteins. General protocols from the Invitrogen (San Diego, CA) manual and the desk reference, "*Pichia* Protocols - Methods in Molecular Biology"[72], were followed with some notable exceptions. For reviews on the *P. pastoris* expression system, see references [73 and 74].

Construction of pPicZA/pep11 Plasmids

CCMV coat protein genes (with integrated peptide 11 sequence and without superfluous 3'untranslated regions) were subcloned from their original plasmid vector (pCC3/pep11) into the *P. pastoris* expression plasmid pPicZA to construct pPicZA/pep11-63, -102, -114 and –129 (see Fig. 17). To accomplish this, PCR was used to amplify the modified genes from pCC3/pep11 plasmid constructs (used as the PCR template) using the deoxynucleotide "nested set" primers: 5'GAT AGT AAG AAT TC*A* *TG*T CTA CAG TCG G; 3'GTA ACG **GTC GAC** AGC GGG C. The 5' primer is

complimentary to a region surrounding the coat protein start codon, *ATG* (italicized in above sequence), and was built to introduce an EcoR1 endonuclease restriction site (underlined in above sequence) into the 5' end of the amplified fragment. The 3' primer is complimentary to the region immediately 3' to the coat protein stop codon, TAG. The CCMV 3' untranslated region (3' of the coat protein stop) was not included in the amplified PCR product since it is not needed for heterologous expression. The native coat protein stop was maintained to ensure that gene fusions from the pPicZA vector (a C-terminal MYC epitope tag or a C-terminal (His)6 tag; Invitrogen) did not occur. Also, the 3' primer was built to introduce a Sal1 endonuclease restriction site (bold in above sequence) into the 3' end of the amplified fragment. This fragment was digested with EcoR1 and Sal1 and then purified from gels using QIAquick DNA /gel purification kits (Qiagen, Valencia, CA).

The digested and purified PCR fragment was ligated (as was done above with peptide 11 ligation into pCC3 vectors) into EcoR1 and Xho1 digested (Xho1 creates compatible ends to Sal1 restriction site) and CIP-dephosphorylated pPicZA *P. pastoris* yeast expression vectors (Invitrogen). The ligation reaction was digested with 1/10 vol Kpn1 restriction endonuclease to linearize the plasmids without insert (and thereby decrease transformants that do not contain correct plasmid constructs). XL-2 Blue ultracompetent *E. coli* cells were transformed as described (Stratagene). Transformed *E. coli* colonies were selected after ~16 hrs growth on low salt LB agar plates (1% w/v Trytone, 0.5% w/v yeast extract, 0.5% w/v NaCl, 15g/L agar, pH7) containing 0.025 mg/ml Zeocin. Single colonies were grown overnight in low salt LB media (as above,

without agar) under Zeocin selection (0.025 mg/ml Zeocin). Cultures were Perfect Prep'd (Eppendorf, Westbury, New York) to extract plasmid DNA and screened by enzymatic digestion with SmaI (a SmaI site is present in both the peptide 11 sequence and the Zeocin resistance gene in the pPicZA base plasmid). Correct incorporation of inserted DNA into pPicZA plasmids (in frame and forward orientation of peptide 11 sequence; no detrimental mutations) was confirmed by sequencing (using the protocol described above) with pPicZA-based primers (5'*Pichia/AOX1* primer: GAC TGG TTC CAA TTG ACA AGC; 3'*Pichia/AOX1* primer: GCA AAT GGC ATT CTG ACA TCC, Invitrogen). Alternatively, some constructs were only sequenced after integration into the *P. pastoris* chromosome (see below). CCMV-peptide 11-based *P. pastoris* expression vectors were termed pPicZA/pep11-63, -102, -114, and –129 for the specific loop with peptide 11-based insert (see Fig. 17).

Transformation into *P. Pastoris*

After linearization with XbaI, pPicZA/pep11 expression vectors were transformed into *P. pastoris* X-33 mut$^+$ yeast (Invitrogen). The standard electroporation transformation method was followed (Invitrogen) using a T800L Electro Cell Manipulator (BTX, San Diego, CA) electroporation unit according to the manufacturer's specifications. Zeo^r transformants were selected after 2-3 days growth on YPDS agar plates (1% w/v yeast extract, 2% w/v peptone, 2% w/v dextrose (glucose), 1M sorbitol, 2% w/v agar) containing Zeocin (0.1 mg/ml). Single colonies were inoculated into 3 ml

of liquid YPDS (as above, without agar) containing Zeocin (0.1 mg/ml) and grown overnight. Direct PCR-based screening of clones was performed using the standard lyticase-PCR method (Invitrogen). Briefly,10 µl of 1:10 diluted (with water) *P. Pastoris* culture was disrupted using 25 U lyticase (Sigma). Internal CCMV coat protein deoxynucleotide primers that flank the peptide 11 sequence were used (5'GCG TAT CGA AGT GGA CCG CC; 3'GTT GTT CAA GGG TTA TAC CCC TAA AGG CCT CGG G). PCR cocktail and thermocycler parameters were followed exactly as described in the Invitrogen *P. pastoris* manual. Lyticase-PCR fragments were digested with SmaI to confirm the presence of peptide 11. PCR fragments were also sequenced (as above) to confirm proper integration.

Small Scale Time Course Expression Screening

25 ml of MGY (1.34% w/v yeast nitrogen base without amino acids (YNB), 1% v/v glycerol, 0.00004% w/v biotin) was inoculated with a single colony of *P. pastoris* carrying the CCMV coat protein gene (plus peptide 11 insert in a single loop). This was grown for 18 hrs at 30°C with vigorous shaking. For induction of the coat protein construct expression, the cells were harvested by centrifugation, washed once with MM medium (1.34% w/v YNB, 0.00004% w/v biotin, 0.5% v/v methanol) and resuspended in 150 ml of MM medium. During the incubation, 100% methanol was added to the culture at 24-hr intervals, each time to give a final concentration of 0.5%. At 0, 24, 36, 60, 84,

108, and 132 hr after induction, 1 ml cell samples were harvested by centrifugation and pellets were stored at –80°C until the time course series was completed.

Time course culture pellets were disrupted in standard SDS-PAGE gel-loading buffer[56] (for dot blots, bromophenol blue was not added) using approximately 20 µl acid-washed glass beads (Biospec Products – BSP, Bartlesville, OK) and vortexing vigorously for 10min. Samples were boiled for 5 min then centrifuged at low speed. 1 µl of the supernatant was blotted onto nitrocellulose membrane and visualized using the Western blot visualization protocol (see above). Samples with greatest protein concentration (determined by visual inspection of time course dot blots) were separated by SDS-12.5% PAGE, Western blotted, and visualized as described above.

Intermediate-Scale Expression of
CCMV-Peptide 11 Coat Protein in *P. Pastoris*

Having optimized expression using the time course study, cultures were scaled-up in an effort to obtain greater amounts of protein that could readily purified. 50 ml of MGY was inoculated with a single colony *P. pastoris* isolate. This culture was grown for 16-18 hrs at 30°C with vigorous shaking. Cells were harvested by centrifugation, washed once with MM medium and resuspended in 500 ml of MM medium (in a 1 L baffled flask) to start induction. During the incubation, 100% methanol was added to the culture at 24-hr intervals, each time to give a final concentration of 0.5%. Cells were harvested by centrifugation and pellets were stored at –80°C.

Intermediate-scale culture pellets were disrupted in equal volume virus buffer (see above) with acid-washed glass beads. This mixture was vortexed vigorously 3-4 times for 5 min each, with 2 min on ice between each vortexing. Samples were then centrifuged at low speed to remove unwanted cell material and glass beads. Virus in the partially clarified supernatant fluid was precipitated by addition of 10% (w/v) polyethylene glycol (PEG) MW 8000, followed by slow overnight stirring. The precipitate was pelleted by low speed centrifugation and then resuspended in 1/10 initial supernatant volume virus buffer. At this point, some undissolved matter remained in the resuspensions. Samples were therefore centrifuged for 10 min at 10,000 r.p.m. to remove this undissolved material. Virus (presumably in the supernatant) was detected from samples using ELISA (as described above). If a sample gave an ELISA signal that was significantly above background, it was concentrated using a Centricon™ 100 microconcentrator after dialyzing against three volume changes of virus buffer. Virus capsids in concentrated samples were visualized by transmission electron microscopy (TEM).

Electron Microscopy. Liquid samples were placed on the surface of carbon-coated, Formvar-covered copper TEM grids. Excess solution was removed with filter paper and the grid was allowed to air dry. Virions were negatively stained with 1% uranyl acetate (UA) which was pipetted onto sample-coated grids. Excess stain was removed with filter paper and the grid was allowed to air dry. Samples were imaged on a Zeiss EM 100CA transmission electron microscope.

Expression and Purification of CCMV-Peptide 11 Chimeras from High-Density Fermenter Cultures

To obtain greater amounts of our CCMV-peptide 11 chimeras from *P. pastoris*, we scaled-up to a four-liter fed-batch fermenter culture. The fermentation was done according to Invitrogen and "*Pichia* Protocols - Methods in Molecular Biology" guidelines in a Bioflo 3000 bench-top fermenter (New Brunswick Scientific) equipped with a 5L water-jacketed glass fermentation vessel with microprocessor control of pH, dissolved oxygen (DO), agitation, temperature, and nutrient-feed. The vessel containing medium composed of fermentation basal salts medium (85% v/v phosphoric acid, 0.93 g/L CaSO4, 18.2 g/L K2SO4, 14.9 g/L MgSO4· 7H20,4.13 g/L KOH, and 4% v/v glycerol, all brought to 3L volume with water) was sterilized by autoclaving, then 9 g/L (NH4)2SO4, 25g/L sodium hexametaphosphate, and 4.35 ml/L trace salts (Invitrogen, PTM1) were added aseptically. Throughout the fermentation cycle, foam was controlled by manually adding an antifoam agent, silwel L-7602 (Witco Corporation) and temperature was kept at 30°C. Fermentation proceeded in three phases:

Glycerol Batch Phase. To begin the initial "glycerol batch phase", 300 ml of inoculum seed culture was added. The seed culture consisted of cells (5-10% of initial fermentation volume) grown in flasks in MGY, as described above. Importantly, these cells were *not* cultured under Zeocin selection, since the inoculum had already been selected for Zeo^r. Agitation was maintained at a minimum of 500 r.p.m. and was

programmed to increase up to 1000 r.p.m. to maintain a DO content of 35%. Initially, compressed air, and later, pure compressed oxygen was used to help reach and maintain this level. A pH of 5.0 was maintained using 50% ammonium hydroxide. The glycerol-fed batch-phase was run until glycerol was completely consumed (18-24 hrs). This was indicated by an increase in the DO to 100%.

Glycerol-Fed Batch Phase. Once all the original glycerol was consumed from the batch growth phase, a glycerol feed was initiated by continuously adding to the culture a medium containing 50% (v/v) glycerol. This phase was used to increase the cell biomass under limiting conditions while derepressing the alcohol oxidase (AOX1) promoter before induction. The feed rate was adjusted to maintain 35% DO for 4 hrs or at a slower rate overnight. The feed was stopped for 30 min to ensure total glycerol depletion before methanol induction.

Methanol-Fed Batch Phase. A methanol feed was initiated at a vary slow rate to maintain ~35% DO using a MBP peristaltic pump for continuous delivery of 100% methanol to the culture. PMT1 trace salts were fed intermittently at 8 ml/hr using one of the fermenter automated feed pumps. As the culture became adapted to methanol utilization, the feed rate was increased as quickly as possible while still maintaining a 20-35% DO reading. Pure compressed oxygen and agitation at 1000 r.p.m was used to maintain this level of DO. It usually took about 1 day to reach this point. This feed rate varied slightly throughout the remainder of the fermentation cycle depending on the

amount and frequency of antifoam agent addition and the age of the culture. DO was maintained throughout the remainder of the fermentation to obtain adequate expression of CCMV coat protein chimeras. 3.5 to 4.0 L of 100% methanol was consumed during a typical fermentation run. The entire vessel was harvested after ~110 hrs on 100% methanol by collecting the culture into 250 ml centrifuge tubes and centrifuging at low speed. 100 g aliquots were stored at –80°C. Small (10 µl) samples were stored separately and later analyzed by lyticase-PCR (see above) to determine *P. pastoris* isolates' genetic homogeneity and stability.

Lysis of Cells and Sample Purification. 100 g frozen aliquots of stored cells were lysed by grinding with 235 g acid-washed glass beads in a bead-beater (BioSpec Products – BSP, model 1107900) in the presence of homogenization buffer (see above). Samples were then centrifuged at low speed to remove unwanted cell material and glass beads. In our first attempts, we tried (with some success) to purify virions from the supernatant using the same method that we used for isolating particles from intermediate scale cultures (see above). Later, in an effort to obtain samples of equal or greater purity, but with increased yields of virion particles, it became necessary to avoid using PEG precipitations. The supernatant (after bead-beating and multiple low speed centrifugations) was filtered (stepwise) through a 0.65 µm pore size low protein binding filter and then a 0.2 µm pore size low protein binding filter. Filtrate was centrifuged using a Beckman Type 30 Rotor at 25,000 r.p.m. for 2 hr. Tubes were top-loaded with ~22 ml filtrate and then a 3 ml 15% (w/v homogenization buffer) sucrose cushion was

bottom loaded using a hand syringe pump. Resultant pellets were resuspended in 1-2 volumes (v/v) of virus buffer and stored at 4°C for later use.

Sample Purification and Virion Characterization. Samples were purified using CsCl isopycnic ultracentrifugation exactly as described above for virion purification from plants. Alternatively, samples were separated on 10-40% sucrose (w/v virus buffer) sedimentation velocity ultracentrifugation gradients. Gradients were poured in Beckman SW41 centrifuge tubes (10-11 ml/tube) using a Gradient Master (BioComp Instruments, NewBrunswick, Canada) as indicated by manufacturer guidelines. 400 µl samples were top loaded and then centrifuged at 37,000 r.p.m. for 3 hrs. The gradients were fractionated using an ISCO model 640 gradient fractionator while monitoring absorption at 280 nm. If no signal was detectable by absorbance, 500 µl fractions were analyzed by ELISA (as described above). Desired fractions (or bands from CsCl separation) were collected and dialyzed against three volume changes of virus buffer, concentrated in Centricon™ 100 microconcentrators, and then stored at 4°C in virus buffer. Samples (purified virion particles) were characterized by ELISA, TEM, and 12.5% SDS-PAGE/western blot (using methods described above). Sucrose gradient purification was also used to determine homogeneity and particle character, based on sedimentation velocity relative to wild type standards.

CHAPTER 4

RESULTS AND DISCUSISON

Design of CCMV-Peptide 11 Chimeras

We chose to present the nine amino acid tumor-targeting peptide 11 sequence (CDPGYIGSR) on each of five exterior loops of CCMV (a virus based protein scaffold). The insertion sequence included five additional residues needed to introduce Bamh1 restriction sites for cloning plus a C-terminal cysteine that we felt could be used to stabilize the cloned loop through a disulfide bond with the first cysteine from the peptide 11 insert. This "closed loop" structure might also more closely mimic the structure of peptide 11 within the laminin-1 LE repeat (see Chapter 2, Why peptide 11?). Together, 15 residues (GICDPGYIGSRCGIL) were selected for insertion into each of five CCMV-coat protein surface loops.

Structure and Sequence Analysis

Precise loop insertion sites were determined by structural analysis, including B-factor and surface accessibility mapping (see below). Analysis of the 3.2 Å resolution CCMV structure reveals five loops (βB-βC, βD-βE, βF-βG, βE-αEF, and βH-βI) that appear to be on the surface of each capsid monomer (Fig. 6). These loops represent

possible graft sites from which to display a short peptide. The structure Cucumber mosaic virus (CMV) (also of the *Bromoviridae* virus family) was recently determined to 3.2 Å resolution (unpublished report by Smith T., Purdue University). Despite having low sequence identity (34% similarity, Fig. 7) to CCMV, the core structures of both virus capsid subunits are highly homologous. Three of the surface loops projecting from these cores are similar, while two are significantly different: the βE-αEF (114) and the BF-BG (129) loop are seven residues and six residues, respectively, longer than that in CCMV. The subunit from another class of virus, the DNA tumor papovaviruses (polyoma and SV40), have a core that is very similar to that of the CCMV coat protein, but have exterior loops that vary considerably in structure (they are generally larger) from those that extend from the CCMV coat protein core (Fig. 6). Comparing capsid structures that are related to the CCMV capsid lends credence to the idea that the exterior loops of CCMV coat protein can be modified to derive functionalities as long as the core fold is maintained.

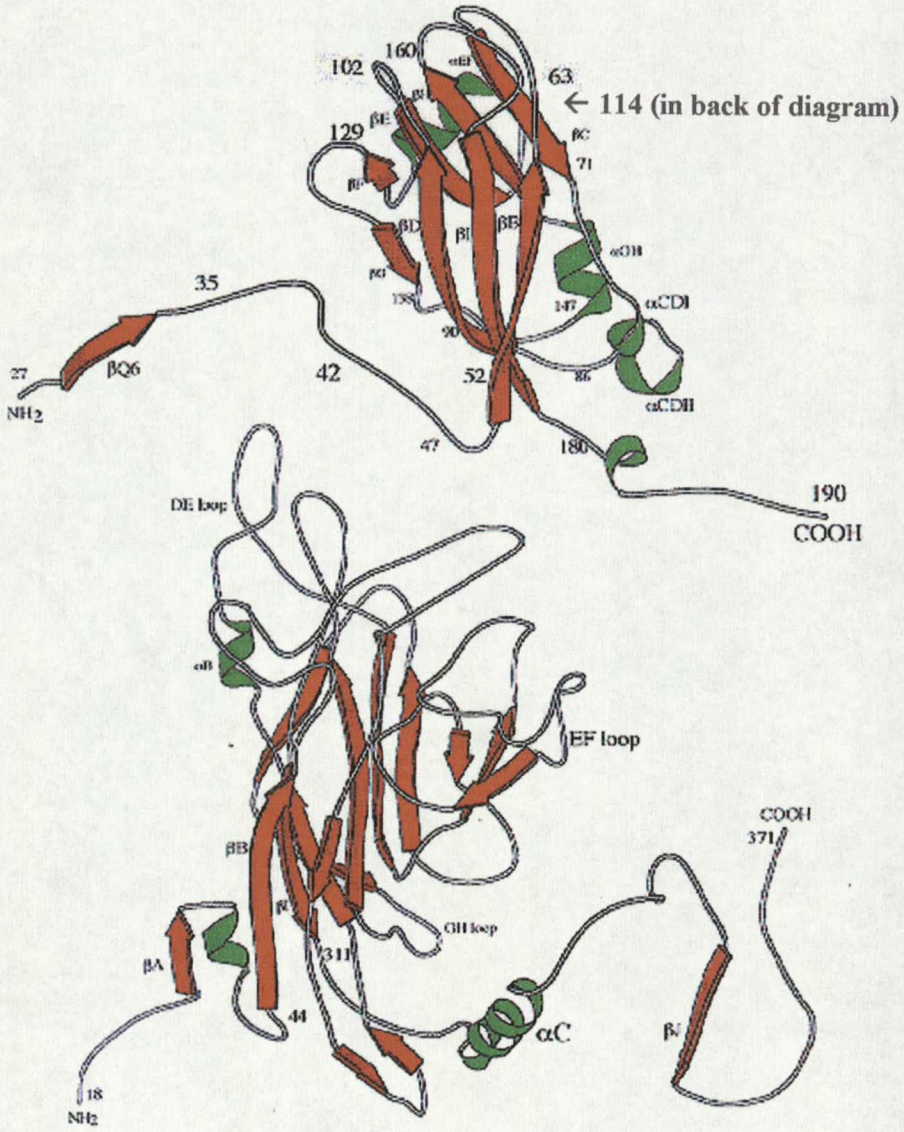
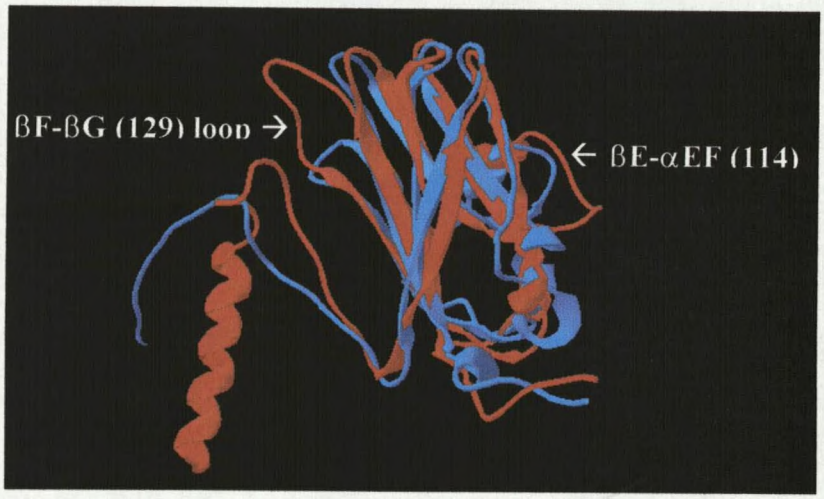


Fig. 6. Structural comparison of the CCMV capsid subunit to the Polyoma virus VP1 capsid subunit. Structural analysis reveals five candidate surface loops on the scaffolding protein, Cowpea chlorotic mottle virus (CCMV) coat protein (top), that could support heterologous expression of peptide 11: βB-βC loop: labeled by amino acid position 63, βD-βE:amino acid position 102, βE-αEF:amino acid position 114, βF-βG:amino acid position 129, and βH-βI:amino acid position 160 and 161. The structurally related Polyoma virus VP1 capsid subunit (bottom) has a similar core fold as CCMV coat protein, but has significantly different loops extending from this core.




```
      1           11          21          31          41
CMV   MDKSESTSAG RNRRRRPPRG SRSAPSSADA NFRVLSQQLS RLNKTLAAGR
CCMV  MSTVGTGKLT RAQRRAAARK NKRNTRV... .........V QPVIVEPIA.
      1           11          21                  29ßQ6

      51          61          71          81          91
CMV   PTINHPTFVG SERCRPGYTF TSITLKPPKI DRGSYYGKRL LLPDSVTE..
CCMV  .....SGQGK AIKAWTGYSV SKWTASCAAA EAKVTSAITI SLPNELSSER
                  43          53ßB        63   ßC     73 αCD1 αCD2

      99          109         119         129         139
CMV   YDKKLVSRLQ IRVNPLPKFD STVWVTVPHV PASSDLSVAA ISAMFADGAS
CCMV  NKQLKVGRVL LWLGLLPSVS GTVKSCVTET QT.......T AAASFQV..A
      83      ßD  93          103 ßE      113         116αEF

       149        158         168         177         187
CMV   .PVLVYQYAA SGVQANNKLL YDLSAMR.AD IGMMRKYAVL VYSKDDALET
CCMV  LAVADNSKDV ......VAAM YPEAFKGITL EQLAADLTIY LYSS.AALTE
      124ßF             ßG  138         148αGH  ßH  158

      197         207
CMV   DELVLHVDIE HQRIPTSGVL .PV.       HELIX  ß-SHEET
CCMV  GDVIVHLEVE HVRPTFDDSF TPVY          Disordered
      167   ßI  177 αCT     187
```

Fig. 7. Structural comparison between C subunits of Cucumber mosaic virus (CMV) and CCMV. (Top) The C-α backbone of CMV is shown in red and CCMV in blue. The SwissPDB (v3.63b) program was used for this alignment (CMV PDB ID, 1F15). The structurally related CMV C capsid subunit has a similar core fold as CCMV coat protein, but has different loops (βE-αEF:approximate CCMV amino acid position 114, βF-βG:approximate CCMV amino acid position 129) extending from this core. (Bottom) Sequence homology based on structural alignments (provided by Smith, T, Purdue University). Gray regions represent disordered regions, helices are in red, β-strands in blue. The nomenclature used for secondary structure is the same as that used for CCMV.

Selection of Insertion Sites

Using the program, Swiss-PDBViewer (version (v) 3.6b3), individual residues within CCMV subunits and capsomers (pentamer and hexamer) were colored according to their temperature factor (B-factor), from dark blue for low B-factor to red for high B-factor (Fig. 8). Residues in exterior loops clearly have greater B-factors than core residues (dark blue) as seen in both the pentamer and hexamer forms. This "B-factor mapping" suggests that the five exterior loops observed in the coat protein structure (Figs. 6 and 7) are generally less constrained than the coat protein core. They might therefore be amenable to insertions. Certain residue pairs within loop regions (e.g. βB-βC loop: amino acid positions 62 and 63, βD-βE:amino acid positions 101 and 102, βF-βG:amino acid positions 129 and 130, βE-αEF:amino acid position 114 and 115, and βH-βI:amino acid position 160 and 161) have noticeably high B-factors. These regions were selected for insertion of the peptide 11-based sequence, GICDPGYIGSRCGIL (Fig. 10).

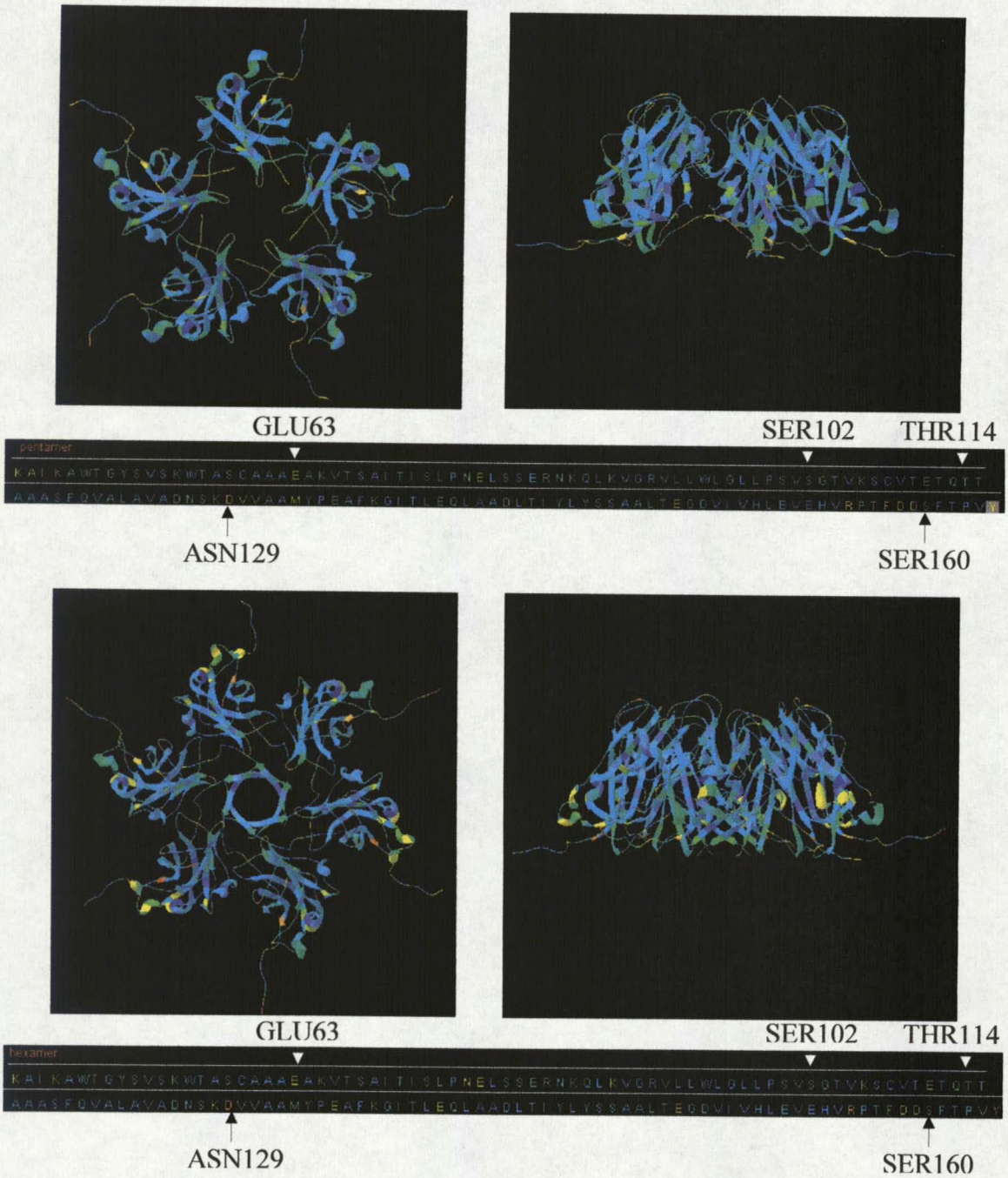


Fig. 8. B-factor mapping of CCMV coat protein capsomers. Top: two views of pentamer, colored by calculated residue-wise B-factor. Bottom: two views of hexamer, colored by calculated residue-wise B-factor. Each residue (single letter amino acid code) is colored for associated B-factor below each corresponding pair of modeled structures. Residues with especially high B-factors within exterior loops are indicated. Images were rendered using SwissPDB Viewer (v3.63b).

Again using Swiss-PDBViewer (v3.6b3), each amino acid in the coat protein was colored by its relative accessibility ("accessibility mapping") to differentiate core amino-acids from surface ones (Fig. 9). Dark blue correlates to completely buried amino-acids, red color correlates to amino-acids with at least 75% of their relative surface "totally" accessible. Although none of the loop residues are "totally" accessible, exterior loop positions appear less buried than core residues, as would be expected. Certain residue pairs within each loop appear especially accessible (relative to neighboring residues). These are: βB-βC:amino acid positions 63 and 63, βD-βE:amino acid position 101 and 102, βF-βG:amino acid position 129 and 130, βE-αEF:amino acid position 114 and 115, and βH-βI:amino acid position 160 and 161. These regions (also identified by B-factor mapping; see above) were selected for insertion of the peptide 11-based sequence, GICDPGYIGSRCGIL (Fig. 10).

Inter-subunit association energies were calculated based on the atomic buried surface areas multiplied by solvation parameters (Appendix A). Low residue-wise contributions to these association energies correlate well with a high accessibility color. Specific amino acids within exposed loops that were noted for their high accessibility had concomitantly low inter-subunit association energies.

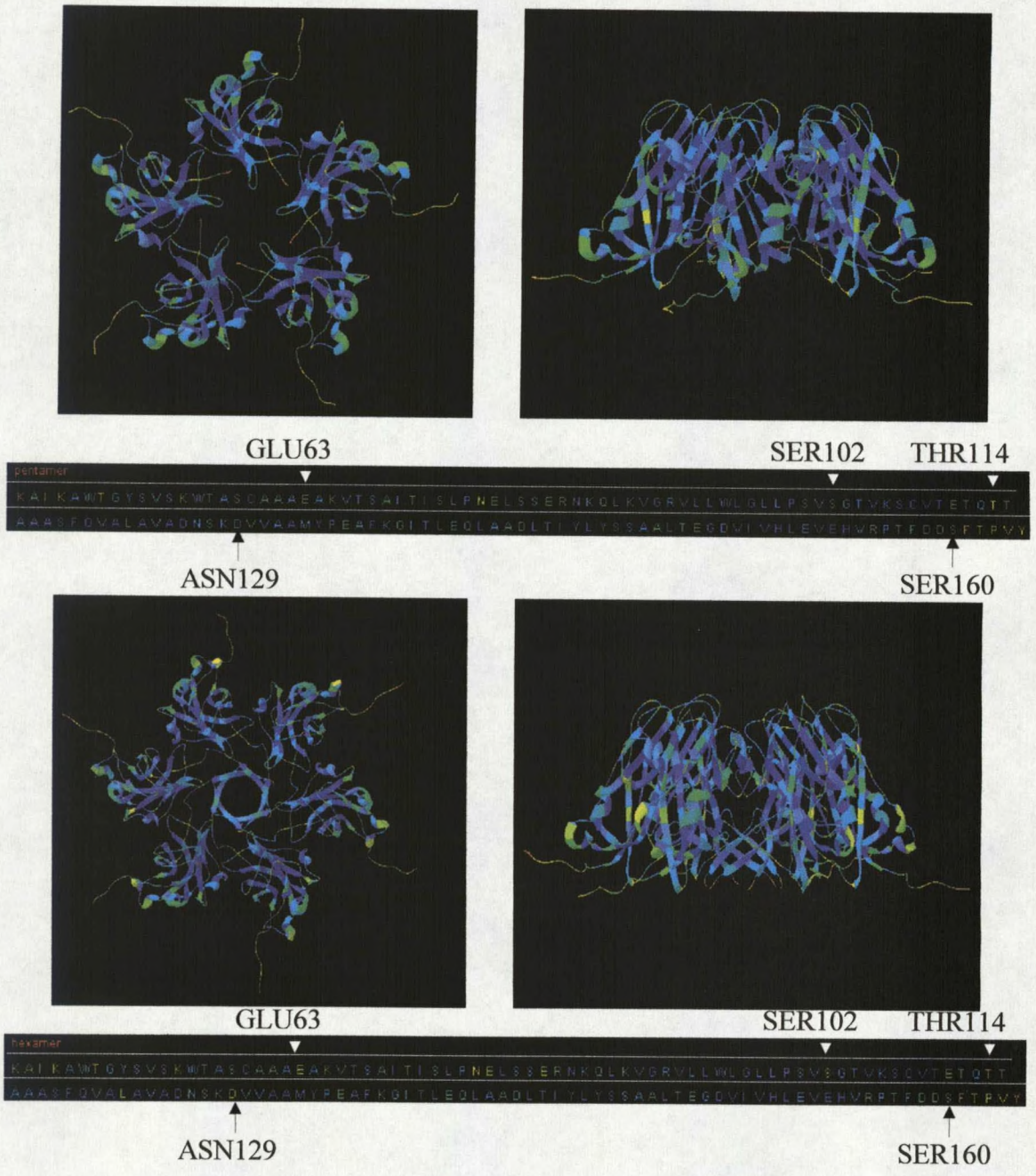


Fig. 9. Accessibility mapping of CCMV coat protein capsomers. Top: two views of pentamer, colored by relative accessibility. below. Bottom: two views of hexamer, colored by relative accessibility. Each residue (single letter amino acid code) is colored for accessibility below each corresponding pair of modeled structures. Residues within exterior loops that are especially accessible (exposed) are indicated. Images were rendered using SwissPDB Viewer (v3.63b).

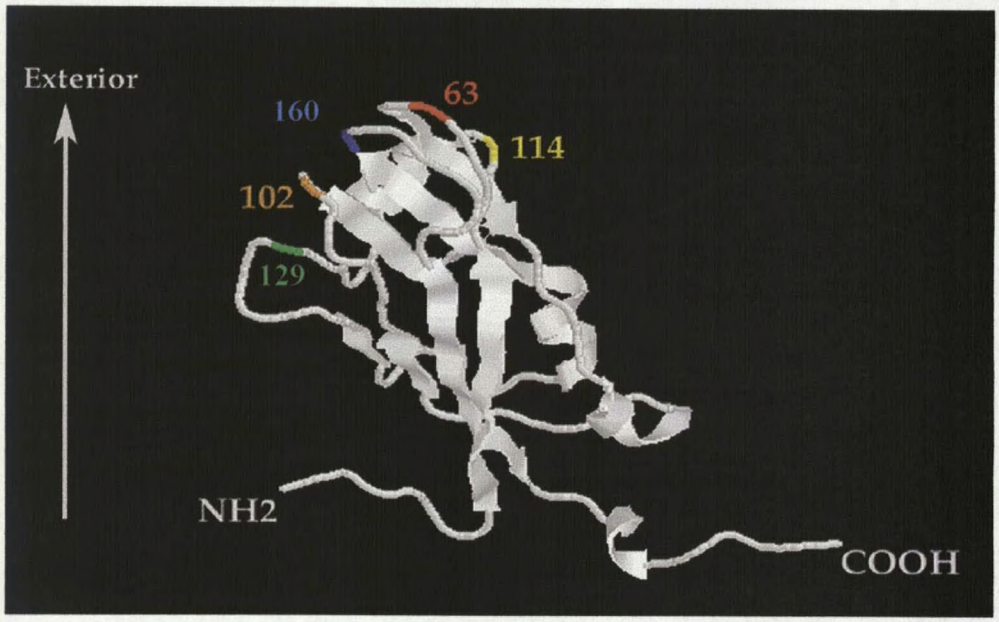


Fig. 10. CCMV coat protein with exterior loop insertion sites labeled according to wild type CCMV coat protein amino acid positions. Image was rendered using SwissPDB Viewer (v3.63b).

Engineering CCMV-Peptide 11 Chimeras (Plants)

Insertion of Peptide 11 Sequence into Loop Positions

As a first attempt to express exogenous peptide 11 in each of the five surface loops discussed above, we decided to insert the peptide 11 sequence (CDPGYIGSRC) *within* the loop sequence (as opposed to replacing CCMV sequence with a foreign one). This was accomplished using PCR oligonucleotide-directed mutagenesis to introduce a unique BamH1 restriction site into selected regions of the CCMV coat protein cDNA. These sites (amino acid positions 63, 102, 114, 129, and 160 in the wild-type coat protein) were identified during the design stage (above). An oligonucleotide encoding the peptide 11 insert plus additional residues to enable cloning with engineered BamH1 ends (GICDPGYIGSRCGIL) was ligated into each of the BamH1 plasmid constructs using standard techniques. Construction of the chimeric virus with the peptide 11-derived sequence in each respective loop site is illustrated schematically as pCC3/pep11 in Fig. 11.

Proper cloning (in frame, and forward orientation of peptide 11 sequence; no detrimental mutations) of each construct, pCC3/pep11-63, -102, -114, -129, and –160, was first confirmed by sequencing purified plasmid constructs. Sequencing data are show in Appendix B. Peptide 11 sequences shown are clearly in a forward orientation, are monomers, and are in frame with the coat protein gene. There were no apparent unintended mutations in the clones we selected. We had difficulties sequencing

pCC3/pep11-160 clones and due to time constraints (and other problems, see below) we dropped this construct from the design/engineering cycle.

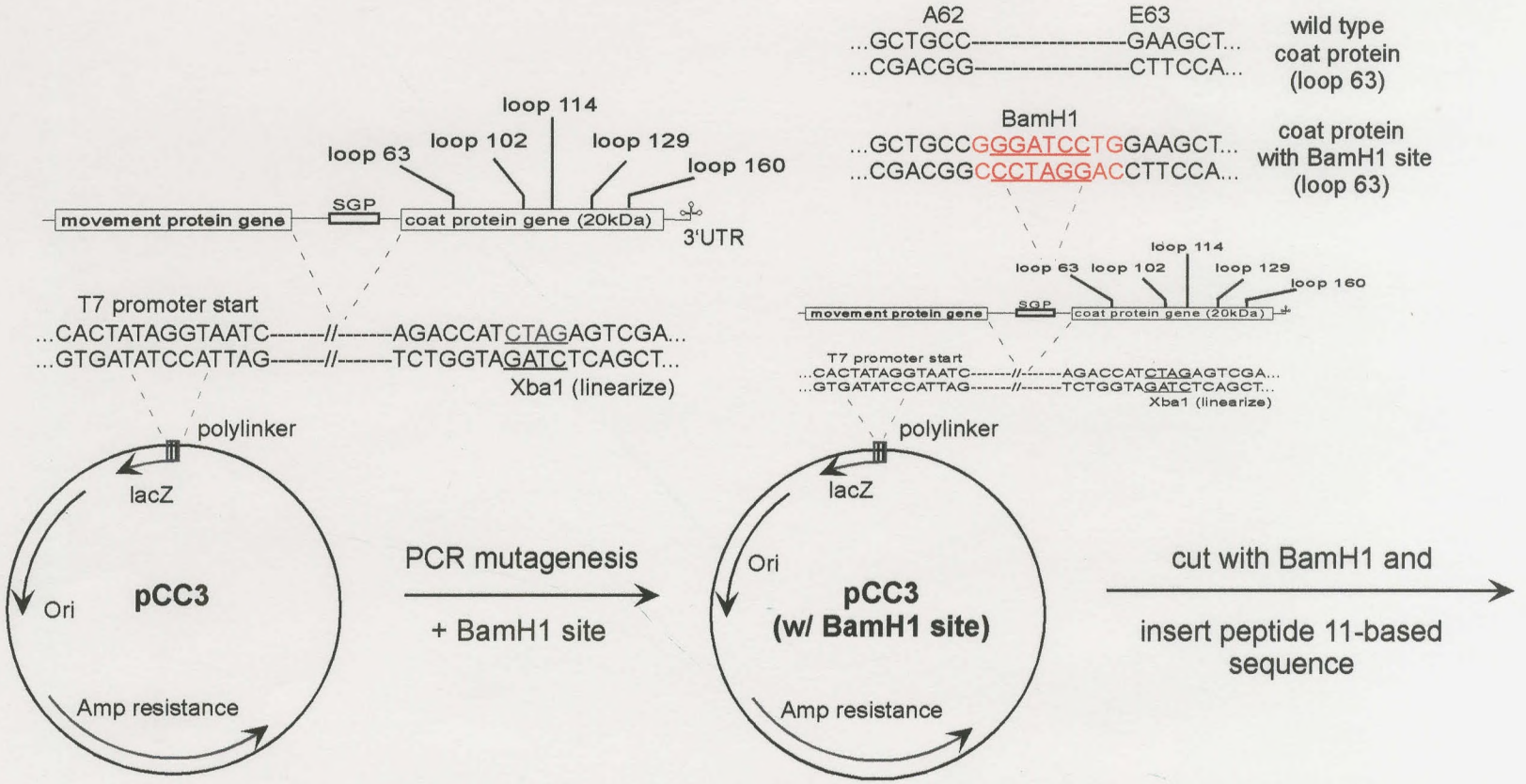


Fig. 11. Schematic representation of plasmid-based vectors used as templates for making infectious RNA. The viral subgenomic promoter element (SGP) and the 3' untranslated region (3'UTR) are indicated. Nucleotide additions (made to wild type base plasmid) are shown in red. Xba1 and BamH1 restriction sites are underlined. The BamH1 restriction site was engineered into the CCMV coat protein gene at five positions: amino acid 63, 102, 114,129, and 160. These positions correspond to exterior loop insertion sites. CCMV coat protein gene with a BamH1 restriction site between corresponding amino acid positions A62 and E63 is shown as an example of the constructs that were made.

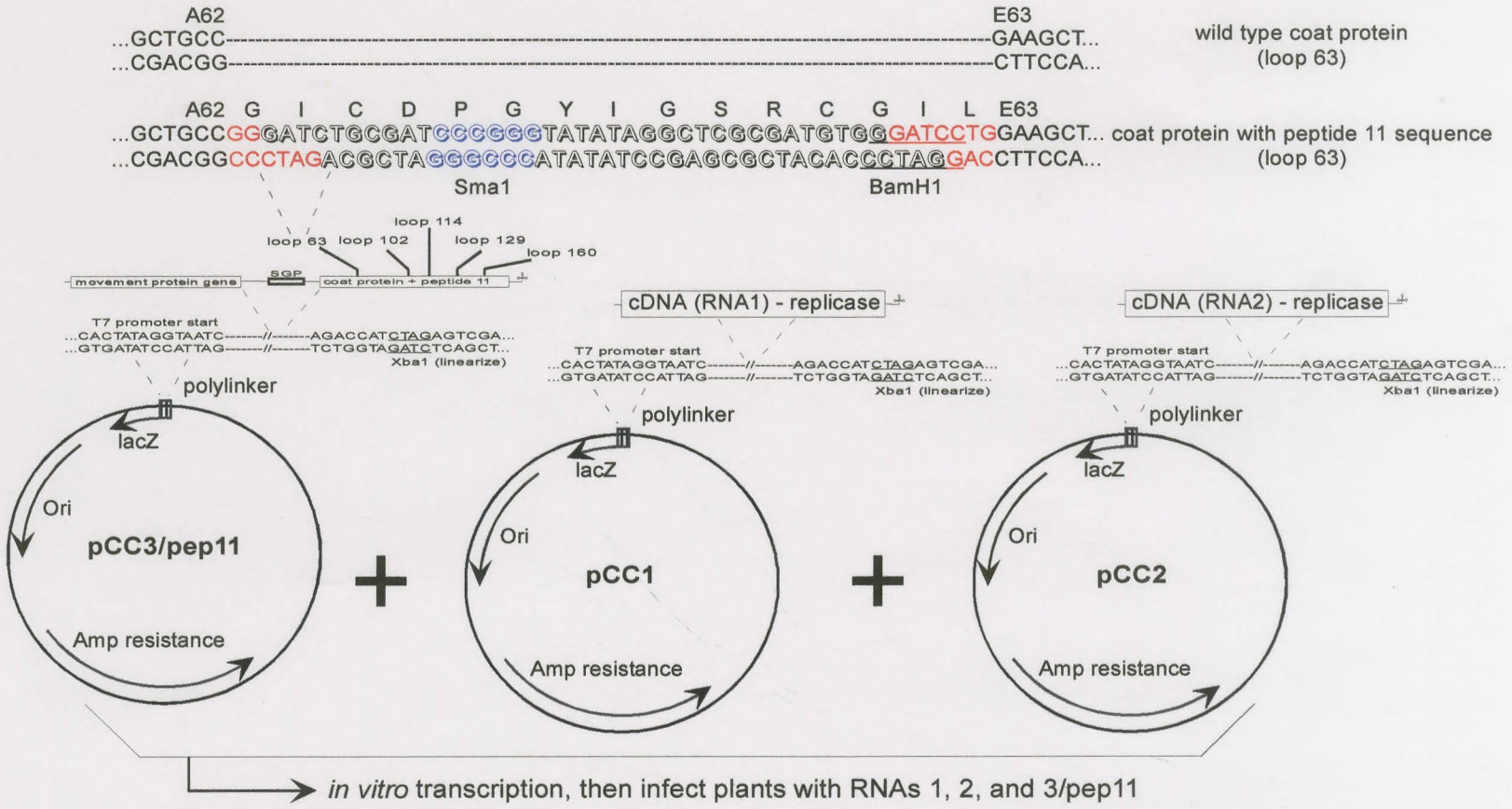
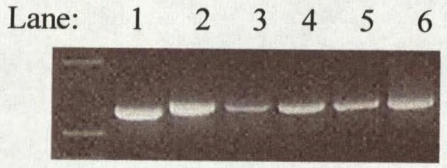


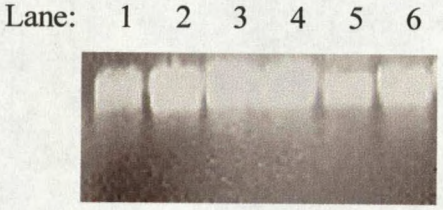
Fig. 11 (continued). Schematic representation of plasmid-based vectors used as templates for making infectious RNA. Nucleotide additions (made to wild type base plasmid) are shown in red (BamH1 cloning site) or are embosed (peptide 11-based sequence). Xba1 and BamH1 restriction sites are underlined; the Sma1 restriction site is in blue. The peptide 11-based sequence (amino acids shown above DNA sequence) was inserted into the CCMV coat protein gene at each of five previously engineered BamH1 sites (loops 63, 102, 114,129, and 160). CCMV coat protein gene with peptide 11 between corresponding amino acid positions A62 and E63 is shown as an example of the pCC3/pep11 constructs that were made.

Proper cloning of each construct, pCC3/pep11-63, -102, -114, -129, and -160, was confirmed independently from sequencing results using a coupled *in vitro* transcription/translation assay for full-length protein production. Fig. 12 (top) shows the PCR amplified fragments generated from pCC3/pep11 plasmids that were used for *in vitro* transcription. Notice the characteristic band shift in chimeric DNAs due to 45 nucleotide peptide 11-derived insert. Fig. 12 (middle) verified RNA integrity prior to *in vitro* translation. Fig. 12 (bottom) shows *in vitro* translation coat protein products visualized by autoradiography. The protein products from all constructs except pCC3/pep11-160, which did not express at all (Fig. 12 bottom, lane 6), were easily detectable. Coat proteins with peptide 11 inserts can be differentiated from wild type coat protein by the slight mobility shift (due to the 15 amino acid insert). There were few, if any, noticeable degradation products on these gels. Because of difficulties expressing pCC3/pep11-160 *in vitro*, and because of sequencing difficulties (above), we decided not to further pursue this construct.

T7l-PCR (coat protein gene) product:



In vitro TXN (capped RNA) product:



In vitro TLN (coat protein) product:

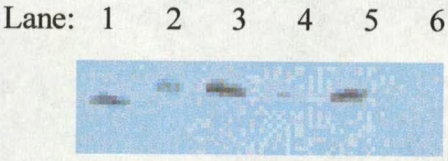


Fig. 12. Coupled *in vitro* transcription/translation from pCC3/pep11 constructs. Lanes 1-6 (all figures): pCC3-wild type control; pep11-63; pep11-102; pep11-114; pep11-129; pep11-160. Top: Nested set complimentary deoxyoligonucleotide primers that flank the coat protein gene and add a bacteriophage T7 RNA polymerase promoter element to the 5' end of amplified products were used for site-specific DNA amplification. Products were phenol-chloroform extracted and ethanol precipitated prior to separating on a 1.5% agarose gel. DNA was visualized by staining with ethidium bromide and illuminating with UV light. Middle: Chimeric CCMV-peptide 11 RNAs were synthesized from approximately 1 μg of T7l-PCR amplification product by *in vitro* transcription using T7 mMessage mMachine (Ambion, Austin, TX). Products were phenol-chloroform extracted and ethanol precipitated prior to separating on a 1.2% agarose/6.6% formaldehyde gel. RNA was visualized by staining with ethidium bromide (prior to electrophoresis) and illuminating with UV light. Bottom: Capped *In vitro* transcribed RNAs were translated using cell free wheat germ extracts. Radioactive [^{35}S]methionine was incorporated into translation products which were separated by 12.5% SDS-PAGE and visualized using autoradiography (12 hour exposure). Notice mobility shifts in PCR products and *in vitro* translation products from peptide 11 constructs. Also note that the pCC3/pep11 construct did not translate (bottom image, lane 6).

Expression of CCMV-Peptide 11 Chimeras in Plants

Capped transcripts from wild type and pCC3/pep11-63, -102, -114, -129 plasmid constructs (along with pCC1 and pCC2 transcripts for viral replication) were inoculated onto cow pea plants (*Vigna unguiculata* (L.) var. California Blackeye) and tobacco plants (*Nicotiana benthamiana*). Representative examples of the transcripts used for inoculation are shown in Fig. 13. Normally, capped RNA transcripts made from wild type CCMV DNA template are able to initiate a local infection in the leaves onto which they are inoculated (primary leaves). Assembled virions spread cell-to-cell within primary leaves, then spread to other plant tissues (mostly secondary leaves) via the plant vasculature. Infected leaves can be homogenized (to release virus particles) and "passaged" to other, naïve (uninfected) plants. Often, the host shows signs that are symptomatic of infection. Leaves can become mottled in color, develop necrotic lesions, and in some instances, tissues will crinkle. If noticeable, these symptoms can be used to characterize the infection.

The course of infection and virus passaging can be also be followed using two independent analytical assays: ELISA-based detection of the protein capsid and an RT-PCR assay, which detects viral (v)RNA from plant extracts. Using these two assays, CCMV-peptide 11 constructs were evaluated (relative to wild type CCMV virus) for their ability to initiate an infection in primary leaves and for their ability to establish a systemic infection (spreading to secondary tissues). Constructs were also tested for their ability to passage from plant-to-plant. Figures 14-16 show the results from these tests.

CCMV coat protein-peptide 11 constructs were able to infect plant hosts with only limited success: Infected plants showed no symtomology. Based on ELISA values, local infection resulted in CCMV-peptide 11 (loop 129) coat protein accumulation at or near wild type levels in primary leaves (Fig. 14). Western blotting of protein extracts (Fig. 15) was performed using the same anti-CCMV coat protein polyclonal serum that was used to detect virions by ELISA. The western blot shows a mobility shift that is characteristic of the larger CCMV-peptide 11 chimera (chimera in left lane compared to wild type in the right lane). There are no signs of major degradation (cleavage products) observed. Both the CCMV-peptide 11 chimera and wild type coat protein showed apparent dimers (banding at approximately 40 kDa) on this 20% SDS-PAGE/Western blot. While dimers are commonly observed in wild type sample preparations, it is important to note that the dimer band from peptide 11 chimera preparations was apparently darker, possibly indicating a higher concentration of the dimer in these extracts relative to wild type extracts. The formation of dimers (as well as other higher-order aggregates) in peptide 11 chimera sample prepartations is an issue that is currently being investigated (see Chapter 5, Conclusions).

Protein from pCC3/pep11-102, -114, and –63 constructs was not detectable by ELISA (data not shown). Based on ELISA, CCMV-peptide 11 (loop 129) coat protein accumulation was not significant in secondary tissues and was not detectable in naïve plants that were inoculated with homogenate from infected primary leaves (Fig. 14).

RT-PCR analysis (Fig. 16) gave independent results that were similar to those found from ELISA-based assays. RT-PCR products (cDNA) using vRNA from

pCC3/pep11-102, -114, and –63 constructs as template was not readily detectable in primary or secondary leaves (data not shown). Primary leaves infected with CCMV-peptide 11 (loop 129) gave a RT-PCR product that was shifted relative to wild type control (top gel, left lane, Fig. 4d). This cDNA product was fully digested with SmaI restriction endonuclease (bottom gel, left lane, Fig. 16). Since the SmaI restriction site is unique to the peptide 11 DNA sequence, this result confirmed the presence of peptide 11 DNA within the CCMV coat protein coding sequence. As expected, the wild type derived coat protein RT-PCR product did not cut with SmaI (bottom gel, middle lane, Fig 16). Sequencing (Appendix B) of the cDNA product, (with presumed peptide 11 insert) isolated from primary tissue confirmed its pCC3/pep11 (loop 129) origin. Sequencing also independently demonstrated that the CCMV-peptide 11 (loop 129) construct was genetically stable with no evidence of reversion to wild-type sequence.

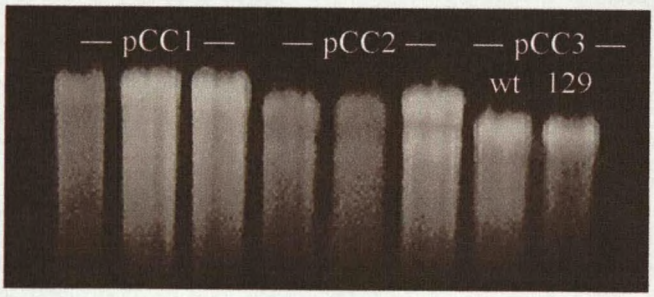


Fig. 13. Examples of RNA used to inoculate plants. Plasmids (pCC1, pCC2, pCC3 (wild type), and pCC3/pep11 plasmid constructs) were prepared for *in vitro* transcription by linearization with Xba1 followed by Proteinase K treatment. Chimeric CCMV-peptide 11 RNAs were synthesized by *in vitro* transcription from approximately 1 μg of linearized DNA using T7 Megascript (Ambion, Austin, TX). Transcription products were run on a 1.2% agarose/6.6% formaldehyde gel. Samples were stained with ethidium bromide and visualized by UV illumination. Samples above are labeled according to the plasmid DNA template used to make them: the right two lanes (pCC3) are from a wild type coat protein DNA template (wt) and from pCC3/pep11-129 template (129).

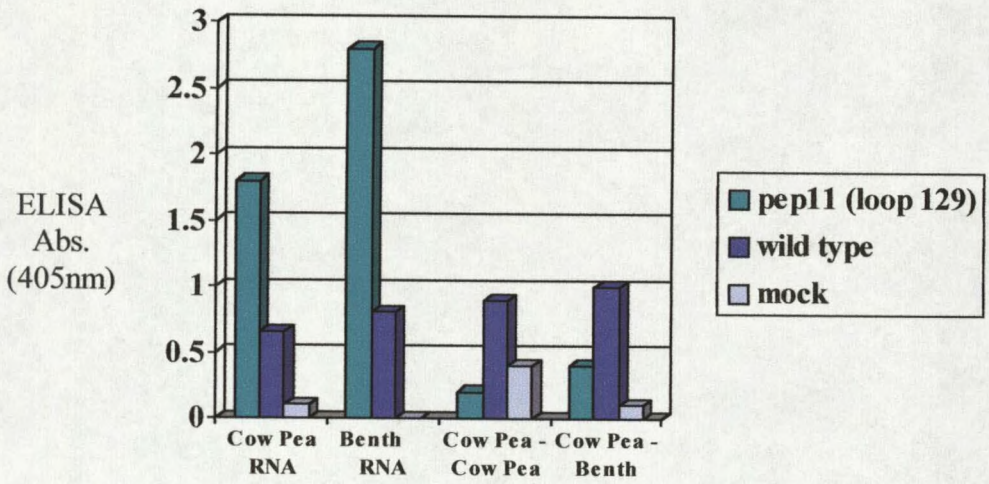


Fig. 14. ELISA results from plants. *In vitro* RNA transcripts were used to inoculate cow pea plants (*Vigna unguiculata* (L.) var. California Blackeye) and tobacco plants (*Nicotiana benthamiana*, "Benth") (left two bar sets, respectively). If a productive infection was established, infected leaves were homogenized and used as inoculum for "passaging" the infection to naïve plants (right two bar sets). 100 µl of antigen, prepared by grinding leaf tissue in water or antigen buffer, was used to coat wells in a "double-sandwich" ELISA. A 1:500 dilution (in antigen buffer) of alkaline phosphatases (AP)-conjugated rabbit anti-CCMV polyclonal antibody was used to detect CCMV capsid protein (wild type or peptide 11 chimeras) with a color readout (registered by absorbance at 405nm, Y-axis). RNA (primary) inoculation resulted in wild type and chimeric (peptide 11 in loop 129) protein production. The peptide 11-loop 129 construct did not move to secondary tissue (not shown) and did not passage to other plants as well as wild type does (right two bar sets).

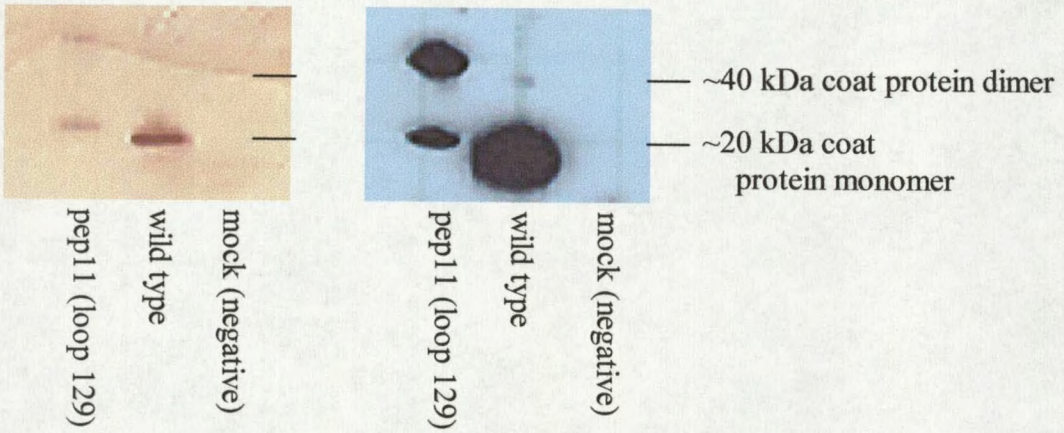


Fig. 15. Western blot of coat protein products from plants. 100 mg plant extract from RNA inoculated cow pea plants was separated by 20% SDS-PAGE and then Western blotted: 1:5000 1^{o} - Rabbit-anti-CCMV (polyclonal); 1:5000 2^{o} - goat-anti-rabbit (horseradish peroxidase conjugated). Blot was visualized using Opti-4CN cholorimetric detection (left) and pierce "supersignal" chemiluminescence - 10 sec exposure (right). Notice that the chimeric coat protein (pep11-loop 129) runs slightly higher than wild type, indicating incorporation of peptide 11. No significant protein cleavage is apparent in the coat protein loop 129 chimera as a result of the insertion. Also note the seemingly predominant ~ 40 kDa dimer form of the chimeric coat protein (upper band in left lanes).

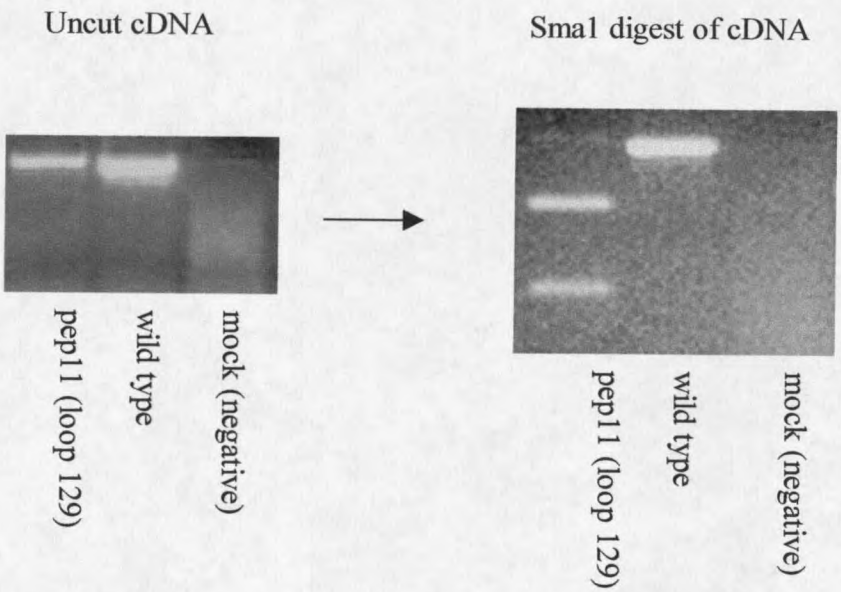


Fig. 16. RT-PCR amplification and diagnostic restriction digestion. RT-PCR cDNA synthesis was performed using viral RNA extracted from plant tissue. Nested set complimentary deoxyoligonucleotide primers that flank the coat protein gene were used for site-specific amplification. Sma1 restriction digestion was used to confirm presence of peptide 11 sequence in cDNA isolates. DNA products were phenol-chloroform extracted and ethanol precipitated prior to separating on 1.5% agarose gels. DNA was visualized by staining with ethidium bromide and illuminating with UV light.

Taken together, these results indicate that the CCMV-peptide 11 (loop 129) construct is able to establish an infection in the primary leaves of host plants. Other constructs were unable to initiate a detectable infection. The initial infection from the loop 129 construct was generally non-productive since efficient movement within the plant and passaging from plant-to-plant appeared to be compromised. It seems that the base construct did not significantly affect coat protein expression, but may have altered

stable capsid assembly. Presumably, unstable virions would be limited in their ability to move and to be passaged since stable packaging of the RNA genome is considered essential to virus infectivity. Although some stable capsids might be retrievable from plants (we attempted, with limited success, to purify the loop 129 construct from primary leaves; data not shown), this system would likely be costly, time-consuming, and technically inconvenient. Plant hosts are therefore impractical for large-scale production of CCMV-peptide 11 constructs. For this reason, we chose to express the chimeric capsids in *Pichia pastoris*.

Engineering CCMV-Peptide 11 Chimeras (*P. pastoris*)

In order to meet the demand for large amounts of easily to purify capsid constructs, we chose to transfer the core of the pCC3/pep11 expression cassettes into the *P. pastoris* heterologous expression system. This system has been used reliably for its relatively simple molecular genetic manipulation, the ability of *P. pastoris* to produce foreign proteins at high levels, the capability of performing numerous eukaryotic posttranslational modifications, and because of its similarity to the well-characterized experimental yeast system, *Saccharomyces cerevisiae*. It was previously demonstrated that *S. cerevisiae* can function as a surrogate host to brome mosaic virus (BMV), a close relative to CCMV[75]. Also, initial attempts at expressing wild-type-derived CCMV coat protein in *P. pastoris* resulted in high level expression and stable assembly of soluble virion particles (data to be published elsewhere).

Construction of pPicZA/pep11
Plasmids and *P. pastoris* Transformation

The stable genomic integration of modified coat protein genes (CCMV with peptide 11 inserted into loops 63, 102, 114 and 129), into the *P. Pastoris* chromosome was performed in two steps (see Fig. 17). First, using standard PCR and cloning techniques, the CCMV coat protein genes (with integrated peptide 11 sequence and without superfluous 3'untranslated regions) were subcloned from their original plasmid vector (pCC3/pep11) into the *P. pastoris* expression plasmid pPicZA. The resulting plasmid constructs were called pPicZA/pep11-63, -102, -114 and -129 (Fig 17). Correct insertion of the subcloning sequence into pPicZA was confirmed by SmaI restriction analysis (Fig. 18) and by sequencing (Appendix B).

After linearization by SacI restriction digestion, plasmids were transfected into *P. pastoris* X-33 mut^+ yeast, and Zeo^r transformants were selected. Appropriate gene integration was confirmed by a lyticase-PCR assay (Fig. 19). Modified coat protein gene fragments were amplified from *P. pastoris* genomic DNA isolates using primers internal to the coat protein gene and flanking the peptide 11 insertion. PCR products from yeasts transformed with pPicZA/pep11 constructs had mobility shifts relative to a wild type control characteristic of the peptide 11 insert (Fig. 19, top). SmaI digestion of PCR products validated the presence of the peptide 11 sequence within the coat protein gene (Fig. 19, bottom). Sequence analysis (Appendix B) confirmed this result.

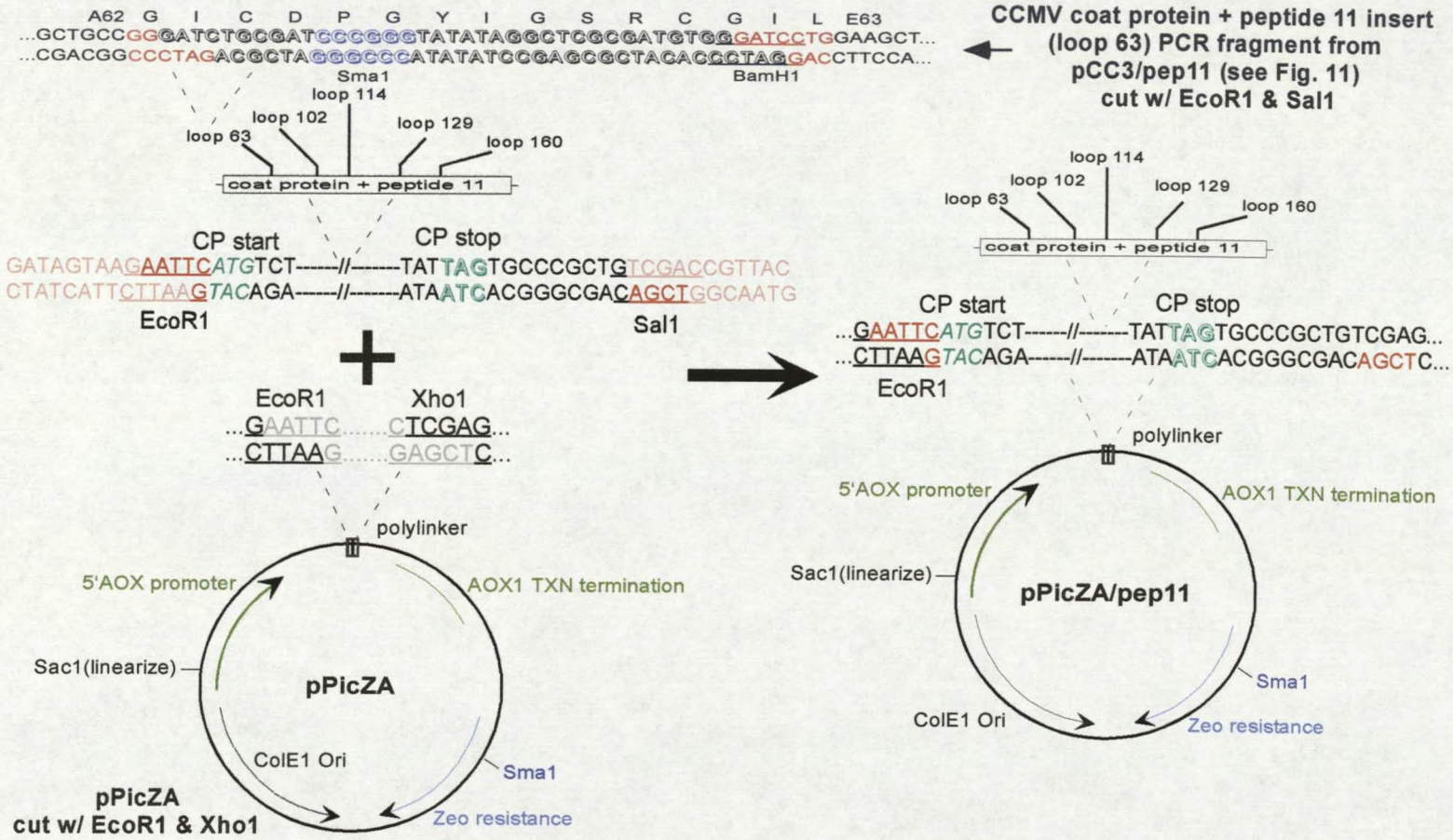


Fig. 17. Schematic representation of pPicZA/pep11 construction. The CCMV coat protein gene, with the peptide 11 sequence inserted into exterior loops (loop 63 is illustrated as an example), was cloned into pPicZA to construct pPicZA/pep11. Nucleotide additions are shown in red. EcoR1, Sal1, Xho1 and BamH1 restriction sites are underlined; the Sma1 restriction site (in the coat protein + peptide 11 PCR fragment and in the Zeocin resistance gene) is blue. Nucleotides that were digested and lost during cloning are indicated by faded coloring. The coat protein start codon, ATG, is italicized and colored green. The coat protein stop codon is embosed and colored green. pPicZA/pep11 constructs were linearized (at Sac1 restriction site) and transfected into *P. pastoris* for coat protein expression.

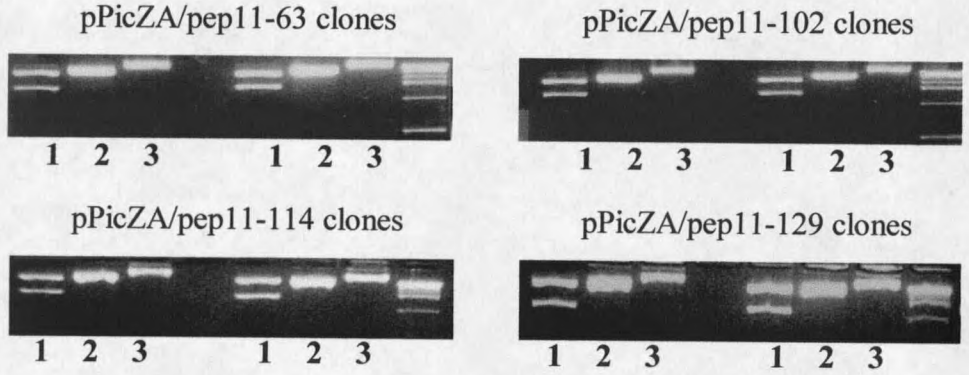


Fig. 18. Screening for coat protein plus peptide 11 insertion into pPicZA. Transformed *E. coli* colonies (grown under Zeocin selection) were miniprep'd and screened by enzymatic digestion with SmaI (lane 1). The SmaI restriction site is found in peptide 11 and in the Zeocin resistance gene of the pPicZA plasmid, so that cleavage results in two distinct bands. If the insert was present, miniprep'd plasmid DNA was linearized with SacI (lane 2), purified and electroporated into *P. pastoris* X-33 mut$^+$ yeast (see transformants, Fig. 19). Lane 3 shows uncut miniprep'd pPicZA plasmid for comparison. DNA products were phenol-chloroform extracted and ethanol precipitated prior to separating on 1.5% agarose gels. DNA was visualized by staining with ethidium bromide and illuminating with UV light.

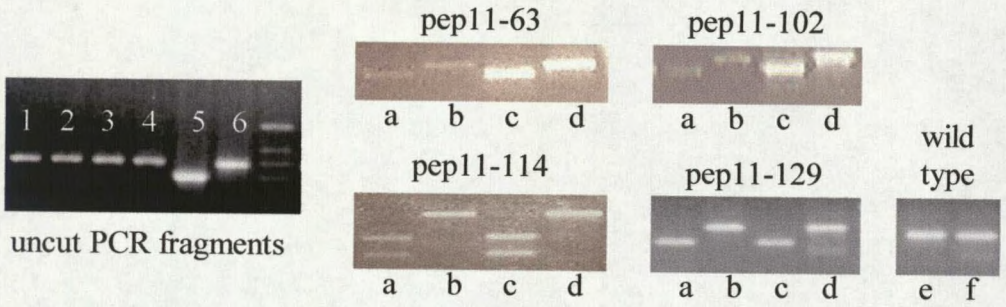


Fig. 19. Lyticase PCR-based screening (integration into *P. pastoris* genome). PCR amplification products were generated from electroporated *P. pastoris* genomic DNA using internal primers to the coat protein gene flanking the peptide 11 insertion (left panel: coat protein + peptide 11 DNA in position –63 (lane 1), -102 (lane 2), -114 (lane 3), -129 (lane 4), wild type coat protein DNA PCR product amplified from pCC3 template (lane 5), PCR product from pPicZA/pep11-129 template for positive control (lane 6)). These fragments were digested with Sma1 to confirm identity (right five panels: Sma1 digested PCR product (lane a), uncut PCR product (lane b), Sma1 digested PCR products from corresponding pPicZA/pep11 constructs (lane c), uncut PCR products from corresponding pPicZA/pep11 constructs (lane d), Sma1 cut and uncut (lanes e and f, respectively) wild type coat protein DNA PCR product amplified from pCC3 template.) DNA products were phenol-chloroform extracted and ethanol precipitated prior to separating on 1.5% agarose gels. DNA was visualized by staining with ethidium bromide and illuminating with UV light.

Small Scale Time Course Expression Screening

The expression of modified coat proteins in transformed *P. pastoris* was examined by regularly adding methanol over five days after an initial (three day) glycerol-fed growth phase. Methanol should actively induce the alcohol oxidase (*AOX*) promoter, which controls expression of stably integrated coat protein constructs. Induced cell lysates were followed for maximum accumulation of coat protein over the five-day methanol feed by dot blot. All *P. pastoris* isolates expressed the integrated coat protein construct, reaching steady-state protein accumulation by day two and maintaining these levels for 2-4 days depending on the isolate (Fig. 20). The 129-50c isolate produced coat protein beyond 5 days (Fig. 20, bottom). Coat protein production stopped, and accumulated coat protein levels dropped off, between 2-4 days for most cultures (all those except the 129-50c culture). A possible explanation for this observation is that, due to disruptions in the culturing environment such as inadequate oxygen diffusion or inadequate methanol induction of the AOX1 promoter, these cultures may have died at this time point in the feeding. Accumulated protein may then have degraded (beyond the point of detectable levels) before the next time point sample was collected. The 129-50c culture may have had greater oxygen diffusion and/or more consistent induction of the AOX1 promoter, explaining the apparent continued production of chimeric coat protein beyond day 5. Alternatively, greater protein production (and longer-lived accumulation levels) in this culture may have been due to multiple integrations of the expression

cassette into the *P. pastoris* genome, although further investigation would be needed to prove this.

Western blotting using polyclonal CCMV antibodies raised in rabbits verified the integrity of chimeric protein from cell lysates. As shown in Fig. 21, the induced cell lystates gave single immunostained bands with a characteristic mobility shift (greater relative molecular mass) when compared to wild type coat protein expressed in *P. pastoris*. This result indicates that full length CCMV-peptide 11 chimeric protein was being produced over a five day induction period with negligible degradation (Fig. 21).

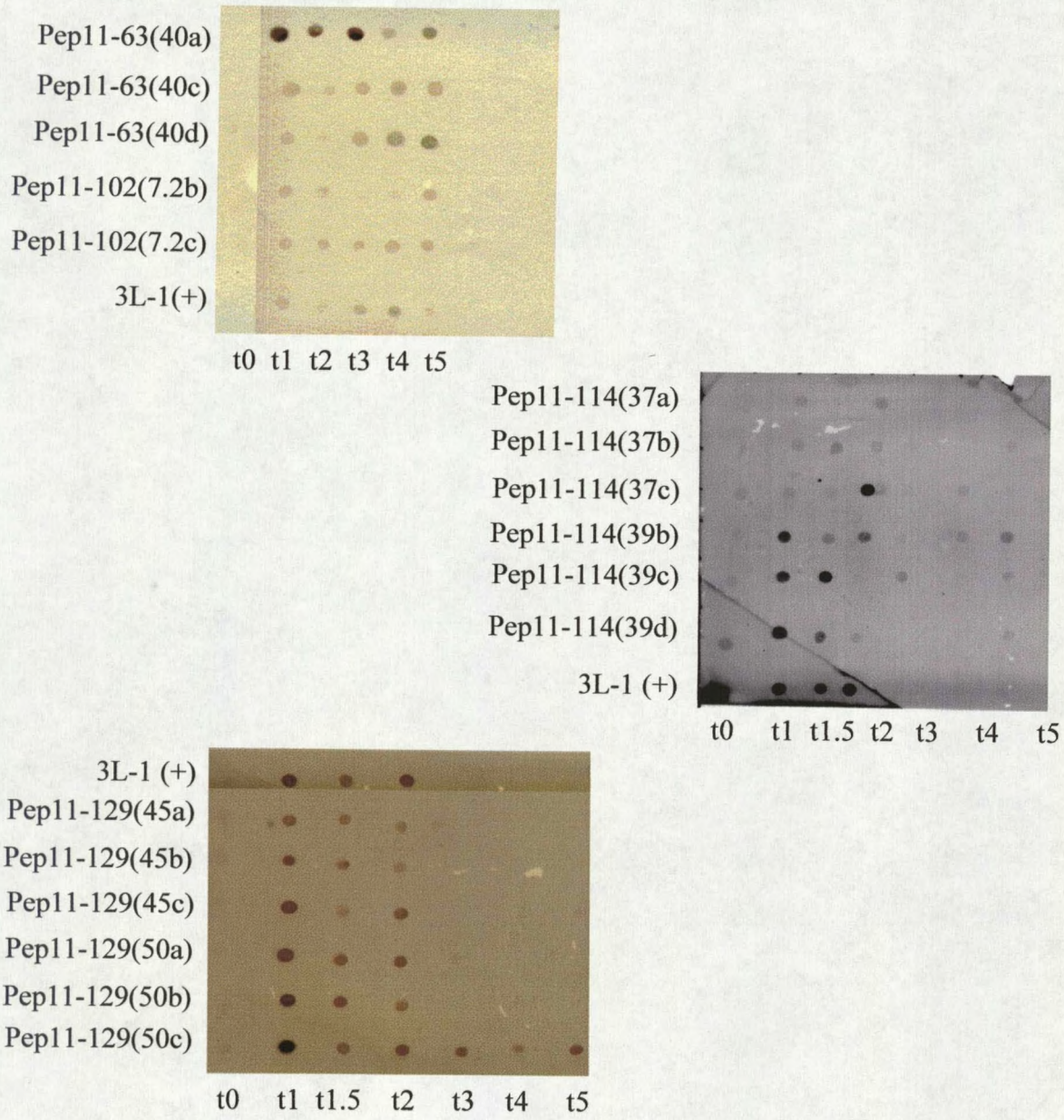


Fig. 20. Small scale time course expression screening. Single colonies (clone ID in parentheses) of *P. pastoris* carrying the CCMV coat protein gene (plus peptide 11 insert in a single loop) were grown under methanol induction for 132 hr. Expression of "wild type" CCMV coat protein construct without peptide 11 insert (expressed in *P. pastoris* isolate, 3L-1) was followed as a positive control. Samples were harvested 0(t1), 24(t1.5), 36(t2), 60(t3), 84(t4), 108(t5), and 132 hrs (t6) after induction, and then disrupted. 1 µl samples were blotted onto nitrocellulose membrane and visualized using the western blot visualization protocol (see Chapter 3, Methods). Samples with greatest protein concentration (determined by visually inspecting time course dot blots) were separated by 12.5% SDS-PAGE and then were western blotted (Fig. 21).

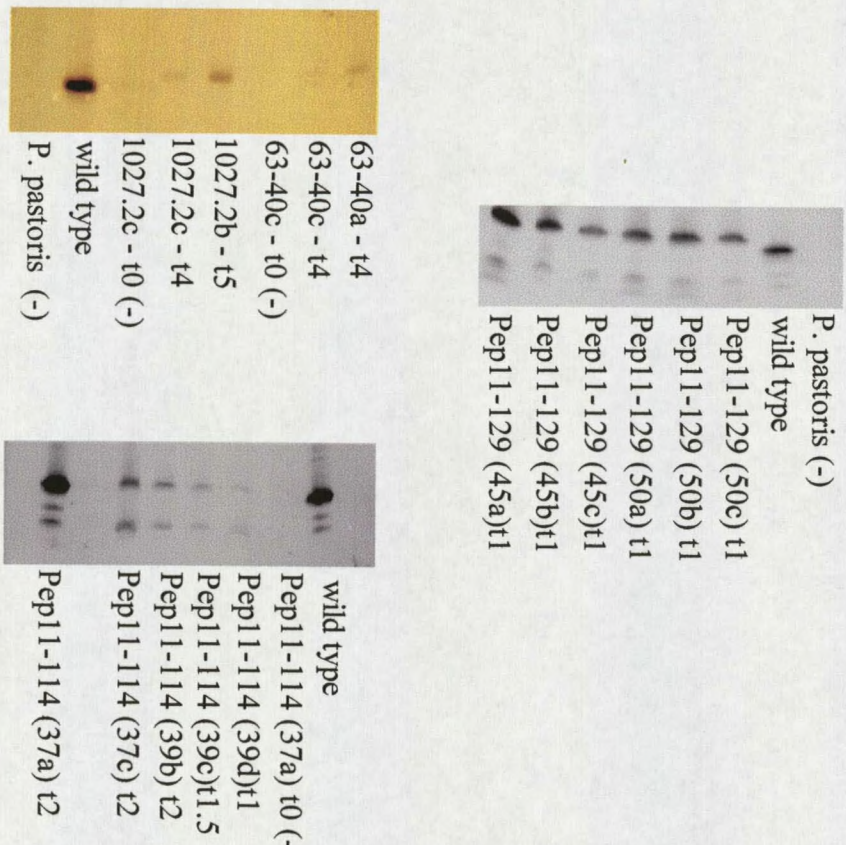


Fig. 21. Small scale time course expression screening (Western blots). Samples with greatest protein concentration (determined by visual inspection of time course dot blots, Fig. 20) were separated by SDS-12.5% PAGE, western blotted, and visualized as described in methods. A sample prepared from non-transformed *P. pastoris* X-33 mut$^+$ isolates was used as a negative control. "Wild type" coat protein without peptide 11 insert was used as a positive control. Notice that coat protein chimeras run slightly higher than wild type coat protein, indicating incorporation of peptide 11 in the chimera. No major protein cleavage appears to occur as a result of the insertion.

Intermediate Scale Expression of
CCMV-Peptide 11 Coat Protein in *P. Pastoris*

Having optimized expression in small-scale batches, we attempted a "scaled-up" expression to produce more protein. This was accomplished by increasing the culture volume in larger (1 L) baffled flasks. 500 ml cultures were grown as small-scale cultures (see Chapter 3, Methods) and yeasts were harvested when protein expression/accumulation was at a peak (as indicated by time course studies; approximately 3 days). Virions were isolated and purified by glass bead disruption followed by PEG precipitation. ELISA was used to determine the amount of isolated capsid protein (Fig. 22a). TEM was used to visualize assembled particles and to characterize the capsid morphologies (Fig. 22b).

Fig. 22a shows ELISA values from a sample isolated from an "intermediate scale" (500 ml) batch. This isolate was expressing CCMV coat protein with the peptide 11-based sequence cloned into loop 63. The purified sample had A260/A280 ratios of ~ 1.0. The sample was concentrated in Centricon™ 100 microconcentrators after dialyzing against three volume changes of virus buffer. Protein concentration increased noticeably (Fig. 22a). Electron microscopy of the concentrated sample showed that the viral protein was native-like virion particles with normal dimensions (~28 nm diameter) and morphology (Fig. 22b). There were very few particles in the viewing field. Compared to the reported yields of other *P. pastoris* expressed proteins (up to 10 g/L culture[66]), the yields of purified capsid product from this isolate were extremely low (less than 10 µg/L culture). *P. pastoris* isolates that expressed other loop constructs (loops 102, 114, 129) in

small-scale cultures did not produce protein when scaled-up to 500 ml cultures; based on ELISA, coat protein concentrations were consistently below the 0.25 ng/100 µl standard.

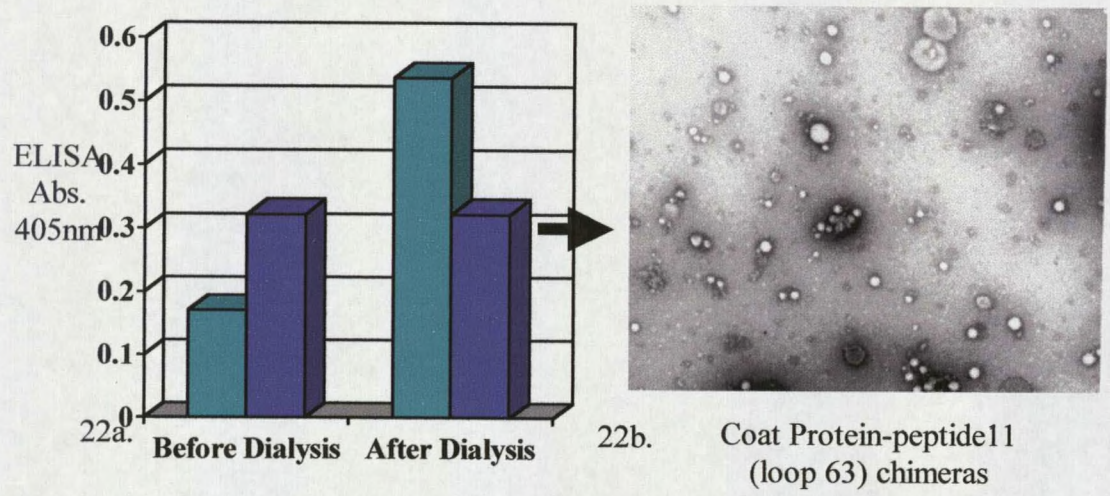


Fig. 22. Intermediate scale expression culture of CCMV-Peptide 11 Coat Protein in *P. Pastoris*. 22a. ELISA-based assay (see Fig. 14 and Chapter 3, Methods) used to detect and quantitate expressed coat protein constructs. 22b. TEM (25k 22610, 4.0sec, 2.6x) used to visualize virion particles.

These results were encouraging because particles were visualized. However, the scale-up expression and purification protocols needed to be further optimized in order to obtain higher amounts of capsid protein that could be readily purified (as virion particles) and characterized. One solution to the problem of low capsid yield was to express our *P. pastoris* isolates in high-density fermenter cultures.

Expression, Purification and Characterization of
CCMV-Peptide 11 Chimeras from High-Density Fermenter Cultures

A hallmark of the *P. pastoris* system is the ease with which expression strains scale-up from shaker-flask to large-scale high-density fermenter cultures. In fermentation, *P. pastoris* can be cultured to densities on the order of 500 OD600 units/ml with high levels of induced transcription from the *AOX* promoter during methanol feeds[66]. To obtain greater amounts of our CCMV-peptide 11 chimeras from *P. pastoris*, we scaled-up to a four-liter fed-batch fermenter culture. Strains were grown for bio-mass in a glycerol-base media and then starved from carbon source for ~30 min. A minimal methanol media was then fed to the culture to induce expression. Cultures were harvested after ~90 hrs of methanol feeding.

To purify chimeras, fermenter cultures were pelleted at low speed and then lysed by grinding with glass beads in the presence of homogenization buffer (see Chapter 3, Methods). We first tried to isolate assembled capsids from the crude lysates using a series of PEG precipitations, as was done to isolate capsids from intermediate scale preparations (see Chapter 3, Methods). Yields using this technique were low, typically 1-2 mg of CCMV capsid per 100 g of pelleted yeast (based on ELISA values, Fig. 23). To characterize and purify capsids, samples (before and after PEG precipitations) were separated on 10-40% sucrose gradients (Fig. 23). The gradients were fractionated in 500 µl aliquots while monitoring absorption at 280 nm (A280). However, protein could not be detected from sucrose gradient fractions by absorbance (protein concentrations were

below the limits of detection). Fractions were therefore analyzed by ELISA (as described above), to generate sucrose gradient profiles (Fig. 23). At least two distinct peaks can be seen in these profiles. The intensity (height) of these peaks increased with PEG precipitations, indicating that there is some capsid selectivity gained with this purification scheme. Capsid particles (the left peak) appear to be T=3-like particles since they sedimented to approximately the same level in the gradient as do native T=3 virions (Fig 23). The left peak in the gradient with the peptide 11 loop 129 construct is shifted slightly to the right (down the gradient). Other gradients run with this construct and TEM images of capsids isolated from this peak (see below) suggest that this shift is probably artifactual, and that the shift it is not due to a significant mass or shape difference in these capsids. The right peak in both gradients might correspond to a more massive or differently shaped particle. We speculated that this peak might represent RNA filled capsids. However, A260/A280 readings for the fractions corresponding to these peaks were $\simeq$ 1 (empty of RNA). While the right gradient peaks are certainly composed of CCMV coat protein (since the profiles were generated using a polyclonal antiserum to CCMV coat protein), the reason that this protein sediments separately (further down the gradient) from protein in the left peak is unknown.

Fractions from both peaks of PEG precipitated samples were isolated, dialyzed, concentrated in Centricon™ 100 microconcentrators, and analyzed by TEM to confirm their capsid identity. The fractions from the left peak contained virions that were similar in size (26-28 nm in diameter) and shape (roughly spherical) to native T=3 CCMV virions (Fig. 23). Particles were not detected by EM in fractions from the right peak.

Western blot analysis (not shown) and sequencing (Appendix B) confirmed that constructs were peptide 11 loop 63 and loop 129 constructs.

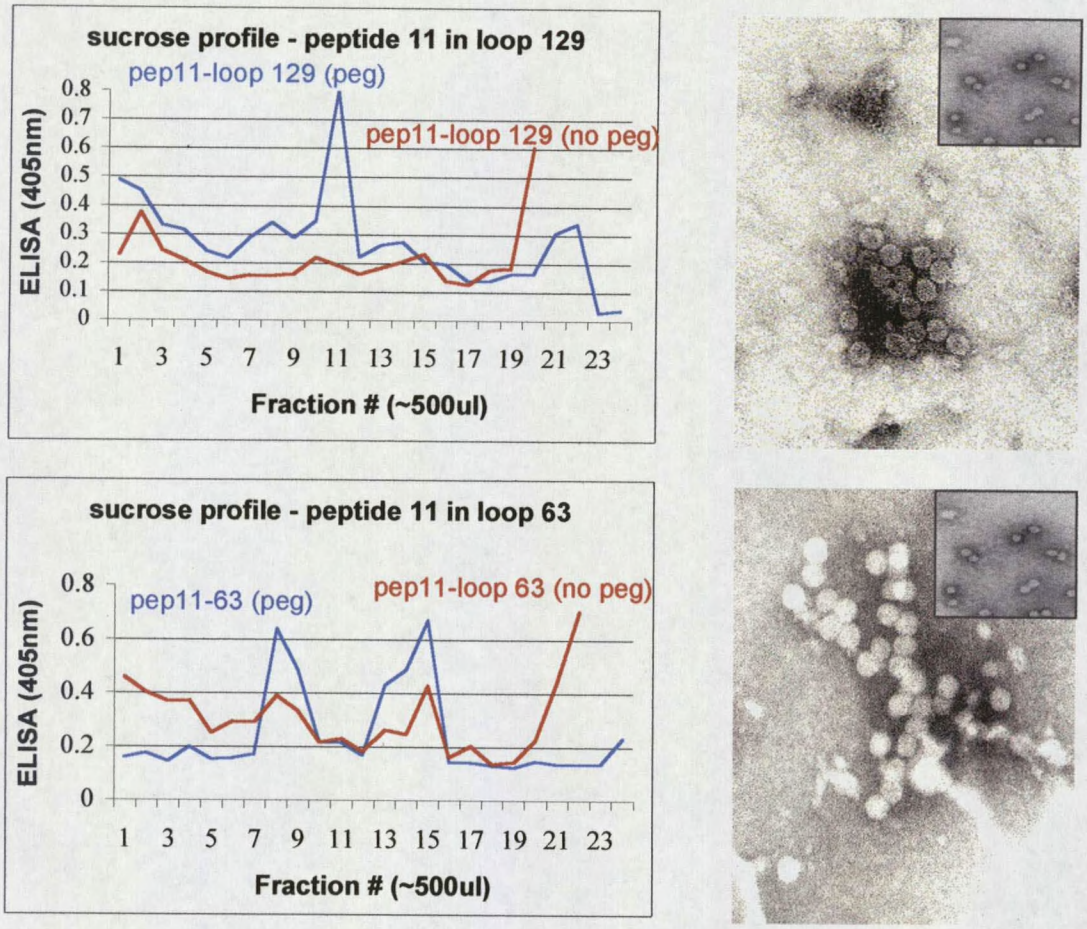


Fig. 23. Isolation of chimeras by PEG precipitation. Virions were isolated from *P. pastoris* fermenter cultures by disrupting with glass beads and precipitating the crude extract twice with PEG (see Chapter 3, Methods). Capsids composed of coat protein with peptide 11 inserted into loop 129 (top panels) or loop 63 (bottom panels were separated on 10-40% sucrose gradients. Gradients (left panels) were divided into successive 500 µl fractions (the left side of the graphs corresponds to the top of the gradient; the right side to bottom), which were analyzed separately by ELISA (as described in Methods). The 2-3 fractions composing each peak were pooled, concentrated in Centricon™ 100 microconcentrators, negatively stained and then analyzed by TEM (right panels: top: peptide 11 (loop 129) chimera 31.5K, 2sec, 5.3X; bottom: peptide 11 (loop 63) chimera 25K, 3.2sec, 4.5x) as described in Chapter 3, Methods. TEM images of native T=3 CCMV virions are shown as comparison (insets).

In an effort to obtain pure samples with high yields of assembled virion particles, it became necessary to avoid using PEG precipitations. Based on other experiments in the Young laboratory, it appeared that a significant amount of coat protein was lost during these precipitations. Instead, crude lysates were filtered and spun at high speed to separate and pellet assembled virions particles. Pelleted virions were resuspended in virus buffer (see Chapter 3, Methods) and then purified by either isopycnic centrifugation in 38% cesium chloride (CsCl) or by separating on 10-40% sucrose sedimentation velocity gradients. Virus bands/peaks were dialyzed against virus buffer, concentrated in Centricon™ 100 microconcentrators, and then stored at 4°C in storage buffer. Using this protocol, we were able to isolate assembled capsids from CCMV-peptide 11 constructs (loops 129 and 63) (Figs. 24 and 25). Typical yields were 10 - 20 mg of CCMV constructs per 100 g of pelleted yeast. Virus purity was estimated by UV spectrophotometry; for pure (and empty) capsid constructs, A260/A280 ≃ 1.0. The loop-63 and loop-129 constructs were found to be empty of RNA (A260/A280 ≃ 1) and were relatively pure.

Peptide 11 modified CCMV capsids were analyzed using 10-40% sucrose gradient centrifugation. (These gradients were also used to purify small amounts of coat protein-peptide 11 chimeras). Fractions were followed by A280. CCMV-peptide 11 loop 129 chimeras sedimented on sucrose gradients as a single peak which corresponded to a T=3-like particle (Fig. 24, top left). The fractions correlating to this peak were concentrated in a Centricon™ 100 microconcentrator, and then analyzed by TEM (Fig. 24, top right). Particles are similar in size and shape to native T=3 CCMV virions.

CCMV-peptide 11 loop 63 chimeras did not give a noticeable peak when purified by sucrose gradient centrifugation. Fractions were collected from where one might expect a peak, based on the corresponding position in the gradients from loop 129 constructs. These fractions were concentrated and analyzed by TEM. No particles were observed. By increasing the Y-axis scale ten-fold (inset in profiles), a shoulder becomes apparent on the right side (the lower third of the gradient) of the profiles for loop 63 and loop 129, but not the wild type profile. This shoulder seems especially predominant in the loop 63 sample. We speculate that this shoulder might consist of CCMV-peptide 11 coat protein aggregates, however this has not yet been confirmed. We are currently investigating the nature of the shoulder protein, and the reason why the loop 63 construct does not have the same sucrose gradient profile that the loop 129 construct does.

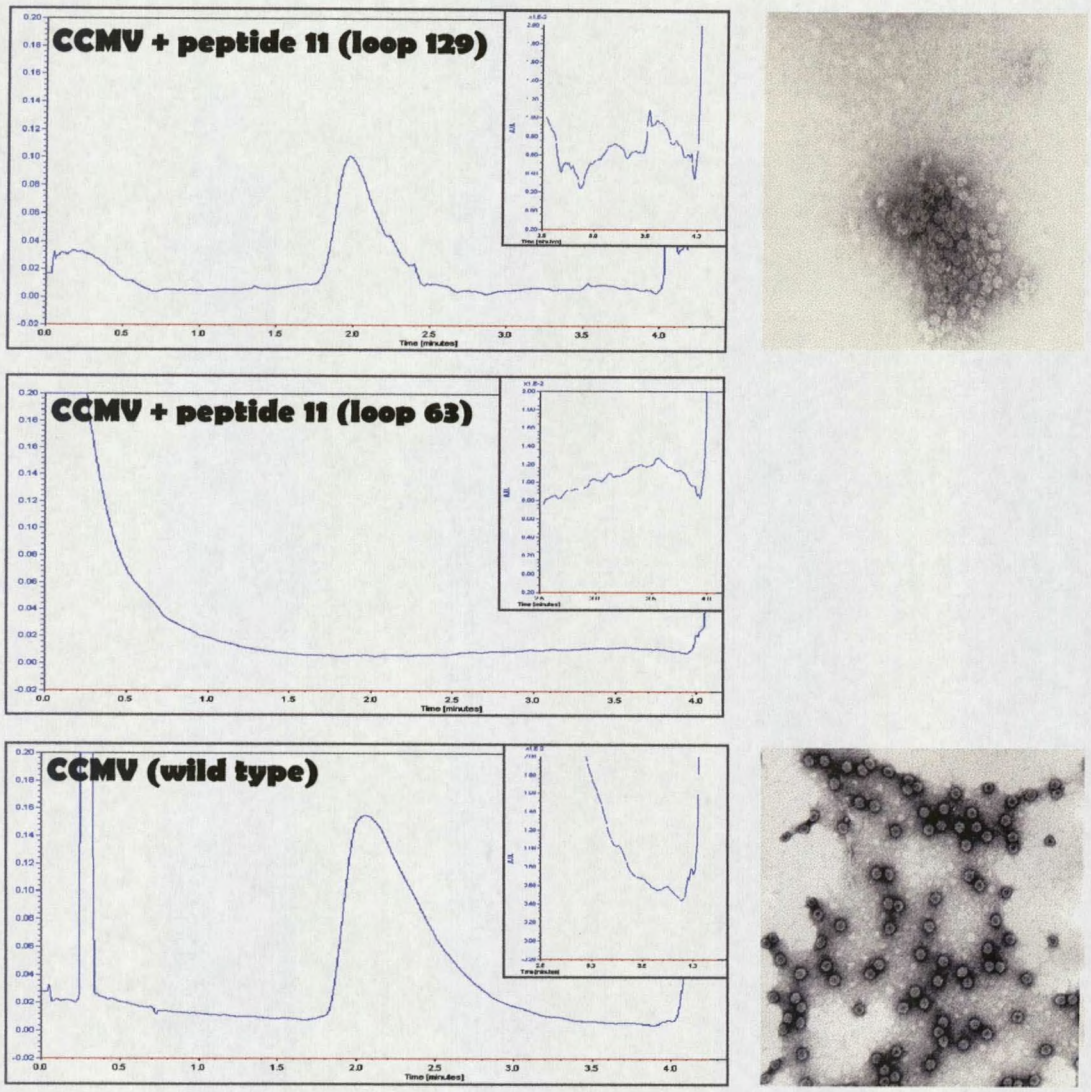


Fig. 24. Isolation of chimeras by ultracentrifugation. Virions were isolated from *P. pastoris* fermenter cultures by disrupting with glass beads, filtering and then ultracentrifuging (25K, 2 hrs) the filtrate (see Chapter 3, Methods). Virion particles (capsids) were resuspended from pellets in virus buffer. Capsids composed of coat protein with peptide 11 inserted into loop 129 (top panels), loop 63 (middle panels), or with no insert (wild type control; bottom panels) were separated on 10-40% sucrose gradients. Gradients (left panels) were fractionated while monitoring absorption at 280 nm (vertical axes) as described in Methods. The 2-3 fractions composing each peak were pooled, concentrated in Centricon™ 100 microconcentrators, negatively stained and then analyzed by TEM (right panels: top: capsids (coat protein with peptide 11 in loop 129), 31.5K, 4sec, 6.6X,); bottom: CCMV native T=3 capsids.

Capsid particles isolated by ultracentrifugation were also purified using CsCl centrifugation. Two bands were observed in each gradient from peptide 11 constructs (loop 63 and loop 129). A single band was observed in a wild type control CsCl gradient at approximately the same position as the bottom of the two bands seen in the gradients from the peptide 11 constructs. All bands were dialyzed and concentrated. Samples were directly quantitated by ELISA and capsids were visualized by TEM (Fig. 25).

By ELISA, the top band from the both peptide 11 loop constructs contained approximately 0.1 μg/μl of coat protein, approximately 10-times more coat protein than was found in the bottom bands. The single band from the wild type control contained more than 1 μg/μl coat protein. All capsids, isolated from each band, looked like native particles (Fig 15). Both native and modified particles were approximately 28-29 nm in diameter, as expected for T=3 capsids. There were far fewer particles in the viewing field from samples of the top bands of chimeric constructs compared to the bottom bands. This would be predicted based on ELISA values. Why there were two CsCl bands from the chimeric constructs is still an unresolved question. It might be that there are two size classes of particles formed from the assembly of peptide 11 chimeric coat proteins. A population of larger particles (with lower relative density), would form the top band in CsCl. To test this theory, a noticeable size difference among virions from different bands might be expected to be seen under TEM. However, we have not observed a consistent size difference in particles from the different bands. It is possible that smaller capsids re-assemble (disassemble and assemble) into more stable full-size (T=3) particles during dialysis and concentration, and therefore cannot be observed. There could also be

particulate matter (other than coat protein) in the sample preparations that co-sediments with coat protein chimeras, possibly giving the second band. This theory is supported by the appearance of non-virus particles in loop 129 construct preparations under TEM (Fig. 25). It is important to note that subsequent CsCl purifications with capsid construct (loops 63 and 129) have given noticeably different results from these reported CsCl runs. Two bands are usually apparent. However, the relative contributions of coat protein in each band vary significantly from run to run. The character of the two bands that are found when chimeric constructs (loop 63 and 129) are purified using 38% CsCl centrifugation, and why these form, are issues currently under investigation.

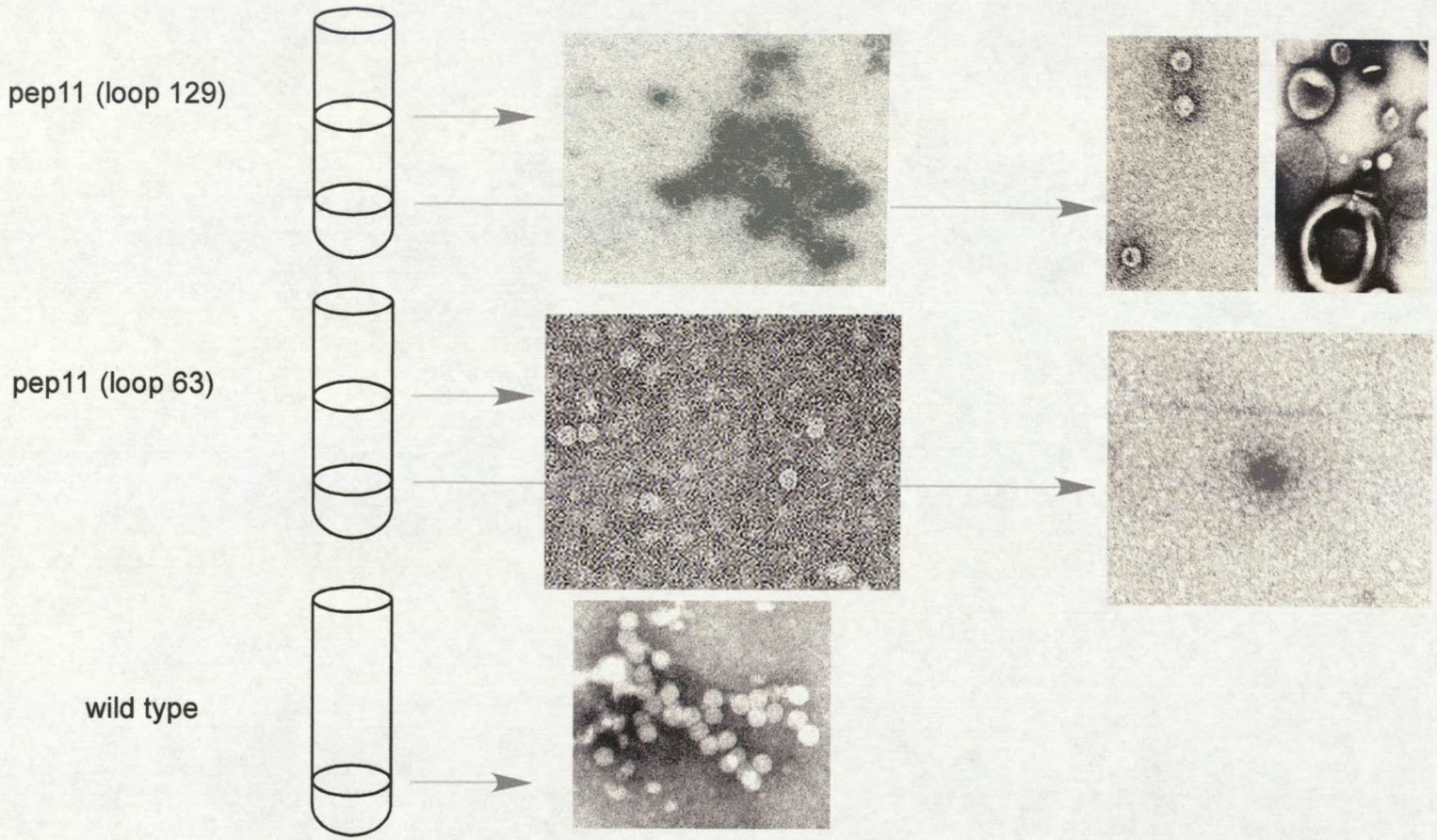


Fig. 25. CsCl purification of chimeric capsids (loops 129 and 63). Virions were pelleted from crude yeast lysates after filtering. Pellets were resuspended in virus buffer (see Chapter 3, Methods). Capsids were purified using 38% CsCl isopycnic gradients. Two gradient bands formed from samples with chimeric constructs (coat protein with peptide 11 insert in loop 129 (top) and loop 63 (middle)). Bands were concentrated, negatively stained and then analyzed by TEM (top left: pep11 (loop 129) top band-31.5K, 7.7X, 2.5sec; top middle: pep11 (loop 129) bottom band-31.5K, 7.7X, 3sec; top right: nonviral particles from bottom band pep11 (loop 129) sample-31.5K, 7.7X, 4sec; middle left: pep11 (loop 63) top band-25K,5.3X,5sec; middle right: pep11 (loop 63) bottom band-31.5K, 7.7X, 2.5sec; bottom: wild type (native T=3) virions-25K, 44X, 2.2sec).

CHAPTER 5

CONCLUSIONS

The overall goal of this thesis project was to begin development of a tumor-targeted bio-imaging system based on a plant virus scaffold – peptide 11 chimera. Viruses have previously been shown to be useful presentation systems for the expression of foreign peptide sequences (see Chapter 2, Why CCMV?). In particular, different plant viruses have been used to present antigenic epitopes for use in vaccine development. These systems are especially effective in their ability to stabilize short peptides in a constrained conformation often similar to the conformation of the "native" epitope that the insertions are meant to mimic. Also, the peptide is presented in multiple copies on the surface of the virion chimera, resulting in a higher local concentration of the displayed peptide relative to that found in a solution of isolated peptides. This is due primarily to the small size and the innate symmetry of the virion particle[52].

Taking advantage of the attributes that make plant viruses effective epitope presentation scaffolds, we endeavor to create a tumor-targeted plant virus-based chimera. To our knowledge, plant viruses have not yet been utilized in this fashion; to target a human cellular receptor. The system most similar to the one that we envision is probably that of tumor-targeted adenovirus-based gene delivery vectors (see Chapter 2, Why CCMV?). Possibly the most important lesson to learn from previous research that was done with these chimeras is the approach that researchers have recently taken in

designing them. Specifically, they have taken a "rational design approach". Krasnykh *et al*[36] first analyzed the high resolution structure of an adenoviral coat protein (the Ad5 fiber protein) in order to identify appropriate positions to incorporate a cell-targeting ligand. This group then went through several iterations of design, construction, and target validation in an effort to create a vector with a novel and intended tropism. Since these "first generation" chimeras, several groups have adapted this system to carry ligands that target more biologically important receptors[33]. Of note, a tumor-targeting RGD sequence has been successfully incorporated into the HI loop of the fiber knob. The adenovirus assembles with this insert and it targets to α_v-type integrin receptors on the surface of cells from ovarian cancer patients[38]. We are taking essentially the same "rational design approach" to construct a tumor-targeted plant virus (CCMV)-based chimera. In this thesis, I have detailed our initial efforts to build and characterize such a construct.

Summary of Results

In the first stage of this project, the 3D structure of CCMV was analyzed. There are five surface exposed loops which, according to the crystal structure, do not contribute significantly to interactions important for capsid assembly and stability. These loops appear to be good choices for exogenous peptide incorporation. Mapping of the CCMV capsid surface (by residue-wise B-factor and surface accessibility-factor mapping) revealed specific insertion sites (residues 63, 102, 114, 129, and 160) with relatively high

flexibility and surface accessibility. We speculate that the peptide 11-based sequence, GICDPGYIGSRCGIL, could be stably incorporated into these surface loop positions. We hypothesize that the core β-barrel fold of the CCMV coat protein scaffold would be maintained while peptide 11 would be left to adopt a conformation within a loop-like surface exposed framework where it can interact (bind) with the peptide 11 receptor, LBP.

Based on insertion positions identified in the "design stage", BamH1 cloning sites were added to loop regions in the coat protein gene (previously cloned into a pUC-based plasmid). Peptide 11-based oligonucleotide duplexes were inserted into these BamH1 cloning sites, creating pCC3/pep11 plasmid constructs (Fig. 11). The loop-63, 102, 114, and 129 constructs were sequenced to confirm correct insertion of peptide-11. *In vitro* transcription/translation of chimeric coat proteins from the plasmid constructs further demonstrated that inserts were in frame and that no deleterious mutations had arisen during cloning (Fig. 12). We had trouble sequencing the loop-160 construct. This may have been because the loop-160 position is downstream in the coat protein gene from the other loop positions (and therefore closer to the 3' sequencing primer). This downstream position may have negatively affected efficient PCR amplification prior to cycle sequencing. Also, the 160-loop clones that were constructed did not translate well in the *in vitro* transcription/translation assay. It is difficult to postulate why these clones did not translate since sequencing data was not available to identify potential problems with cloned plasmid DNA. The loop-160 constructs are currently being re-made and new primers (made with greater attention to the 160 position) have been ordered to sequence

these constructs. "Positive" clones will be sub-cloned directly into the pPicZA expression vector for heterologous expression in *P. pastoris*. Loop 160 constructs will be evaluated as other constructs have been.

In the next stage of this project, CCMV coat protein-peptide 11 constructs (loops 63, 102, 114, and 129) were evaluated for their ability to be expressed in plants. Capped RNA transcripts were made from pCC3/pep11 plasmids. These were inoculated onto cowpea plants, along with capped transcripts from pCC1 and pCC2 (these provide the other RNA elements essential for CCMV replication in plants). Based on ELISA and RT-PCR analysis (Figs. 14 and 16, respectively), only the loop 129 construct was able to establish any detectable infection in plants. Infections were characteristically localized and did not spread throughout the plant, nor could isolated chimeras be efficiently passaged from plant-to-plant. Although the peptide 11 loop 129 construct seemed genetically stable, this construct can be deemed not viable in plants since it was not able to "move" from cell-to-cell or systemically within the plant. Also, this construct could not generate infectious chimeras; the local infection could not be passaged to other plants.

The loop 129 construct's lack of viability in plants implies structural instability in the chimera. Movement of wild type CCMV within the plant appears to be dependant on stable packaging (and release) of the RNA genome[76]. It can be proposed that the loop 129 chimera significantly disrupted normal RNA packaging/release by either changing the structure of the RNA (and requisite RNA-coat protein contacts) or by disrupting the coat protein structure and its ability to assemble/disassemble as is required for viral movement in plants. In a review of plant virus based scaffolds[52], Johnson J. *et al* suggest

that chimeras carrying large inserts are genetically unstable and that inserts that are reduced in size are more viable in host plant expression systems. There therefore appears to be certain constraints on the packaging of viral genetic material that are imposed on modified plant viruses when they are in their "native" (host) environment. These constraints are probably not due to the internal volume of the capsid since there appears to be more than enough internal space to accommodate RNAs of larger size. Whether apparent constraints are imposed through protein structure limitations, genetic limitations (genome size and RNA structure), or both, is an issue that deserves further study. Visualization of a predominant dimer band by Western blot analysis (see Fig. 15) may lend support to the idea that the coat protein structure (and *in vivo* viral function) was significantly altered due to peptide 11 incorporation. Coat protein dimers are normally present in wild type coat protein preparations from plants. Most of these dimers are thought to be due to non-covalent association of monomers, although it has been proposed that one covalent (disulfide linked) dimer is formed per 180 coat protein subunits in an assembled capsid. It is possible that dimers form from coat proteins with the peptide 11 insert in loop 129 as a result of non-covalent interactions between subunits. Alternatively, covalent bonds between monomers could form via disulfide links. These covalent dimers might be visible by SDS-PAGE (Fig. 15) if disulfide reduction (in SDS-sample buffer with 2.5% β-mercaptomethanol) were inefficient. The addition of the two extra cysteine residues flanking the peptide 11 insertion sequence might potentiate disulfide bond formation. While a predominant dimer band was not observed by Western blot with protein extracted from yeast (see Fig. 21), chimeric coat

protein did seem to aggregate during purifications from yeast fermenter cultures (see Fig. 23 and Chapter 4 Results and Discussion). It is possible that the dimer band observed in Western blots of protein isolated from plants and the aggregation of coat protein from yeast are both due to similar intersubunit interactions. Further investigation is needed to resolve this issue (see suggestions for future work below).

The direct synthesis of protein chimeras in heterologous systems (like *E. coli* or yeast) does not seem to be limited by the same constraints (i.e. virus infection, replication, movement, and propagation) found in homologous (plant) expression systems. The *P. pastoris* heterologous expression system may be particularly well suited to the expression of chimeric CCMV coat proteins. *P. pastoris* is well-characterized biochemically, has been shown to produce many different foreign proteins at high levels, and is capable of performing eukaryotic posttranslational modifications such as glycosylation and disulfide bond formation. For these reasons, we chose to try to express the above constructs (CCMV coat protein with peptide 11 in loops 63, 102, 114, 160) in *P. pastoris*.

To accomplish this, pPicZA/pep11 *P. pastoris* expression vectors were constructed (see Fig. 17). Successful cloning was demonstrated by diagnostic (SmaI) restriction digestion of purified clones and then by sequencing "positive" clones (Fig. 18). Vectors were successfully used to stably transform X-33 mut$^+$ yeast with the chimeric coat protein genes (Fig. 18). Small scale time course expression studies were used to demonstrate coat protein expression in *P. pastoris* isolates and to optimize

culturing strategies (Figs. 20 and 21). Single isolates were chosen for intermediate-scale (500 ml) expression.

Only the loop 63 construct was able to express detectable (by ELISA) amounts of chimeric coat protein from intermediate-scale cultures, albeit still at very low levels (Fig. 22). Capsids purified from this culture looked like native T=3 CCMV capsids. They appeared to be empty of RNA (A260/A280 ~ 1). The reason(s) why constructs generally did not express well in 500 ml cultures is not clear. Possible explanations include inadequate oxygen supply in the shaker-flasks and/or poor methanol induction.

Fermentation can be a good alternative to shaker-flask culturing. More than providing larger (4 L) culture volumes, fermentation offers greater control of the culture environment, such as oxygen supply, pH, and nutrient load, making methanol induction more effective. In an attempt to obtain greater amounts of chimeric coat proteins, we expressed two constructs, loop 63 and loop 129, by fermentation. Constructs were isolated from fermenter batches in two ways. First, we tried a two-step PEG precipitation (see Chapter 3, Methods). This purification scheme was somewhat effective in isolating chimeric capsids. Sucrose gradient profiles showed two peaks from both constructs (Fig. 23). The left peak (higher in the gradient) contained what looked like native T=3 particles. The right peak was composed of coat protein, but capsids from this peak were not readily observed by TEM. We suspect that expression levels from fermenter batches were adequate for obtaining relatively large amounts of assembled virion particles. Low yields may have been a result of significant losses of coat protein during the PEG purification. Since 15% PEG specifically precipitates assembled capsids, the loss of coat

protein during PEG precipitations may be a reflection of general capsid instabilities. Further studies (see below) are needed to clarify this issue.

To isolate greater numbers of assembled virion constructs, we used a high-speed centrifugation-based purification scheme (see Chapter 3, Methods). Samples processed in this way were purified by sucrose gradient (Fig. 24) and CsCl gradient centrifugation (Fig. 25). A single peak was visualized in sucrose gradients from the loop 129 construct. This peak contained native T=3-like capsids. There were no apparent peaks (and no particles) observed in sucrose gradients from the loop 63 construct. However, a predominant "shoulder" was visible in the sucrose gradients from this sample. This shoulder may be aggregated coat protein. Its exact composition is currently being investigated.

CsCl purification of sample preparations of both loop constructs each gave two distinct bands (Fig. 25). Both top and bottom bands contained T=3-like capsids. The bottom bands are localized at approximately the same level as the single band obtained from wild type sample preparations. This lower band usually contained less coat protein and had fewer assembled capsids. However, subsequent CsCl purifications of the same constructs have given the opposite results: T=3-like particles in both bands, but fewer in the top band than the bottom band. Non-viral particles were observed in the bottom band from the loop 129 construct sample. It is possible that this material affects the separation of capsids. The nature of the non-viral material is currently under investigation.

To recapitulate: In this report, we have described the generation and characterization of recombinant CCMV virions with a peptide 11-derived sequence

genetically incorporated within two different surface loops. An effort to create such constructs was undertaken in order to demonstrate the utility of surface loops of CCMV as potential sites for incorporation of short peptide ligands, which (eventually) would allow the virion to bind to ligand-specific cellular receptors.

Further Studies

The preliminary results presented in this thesis already indicate that laminin peptide 11 can be expressed on the surface of the CCMV protein scaffold. Future experiments are aimed at extending these initial results to 1) determine which of the five selected surface exposed loops is most suitable for high-level stable expression of CCMV-peptide 11 capsid chimeras, and 2) determine which of the five peptide 11 incorporation positions is most effective for *in vivo* targeting the cellular receptor, LBP.

First, peptide 11 expression in *P. pastoris* (by fermentation) will be assessed in the three remaining surface loops (positions 102, 114, 160) as was done for loops 63 and 129. Improvements in the *P. pastoris* expression system should increase yield of assembled capsids. To do this, further characterization of yeast isolates is currently under way. The *P. pastoris* expression system has the possibility of isolating high-yield clones with multiple copies of the inserted gene. It is possible that large-scale screening under varying (increasing) concentrations of Zeocin will lead to the isolation of *P. pastoris* clones with a higher coat protein-peptide 11 expression level than those reported here. Improvements in the fermentation protocol, generally aimed at increasing oxygen and

methanol utilization (and presumably increased protein induction) are also under way. Also, it may be helpful to incorporate a reporter construct as a transcriptional fusion product in order to follow coat protein expression patterns in *P. pastoris* cultures. Recently, the beta-D-glucuronidase enzyme demonstrated feasibility as a functional reporter in *P. pastoris*[77].

The stability of the chimeras that are produced from *P. pastoris* fermentation may be improved in several ways. An extra cysteine was engineered into the C-termini of peptide 11 inserts (CDPGYIGSRC). This cysteine may form a covalent disulfide bond with its close (in sequence) neighbor, the cysteine on the N-termini of the insert. We will asses disulfide bond formation in purified constructs using a thiol conjugation assay already being employed in the Young laboratory. It is possible that this covalent linkage may stabilize particles and could stabilize peptide 11 insert within coat protein loops in a conformation that is more readily recognized by its LBP receptor.

Hypothetically, virions could also be stabilized by chemical fixation with glutaraldehyde or paraformaldehyde. This type of chemical stabilization is being explored with other capsid constructs in the Young laboratory. Stabilizing mutations like the K42R substitution (see Chapter 2, Why CCMV?) could also be introduced to stabilize capsid constructs. These mutations are currently being introduced into pPicZA/pep11 plasmids and will be evaluated for expression in *P. pastoris*. Stabilized CCMV capsid constructs may be more readily purified since the current purification protocol generally purifies assembled capsids. Also, a stable capsid is likely to be more effective in presenting peptide 11 from its surface, and may be less susceptible to degradation by

cellular peptidases than unstable capsids. Virions stably expressing peptide 11 will be crystallized for structural determination using X-ray crystallography. A high-resolution crystal structure would be helpful in determining structure-function relationships with chimeric constructs.

Coat-protein-peptide 11 chimeras will be analyzed for cell-targeting activity using both *in vitro* and in animal assays. To find out whether the peptide 11 ligand introduced into the four loops of CCMV (63, 102, 114, 129) is available for binding their LBP target *in vitro*, two independent assays will be used. First, an ELISA-based assay will be used to demonstrate specificity and establish a dose response curve. Dilutions of recombinant LBP (purified from *E. coli*) ranging in concentration from 5 to 100 μg/ml (100 μl volume) will be bound to high (protein)-binding microtiter plates. After blocking, CCMV constructs, purified by 38% CsCl centrifugation, will be applied. Specific binding will be detected using CCMV polyclonal antiserum (see Methods). Wild type CCMV will be applied to LBP-coated plates as a negative control. Anti-LBP polyclonal antiserum will be used as positive control to ensure that coated LBP is available for binding. This anti-LBP antibody may be used to inhibit binding of CCMV-peptide 11 constructs to demonstrate specificity. Synthetic peptide 11 photoprobe could also be used in this capacity.

For another independent *in vitro* test of the accessibility of peptide 11 within exterior loops, we will use an assay based on the specific interaction of the peptide 11 constructs with an affinity matrix containing an anti-peptide 11 monospecific antibody. The specificity of this antibody will first be tested by Western blot. Coat protein-peptide

11 constructs purified on CsCl gradients will be Western blotted as before and protein membranes will be probed with the peptide 11 antibody, using wild type CCMV as negative control and purified laminin-1 as positive control. If this antibody shows specificity, CCMV-peptide 11 constructs purified on CsCl gradients will be immunoprecipitated with M2-affinity gel (with anti-peptide 11 M2 antibodies). Protein capsids that bind to the matrix will be eluted with synthetic peptide 11 and then analyzed by TEM and by SDS-PAGE/Western blot (probing independently with anti-peptide 11 antiserum and anti-CCMV anti-serum) as previously described.

Synthetic peptide 11 and various peptide 11 analogues have demonstrated ability to block the invasion of LBP expressing tumor cells through EHS basement membrane matrix (see introduction). Based on this concept, CCMV-peptide 11 capsids will be tested for their ability to block tumor cell invasion. To measure invasion, a "Transwell" two-chamber assay system in which the 8μm pore barrier is impregnated with EHS matrigel basement membrane matrix will be used as described[21]. 5 X 10^4 tumor cells are seeded into the upper well and the chambers are incubated for 3 days or one week depending on the test line. Upper transwells are removed and the number of cells which have invaded into the lower well is quantitated. Positive control assays with synthetic peptide 11 and negative controls (CCMV protein capsids without peptide 11 insert) will be compared with CCMV-peptide 11 capsid constructs.

A direct competition assay for binding to cells up-regulated in LBP on their cell surface will be used to further asses the ability of CCMV-peptide 11 chimeras to target tumor cells. Briefly outlined, an engineered DG44CHO variant cell line with up-

regulated LBP surface expression will be used[21]. The peptide 11-based biotinylated photoprobe (mentioned in the Chapter 2, Why Peptide 11?) can be used to directly image the specific binding of peptide 11 to the surface of these cells by confocal microscopy using FITC-conjugated avidin. Also, FACScan can be used to quantitatively follow binding of peptide 11 to DG44CHO cells. Increasing concentrations of the CCMV-peptide 11 capsid chimeras will be added to a fixed amount of free peptide 11 based photoprobe and qualitatively assayed for inhibition of the photoprobe using confocal microscopy. FACScan analysis will be used to quantitate results. A corresponding analysis will also be carried out using a constant concentration of CCMV-peptide 11 chimeras while varying the concentration of free synthetic peptide 11 (not the photoprobe analogue). The negative control will be CCMV capsids without peptide 11 inserts.

The assays outlined here are aimed at determining specificity and does-dependence of the binding of chimeric constructs to the tumor cell surface receptor, LBP. It is quite possible that the peptide 11 insert is not the best choice for an incorporated ligand. In the event that peptide 11-based chimeras are unsuccessful, several other ligands should be considered. Possibilities include (YIGSR)n-based insertions; the integrin receptor binding motif, RGD; HA or FLAG tags (and subsequent testing of binding to surface expressed ScFv HA or FLAG pseudoreceptors) (see Chapter 2, Background and Project Rational for further discussion). The exterior loops of the CCMV capsid protein may be beneficial for tumor targeting peptide ligands obtained from phage display libraries containing random peptide sequences flanked with two cysteine residues forming a disulfide bridge[78],[79]. An alternative approach to targeting

tumor cells could be the use of bi-specific monoclonal antibodies. In theory, one end of the bi-specific molecule would be directed to bind the CCMV capsid directly (or possibly to bind to a high-affinity epitope displayed within coat protein loops), while the opposite end of the bi-specific molecule is directed to bind tumor cells. This approach has successfully been used to re-target adenoviral vectors to α_v integrins[44] and has been used in immuno-based tumor therapies[80] (targeting effector cells to tumor cells while triggering effector cell responses).

Once we can demonstrate stable virion assembly and specific cell targeting, constructs will be evaluated for their ability to image tumor cells. Appropriate collaborations have already been arranged to evaluate the effectiveness of our constructs for bio-imaging applications. A discussion of these studies is beyond the scope of this thesis.

Conclusions

The results presented in this thesis indicate that laminin peptide 11 can be expressed on the surface of the CCMV protein scaffold. Future experiments are aimed at determining which of the five selected surface exposed loops is most suitable for high-level stable expression of CCMV-peptide 11 capsid chimeras, and which of the five peptide 11 incorporation positions is most effective for *in vivo* targeting the cellular receptor, LBP. The long-term goal of this work is to create a tumor-targeted virus-based bio-imaging agent for use in the early detection of metastases. We therefore hope to

eventually extend the utility of these "first generation" constructs to cancer bio-imaging applications. An agent of this sort may be helpful in estimating the metastatic potential of tumors, thereby aiding in staging and prognostic-related decisions on cancer treatment.

REFERENCES CITED

[1] Malinoff, HL and Wicha MS. 1983. Isolation of a cell surface receptor protein for laminin from murine fibrosarcoma cells. J. Cell. Biol. 96:1475-1479.

[2] Rao et al. 1983. Isolation of a tumor cell laminin receptor. Biochem. Biophys. Res. Commun. 111:804-808.

[3] Sobel ME. 1993. Differential expression of the 67-kDa laminin receptor in cancer. Cancer Biol. 4:311-317.

[4] Hand PH et al. 1985. Expression of laminin receptor in normal and carcinomatous human tissues as defined by a monoclonal antibody. Cancer Research. 45:2713-2719.

[5] Cioce V et al. 1991. Increased expression of the laminin receptor in human colon cancer. J. Natl. Cancer Inst. 83:29-36.

[6] Castronovo V et al. 1990. Immunodetection of the metastasis-associated laminin receptor in human breast cancer cells obtained by fine-needle aspiration biopsy Am. J. Pathol. 6:1373-1381.

[7] Castronovo. 1993. Laminin receptors and laminin-binding proteins during tumor invasion and metastasis. Invasion Metastasis. 13(1):1-30. Review.

[8] Martignone S et al. Characterization of two monoclonal antibodies directed against the 67 kDa high affinity laminin receptor and application for the study of breast carcinoma progression. 1992. Clin. Exp. Metast. 10:379-386.

[9] Martignone S et al. 1993. Prognostic significance of the 67-kilodalton laminin receptor expression in human breast carcinomas. J. Natl. Cancer Inst. 85:398-402.

[10] Brule van den FA et al. 1994. Inverse correlation of two laminin binding proteins, 67LR and galectin-3, correlates with the invasive phenotype of trophoblastic tissue. Biochem. Biophys. Res. Commun. 201:388-393.

[11] Bryant G. et al. 1987. A role for the laminin receptor in leukocyte chemotaxis. Journal of Leukocyte Biology. 41:220-227.

[12] Crawford BD and Burke RD. 1994. YIGSR domain of laminin binds surface receptors of mesenchyme and stimulates migration during gastrulation in sea urchins. Development. 120:3227-3234.

[13] Rao et al. 1989. Evidence for a precursor of the high-affinity metastasis-associated murine laminin receptor. Biochemistry 28:7476-7486.

[14] Castronovo V et al. 1991. Biosynthesis of the 67 kDa high affinity laminin receptor. Biochem. Biophys. Res. Commun. 177:177-183.

[15] Landowski TH, Dratz EA, Starkey JR. 1995. Studies of the structure of the metastasis-associated 67kDa laminin binding protein: fatty acid acylation and evidence supporting dimerization of the 32 kDa gene product to form the mature protein. Biochemistry. 34:11276-11287.

[16] Rosenthal ET and Wordeman L. 1995. A protein similar to the 67 kDa laminin binding protein and p40 is probably a component of the translational machinery in Urechis caupo oocytes and embryos. J Cell Sci. 108 (Pt 1):245-56.

[17] Auth D and Brawerman G. 1992. A 33-kDa polypeptide with homology to the laminin receptor: component of translational machinery. PNAS 89:4368-4372.

[18] Ford CL, Randal-Whitis L and Ellis SR. 1999. Yeast proteins related to the p40/laminin receptor precursor are required for 20S ribosomal RNA processing and the maturation of 40S ribosomal subunits. Cancer Research. 59:704-710.

[19] Ardini E et al. 1998. The 67-kDa laminin receptor originated from a ribosomal protein that acquired a dual function during evolution. Mol. Biol. Evol. 15(8):1017-1025.

[20] Kazmin DA et al. 2000. Phage display mapping for peptide 11:sensitive sequences binding to laminin-1. J. Mol. Biol.00(unpublished)1-15.

[21] Starkey JR et al. 1999. Cell surface and substrate distribution of the 67-kDa Laminin-Binding Protein Determined by using a ligand photoaffinity probe. Cytometry.35:37-47.

[22] Terry H. Landowski et al. 1995. Control pathways of the 67kDa laminin binding protein: surface expression and activity of a new ligand binding domain. Clin. Exp. Metastasis. 13:357-372.

[23] Magnifico A et al. 1996. Peptide G, containing the binding site of the 67-kDa laminin receptor, increases and stabilizes laminin binding to cancer cells. J. Biol. Chem. 271:31179-31184.

[24] Unpublished data from Starkey laboratory.

[25] Iwamoto Y et al. 1987. YIGSR, a synthetic laminin pentapeptide, inhibits experimental metastases formation. Science. 238:1132-1134.

[26] Barsky S et al. 1984. Laminin molecular domains which alter metastasis in a murine model. The Journal of Clinical Investigation. 74:843-848.

[27] Nelson J et al. 1996. Murine epidermal growth factor peptide (33-42) binds to a YIGSR-specific laminin receptor on both tumor cells. J Biol Chem. 271(42):26179-26178.

[28] Graf J et al. 1987. Identification of an amino acid sequence in laminin mediating cell attachment, chemotaxis, and receptor binding. Cell. 48:989-996.

[29] Starkey JR, Dai S, Dratz EA. 1998. Sidechain and backbone requirements for anti-invasive activity of laminin peptide 11. Biochemica et Biophisica Acta. 1-21.

[30] Sakamoto N et al. 1991. Inhibition of angiogenesis and tumor growth by a synthetic laminin peptide, CDPGYIGSR-NH2. Cancer Research. 51:903-906.

[31] Ostheimer GJ. 1992. NMR constrained solution structures for laminin peptide 11. Analogs define structural requirements for inhibition of tumor cell invasion of basement membrane matrix.J Biol Chem.267(35):25120-25128.

[32] Matsuoka T et al. 1998. Adhesion polypeptides are useful for the prevention of peritoneal dissemination of gastric cancer. Clin. Exp. Metastasis. 16:381-388.

[33] Wickham TJ. 2000. Targeting adenovirus. Gene Ther. 7(2):110-114. (review)

[34] Wickham et al. 1997. Increased in vitro and in vivo gene transfer by adenovirus vectors containing chimeric fiber proteins. Journal of Virology. 71(11):8221-8229.

[35] Xia et al. 1995. Structure of the receptor binding domain of adenovirus type 5 fiber protein. Curr. Top. Microbiol. Immunol. 199:39-46.

[36] Krasnykh V et al. 1998. Characterization of an adenovirus vector containing a heterologous peptide epitope in the HI loop of the fiber knob. Journal of Virology. 72(3).1844-1852.

[37] Einfeld DA et al. 1999. Constructuion of a pseudoreceptor that mediates transduction by adenovruses expressing a ligand in fiber or penton base. Journal of Virology. 73(11):9130-9136.

[38] Dmitriev I et al. 1998. An adenovirus vector with genetically modified fibers demonstrates expanded tropism via utilization of a coxsackievirus and adenovirus receptor-independent cell entry mechanism. Journal of Virology. 72(12):9706-9713.

[39] Rossman MG and Johnson JE. 1989. Icosahedral RNA virus structure. Annu. Rev. Biochem. 58:533-573.

[40] Johnson JE. 1996. Functional implications of protein-protein interactions in icosahedral viruses. Proc. Natl. Acad. Sci. 93:27-33.

[41] Johnson JE and Speir JA. 1998. Quasi-equivalent viruses: a paradigm for protein assemblies. J Mol Bio. 269:665-675.

[42] Lane LC. 1981. The Bromoviruses. In E. Kurstak (ed.), Handbook of plant virus infection and comparative diagnosis. Elsevier/North-Holland, Amsterdam.

[43] Fields Virology, Third Edition. Edited by B.N. Fields, D.M. Knipe, P.M. Howley, et al. Lippincott – Raven Publishers, Philadelphia © 1996.

[44] Zhao X et al. 1995. In vitro assembly of cowpea chlorotic mottle virus from coat protein expressed in Escherichia coli and in vitro-transcribed viral cDNA.Virology. 207:486-494.

[45] Speir JA et al. 1995. Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by X-ray crystallography and cryo-electron microscopy. Structure.3:63-78.

[46] Zhao X. 1998. PhD. Purdue University.

[47] B. Steipe. 1999. Evolutionary approaches to protein engineering. Curr Top Microbiol Immuno.243:55-86.

[48] Vita C et al. 1995. Scorpion toxins as natural scaffolds for protein engineering. PNAS. 92:6404-6408.

[49] Vita C et al. 1998. Novel miniproteins engineered by the transfer of active sites to small natural scaffolds. Bioploymers (Peptide Science). 47:93-100.

[50] Sugiyama Y.1995.Systemic production of foreign peptides on the particle surface of tobacco mosaic virus. FEBS lett. 359(2-3):247-50.

[51] Turpen TH et al. 1995. Malarial epitopes expressed on the surface of recombinant tobacco mosaic virus. Bio/Technology. 13:53-57.

[52] Johnson J et al. 1997. Presentation of heterologous peptides on plant viruses: Genetics, structure, and function. Annu. Rev. Phytopathol. 35:67-86.

[53] Joelson T et al.1997.Presentation of a foreign peptide on the surface of tomato bushy stunt virus.78(Pt 6):1213-1217.

[54] Fernandez-Fernandez MR et al. 1998. Development of an antigen presentation system based on plum pox potyvirus. FEBS Lett. 427(2):229-235.

[55] Tianwei L et al. 1996. Structure-based design of peptide presentation on a viral surface: the crystal structure of a plant/animal virus chimera at 2.8 A resolution. Folding and design. (1):179-187.

[56] Cancer Facts and Figures 1991. American Cancer Society, Atlanta, GA, 1991.

[57] Jensen OM et al. 1990. New cases per annum in the period 1978-1992. Eur. J. Cancer. 26:1167.

[58] Zlotnick A et al. 2000. Mechanism of capsid assembly for an icosahedral plant virus. Unpublished draft 4/16/00.

[59] VIPER site – oligomer generation: http://mmtsb.scripps.edu/viper/oligomer.html

[60] Scripps site with CCMV PDB files: http://www.scripps.edu/~chunxuqu/ccmv.html

[61] Swiss-PDBViewer site: http://www.expasy.ch/spdbv/

[62] Guex N. and Peitsch MC. 1997. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis. 18: 2714-2723.

[63] Eisenberg D, McLachlan AD. 1986. Solvation energy in protein folding and binding. Nature. 319(6050):199-203.

[64] VIPER site: http://mmtsb.scripps.edu/viper/viper.html

[65] Sambrook J, Fritsch, EF and Maniatis, T. 1989. "Molecular Cloning: A Laboratory Manual," 2nd ed. Cold Springs Harbor Laboratory Press, Cold Springs Harbor.

[66] Allison RF, Janda M, and Ahlquist P. 1988. Infectious *in vitro* transcripts from cowpea chlorotic mottle virus cDNA clones and exchange of individual RNA components with brome mosaic virus.

[67] "QuikChange" Site-Directed Mutagenesis protocol. From kit with same name. Copyright©1996 by Stratagene.

[68] Perkin-Elmer – Applied Biosystems, Foster City, CA (Standard double-stranded template DNA protocol).

[69] Sambrook J, Fritsch, EF and Maniatis, T. 1989. "Molecular Cloning: A Laboratory Manual," 2nd ed. Cold Springs Harbor Laboratory Press, Cold Springs Harbor.

[70] Bancroft JB and Flack IH. 1972. The behavior of cowpea chlorotic mottle virus in CsCl. J Gen Virol. 15:247-251.

[71] Harlow E and Lane D. 1998. Antibodies, A Laboratory Manual. Cold Springs Harbor Laboritory Press, Cold Springs Harbor.

[72] Higgins, DR and Cregg JM (eds.). 1998. Methods in Molecular Biology. Vol 103. *Pichia* Protocols. Humana Press, Totowa.

[73] Sreekrishna K et al. Strategies for optimal synthesis and secretion of heterologous proteins in the methylotrophic yeast Pichia pastoris. Gene. 1997 Apr 29;190(1):55-62.

[74] Cereghino JL and Cregg JM. 2000. Heterologous protein expression in the methyltrophic yeast *Pichia Pastoris*. FEMS Microbiology Reviews. 24:45-66.

[75] Janda M and Alqhuist P. 1993. RNA-dependant replication, transcription, and persistence of brome mosaic virus RNA replicons in S. cerevisiae. Cell. 72:961-970. (And unpublished results.)

[76] Correspondence with Mark Young, Montana State University, Bozeman, June 20, 2000.

[77] Ayra-Pardo C et al. 1999. Beta-D-glucuronidase gene from Escherichia coli is a functional reporter in the methylotrophic yeast Pichia pastoris. Lett Appl Microbiol. 29(5):278-83.

[78] Koivunen, E et al. 1994. Peptides in cell adhesion research. Methods Enzymol. 245:346-369.

[79] Koivunen, E et al. 1994. Isolation of a highly specific ligand for the alpha 5 beta 1 integrin from a phage display library. J Cell Biol. 124:373-380.

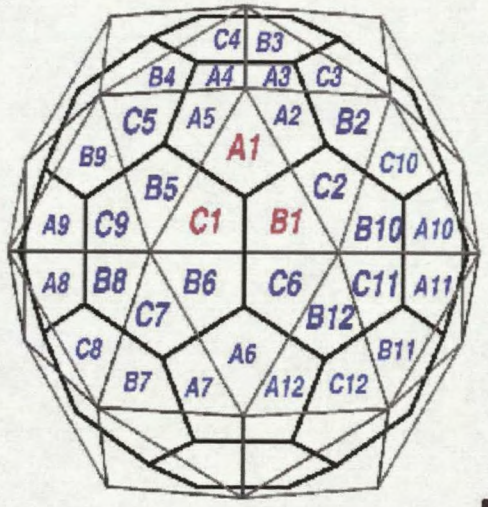
[80] Martin K. Wild et al. 1999. Tumor therapy with bispecific antibody: the targeting and triggering steps can be separated employing a CD2-based strategy. The Journal of Immunology. 163:2064-2072.

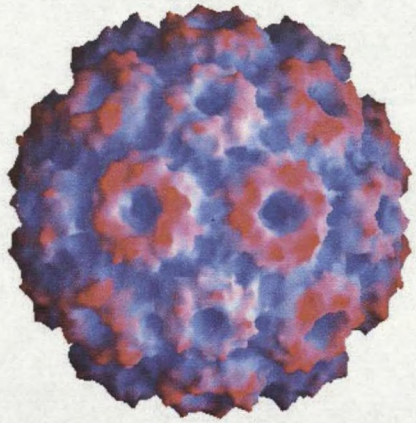
APPENDICES

APPENDIX A

1CWP: RESIDUE-WISE INTERACTION ENERGIES

Subunit Interface	Quasi/Icosahedral symmetry	Association energies (kcal/mol)	Buried Surface Area (A^2)**
A1_B1	QUASI_3-FOLD	-21.0	1229.0
B1_C1	QUASI_3-FOLD	-18.0	1092.0
C1_A1	QUASI_3-FOLD	-16.0	1018.0
B5_A1	QUASI_2-FOLD	-69.0	3250.0
C1_C6	ICOS._2-FOLD	-69.0	3253.0
A1_A5	ICOS._5-FOLD	-23.0	1267.0
B5_C1	QUASI_6-FOLD	-47.0	2394.0
C1_B6	QUASI_6-FOLD	-47.0	2364.0
C1_C7	ICOS._3-FOLD	3.0	147.0
B1_B10	ICOS._3-FOLD	3.0	138.0
C1_B8	QUASI_6-FOLD	3.0	134.0
A1_R1	A1-RNA	-1.0	363.0
B1_R1	B1-RNA	-6.0	547.0
C1_R1	C1-RNA	-2.0	405.0

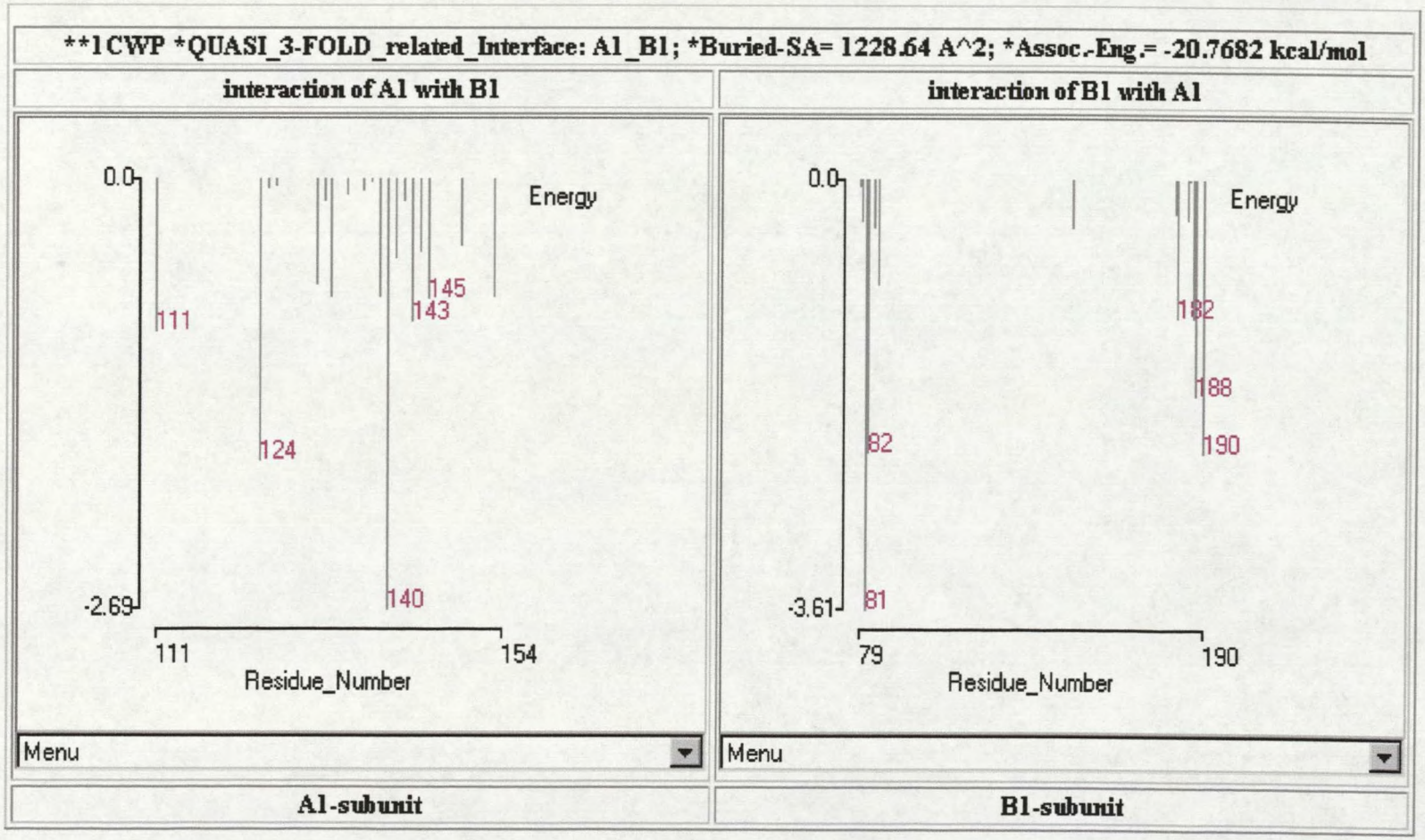


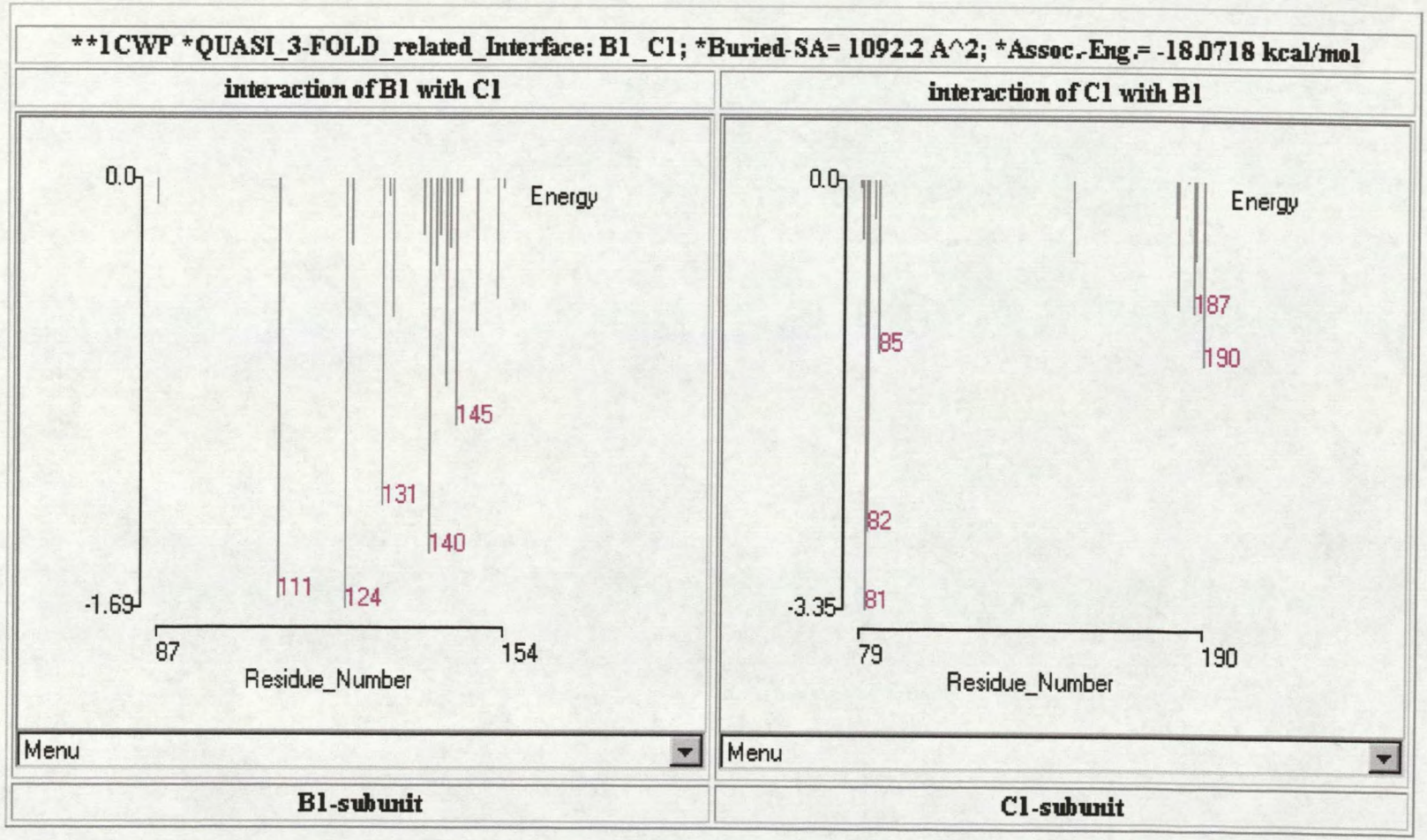


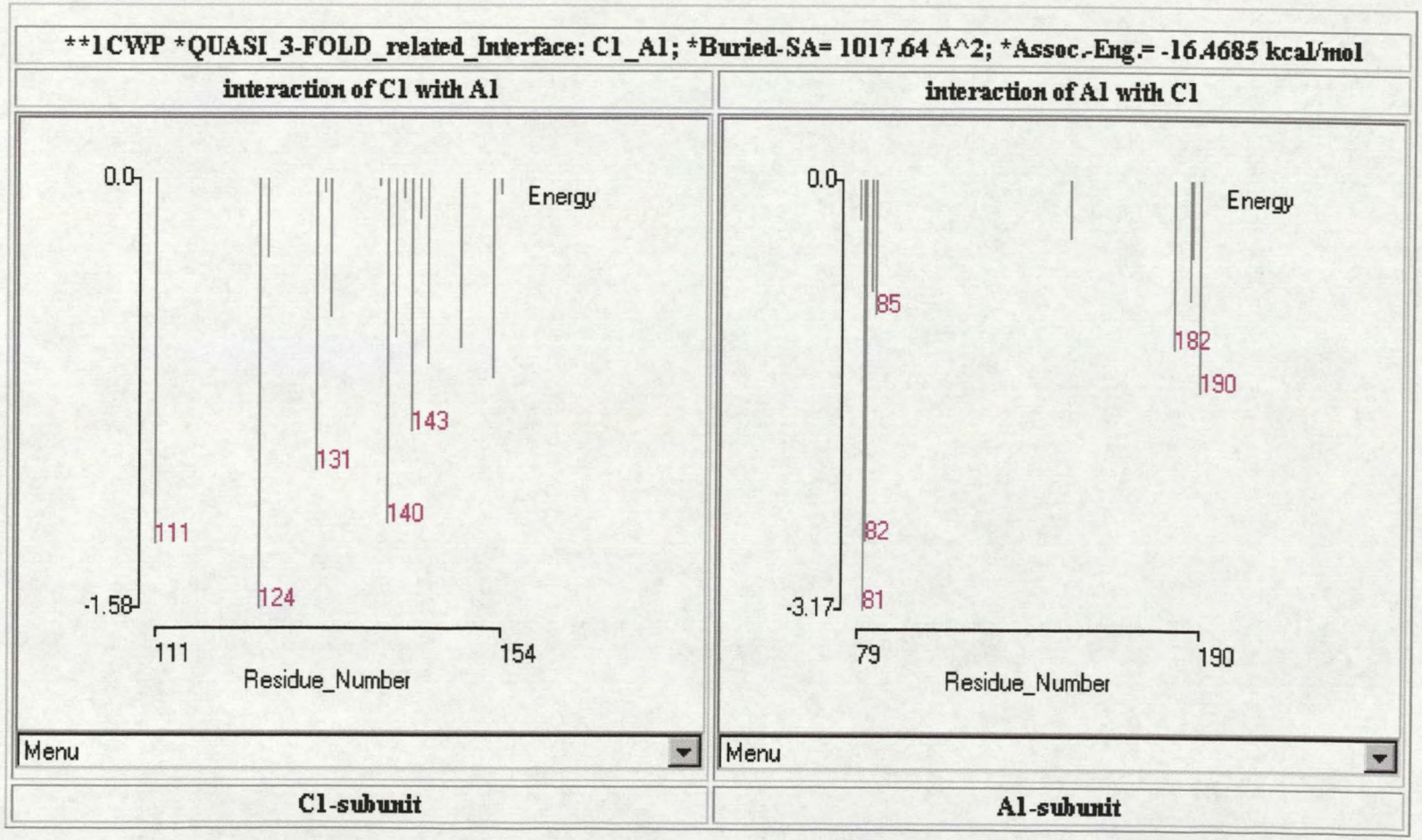
**Total buried of surface area = (A.S.A of A + A.S.A of B) - A.S.A of the complex(A-B).
Sometimes the buried surface areas are reported as 1/2 the the values reported here.

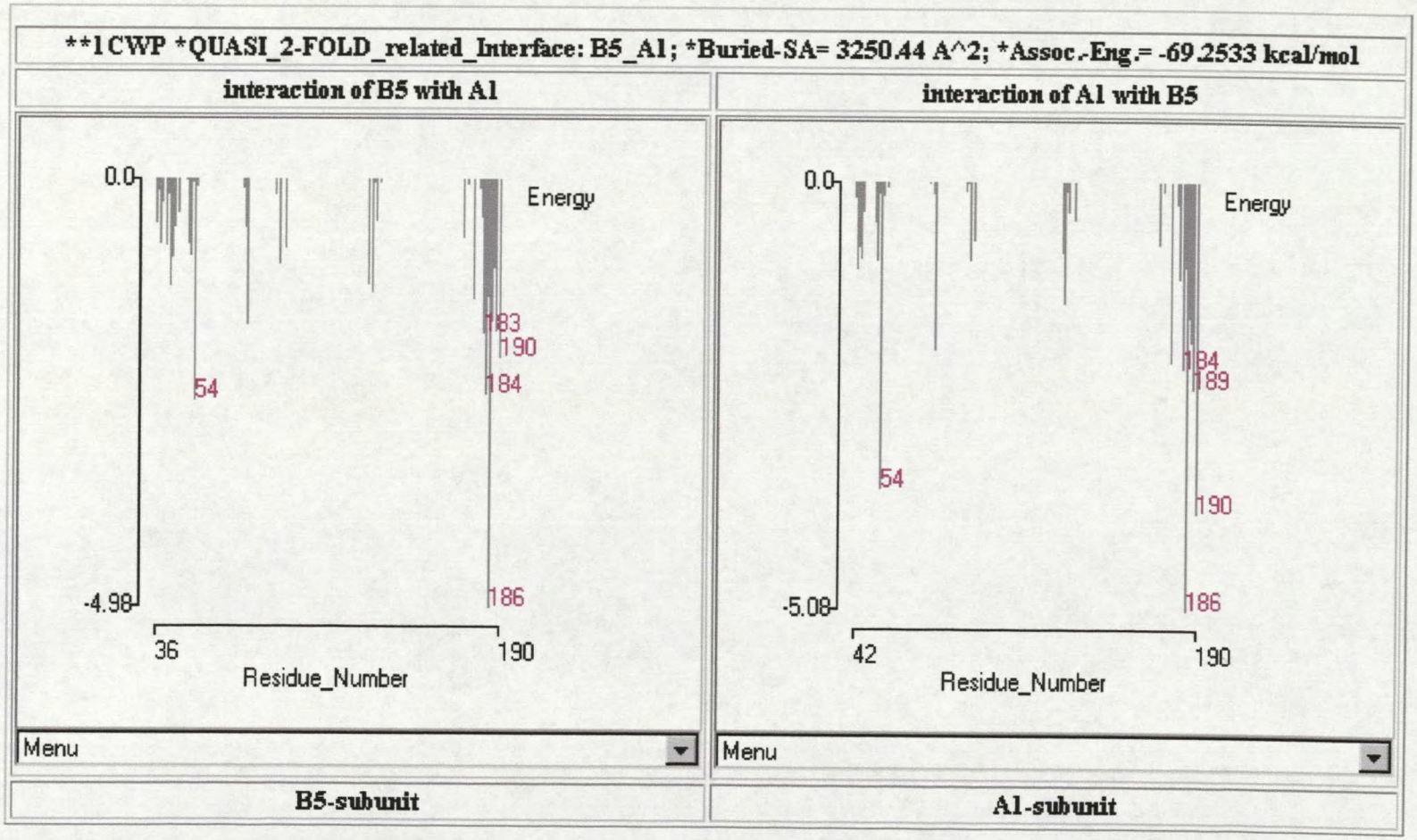
VIPER

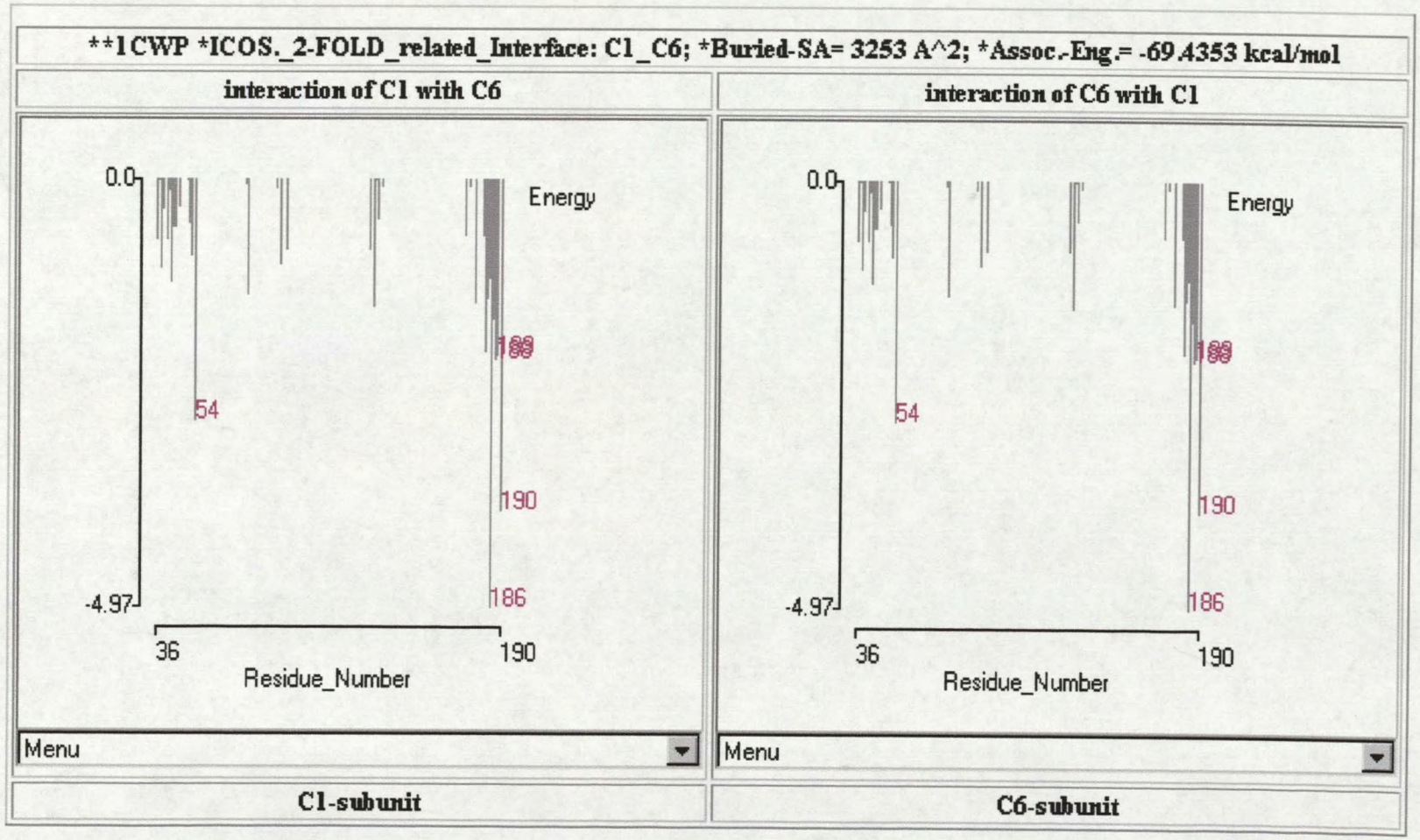
http://mmtsb.scripps.edu/viper/viper.html

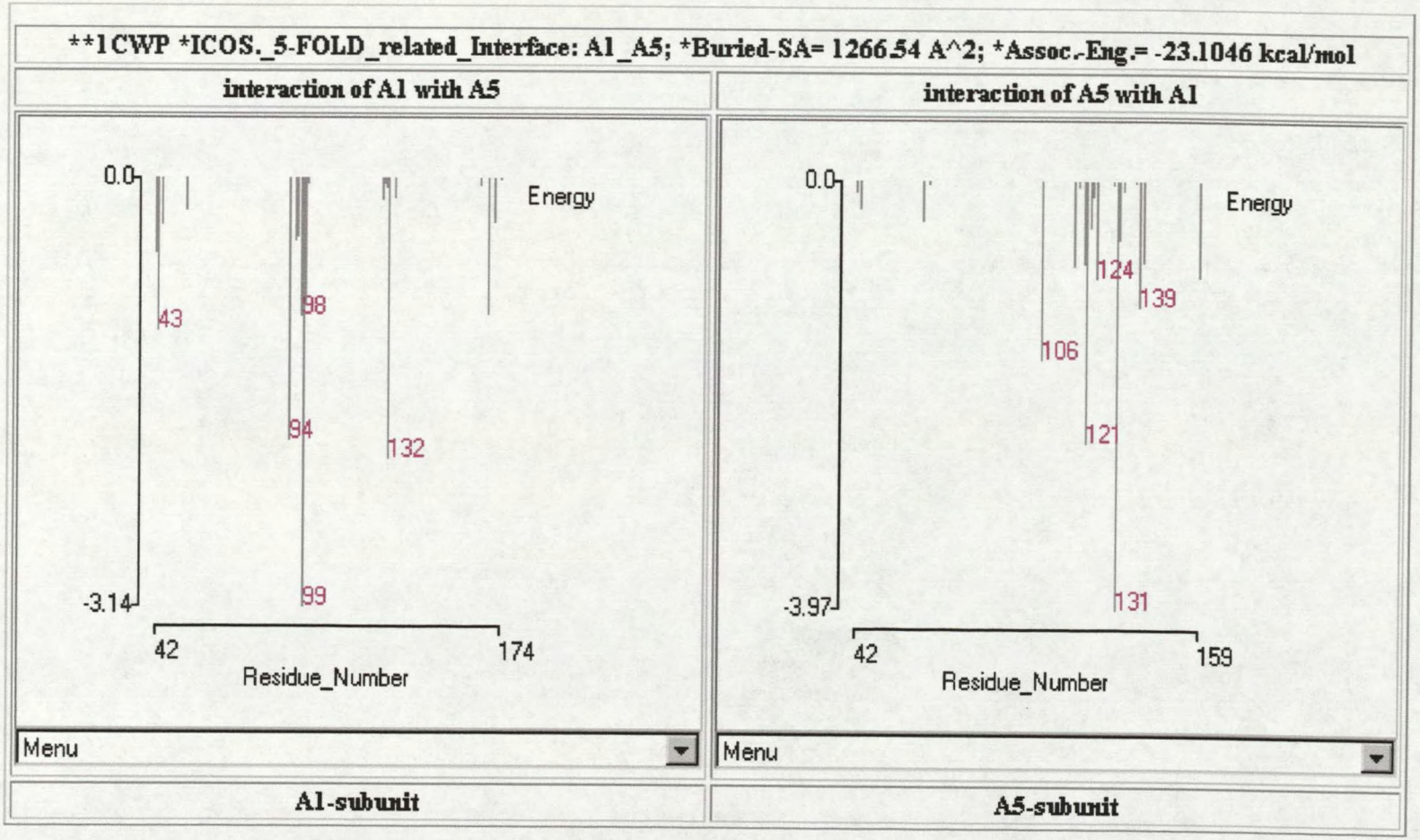

**1CWP *QUASI_3-FOLD_related_Interface: A1_B1; *Buried-SA= 1228.64 A^2; *Assoc.-Eng.= -20.7682 kcal/mol
interaction of A1 with B1
0.0
Energy
111
124
140
143
145
-2.69
111
154
Residue_Number
Menu
A1-subunit
interaction of B1 with A1
0.0
Energy
82
81
182
188
190
-3.61
79
190
Residue_Number
Menu
B1-subunit

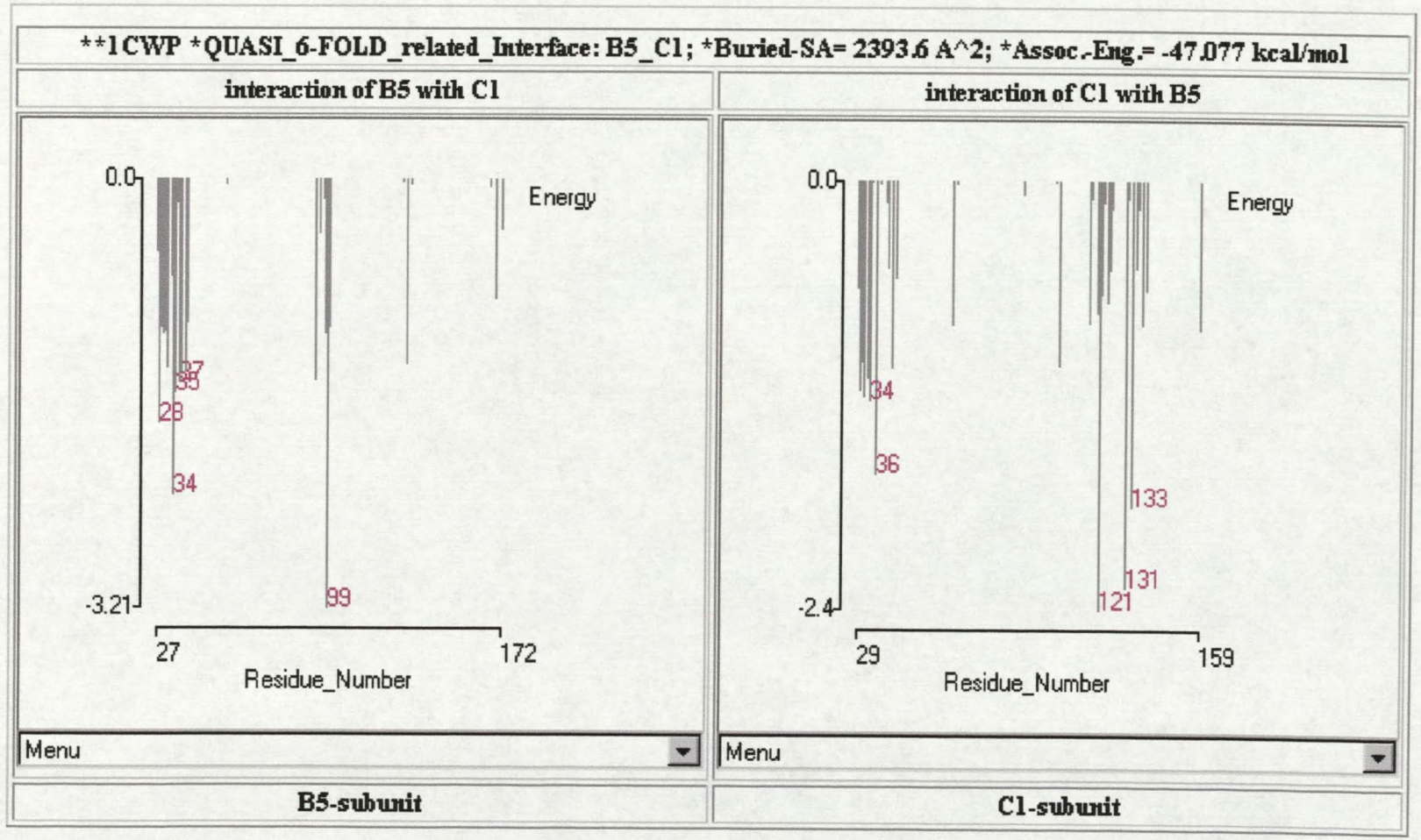

**1CWP *QUASI_3-FOLD_related_Interface: B1_C1; *Buried-SA= 1092.2 A^2; *Assoc.-Eng.= -18.0718 kcal/mol
interaction of B1 with C1
0.0
Energy
145
131
140
111
124
-1.69
87
154
Residue_Number
Menu
B1-subunit
interaction of C1 with B1
0.0
Energy
187
85
190
82
-3.35
81
79
190
Residue_Number
Menu
C1-subunit

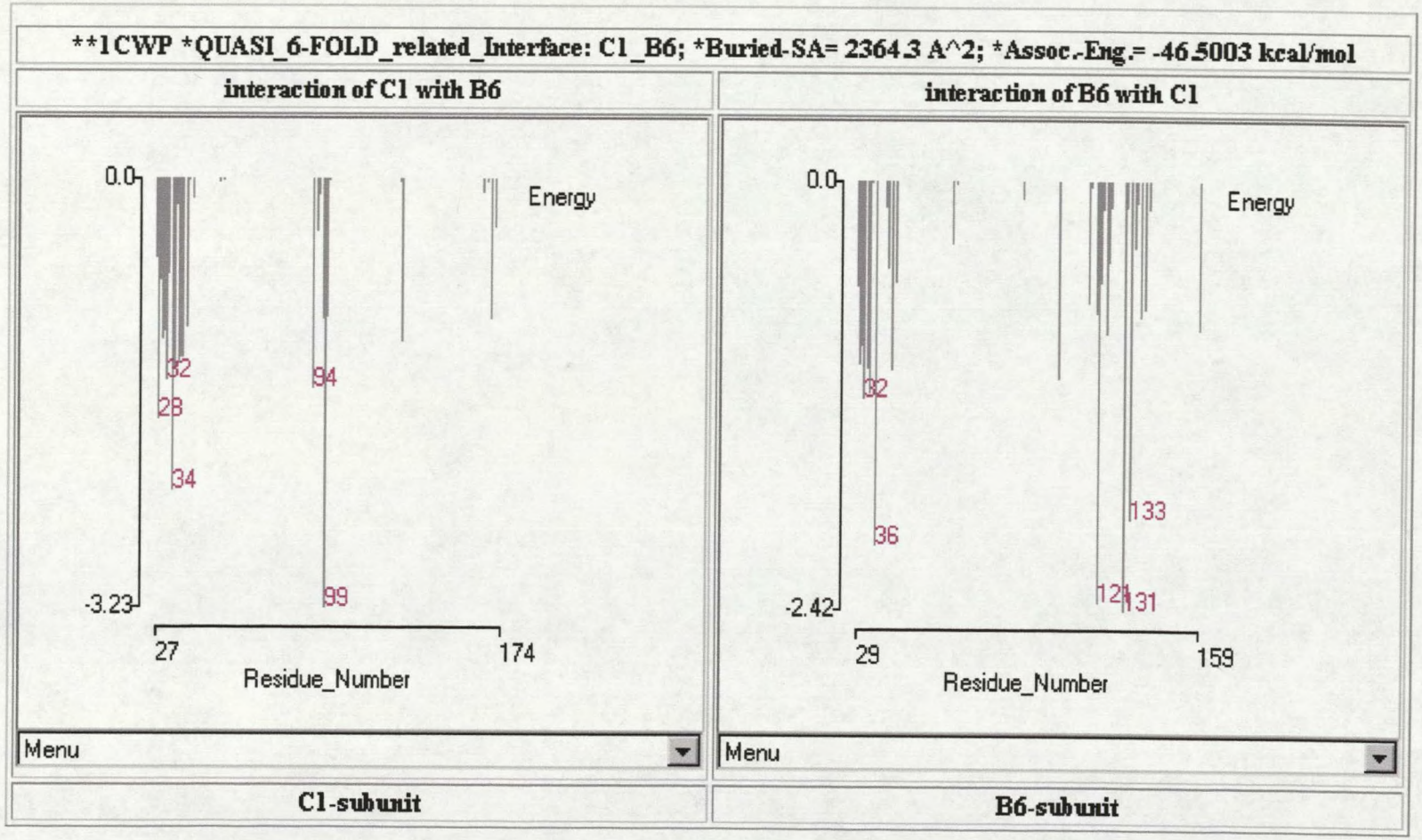

**1CWP *QUASI_3-FOLD_related_Interface: C1_A1; *Buried-SA= 1017.64 A^2; *Assoc.-Eng.= -16.4685 kcal/mol
interaction of C1 with A1
0.0
Energy
111
124
131
140
143
-1.58
111
154
Residue_Number
Menu
C1-subunit
interaction of A1 with C1
0.0
Energy
85
182
190
82
81
-3.17
79
190
Residue_Number
Menu
A1-subunit

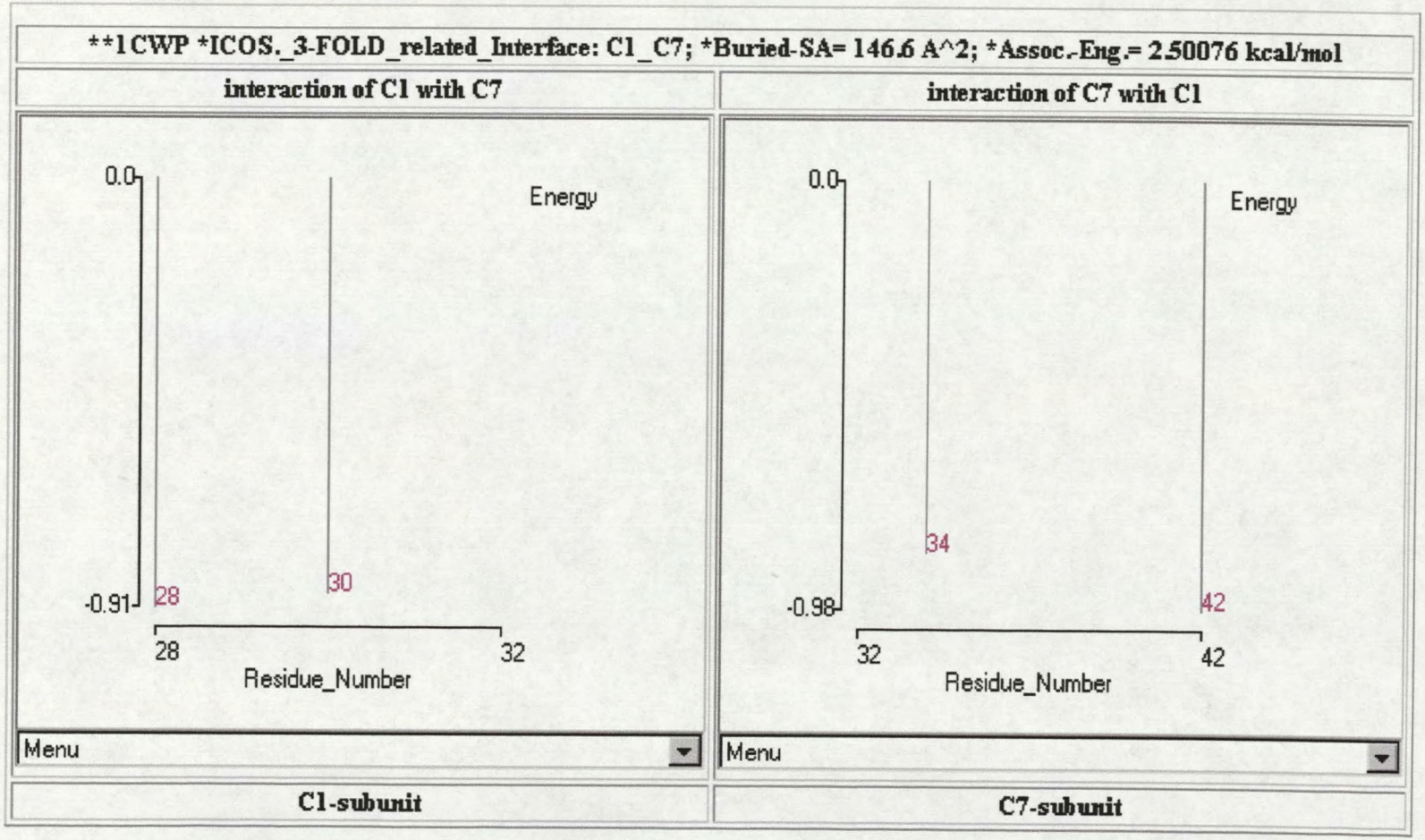

**1CWP *QUASI_2-FOLD_related_Interface: B5_A1; *Buried-SA= 3250.44 A^2; *Assoc._Eng.= -69.2533 kcal/mol
interaction of B5 with A1
0.0
-4.98
Energy
54
183
190
184
186
36
190
Residue_Number
Menu
B5-subunit
interaction of A1 with B5
0.0
-5.08
Energy
54
184
189
190
186
42
190
Residue_Number
Menu
A1-subunit

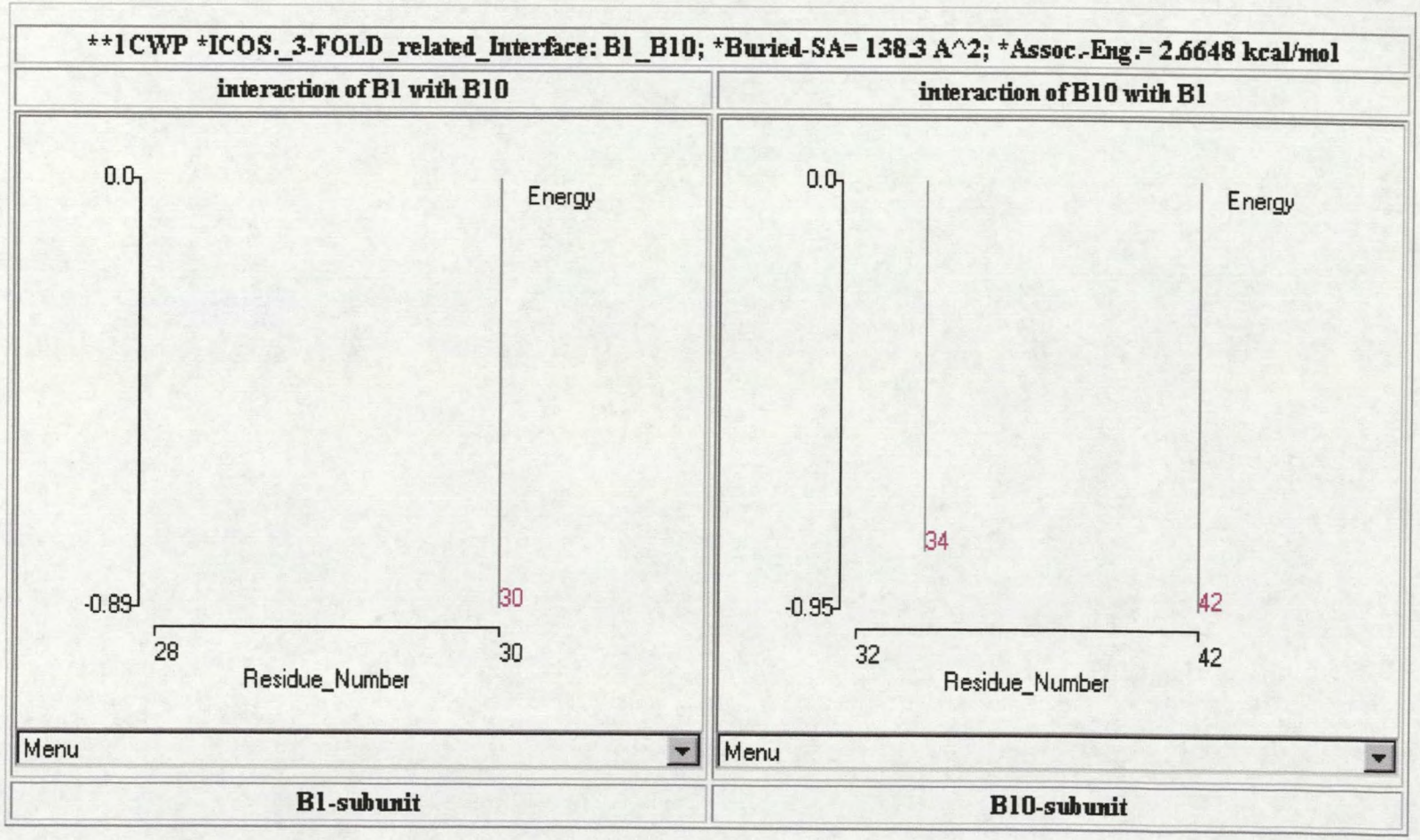

**1CWP *ICOS._2-FOLD_related_Interface: C1_C6; *Buried-SA= 3253 A^2; *Assoc.-Eng.= -69.4353 kcal/mol
interaction of C1 with C6
interaction of C6 with C1
0.0
-4.97
Energy
54
188
190
186
36
190
Residue_Number
Menu
C1-subunit
C6-subunit

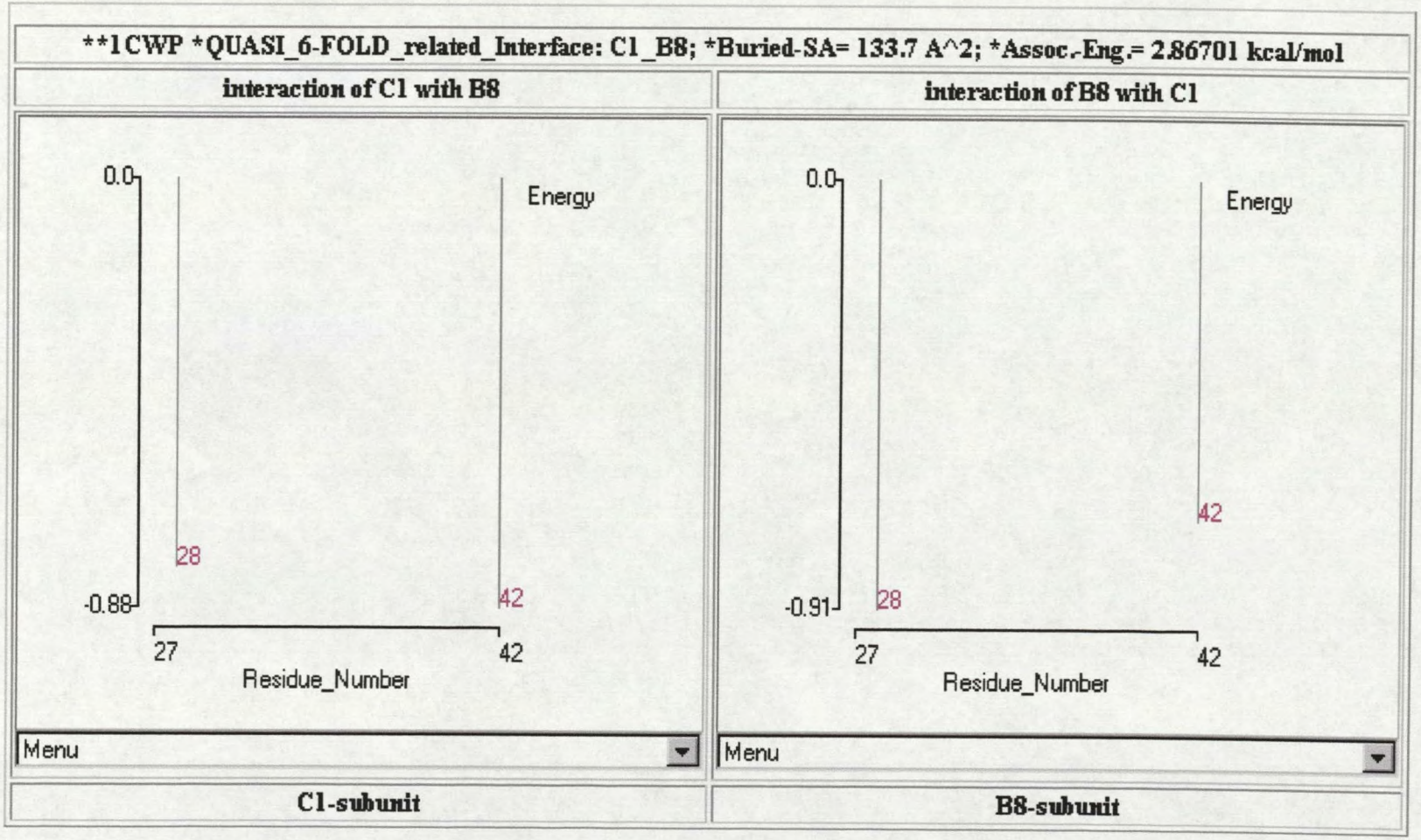

**1CWP *ICOS._5-FOLD_related_Interface: A1_A5; *Buried-SA= 1266.54 A^2; *Assoc._Eng.= -23.1046 kcal/mol
interaction of A1 with A5
0.0
Energy
43
88
94
132
99
-3.14
42
174
Residue_Number
Menu
A1-subunit
interaction of A5 with A1
0.0
Energy
124
139
106
121
131
-3.97
42
159
Residue_Number
Menu
A5-subunit

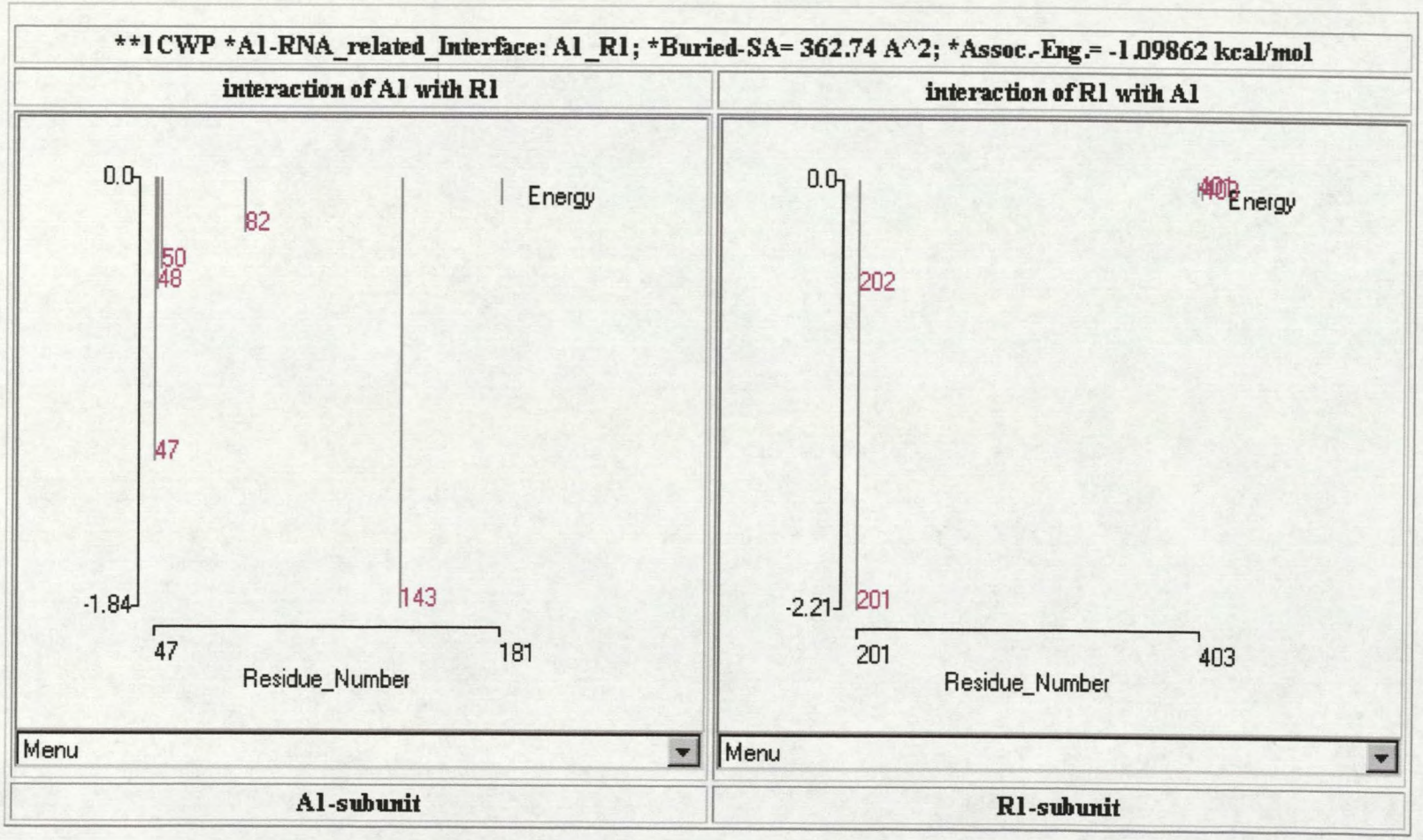

**1CWP *QUASI_6-FOLD_related_Interface: B5_C1; *Buried-SA= 2393.6 A^2; *Assoc.-Eng.= -47.077 kcal/mol
interaction of B5 with C1
0.0
Energy
-3.21
27
28
34
99
27
172
Residue_Number
Menu
B5-subunit
interaction of C1 with B5
0.0
Energy
34
36
133
131
121
-2.4
29
159
Residue_Number
Menu
C1-subunit

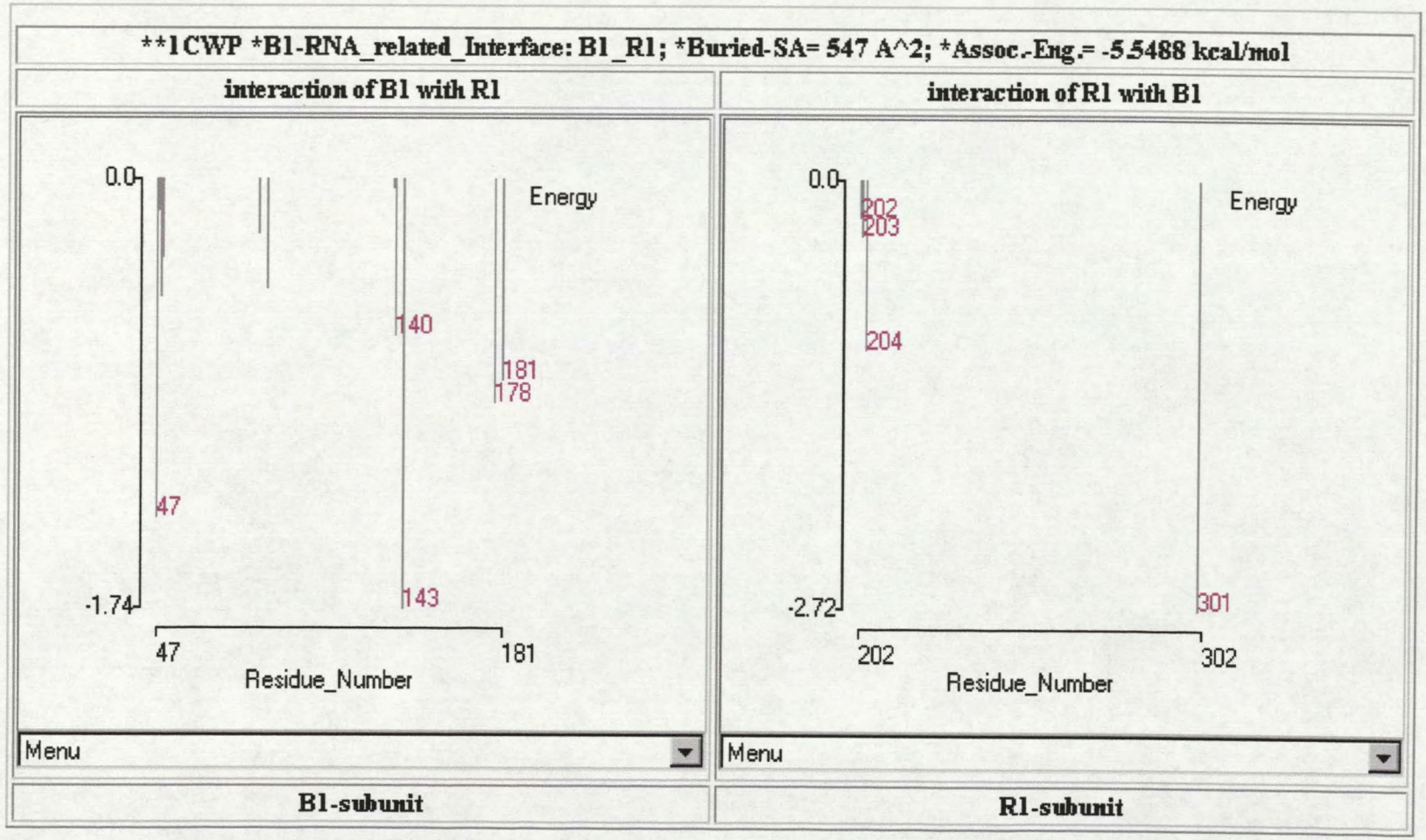

**1CWP *QUASI_6-FOLD_related_Interface: C1_B6; *Buried-SA= 2364.3 A^2; *Assoc.-Eng.= -46.5003 kcal/mol
interaction of C1 with B6
0.0
-3.23
Energy
32
28
34
94
99
27
174
Residue_Number
Menu
C1-subunit
interaction of B6 with C1
0.0
-2.42
Energy
32
36
133
121
131
29
159
Residue_Number
Menu
B6-subunit

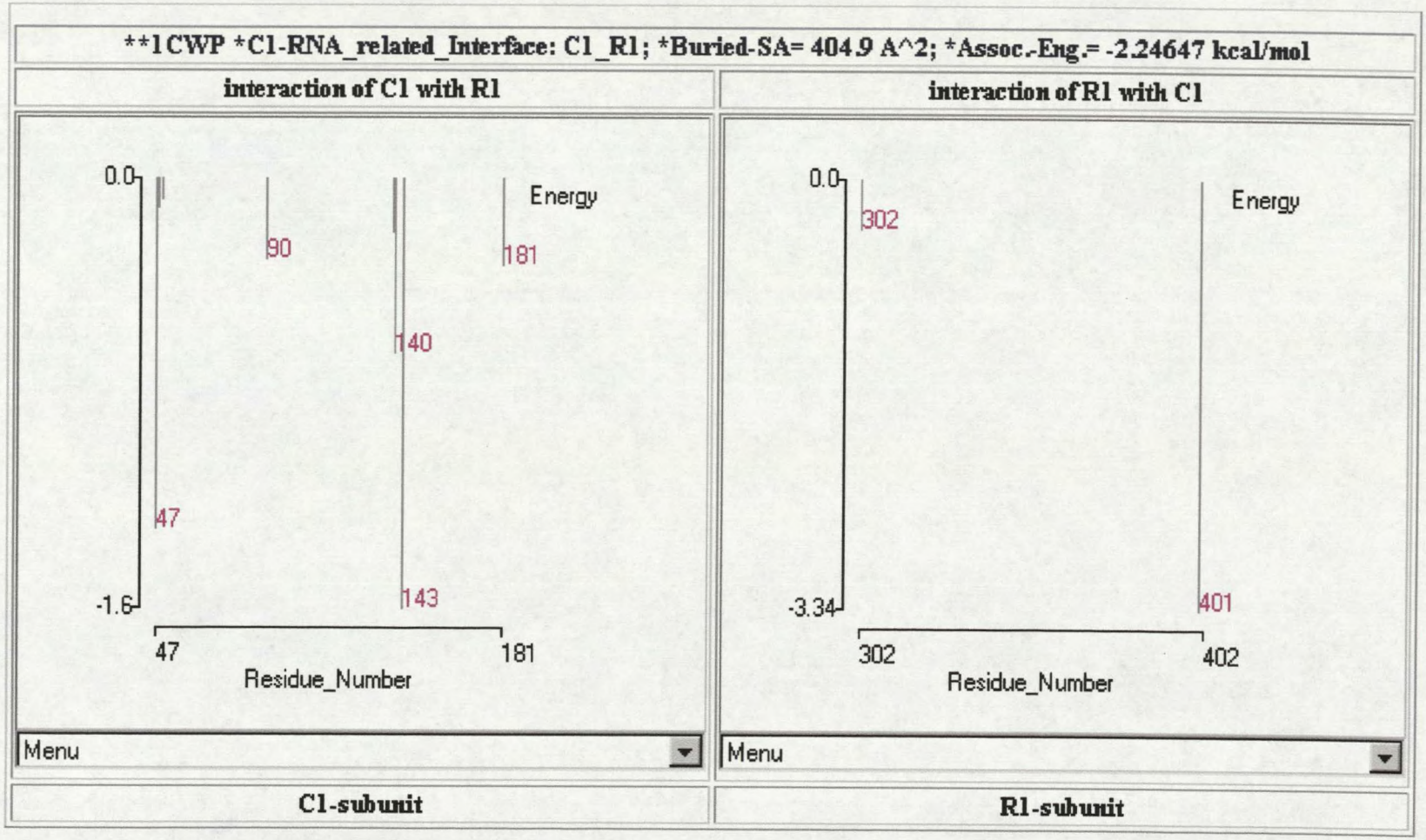

**1CWP *ICOS._3-FOLD_related_Interface: C1_C7; *Buried-SA= 146.6 A^2; *Assoc._Eng.= 2.50076 kcal/mol
interaction of C1 with C7
0.0
Energy
-0.91
28
30
28
32
Residue_Number
Menu
C1-subunit
interaction of C7 with C1
0.0
Energy
34
-0.98
42
32
42
Residue_Number
Menu
C7-subunit


**1CWP *ICOS._3-FOLD_related_Interface: B1_B10; *Buried-SA= 138.3 A^2; *Assoc.-Eng.= 2.6648 kcal/mol
interaction of B1 with B10
0.0
Energy
-0.89
30
28
30
Residue_Number
Menu
B1-subunit
interaction of B10 with B1
0.0
Energy
34
-0.95
42
32
42
Residue_Number
Menu
B10-subunit


**1CWP *QUASI_6-FOLD_related_Interface: C1_B8; *Buried-SA= 133.7 A^2; *Assoc.-Eng.= 2.86701 kcal/mol
interaction of C1 with B8
0.0
-0.88
Energy
28
42
27
42
Residue_Number
Menu
C1-subunit
interaction of B8 with C1
0.0
-0.91
Energy
42
28
27
42
Residue_Number
Menu
B8-subunit


**1CWP *A1-RNA_related_Interface: A1_R1; *Buried-SA= 362.74 A^2; *Assoc.-Eng.= -1.09862 kcal/mol
interaction of A1 with R1
interaction of R1 with A1
0.0
-1.84
Energy
82
50
48
47
143
47
181
Residue_Number
Menu
0.0
-2.21
Energy
202
201
201
403
Residue_Number
Menu
A1-subunit
R1-subunit


**1CWP *B1-RNA_related_Interface: B1_R1; *Buried-SA= 547 A^2; *Assoc.-Eng.= -5.5488 kcal/mol
interaction of B1 with R1
0.0
-1.74
Energy
47
140
143
181
178
47
181
Residue_Number
Menu
B1-subunit
interaction of R1 with B1
0.0
-2.72
Energy
202
203
204
301
202
302
Residue_Number
Menu
R1-subunit


**1CWP *C1-RNA_related_Interface: C1_R1; *Buried-SA= 404.9 A^2; *Assoc.-Eng.= -2.24647 kcal/mol
interaction of C1 with R1
0.0
-1.6
Energy
90
181
140
47
143
47
181
Residue_Number
Menu
C1-subunit
interaction of R1 with C1
0.0
-3.34
302
Energy
401
302
402
Residue_Number
Menu
R1-subunit

APPENDIX B

SEQUENCING DATA

Below are the DNA sequences that were generated from sequencing indicated plasmids or lyticase PCR amplification products from *P. pastoris* genomic DNA samples. Only regions surrounding the peptide 11 insertion sequence are shown (wild type CCMV sequence is shown above each construct sequence with corresponding base pair number in bold for comparison). Sequencing protocols are outlined in Chapter 3, Methods.

pCC3/pep11-loop 63, pPicZA/pep11-loop 63, and *P. pastoris* isolates with integrated chimeric CCMV coat protein-peptide 11 (in loop 63) gene :

(1540) **(1551)**
GCTGCC GAAGC**T**
GCTGCCGGGATCTGCGATCCCGGGTATATAGGCTCGCGATGTGGGATCCTGGAAGCT

pCC3/pep11-loop 102, pPicZA/pep11-loop 102, and *P. pastoris* isolates with integrated chimeric CCMV coat protein-peptide 11 (in loop 102) gene:

(1657) **(1668)**
AGTGTT AGTGG**C**
AGTGTTGGGATCTGCGATCCCGGGTATATAGGCTCGCGATGTGGGATCCTGAGTGGC

pCC3/pep11-loop 114, pPicZA/pep11-loop 114, and *P. pastoris* isolates with integrated chimeric CCMV coat protein-peptide 11 (in loop 114) gene:

(1696) **(1707)**
CAGACT ACTGC**T**
CAGACTGGGATCTGCGATCCCGGGTATATAGGCTCGCGATGTGGGATCCTGACTGCT

pCC3/pep11-loop 129, pPicZA/pep11-loop 129, and *P. pastoris* isolates with integrated chimeric CCMV coat protein-peptide 11 (in loop 129) gene:

(1741) **(1752)**
GACAAC TCGAA**A**
GACAACGGGATCTGCGATCCCGGGTATATAGGCTCGCGATGTGGGATCCTGTCGAAA

APPENDIX C

THE CANCER PROBLEM

TABLE OF CONTENTS FOR APPENDIX C

INTRODUCTION 137
WHAT IS CANCER? 138
HOW DOES CANCER ARISE? 142
Viruses and the Oncogene Theory 142
Genetic Alteration and the Aneuploidy Hypothesis 147
Types of Genetic Instability in Cancer 149
Subtle sequence instabilities 149
Gross chromosomal abnormalities 153
TUMOR CELL METASTASIS 161
A Closer Look at Tumor Metastasis and Invasion 166
Tumor Cell Adhesion 166
Proteolytic Digestion and 169
Tumor Cell Migration 171
APPENDIX C – MIDWAY SUMMARY 171
OUR NATURAL DEFENSES AGAINST CANCER 172
Internal Mechanisms Inducing Apoptosis (Suicide) 173
External Mechamisms Inducing Apoptosis (Murder) 175
Caspases - the Common Pathway to Induced Cell Death 176
CANCER - NEW PROMISES IN THERAPY 178
Surgery 178
Radiation Therapy 179
Chemotherapy 181
Endocrine Therapy 183
Immuno-Based Therapies 184
Nonspecific Immunotherapies 184
Passive Immunization and Antibody-Based Therapies 184
Active Immunotherapy and Vaccine-Based Therapies 185
Adoptive Immunotherapy 185
New Molecular Therapies 186
CANCER PREVENTION AND DETECTION 188
Basic Early Detection 188
Internal Markers for Early Detection of Cancer 190
Genetic Testing 191
Tumor Staging and Grading 193
Tumor Staging 193
Tumor Grading 194
Cancer Imaging 195
X-ray Imaging 196

Computed Tomography (CT) 196
Directed radiography 197
Single Photon/Positron Emission Computed Tomography (SPECT/PET) 198
Ultrasound 199
Magnetic Resonance Imaging (MRI) 199
Providing contrast in MRI 201
REFERENCES CITED FOR APPENDIX C 207

Introduction

As early as the days of Hippocrates, the ancient Greek physicians had a clear understanding of the ways in which a malignant growth so often pursues its inexorable determination to destroy life. Based on physical findings, they gave a name to the hard swellings and ulcerations we now know as cancer. To the term karkinos, or "crab" they added the suffix oma, referring to "tumor"; karkinoma was used to describe a malignant tumorous growth.

Cancer, the Latin word for "crab", eventually came into more common usage. In fact, Galen, an early Roman codifier of Greek medicine, described cancerous lesions he observed in the breasts of women as "just like a crab's legs extending outward from every part of its body."[1]

Since Galen's early description of the disease, cancer has received enormous biomedical attention, and today is recognized as a major human health problem. The disease has become the second-ranking cause of death in the modern Western world, led only by heart disease. Current estimates project that one person in three in the United States will develop cancer; one in five will die from it. That adds up to about 550,000 cancer related deaths every year. More than 100 different types of human cancers have been characterized, advanced methods of cancer detection and treatment are continually being developed, and cancer epidemiology has been extensively characterized. All this considered, we are nevertheless only beginning to understand this grave disease.

The goal of this thesis Appendix section is to provide a surface-level review of cancer: what cancer is, what causes it, our "natural" defenses against it, and finally how

cancer is detected and treated. Particular attention will be paid to two topics: tumor metastasis and modern cancer imaging techniques.

What is cancer?

The multiplication of cells in higher multicellular organisms is normally carefully regulated and responsive to the specific needs of the animal. In a young animal, cell multiplication exceeds cell death, so the animal increases in size; as adulthood is approached, a homeostatic balance is usually established between cell renewal and cell death in most organs and tissues. The various types of mature cells in the body have a given life span; as these cells die, new cells are generated by the proliferation and differentiation of various types of stem cells. For some cell types, like intestinal cells with a half-life of a few days, renewal is rapid. Other cells live longer lives: red blood cells live approximately 100 days, healthy liver cells rarely die, adult brain cells die a slow death and their regeneration is virtually non-existent. This turnover of cells is normally tightly regulated so that the numbers of any particular cell type remains constant (outside of normal growth during development). Occasionally however, some influence, or combination of influences, whether genetic or environmental, acts to interfere with the maturation of a cell early enough in developmental pathways endowing it with the ability to divide uncontrollably. When such a cell has descendants that inherit the propensity to proliferate the result is a clone of autonomously growing cells with the inability to further differentiate into fully mature cells. These cells can eventually give

rise to the hard mass called a tumor (or neoplasm) that Galen noticed in his patients. A tumor that is not capable of indefinite growth and does not invade the healthy tissue that surrounds it is considered benign. A wart is a common example of a benign neoplasm. A tumor that continues to grow and progressively becomes invasive is considered malignant. In clinical terms, cancer specifically refers to a malignant neoplasm. And as if uncontrolled growth were not enough, unruly clusters of cancer cells can dislodge from the primary tumor and spread to distant sites. This process, called metastasis (discussed further below), is the major characteristic that differentiates malignancies from benign tumors.

For several thousand years cancer was visible only in its outward manifestations. It was the invention of the microscope that revealed the cancer cell itself, enabling its differentiation from normal cells. Under the microscope, the cancer cell retains an immature and different appearance – they may "bulge, flatten, elongate, round themselves out...; the nucleus is larger and more prominent than that of mature relatives...; and chromosomes... align themselves in bizarre patterns, attempting with varying degrees of success to rapidly multiply."[1] In a specific tissue, malignant cells look a lot like rapidly growing cells: they have a high nucleus-to-cytoplasm ratio, prominent nucleoli, aneuploidy (the abnormal balance or number of chromosomes), and relatively little specialized structure (Fig. 26).



Fig. 26. Cancer tissue has a distinctive appearance under the microscope. Among the traits the doctor looks for are a large number of dividing cells, variation in nuclear size and shape, variation in cell size and shape, loss of specialized cell features, loss of normal tissue organization, and a poorly defined tumor boundary. (Images adapted from NCI, library of figures).

Instead of finding a benign or malignant tumor, microscopic examination of a biopsy specimen will sometimes detect a condition called hyperplasia. Hyperplasia refers to tissue growth based on an excessive rate of cell division, leading to a larger than usual number of cells. Nonetheless, cell structure and the orderly arrangement of cells within the tissue remain normal, and the process of hyperplasia is potentially reversible and is not necessarily dangerous. In addition to hyperplasia, examination of a specimen may reveal another noncancerous condition called dysplasia. Dysplasia is an abnormal type of excessive cell proliferation characterized by loss of normal tissue arrangement and cell structure. Often such cells revert back to a normal phenotype, but occasionally, they become malignant. The most severe cases of dysplasia are sometimes referred to as

carcinoma *in situ*, an uncontrolled growth of cells that remains in its original location. Carcinoma in situ can develop into an invasive, metastatic malignancy (see below).

Like normal animal cells, tumors are generally classified according to the embryonic origin of the tissue from which they are derived. Carcinomas are tumors arising from endodermal or ectodermal tissues such as skin or the epithelium of internal organs and glands. Sarcomas, which arise less frequently, are derived from the mesoderm (connective tissues such as bone, fat, and cartilage). The leukemias and lymphomas are malignant tumors of hematopoietic cells of the bone marrow. Leukemias proliferate as single cells in the blood, whereas most other tumors including lymphomas tend to grow as solid masses. A variety of prefixes are used to distinguish among the many different types of carcinomas, sarcomas, lymphomas, and leukemias. For instance, a cancer arising in bone is called an osteosarcoma; a cancer of gland cells is called adenocarcinoma.

PREFIX	MEANING
adeno-	gland
chrondro-	cartilage
erythro-	red blood cell
hemangio-	blood vessels
hepato-	liver
lipo-	fat
lympho-	lymphocyte
melano-	pigment cell
myelo-	bone marrow
myo-	muscle
osteo-	bone

Fig. 27. Some prefixes used in naming cancers.

Having presented a basic picture of what cancer is and what it looks like, the next logical question might be: what causes it; how does cancer arise? In the next section of this chapter, evidence will be presented regarding three hypotheses commonly used to answer this question. They are: the oncogene theory, the gene mutation hypothesis, and the aneuploidy hypotheis. The purpose of discussing these theories is to fortify the conceptual foundation that underlies my thesis research.

How Does Cancer Arise?

John Hunter, the preeminent 17th Century British surgeon, believed there were three major causes of cancer: age, hereditary disposition and possibly climate. Perhaps he wasn't so far off. Biotechnological developments over the past two decades have enabled scientists to pursue, more deeply, the mechanisms that trigger cancer's uncontrolled and deadly growth.

Viruses and the Oncogene Theory

In 1911, Peyton Rous put forth the theory that viruses could cause cancer. This was based on studies in which he passed a cell-transforming agent from chicken fibrosarcomas to young chicks, inducing tumor cell formation . The agent in the filtrate was eventually characterized as a virus (duly named Rous Sarcoma Virus – RSV). The fields of virology and oncology have since remained intertwined.

What was in RSV that enabled it to turn normal cells into cancer cells? Further studies later demonstrated that RSV is a retrovirus that, like the human immunodefiency virus, reverse transcribes its RNA genome into DNA, which is subsequently incorporated into the host's genome. The base sequence of the RSV genome reveals that it contains four genes of which only three are critical for viral replication: *gag*, which encodes an inner capsid protein; *pol* encodes a reverse transcriptase; and *env* codes for the outer envelope protein. The fourth gene in RSV, named *v-src* (*v* for viral, *src* for sarcoma), encodes a 60 kDa protein tyrosine kinase which maintains the host cell in a proliferative (transformed) state, presumably enhancing viral replication rates. Because of its transforming ability, *v-src* was therefore termed an oncogene (literally, "tumor gene"). In 1969 George Todaro and Robert Huebner of the National Cancer Institute formulated the oncogene concept by which a cancer-causing gene produced by the virus initiates the uncontrolled cell growth that is cancer[2]. In 1971 Howard Temin and collaborators suggested that oncogenes might be acquired by a virus from the genome of an infected cell[3,4,5]. These cellular host genes, termed proto-oncogenes, or cellular oncogenes, were later verified by Michael Bishop and Harold Varmus, when they discover that chicken cells not infected with RSV contained a gene that is homologous to *v-src*[6]. This gene, named *c-src* (*c* for cellular) is highly conserved in a wide variety of eukaryotes including humans.

The proteins encoded by many viral oncogenes are either analogs of various cell-cycle, growth factor receptor and endocrine system components, or directly interact with these elements. Examples abound:

The *v-sis* oncogene of simian sarcoma virus encodes a secreted protein homologous to one the two subunits of the mitogen (a substance that induces mitosis), platelet derived growth factor (PDGF).

Oncogene-encoded protein kinases, like *v-src* from RSV, inappropriately phosphorylate target proteins normally recognized by growth factor receptors, thereby inducing proliferation. The *v-ras* oncogene encodes a protein, p21, that functionally resembles G-proteins, activating a number of downstream kinases through Ser/Thr posphorylation. *v-ras* hydrolizes GTP, but more slowly than its cellular analog. The restraint to protein phosphorylation normally afforded by fast(er) hydrolysis of GTP is therefore relieve by *v-ras*, thereby transforming the cell.

Numerous retroviral oncogenes (e.g. *v-jun* and *v-fos*) encode nuclear transcription factors whose cellular analogs are normally synthesized in response to mitogenic signals then, along with a variety of other elements, bind DNA and influence transcription and/or replication.

Spleen focus-forming virus (SFFV) is a retrovirus that induces erythroleukemia in adult mice. SFFV encodes a mutant envelope glycoprotein called gp55. This protein is missing most of its extracellular domain and the entire cytosolic domain. It therefore cannot function as a normal envelope protein would, mediating viral budding and

infection, but rather it binds to and inactivates Epo receptors expressed in the same cell. Like the Epo protein, gp55 stimulates cellular proliferation. Its constituitive action eventually potentiates the formation of malignant clones of erythroid progenitor cells.

Unlike the retroviral oncogenes, which appear to derive from normal cellular genes and have no direct function for the virus, known oncogenes from DNA-viruses are integral components to the viral genome and are required for viral replication. The DNA virus, human papillomaviruses (HPV) utilizes one of its oncogene products in a fashion that is similar to that of SFFV to induce transformation. The protein encoded by the early region gene, *E5* forms a transmembrane dimer or trimer and can aggregate PDGF receptors on the cell surface membrane. This mimics hormone-mediated receptor dimerization, causing sustained receptor activation which may potentiate cellular transformation.

HPV has a variety of other strategies for transforming cells beyond just PDGF receptor mimicry: The HPV E6 protein is involved in cellular transformation, partly through inactivation of cellular tumor suppressing protein p53 (described below). E6 is reported to act in conjunction with E6 associated protein (E6-AP), tagging p53 for degradation via the ubiquitin-protease degradation pathway[7]. E6-AP is a 100 kDa cellular protein with ubiquitin ligase activity. Many p53 mutants are resistant to E6-mediated degradation, suggesting that the interaction between these two proteins can be easily disrupted[8]. E6 is associated with cellular minichromosome maintenance (MCM) proteins, which help to regulate cellular DNA replication. E6 may also activate

telomerase, bind to paxillin, transactivate the c-myc promoter, upregulate epidermal growth factor receptor, and more recently has been shown to bind to a novel GTPase-activating protein (GAP) called the E6 targeting protein 1[7,9,10]. The common theme here is that E6 mediates regulation of cell signaling and/or the cell cycle during infection and pathogenesis.

Another HPV protein, E7, helps to regulate cellular transformation. E7 interacts with the tumor suppressor protein, retinoblastoma (Rb) (see below). Like the E6 interaction with p53, E7 tags Rb for degradation through the ubiquitin-protease pathway. One consequence of the E7-Rb interaction is the disruption of a complex between Rb and the E2F-1 transcription factor. E7-mediated release of E2F from these complexes is thought to influence the expression of genes involved in cell cycle progression. Other cell cycle regulatory elements that are targeted by HPV E7 include p21, p27, cyclin A, TATA binding protein (TBP), and members of the AP-1 family of transcription factors[7]. A recent paper describes the binding and modulation of the glycolytic enzyme, M2 pyruvate kinase (M2-PK), by the HPV16 E7 protein[11]. Pyruvate kinase (PK) is a key enzyme involved in regulating the glycolytic flux in tumor cells. It is hypothesized that the expansion of phosphometabolite pools resulting from the down-regulation of M2-PK substrate affinity (through interaction with HPV E7) is required for the increased biosynthetic activity of rapidly proliferating tumor cells.

The different examples of viral-induced cellular transformation demonstrate that viral factors (often viral oncogenes) generally function to modify or inappropriately

express components of the elaborate control networks regulating cell growth and differentiation. The complexity of these networks (the cross-talk pathways involved in DNA replication and repair, transcription, translation, and cellular response to growth factors and hormones) is part of the reason why malignant transformation is thought to include several carcinogenic events.

Genetic Alteration and the Aneuploidy Hypothesis

The study of viruses, particularly retroviruses, provided an early mechanistic model for cancer, putting scientists onto the trail of oncogenes and other theories that could explain the transformation from a normal cell to a tumor cell. Retroviruses, by way of their oncogenes, seem to induce a very specific type of oncogenesis characterized by: (i) transformation within one or a few cell generations; (ii) transformation depending on continued function of viral protein; (iii) transforming function of chromosomal DNA; (iv) polyclonal tumors; and (v) diploid tumors (at least initially)[12]. These characteristics meet all of the predictions of what became the "gene mutation hypothesis": the idea that oncogenesis is a multistep process of clonal evolution driven by a series of somatic (and sometimes germline) mutations that accumulate and override the normal mechanisms controlling cellular proliferation. Progressively, the cell is converted from normal growth to a precancerous stage and eventually into a cancerous state[13,14].

The search for specific genes implicated in oncogenesis has turned many promising candidates. However, none of these cellular oncogenes or tumor suppressor

genes are activated or deleted from *all* cancers. Even tumors of a single organ rarely have uniform genetic alterations, although tumor types from one specific organ have a tendency to share mutations in certain genes or in different genes within a single growth-regulatory pathway[15]. The mechanism(s) through which "specific" mutations are generated (activating proto-oncogenes and/or inactivating tumor suppressor genes) and accumulate have therefore become the subject of considerable debate.

Recently, a link between carcinogenesis and general genomic instability has been established, possibly resolving (at least in part) this debate. It is argued that underlying genetic instability is required for the generation of the multiple and heterogeneous mutations that seem to underlie most cancers[14]. Genetic instability can be conceptually organized into two major categories: *subtle sequence instabilities* and *gross chromosomal abnormalities*. Below, I will review each of these and describe the downstream effects of instability in terms of tumor formation and progression.

Types of Genetic Instability in Cancer

Subtle sequence instabilities:

These changes involve base substitutions, deletions, or insertions of a few nucleotides and unlike the other types of instabilities (below), cannot be detected through cytogenetic analysis[14]. Types of subtle sequence instabilities include nucleotide-excision repair- (NER)-related instability - NIN, APC instability, and microsatellite instability - MIN.

Nucleotide-excision repair- (NER)-related instability (NIN): Two major processes contribute to controlling errors in, and maintaining fidelity of, DNA replication: DNA polymerization (and intrinsic proofreading activity of DNA polymerases) and DNA repair mechanisms. Nucleotide-excision repair (NER) is one of these mechanisms; defects in NER leads to subtle sequence instabilities (NIN). NER is normally responsible for repairing damage caused by many exogenous mutagens. The basic mechanism of NER action involves endonuclease-mediated nicking of the damaged strand (usually bulky lesions, 1- 20bp, from carcinogenic damage and ultraviolet (UV) light induced pyrimidine dimers) on both sides of the damage, 10 to 20 nucleotides away from it. Damage is removed by exonucleases then DNA polymerase and DNA ligase fill in the gaps. Transcription coupled repair (TCR), a subset of NER, uses a similar mechanism to repair mistakes generated during transcription. TCR could account, in

part, for the observation that active genes generally contain fewer mutations than non-transcribed DNA elements and could speculatively explain why certain gene regions appear to be mutation "hot spots" (e.g. the predominance of mutations in certain regions of the p53 tumor suppressor gene)[16]. The role of NER (and DNA repair in general) in cancer first came to light in studies of the group of rare autosomal recessive genetic diseases collectively termed xeroderma pigmentosum (XP). Cells from patients with XP are hypersensitive to the killing and mutagenic effects of UV light and are defective in the excision repair of their DNA[16]. These patients develop numerous skin tumors in sun-exposed areas; although the incidence of internal cancers is not raised to the same degree[14]. This may be do to an underlying link between UV damage and NER-specific activity. It is interesting that patients heterozygous for a NER-gene mutation are not especially tumor prone. This is probably because inactivation of the normal allele does not always lead to a high mutation rate; exposure to an environmental carcinogen (e.g. UV light) is probably required to engender such mutations[14].

APC instability: Another defect that could be lumped under the category of subtle sequence instabilities does not involve NER machinery. Inheritance of, or somatic alterations in, the adenomatous polyposis coli (*APC*) gene is causally linked with familial adenomatous polyposis (FAP). Interestingly, a single base change (adenine for thymine at position 1307 in *APC*) may render the surrounding DNA susceptible to mistakes. The polymorphism introduces a one-to-two base pair bubble that allows the DNA polymerase to incorporate extra bases without correcting its mistake. This type of instability,

different from microsatellite instability (see below), ultimately results in a defunct APC protein[17]. With APC knocked out, β-catenin, normally held by E-cadherin or destroyed, accumulates in the cell and reaches the nucleus were it relieves CBP- (Creb Binding Protein-) mediated repression of *c-MYC* (Lef and Tcf can associate with the transcription initiation complex and attach to the *c-MYC* promoter when CBP is not associated)[18,19]. This pathway provides rational for the observation that *c-MYC* is overexpressed in colorectal cancers (of which 85% are associated with mutant *APC*).

Microsatellite instability (MIN): Microsatellite instability (MSI or MIN) refers to the expansion or deletion of tandem repeat elements (e.g. poly(CA) repeats) in regions outside the centromeres and telomeres. They can be detected by comparing, on a DNA gel, PCR amplified fragments (usually using a panel of primers to various MS markers) from pared samples of normal DNA (usually from blood lymphocytes) and DNA from a clinical sample from the same individual. Mobility shifts in the clinical sample preparation are presumably due to length changes in microsatellite regions on one allele and is indicative of MIN. (The loss of one allele - loss of heterozygosity; LOH - from the clinical sample preparation results from deletion or mitotic recombination (see below) and is sometimes considered a form of MIN).

It is thought that MIN arises from a genetic defect in DNA, such as mismatched nucleotides (NTs), which destabilize double stranded (ds)DNA. There are at least three ways to get damage where mismatched NTs arise: 1.) Physical or chemical damage (resulting in nicks, gaps, dimers, etc.) including deamination, oxidative damage, and UV

damage; 2.) Misincorporation of NTs during DNA replication or repair to produce a mismatch, an insertion or a deletion; 3.) Genetic recombination to produce heteroduplex DNA with mismatched NTs from two different parental sequences[20]. These defects are thought to cause strand slippage (on the nascent lagging strand) and associated formation of either hairpins, triplexes or dissociation flaps (capable of forming other tertiary structures).

Although MIN is unusual in sporadic colorectal cancers (CRCs), MIN occurs in most tumors from patients with hereditary non-polyposis colorectal cancer (HNPCC). People with HNPCC are at an increased risk for early onset colorectal cancer and cancer at other sites (uterus, ovary, stomach, urinary tract, small bowel, and bile duct). About 75 percent to 80 percent of persons with HNPCC develop colon cancer. HNPCC is associated with mutations in any of five genes involved in DNA mismatch repair (MMR). Patients with HNPCC have one normal allele of the relevant MMR gene; this wild-type allele is sufficient to maintain normal levels of MMR. It is only when the wild-type allele is inactivated during tumorigenesis - through a gross chromosomal event (below) or through mutation – that MMR activity is abolished. Active MMR machinery binds to a DNA lesion, creates a nick, and then removes, via exonuclease activity, DNA between the nick and the mismatch (up to 256 bp). Then, DNA polymerase III and DNA ligase fill in and repair the gap. Without this important repair function, DNA mutations accumulate. In theory, the progeny of this cell then accumulate mutations in oncogenes and tumor-suppressor genes, thus resulting in the clonal dominance seen in tumorigenesis[14].

One gene frequently mutated in colon cancers as a result of defunct MMR encodes the type II receptor for Transforming Growth Factor β (TGFβ). Apparently, strand slippage causes a frameshift mutation in the coding-sequence that abolishes the normal functioning of the TGFβ receptor. This leads to cells that are resistant to the growth inhibiting action of this important cytokine, thereby contributing to the unregulated growth characteristic of these tumors.

Many cancers may contain a few microsatellite changes, however the extent of these changes is much less pronounced than in MMR-deficient cancers. MIN in cancer cells with proper functioning MMR is similar to that in normal cell lines and is probably due to near-normal mutation rates that occur during the multiple rounds of cellular division (during cancer cell clonal expansion)[14]. MMR deficiency therefore provides a mechanistic explanation for MIN (and the generation of subsequent tumor transforming mutations) in some hereditary cancers. However, most (sporadic) cancers are rarely (<2%) MMR deficient[14]. So the question remains: what is are the underlying instabilities that cause cancer (those that maintain active DNA repair mechanisms and do not contain an abundance of subtle sequence instabilities)?

Gross chromosomal abnormalities:

Human cells normally have 23 pairs of chromosomes that are recognized by their well-defined structure, but tumor cells usually contain major chromosomal abnormalities, making their karyotypes appear quite different from their parental cells. As a rule,

chromosomal abnormalities are not the same from tumor to tumor; each tumor has its own set of anomalies. Certain general anomalies recur however, allowing for their categorization. Types of gross chromosomal abnormalities include chromosomal translocations, gene amplifications (and telomere maintenance), and chromosomal number instabilities.

Chromosomal translocations: Chromosomal translocations can be detected cytogenetically as fusions of different chromosomes or segments of normally non-contiguous single chromosomes. Two general types of chromosome translocations are observed: simple and complex. Simple translocations are characterized by distinctive rearrangements of chromosomal segments in specific neoplastic diseases, commonly leukemias and lymphomas. The specific gene rearrangements found in simple translocations generally lead to the activation (gain of function) of an oncogene through its positioning near a strong promoter or its fusion with another gene[14]. One of the first simple translocations to be discovered, the so called Philadelphia chromosome, is recurrently found in hematopoietic cells of patients with chronic myelogenous leukemia (CML). This chromosome results from a translocation between chromosomes 9 and 22, causing the a gene fusion between a portion of the *bcr* gene (of unknown function) and *c-abl* gene (encoding a protein-tyrosine kinase whose normal substrates are not known). The chimeric *bcr-abl* oncogene product exhibits constitutive and promiscuous kinase activity, activating JAK-STAT signaling proteins and PI-3 kinase. This aberrant activity probably plays a role in the characteristic expansion of well-differentiated granulocytes

that is seen in CML. Additional mutations in these cells (e.g. in p53) may lead to acute leukemia.

Another common simple translocation, seen in tumor cells from patients with Burkitt's lymphoma, is that of the *c-myc* oncogene moving from chromosome 8 to chromosome 14, near an antibody gene enhancer region. This translocation results in the increased expression of the Myc nuclear transcription factor, which contributes to unregulated cell division (c-Myc works with the c-Jun nuclear transcription factor to stimulate transcription of genes encoding proteins involved in cell cycle progression through the G1 phase and the G1/S transition). Other simple translocations found in cancerous immune system progenitor cells (i.e. leukemias and lymphomas) may be generated by the deviant behavior of RAG proteins, which are involved in immunoglobulin and T-cell receptor gene rearrangements.

The other type of chromosomal translocations, complex translocations, are seemingly random and heterogeneous (among individual tumors and among tumor classes) translocations in which large portions of chromosomal material are either deleted (resulting in loss of heterozygosity; LOH) or cut and adjoined to other chromosomes. The en mass movement of this DNA can result in gains and losses of coding sequence, as well as the generation of new gene products[14]. The molecular mechanisms underlying complex chromosomal translocations are not well understood, however, DNA "caretaker genes" (genes that are involved in detecting and repairing DNA damage) may be involved. For example, genes like *ataxia telangiectasia mutated* (*ATM*), the *ATM* related gene (*ATR*), p53 (a downstream effector of *ATM*), and the *BRCA1/2* genes, may help to

safeguard the cell against chromosome translocations. The DNA repair protein, Ku80, has recently been recognized for its ability to suppress certain chromosomal aberrations; mouse cells deficient for this the *Ku80* gene display a marked increase in breakage, complex translocations and aneuploidy[21]. Despite these genetic instabilities however, the mice have only a slightly earlier onset of cancer, indicating the role of multiple mutations and generalized genetic instability in cancer. The p53 "genome guardian" may be involved: p53 increases the susceptibility of *Ku80*$^{-/-}$ mice to early tumorigenesis[21] possibly by increasing the rate of homologous recombination[14].

Gene Amplifications: Another common gross chromosomal abnormality in tumor cells is the localized (re)duplication of DNA to produce as many as 100 copies of a given 0.5 – 10 megabase spanning region. Gene amplifications can be detected cytogenetically as homologously stained regions (HSR) – the duplicated DNA is tandemly organized at a single site on a chromosome, or double minute chromosomes – small, independent chromosome-like structures that pepper a stained chromosomal preparation[12,14]. At the molecular level, the multiple copies of an 'amplicon' (the genomic region that is amplified) often contain a growth-promoting gene(s). Oncogene amplification occurs in a subset of late-stage cancers of many organs (e.g. N-myc amplifications occur in ~ 30% of advanced neuroblastomas[22]), and amplification of genes involved in metabolism or inactivation of drugs represents a common way for cultured cells to acquire resistance to chemotherapeutic agents[14,23]. The mechanisms of gene amplification are largely unknown. Caretaker genes like p53 may play some role: in normal cells, the presence of

an amplicon could signal DNA damage has occurred, thus setting off p53 dependant apoptosis (see below). In the absence of normal p53, cells with nascent amplifications of a locus may survive and accumulate additional amplicons during rounds of cell division, especially if these amplicons select for a hardier tumor cell population.

Telomeric maintenance can be considered an "amplification event" and could be important in oncogenesis. The physical ends of linear chromosomes (telomeres) consist of G-rich tandem arrays of short DNA sequences that are amplified in order to maintain the overall length of the chromosome. Telomeres provide the solution to the "end replication problem" – the inability of DNA polymerases to completely replicate the end of double-stranded DNA (the lagging strand 5'-primer that normally primes replication has no place to bind at chromosome ends leading to successively shortened chromosomes). In germ-line, rapidly dividing somatic cells, and in some cancer cells, a reverse transcriptase called telomerase "counts" the number of repeating G-rich DNA elements (via the DNA binding protein, Rap1p), thereby judging the overall length of the chromosome ends[24]. When the telomeric region appears shortened, heterogeneous TG_{1-3} repeats are inserted and amplified via telomerase action. Shortened telomeres are thought to play a key role in signaling cellular senescence and may be important in protecting chromosome ends from fusions and other rearrangments[25,25]. Maintaining telomeres is therefore thought to be important in some cancers. Telomerase activation has been demonstrated to be a potentially important event during multistage carcinogenesis of various cancers. After in vitro immortalization of certain cell types, telomere lengths are observed to be stabilized and telomerase is active[26]. However, it has been shown that

mouse cells lacking the essential RNA subunit of telomerase could still undergo malignant transformation and form tumors after transformation with viral oncoproteins[27,28]. Apparently, activation of telomerase might not be essential for tumoregsis (in mice). Whether the same applies to humans is a matter of controversy, since mice have much longer telomeres than humans and mouse cells may not require continuous telomerase expression, and may access alternative mechanisms to maintain their telomeres. Nevertheless, telomerase activation and chromosomal end amplification probably contributes to the immortalization of certain tumor cells.

Chromosome number instabilities: Alterations in chromosome number involve losses or gains of whole chromosomes (aneuploidy). Such changes occur in nearly all major tumor types; rare exceptions to this include early "diploid cancers" with MMR deficiency and viral transformed tumor cells[12]. Aneuploidy seems to be due to chromosomal instability (CIN)[12,14,15]. A recent model describes a two stage mechanism whereby carcinogens/mutagens cause cancer via CIN and aneuploidy[12,14]:

(i) Stage one: Carcinogens (by chemically or physically altering one or more of the many proteins of the cell spindle apparatus or the chromosomes) and/or mutations induce CIN and early aneuploidy. Types of mutations that could bring about this effect include genes involved in chromosome condensation, sister-chromatid cohesion, kinetochore structure and function and centrosome/microtubule formation and dynamics, as well as checkpoint genes

that monitor the proper progression of the cell cycle. Some promising candidates include hMAD2, hBUB1 or hBUBR1, *ATM, ATR*, p53, *BRCA1/2*, and Ku80.

(ii) Stage two: Early aneuploidy destabilizes the karyotype and thus initiates an "autocatalytic karyotype evolution", which generates lethal, preneoplastic and eventually neoplastic karyotypes. The source of karyotype instability is the imbalance that aneuploidy imparts on the spindal apparatus (e.g. abnormal ratios of spindle proteins and chromosomal proteins and abnormal structures and numbers of centrosomes). In theory, the imbalanced spindle will cause chromosome nondisjunction and thus autocatalytically regroup the karyotype in a process termed chromosome error propagation.

By this model, cancers are clonal for aneuploidy but not for a *particular* karyotype. Over successive generations of dividing tumor cells (karyotype evolution), certain gene mutations may begin to predominate. Cancers thus may become clonal for several specific mutations, but no "minimal set" of mutations (at least of known cancer-related genes) are adequate to always induce cancer in all cell types. These "downstream" gene mutations would include, by default, many of the altered genes that are commonly found in cancer.

This model predicts that aneuploidy is the *cause* of cancer. In theory, the activation of certain oncogenes or the inactivation of certain tumor suppressor genes

would be buffered by unmutated alleles (dosage compensation) and/or by functional upstream and downstream effectors and redundant capacity. However, when you disrupt the fundamental genomic architecture of the cell (aneuploidy) by carcinogens (or mutagens), you, at once "alter the dosage of thousands of structural and regulatory genes". Slowly, karyotypes evolve, as would a species in nature. With a growth advantage, aneuploidies (along with certain mutations) dominate and pre-cancerous cells begin to slowly out-compete their neighboring normal cells. Cancer ensues. Considering the number and heterogeneity of genes likely to be involved in tumor cell transformation as well as the myriad alternative checkpoints for recognizing and correcting or killing the mutant cell, the chance that subtle sequence instabilities spontaneously leading directly to cancer is arguably low. That aneuploidy (chromosomal instability leading to karyotype instability and evolution leading to cancer – the aneuploidy hypothesis) is in fact the *cause* of cancer is certainly a possible explanation. To prove this hypothesis some investigation is needed. Remaining questions might include: What is the break-point (number and type of extra chromosomes) for neoplastic aneuploidy (aneuploidy resulting in cells of a cancerous phenotype versus nonneoplastic aneuploidies)?; Why don't nonneoplastic aneuploidies, like Down's syndrome with a trisomy or monosomy of chromosome 21, self propagate (by "autocatalytic karyotype evolution") and become neoplastic aneuploidy?; What is the explanation for apparent diploid cancers (tumor cells that are diploid and not aneuploid) like those caused by retroviruses and possibly certain hereditary cancers like FAP or HNPCC? In support of the aneuploidy hypothesis, different chromosomes, regions, and bands have been found to be preferentially involved

in different neoplasia. There is clinical cytogenetic correlation with specific human neoplasia and the specificity of chromosome involvement is significant. A WWW page from the UW Waisman Center Cytogenetics Laboratory[29] reviews these cytogenetic correlates.

There is obviously still quite a lot of work to be done in order to understand what conditions must exist to inevitably cause cancer. Above, I have presented three possible explainations. While the evidence could be presented in the guise of one of these hypotheses, realize that other ways of looking at the data are possible and that causality is very difficult to determine at our current level of understanding. Anyway, the purpose of laying out these theories is not to prove causality, it is to create a clearer picture of how cancer might arise. The next section will review the process termed metastasis. It is this process that makes some tumors so malign once they do come about. Understanding metastasis is important for creating rational approaches to treat, cure, and prevent neoplastic diseases.

Tumor Cell Metastasis

A lot of research has gone into trying to understand the molecular basis of cancer. The discussion above outlines some of the alterations that may cause cancer, providing models (i.e. mutation/oncogene theory or aneuploidization) for the development of tumor cells. However, tumors, as they occur in experimental animals and in humans, that is, the aggregate behavior of groups of tumor cells, are distinctly characteristic in and of

themselves. Tumors arise with great frequency, especially in older animals and humans, but most pose little risk to their host because they are small and localized. These tumors are appropriately termed benign. It appears that the cells making up benign neoplasms closely resemble, and may function like, normal cells.

The surface interaction molecules and physical barriers that hold tissues together are what keep benign tumor cells localized to "appropriate" tissues. Primary among the physical barriers that keep tissues separated is the basal lamina (also known as the basement membrane or extracellular matrix), which underlies layers of epithelial cells as well as surrounds the endothelial cells of blood vessels. The extracellular matrix consists of a supramolecular aggregate of connective tissue proteins including collagens, elastin, glycoproteins (laminin, fibronectin, enactin, nidogen) and glycosaminoglycans (ground substances-heparin sulfate proteoglycans) which interact with one another through covalent and noncovalent bonds to form highly insoluble materials. The other major physical barrier to invasion of secondary tissues by benign tumors is the fibrous capsule often found surrounding them. With these barriers (tissue localization by basement membrane and fibrous capsulation), benign tumors usually only become serious medical problems if they grow large enough to interfere with normal functions of the tissue/organ where they are localized or if they secrete excessive amounts of biologically active effector molecules like hormones or cytokines.

In contrast, malignant tumor cells express some proteins characteristic of its tissue origin, but due to clonal evolution and selective pressures, malignant tumor cells start taking on their own distinct identities. Were a single characteristic to be chosen that most

distinctively differentiates malignancies from any other tissues or growths, it is their ability to spread to sites not directly in contact in them[30]. This dissemination has been given the name metastasis (meta is a Greek preposition meaning "beyond" or "away from" and stasis connotes "position" or "placing"). A metastasis is, in effect, a transplant of a sample of primary tumor to another, secondary site in the body.

This process can be conceptualized as a complex series of sequential steps that involve:

(1) the initial transforming event (i.e. oncogene activation or aneuploidization); (2) proliferation of transformed cells; (3) the ability of cancer cells to avoid destruction by immune mechanisms (see below); (4) nutritional supply to the tumor mass (angiogenesis; see below); (5) metastatic cells from a primary tumor erode (intravasate) through the wall (basement membrane/extracellular matrix and parenchymal cells) of a blood vessel and/or lymph channel; (6) migration of tumor cells away from the primary tumor mass; (7) penetration of cancer cells through the blood vessel wall; (8) either individually or clumped into an embolus, the cells are then carried to some other tissue; (9) arrest of cancer cells in the lumen of small blood vessels or lymphatics; (10) reverse penetration (extravisation) of blood vessels; (11) repetition of the process beginning at step 2 resulting in the formation of secondary tumor metastases.

In order to develop metastases to different organs, tumor cells must possess or develop many highly differentiated characteristics which define the metastatic and invasive phenotype. These characteristics include a larger number of matrix receptors and a greater capacity to bind to and degrade basement membranes[31]. The development of considerable biological diversity by cancer cells has been attributed to their high mutation rate and genetic/karyotype instability which leads to a heterogeneous population of cells (as discussed above). High mutation rates and genomic instabilities enhance the ability of cancer cell populations to develop clones with metastatic capacity. For most tumors, the neoplastic population must expand considerably before invasion and metastasis occur[12]. This expansion may be clonal (at least for aneuploidy and certain key mutations) and mainly limited to a more aggressive subpopulation of neoplastic cells (clonal dominance)[12].

Metastasis is an inefficient process; only a few cancer cells are able to form tumors after being released into the circulation. Some cells that are shed from primary tumors are probably destroyed by mechanical and immunological mediators (see below) before they can escape out of the microcirculation into tissue. However, the majority of traveling cancner cells survive their journey to secondary sites[32]. The major metastatic control in the body is likely the control of tumor growth at secondary sites: "Only a small fraction of extravasated cells begin to divide to form micrometastases, and only a very small fraction of these micrometastases continue to grow to form tumors."[33] Cells that establish themselves at the secondary site, stimulate the growth of new blood vessels (angiogenesis or neovascularization) to supply the nutrient demands of the new tumor.

Angiogenesis is thought to take place as a series of sequential steps[6,34,9], recruiting the growth of new vessels first at the outer rim and later within the interstitial mass of the growing tumor[6]:

> (1) an early morphological step is the local degradation of the basement membrane of the parent vessel; (2) endothelial cells psuedopod, protrude through these holes in the basement membrane; (3) endothelial cells migrate toward an angiogenic stimulus; (4) the cells align in a bipolar configuration to form a spout; (5) DNA synthesis and mitosis of endothelial cells begin; (6) a lumen is then formed by each endothelial cell in the sprout and individual sprouts soon join with each other to form a loop through which blood begins to flow[6].

It is apparent that the two processes of metastasis and angiogenesis share some common characteristics – in particular, the attachment to, local degradation of, and subsequent invasion through the basement membrane of neighboring cellular compartments. Often the similar molecular interactions mediate these events in both metastasis and angiogenesis.

The process of metastasis (and the subsequent establishment of secondary tumors) is responsible for a large proportion of the mortality caused by cancer. For instance, the most frequent cause of death in patients with breast cancer can be attributed to metastasis of tumor cells and their growth at distal regions rather than the *in situ* effects of the malignancy at the primary site[35].

A Closer Look at Tumor Metastasis and Invasion

The overall process of metastasis/invasion involves the adhesion of tumor cells to epithelial and endothelial cells (depending on the direction of migration) and to basement membrane matrices, partial proteolytic digestion and "reorganization" of basement membrane layers, followed by cell penetration (migration) through the disrupted membranes[36]. Below, I will review each step in this sequence.

Tumor Cell Adhesion. The heamatogenous phase of cancer metastasis facilitates the transport of metastatic cells within the blood and incorporates a sequence of interactions between circulating intravascular cancer cells and the endothelium of blood vessels at the sites of tumor arrest often at the first physically tight junctions encountered, like pulmonary capillary beds (the lungs are the most common site of metastasis[37]; only the intestines send their blood to the liver first). In part, cancer cells lodge in capillaries because these cells tend to be relatively large. Also, some cancers produce procoagulant factors that cause platelets to aggregate around them. These aggregates effectively make metastatic tumor cell emboli even larger and stickier. (Platelets may also aid in the survival of tumor cells in the blood by supplying growth factors and possibly by sequestering tumor cells from immune surveillance).

Once lodged in a secondary site, contacts with cells and/or basement membrane are established. Initial interactions involve mechanical contact and transient cell-to-cell and/or cell-to-matrix adhesion[38]. Cellular adhesion involves multiple ligands and cell-

adhesion molecules (CAMs). CAMs, a diverse group of integral membrane proteins, fall into six main classes: the cadherins, Ig-superfamily CAMs, selectins, mucins, integrins and non-integrin surface receptors, of which cadherins, selectins, integrins and non-integrin receptors are especially relevant to cancer.

CAMs of the cadherin family have an inhibitory role in the process of metastasis. Cadherins are Ca^{2+}-dependent CAMs that mediate cell-cell binding. Three major types have been identified in mammals: E-cadherin, P-cadherin, and N-cadherin. E-cadherin is a 120 kDa molecule with an intracellular, a plasma membrane spanning, and an extracellular domain. The intracellular domain is connected with a group of proteins called catenins which function as a link to the actin cytoskeleton (referenced above in connection with the *APC* tumor suppressor gene). These connections not only reinforce the cell adhesive capacity of E-cadherin but also suggest that E-cadherin is integrated in signal response circuits[39,40]. Cadherin proteins are (and catenin signaling is) diminished in some cancer cells. This is thought to loosen intercellular adhesions and concomitantly provide appropriate signaling (or loss of signaling), freeing tumor cells from their local restraints and encouraging them to invade and migrate.

Adhesion to extracellular matrix and endothelial cells (of metastatic secondary sites), on the other hand, allows cells to survive, proliferate and invade, creating metastases. Members of the immunoglobulin (Ig) superfamily that have been studied for their putative role in invasion-related homotypic cell-cell adhesion are N-CAM, DCC (deleted in colon cancer), and CEA. Endothelial selectins and their ligands on neoplastic cells are also involved in initial adhesion events[37]. Selectins (and integrins; see below)

are understood to be "trafficking" molecules, telling lymphocytes and leukocytes where to go in the body. By analogy, selectins and their ligands (addressins) could play a role in directing cancer cells to specific distant sites where they help to initiate primary contacts. (Other surface molecules and soluble factors like chemokines, growth-promoting factors and hormones likely play a role here too[36]). Indeed, some types of cancer show a striking preference for organs other than those that receive their venus blood (e.g. the tendency of metastic prostate cancer to move into the bones). Initial contact is thought to initiate a sequence of activation pathways that involve cytokines, growth factors, bioactive lipids, and reactive oxygen species produced either by the cancer cell or by the endothelium. These molecules elicit expression of active integrin adhesion molecules on the surface of the cancer cells (and the endothelium).

Attachment to specific glycoproteins of the extracellular matrix such as fibronectin, collagen, and laminin, is mediated through tumor cell receptors of the integrin and non-integrin variety. Integrins are a large class of αβ heterodimer receptors for adhesion molecules with broad specificity and low-to-moderate binding affinity. Numerous different α and β subunits have been identified which noncovalently associate to produce more than 20 integrin types. Integrins serve as integral cell membrane receptors that form focal adhesion contacts with various extracellular matrix ligands (i.e. fibronectin, laminin, vitronectin, the collagens, thrombospondin, entactin, fibrinogen[6]. The process of binding ligands to their integrins involves "outside-in" signaling initiated by receptor clustering accompanied by conformational changes in the α/β chains culminating in affinity modulation for the ligand. During this process, adhesion plaques

form at the cytoplasmic face of the cell membrane that serve as focal points for recruitment of proteins (e.g. talin, venculin, paxillin) to provide cascade interfaces for actin, G-proteins, tyrosine kinases, and general transcription factors[6]. Bidirectional crosstalk occurs at the cell-to-extracellular matrix or cell-to-cell interface, transducing signals, inside-to-outside and outside-to-inside, that further activate surface receptors and also modify various intracellular target proteins involved in survival, growth, morphology, movement and differentiation of cells.

During cellular transformation and cancer growth, integrin structure and integrin expression patterns are modified: some integrins are either overexpressed or no longer expressed, whereas others become phosphorylated, affecting their cytoskeletal and extracellular binding properties[6]. Changes in integrin cell-surface distribution can affect ligand binding affinity, and correlate with a concomitant dis- or re-organization of the basement membrane structure itself. The loss or altered patterns of the extracellular matrix-binding integrins appears to be one of the "abnormalities" underpinning tumor progression[6].

The 67 kDa LBP is a major nonintegrin cell surface receptor that mediates high-affinity interactions between cells and the basement membrane molecule laminin-1. This receptor and its major ligand binding domain in laminin-1, peptide 11, is discussed in Chapter 2 of this thesis.

Proteolytic Digestion and "Reorganization" of Basement Membrane. During the process of tumor cell adhesion to other cells and to the basement membrane, the outside-to-in signaling mentioned above induces the production of numerous proteinases and

other factors (LBP may play a role in "reorganizing" basement membrane laminin; see above) that degrade or reorganize the basement membrane and facilitate subsequent tumor cell invasion. Although considerable information is available on the digestion of isolated connective tissue proteins by tumor enzymes, less is known about the mechanisms by which tumor cells digest highly complex mixtures of these proteins. The degree to which the extracellular matrix has to be broken down in order to accommodate invasion by tumor cells is unknown and thus, it is a matter of speculation as to the amount of proteolytic activity required for the process.

All four categories of proteinases (serine, cysteine, aspartic proteinases, and metalloproteinases) have been implicated in the invasive process. The proteinase theory of cancer invasion is supported by the finding that normally invasive cells, such as activated macrophages and trophoblast cells produce proteinases. Tumor enzymes that hydrolyze the extracellular matrix components (glycosidase, hyaluronidase, heparanase) have also been described[37].

An often noted feature of the tumor fibrinolytic system is the catalytic nature of plasminogen activator (PA) which, when in the presence of high concentrations of substrate plasminogen in the serum or tissues, yields high levels of local plasmin (a broad spectrum serine proteinase) at the leading edge of cancer invasion. Cell surface receptors for urokinase-type PA and plasmin have been described[41]. Plasminogen activator inhibitors are also produced in tissues and counteract uPA activity, but may be less efficient against receptor bound uPA[42]. uPA is able to activate matrix metalloproteinases (a family of closely related, metal-dependent endopeptidases which, once activated,

degrade a variety of extracellular matrix components) which may represent a major function of the enzyme[68].

Tumor Cell Migration. As metastatic tumor cells mediate the diruption of the basement membrane, they migrate through it prior to establishing itself in its new its local. Migration is thought to involve a variety of motility factors. Many motility factors for cancer cells and non-malignant cells were described first as growth factors. A motility factor converts a cell from a static to a motile status, a transition that is characterized by the appearance of membrane ruffling, lamellae, and pseudopodia. Several motility factors have been described for cancer cells including: (1) autocrine motility factor (AMF) which stimulates chemokinesis and chemotaxis of metastatic melanoma cells in an autocrine fashion; an AMF receptor which becomes phosphorylated has been identified; (2) Autotaxin; (3) Scatter factor/hepatocyte growth factor (ligands for the c-met oncogene product, a tyrosine kinase receptor family member) may also act as a positive or negative growth factor; (4) TGF-a and EGF binding to the same receptor (EGFR); (5) Insulin like growth factors stimulate chemokinesis and chemotaxis.

Appendix 3 - Midway Summary

To recapitulate (up to this point): A cancer arises because DNA in a cell is altered by a genetic, chemical, and/or physical influence. This leads to the production of altered or unregulated genetic products that cause uncontrolled multiplication of the involved

cell. If not stopped, these cells can clonally develop malignant characteristics: they can break from their local constraints and spread to distant sites in a process called metastasis. But even though it loses the battle to prevent cancer from forming in the first place, the body is not without defense. At first, it attempts to keep the growth contained. Few cancer cells can effectively overcome an organ's resistance and succeed in implanting themselves well enough to achieve independent growth[36]. There are a number of restraints that keep tumors in check, the mechanisms of which converge in the process of induced cell killing. These restraints represent our natural defenses against cancer and tumor progression.

Our Natural Defenses Against Cancer

The progression of a cell to a cancerous phenotype is usually checked at various points. As described above, genomic mutations early in tumor development can be corrected by numerous DNA repair mechanisms. If mutations go uncorrected, other intracellular effectors recognize the altered DNA and can initiate programmed cell death (apoptosis). Apoptosis is a Greek word meaning dropping or falling off. Cells undergoing apoptosis: shrink; have their mitochondria break down with the release of cytochrome c; develop bubble-like blebs on their surface; exhibit specific condensation and degradation of DNA (leaving 180 bp fragments indicative of nucleosomal spacing); break into small, membrane-wrapped, fragments (called apoptotic bodies) which are generally engulfed by nearby phagocytic cells like macrophages.

Importantly, the intracellular constituents are not released into the extracellular milieu where they might have deleterious effects (via inflammation) on neighboring cells. In contrast, cells that die in response to tissue damage (necrosis) exhibit very different morphological changes. They (and their organelles like mitochondria) swell because the ability of the plasma membrane to control the passage of ions and water is disrupted. Highly noxious intracellular contents leak out, damaging surrounding cells, often leading to inflammation.

Many cancer cells characteristically loose the intrinsic ability to initiate apoptosis on their own. In these cases, a second line of defense is afforded by the immune system which, using a variety of strategies, circumvents intracellular apoptosis signaling pathways to induce apoptosis. Whether cells commit suicide due to internal cues or are "murdered" by killing signals from other cells, recent studies suggest that both suicide and murder pathways converge into a common molecular pathway.

Internal Mechanisms Inducing Apoptosis (Suicide)

One possible mechanism to suppress tumor progression involves turning on internal cues that, in turn, activate an apoptosis pathways. A major player in this pathway is the p53 tumor suppressor gene. p53 (encoding a 53 kDa nuclear phosphoprotein) has emerged as the most commonly altered gene in cancer: approximately 50% of cancers contain a mutation in p53. p53 acts as a monitor of genome integrity. If the genome is damaged, p53 is stabilized and accumulates in cells. The precise signaling pathway that

senses DNA damage and recruits p53 is unclear, however genes like ATM and ATR (mentioned above) and MDM2 may be involved. Apparently, ATM kinase, which phosphorylates and activates the checkpoint kinase CDS1/CHK2, NBS1 (a protein involved in Nijmegen breakage syndrom; NBS), and BRCA1, is activated upon recognition of DNA damage[43,44,45]. These proteins in turn activate p53 (via phosphorylation) and the p53-dependant transcription of various genes (p53 activates the nuclear transcription factor, NF-κB, which may induce expression of various apotosis effectors[46]). Physical interaction between p53 and MDM2 enhances p53 degradation through the ubiquitination pathway and conceals the activation domain of p53 to regulate transcription. MDM2 restraints are released upon p53 phosphorylation[47]. The targeting of p53 by ATM and CHK2 sets in motion other pathways that result in a prolonged (possibly irreversible) G1 phase arrest. For instanced, activated and accumulated p53 activates transcription of Pic1 ($p21^{cip1}$), a cyclin dependant kinase inhibitor gene which is also partly responsible G1 arrest. Pic1 maintains pRb in its active (dephosphorylated) state, thereby blocking the action of E2F and cell cycle progression[48]. Cell cycle arrest allows time for DNA repair and/or p53 helps to maintain basal cell cycle progression if no damage is detected. Other p53-dependant responses such as transcription of repair genes or genes encoding 14-3-3 proteins may also facilitate DNA repair and prolong G2 arrest[47]. Activation of p53 can lead to the induction of pro-apoptosis genes like *Bax* and modulators of the anti-apoptosis protein Bcl-2. Inducing BAX protein and inactivating Bcl-2 may play a role in the release of cytochrome c from mitochondria, which in turn

may activate the caspase cascade and apoptosis (see below). However, the signaling pathway that relays recognition of DNA damage to apoptosis is not well understood.

External Mechanisms Inducing Apoptosis (Murder)

Cells that escape internal control mechanisms and maintain viability are policed by the immune system. Cells of the immune system recognizes tumor cells as non-self (or "altered self") because they display seemingly foreign "tumor antigens" on their surface. Both early innate (e.g. core pattern recognition and early cellular responders like natural killer (NK) and NK-T cells) as well as downstream acquired responses (T-cell mediated and hummoral) play critical roles in eliminating these cells.

There are at least five ways that the immune system kills tumor cells: 1.) Cytotoxic T-cells (CTLs), γδ T-cells and possibly NK-T cells recognize and bind target cells via an immunological synapse Synapse formation stimulates golgi reorientation and Ca^{2+}-dependant release of perforin, granzymes and lyticases. Perforin oligimerizes in the plasma membrane of the target cell, forming a pore through which mediators (like granzymes which include "initiator" caspases (see below) and other proteases, nucleases, and lyticases) can enter and induce apoptosis. 2.) Fas-ligand on CTLs binds to Fas receptor on target cells and initiates Fas-mediated apoptosis (see below). 3.) In a two-receptor model of killing, an altered self (target) cell may have reduced surface expression of MHCI which normally would bind Ly49 on NK and NK-T cells. MHCI interaction with Ly49 sends an anti-apoptosis signal to the effector cell. Without this

inhibitory signal, GlyPr (on the target cell) binds NKR-P1 on the effector cell, inducing an unchecked death signal. In this case, apoptosis is mediated via perforin and grazyme release. 4.) Antibody dependant cell-mediated cytotoxicity (ADCC) again induces perforin and granzyme release, here via receptors recognizing antibodies bound to the target cell surface. 5.) Sometimes target cells (including tumor cells) will upregulate surface expression of the "stress ligand", MICA/B. γδ-T cells, NK cells and CTLs recognize this ligand via the NKG2D receptor and relay a signal through the adaptor protein, DAP10, to phosphatidyl inositol 3 phosphate kinase (PI3K) which, in turn, cab signal release of perforin and granzymes to induce apoptosis in the target cell[77].

Caspases – the Common Pathway to Induced Cell Death

Death signals, whether they originate from activation of death receptors, granzymes, DNA damage, or metabolic or cell cycle perturbations, all converge at the activation of the caspase cascade. Caspases (also known as interleukin-1β-converting enzymes (ICEs)) are 30-50 kDa cysteine proteases with aspartate specificity. They are normally present in cells as catalytically inactive proenzymes and are proteolytically processed and activated during the induction of apoptosis[49,50]. (Cleavage results in the dimerization of the large and small subunits of the proenzyme; a dimer of dimers forms the active caspase.) The caspase proteolytic cascade is activated via receptor signaling (Fas-associated protein with death domain (FADD), through the death effector domain (DED), activates "initiator" caspase 8) or through interaction of cofactors with other

"initiator" caspases. For instance, procaspase-9 activation involves the formation of a complex with the cofactor APAF-1 through CARD (caspase recruitment domain)[54]. Activation of caspase-9 also requires cytochrome c and deoxyadenosine triphosphate, indicating that caspase activation may require multiple cofactors. Very little is known about the interaction between procaspases and their cofactors. Certain caspase inhibitors (e.g. FADD-like ICE inhibitory proteins (FLIPS), apoptosis repressor with caspase recruitment domain (ARC), inhibitors of apoptosis (IAPs), and viral inhibitors like CrmA and p35) have been identified[54]. Compartmentalization of caspases and their cofactors is likely to be another way of regulating caspase activation (e.g. suquestration of chytochrome-c in the mitochondria).

Once activated, "initiator" caspases like caspase-8 and caspase-9 initiate a cascade which activates "effector" caspases, resulting in cellular disassembly and apoptosis[54]. The direct effects of these downstream caspases include reorganization of the cytoskeleton (e.g. gelsolin, FAK, p21-activated kinase), shut down of DNA replication and repair (e.g. replication factor C, DNA-cytoskeletal protein kinases (DNA-PKcs), U1-70K (mRNA splicing)), specific DNA cleavage (caspase-3-mediated cleavage of ICAD (inhibitor of caspase-activated deoxyribonuclease (CAD), thereby activating CAD[51]), disruption of nuclear structure (e.g. nuclear lamina), marking of apoptotic cells (e.g. loss of plasma membrane phospholipid asymmetry exposes phosphatidylserine externally, which is specifically recognized by active phagocytes[52]), and the controlled release of apoptotic bodies from the cell.

Cancer – New Promises in Therapy

Cancer is not a single disease. It encompasses a large group of disorders that share certain characteristics. At least three common features of cancer can be gleaned from the information presented above: 1.) cancerous tissue continuously grows as cancer cells proliferate autonomously; 2.) as this growth continues, tumors can invade into surrounding normal tissue; 3.) tumors can metastasize. Most current cancer therapies are aimed at stopping three features of cancer.

Surgery

Probably the earliest established therapy for cancer, and still a widely used approach, is surgery. Surgical therapy for cancer offers several advantages: It is relatively quick and effective; it provides the opportunity to confirm that the local tumor has been fully excised; and surgical biopsies are important for tumor staging. Tumor grading and staging (see below).

There are however several shortcomings to surgical intervention: There is no guarantee that surgery can excise all of the primary tumor (including all microscopic extensions). Often the surgeon is forced to cut out large portions of healthy tissue with the diseased tissue leaving severe damage and sometimes disfigurement. The procedure

itself (usually a major operation requiring anesthesia) can be traumatizing. And importantly, surgery cannot effectively treat cancer that has extensively metastasized.

A variety of new surgical techniques are being developed to fortify this type of cancer treatment. For instance, cryosurgery (also called cryotherapy), the use of extreme cold to destroy cancer cells, is currently being explored as a treatment option. Cryosurgery has been used to effectively treat retinoblastoma, early-stage skin cancers (both basal cell and squamous cell carcinomas), precancerous skin growths known as actinic keratosis and the precancerous condition cervical intraepithelial neoplasia (abnormal cell changes in the cervix that can develop into cervical cancer). New applications also include the treatment of some internal tumors, where liquid nitrogen is circulated through an instrument called a cryoprobe, which is then placed in contact with the tumor. To guide the cryoprobe and to monitor the freezing of the cells, the physician uses ultrasound (see below). Cryosurgery often involves a cycle of treatments in which the tumor is frozen, allowed to thaw, and then refrozen[53].

Laser microdissection is another surgical approach that may be useful in obtaining pure cancer cell samples from a biopsy in order to more effectively characterize a tumor[54].

Radiation Therapy

Radiation therapy (or radiotherapy) refers to the treatment of cancer with radiation. Linear accelerators, cobalt units or orthovoltage x-ray machines produce

ionizing radiation, such as x-rays, gamma rays, and electron beams, which destroy cancer cells[55]. Radiation treatments are thought to act either by inflicting genetic damage sufficient to kill cells directly or by inducing apoptosis (see below). Because healthy tissues can recover from radiation exposure more readily than cancerous cells, radiation therapy can preserve the anatomical structures that surround a cancerous growth, thus curing the cancer without sacrificing the patient's ability to function. Radiation therapy is also advantageous in that it can destroy microscopic primary tumor extensions and it can be safer for, and better tolerated by frailer patients since radiation therapy routinely receive five to eight weeks of daily out-patient treatments. However, radiation therapy can sometimes fail to eliminate all of a tumor and widespread metastases are difficult to treat with radiation alone.

There are several new advances in radiation-based cancer treatment[59]. Conformal therapy, the surgical directing of X-ray beams, based on computer tomographic scans and magnetic resonance imaging of the patient, maximizes the dose of radiation absorbed by the tumor while minimizing the exposure of the surrounding tissue. Brachytherapy (implants of radioactive sources directly into or in the proximity of the tumor) and ingestion (radioactive iodine can be taken to treat cancers of the thyroid) can increase radiotherapy tumor selectivity. Proton- and neutron-based therapies are also available. Protons target tumor-bearing sites better then do X-rays, and neutrons seem to have more potency against some cancers. Proton beam therapy can selectively be used for uncommon tumors around the eye, some spinal tumors and advanced prostate cancer. Neutron beam treatment can be used to treat advanced prostate, and salivary glands with

excellent results. A technique called neutron capture therapy combines brachytherapy with neutron therapy. A tumor-seeking compound containing boron-10, a non-radioactive isotope, or gadolinium (a magnetic resonance imaging contrast agent; see below) is introduced into the brain and given time to accumulate in the tumor. The tumor is then exposed to a beam of neutrons, which are "captured" or absorbed by the metal. Capturing neutrons causes the boron or gadolinium nuclei to break down, resulting in the emission of high energy intermediates (alpha radiation and/or high energy nuclei). Intermediates are reactive but short in range, which means they destroy the malignant cells where they are localized without hurting the adjacent healthy cells[56]. Research also is underway using radiolabeled antibodies that locate tumors and kill them. This technique called radioimmunotherapy can also be used diagnostically (see below). Several factors including heat and several types of drugs increase the sensitivity of tumors to radiation (radiosensatizers) while some drugs are used to protect normal tissue near the site of treatment (radiation protectors)[59].

Chemotherapy

Chemotherapy is the use of drugs to kill tumor cells. There are several classes of chemotherapeutic drugs: Antimetabolites usually act as false substrates in a cell's metabolic or biosynthetic pathways. A good example is methotrexate, a chemical analogue of folic acid, which interferes with purine biosythesis and prevents cells from dividing. Other antimetabolites include flourouracil and gemcitabine[59]. Topoisomerase

inhibitors inhibit proper strand separation during replication and leave DNA strand breaks, inducing cell death. Examples include doxorubicin and CPT-11. Alkylating agents like cyclophosphamide and chlorambucil covalently bind to DNA, producing nicks and breaks, which may induce apoptosis. Plant alkoids can prevent proper cell division by binding to tubulin. Examples of plant alkoids include vinblastine, vinorelbine, and taxane such as pacliaxel (Taxol) or docetaxel (Taxotere)[59, 57]. Some chemotherapeutics are most effective when used in combination with other agents (combination therapy), with other types of therapy, like hormone-based therapy (combined modality therapy) or with radiotherapy (concomitant chemoradiotherapy).

It is commonly held that most chemotherapeutic drugs (as well as radiation) kill cancer cells by causing them to self-destruct through the p53-directed apoptotic pathway (described above). Although many blood cell tumors seem to follow this mechanism, a definitive answer about how p53 influences the response of solid tumors to therapies is not as obvious. There appear to be numerous interdependent pathways that detect genomic damage and instability and then relay this information to the cells' apoptosis machinery to induce cell death. Moreover, the pathway that is elicited (and the cells' response to therapy) may vary depending on the drug regimen and/or radiation protocols used[58] as well as the specific type of cancer (lineage and pheotype).

The extensive bioavailablitiy of most chemotherapeutics often leads to unwanted side-effects by killing not only rapidly dividing tumor cells, but also other rapidly dividing cells like digestive tract epithelial cells (causing diarrhea, nausia, and vomiting) and bone marrow cells (causing anemia). Also, some tumors are able to develop

resistance (or are innately resistant to) certain treatments, limiting chemotherapy's effectiveness. Possible mechanism leading to resistance include "the drug efflux pump" (a protein, P-glycoprotein, that resides in the cell membrane and is thought to pump anticancer drugs out of cancer cells along with other substances perceived as toxic. Identification of the human gene that encodes the drug pump has allowed researchers to consider possibilities for circumventing the pump and enhancing drug therapy) and general tumor instability (including amplifications; see above).

Endocrine Therapy

Another kind of therapy, available to a subset of cancers, sidesteps many of the difficulties associated with chemotherapy. Hormone-based therapies manipulate the endocrine system to kill tumor cells. The breast and prostate are glands regulated by sex hormones; malignancies that arise in these tissues may respond to therapies based on these signaling molecules. For instance, antiestrogens or "androgen blockade" can keep some breast cancers and prostate cancers, respectively, in check. This approach sometimes proves ineffective when the tumors are hormone-independent (like estrogen receptor negative tumor cells)[59].

Immuno-Based Therapies

In some promising new therapies, the body's internal weapons are trained on cancerous growth, helping the immune system to do its job. General types of immuno-therapies include[59]:

Nonspecific Immunotherapies. Examples range from the administration of Coley's (the early oncologist and surgeon) toxins (inducing bacterial LPS-related tumor necrosis) to the delivery of cytokines like INF, TNF, IL-2, IL-12 and GM-CSF (each which mediates immune responsiveness to tumors).

Passive Immunization and Antibody-Based Therapies. Another class of novel cancer treatments that induce tumor cell destruction (and possibly solid tumor regression) are immuno-based therapies. Acting alone, antibodies can bind to tumor antigens on the surface of the tumor cell, helping to target it for antibody dependant cell-mediated cytotoxicity (ADCC) induced cell death. Antibodies can similarly target and attack the blood vessels feeding a tumor or the connective tissue supporting it. And antibodies can neutralize or block the action of growth factors and other tumor enhancers. Finally, antibodies can be used to deliver an array of damaging compounds to tumor sites. Examples include antibody-mediated deliver of radioactive isotopes or toxins like ricin (immunoradiotherapies and immunotoxins, respectively) and the delivery of therapeutics.

Developments in the production of monoclonal antibodies and the humanization of antibodies should help to make antibody-based therapies more effective.

Active Immunotherapy and Vaccine-Based Therapies[60]. All sorts of antigens have been found to induce immediate (and sometimes protective) immune responses against specific tumor cells. Whole cancer cells, antigenic proteins and small peptides, antigen presenting cells (like dendretic cells) pulsed with tumor antigen[61], nucleic acids (DNA vaccines), and viral and bacterial delivered genes encoding tumor antigens are a few ways in which specific antigens can be delivered and can potentially elicit an immune response against the tumor. Other treatments potentiate cell-mediated (CD8+-CTL, CD4+-CTL, NK and/or NKT cell) cytotoxicity using tumor-antigen based vaccines and other molecules (e.g. α-galactocerimide induction of NKT cell tumor killing[7] or HSP- and ganglioside-induced immune responses) that can induce innate and acquired immune responses against tumor cells[6].

Adoptive Immunotherapy[62]. Yet another approach to immuno-based therapy is under investigation. Adoptive immunotherapy involves stimulating (priming) T cells by exposing them to tumor cells or antigens in the laboratory (*ex vivo*), expanding these tumor-targeted T cells by adding other mediators (costimulatory molecules and cytokines), and then injecting the expanded populations back into the patient. Adoptive immuno-therapies may have its greatest value in treating patients who are (or will become) immunocompromised. For instance, before leukemia patients can be treated *ex*

vivo for threatening viral infections and for cancer after receiving chemotherapy, radiotherapy and bone marrow transplants. Adaptive immuno-therapies also seem to be efficacious in the treatment of melanoma, lung and hepatic cancers[66].

New Molecular Therapies

The promise of new molecular-based medicines founded on a genetic understanding of cancer is in the process of being realized, the clinical results of which should reflect our newest cancer research imperatives. New molecular therapies attack each level of tumorigenesis: cell-cycle regulation, signal transduction, apoptosis, telomere biology, angiogenesis and metastasis. Examples abound: The *ras* oncogene (mutated in 20-30% of all human tumors) activation pathway is targeted by a variety of farnesyl transferase inhibitors and clinical trials are under way with a small-molecule inhibitor of Bcr-abl (mentioned above)[59]. Protein kinase inhibitors may effectively target cyclin dependant kinases (e.g. the drug Flavopiridol), protein receptor kinases (drug: CP358,744), and protein kinase C (drug: UCN-01) among others. New estrogen receptor antagonists include Raloxifene and GW 5638. Advanced apoptosis-based therapies could activate caspases directly, either through activation of death receptor complexes or through caspase cofactors (inhibitors and enhancers)[53]. The upregulation of the anti-apoptosis gene *Bcl2* is proving to be effectively inhibited by G3139, an antisense gene therapy. Therapeutic antibodies to EGF receptor (Ab: C225), VEGF (Ab: Anti-VEGF), KDR (Ab: IMC-1C11), and integrins (Ab: Vitaxin) represent another new line of

mechanism-based immunotherapies. Antisense therapies targeting Ras, Raf and protein kinase C, are being explored. Recently, inhibition of telomerase activity was shown to limit the growth of cancer cells, thus validating this enzyme as an important target for the development of anti-neoplastic therapies[63].

Small molecule-based anti-cancer drugs are another type of modern cancer therapy. These are discussed in Chapter 2 of this thesis.

Cancer as a disease in the human population is becoming an ever larger health problem as the population ages. Although the medicines and interventions which are used as treatments are clearly limited, they are certainly advancing in design and efficacy. Most current cancer therapies are aimed at stopping uncontrolled growth, tissue invasion and metastasis, three features common to many cancers. Primary tumors can often be surgically excised. At the same time, the surgeon might also remove enough of the axillary draining lymph nodes for microscopic analysis – to determine the metastatic potential of the tumor. Chemotherapy, often the next step in treatment after resection, is used to help "sterilize" the blood and lymph streams and to chemically interfere with the growth of (or kill outright) any undetectable metastases. X-ray therapy can be used to ensure that there are no other areas of malignant cells in the remaining tissue[64]. The past twenty years has lead to a tremendous increase in our understanding of the molecular mechanisms and pathophysiology of human cancer. A number of these mechanisms have already been exploited, leading to promising new molecular-based anticancer drugs and advanced interventions, whatever the stage of treatment.

Cancer Prevention and Detection

Strategies aimed at keeping cancer from occurring, progressing or recurring remain invaluable weapons. Research over the last decade, on the mechanisms involved in cell growth regulation and tumor spread, has provided us better tools for screening, diagnosing, and monitoring the progression of tumors and the efficacy of treatment.

Prevention and early detection are two of the most important and effective strategies for saving lives lost from cancer, diminishing suffering due to cancer, and eliminating cancer as a major health problem. Prevention (sometimes called primary prevention by public health professionals) includes measures that stop cancer from developing. Prevention is still an idea with plenty of untapped potential. Approximately one third of fatal cancers can be blamed primarily on smoking, and an equal number on lifestyle, especially dietary practices and lack of exercise. Prevention is far more efficient as well as economical when compared to treatment. By some estimates, if risky behaviors were curtailed, upward of 200,000 lives could be saved from cancer annually even if no new treatments were available[65].

Basic Early Detection

Early detection (sometimes called secondary prevention or screening) includes examinations and tests intended to find cancer as early as possible (ideally, before it has spread) when it can be treated most effectively and with the fewest possible side effects.

Early detection examinations and tests can help save lives and reduce suffering from cancers of the breast, colon, rectum, cervix, prostate, testis, oral cavity, and skin. Some of these cancers can be found early by self examinations (such as breast self-examination) or physical examinations by a health professional (such as examinations of the breast, thyroid gland, skin, colon and rectum, testicles, and prostate).

X-ray and laboratory tests are also proving to be quite useful in early cancer detection. The Pap test allows early detection of cancer of the uterine cervix (almost always caused by the human papilloma virus)[66]. In this procedure, a doctor uses a small brush or wooden scraper to remove a sample of cells from the cervix and upper vagina. The cells are checked for major abnormalities indicative of cervical cancer. Since the 1930s, early detection using the Pap test has helped lower the death rate from cervical cancer more than 75 percent.

Breast cancer can sometimes be detected in its early stages using a mammogram, an X-ray (or sometimes ultrasound) of the breast. Mammography is most beneficial for women as they age and undergo menopause. Mammography is a screening tool that can detect the possible presence of an abnormal tissue mass. By itself, it is not accurate enough to provide definitive proof for either the presence or absence of breast cancer. If a mammogram indicates the presence of an abnormality, further tests must be done to determine whether breast cancer actually is present and, if found, to what degree the disease has progressed.

Internal Markers for Early Detection of Cancer

Recently, several internal cancer markers have been identified and various commercial assays developed. The utility of any marker depends predominantly on its sensitivity and specificity. To be useful, the concentration of any cancer marker must reflect the changes in tumor burden and allow for the prediction of response to therapy and the course of disease. A single marker or a panel of several individual markers is often used to provide a clear picture of cancer progression.

Common internal cancer markers include[67]: Prostate-Specific Antigen (PSA; used in detecting and following prostate cancer); Fecal Occult Blood (FOB test; used in colorectal cancer screening and diagnosis); Cancer Associated Antigen-Ovarian Tumor (CA-125; A cell surface glycoprotein widely used as a marker for ovarian epithelial cancer); Cancer Associated Antigen-Breast Tumor (CA 15-3; a serum glycoprotein, used in prognosting patients with advanced breast cancer); Cancer Associated Antigen-Gastrointestinal (CA 19-9; a serum glycoprotein that is elevated in individuals with colorectal, gastric and pancreatic cancer); Cancer Associated Antigen-TAG-72 (CA 72-4; a complex glycoprotein that is elevated in serum of individuals with breast, pancreatic, ovarian, colon, and stomach cancer); A-Fetoprotein-Human Cord Serum (hAFP; a major fetal serum glycoprotein indicative of hepatoma, yolk sac and gastric tumors); and Tissue Polypeptide-Human Breast Epithelium Carcinoma Cell Line (hTPA; used in diagnosing and prognosing patients with gastrointestinal, gynecological, lung and bladder cancers).

Detection of high levels of active telomerase in certain tissues may be a sing of certain cancers and telomerase-based tests may aid in making prognoses[68].

Genetic Testing

For people with a possible hereditary disposition to a certain type of cancer, genetic testing can determine whether that person carries some of the genetic alterations indicative of the disease. Predictive gene tests for sporadic cancer are still primarily a research tool, difficult to execute and available only through research programs. However, the field of gene testing is evolving rapidly, with new genes being discovered almost daily and innovations in testing arriving almost as quickly. Tests for a few rare cancers are already in clinical use[69,70]. Examples include: predictive tests for retinoblastoma and Wilms' tumor; a gene test available for persons in the rare cancer-prone Li-Fraumeni families; a test for *APC*, the gene involved familial adenomatous polyposis (FAP); genetic testing for the genes involved in HHPCC; mutations in the *BRCA1*, and possibly *BRAC2*, genes predisposes a person to hereditary breast cancer and ovarian cancer and can be screened for in a blood test.

There are a variety of new nucleic-acid based methods for cancer detection that appear to be promising[74]. Detection of microsatellite instability (MIN) and loss of heterozygosity (LOH) is common in certain hereditary cancers (e.g. HNPCC; see above) and may appear in various sporadic cancers. Interestingly, it appears that high-frequency microsatellite instability in colorectal cancer is independently predictive of a relatively *favorable* outcome and, in addition, correlates with reduced likelihood of metastases[71].

Recently, mitochondrial DNA (mtDNA) mutations were detected in human bladder, head and neck, and lung primary tumors. The mutated DNA was readily detectable in bodily fluids from each type of cancer, suggesting that mtDNA mutations may provide a new molecular marker for noninvasive detection of cancer[72]. Mutations in proto-oncogenes and tumor suppressor genes can be produce vast changes in the expression of many other downstream genes. These changes can be assessed at the RNA level directly using differential display techniques and with newer RNA-based techniques like serial analysis of gene expression (SAGE) and cDNA microarray chips[74]. A major hurdle to the success of these techniques is obtaining a pure sample of tumor tissue. Pure samples are especially difficult to obtain from solid tumors (ninety percent of cancers are solid tumors)[74]. Techniques like cell-sorting and laser-capture microsurgery may be helpful in obtaining purer samples. Another problem with nucleic acid-based tests is that the information obtained is often complex and difficult to interpret. The decision to undergo genetic testing should therefore be a personal, voluntary one and should only be made in conjunction with appropriate genetic counseling. For now, genetic testing for cancer seems most efficacious for people within defined high-risk groups; screening for sporadic cancers in the population at large and even using genetic tests to follow treatment in patients already diagnosed with cancer does not yet seem realistic.

Tumor Staging and Grading

Once cancer is found, steps are taken to determine what type it is and how fast it is growing. A physician will also try to figure out whether cancer cells have invaded nearby healthy tissue or spread (metastasized) to other parts of the body. This is called staging and grading. These factors are important in trying to make predictions about cancer development and appropriate treatment.

Tumor Staging. Tumor staging systems describe how far cancer has spread anatomically. There are two related staging systems. In one system, cases are grouped into four stages denoted by Roman numerals I through IV. Stage I cancers are small localized cancers that are usually curable, while stage IV usually represents inoperable or metastatic cancer. Stage II and III cancers are usually locally advanced and/or with involvement of local lymph nodes. It is important to realize that the prognosis also depends on what kind of cancer it is, so that a stage II non small cell lung cancer has a different prognosis from a stage II cervical cancer.

The second system is called the Tumor, Nodes and Metastases ("TNM") system. Each of these categories, T, N, and M, is classified with a number to give the total stage. T Classifies the primary tumor and is normally given as T0 through T4. T0 represents *in situ* carcinoma; T4 represents a large primary tumor that has probably invaded other organs by direct extension and is usually inoperable. N classifies the amount of *regional* lymph node involvement. (Involvement of distant lymph nodes is considered to be

metastatic disease and falls under the M category.) N0 means no lymph node involvement while N4 means extensive involvement. M is either M0 if there are no metastases or M1 if there are metastases. A T1N1M0 cancer is defined as having a T1 tumor, N1 lymph node involvement, and no metastases.

The TNM system is more precise than the I through IV system and certainly has a lot more categories. The two systems are actually related. The I through IV groupings are actually defined using the TNM system. For example, stage II non-small cell lung cancer means a T1 or T2 primary tumor with N1 lymph node involvement, and no metastases (M0). Also, some types of cancers use alternative staging systems. Prostate and colon cancer are often staged as A through D rather than I through IV. For many other cancers the staging is further divided with classifications like IIa, and IIIb. It's important to keep in mind that there is considerable variation in outcome for every type and stage of cancer. A prognosis associated with a cancer stage is only a general guide, not an infallible prediction.

Tumor Grading. Tumor grade is based on microscopic analysis and generally refers to the appearance of the cells or to the percentage that appear to be dividing. The higher the grade, the more aggressive and fast growing the cancer. Microscopic examination also provides information regarding the likely behavior of a tumor and its responsiveness to treatment. Cancers with highly abnormal cell appearance and large numbers of dividing cells tend to grow more quickly, spread to other organs more frequently, and be less responsive to therapy than cancers whose cells have a more

normal appearance. Based on these differences in microscopic appearance, doctors assign a numerical "grade" to most cancers. In this grading system, a low number grade (grade I or II) refers to cancers with fewer cell abnormalities than those with higher numbers (grade III, IV).

Finding cancer early may decrease a person's risk of dying from the cancer, or at least provide better palliative therapies if the cancer is too far progressed for curative therapies. For this reason, improving methods for early detection and for following the progression of cancer (and cancer treatment) is currently a high priority. Modern bioimaging techniques occupy an important place in cancer detection.

Cancer Imaging

Cancer is difficult to detect, yet for some types of cancer the sooner it is caught, the better are the chances of treating it effectively. Imaging techniques - methods of producing pictures of areas inside the body - are key to early detection. But physicians and researchers do not use imaging simply for detection. Imaging is also an important tool used in determining the stage/grade and the precise locations of cancer, to aid in directing surgery, to monitor therapy, and to check to see if cancer is recurring. Advances in computer technology in the 1980s transformed diagnostic imaging, now making it possible to visualize organs and soft tissues in a detail that had previously been available only with anatomical dissections. Imaging techniques such as CT, SPECT/PET, MRI and ultrasound provide windows into the body, allowing tumors or

other abnormalities to be detected in areas not accessible with a physical examination or x-rays alone.

X-ray Imaging

A beam of x-rays is shot straight through the tissue. As it comes out the other side, the beam is blunted or "attenuated" slightly because it has hit dense living tissues on the way through. Very dense tissues like bone or solid tumors block a lot of x-rays; grey matter of the brain blocks some and fluid even less. Basic (single angle) X-ray technology is used in initial diagnoses of breast tumors in a mammogram screening and to see if cancer has spread to the lungs or other areas in the chest.

Computed Tomography (CT)

X-ray detectors positioned around the circumference of the scanner collect attenuation X-ray images from multiple angles. A computerized algorithm reconstructs an image of each slice ("tome" is Greek for "slice"). CT generates a three-dimensional picture of a cross-section of the body, allowing doctors to target tumors and avoid vital structures. A conventional CT scan takes picture slices of the body, about 3-5 mm apart (figure..). Spiral (also called helical) CT scan takes continuous pictures of the body in a spiral motion, so that there are no gaps in the pictures collected. Spiral CT image slices can be made thinner than in conventional CT scans.

Directed radiography.

Tumor specific antibodies, linked to radioactive isotopes, can be used in CT to seek out and specifically identify cancerous growths. This technique is called immunodiagnosis.

Radiolabeled ligands that bind specifically to surface receptors overexpressed in tumor cells, are also being evaluated for their ability to specifically image metastatic tumor cell populations. Promising ligands include bioactive endogenous peptides or their analogues, inhibitors of glucose transport proteins, estrogen and sigma receptor ligands, growth factors, and cytokines[73]. Ligands for the laminin binding protein also show potential for imaging applications. In a group of studies, YIGSR, which has been shown to target the metastasis-associated laminin binding protein (LBP; see above), was radioiodinated and administered to mice with Lewis Lung carcinoma[74] and to B16 melanoma-bearing mice[75]. In both cases, the radiolabeled LBP ligands were able to bind tumor cells and were frequently found in sites of metasasis. Indium-111 labeled laminin peptide fragments, GYIGSR, $(GYIGSR)_2$, CDPGYIGSR ("peptide 11"), and $(CDPGYIGSR)_2$, were evaluated for their tissue distribution and ability to target (and image) metastatic tumor cell populations[76]. The peptide 11-based radiopeptides ((CDPGYIGSR and $(CDPGYIGSR)_2$) showed the highest specific tumor localization. Although extensive research is still required, peptide 11- and YIGSR-based imaging techniques may prove to be useful for the in vivo estimation of the metastatic potential of tumors and in staging and prognostic-related decisions on cancer treatment. This thesis is

concerned with extending peptide-based targeted imaging approaches by utilizing magnetic resonance imaging (MRI; see below).

Single Photon/Positron Emission Computed Tomography (SPECT/PET)

In SPECT/PET, isotopes which emit positrons (PET) or agents that release photons (SPECT) (gamma rays/photons are produced when positrons collide with electrons) are attached to analogues of naturally occurring substances. Compounds similar to glucose, water, ammonia, and certain drugs may be used. The radioactive drug is administered to the patient, usually by injection and a PET scanner images how the body processes the drug. The absorption and processing of the labeled substance may reflect, for example, blood flow, oxygen or glucose metabolism, or dopamine transporter concentration. Often images are shown with a color scale. SPECT/PET is unique because it produces images of the body's basic biochemistry or function compared to other diagnostic techniques, such as x-rays, CT scans or MRI (see below), which produce images of the body's anatomy or structure. The premise with these techniques is that the change in structure or anatomy that occurs with disease can be seen. Biochemical processes are also altered with disease and may occur before there is a change in gross anatomy. SPECT/PET is an imaging technique that is used to visualize some of these processes that change. For instance, in cancer, malignant tumor cells have been found to have a high rate of glycolysis to be associated with increased membrane glucose transfer capability. (F-18) fluorodeoxyglucose (FDG; Fig. 19) is a tracer compound (a radioactive

glucose analogue) that has proven useful for tracing glucose metabolism, and therefore for detecting malignant tissue and quantifying changes in tumor glycolysis during and after treatment.

Ultrasound

This constantly improving technology sends high-frequency sound waves (the propagation of pressure waves through some physical elastic medium) into tissues. The tissue interfaces reflect the sound, and the resulting pattern of sound reflection is processed by a computer to produce a photograph or a moving image on a television. Ultrasound can be used to examine many parts of the body, but its best known application is the examination of the fetus during pregnancy. Ultrasound can also distinguish between solid tumors and cysts, and is frequently used to detect tumors in the breast and pelvic region.

Magnetic Resonance Imaging (MRI)

Magnetic Resonance Imaging (MRI) is a noninvasive medical diagnostic tool which utilizes proton nuclear magnetic resonance (NMR). Recent studies investigating the use of PET, gamma cameras, and SPECT for imaging applications indicate ranges of detection on the order of cubic millimeters and larger. MRI provides an alternative to these techniques, allowing the interrogation of intact, opaque tissue sections at cellular

resolution (~10 μm)[77]. Other advantages of MRI include superior contrast resolution, superior tissue characterization, superior multiplanar capabilities, no exposure to ionizing radiation, and possibly safer intravenous contrast agents.

The magnetic resonance machine is somewhat similar to a CT machine in configuration. It consists of a large frame with a center aperture in which a patient is placed on a table which can be precisely controlled. The patient is in the bore of a large magnet which produces the background magnetic field, aligning proton nuclei. Most of the magnets in clinical use are superconducting magnets although permanent magnets and resistive electromagnets are also used. Gradient coils and radio frequency coils are used to transmit pulses of radio-frequency energy (RF) causing populations of aligned nuclei to move from alignment with to alignments against the magnetic field. Intrinsic spin causes nuclei to precess. With time, the angle of precession is reduced as the nuclei return to alignment with the magnetic field and, in doing this, a radio frequency signal is emitted according to standard T_1 and T_2 relaxation kinetics. Multiple crosssectional scans of the patient's body are taken, and the various chemical shifts of absorbing protons are translated into specific colors. A tilt series provides a three dimensional picture showing the relative density of specific types of protons; for instance, a dark blue area on the MRI output screen indicates a high concentration of the types of protons giving rise to the range of resonances that correlate to dark blue. Comparison to normal MRI then allows the diagnostician to detect abnormalities in the scanned region.

The timing of the RF pulse and signal sampling can be varied so that images (or numerical matrixes relative to signal) can be produced showing different characteristics

of the tissue. Commonly used methods for medical images include signals which are weighted for proton density, T_1 (spin-lattice) relaxation, and T_2 (spin-spin) relaxation. The signal intensity arising from any one-volume element (voxel) in the 3D image is typically a function of the water concentration and relaxation times. Local variations in these parameters provide the vivid contrast seen in the *in vivo* images obtained in MRI. T_1 images demonstrate overall signal intensity and are more "anatomical". T_2 images have high signal return from areas with increased water (this is fortunate since most pathologic processes have increased water content within the tissue).

Providing contrast in MRI.

Although MRI was initially hoped to provide a means of making definitive diagnoses noninvasively, it has been found that the addition of contrast agents in many cases improves sensitivity and/or specificity. MRI contrast agents act by modulating the local T_1 and/or T_2 relaxation times[78]. T_1 agents reduce the spin-lattice relaxation times of nearby water, increasing the signal from these protons and making the effected voxel seem brighter in a T_1-weighted image. Typical T_1 agents provide at least one coordination site for the rapid exchange of water molecules, and the decrease in T_1 is mediated by the direct interaction of water molecules with the unpaired electrons of a paramagnetic metal ion. T2 agents work by altering the local homogeneity of the magnetic field.

Intravenous (IV) contrast agents include chelates of paramagnetic ions, both ionic and nonionic. Paramagnetic metal ions suitable as MRI contrast agents are all potentially toxic when injected IV at or near doses needed for clinical imaging. With chelation of these ions, acute toxicity is reduced and elimination rate is increased thereby reducing the chance of long term toxicity. Chelates of paramagnetic ions with EDTA were first used. However, Gadolinium (Gd)-EDTA was toxic in animals. Chelates with a higher stability constant have since been used successfully. Gd-DTPA was the first intravenous MRI contrast agent to be approved for human use (Magnevist, Berlex Labs). Gd (III) has a large magnetic moment, exceeded only by Dysprosium(III) and Holmium(III), explaining its paramagnetic properties at low concentrations and its ability to provide good T_1 contrast. This large magnetic moment is related to its seven unpaired orbital electrons. Nonionic Gd-based contrast agents include gadodiamide (Omniscan, Winthrop Pharm.) and gadoteridol (Prohance, Squibb).

There are a variety of tumor-specific MRI contrast agents[79]. Tumor specific agents are pharmaceuticals that are targeted to tumors, either specifically or nonspecifically. Examples include metalloporphyrins, nitroxides, ferrioxamine, gene-inducible contrast agents, and tumor-targeted contrast agents.

Metalloporphyrins exhibit affinity for many tumor types including carcinoma, sarcoma, neuroblastoma, melanoma and lymphoma. The metalloporphyrin most commonly used as a MRI contrast agent is Mn(III)TPPS4 (manganese(III) tetra-[4-sulfanatophenyl] porphyrin) because of its low toxicity (compared to Fe(III)TPPS4 for example). This material appears to work best with tumors that are isointense to

surrounding structures on T_1-weighted sequences. Incidentally, the fluorescent and tumor localizing characteristics of porphyrin derivatives have been exploited in phototherapy of tumors.

Nitroxide stable free radicals (also called nitroxyl spin labels) are chemically stable organic compounds that have an unpaired electron that results in paramagnetic properties. They generally consist of a six-member ring piperidine derivative or a five-member ring pyrroxamide derivative. The pharmacokinetics of nitroxides are similar to iodinated contrast agents and Gd-DTPA. They do not cross an intact blood brain barrier and undergo glomerular filtration as a dominant route of elimination. Their ease of conjugation to various biomolecules makes them attractive for targeting to various organ systems. Nitroxides are chemically stable and show limited in vivo metabolism. Their relaxation effects in vivo can be eliminated almost immediately by IV injection of sodium ascorbate, a strong reducing agent. This will allow an unenhanced MR study to be performed immediately after a contrast enhanced study, if the contrast study is not satisfactory alone. Larger molecular weight nitroxides exhibit increased relaxation rates (as do paramagnetic ions attached to macromolecules). This phenomenon occurs when attaching five-membered nitroxide rings to fatty acids. The fatty acids attach to human serum albumin, either in vitro or in vivo, resulting in a significant increase in relaxation rate. Safety studies and clinical trials need to be performed before nitroxides will be available for use.

Ferrioxamine methanesulfonate is a paramagnetic contrast agent that has undergone phase I and phase II clinical trials for use as an IV and retrograde contrast

agent for the kidneys, ureters and bladder. It is more stable than Gd-DTPA, though its relaxivity is somewhat less, as expected from it having 5 unpaired electrons, vs. 7 unpaired electrons for Gd-DTPA.

A recent study investigated the utility of an enzyme activatable Gd-based MRI contrast agent. They found that a reporter enzyme (β-galactosidase) was able to modify the agent, making Gd accessible to water molecules, thereby providing T_1-contrast in MRI. When the reporter enzyme was expressed in developing *Xenopus laevis* embryos, regions of higher intensity in the MR image correlated with regions expressing marker enzyme. Images were obtained at cellular resolution. This enzyme activatable imaging method holds promise for a wide variety of applications. It was postulated that this technique could monitor the effectiveness of gene therapy, and it might provide a technique for early detection of new tumors[80]. However, considering the heterogeneity of mutations and downstream mutant proteins that exist across the gamut of solid human tumors, and the general unreliability of genetic markers in detecting early sporadic cancers, the use of Gd-based agents targeted to gene expression is, for now, unrealistic. This technique awaits the identification of specific genes causally linked to numerous cancer types whose protein products are somehow able to activate the enzyme-linked contrast agent. Gene-activated bio-imaging agents do show promise in following gene-based cancer therapies (in which a reporter gene could be co-transfected with the therapeutic gene).

Tumor-targeted MRI contrast agents are also becoming popular. Monoclonal antibodies (Mabs) are used successfully in nuclear medicine for localization of tumors

(immunoradiotherapy; see above) but initial attempts at extending this use to MRI with paramagnetic (Gd (III)) labeled antibodies was unsuccessful because of the estimated 800-fold lesser sensitivity of MRI. This problem could be addressed in several ways: The number of paramagnetic ions attached to the Mab could be increased. More antibodies or those with an affinity to multiple antigenic sites per cell (or both) could be used. A superparamagnetic particle (composed of individual domains of elements that have ferromagnetic properties in bulk) could be attached to the MAb. Implanted human colon carcinoma tumors in mice have been successfully imaged by using monoclonal antibodies with a large number of Gd-DTPA molecules attached. Additional studies report the use of very small magnetite particles coated with MAb. The magnetite cores are 10-20 nm in diameter with a total particle diameter of 20-32 nm. The magnetic moment of these superparamagnetic particles is about 1000 times that of comparable paramagnetic particles. This allows the use of 1-10 nmol concentrations of the Mab coated magnetite particles.

Numerous paramagnetic ions could also be attached to a macromolecule. Association with a host molecule increases T_1 relaxation times and other attributes of the macromolecule can be exploited. Spherical polymers (starburst dendrimers), which incorporate significant amounts of Gd (from 10's to 1000's of Gd ions)[81,82], carry a large Gd payload and provide strong MRI contrast. The contrast-agent labeled macromolecule could in turn be attached to a tumor-targeting Mab for use in cancer imaging. An extension of this approach represents the long-term aim of this thesis project. Our idea is

to engineer an ordered protein cage to target large numbers of Gd ions to metastatic tumors (see main text of this thesis).

REFERENCES CITED FOR APPENDIX C

[1] Nuland SB. 1995. How we die. Vintage Books – Random House, Inc., New York.

[2] Huebner RJ, Todaro GJ.1969. Oncogenes of RNA tumor viruses as determinants of cancer. Proc Natl Acad Sci U S A.64(3):1087-94.

[3] Temin HM. 1971. Mechanism of cell transformation by RNA tumor viruses. Annu Rev Microbiol.25:609-48.

[4] Temin HM. 1971. The protovirus hypothesis: speculations on the significance of RNA-directed DNA synthesis for normal development and for carcinogenesis. J Natl Cancer Inst.46(2):3-7.

[5] Mizutani S, Temin HM. 1971. Enzymes and nucleotides in virions of Rous sarcoma virus. J Virol.8(4):409-16.

[6] Stehelin D, Varmus HE, Bishop JM, Vogt PK. 1976. DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA.Nature.260(5547):170-173.

[7] Sonnex, C.. 1998. Human papillomavirus infection with particular reference to genital disease. J Clin Pathol. 51:643-648.

[8] Storey, A., et al. 1998. Role of a p53 polymorphism in the development of human papilloma-virus-associated cancer. Nature. 393:229-234.

[9] Gao, Q., et al. 1999. The E6 oncoproteins of high-risk papillomaviruses bind to a novel putative GAP protein, E6TP1, and target it for degradation. Mol Cell Biol. 19(1):733-44.

[10] Fields Virology, Third Edition. Edited by B.N. Fields, D.M. Knipe, P.M. Howley, et al. Lippincott – Raven Publishers, Philadelphia © 1996.

[11] Zwerschke, W., et al. 1999. Modulation of type M2 pyruvate kinase activity by the human papillomavirus type 16 E7 oncoprotein. PNAS. 96:1291-1296.

[12] Li R, Sonik A, Stindl R, Rasnick D, Duesberg P. 2000. Aneuploidy vs. gene mutation hypothesis of cancer: recent study claims mutation but is found to support aneuploidy. Proc Natl Acad Sci.97(7):3236-41.

[13] Hahn WC et al. 1999. Creation of human tumor cells with defined genetic elements. Nature. 400(6743):464-8.

[14] Lengauer C, Kinzler KW, Vogelstein B. 1998. Genetic instabilities in human cancers. Nature. 396(6712):643-9. Review.

[15] Boland CR, Ricciardiello L. 1999.How many mutations does it take to make a tumor? Proc Natl Acad Sci.96(26):14675-14677.

[16] Mellon I et al. 1996. Transcription-couples repair deficiency and mutations in human mismatch repair genes. Science. 272:557-560.

[17] Gura T. 1997. New kind of cancer mutation found. Science (Research News). 277:1201-1202.

[18] Pennisi E. 1998. How a growth control path takes a wrong turn to cancer. Science.281:1438-1441.

[19] Tong-Chuan He et al. 1998. Identification of c-MYC as a target of the APC pathway. Science.281:1509-1512.

[20] Fishel R. 1998. Mismatch repair, molecular switches, and signal transduction. Genes and Development. 12:2096-2101.

[21] Michael J. Difilippantonio et al. 2000.DNA repair protein Ku80 suppresses chromosomal aberrations and malignant transformation. *Nature* 404, 510 - 514

[22] Seeger RC et al. 1985. Association of multiple copies of the N-myc oncogene with rapid progression of neuroblastomas. N. Engl. J. Med.313:1111-1116.

[23] Brodeur GM and Hogerty MD. 1998. In The genetic basis of human cancer 1st edition vol.1.161-179. (McGraw-Hill, New York). (referenced from 14 and personal notes)

[24] Ray A and Runge K. 1999. The yeast telomere length counting machinery is sensitive to sequences at the telomere-nontelomere junction. Mol Cel Biol.19(1):31-45.

[25] Rudolph KL et al. 1999. Longevity, stress response, and cancer in aging telomerase-deficient mice. Cell.96:701-712.

[26] Greider CW. 1998. Telomerase activity, cell proliferation, and cancer. PNAS.90-92.

[27] Lange T. 1998.Telomeres and Senescence: ending the debate. Science.270:334-335.

[28] Blasco MA et al. 1997. Telomere shortening and tumor formation by mouse cells lacking telomere RNA. Cell.91:25-34.

[29] http://www.waisman.wisc.edu/cytogenetics/BMproject/CancerCyto.htmlx

[30] Nuland SB. 1998. How we live. Vintage Books – Random House, Inc., New York.

[31] Terranova VP et al. 1984. Modulation of the metastatic activity of melanoma cells by laminin and fibronectin. Science 226:982-985.

[32] Koop S et al. 1995. Fate of melanoma cells etering the microcirculation: over 80% survive and extravasate. Cancer Research. 55:2520-2523.

[33] Morris VL et al. 1997. Sequential steps in hematogenous metastasis of cancer cells studied by in vivo videomicroscopy. Invasion Metastasis. 17:281-296.

[34] Sakamoto N et al. 1990. Inhibition of angiogenesis and tumor growth by a synthetic laminin peptide, CDPGYIGSR-NH_2. Cancer Research 51:903-906.

[35] Lin Z, Brattain MG, Appert H. 1997. Differential display of reticulocalbin in the highly invasive cell line MDA-MB-435, versus the poorly invasive cell line, MCF-7. Biochem Biophys Res Commun 231:283-289.

[36] Mizejewski GJ. 1999. Role of integrins in cancer: survey of expression patterns. PSEBM 222:124-138.

[37]Ruoslahti E. 1996. How cancer spreads. *Sci Am.* 275(3):72-77.

[38] Orr FW et al. 2000. Interactions between cancer cells and the endothelium in metastasis. J Pathol.190(3):310-329.

[39] Christofori G and Semb H. 1999. The role of the cell-adhesion molecule E-cadherin as a tumor-suppressor gene. Trends Biochem Sci.24(2):73-76.

[40]Behrens J. 1999.Cadherins and catenins:role in signal transduction and tumor progression.18(1):15-30.

[41] Goldberg GI et al. 1990. Secreted proteases. Regulation of their activity and their possible role in metastasis. *Ann N Y Acad Sci.*580:375-384

[42] DeClerck YA et al. 1997.Proteases and protease inhibitors in tumor progression. *Adv Exp Med Biol.*425:89-97

[43] Carr AM. 2000. Peicing together the p53 puzzle. Science. 287:1765-1766.

[44] Hirao A et al. 2000. DNA damage-induced activation of p53 by the checkpoint kinase Chk2. Science. 287:1824-1827.

[45] Wang JYJ. 2000. New link in a web of human genes. Nature.405:404-405.

[46] Ryan KM et al. 2000. Role of NF-κB in p53-mediated programmed cell death. Nature. 404:892-896.

[47] Momand J et al. 2000. MDM2-master regulator of the p53 tumor suppressor protein. Gene. 242:15-29.

[48] Hartwell LH and Kastan MB. 1994. Cell cycle control and cancer. Science. 266:1821-1827.

[49] Thornberry NA and Lazebnik Y. 1998. Caspases: enemies within. Science. 281:1312-1316.

[50] Cryns VL et al. 1997. Specific proteolysis of the kinase protein kinase c-related kinase 2 by caspase-3 during apoptosis. J Biol Chem. 272(47):29449-29453.

[51] Hakahira H et al. 1998. Cleavage of CAD inhibitor in CAD activation and DNA degradation during apoptosis. Nature. 391:96-99.

[52] Fadok VA et al. 2000. A receptor for phosphatidylserine-specific clearance of apoptotic cells. Nature. 405:85-90.

[53] http://www.nci.nih.gov/cancerinfo/index.html

[54] Masters JRW and Lakhani SR. 2000. How diagnosis with microarrays can help cancer patients. Nature (correspondence). 404:921.

[55] Hellman S and Vokes EE. 1996. Advancing current treatments for cancer. Sci Amer. 118-123.

[56] Shih JL, Brugger RM. 1992. Gadolinium as a neutron capture therapy agent. Med Phys. 19(3):733-44. Review.

[57] Gibbs J. 2000. Mechanism-based target identification and drug discovery in cancer research. Science. 287:1969-1973.

[58] E. Finkel. 2000. Does cancer therapy trigger cell suicide? Science. 286:2256-2258.

[59] Old LJ. 1996. Immunotherapy for cancer. Sci Amer. 136-143.

[60] Lord EM and Frelinger JG.1998. Tumor immunotherapy: cytokines and antigen presentation. Cancer Immunol Immunother. 46:75-81.

[61] Hermans IF et al. 1997. Tumor-peptide-pulsed dendretic cells isolated from spleen or cultured in vitro from bone marrow precursors can provide protection against tumor challenge. Cancer Immunol Immunother. 44:341-347.

[62] Semino C et al. 1999. Adoptive immunotherapy of advanced solid tumors: and eight year clinical experience. Anticancer Res. 19(6C):5645-5649.

[63] Hahn WC et al. 1999. Inhibition of telomerase limits the growth of human cancer cells. Nat Med. 5(10):1164-1170.

[64] BD Michael and P O'Neill. 2000. A sting in the tail of electron tracks. Science. 287:1603-1604 and B Boudaiffa et al. 2000. Resonant formation of DNA strand breaks by low-energy (3 to 20 eV) electrons. Science. 287:1658-1660.

[65] Rennie J and Rusting R. 1996. Making headway against cancer. Sci Amer. 56-59.

[66]Sonnex C. 1998. Human papillomavirus infection with particular reference to genital disease. J Clin Pathol. 51:643-648.

[67] http://www.bmj.com/cgi/content/full/320/7232/424.

[68] Hiyama E et al. 1996. J Nat Canc Instit. (88):116-122. therapy.

[69]http://rex.nci.nih.gov/PATIENTS/INFO_TEACHER/bookshelf/NIH_gene_testing/gene00.html

[70] Sidransky D. 1997. Nucleic acid-based detection methods for the detection of cancer. Science. 278:1054-1058.

[71] Gryfe R et al. 2000. Tumor microsatellite instability and clinical outcome in young patients with colorectal cancer. N Engl J Med. 342(2):69-77.

[72] Fliss MS et al. 2000. Facile detection of mitochondrial DNA mutations in tumors and bodily fluids. Science. 287(5460):2017-2019.

[73] Mcafee JG and Neumann RD. 1996. Radiolabeled peptides and other ligands for receptors overexpressed in tumor cells for imaging neoplasms. Nucl Med Biol. (6):673-676.

[74] Koliakos G et al. 1997. Lung carcinoma imaging using a synthetic laminin derivative radioiodinated peptide YIGSR. The Journal of Nuclear Medicine. 38(12):1940-1944.

[75] Kouzi-Koliakos K et al. 1997. In vivo binding of the rasioiodinated peptide YIGSR on B16 melanoma cells. Invasion and Metastasis. 16:322-329.

[76] Mokotoff M et al. 1997. Evaluation of laminin peptide fragments labeled with indium-111 for the potential imaging of malignant tumors. J Pept Res. 49(6):510-516.

[77] *http://www.mritutor.com/*

[78] Louie AY et al. 2000. In vivo visualization of gene expression using magnetic resonance imaging. Nature Biotech. 18:321-325.

[79] http://128.227.164.224/mritutor/tumorcon.htm

[80] Straus E. 2000. Detecting enzyme activity in live animals. Science.287:1575.

[81] Bryant LH et al. 1999. Synthesis and relaxometry of high-energy (G=5,7,9, and 10) PAMAM dendrimer-DOTA-gadolinium chelates. J Magn Res Imag. 9:348-352.

[82] Fischer M and Vogtle F. 1999. Dendrimers: from design to application – a progress report. Angew Chem Int Ed. 38:884-905.

MONTANA STATE UNIVERSITY - BOZEMAN
3 1762 10337681 8