

Crystal Structure of the Vinculin Tail Domain

and

A Model For Activation

Thesis submitted for the degree of

Doctor of Philosophy

at the University of Leicester

by

Constantina Bakolitsa BSc (Salford)

Department of Biochemistry

University of Leicester

November 1999

UMI Number: U125615

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U125615

Published by ProQuest LLC 2014. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

Title: Crystal structure of the vinculin tail domain and a model for activation.

Author: Constantina Bakolitsa

Cell adhesion is a fundamental requirement in multicellular organisms for a variety of biological processes, such as tissue integrity, cell growth and differentiation, and wound healing. At sites of cell-matrix adhesion, termed focal adhesions, protein networks converge in a variety of multimolecular complexes composed of transmembrane adhesion receptors, anchoring intracellular cytoskeletal structural proteins and signal transduction molecules. Basic physical mechanisms underlying the regulation of focal adhesion formation and function include aggregation, phosphorylation and conformational changes of the molecules involved.

Vinculin is a protein located at sites of cell-cell and cell-matrix adhesion where it plays a dynamic role in the assembly of the actin cytoskeleton. A strong interaction between its head and tail domains that regulates binding to other cytoskeletal components is disrupted by acidic phospholipids. The crystal structure of the vinculin tail, residues 879-1066, was determined by the method of single isomorphous replacement with anomalous scattering. Five amphipathic helices form an antiparallel bundle that resembles exchangeable apolipoproteins. A C-terminal arm wraps across the base of the bundle and emerges as a hydrophobic hairpin surrounded by a collar of basic residues, adjacent to the N-terminus.

A series of structure-designed solution studies were carried out on the vinculin tail and a C-terminal deletion mutant. Co-sedimentation assays showed that the C-terminal arm is required for binding to acidic phospholipids but not to F-actin, while limited proteolysis and stopped-flow measurements indicated that binding to either ligand induces conformational changes in the N- and C-terminal regions of the molecule. These changes may represent the first step in a model for the activation of the intact vinculin, where PIP₂-triggered unfurling of the tail domain would lead to subsequent sequestration of the head and neck regions allowing binding to F-actin.

Crystals of the intact vinculin diffracting synchrotron radiation to medium resolution have also been obtained.

Acknowledgments

I am grateful to my supervisors, David Critchley and Bob Liddington, for offering me the chance to work on this project - Bob especially for his special brand of light-handed supervision which allowed me to explore all possible experimental avenues at my own risk, and friendly take-over tactics which made for some high-speed results. Without this freedom and timely assistance, my years in his lab would not have been as enjoyable or as productive, and this thesis would probably have had a very different outcome.

Thanks to all the following: Bipin Patel for providing the original Vt vector and for carrying out the actin co-sedimentation and Vt Δ C/F-actin limited proteolysis experiments; Robert Woolley for assistance with the stopped-flow measurements; Clive Bagshaw for critical discussions of the stopped-flow results; Kathryn Lilley for prompt delivery of mass spectra, DNA and N-terminal protein sequencing; Lu Yun Lian for conducting and interpreting the NMR measurements.

Special thanks to José Maria de Pereda for his everyday experimental and moral support, in particular his work with the Vt Δ C mutant, PIP2 and CD experiments, and unrivaled Spanish omelettes. Thanks also to Andrew Pannifer for help with the molecular replacement of Vt and for providing me with wine and shelter in times of need, Carlo Petosa for collecting the very first diffraction images of Vt and for being both a friend and a mentor, Jonas Emsley for assistance with Vg data processing, Geoff Briggs for his good cheer and patient instruction in all things biochemical, computational and generally transgressive, and Mike Wilson for discussions that could (and invariably did) lead to places not many had trodden before. Finally thanks to all the rest of the G5/G8 team for providing a great work and social environment and for putting up with my everyday moods and antics.

Leicester

November, 1999

To ET.

- Table of Contents -

Chapter 1: Introduction	1
Cell adhesion and migration	1
Cadherins	2
Integrins	5
Focal adhesions	8
Signal transduction	9
Vinculin	10
Domain structure	12
Gene disruption	13
Isoforms and post-translational modifications	14
Interactions with acidic phospholipids	15
Interactions with actin	16
Evidence for oligomerisation	17
Summary	17
Scope of thesis	19
Chapter 2: Crystallisation of Vt	32
Introduction	32
Production of Vt and SeMVt	
Expression and purification of Vt using a GST-fusion	33
Subcloning of Vt	35
Expression and purification of Vt using a His-tag	38
Expression and purification of SeM-Vt	39
Crystals of Vt	
Tetragonal form	41
Monoclinic form	41
Orthorhombic form	42
Materials	43
General methods	
SDS-PAGE	43
Protein concentration assays	44
Summary	45

Chapter 3: Structure Determination of Vt	53
Introduction	53
Data collection and processing	54
Heavy metal derivatives	55
Model building and refinement	61
Model refinement using a monoclinic crystal	62
Assessment of the reliability of the current model	64
Agreement with crystallographic data	64
B factors and geometry	65
Crystal packing, heavy metal sites, PDB validation	66
Summary	67
Chapter 4: The Crystal Structure of Vt	81
An overview of the crystal structure	81
Amphipathicity	82
Charge distribution and surface properties	83
Methionine zipper	84
Comparison of the two crystal forms	85
Dimer contacts	85
N-terminal arm	86
C-terminal arm	87
Similarity to exchangeable apolipoproteins	88
Comparison with EM structures	89
Hypothesis of C-terminal involvement in membrane interaction	89
Summary	90
Chapter 5: Structure-Based Solution Studies of Vt	101
Introduction	101
Vt Δ C	102
Interactions with actin	103
Co-sedimentation assays	103
Stopped-flow fluorescence	103
Limited proteolysis	105
Interactions with phospholipids	106
Co-sedimentation assays	106
Limited proteolysis	107

NMR studies	
¹ H spectra	107
¹ H- ¹³ C spectra	107
Materials and methods	109
Design, expression, and purification of VtΔC	109
Co-sedimentation assays	109
Limited proteolysis	110
Stopped-flow fluorescence	111
NMR spectroscopy	111
Summary	112
Chapter 6: Discussion	123
Introduction	123
Role of C-terminal arm	124
Conformational changes	124
Relationship between actin and phospholipid binding	127
Other binding sites	128
Head-binding region	128
VASP, vinexin and ponsin	129
A pathway for activation	130
Perspectives	130
Summary	133
Chapter 7: Crystallisation of Vg	135
Introduction	135
Crystals of Vg	136
Data collection and processing	136
Heavy metal derivatives	136
Discussion	137
Materials and methods	138
Purification of Vg	138
Summary	139
References	144
Publications	165

- List of Tables -

Chapter 1

1.1	Integrin α - β subunit combinations and ligands	20
1.2	Common structural FA components	21
1.3	Vinculin binding proteins and ligands	22

Chapter 2

2.1	A summary of crystal forms	45
-----	----------------------------	----

Chapter 3

3.1	Data collection statistics for crystals of Vt	68
3.2	Some of the heavy metal compounds screened using monoclinic crystals	69
3.3	Heavy metal sites	70
3.4	Molecular replacement solutions of the monoclinic crystal	71
3.5	Model building and refinement statistics	72
3.6	Summary of all data collection, phasing and refinement statistics	73

Chapter 4

4.1	Salt-bridges in Vt	91
4.2	Hydrophobic moments of the helices of Vt	91
4.3	Dimer contacts in the two crystal forms	92

Chapter 7

7.1	Data collection statistics for a native crystal of Vg	140
7.2	Summary of data collection statistics for native and putative derivative data sets of Vg	140

- List of Figures -

Chapter 1

1.1	Schematic representation of the domain structure of a typical type I classical cadherin	23
1.2	Model of cadherin rearrangements during cell adhesion	23
1.3	Model of cadherin-mediated cell adhesion	24
1.4	Schematic representation of integrin α and β subunits	25
1.5	Proposed model for integrin outside-in signalling	25
1.6	Major integrin-mediated signal transduction pathways	26
1.7	Model for Rho-mediated signalling	27
1.8	Proposed models for the interaction of vinculin at the sites of focal adhesions	28
1.9	Comparison of homology domains and ligand binding sites in vinculin and α -catenin	29
1.10	Schematic representation of the domain organisation of vinculin	30
1.11	Galery of individual ESI views of vinculin molecules	31

Chapter 2

2.1	Purification of Vt using a GST-affinity column and ion-exchange	46
2.2	Vt in pET15b	47
2.3	Purification of Vt using a His-tag	48
2.4	Mass spectroscopy of Vt and SeMVt	49
2.5	Crystals of Vt	51

Chapter 3

3.1	Harker sections of EMTS2 isomorphous difference Patterson	74
3.2	$z=1/2$ Harker sections of EMTS1 difference Patterson	75
3.3	Difference Fourier of EMTS2	76
3.4	Stereo-view of the first experimental electron density map of Vt	77
3.5	Tryptophan stacking	78
3.6	Average B factors for main chain atoms in the monoclinic model of Vt	79

3.7	Ramachandran plot	80
Chapter 4		
4.1	Stereo ribbon representation of Vt	93
4.2	Secondary structure assignment and sequence alignments	94
4.3	Molecular surface of Vt	95
4.4	Methionine zipper	97
4.5	Heptad repeats in Vt	98
4.6	Stereo view of a backbone superposition of the four copies of Vt	99
4.7	Structural comparison of Vt with apolipoproteins	100
Chapter 5		
5.1	Co-sedimentation of Vt and Vt Δ C with actin	113
5.2	Change in tryptophan fluorescence during the interaction of Vt and Vt Δ C with F-actin as measured by stopped-flow	114
5.3	Limited proteolysis by elastase	115
5.4	Limited proteolysis by chymotrypsin	116
5.5	Limited proteolysis of Vt Δ C by elastase	117
5.6	Limited proteolysis of Vt Δ C by chymotrypsin	118
5.7	Co-sedimentation of Vt and Vt Δ C with phospholipids	119
5.8	Limited proteolysis of Vt by a) elastase and b) chymotrypsin in the presence of PS or PC MLVs	120
5.9	^1H spectra of Vt at two pH values	121
5.10	^1H - ^{13}C HSQC spectrum of ^{13}C -labelled-Met Vt in the a) presence and b) absence of TEMPO	122
Chapter 6		
Figure 6.1	A model for vinculin activation	134
Chapter 7		
Figure 7.1	A crystal of Vg	141
Figure 7.2	1° oscillation image of a native Vg crystal	142
Figure 7.3	Purification of Vg by ion-exchange	143

List of Abbreviations & Recipes

Amp	Ampicillin
AS	Ammonium sulphate
A280	Absorbance at 280 nm
A595	Absorbance at 595 nm
Binding Buffer	5 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl, pH 7.9
Charge Buffer	100 mM NiSO ₄
2xYT	Pet liter: 16 g bactotryptone, 10 g yeast extract, 5 g NaCl
DTT	Dithiothreitol
EDTA	Ethylenediamine tetraacetic acid
Elute Buffer	1M imidazole, 0.5 M NaCl, 20 mM Tris-HCl, pH 7.9
EM	Electron microscopy
EMTS	Ethyl mercury thiosalicylate
ESI	Electron spectroscopic imaging
FFT	Fast Fourier Transform
FPLC	Fast-protein liquid chromatography
GST	Glutathione-S-transferase
HSQC	Heteronuclear single quantum correlation
IPTG	Isopropyl β -D-thiogalactopyranoside
kDa	Kilodalton
LB	Per liter: 10g bactotryptone, 5 g yeast extract, 10 g NaCl
β-ME	2-mercaptoethanol
MLV	Multilamellar vesicles
MR	Molecular replacement
M9ZB	Per liter: 10 g bactotryptone, 5 g NaCl. After autoclaving and cooling, 1mM MgSO ₄ , 0.4% glucose added from autoclaved stocks
NCS	Non crystallographic symmetry
NMM	Per liter: 7.5 mM NH ₄ Cl, 8.5 mM NaCl, 55 mM KH ₂ PO ₄ , 100 mM K ₂ HPO ₄

NMMPlus	NMM including 1mM MgSO ₄ , 60 mM CaCl ₂ , 20 mM glucose, 10 mg/L thiamine, 10 mg/L biotin, 50 mg/L each of valine, isoleucine, and leucine, 100 mg/L threonine, phenylalanine and lysine, 59 mg/L SeM
NMR	Nuclear magnetic resonance
OD₆₀₀	Optical density at 600 nm
PBS	Phosphate buffered saline (10g NaCl, 0.25g KCl, 1.44g Na ₂ HPO ₄ , 0.25g KH ₂ PO ₄ in 1L)
PC	L- α -phosphatidyl-L-choline
PEG	Polyethylene glycol
PEG MME	Polyethylene glycol monomethyl ether
PIP2	Phosphatidylinositol 4,5-bisphosphate
PMSF	Phenylmethylsulfonylfluoride
PS	L- α -phosphatidyl-L-serine
PVDF	Polyvinylidene difluoride
RMS	Root-mean-square
RT	Room temperature
SA	Simulated annealing
SAD	Single anomalous dispersion
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate - polyacrylamide gel electrophoresis
SeM	Seleno-L-methionine
SIRAS	Single isomorphous replacement with anomalous scattering
SRS	Synchrotron radiation source
Strip Buffer	100 mM EDTA, 0.5 M NaCl, 20 mM Tris-HCl pH 7.9
TCB	Thrombin cleavage buffer: 150 mM NaCl, 20 mM Tris-HCl pH 8.0
TEMED	N, N', N', N',-tetramethylethylenediamine
Tris	Tris hydroxymethyl aminomethane
Vt	Vinculin 878-1066
VtΔC	Vinculin 878-1051
Vg	Gizzard smooth-muscle vinculin (1-1066)
Wash Buffer	60 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl, pH 7.9
w/v	Weight to volume

Chapter 1: Introduction

Cell adhesion and migration

Cadherins

Integrins

Focal adhesions

Signal transduction events

Vinculin

Domain structure

Gene disruption

Isoforms and post-translational modifications

Interactions with acidic phospholipids

Interactions with actin

Evidence for oligomerisation

Summary

Scope of thesis

Cell adhesion and migration

Establishing contacts with neighbouring cells and with the extracellular matrix (ECM) is a prerequisite in multicellular organisms for many fundamental cellular processes, such as the maintenance of tissue integrity, cell motility and cell growth. Normal cells only proliferate when they are anchored to a substratum, and adhesion has been shown to inhibit apoptosis (programmed cell death). In animals, cell adhesion is an integral part of the cell migration observed during normal biological processes, such as embryonic development,

wound healing and the immune response, as well as in the development of a number of pathological conditions, including inflammatory disorders, atherosclerosis and metastatic tumour spread (Giancotti and Mainiero, 1994). Cell surface adhesion receptors serve to connect the substratum with the cytoskeleton and adhesive interactions are used to generate the traction and force required for cell movement. Although sufficiently stable to guarantee firm cellular adherence, both cell-cell and cell-ECM contacts can be reversibly assembled and disassembled within minutes. The switch from adherent to non-adherent state can be initiated by a response to either cytoplasmic or external signals, and requires a highly dynamic interaction between cell-surface receptors, their ligands and the cytoskeleton-associated motile apparatus. These interactions are mediated by families of cell surface receptors, of which two, the cadherins and integrins, have been extensively studied.

Cadherins

Cadherins were first classified as a family of transmembrane glycoproteins that mediate Ca^{2+} -dependent homo- and heterotypic cell-cell adhesion (Takeichi, 1995). Historically, cadherins' function has been perceived to be to link cells and to stabilise these links through interaction with the cytoskeleton. The consequences of such adhesions can have dramatic effects on tissue stability and morphogenetic processes such as cell rearrangement (Steinberg, 1996), cell migration (Nakagawa and Takeichi, 1998), tissue formation (Linask et al., 1998), and the establishment and maintenance of neural networks (Suzuki et al., 1997). Defective cadherin function can render non-invasive cells invasive (reviewed in Yap, 1998; Christofori and Semb, 1999) and induced expression of various

cadherins in invasive cells (Lee et al., 1998) can suppress their invasiveness in a cell-contact-dependent manner (Foty and Steinberg, 1997).

Members of the cadherin family can be divided into four subgroups: classic cadherins, desmosomal cadherins, protocadherins, and atypical cadherins. Proteins are designated as members of the broadly defined cadherin family if they have one or more 'cadherin repeats', sequences of approximately 110 amino acids containing amino acid motifs with the conserved sequences LDREE, DXNDN, and DXD (Kreis and Vale, 1999). The type I classical cadherin family, of which E-cadherin is a prime representative, comprises about 30 members composed of a highly conserved carboxy-terminal cytodomain (~25 amino-acids), a single-pass transmembrane domain and five similar extracellular cadherin-motif subdomains (C1-C5, where C1 is the most distant from the membrane) with an HAV sequence (in the single letter code for amino acids) in C1 argued to be essential to cell-cell adhesion (figure 1.1) (Gallin, 1998). Recent high-resolution crystal analysis (Pertz et al., 1999) has provided a mechanistic basis for some of the behavioural characteristics of cadherin-expressing cells, suggesting that the dimerisation interface is proximal to the calcium-binding domain bridging subdomains C1 and C2 (see figure 1.2). When C1 becomes organised at higher calcium concentrations, steric rearrangements shift W2 (portrayed in figure 1.2 by pivoting amino acid structure) to dock into the intramolecular hydrophobic pocket formed in part by the highly conserved HAV sequence. Mechanistically, these findings suggest that inhibition of adhesion by the addition of HAV peptides may occur through their competitive binding to Trp2 and the consequent inability of Trp2 to dock in the hydrophobic pocket, blocking *trans* dimerisation and therefore cell-cell adhesion.

A downstream effect of the initiation of cell-cell adhesion is the recruitment of additional cadherins to the adhesion site and the assembly of the CAC (cadherin-associated complex) around the cytodomain/actin linkage (reviewed in Steinberg and McNutt, 1999). An increasing number of actin-binding or cadherin-associated molecules have been implicated as members of the CAC. In general, the classical cadherin cytodomain interacts with β -catenin or plakoglobin (γ -catenin), which in turn bind α -catenin (Aberle et al., 1994, 1996). α -Catenin is in turn attached to the actin cytoskeleton directly and/or indirectly through proteins including, α -actinin, vinculin, ZO-1 and a number of other molecules that associate with the CAC (reviewed in Yamada and Geiger, 1997). p120^{cas}, β -catenin and plakoglobin are targets of tyrosine phosphorylation, which causes them to dissociate from the CAC and thus disrupts adhesion. Such phosphorylation can be induced, directly or indirectly (e.g. through Src tyrosine kinases) by exposure of cells to a number of growth factors, including EGF, TGF β , PDGF and HGF. A variety of GTP-associated proteins, such as Rho, Rac (Braga et al., 1999) and IQGAP1 (Kuroda et al., 1998) have also been shown to modulate adhesion. Rac is necessary for the accumulation of actin near areas of cell-cell contact and Rac1 acts in conjunction with Cdc42 upstream of IQGAP1. Dephosphorylation of β -catenin is correlated with upregulation of protein kinase C (PKC) and the concomitant disruption of tight junctions. It is clear that the formation and maintenance of CACs is a dynamic, multiply regulated process where the metabolic state of the cell acts in concert with the external environment to modulate cell-cell adhesion (figure 1.3).

Integrins

The best-studied family of adhesion receptors are integrins (Hynes 1992), $\alpha\beta$ heterodimeric transmembrane glycoprotein receptors that bind to many components of the ECM as well as to soluble multivalent plasma proteins such as fibrinogen, and some cell surface adhesion molecules of the immunoglobulin superfamily (such as intracellular adhesion molecules, ICAMs). Specific, non-covalent pairing of the 16 known α chains and the 8 known β chains results in at least 22 distinct human integrin heterodimers. Integrins are classified into several families according to their β subunits, as most β chains can pair with multiple α subunits, but only a few α subunits pair with more than one β subunit (Hynes 1992). A list of integrin α - β subunit combinations and ligands is given in table 1.1.

A schematic diagram of integrin α and β subunits is shown in figure 1.4. Each α subunit has ~1000-1150 amino acids and a mature size of 140-210 kDa. All integrin α subunits have seven amino terminal repeating segments that may fold into a seven unit β -propeller motif, resembling that found in G protein β subunits (Springer, 1997). Some α subunits (α_1 , α_2 , α_L , α_M , α_X , α_D , α_E) have an inserted 'I-domain' region (sometimes called A domain) between amino terminal repeating units 2 and 3. This I domain can function as a discrete unit, with divalent cation and ligand binding activity (Michishita et al., 1993). Analysis of the I-domain crystal structure revealed a 'Rossmann' folding pattern, with an unusual divalent cation coordination site at its surface (Lee et al., 1995). This site has been called 'MIDAS', for metal ion-dependent adhesion site. One of the coordination sites of the MIDAS divalent cation may be provided by an acidic residue as found within key regions of most integrin ligands (Bergelson and Hemler, 1995).

Integrin β subunits (with the exception of $\beta 4$) are composed of ~730-800 amino acids and are ~90-130 kDa in size. Many of the 48-56 cysteine residues conserved among all β subunits are within four repeating motifs in a 'cysteine-rich' region. A conserved region in the amino-terminal half of the β subunit has some resemblance to the MIDAS motif found in I domains. As in I domains, β subunit MIDAS-like regions also contain amino acids essential for ligand and divalent cation binding (Takada et al., 1992). However, this region in the β subunit has not been shown to function as a discrete unit and may also differ from I-domain MIDAS regions in other key respects (Lin et al., 1997).

Cytoplasmic tails of integrin $\beta 1$, $\beta 3$, $\beta 5$, $\beta 6$, and $\beta 7$ subunits have 'NPXY' sequences that may be important for integrin localisation, endocytosis, and signalling (Reszka et al., 1992; Van Nhieu et al., 1996). Also, cytoplasmic domains of $\beta 1$ can be alternatively spliced, yielding $\beta 1A$, $\beta 1B$, $\beta 1C$, and $\beta 1D$ isoforms. The $\beta 1A$ form is expressed ubiquitously and contains all of the sequences needed for localisation into focal adhesions (see below) (Reszka et al., 1992). The $\beta 1B$ variant, expressed in skin and liver, fails to localise into focal adhesions because key residues have been replaced (Balzac et al., 1993). The $\beta 1C$ isoform may regulate cell proliferation. It correlates with a benign, non-proliferative phenotype in epithelial cells in vivo and is also found in platelets, and other cell lines (Meredith et al., 1995). The $\beta 1D$ form is found in cardiac and skeletal muscle, where it is regulated during myoblast differentiation (Belkin et al., 1996). The cytoplasmic tails of both α and β subunits are usually highly conserved between species, consistent with their having critical functions.

Several studies have shown that single subunit chimerae containing integrin β cytoplasmic domains can mimic or interfere (depending on the expression level) with a number of integrin receptor functions such as FAK phosphorylation (Akiyama et al., 1994),

inside-out activation (Chen et al., 1994) and actin association (LaFlamme et al., 1994), suggesting that the activity of the β cytoplasmic domain must be regulated upon ligand binding. Deletion of the α cytoplasmic domain in two integrins, $\alpha 1\beta 1$ (Briesewitz et al., 1993) and $\alpha IIb\beta 3$ (Ylanne et al., 1993) resulted in constitutive recruitment of these mutant integrins to focal adhesions, independently of ligand binding.

The notion that the α cytoplasmic domain can regulate the ligand dependency of integrin function suggests that the α and β cytoplasmic domains may be associated. This hypothesis is supported by experiments in which exchange of the intracellular domains of $\alpha 1$ and $\beta 1$ subunits led to a complete lack of either cytoskeletal organisation or phosphotyrosine signalling, despite intact ligand binding ability (Briesewitz et al., 1995). In addition, insertion of an α -helix-promoting spacer in the β cytoplasmic domain of the swapped $\alpha 1\beta 1$ receptor led to a complete rescue, while replacement of the spacer with a sequence predicted to be less α -helical resulted in only a partial rescue, leading to a model of α -helix propagation being proposed for integrin outside-in signalling (figure 1.5) (Marcantonio and David, 1997). This model proposes that the two membrane proximal regions maintain integrin latency by prevention of the formation of α -helical structures within the integrin β cytoplasmic domains. Upon ligand binding, the relationship of the cytoplasmic domains is altered, leading to either permissive or active promotion of α -helical structure in distal regions of the β cytoplasmic domain. It is the formation of this helical structure that has been proposed to be critical for interaction with downstream elements and integrin mediated signalling (Marcantonio and David, 1997).

Focal adhesions

A key feature of integrin signalling involves the organisation of integrins into focal adhesion complexes. FAs are specialised sites of adhesion developed by many cell types in culture (Burrige and Chrzanowska-Wodnicka, 1996). They consist of aggregated ECM receptors (integrins) interacting on the outside with ECM components and on the inside with bundles of actin filaments (stress fibers) and thus represent a model system for studying cell-ECM adhesion. A list of structural proteins identified at focal adhesions is given in table 1.2. The majority of these proteins are multidomain molecules capable of interacting with several ligands (Figure 1.8*a*). The number of potential connections and the fact that some can be regulated by phosphorylation, the state of integrin aggregation or occupancy, and signalling molecules (e.g. PIP2, see below) suggests the likelihood of variable compositions of adhesion complex.

The cytoskeletal elements of FAs are part of the microfilament system. Their relative abundance varies considerably; paxillin, tensin, zyxin and many signalling molecules are minor components compared with vinculin and talin. The picture is further complicated by the finding that in a number of cases, such interactions depend on conformational changes in the structural components themselves, changes which appear to be under the control of signalling pathways (see p.15).

A large number of FA proteins bind actin (table 1.2). This observation hints at redundancy in the linkage of actin filaments to integrins although it has been proposed that some of the actin-binding proteins may function at different stages in the life of a FA (Burrige and Chrzanowska-Wodnicka, 1996). For example, talin and vinculin may be involved in FA formation, whereas α -actinin may be more important in maintaining or

stabilising microfilament attachment in mature FAs (Nuckolls et al., 1992). Multiple linkages may also facilitate actin polymerisation, permitting addition of actin monomers while simultaneously maintaining attachment to the membrane. Some FA proteins, such as tensin and talin, have multiple actin-binding sites, making it possible to cross-link actin filaments and thus stabilise attachments to the membrane.

In addition to a role in cell adhesion to the ECM, FAs play a role in transducing extracellular and intracellular signals, and appear to provide a point of convergence of growth factor and integrin signalling (Rozengurt, 1995).

Signal transduction events

Ligation of integrins results in a number of intracellular events, including tyrosine phosphorylation of intracellular proteins, activation of PKC and MAPK, pH alterations, elevation of intracellular Ca^{2+} concentrations, and changes in lipid metabolism (Parsons, 1996). A diagram of the major integrin-mediated signal transduction pathways is given in figure 1.6. Integrins possess no intrinsic tyrosine kinase activity and must therefore signal via cytoplasmic kinases, such as the Src family kinases and focal adhesion kinase (FAK). Through these tyrosine kinases, integrins activate the GTPase Ras, leading to the activation MAPK cascade which regulates immediate-early gene expression (Giancotti, 1997). In addition to activating tyrosine kinase pathways, leading to the activation of Ras, integrins also affect the Rho family GTPases Rac and Rho. These proteins have been implicated in regulating the cytoskeleton as well as cooperating with Ras to activate immediate-early genes (Vojtek and Cooper, 1995). Finally, it is becoming increasingly evident that the

integrin-dependent pathways which regulate cell cycle progression and those which regulate the cytoskeleton do not function independently, but 'cross-talk' extensively.

A major regulator of FAs and stress fibers is the small GTP-binding protein Rho, which is a member of the Ras superfamily of proteins (Hall 1994). In addition to other functions (reviewed in Machesky and Hall, 1996 - see figure 1.7), Rho regulates FAs and stress fibers (Ridley and Hall 1992). Several potential targets for Rho have been identified, including phosphatidylinositol 5-kinase (PIP5K). Activated Rho stimulates phosphatidylinositol 4,5-bisphosphate (PIP₂) synthesis from its precursor PIP, a reaction catalysed by PIP5K, resulting in elevated PIP₂ levels in response to cell adhesion (Chong et al., 1994).

PIP₂ acts on several cytoskeletal proteins, promoting actin polymerisation and enhancing FA formation. The atomic structure of gelsolin indicates that polyphosphoinositides compete with F-actin for binding to overlapping, solvent-exposed sites on plasma gelsolin (Burtnick et al., 1997). Acting on profilin and gelsolin, PIP₂ stimulates dissociation of these proteins from actin, thereby promoting actin polymerisation. PIP₂ has also been reported to enhance the binding of α -actinin to actin (Fukami et al., 1992) and to affect the conformation of vinculin (see p.15). Antibodies to PIP₂ block the assembly of FAs in mouse fibroblasts (Gilmore and Burridge, 1996) underlining the importance of this signalling molecule in FA assembly.

Vinculin

Vinculin is a 117 kDa microfilament-associated protein found in higher eukaryotes and located at the cytoplasmic aspects of cell adhesion sites. Originally isolated from

chicken gizzard extract as a by-product during the purification of α -actinin (Geiger, 1979), vinculin was subsequently shown to be ubiquitously associated with both cell-cell and cell-ECM adhesions (Geiger et al., 1980) in a wide variety of cell types. Consequently, vinculin is commonly used as a hallmark for these families of adhesion sites.

Microinjection of fluorescently tagged protein into cells, indicates that vinculin is associated with two cytoplasmic pools; a diffusible fraction and a junctional membrane-associated fraction, which are in a dynamic equilibrium (Burrige and Feramisco, 1980). Like a number of other cytoskeletal molecules, vinculin requires both integrin clustering and integrin occupancy by a ligand to accumulate in FAs. Electron microscopic studies on the assembly of nascent FAs in fibroblasts have shown that, talin precedes the visual accumulation of vinculin into F-actin-rich foci (De Pasquale and Izzard 1987, 1991).

From nematodes to humans, only one vinculin gene has been identified (Weller et al., 1990). Sequence conservation in all vinculins is high (Weller et al., 1990) and sequence comparisons reveal three extended stretches of homology with α -catenin (Herrenknecht et al., 1991), located for the most part at the N- and C-terminal regions of vinculin (figure 1.9), which contain the major established binding sites of the molecule (see p.12).

Based on the sequence homology between the two proteins and on the fact that vinculin forms homo-oligomers (see below), heterodimer formation between vinculin and α -catenin has been proposed. Binding of the C-terminal region of α -catenin to the vinculin N-terminal domain has been demonstrated *in vitro* (Weiss et al., 1998). However, binding of the N-terminal region of α -catenin to the vinculin C-terminal domain and formation of vinculin/ α -catenin heterodimers *in vivo* remain to be observed.

Domain structure

Vinculin is a multi-ligand protein that interacts *in vitro* with a variety of structural and regulatory focal adhesion proteins. These binding partners are listed in table 1.3 and a schematic representation of the domain structure of vinculin is given in figure 1.10.

Electron microscopic images of chicken vinculin (figure 1.11) show a "head" 80 Å in diameter, consisting of three globular centres of mass, surrounding a protein-deficient centre, and of a short, stem-like fragment, which is highly flexible, and serves to link the head to the "tail" (Winkler et al., 1996). The "tail" contains four spherical protein masses arranged like pearls on a string. Other electron microscopic images of intact pig vinculin show a globular structure with no "tail" evidence (Gimona et al., 1987). The chicken vinculin images correlate with biochemical studies showing that cleavage by V8-protease within a proline-rich region (see Figure 1.9) generates a 95 kDa globular N-terminal head (Vh) and a 22 kDa C-terminal tail (Vt) (Price et al., 1989). Vh contains binding sites for the cytoskeletal proteins talin (Gilmore et al., 1992; Jones et al., 1989), tensin (Wilkins et al., 1987), α -actinin (Kroemker et al., 1994), α -catenin (Weiss et al., 1998) and possibly β -catenin (Hazan et al., 1997), whereas Vt binds paxillin (Wood et al., 1994), F-actin (Huttelmaier et al., 1997; Menkel et al., 1994) and acidic phospholipids (Johnson et al., 1998). The proline-rich region contains distinct binding sites for VASP (Brindle et al., 1996), vinexin (Kioka et al., 1999) and ponsin (Mandai et al., 1999).

The binding sites for talin (Johnson and Craig, 1994), α -actinin (Kroemker et al., 1994), VASP (Huttelmaier et al., 1998) and F-actin (Johnson and Craig, 1995) are blocked by an intramolecular interaction between Vh and Vt. Acidic phospholipids, including PIP₂, have been shown to relieve this interaction and induce binding for talin, α -actinin and

VASP. In the case of F-actin, the role of PIP2 with respect to F-actin/vinculin binding appears controversial (see p.16).

Gene disruption

Vinculin-null mice die in early embryogenesis, suggesting that vinculin is essential for the complex array of cell contacts and movements that occur during early development (Xu et al., 1998). Transfection of antisense vinculin cDNA into fibroblasts can produce cells expressing decreased levels of vinculin. Cells expressing up to a third of vinculin control levels, display a phenotype characteristic of transformed cells, i.e. a rounded shape, decreased adhesion and number of stress fibers, increased motility, and anchorage independent growth. Dramatic decreases in vinculin levels have been observed in SV40-transformed cells, and a spontaneous rat adenocarcinoma has been found to lack any detectable vinculin. Transfecting full length chick vinculin cDNA into these two cell lines led to loss of the characteristic transformed phenotype, with the cells becoming unable to form tumours when injected into syngeneic animals and nude mice (Fernandez et al., 1993). The above results led to vinculin being characterised as a tumour suppressor and indicate that modulation of vinculin expression may constitute an effective control mechanism for suppressing or allowing cell locomotion. However, the molecular mechanisms which link the changes induced in vinculin levels with altered cell shape and motility, and alterations in anchorage dependent growth remain to be defined.

Isoforms and post-translational modifications

Smooth and cardiac muscle synthesise two alternative splice isoforms, vinculin and metavinculin. Metavinculin, which has also been identified in platelets (Turner and Burridge, 1989), contains an additional sequence stretch of 68 residues as an insert (Gimona et al., 1988) near the proline-rich region, between Vh and Vt. The tail domain of metavinculin (MVt) has been shown to display a different phenotype to that of Vt in interacting with actin (Rudiger et al., 1998); while Vt causes F-actin needles or coils, MVt-expressing cells display a diffuse F-actin distribution. Thus it appears that the MVt-specific insert induces an F-actin supraorganisation different from the Vt-based form, suggesting that metavinculin has a specific role in muscle, although the exact functional significance of this isoform heterogeneity remains to be determined.

Phosphorylation of the intact vinculin occurs in the presence of pp60src (Ito et al., 1983) and PKC α (Werth et al., 1993), a Ca²⁺ and phospholipid-dependent serine/threonine kinase. In mammalian fibroblasts, three vinculin iso-electric variants have been identified that may be the result of different phosphorylation states. Focal adhesions appear to contain primarily the most acidic form of vinculin (Geiger 1982) and phosphorylation of vinculin is dependent on cells being adherent to a surface and on actin polymerisation (Hagmann and Burger, 1992), indicating that phosphorylation plays a role in anchoring vinculin at sites of microfilament-membrane interaction. In addition, the phosphorylation of vinculin by PKC appears to be regulated by the head-tail intramolecular interaction. While intact vinculin and the isolated head domain are only weakly labelled, the isolated tail fragment is much more strongly phosphorylated. Conversely, in the presence of the tail, the head is fully protected from PKC (Schwienbacher et al., 1996; Weekes et al., 1996). These data are consistent with

observations that native vinculin is primarily phosphorylated within the tail domain and suggest a function for vinculin phosphorylation in the regulation of the vinculin conformation.

In normal cells, the level of phosphotyrosine in vinculin is minimal. However, two of the proteins that become tyrosine phosphorylated in response to adhesion, paxillin and tensin, bind vinculin. The functional significance of these phosphorylations remains to be tested.

Other post-translational modifications of vinculin include palmitoylation and myristoylation reactions (Burn and Berger, 1987). Transformation of chick embryo fibroblasts by the Rous sarcoma virus leads to a specific reduction in the acylation of vinculin by these lipids, an observation that could be directly or indirectly linked with the dissociation of vinculin from the plasma membrane and could thus account for the phenotypic changes characteristic of the transformed state.

Interactions with acidic phospholipids

Vinculin immunoprecipitated from adherent cells contains PIP₂, but not other phospholipids, consistent with PIP₂ being the physiological effector of the conformational change in this protein (Gilmore and Burridge 1996). Acidic phospholipid binding regions have been mapped to residues 916-970 (Johnson et al., 1998) and 1012-1066 (Tempel et al., 1995). Using expressed peptides, residues 916-970 have been shown to contain two binding sites for PIP₂ (913-940 and 940-970), with hydrophobic photolabelling indicating that both sites are necessary for insertion into the lipid bilayer (Johnson et al., 1998). How many of these binding sites are occupied in the intact protein, remains to be demonstrated in vivo.

Binding to PIP2 and other acidic phospholipids has been shown by chemical cross-linking to induce oligomerisation of Vt (Huttelmaier et al., 1998) (see below) indicating a model for the interactions of vinculin at FA sites (figure 1.8b).

Interactions with actin

From the early 1980s to mid-1990s, there was controversy over whether vinculin interacted directly with actin. Vinculin was shown to bind to actin-capping (tensin) and actin cross-linking proteins, such as talin and α -actinin, but did not bind to F-actin, leading to the hypothesis that vinculin acted as a linker between other prominent FA proteins. However, studies of bacterially expressed vinculin fragments, indicated that the C-terminal tail of vinculin binds F-actin (Menkel et al., 1994).

These apparently conflicting results were resolved when the actin-binding site was found to be masked by the intramolecular head-tail interaction in the intact molecule, and relieved by PIP2 (Weekes et al., 1996; Gilmore and Burridge, 1996). Actin-binding sites in vinculin were subsequently mapped to two regions of Vt, residues 893-985 and 1016-1066 (Huttelmaier et al., 1997). Both sites were shown to be necessary to cross-link or bundle actin filaments as demonstrated by low shear viscometry, and both actin-binding sequences were shown to be capable of oligomer formation (see below).

Subsequent work contradicted these results, indicating that PIP2 inhibits the interaction of Vt with F-actin (Steimle et al., 1999). The hypothesis subsequently put forward was that activation of the actin-binding activity of vinculin requires steps other than or in addition to the binding of PIP2.

Evidence for oligomerisation

Vt has been shown by chemical cross-linking to form oligomers in the presence of F-actin (Huttelmaier et al., 1997) or acidic phospholipids (Huttelmaier et al., 1998). Gel filtration studies indicate that, in the absence of these ligands, Vt remains monomeric. However, residues 916-970 and 1012-1066 migrate as dimeric species when expressed as separate fusion peptides (Johnson and Craig, 1998).

Electron microscopy (Milam, 1985) and hydrodynamic calculations (Eimer et al., 1993) of the intact vinculin, indicate that oligomerisation of the whole molecule also occurs at the Vt regions, resulting in “parachute”-like aggregates. However, other EM studies have not come up with evidence in support of this occurrence (Winkler et al., 1996) and no vinculin oligomers have been observed in vivo so far.

Summary

Cell adhesion plays a pivotal role in both normal and pathological processes and requires a dynamic interaction between the interior of the cell and its external environment. This interaction is mediated by transmembrane cell surface receptors, of which cadherins and integrins are prime representatives. Cadherins mediate cell-cell adhesion and their activation is Ca²⁺-dependent, while integrins mediate cell-ECM adhesion and can also be dependent on divalent cations for activation. Activation of both cadherins and integrins leads to recruitment of many cytoskeletal and signal transduction molecules and the formation of elaborate three-dimensional networks of varying compositions. These adhesion complexes can serve as both recipients and generators of signalling information.

FAs provide an *in vitro* model for the study of integrin-mediated adhesion. The cytoskeletal components of FAs are part of the microfilament system and are for the most part multidomain proteins capable of interacting with several ligands. A redundancy is observed in the linkage of actin filaments to integrins in FAs, which serves to facilitate actin polymerisation and stabilise the attachments of actin filaments to the membrane. Basic physical mechanisms that underlie the regulation of FA formation and function include aggregation of individual signalling molecules and conformational changes that lead to altered accessibility of key molecular binding sites. PIP2 is a prominent signalling molecule implicated in promoting actin polymerisation and enhancing FA formation. In addition, FA formation is regulated by tyrosine phosphorylation and other protein phosphorylation events.

Vinculin is a 117 kDa microfilament-associated protein involved in both cell-cell and cell-ECM adhesion. *In vivo*, a dynamic equilibrium exists between diffusible and membrane-associated vinculin fractions with phosphorylation thought to play a role in anchoring vinculin at the sites of microfilament-membrane-interaction. Electron microscopic images of vinculin distinguish three regions: a trilobar "head" (Vh) with a protein deficient centre, a short hinge region, and a long flexible "tail" (Vt). Phosphorylation of vinculin and binding to a number of its ligands is blocked by an intramolecular interaction between Vh and Vt, an interaction relieved by acidic phospholipids. Immunoprecipitated vinculin contains PIP2 and PIP2 has been shown to control the activation of vinculin *in vitro*, and to induce oligomerisation of isolated Vt. Vinculin binds F-actin directly, and indirectly through a number of F-actin binding proteins such as talin, α -actinin and tensin. F-actin can also induce oligomerisation of the vinculin tail domain and its binding site overlaps with the PIP2 binding site in Vt. A schematic representation of the domain organisation of vinculin is

given in figure 1.10 and two proposed models for its interaction at the sites of FAs are presented in figures 1.8*a* and *b*.

Scope of thesis

The aim of this research was to determine the crystal structure of the vinculin tail domain (Vt) by X-ray crystallography. Initial and optimised purification and crystallisation conditions of Vt are presented in chapter 2. The details of the structure determination are discussed in chapter 3, and the structure itself is described in chapter 4. The ultimate aim of the structure determination, was to provide a better understanding of the activity and regulation of Vt in the context of both the full-length molecule and of its interacting partners at the sites of cell adhesion. To this effect, a series of structure-based solution experiments were carried out as presented in chapter 5. When reviewed in concert with the Vt structure, these results led to the formulation of a hypothesis for vinculin activation which is outlined in chapter 6. Finally, crystallisation of the full-length vinculin (Vg) and characterisation of Vg crystals is presented in chapter 7. It is hoped that the work described in this thesis will stimulate efforts to further elucidate the nature of cell adhesion at the molecular level.

	$\beta 1$	$\beta 2$	$\beta 3$	$\beta 4$	$\beta 5$	$\beta 6$	$\beta 7$	$\beta 8$
$\alpha 1$	Coll, Lm							
$\alpha 2$	Coll, Lm							
$\alpha 3$	Lm							
$\alpha 4$	VCAM-1 Fn						MAdCAM-1, VCAM-1, Fn	
αE							E-Cadherin	
$\alpha 5$	Fn							
$\alpha 6$	Lm Fertilin			Lm				
$\alpha 7$	Lm							
$\alpha 8$	Fn, Vn, Tn							
$\alpha 9$	Tn							
αV	Fn, Vn, Opn, Fb		Fb, Vn, Opn, vWF, Tn		Vn (Opn, Fn)	Fn, Tn		Coll, Lm, (Vn)
αIIb			Fb, Vn, Fn, vWF					
αL		ICAM-1, 2, 3						
αM		FX, Fb, ICAM-1						
αX		ICAM-1, Fb						
αD		ICAM-3						

Table 1.1. Integrin α - β subunit combinations and ligands. Abbreviations used are : Coll, collagen; Fb, fibrinogen; Fn, fibronectin; FX, factor X; ICAM-1, 2, 3, intracellular adhesion molecule-1, 2, 3; Lm, laminin; MAdCAM-1, mucosal addressin cell adhesion molecule-1; Opn, osteopontin; Tn, tenascin, VCAM-1, vascular cell adhesion molecule-1; Vn, vitronectin; vWF, von Willebrand's factor. Empty boxes indicate $\alpha\beta$ subunit combinations that have not been observed. Many additional ligands have been proposed, but often these are not as well established, or interact more weakly, compared to the ones listed (reproduced from Kreis and Vale, 1999).

Transmembrane	Cytoplasmic	Actin- Binding	Regulated Binding
Integrins	actin	–	nd
Syndecan IV	α -actinin	+	+, PIP2 ³
	filamin	–	nd
	fimbrin	+	nd
	gelsolin	+	–, PIP2 ⁴
	ponsin ¹	nd	nd
	profilin	+	–, PIP2 ⁵
	radixin	+	nd
	talin	+	nd
	tensin	+	nd
	VASP	+	nd
	vinculin	+	+/-, PIP2 ^{6,7,8}
	vinexin ²	nd	nd
	zyxin	–	nd

Table 1.2. Common structural FA components. The third column indicates actin-binding cytoplasmic proteins; the fourth column indicates proteins for which some binding functions are regulated; + indicates an increase in actin-binding, – a decrease. In the case of vinculin, there is conflicting evidence as to the role of PIP2 in binding actin. Adapted from Burridge and Chrzanowska-Wodnicka (1996). 1. Mandai et al., 1999; 2. Kioka et al., 1999; 3. Fukami et al., 1992; 4. Burtnick et al., 1997; 5. Pollard et al., 1994; 6. Weekes et al., 1996; 7. Gilmore and Burridge, 1996 8. Steimle et al., 1999; *VASP* = vasodilator-stimulated phosphoprotein; *nd* = none determined.

Ligand	Kd (mM)	Vinculin site	Regulated
F-Actin	~ 1	893-985, 1016-1066	+/-, by PIP2
α -Actinin	13	1-107	+, by PIP2
α -Catenin	0.2-0.4	1-258	nd
β -Catenin	nd	nd	nd
Paxillin	0.06	979-1028	No
PIP2	0.03	916-970, 1012-1066	No
Ponsin	nd	839-850	nd
Talin	0.4	1-208, 1-258	+, by PIP2
Tensin	nd	1-258	nd
VASP	nd	839-850	+, by PIP2
Vinculin	(0.05)	1-258, 1029-1036	+, by PIP2
Vinexin	nd	839-850	nd

Table 1.3. Vinculin binding proteins and ligands. Binding sites have been determined using deletion mutants. Regulation studies have been determined using intact vinculin. The head-tail association constant of vinculin (in brackets) has been measured using separate domains of the molecule and is therefore not indicative of the intramolecular association constant. Binding of β -catenin to vinculin has been postulated but remains to be established. See text for references and details.

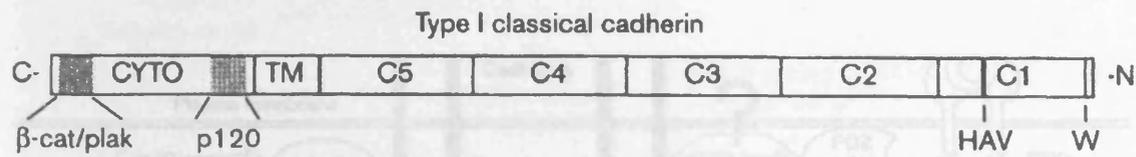


Figure 1.1. Schematic representation of the domain structure of a typical type I classical cadherin. The primary sequence depicts the carboxy-terminal cytodomain (CYTO) with its p120 and α -catenin/plakoglobin sites (in grey), the transmembrane region (TM) and subdomains C1-C5 of the ectodomain. Also depicted are the HAV and Trp2 (W) sequences in subdomain C1. Figures 1.1-1.3 are reproduced from Steinberg and McNutt, 1999.

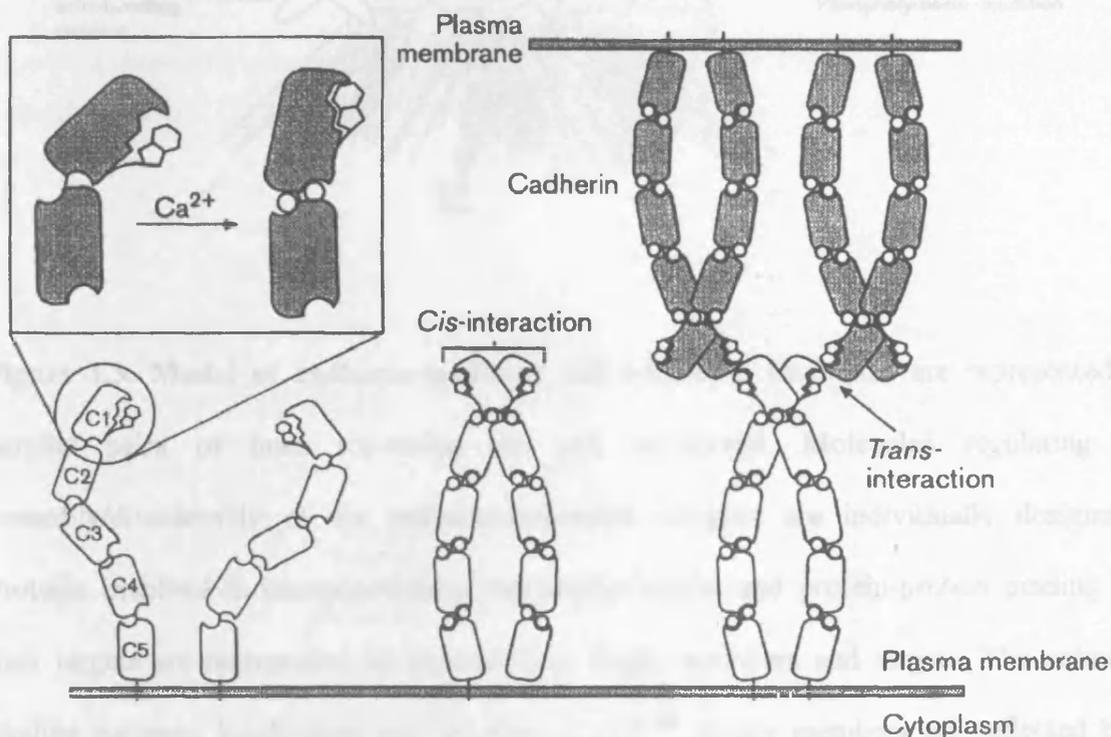


Figure 1.2. Model of cadherin rearrangements during cell adhesion. In the absence of Ca^{2+} (circles), the cadherin structure is disorganised and incapable of participating in adhesion. With increasing Ca^{2+} concentrations, the C1-C5 subdomains become organised (C1, which has the lowest affinity for calcium, becomes organised last) and the cadherin ectodomain becomes rigid and competent to participate in *cis*-dimerisation. The HAV-W structure is essential for *trans*-dimerisation but not *cis*-dimerisation.

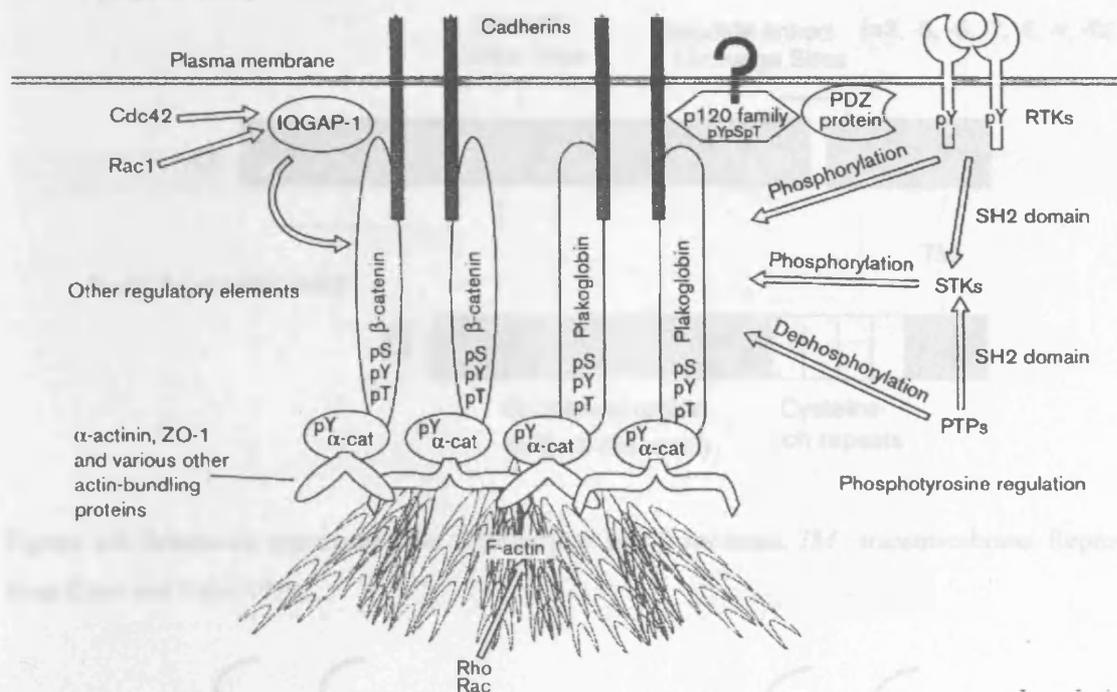


Figure 1.3. Model of cadherin-mediated cell adhesion. Cadherins are represented by parallel pairs of lines traversing the cell membrane. Molecules regulating the assembly/disassembly of the cadherin-associated complex are individually designated. Proteins involved in phosphorylation, dephosphorylation and protein-protein binding and their targets are represented by characteristic family members and targets. The unknown binding partners, localisation and function of p120^{cas} family members are reflected by a question mark. *α-cat*: α-catenin; *pX*: phosphoamino acid. *STK*: serine threonine kinase.

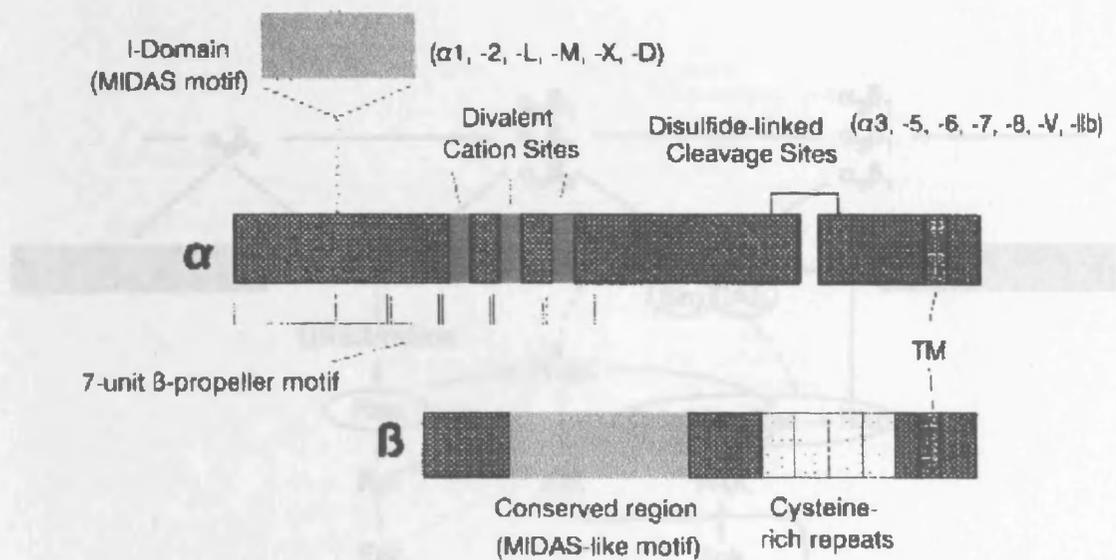


Figure 1.4. Schematic representation of integrin α and β subunits. *TM* : transmembrane. Reproduced from Kreis and Vale (1999).

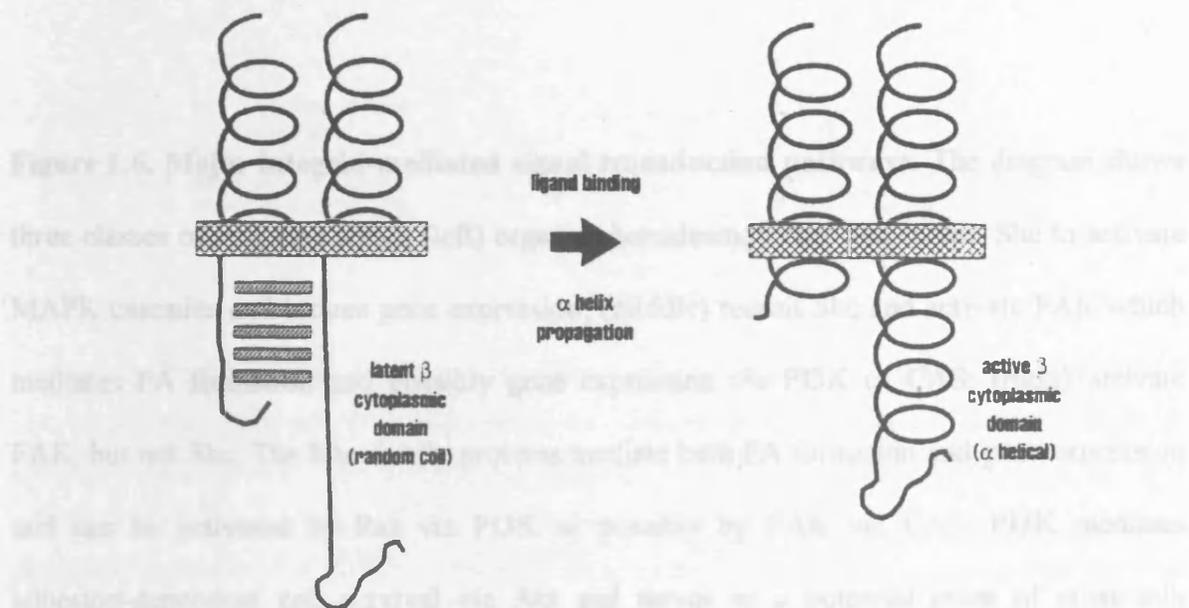


Figure 1.5. Proposed model for integrin outside-in signalling. In the absence of ligand, the conserved membrane proximal regions of the α and β cytoplasmic domains interact via their side chains and keep the proximal β cytoplasmic domain in an unfolded, random coil state. Upon ligand binding, the relationship between the two cytoplasmic domains is altered, possibly by a sliding of the α chain. This conformational change disrupts the association of the two chains, allowing an α -helix to be propagated from the β transmembrane domain through the proximal β cytoplasmic domain. Reproduced from Marcantonio and David, 1997.

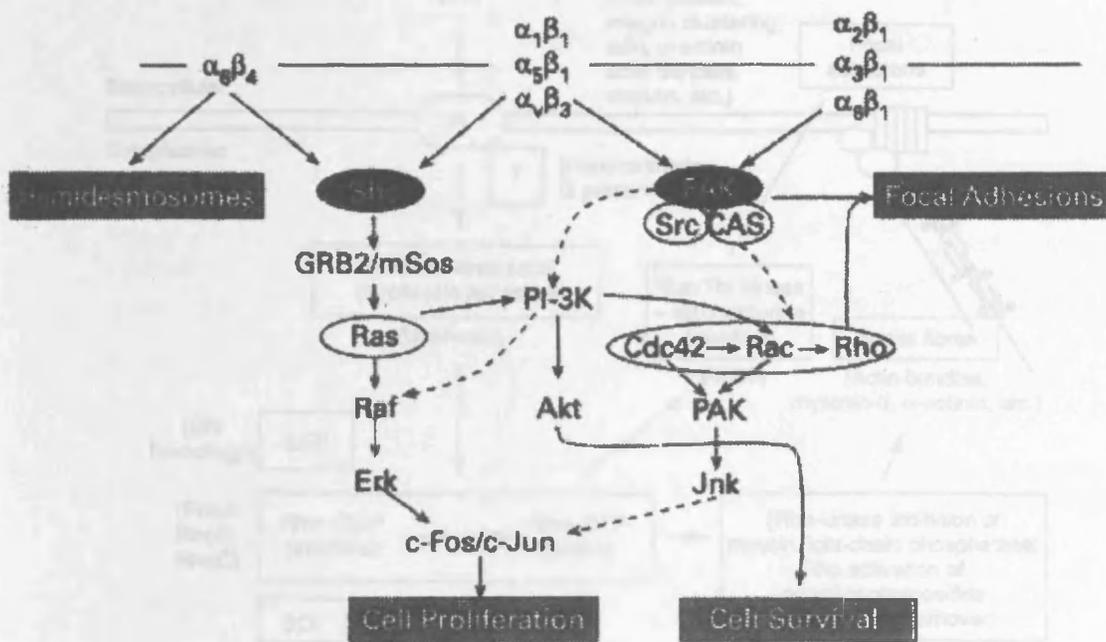


Figure 1.6. Major integrin-mediated signal transduction pathways. The diagram shows three classes of integrins which (left) organise hemidesmosomes and recruit Shc to activate MAPK cascades and induce gene expression; (middle) recruit Shc and activate FAK which mediates FA formation and possibly gene expression via PI3K or CAS; (right) activate FAK, but not Shc. The Rho family proteins mediate both FA formation and gene expression and can be activated by Ras via PI3K or possibly by FAK via CAS. PI3K mediates adhesion-dependent cell survival via Akt and serves as a potential point of cross talk between the Shc and FAK pathways. Reproduced from Kreis and Vale, 1999.

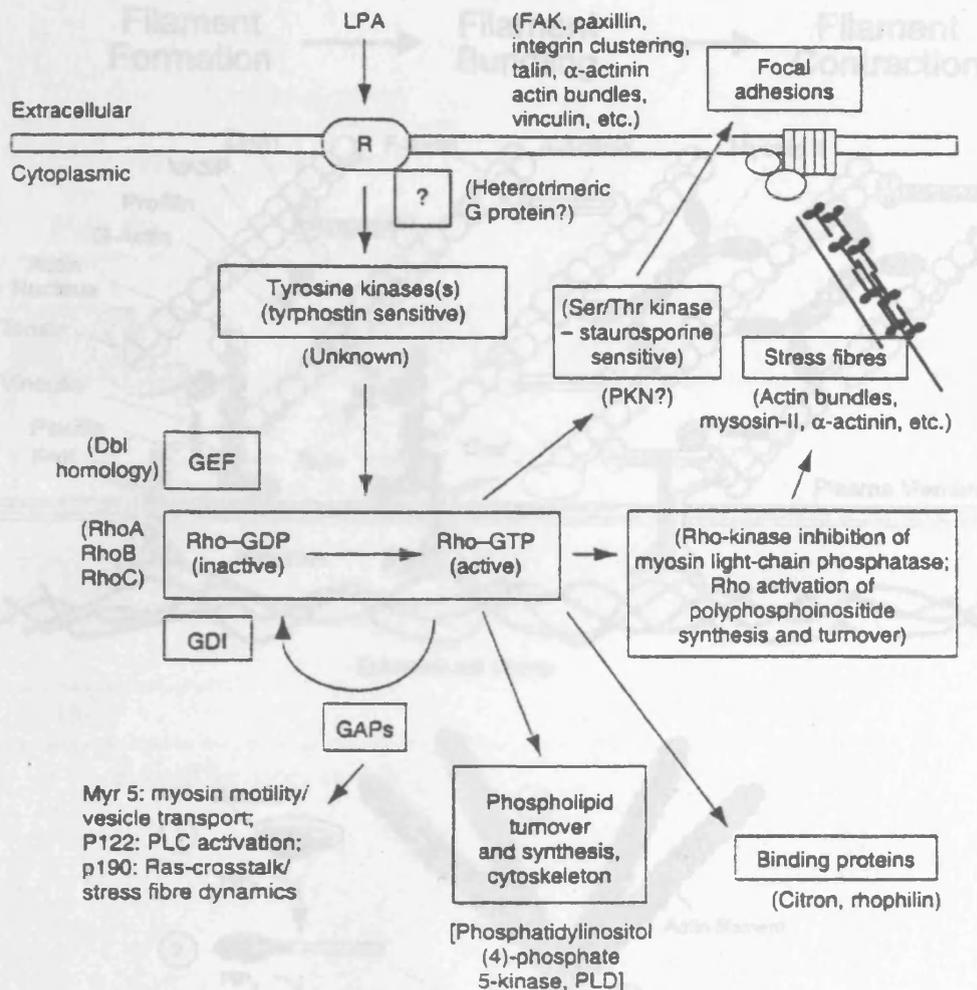


Figure 1.7. Model for Rho-mediated signalling. Lysophosphatidic acid (LPA), a growth-factor-like water-soluble phospholipid, is produced by platelets during aggregation and is thought to activate a heterotrimeric G-protein-coupled receptor. Rho is activated in cells by LPA, and this activation is sensitive to tyrphostin, an inhibitor of tyrosine kinases. Following receptor activation, Dbl-family proteins act as GTPase exchange factors for Rho and cause it to exchange GDP for GTP and release it from GDP-dissociation inhibitor (GDI), a cytosolic inhibitor protein. Once bound to GTP, Rho is active and binds to effectors, causing activation of phospholipid synthesis and turnover, kinase activation and the formation of stress fibres and FA complexes. Rho returns to the inactive state by interaction with GTPase-activating proteins (GAPs), which may also localise Rho or cause downstream effects themselves. Following hydrolysis of GTP to GDP, catalysed by GAPs, Rho rebinds to GDI and is inactive until a further stimulus restarts the cycle. FAK, focal-adhesion kinase; GEF, GDP-GTP exchange factor; PKN, protein kinase N; PLD, phospholipase D; R, transmembrane receptor for LPA. Reproduced from Machesky and Hall, 1996.

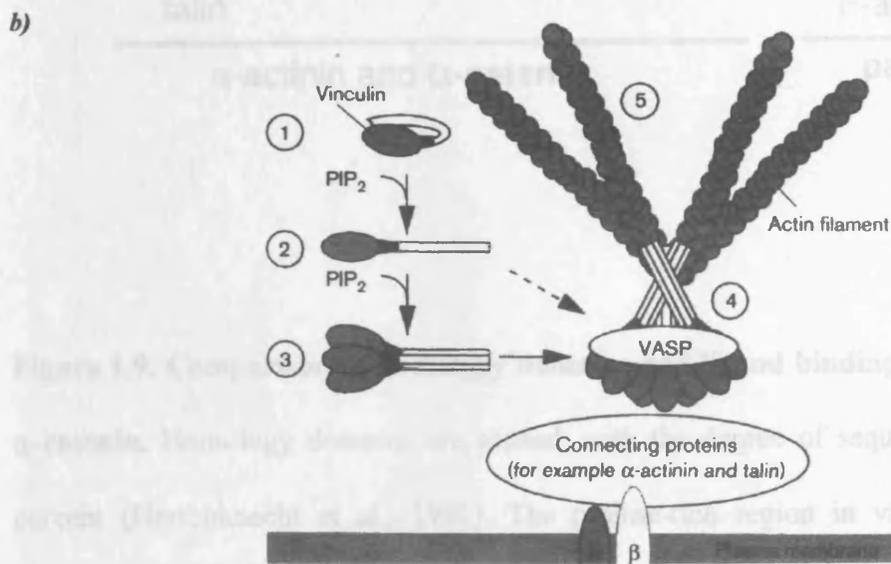
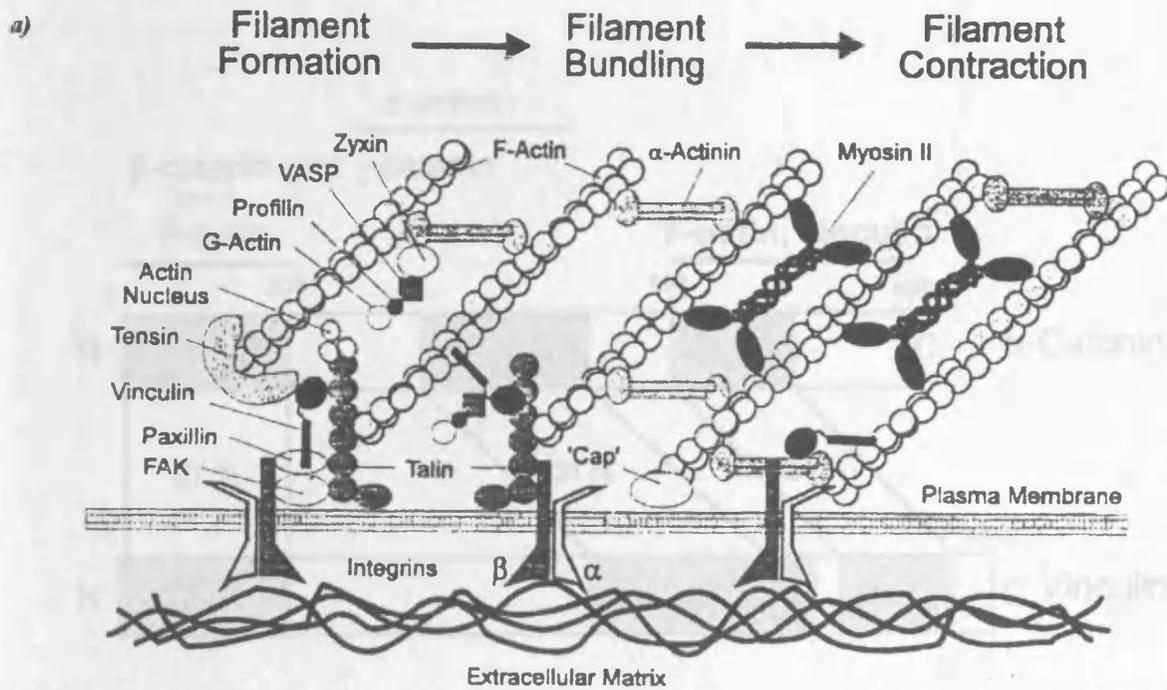


Figure 1.8. Proposed models for the interaction of vinculin at the sites of focal adhesions. *a)* During filament formation, Vt is bound to paxillin, and Vh to tensin and talin; during filament bundling, binding of VASP causes Vt to bind F-actin; during filament contraction Vh switches to binding α -actinin (reproduced from Critchley et al., 1991). *b)* The cytoplasmic, closed form of vinculin (1) is opened and thus activated by PIP₂ (2) and induced to form oligomers via Vt (3). Concomitantly or subsequently, vinculin oligomers form complexes with VASP (4) with Vts engaged in bundling the distal ends of actin filaments (5), close to the plasma membrane (reproduced from Hüttelmaier et al., 1998).

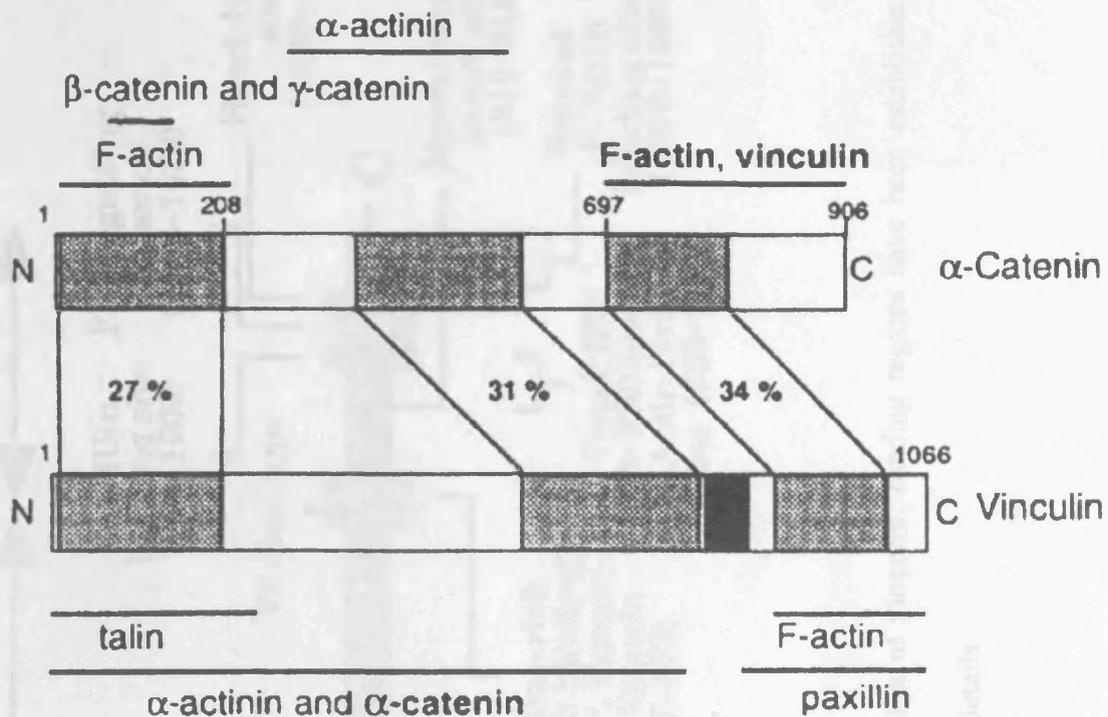


Figure 1.9. Comparison of homology domains and ligand binding sites in vinculin and α -catenin. Homology domains are shaded, with the degree of sequence identity given in percent (Herrenknecht et al., 1991). The proline-rich region in vinculin, which has no correspondence in α -catenin, is shown in black. Ligand-binding sites are indicated (see text for details). Note that the highest degree of identity is found within the C-terminal regions of both molecules. Reproduced from Weiss et al., 1998.

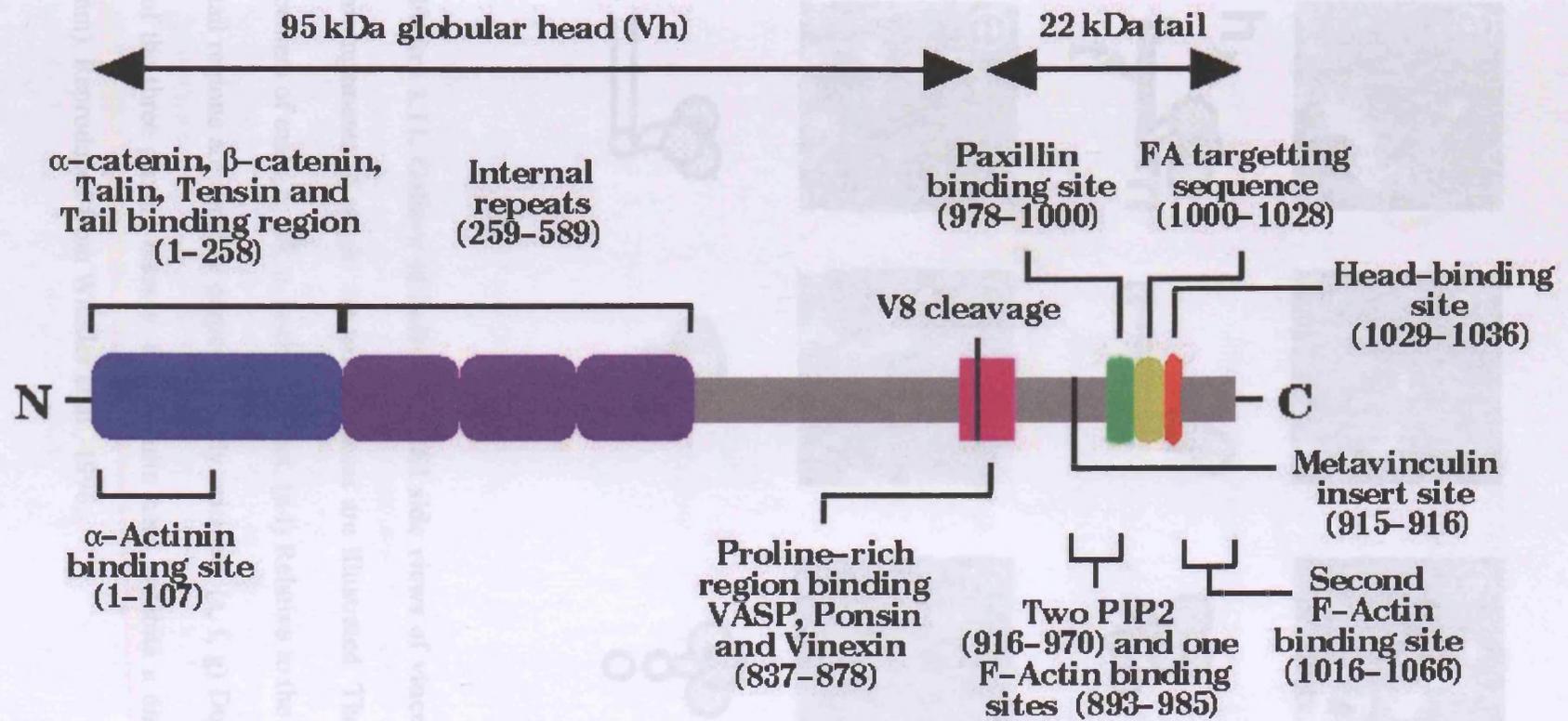


Figure 1.10. Schematic representation of the domain organisation of vinculin. Binding regions have been established using deletion mutants and immunoprecipitation assays. See text for more details.

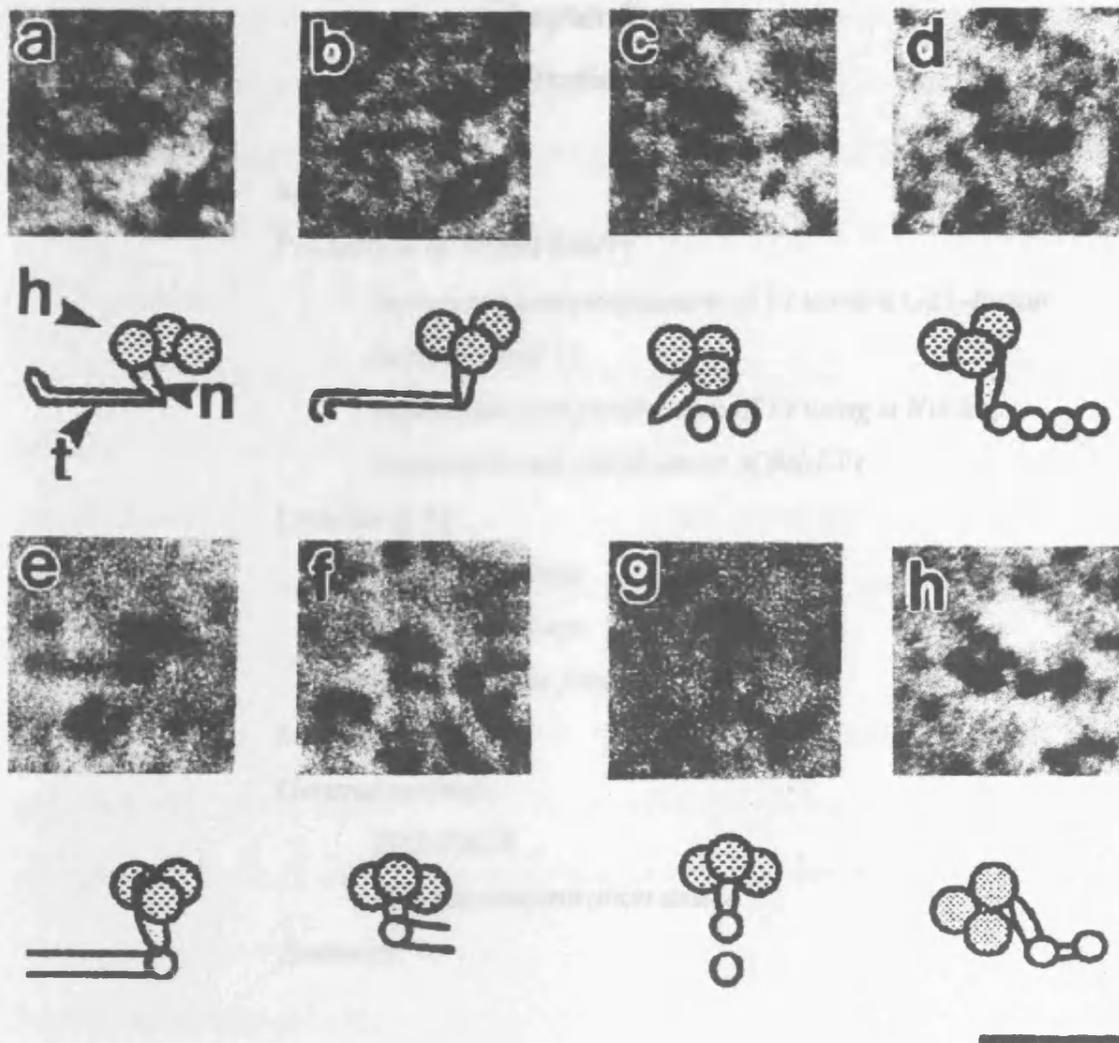


Figure 1.11. Gallery of individual ESI side views of vinculin molecules. The subunit arrangements of single vinculin particles are illustrated. The head (h) consists of three centers of mass, t, tail; n, neck fragment. (a-f) Relative to the neck fragment the head and tail regions are roughly perpendicularly oriented. (a, f, g) Due to the planar arrangement of the three protein masses, the vinculin head exhibits a disc-shaped structure (bar, 20 nm). Reproduced from Winkler et al., 1996.

Chapter 2:

Crystallisation of Vt

Introduction

Production of Vt and SeMVt

Expression and purification of Vt using a GST-fusion

Subcloning of Vt

Expression and purification of Vt using a His-tag

Expression and purification of SeM-Vt

Crystals of Vt

Tetragonal form

Monoclinic form

Orthorhombic form

Materials

General methods

SDS-PAGE

Protein concentration assays

Summary

Introduction

The isolation and characterization of chicken vinculin cDNA was first reported in 1987 by David Critchley's group (Price et al., 1987). This group was also the first to derive the primary sequence and domain structure of this protein (Price et al., 1989). Subsequently, N- and C-terminal domains of chicken vinculin were cloned by the same group into a variety of vectors expressed with or without an affinity tag for binding assays and other cell biology purposes.

By the mid 1990s, the picture surrounding the binding and regulatory properties of vinculin had become complex enough to merit a structural approach of the protein at the atomic level. In October 1995, Bob Liddington began a collaboration with David Critchley aimed at providing such information on a number of cytoskeletal proteins, including vinculin.

I started working on vinculin in September 1996, and was supplied with a construct which expressed Vt as a GST-fusion. Despite the low yields and variable degree of purity of the protein produced by this system, which finally required a change of vector, it was in this form that I determined crystallisation conditions for Vt, two months after my arrival.

This chapter focuses on the methods used to optimize protein yield, purification, and ultimately crystal growth. Problems encountered during all three stages are presented and strategies aimed at resolving them are analysed and discussed.

Production of Vt and SeMVt

Expression and purification of Vt using a GST-fusion

Expression and affinity purification was carried out according to Smith and Johnson (1988). Overnight 100 mL 2xYT/Amp cultures of frozen 20% glycerol stock *E. coli* strain JM109 cells transformed with the recombinant pGEX-2T plasmid were diluted each to OD₆₀₀ = 0.3-0.4 in 1L of fresh 2xYT/Amp medium and grown for 1 h at 37°C before adding IPTG to 0.25 mM. After an induction time of 4-5 h, cells were pelleted (20 minutes, 7000 rpm, 4°C), resuspended in 20% glycerol and frozen at -80°C. Upon thawing, DTT to 2 mM, EDTA pH 8 to 20 mM, PMSF to 0.5 mM, KCl to 100 mM and 5 mL PBS/1% Triton X-100 were added per litre of culture. Cells were lysed by French

press (two passages at 1,000 MPa) and ultracentrifuged for 1 h at 32-36,000 rpm (4°C). Lysates were mixed with immobilised glutathione-agarose beads (Sulphur linkage, Sigma) in a Flex column (Contes) for 20-60 min at RT, washed with PBS/1% Triton X-100, and the GST-Vt fusion released by competitive elution with excess reduced glutathione (30min, RT on a moving rack). The glutathione beads were regenerated by excess washing with PBS. After overnight dialysis in TCB at 4°C, the fusion protein was thrombin-cleaved after a further incubation (3-4 h, RT on a moving rack) and passed several times through the regenerated immobilised glutathione beads before being collected. Yields, calculated by concentration measurements at 280 nm, showed, at this stage, a production of ~ 5mg of Vt per litre of culture.

The protein was concentrated and buffer exchanged in a stirred ultrafiltration cell (Amicon, Beverly, U.S.A.) against 50 mM NaCl, 50 mM sodium acetate pH 5.0 before being purified by ion-exchange on a Mono-S 5 mL column using a 20% (0.05-0.5 M) NaCl gradient. Vt was finally redialyzed into TCB prior to setting up crystallisation trials. The protein at two stages of purification is shown in figure 2.1. The purified protein includes three residues (Gly-Ser-Met) from the thrombin cleavage site at the N-terminus.

The problems encountered during the last purification step were two-fold. First, the yield decreased by >50% with respect to the yield prior to the ion-exchange step. This was due to heavy precipitation occurring during acidification of the protein. In an attempt to avert the loss due to precipitation, purification of Vt at higher values of pH was attempted (pH 6.2, pH 7.0 - data not shown). The rate of precipitation did decrease with this method, but the protein eluted was of lower purity than the one obtained at pH 5.0. Analysis of the precipitate by SDS-PAGE (data not shown) showed this to contain

Vt, GST-Vt and a smear of bacterial protein impurities, indicating that the precipitation step also resulted in an increase in the purity of the Vt remaining in solution. This avenue was therefore not further pursued.

Secondly, attempts at reproducibly growing crystals from the highly purified Vt (figure 2.1) proved unreliable. The quality of the crystals that did grow also varied, with the predominant monoclinic crystals invariably displaying non-isomorphism, high disorder, twinning or a combination thereof (see p.40-41). Attempts at resolving these problems did not meet with success, so a change in the expression system was eventually realised.

Subcloning of Vt

Primer Design

Primers for Vt were designed to include *Nde I*-*BamH I* restriction sites for cloning from pGEX-2T into pET15b. The primers used were:

< BamHI > < NdeI >

5' TGA GAA TTC CAT ATG GAA GAA AAA GAT GAG G3' (VTU)

and

< EcoRI > < BamHI >

5' GCC GAA TTC GGA TCC TTA CTG ATA CCA TGG 3' (VTL)

VTU=upper (5') primer; VTL=lower (3') primer

Each of the primers was ethanol precipitated using two cycles of storage at -80°C for 10 min, followed by centrifugation on a benchtop microfuge (30 min, 13000 rpm,

4°C). The supernatant was discarded, the primers resuspended in deionised water and their concentration determined by absorbance readings of 1:100 dilutions at 260 nm. 100 µL 5 µM stocks were thus made and stored at -20°C.

PCR of Vt cDNA from pGEX-2T

The template (pGEX-2T vector containing the DNA sequence corresponding to Vt) was purified using a Quiagen kit (QIAprep Spin) and the PCR reaction mixture as per manufacturers' instructions. PfuI was from Stratagene and PCR Buffer and AmpliTaq from Perkin Elmer.

The annealing temperature was calculated using the programme OLIGO (<http://bricoh.coh.org/~oligolab/olicalc.html>) to be 57.7° for VTU and 62.9° for VTL. The Vt cDNA was obtained by PCR using a Biometra Trio Thermoblock and 25 cycles of 60 sec at 94° (splitting temperature), 60 sec at 50° (annealing temperature) and 60 sec at 72° (extending temperature). Paraffin oil was overlaid on the sample to avoid evaporation on heating. The presence of the PCR product was confirmed by visualization of a ~600 bp band on a 1.1% agarose gel.

Ligation into pBluescriptSK(-)

Prior to ligation into pET15b, the cDNA encoding Vt was ligated into pBluescriptSK(-) to ensure production of high DNA yields for DNA sequencing.

The Vt PCR product was digested with *EcoR I* as per manufacturers' instructions. The reaction was left to proceed for 3 h at 37°C. Ligation of the *EcoR I*-digested Vt

PCR product into pET15b (provided as a gift by José M. de Pereda) was left to proceed for 2 h at RT.

Transformation of pBluescriptSK(-)-Vt into DH5 α

The ligation mixture was added to competent *E. coli* DH5a cells (provided as a gift by Clare J. Macleverty) and, together with two control mixtures (one containing competent cells but no plasmid, the other containing competent cells, plasmid but no Vt insert), placed for 30 min on ice. After heat shock (2 min, 42°C waterbath) and a subsequent 5 min on ice, 0.8 mL of 2xYT was added and the samples incubated with shaking for 1 h at 37°C before being plated on fresh Amp plates.

After picking single colonies from the cells transformed with Vt-containing pBluescriptSK(-) and isolating plasmids (Quiagen kit), EcoRI digestions were performed on each plasmid (see above) and the digested products run on a 1.1% agarose gel. The plasmids found to contain the Vt PCR product were then sent for DNA sequencing to confirm the absence of errors before proceeding with subcloning into pET15b.

Subcloning into pET15b

pBluescriptSK(-) vectors containing the Vt PCR product were digested overnight at 37°C with *Nde I* and *Nde I/BamH I* as per manufacturers' instructions. The digestion products were extracted from a 1.3% agarose gel, purified (Quiagen QIAquick gel extraction kit) and ligation into pET15b carried out as per manufacturers' instructions. The ligation was left to proceed for 2.5 h at RT.

Transformation of pET15b-Vt into BL21(DE3)

Transformation of pET15b-Vt into competent *E. coli* BL21(DE3) cells (provided as a gift by Geoff Briggs) was carried out as above (see transformation of pBluescriptSK(-)-Vt into DH5 α). After picking up single colonies from the cells transformed with Vt-containing pET15b and isolation of the plasmids (Quiagen kit), separate digestions with *EcoR V* and *Nco I* were performed on each plasmid as per manufacturers' instructions, and the digested products were run on a 0.8% agarose gel.

Since the Vt PCR product was inserted between positions 5374 (*Nde I*) and 5947 (*BamH I*), cleavage with *EcoR V* (sites at 4040 bp and 6081 bp) should generate a 4229 bp and 2041 bp fragment (or, in the case of no insert, a 2041-573=1478 bp fragment), whereas cleavage with *Nco I* (sites at 5315 bp and 5932 bp generating 5653 bp and 617 bp fragments respectively) should generate the 573 bp Vt PCR fragment.

The successful clones were thus isolated, cultures grown from single colonies and stored as 20% glycerol stocks at -80°C. The map of the final vector containing the Vt insert is presented in figure 2.2.

Expression and purification of Vt using a His-tag

Expression and affinity purification were carried out essentially according to the pET system manual (Novagen, 7th edition). Overnight 10 mL M9ZB/Amp cultures of frozen glycerol stock *E. coli* strain BL21(DE3) cells transformed with the recombinant pET15b plasmid. 1 mL of culture was used to transfect 1L of LB/Amp medium and grown to OD₆₀₀ = 0.6 at 37°C before adding IPTG to 0.25 mM. After an induction time of 4-6 h, cells were pelleted (20 minutes, 7000 rpm, 4°C), resuspended in 1/20 culture

Binding Buffer/0.1% Triton X-100 and frozen at -20°C. Upon thawing, lysozyme was added to 0.1 mg/mL, and PMSF to 2 mM. Cells were lysed by sonication on ice using 5 cycles of pulse 20 sec and wait 30 sec, and pelleted by centrifugation at 16,000 rpm for 15 min at 4°C. Lysates were loaded on a HIS-BIND chromatography column pre-equilibrated in 3 volumes sterile deionized water, 5 volumes Charge Buffer, and 3 volumes Binding Buffer. After loading, the column was washed with 10 volumes Binding Buffer, 6 volumes Wash Buffer and 2 volumes of 180 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl, pH 7.0 before being eluted with 6-8 volumes Elute Buffer. The column was regenerated using 3 volumes of Strip Buffer and stored in 20% ethanol. After dialysis into TCB and thrombin cleavage (3-4 h, RT on a moving rack using 3.5 units of thrombin (Boehringer Mannheim) per L of culture), Vt was re-dialysed in TCB and concentrated in a stirred ultrafiltration cell (Amicon, Beverly, U.S.A). The protein at various stages of purification is shown in figure 2.3. The purified Vt includes three residues (Gly-Ser-His) from the His-tag at the N-terminus.

Expression and purification of SeM-Vt

For the pGEX-2T construct, expression of SeM-Vt was initially carried out according to Budisa et al. (1995). Briefly, B834(DE3) cells which are a methionine auxotrophic strain of *E. coli*, were transformed with the Vt expressing pGEX-2T vector and left to grow in NMM (see below). However no protein expression was observed under these conditions, and the protocol was changed to that of Harrison et al. (1994) as detailed below. This protocol resulted in protein expression, and selenomethionine incorporation was complete (figure 2.4a, measured mass = 21,919 Da; calculated mass = 21,927 Da). No crystallisation of SeM-Vt was observed however, as the low yields

(~1mg/L of culture) and, as with the wild-type Vt, variable degree of purity, did not permit an effective screening of crystallisation conditions.

For the pET15b construct, expression of SeM-Vt was carried out essentially according to Harrison et al. (1994). Briefly, 5 mL of M9ZB inoculated with bacterial stock (*E. coli* strain BL21) transfected with pET15b were used as a starter culture grown throughout the day (~8 hours). Cells were pelleted (3000 rpm, 5 min, RT) and resuspended in 12 mL NMM. Three 500 mL NMMPlus/Amp flasks were inoculated each with 4 mL culture and left to grow overnight. A fourth 500 mL NMMPlus/Amp flask was used to top up the cultures which, after 15 minutes, were induced with IPTG for 8-12 hours. The cells were harvested (8000 rpm, 20 min, 4°C) and protein purification carried out essentially as for Vt using a His-tag (see previous section). Full incorporation was confirmed by electrospray mass-spectrometry (Figure 2.4b; measured mass = 22,057 Da; calculated mass = 22,055 Da).

During the purification of SeMVt, over 50% of the protein precipitated out of solution after being eluted off the HIS-BIND chromatography column, during dialysis into TCB. This decrease in solubility, necessitated a change in the previous protocol with buffer exchange occurring in two steps of decreasing salt concentration (typically from 0.5 to 0.15 M NaCl). Thrombin cleavage was carried out at 0.5 M NaCl. Some of the precipitated protein, was resolubilised in even higher salt buffer (1M imidazole) prior to thrombin cleavage. This last preparation of resolubilised Vt resulted in the production of large, single monoclinic crystals when used in crystallisation trials with A.S. as an additive.

Crystals of Vt

Tetragonal form

Diamond-like crystals were initially grown out of a multi-factorial screen (Hampton Research) at RT and 4°C from PEG 2K MME and PEG 4K as a precipitant respectively. The crystals were typically 10-15 μm on each side and diffracted synchrotron radiation to better than 3 \AA (images collected by Carlo Petosa at Daresbury SRS, station 7.2, $\lambda = 1.488 \text{ \AA}$) showing the symmetry to be P422 with cell dimensions $a = b = 93.53 \text{ \AA}$, $c = 57.85 \text{ \AA}$. Typical crystals of this form are shown in figure 2.4a. However, subsequent attempts at reproducing this form did not meet with success. Instead efforts were concentrated on two other crystal forms which emerged from fine-screening of the initial crystallisation conditions.

Monoclinic form

Native monoclinic crystals grow in sitting drop diffusion trays at 22°C using a 3:2 mixture of Vt (6 mg/mL, in TCB including 4 mM DTT) and reservoir (25% (w/v) PEG 2000, 0.2-0.15 M A.S., 0.1 M CH_3COONa pH 5.0). Crystals have a plate-like morphology and grow, in the case of sitting drops, to dimensions of typically 0.2 x 0.1 x 0.05 mm within two weeks. Crystals belong to space group $P2_1$ with two molecules per asymmetric unit and a solvent content of 44%. The mosaicity of native and derivative crystals varied between 0.5 and 2.0° with 20% glycerol as the cryoprotectant, and stabilized at $\sim 0.8^\circ$ when glycerol was replaced with 30% PEG400.

A crystal is shown in figure 2.5b. This crystal has unit cell parameters $a = 39.6 \text{ \AA}$, $b = 87.8 \text{ \AA}$, $c = 50.6 \text{ \AA}$, $\beta = 91.3^\circ$. For most crystals of this form, non-isomorphism was

observed, this being especially pronounced along c^* and for the β angle both of which were found to vary by up to 5%. Table 2.1 lists some of the unit cell parameters encountered in the monoclinic form.

The diffraction intensity of capillary-mounted monoclinic crystals at RT decayed by ~60% after 24 hours in an X-ray beam from a 100 kW rotating anode source and were therefore flash-frozen using 20% glycerol. However, the use of glycerol as a cryoprotectant resulted in an increase of the mosaic spread of the crystals and was subsequently replaced by 30% PEG 400.

The crystallisation conditions for the SeMVt protein are similar to those of the wild type protein, except that the precipitant is 30% PEG 400 and the protein concentration 11mg/mL. In this case, the cryoprotectant used was 35% PEG 400. Data collected from this crystal form were used to refine the Vt model at 1.8 Å resolution.

Orthorhombic form

Orthorhombic crystals of wild-type Vt are grown under identical conditions to those of monoclinic crystals. Needles of ~0.5-1mm in length and $0.2 \times 0.01 \text{ mm}^2$ in cross-section grow within a period of one to two weeks. Typical crystals are shown in figure 2.5b. The space group is $P2_12_12_1$, with $a = 35.3 \text{ \AA}$, $b = 86.4 \text{ \AA}$, $c = 126.1 \text{ \AA}$. Crystals also have two molecules per asymmetric unit with a slightly higher solvent content (47%) than form 1. Using 30% PEG 400 as a cryoprotectant, these diffract synchrotron radiation to 2.1 Å. Data collected from this crystal form were used to compare the model of Vt with that generated from crystals of SeMVt (see Chapter 3).

Spherulites and orthorhombic form crystals of SeMVt (11mg/mL) were grown from 30% PEG 400, 0.2 M A.S., 0.1 M sodium acetate pH 5.0. The addition of the

detergents n-octyl- β -D-glucoside or HECAMEG used according to manufacturers' instructions (4:5 protein:precipitant ratio, 1:5 detergent:precipitant ratio), decreased the number of spherulites leading to the growth of hollow needles. Native and derivative data sets were collected from these crystals.

Materials

Ampicillin, DTT, thrombin and PMSF were from Sigma (Poole, Dorset, U.K.), reduced glutathione, β -ME, Triton X-100 and PEGs from Fluka (Buchs, Switzerland), EDTA and agarose from Promega (Southampton, U.K.), IPTG from Melford Laboratories Ltd. (Ipswich, Suffolk, U.K.). All ion-exchange columns as well as the FPLC units used were from Pharmacia Biotech (Upsala, Sweden) or BioRAD (www.bio-rad.com) and buffers from Sigma or Fisher Scientific (Loughborough, U.K.).

Manufacturers of materials not mentioned here, are stated at the relevant sections below.

General methods

SDS-PAGE

Protein samples were analyzed by SDS-PAGE according to Laemmli (1970), using the Mini-Protean II electrophoresis cell system (BioRad). 13% and 8% acrylamide gels were used for the analysis of Vt and Vg respectively. Typically 4 mL of separating gel (3.03 mL (Vt) or 2 mL (Vg) 30 % acrylamide, 1.2 mL (Vt) or 2.5 mL (Vg) deionised water, 2.8 mL 1M Tris-HCl, pH 8.8, 75 μ L 20% SDS, 50 μ L of 10% $(\text{NH}_4)_2\text{S}_2\text{O}_8$, 15 μ L TEMED) were poured between the two plates and after polymerization, 1.5 mL of

stacking gel (0.67 mL 30% acrylamide, 3.75 mL deionised water, 0.625 mL 1M Tris-HCl, pH 6.8, 25 μ L 20% SDS, 25 μ L 10% $(\text{NH}_4)_2\text{S}_2\text{O}_8$, 10 μ L TEMED) were poured on top of the separating gel. Sample wells were formed using a 10-well comb. Samples were mixed with sample buffer containing β -ME or DTT, and incubated at 100°C for 5 minutes. After loading of samples into the sample well, the apparatus was run at 20-25 mA per gel for approximately 1 hour.

Protein bands were visualised by staining the gel with Brilliant (Coomassie) Blue R250 for 1-2 hours, followed by destaining in 17.5% isopropanol/10% acetate solution.

Protein concentration

Two methods were used for the estimation of protein concentration. The first was the Coomassie dye-binding protein assay, according to the method of Bradford (1976). The protein samples were incubated in assay mix for 5-10 minutes, and the absorbance at 595 nm was determined. After subtraction of the blank absorbance (of assay mix without the protein), the protein concentration in the sample was determined from a standard curve of absorbance against known concentrations of bovine serum albumin (BSA) from 0-1 mg/mL.

The second method for determination of protein concentration was measuring the absorbance at 280 nm and calculating the concentration using the molar extinction coefficient, $\epsilon^{0.1\%}_{280}$, for the purified protein. This method had the advantage that the protein sample could be recovered after the protein concentration had been determined, but could only be used for relatively pure (>90%) protein samples. The $\epsilon^{0.1\%}_{280}$ value used for Vt was 0.883 and 0.526 for Vg, were determined using the programme

ProtParam (<http://www.expasy.ch/tools/protparam.html>), according to the method of Gill and von Hippel (1989).

Summary

Vt crystallised readily in a variety of crystal forms. The original tetragonal form could not be repeated, while orthorhombic crystals were obtained rarely and did not grow large enough to diffract CuK α radiation to high resolution. The abundant monoclinic form presented problems of twinning and non-isomorphism. Reproducibility problems also arose from the low yields and variable purity of the protein produced from the original vector. Control of protein purity and an increase in protein yield required a change in the expression vector which, once achieved, made it possible to bypass the need for the monoclinic form by growing large orthorhombic crystals capable of high-resolution diffraction. Selenomethionine-incorporated orthorhombic crystals were also grown, with single monoclinic ones being finally obtained from resolubilised protein.

Crystal	Habit	Space Group	a(Å)	b(Å)	c(Å)	β (°)	Diffraction Limit (Å), (Source)
Tetragonal form	Diamonds	P422	93.5	93.5	57.9	90	3.0 (SRS, $\lambda = 1.488$)
Monoclinic form	Plates	P2 ₁	39.6	89.0	53.0	94.3	2.2 (CuK α)
			39.3	90.2	52.1	92.9	3.5 (SRS, $\lambda = 0.87$)*
			39.6	87.8	50.6	91.3	1.8 (SRS, $\lambda = 0.87$)
Orthorhombic form	Needles	P2 ₁ 2 ₁ 2 ₁	35.3	86.4	126.1	90	2.2 (SRS, $\lambda = 1.488$)

Table 2.1. A summary of crystal forms. In all cases, the temperature of the crystals during diffraction was 100K.

SRS: Daresbury Synchrotron Radiation Source. Note the non-isomorphism displayed by the monoclinic form.

* Ce soak

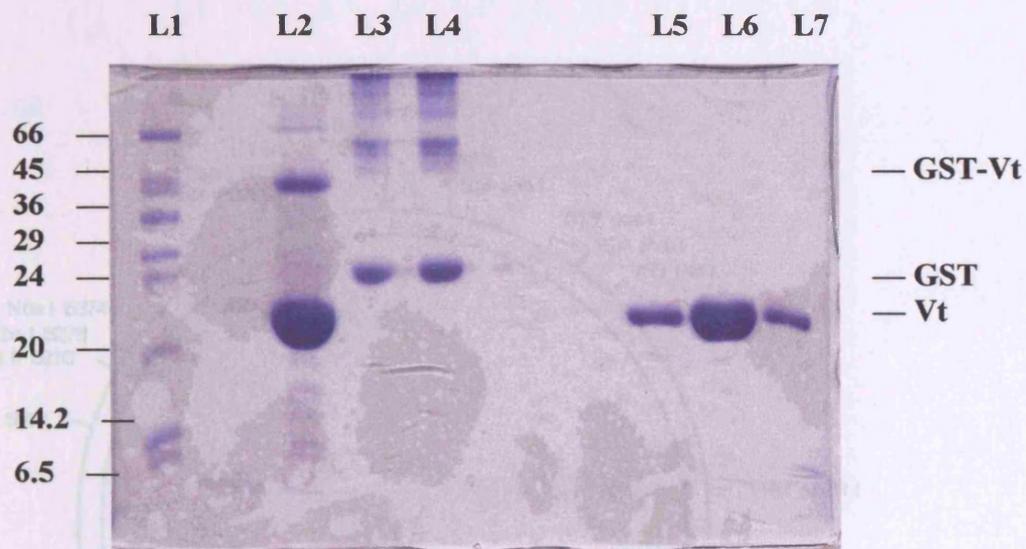


Figure 2.1. Purification of Vt using a GST-affinity column and ion-exchange. L1 = Low molecular weight markers (Sigma), L2 = Vt after (incomplete) thrombin cleavage and passage through the GST-affinity column, L3-L4 = GST and impurities on beads , L5-L7 = fractions of Vt separated by cation-exchange (pH 5.0).

Figure 2.2. Vt in pET15b. Vector size: 3702 bp. Insert size: 573 bp. The figure was generated using the program Plasmid Processor version 1.02/3.7.1996 (<http://www.oku.fi/~livvraun/plasmid/plasmid.html>).

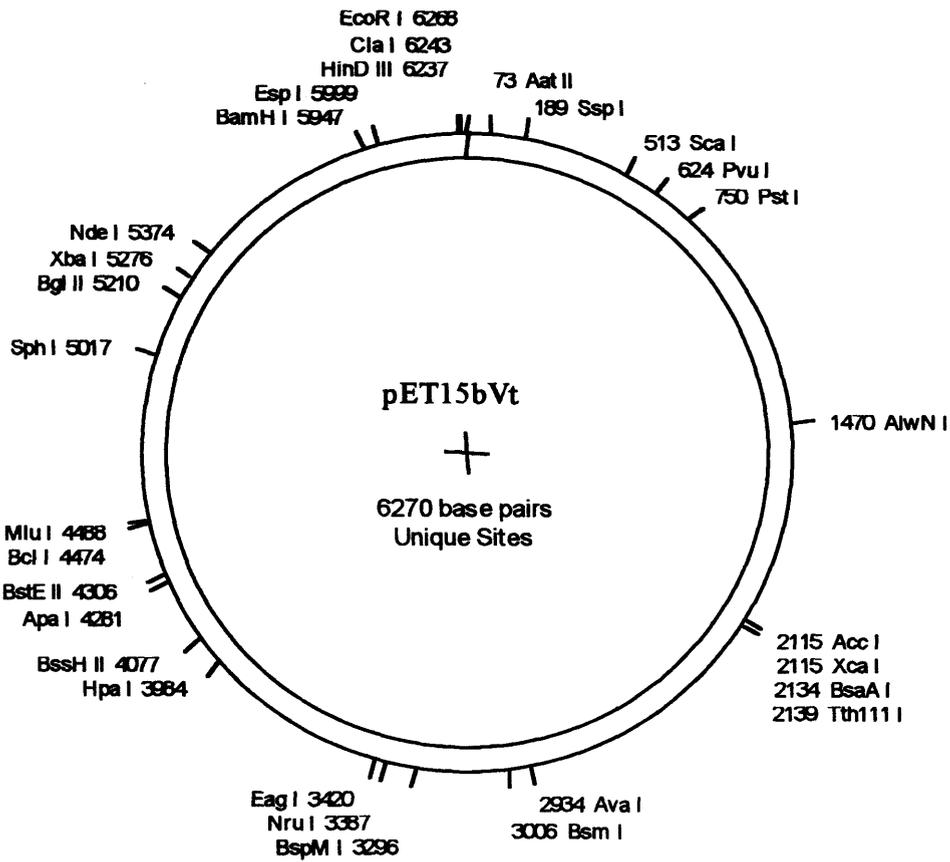


Figure 2.2. Vt in pET15b. Vector size: 5708 bp. Insert size: 573 bp. The figure was generated using the programme Plasmid Processor version 1.02/3.7.1996 (<http://www.uku.fi/~kiviraum/plasmid/plasmid.html>).

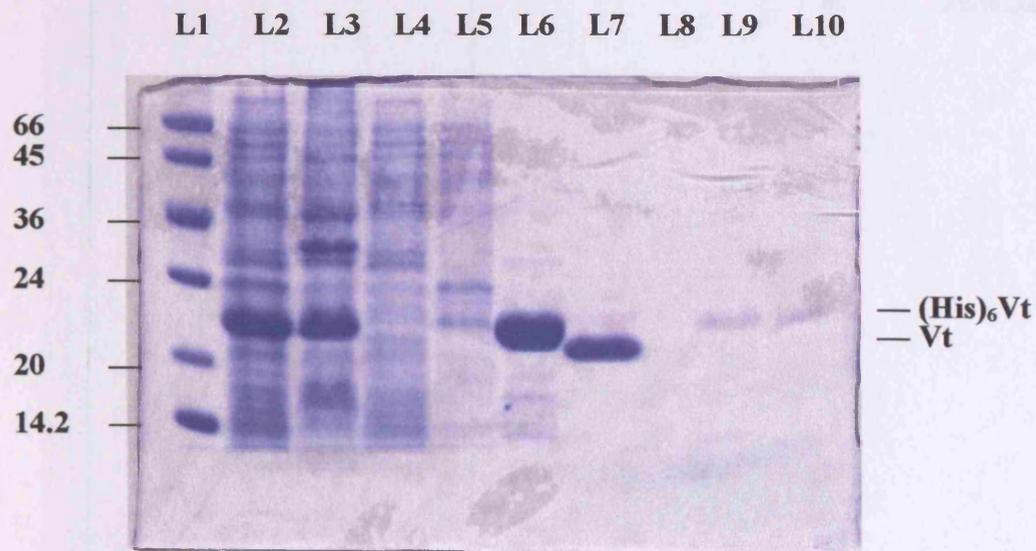


Figure 2.3. Purification of Vt subcloned using a His-tag. L2 = total cell lysate, L3 = cell lysate supernatant, L4 = 5mM imidazole flow-through, L5 = 60mM imidazole flow-through, L6 = 1M imidazole flow-through, L7 = Vt after thrombin cleavage and dialysis, L9-10 = (His)₆Vt in EDTA column wash.

Figure 2.4. ESI mass spectrometry of Vt and ScMVt. Proteins were produced using the vectors a) pGEX-2T (A = Vt, B = ScMVt) and b) pST15b (A = ScMVt, B = Vt).

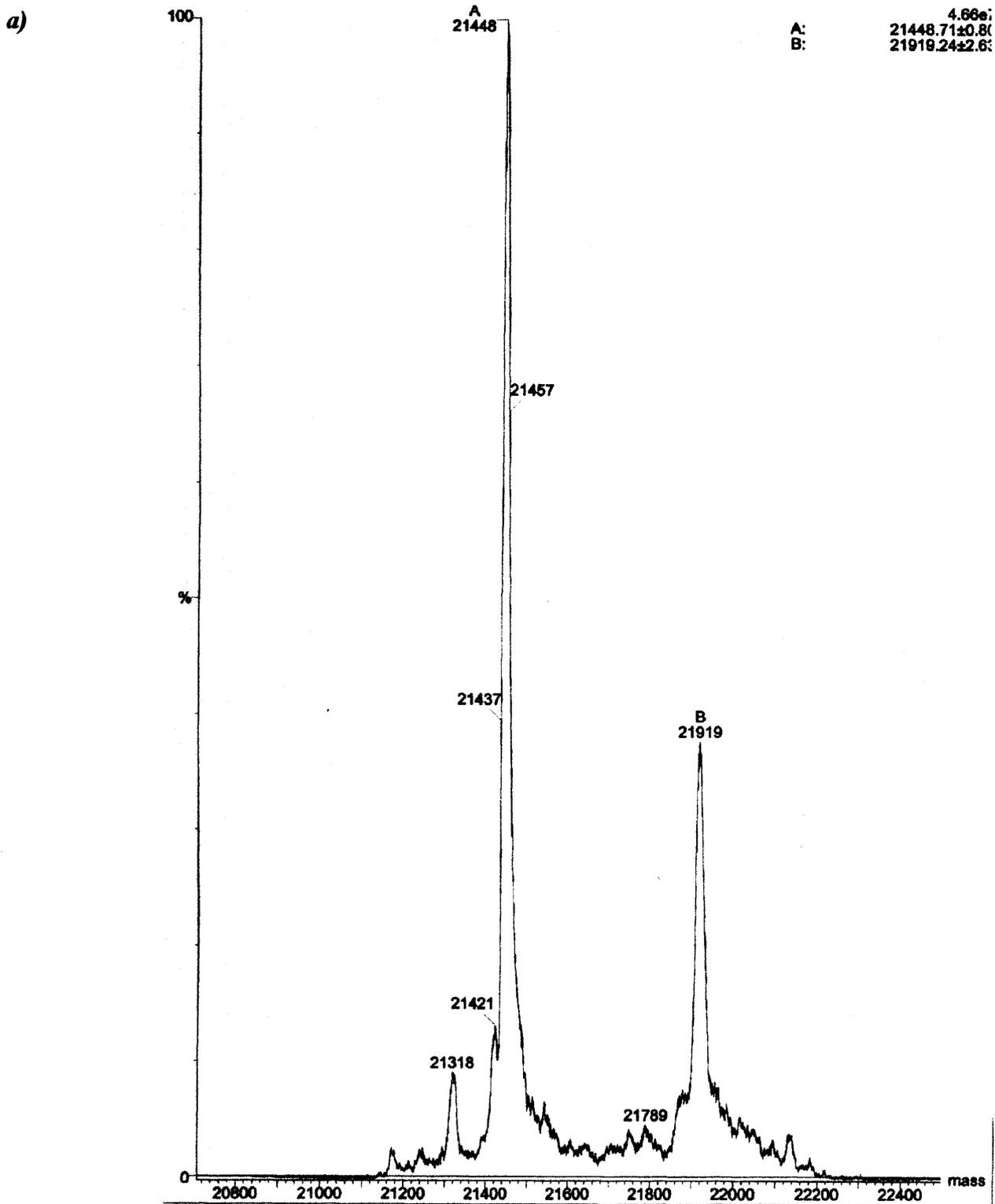
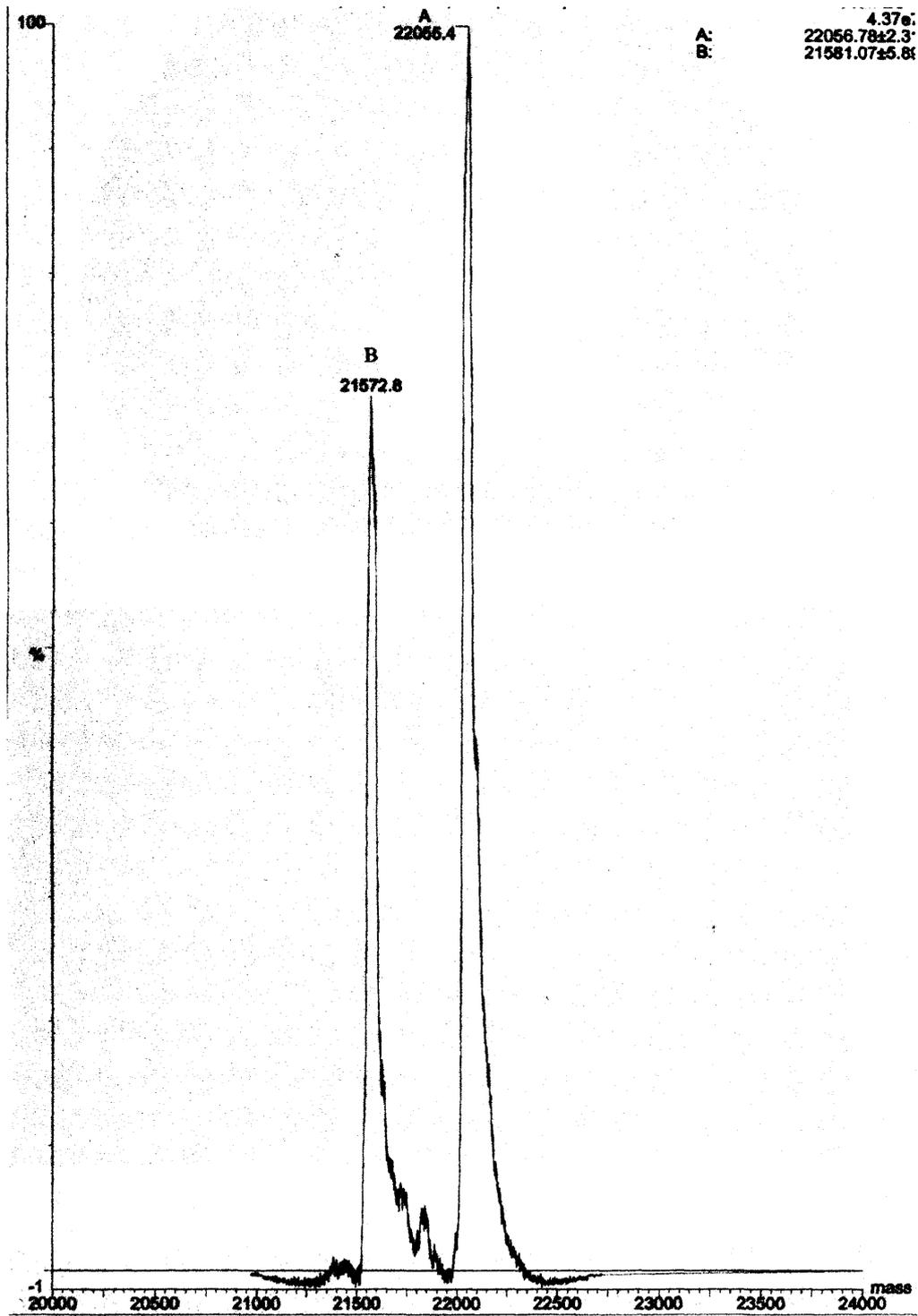
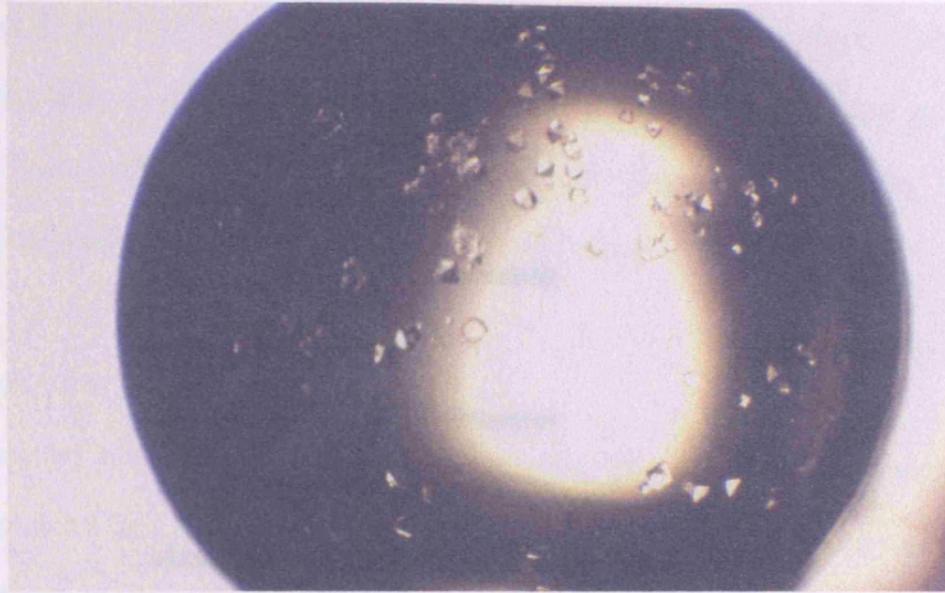


Figure 2.4. ESI mass spectrometry of Vt and SeMVt. Proteins were produced using the vectors *a*) pGEX-2T (A = Vt, B = SeMVt) and *b*) pET15b (A = SeMVt, B = Vt).

b)



a)



b)



Figure 2.5. Crystals of Vt. a) *Tetragonal form*. Crystals grow as diamonds of typical dimensions $0.01 \times 0.01 \times 0.015 \text{ mm}^3$. b) *Monoclinic form*. Crystals grow as plates up to 0.2 mm in length and up to 0.05 mm thick. *Orthorhombic form*. Crystals grow as needles of dimensions $0.5 \times 0.1 \times 0.01 \text{ mm}^3$.

Chapter 3:

Structure Determination of Vt

Introduction

Data collection and processing

Heavy metal derivatives

Model building and refinement

Model refinement using a monoclinic crystal

Assessment of the reliability of the model

Agreement with crystallographic data

B factors and geometry

Crystal packing, heavy metal sites, PDB validation

Summary

Introduction

As outlined in the previous chapter, the search for heavy atom derivatives of Vt was hampered by the twinning and non-isomorphism of monoclinic crystals. This led to efforts being diverted towards the orthorhombic form with selenomethionine-substituted Vt crystals grown for purposes of phasing by the MAD method. Ultimately, this method turned out not to be necessary, although SeM-Vt crystals were used in the initial and final stages of the structure determination process.

This chapter presents the steps involved in the structure determination of Vt from the collection and processing of native and derivative data sets, to the subsequent

phasing, model building, and refinement. Model building is presented from first experimental maps of the native orthorhombic form to the complete, refined structure in the selenomethionine-substituted monoclinic form. Finally, the reliability of the monoclinic model is examined according to various criteria.

Data collection and processing

The first step in a crystal structure determination involves the extraction of intensity data from raw diffraction images. These data are used to determine the crystal and cell parameters and to generate an indexed reflection list. The reflections are subsequently integrated, scaled and merged to give a final set of intensities.

Diffraction data from orthorhombic native (Nat) and derivative crystals (SeM1, EMTS1, EMTS2) were collected on a MAR image plate at SRS Daresbury, station 7.2 at a wavelength of 1.488 Å. Data from a monoclinic crystal (SeM2) were collected on a CCD area detector at SRS Daresbury, station 9.6 at a wavelength of 0.87 Å. An additional dataset of an orthorhombic crystal (Se_SAD) was collected on a MAR image plate at SRS Daresbury, station 9.5 at a wavelength of 1.000 Å.

All data were processed with the programmes DENZO and SCALEPACK (Otwinowski and Minor, 1997). DENZO was used for automatic indexing of the diffraction spots to determine the orientation matrix of the cell and to integrate the measured intensities by both summation (definition of a volume containing an individual spot) and profile fitting (using the radius of the area around each particular spot, containing neighbouring spots, to calculate the average reflection spot; the spot in question is fitted to the average profile of all the spots within the specified radius - the radius is set so that spots on roughly 3-5% of the area of the detector are included in

the averaging); SCALEPACK was used to scale the intensities associated with a given diffraction spot taking into account Lorentz and polarization effects (dependent on the data acquisition system and the direction of X-ray scattering respectively), and air and crystal absorption due to asymmetry in the crystal shape (correction factors L, P and A respectively). From the merged intensities (I_s) and their standard deviations ($\sigma(I_s)$), a best estimate of F_s and ΔF_s (structure factors and anomalous difference structure factors) was calculated using TRUNCATE (CCP4 suite). The truncation procedure has the effect of forcing all negative observations of up to minus four standard deviations to be positive while inflating the weakest reflections (less than 3 standard deviations), thus resulting in significant better values for the weak data.

Data collection statistics for the two data sets used during refinement are summarised in table 3.1 and a summary for the derivative datasets is given in table 3.6a.

Heavy metal derivatives

Theory

The electron density in a crystal can be obtained by calculating the Fourier summation:

$$\rho(xyz) = V^{-1} \sum_{hkl} |F(hkl)| \exp[-2\pi i(hx + ky + lz) + i\alpha(hkl)]$$

in which $|F(hkl)|$ is the structure factor amplitude of reflection (hkl), including the temperature factor (see p.65), and $\alpha(hkl)$ is the phase angle. x , y , and z are coordinates in the unit cell. From the diffraction pattern the values of $I(hkl)$ can be obtained after applying the correction factors L, P, and A. Since $I(hkl) = |F(hkl)|^2$ the amplitudes $|F(hkl)|$ can be found. However, no information is available on the phase angles.

The basic technique of solving the phase problem in protein crystallography is isomorphous replacement. Two crystals are said to be isomorphous if they have essentially the same structure but are composed of chemically different atoms. In protein crystallography, isomorphous replacement usually involves the binding of a heavy atom in a position previously occupied by disordered solvent. The electron density of the isomorphous derivative should differ from the electron density of the native crystal just for the peaks due to one or a few specifically-bound heavy atoms; the unit cell dimensions and the protein molecule conformation must be essentially the same in both.

Crick and Magdoff (1956) compared the average change in intensity produced by adding extra atoms to the crystal with the average change in intensity produced by small shifts of the protein molecules and slight changes in the unit cell dimensions. They showed that the rms fractional change in intensity I produced by stoichiometric binding of N_H heavy atoms of scattering power f_H to a protein of N_P atoms having a mean scattering power f_P is $\langle \Delta I^2 \rangle / \langle I \rangle \sim (N_H/N_P)^{1/2} (f_H/f_P)$. They concluded that small shifts either of the molecule or of the cell dimensions may produce changes in the intensity sufficient to interfere seriously with the changes due to the heavy atoms at high values of $1/d$ and proposed the examination of the variation of $\langle \Delta I^2 \rangle / \langle I \rangle$, or alternatively of $|\langle F_{PH} \rangle - \langle F_P \rangle| / \langle F_P \rangle$, with resolution to identify lack of isomorphism; if the ratio increases at high $1/d$ some movements of the protein have taken place.

In order to use the isomorphous replacement method, the amplitude and phase of the heavy atom contribution to the structure factor of the derivative has to be determined. The structure factors of the native protein F_P , the derivative F_{PH} and the heavy atom structure F_H , are related by the vector equation $F_H = F_{PH} - F_P$. Only the

magnitudes of F_{PH} (F_{PH}) and of F_P (F_P) can be experimentally obtained. For centric reflections, F_{PH} and F_P are collinear ($F_H = |F_{PH} - F_P|$ or $F_H = |F_{PH} + F_P|$), while for acentric reflections the phases of F_{PH} and F_P are not, in general, correlated and $|F_{PH} - F_P|$ is approximately equal to $|F_H \cos(\alpha_{PH} - \alpha_H)|$, where α_{PH} and α_H are the phases of the derivative and heavy atom respectively. Anomalous scattering data (ΔF_{ano}) can also be used to estimate F_H with the approximate formula being $\Delta F_{ano} = 2k^{-1}F_H \sin(\alpha_{PH} - \alpha_P)$, where k is defined as F_H/F_H'' and can be obtained from the International Tables of Crystallography.

The estimated F_H s are used to determine the heavy atom positions. These coordinates will subsequently be used to calculate the vector F_H required to find the protein phases. The most general method for determination of heavy atom positions is the use of Patterson techniques. An isomorphous difference Patterson synthesis is a Fourier summation of the terms $|F_{PH} - F_P|^2 - |F_H \cos(\alpha_{PH} - \alpha_H)|^2$, that is to say a Fourier summation with intensities as coefficients and with all phase angles equal to zero. In the case of centric reflections, $|F_{PH} - F_P|$ really corresponds to F_H and a Patterson with $(F_{PH} - F_P)$ coefficients is a good estimate of a F_H^2 Patterson. For acentric reflections the summation has two terms: $F_H^2/2$ and $F_H^2 \cos 2(\alpha_{PH} - \alpha_H)$. The first term will generate vectors between the heavy atoms at half the ideal height. The second term will contribute only noise, since there is no connection between the F_H and the phases used.

An anomalous difference Patterson synthesis, ΔF_{ano}^2 , is a Fourier summation of the terms $|F_{PH}^+ - F_{PH}| - 2k^{-1}F_H \sin(\alpha_{PH} - \alpha_P)^2$. As $\sin^2 \theta = (1 - 2\cos(2\theta))/2$, the summation of ΔF_{ano}^2 should also give peaks in the Patterson corresponding to vectors between heavy atoms plus noise. There is no lack of isomorphism in the anomalous difference, but in general

the F'' signal is much weaker than the F_H signal, so the random errors inherent in the measurements can swamp it more easily. Better Patterson maps can sometimes be obtained using estimates of F_H which take into account both isomorphous and anomalous data (summed or combined Patterson). Symmetry elements can cause a concentration of peaks in certain lines or planes, termed Harker lines or Harker planes. Automated search procedures for locating heavy atoms in difference Patterson maps have been developed (Terwilliger and Eisenberg, 1987).

When a previous estimate of the protein phases exist, a Fourier synthesis with coefficients $m(F_{PH}-F_P)e^{i\alpha_p}$, where m is the weight associated with this phase, is the simplest way to locate heavy atoms. A difference Fourier map shows positive electron density at the site of atoms that were not present in the native structure and negative density at the positions of atoms present in the native, but not in the derivative structure. It can therefore be used to detect additional weakly occupied sites after the main heavy atom site(s) has been found from a difference Patterson map. The biggest problem with this method is the fact that these Fourier syntheses will also have 'ghost' peaks or halos at or near the heavy atom sites used for phasing. These can be erroneously interpreted as minor sites. It is therefore useful to check any such sites against the derivative Pattersons.

Any Patterson function will be equally well satisfied by coordinates x,y,z or $-x,-y,-z$. In other words, the hand of the solution will not be determined. In addition, for cases of more than one derivative, it is essential to apply the same origin and hand to all derivatives. The use of a difference Fourier based on the phases of the first derivative will give the sites of the subsequent derivatives on the same relative origin and hand (Matthews, 1966).

Results

Two out of the three cysteines per Vt molecule have titratable sulphhydryl groups (reaction with 5, 5' dithio-bis (2-nitrobenzoic acid) (DTNB), data not shown). The existence of reducing agent in the crystallisation buffer additionally inhibited formation of intermolecular disulphide bonds. The small number of free cysteines, crystallisation conditions for both crystal forms (acidic pH, low ammonium sulphate concentration) and the strong scattering power of mercury made it the derivitisation compound of first choice.

Monoclinic crystals were soaked in a variety of mercury and other heavy metal solutions in an attempt to obtain suitable isomorphous derivatives. The compounds screened are listed in table 3.2. Most compounds resulted in crystal cracking, greatly reduced diffraction quality, non-isomorphism or insufficient substitution. Co-crystallisation attempts with mercury, platinum and iridium compounds were unsuccessful, resulting in either a lack of crystal growth or in non-diffracting crystals. SeM-Vt is the only successful derivative of the monoclinic form to date.

EMTS was the only mercury compound that did not appear to result in crystal cracking or significant deterioration in the diffraction quality of orthorhombic crystals. Two soaking conditions were tried: a high soak (10mM, 1h) referred to as "EMTS1", and a low soak (1mM, 12h), referred to as "EMTS2". Using the programme SOLVE (Terwilliger and Berendzen, 1999), 6 mercury sites were identified for EMTS1 with an overall phasing power of 1.33 (centric) and 1.63 (acentric) to 3.15 Å, and a mean figure of merit of 0.43 to 3.5 Å. For EMTS2, 8 major sites were similarly determined, giving an overall phasing power of 1.27 (centric) and 1.57 (acentric) to 2.8 Å, and a mean figure of merit of 0.50 to 3.5 Å. These results are summarised in table 3.3. The

Harker sections of isomorphous difference Pattersons of EMTS2 are shown in figure 3.1, and the Harker section $z=1/2$ of isomorphous and anomalous difference Pattersons of EMTS1 is shown in figure 3.2.

The positional and occupancy parameters for the heavy atom sites in both EMTS1 and EMTS2 were input into MLPHARE (CCP4 suite) and refined to 3.5 Å in five cycles using fixed overall scale, isotropic and individual thermal parameters. Self-difference Fourier maps were used to test for the ability of major sites to locate others. No additional sites were located. A self-difference Fourier of EMTS2 is shown in figure 3.3a.

To establish the correct hand, three sets of phases were calculated from the EMTS1 derivative to 3.15 Å. SIR phases were calculated using only centric reflections, and SIRAS phases for each hand were calculated using both centric and acentric data with anomalous occupancies set to zero. The anomalous occupancies all refined positive, indicating that the original choice of coordinates was correct.

The process was repeated for the EMTS2 derivative to 2.8 Å. Using EMTS2 SIR phases, an anomalous cross-difference Fourier with the Se-SAD dataset (table 3.6a) revealed a cluster of selenium atoms (Figure 3.3b) locating the methionine-zipper (see Chapter 4). In total, 18 selenium atoms were identified by this method. However, no clear indication of hand was obtained with the inclusion of anomalous data. Subsequently, maps for both hands were calculated and after solvent flattening, comprising 100 cycles with phase extension from 15-2.8 Å using the programme DM (CCP4 suite) gave an R_{free} value of 50.1% for one hand and 41.8% for the other. DM (density modification) provides an automatic procedure for applying a low constant value to the disordered solvent encountered in the crystal (Leslie, 1988). R_{free} is a

property of both the density modification technique and the initial data set and is analogous to the R_{free} used in coordinate refinement in that it is calculated using a set of reflections which are excluded from the initial map calculation. However, once phase relationships between structure factors are introduced it is impossible to completely isolate a set of reflections since the absence of a set of reflections leads to a systematic error in the phasing of the rest of the reflections. Thus it is necessary to change the R_{free} from cycle to cycle.

The hand with the lower R_{free} value was chosen and a map calculated to 2.8 Å revealed right-handed α -helices as final proof of correct hand.

Model building and refinement

A map calculated using the EMTS2 derivative was examined and showed clear evidence of helical electron density (figure 3.4). Therefore EMTS1 phases were not further examined and model building on the EMTS2 map performed using FRODO (Jones, 1978). The initial model consisted of 5 continuous polypeptide segments accounting for 150 residues. Localization of the 9 selenium and 4 mercury atoms permitted the unambiguous orientation of the C α trace (3 methionines are continuous in the primary sequence of Vt, and 3 other methionines are positioned on same side of helix). After localisation of the methionines (SeM1 dataset, table 3.6a) the model was refined using the Nat data set (table 3.1) with CNS (Brunger *et al.*, 1998). Simulated annealing (SA) was performed by heating the model to 3000K (1500K in the final rounds of refinement) and letting it cool in steps of 25K. No sigma cutoff or NCS restraints were applied; the second molecule was generated by manual rotation of the first. Phases were then calculated from the refined model to 2.5 Å and used to generate

a new map. This map showed a visible improvement over the starting map, allowing more residues to be modeled and poorly modeled residues to be corrected. The new model was subjected to a second round of refinement, which reduced the R-factor to 0.37.

Successive rounds led to the location of the remaining sequence (loops, N- and C-terminal regions) and to a gradually improving image of the molecule. The progress is summarised in table 3.5.

Model refinement using a monoclinic crystal

Theory

A molecular replacement search was performed using a monomer of the orthorhombic crystal form as the search unit with the program AMORE (Navaza, 1994). The program incorporates rotation and translation steps plus rigid body fitting. The rotation function is used to find the correct orientation of the model in the crystal lattice by testing the agreement between the Patterson functions calculated from the model and from the data (Fs) at various relative orientations. At the relative orientation that superposes the model onto the crystal, the product of the two Patterson functions i.e. the maps of interatomic vectors, should have a large value.

The object of the rotation function is to include vectors between atoms within the same molecule (self vectors) while eliminating vectors between neighbouring molecules (cross vectors) as much as possible. Crowther's FFT self rotation function (Crowther and Blow, 1967) is used to detect NCS from native Fs, and the cross rotation function evaluates externally- or internally-calculated Fs to produce a list of Eulerian angles (α , β , γ) giving all symmetry equivalents (crystallographic and NCS-related).

The translation function investigates the correlation between the observed intensities and the cross-vectors between the symmetry related molecules of the model as it is moved about the unit cell. When the model is positioned correctly, the function should have peaks at values corresponding to the translation vectors (x, y, z) between the symmetry related molecules. Finally, rigid-body refinement (a method assigning a rigid geometry to parts of the structure and seeking to minimise $(|F_{\text{obs}}| - |F_{\text{calc}}|)^2$ with respect to all positional coordinates and thermal parameters by the least-squares method) is applied to selected solutions of the translation search and the refined rotation and translation parameters applied to the model.

Results

The MR results are summarised in table 3.4. A cross-rotation gave two pairs of symmetry-related solutions (1,1' and 2,2') for the original model. After applying the Euler angles from solution 1, a translation search revealed two possible solutions (1-2, 1-2'). After one cycle of rigid body refinement, solution 1-2-i had an R-factor of 0.44 and a correlation coefficient of 57.7.

A complete data set to 1.8 Å resolution was collected from a SeM-Vt monoclinic crystal (SeM2, table 3.1) and used to refine the MR model. Refinement statistics for the model of the monoclinic crystal form are summarized in table 3.6*b* and a detail of the final model is shown in figure 3.5.

Maps of the difference between observed and calculated structure factors ($F_o - F_c$ maps) showed positive peaks of density consistent with the presence of water molecules, of which over 300 have been modeled. An additional peak of density was found and was attributed to the presence of a sulphate ion for the following reasons:

the heights of this peak was significantly greater than that of nearby waters, it possessed tetrahedral coordination geometry, and it was located within 2.7 Å of basic residues (R910, K1061). In the second molecule the basic residues, implicated in coordinating the sulphate ion in the first molecule, are involved in crystal contacts. The ion is not visible in the orthorhombic model.

Assessment of the reliability of the current model

Agreement with crystallographic data

The current model for the monoclinic crystal form consists of residues 881-1061 (molecule 1) and 880-1046/1054-1064 (molecule 2), 319 water molecules and one sulphate ion. The working R-factor is 20.01% and the free R-factor is 24.78% for reflections between 10 and 1.8 Å. No sigma cut-off has been applied. No density is seen for residues 1062-1066 (molecule 1) and very weak and discontinuous density is seen for residues 1047-1053 (molecule 2). An estimation of the average value of the error in the atomic coordinates of the molecular model can be obtained by methods proposed by Luzzati (Luzzati, 1952) and Read (Read 1986; Read 1990). The Luzzati method expresses the average error in the atomic coordinates in relation to the difference between $|F_o|$ and $|F_c|$, as expressed in the crystallographic reliability factor (R_{work}). In the method proposed by Read (1986), the rms value of the coordinate error is obtained from a plot of $\ln\sigma_A$ vs. $\{(\sin\theta)/\lambda\}^2$ where σ_A is a function of the correlation between observed and calculated structure factors. A Luzzati plot (not shown) shows that the mean error of the coordinates is 0.21 Å, and a σ_A plot (not shown) shows the error to be 0.22 Å. Both the Luzzati and σ_A values are within or below the range of

estimated coordinate errors established for well-refined structures (between 0.2 and 0.3 Å) (Drenth, 1999).

B factors and geometry

X-rays do not meet identical atoms on exactly the same position in successive unit cells, as atoms vibrate around equilibrium positions. This is similar to an X-ray beam meeting a smeared atom in a fixed position, the size of the atom being larger if the thermal vibration is stronger. This diminishes the scattered X-ray intensity, especially at high scattering angles. In the simple case in which the components of vibration are the same in all directions, the vibration is called isotropic. Assuming isotropic and harmonic vibration it can be shown that the thermal parameter or temperature factor, B, is related to the mean square displacement $\langle u^2 \rangle$ of the atomic vibration: $B = 8\pi^2 \langle u^2 \rangle$.

In addition to temperature-dependent, dynamic disorder, protein crystals have static disorder, as molecules, or parts of molecules, in different unit cells do not occupy exactly the same position and do not have exactly the same orientation. The effect of this static disorder on the X-ray diffraction pattern is the same as for the dynamic disorder and they cannot be distinguished, unless intensity data at different temperatures are collected.

The average B factors for the monoclinic form of Vt are 18.9 and 22.0 Å² for the main chain atoms of molecules 1 and 2 respectively (figure 3.6). There are two stretches where main chain B factors exceed 50 Å², but these are confined to surface loops (residues 1006-1012, molecule 1) or residues bordering a region of undefined electron density (residues 1045-1053, molecule 2). A marked increase in the B-values

of molecule 2 (residues 960-985) can be explained by the absence of a salt-bridge visible in molecule 1 (see Chapter 4 for details).

Rms deviations from ideal values are 0.010 Å for bond lengths, 1.06° for bond angles and 17.7° for dihedrals. 95.5% of the backbone angles are in the most favoured regions of the Ramachandran plot (Ramakrishnan and Ramachandran, 1965) and none are in disallowed regions (figure 3.7). Almost all side chains fall within allowed regions of a $\chi_1\chi_2$ map (data not shown). Conformations predicted to be the most stable are observed to be the most populated and the overall main-chain and side-chain geometry is equal to or better than that of the average protein structure determined at 1.8 Å resolution.

Crystal packing, heavy metal sites, PDB validation

Although the crystal packing has not been closely examined, there is no overlap among neighbouring molecules in either crystal form, and crystal contacts determined using the programme CONTACT (CCP4 suite) look chemically reasonable. The heavy metal atoms are bound at the expected locations (table 3.3).

The monoclinic structure has been deposited in the Brookhaven Protein Data Bank (PDB), validated with WHAT_CHECK (Hooft *et al.*, 1996) and is currently awaiting release (coordinate entry 1qkr, structure factor entry r1qkrsf).

Summary

The structure of Vt has been solved by the SIRAS method using a mercury derivative in orthorhombic crystals supplemented with the location of Se atoms from SAD data to correctly orient the $C\alpha$ -trace, and by molecular replacement in a monoclinic crystal. The model, according to a number of crystallographic, structural and chemical criteria, is highly reliable. Atomic coordinates and structure-factors for Vt have been deposited with the PDB with accession codes 1qkr and r1qkrsf respectively.

Resolution Shell (Å)		total unique	%complete	Average redundancy	$\langle I \rangle /$ $\langle \text{err} \rangle$	R_{sym}
ORTHORHOMBIC (Nat)						
15.00	4.70	2143	96.7	3.04	17.6	0.025
4.70	3.75	2062	98.3	3.09	19.7	0.028
3.75	3.28	2026	98.1	2.92	17.3	0.042
3.28	2.98	2005	97.9	2.79	13.0	0.064
2.98	2.77	2005	97.3	2.72	10.4	0.088
2.77	2.61	1932	95.1	2.69	8.7	0.112
2.61	2.48	1867	92.2	2.76	7.2	0.142
2.48	2.37	1794	89.6	2.85	5.9	0.176
2.37	2.28	1785	88.1	2.85	4.6	0.218
2.28	2.20	1775	88.3	2.94	4.1	0.262
15.00	2.20	19394	94.2	2.87	14.5	0.048
MONOCLINIC (SeM2)						
100.00	3.88	3053	94.0	2.9	23.4	
3.88	3.08	3161	98.4	2.9	23.0	0.051
3.08	2.69	3163	98.9	2.8	20.6	0.062
2.69	2.44	3156	99.4	2.7	17.7	0.068
2.44	2.27	3170	99.5	2.7	13.5	0.074
2.27	2.13	3186	99.5	2.6	11.4	0.089
2.13	2.03	3144	99.7	2.6	8.1	0.117
2.03	1.94	3169	99.3	2.6	5.7	0.151
1.94	1.86	3108	98.7	2.6	4.0	0.221
1.86	1.80	3118	98.2	2.6	3.0	0.270
100.00	1.80	31428	98.6	2.7	16.6	0.064

Table 3.1. Data collection statistics for crystals of Vt. Nat. 2.2 Å data set collected from a native orthorhombic crystal at SRS Daresbury, station 7.2 ($\lambda=1.488$ Å). *SeM2*: 1.8 Å data set collected from a SeM-Vt monoclinic crystal at SRS Daresbury, station 9.6 ($\lambda=0.87$ Å). Both crystal were flash-frozen and data collected at 100K. $R_{\text{sym}} = \sum |I - \langle I \rangle| / \sum \langle I \rangle$, where I is the observed intensity and $\langle I \rangle$ is the average intensity from multiple observations of symmetry-related reflections.

Compound	Concentration	Soaking time	Result
Ag			
Ag(NO ₃) ₂	20mM	28h ⁺	No diffraction (synchrotron)
	1mM	48h	Crystal turned yellow. Poor diffraction.
Au			
AuCl ₂	10mM	1h	Heavy cracking. Diffraction to 8 Å.
Ce			
Ce(III)	20mM	36h	Flat ΔPatt.
Er			
Er(III)Cl ₃	10mM	12h	Streaky diffraction (synchrotron)
Hg			
MMN	1mM	15min	Diffraction to 10 Å.
	0.1mM	30min	Diffraction to 8-9 Å.
	0.1mM	15min	Flat ΔPatt.
PCMB	0.1mM	1h	Flat ΔPatt.
	1mM	1h	Diffraction to 8-9 Å.
Hg(Ac) ₂	1mM	4h ⁺	Diffraction to ~5 Å but dataset unindexable.
TAMM	1mM	10min	Heavy cracking. Diffraction to 9 Å.
EMTS	10mM	2h	Crystal cracking. Diffraction to 3 Å. Flat ΔPatt.
Ir			
IrCl ₆	10mM	120h	Diffraction to 2.5 Å. Flat ΔPatt.
Mo			
Mo(VI)	10mM	12h ⁺	Crystal turned black. No diffraction.
	1mM	12h ⁺	Diffraction to 8 Å (synchrotron).
Pb			
Pb(NO ₃) ₂	1mM	4h	Diffraction to 2.5 Å. Flat ΔPatt.
Pb(Ac) ₂	1mM	15h	Diffraction to ~4 Å. Flat ΔPatt.
Pt			
K ₂ Pt(II)Cl ₄	1mM (progressive)	12h	Diffraction to 10 Å.
K ₂ Pt(IV)Br ₆	1mM	16h	Diffraction to 4 Å. Flat ΔPatt.
	2mM	48h	No diffraction.
PIP	2mM	36h	Diffraction to ~10 Å.
Sm			
Sm(Ac) ₂	2mM	12h	Diffraction to 7 Å.
	2mM	44h ⁺	Diffraction to 3 Å. Flat ΔPatt.
Tl			
TlNO ₃	20mM	28h	Streaky diffraction (synchrotron).
U			
UO ₂ (NO ₃) ₂	0.1mM	2h	Crystal cracking. Diffraction to ~7 Å.
W			
Na ₂ WO ₄	10mM	12h ⁺	Poor diffraction (synchrotron).

Table 3.2. Some of the heavy metal compounds screened using monoclinic crystals. Crosses indicate soaks carried out in the absence of A.S. (replaced by Li₂SO₄). ΔPatt = isomorphous difference Patterson; *MMN*=methyl mercury nitrate; *PCMB*=p-chloro-mercuribenzoic acid; *TAMM*=tetrakis (acetoxymethyl)mercurimethan; *PIP*=diμiodobis (ethylenediamine) diplatinum (II) nitrate.

Deriv	Site	x	y	z	Occ	Aocc	B	Residue	Molec.	
EMTS1	a	.248	.198	.023	.858	.369	15.9	C950	A	
	b	.874	.968	.027	.601	.312	28.9	C950	A'	
	c	.078	.717	.136	.454	.246	30.3	C985	A	
	d	.455	.479	.232	.444	.208	9.7	C985	A'	
	e	.915	.877	.129	.612	.312	41.3	C972	A	
	f	.210	.595	.076	.436	.189	24.4	C972	A'	
EMTS2	a	.250	.197	.022	.853	.556	19.7	C950	A	
	b	.873	.963	.024	.462	.336	45.0	C950	A'	
	c	.081	.712	.136	.362	.227	25.4	C985	A	
	d	.456	.476	.232	.408	.283	14.6	C985	A'	
	e	.922	.873	.127	.378	.298	44.8	C972	A	
	f	.208	.600	.077	.355	.273	48.0	C972	A'	
	g	.908	.820	.138	.420	.276	22.4	C985	A	
	h	.383	.546	.008	.409	.312	30.4	C985	A'	
Deriv	Resol (Å)	11.99	8.56	6.65	5.44	4.60	3.99	3.52	3.15	Total
EMTS1	PhP_a	1.96	2.03	2.50	2.58	1.81	1.39	1.36	1.53	1.63
	PhP_c	1.49	1.79	1.94	1.86	1.48	1.05	1.08	1.13	1.33
	R _{cullis}	0.95	0.92	0.88	0.90	0.93	0.96	0.98	1.00	0.97
Deriv	Resol (Å)	11.31	7.89	6.05	4.91	4.13	3.57	3.14	2.80	Total
EMTS2	PhP_a	2.66	2.28	2.58	2.12	1.41	1.29	1.40	1.63	1.57
	PhP_c	1.81	1.86	2.01	1.86	1.12	0.94	1.05	1.06	1.27
	R _{cullis}	0.75	0.71	0.70	0.75	0.81	0.88	0.94	0.98	0.89

Table 3.3. Heavy metal sites. Final parameters and statistics are according to MLPHARE (CCP4 suite), and sites are identified by residue. NCS-related sites were identified with the programme FINDNCS (<http://gamma.mbb.ki.se/~guiguang/findncls.html>). A and A' are the NCS-related molecules. *Occ* = real occupancy; *Aocc* = anomalous occupancy; $R_{cullis} = \frac{\sum |F_{der} \pm F_{nat}| - |FH|}{\sum |F_{der} - F_{nat}|}$ for centric reflections only, where F_{nat} is the protein structure factor amplitude, F_{der} the heavy-atom derivative structure factor amplitude, and FH is the rms heavy-atom structure factor amplitude. *PhP* = phasing power, defined as the root-mean-square ($|FH|/E$) with E the residual lack of closure error. *PhP_a* and *PhP_c* is the phasing power for acentric and centric reflections respectively.

	Soln	α	β	γ	peak height/ σ	
Cross-rotation	1	227.76	72.24	118.98	27.9	
	1'	312.24	107.76	298.98	27.9	
	2	312.40	69.65	265.16	24.7	
	2'	227.60	110.35	85.16	24.7	
		x	y	z	C	R
Translation	1-2	0.8606	0.8625	0.6095	39.9	48.9
	1-2'	0.7007	0.5324	0.4200	20.4	55.6
		α	β	γ		
Rigid body refinement	1-2-i	228.53	71.10	118.47		
	1-2-ii	312.28	69.66	265.25	57.7	43.6

Table 3.4. Molecular replacement solutions of the monoclinic crystal. The molecular replacement search was carried out using AMORE. The cross-rotation function was calculated using a sphere radius of 15 Å, resolution limits of 10-3 Å and an angular search step of 2.5°. C = standard linear correlation coefficient between observed and calculated structure factor amplitudes defined as $C = [\sum_h (|F_h(\text{obs})| - \langle |F_h(\text{obs})| \rangle) \times (|F_h(\text{calc})| - \langle |F_h(\text{calc})| \rangle)] / [\sum_h (|F_h(\text{obs})| - \langle |F_h(\text{obs})| \rangle)^2 \times \sum_h (|F_h(\text{calc})| - \langle |F_h(\text{calc})| \rangle)^2]^{1/2}$; $R = R_{\text{free}}$ as defined in the legend to table 3.6.

Crystal Form	Round	Procedure/ Remarks	Resolution (Å)	R _{work}	R _{free}	
Orthorhombic	1	RB	10-2.2	36.7	38.4	
	2	SA BG, BI		29.8 25.8	38.2 32.3	
	3	74 waters SA, BG, BI	25.6	31.8 31.0		
	4	15 waters SA, BG, BI	25.6	31.7 29.6		
	5	13 waters SA, BG, BI	24.8	31.3 28.9		
	Monoclinic	1'	RB	10-1.7	31.2	33.1
		2'	SA, BG, BI 137 waters		25.9 24.9	28.0 26.7
		3'	SA, BG, BI 66 waters	23.4	26.2 25.6	
		4'	SA, BG, BI Weight	22.2	25.9 25.5	
			40 waters, SO ₄ ²⁻ SA, BG, BI	22.8	25.2 25.1	
5'		BG, BI	10-1.8	20.9	24.8	

Table 3.5. Model building and refinement statistics. Each round consisted of a map inspection and model modification with FRODO followed by refinement with CNS. After the first round, 5% of the reflections were excluded from the refinement and used to calculate a free R-factor (R_{work} and R_{free} are defined in the legend to Table 3.6). The maximum SA temperature was 3000 K in the initial rounds and 1500 K by round 4'. Added waters and ions are indicated. Molecular replacement into the monoclinic form took place after round 3 (the orthorhombic form continued to be refined separately). *RB*=rigid body refinement, *SA*=simulated annealing, *BG*=group B-factor, *BI*=individual B-factor, *Weight*= optimisation of X-ray and geometry weight.

a)

	SeM1	EMTS1	EMTS2	Se_SAD
Resolution range (<i>d</i> min) (Å)	25-2.1 (2.2-2.1)	25-2.8 (2.9-2.8)	25-2.8 (2.9-2.8)	20.0-3.5 (3.6-3.5)
Completeness (to <i>d</i> min) (%)	96.4 (68.4)	93.4 (63.3)	97.3 (73.4)	99.5 (99.4)
Reflections (total/unique)	82196/24624	34537/10405	75407/19439	20715/9245
R_{sym} (to <i>d</i> min) (%)	6.3 (22.8)	8.3 (31.6)	6.4 (15.4)	5.4 (10.2)
$I/\sigma I$ (to <i>d</i> min)	19.2 (3.8)	13.0 (2.7)	19.6 (6.3)	14.8 (9.4)
R_{iso} (%)		30.5	28.5	19.8
Phasing power (cen/acen)		1.34/1.63	1.27/1.57	—
R_{cullis} (cen/acen/anom)		0.66/0.71/0.97	0.68/0.73/0.89	—
FOM		0.51	0.51	—
Number of sites		6	8	18 (by XDF)

b) Refinement Statistics

	Nat	SeM2
Non-hydrogen atoms in a.u.	2711	2808
Water molecules	101	319
<i>d</i> -spacings (Å)	10-2.2	10-1.8
R_{work} (%)	22.5	20.0
R_{free} (%)	28.9	24.8
Rmsd bond lengths (Å)	0.006	0.010
Rmsd bond angles (°)	0.98	1.05

Table 3.6. Summary of all data collection, phasing and refinement statistics. a) Statistics for orthorhombic SeM-Vt crystals: *EMTS1* and *EMTS2*, high and low soaks of ethylmercurythiosalicylate; *SeM1* ($\lambda=1.488$ Å), *Se_SAD* ($\lambda=1.000$ Å). Integration and data reduction were carried out using DENZO and SCALEPACK. SeM1 was used for the SIRAS phasing. Heavy atom parameters were refined using MLPHARE. Phasing with EMTS1 was carried out to 3.15 Å resolution. The temperature factors for all sites were refined isotropically. R_{sym} , phasing power and R_{cullis} (centric) are defined in the legends to Tables 3.1 and 3.3. R_{cullis} (anomalous) = $\sum ||\Delta F_{\text{PH}}^{\pm}(\text{obs})| - |\Delta F_{\text{PH}}^{\pm}(\text{calc})|| / \sum \Delta F_{\text{PH}}^{\pm}(\text{obs})$ where $\Delta F_{\text{PH}}(\text{obs})$ is the structure factor amplitude difference between symmetry related reflexions (Bijvoet pairs) and $\Delta F_{\text{PH}}(\text{calc}) = 2f''|FH|\sin(\alpha_{\text{PH}} - \alpha_{\text{H}})/f'$; $R_{\text{iso}} = \sum ||F_{\text{der}}| - |F_{\text{nat}}|| / \sum |F_{\text{nat}}|$ where F_{nat} and F_{der} are the protein and heavy-atom derivative structure factors respectively; *FOM* = figure of merit, defined as the weighted mean of the cosine of the deviation of the phase angle from α_{best} : $\langle \cos\{\alpha - \alpha_{\text{best}}\} \rangle$; *XDF* = cross-difference Fourier; b) Refinement statistics for final Vt models. *Nat*: orthorhombic native form, *SeM2*: SeM-Vt monoclinic form. $R_{\text{work}} = \sum |FP_{\text{o}} - FP_{\text{c}}| / \sum |FP_{\text{o}}|$, where FP_{o} and FP_{c} are the observed and calculated protein structure factor amplitudes respectively; R_{free} is defined as for R_{work} but only for reflexions belonging to a test set of unique reflections. R_{free} was computed using 5% of the total reflections selected randomly and not used in refinement.

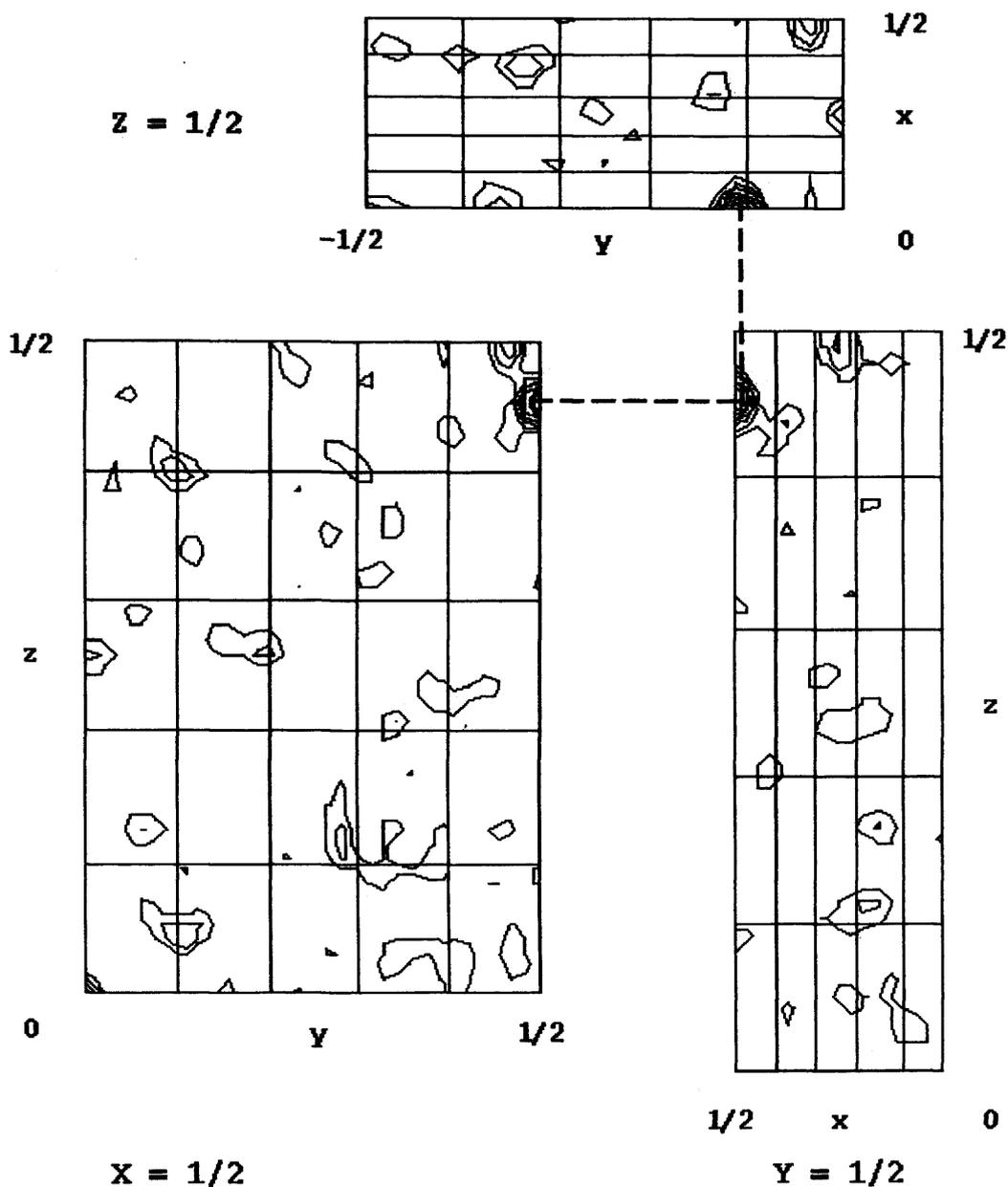


Figure 3.1. Harker sections of EMTS2 isomorphous difference Patterson. Calculated from 30 to 4 Å and contoured at 1,2,3,... sigma above the mean. The limits on each section axis are shown. The peak seen on all three Harker sections is due to a Hg atom bound to Cys 985. For definition of the isomorphous difference Patterson function see legend to Figure 3.2.

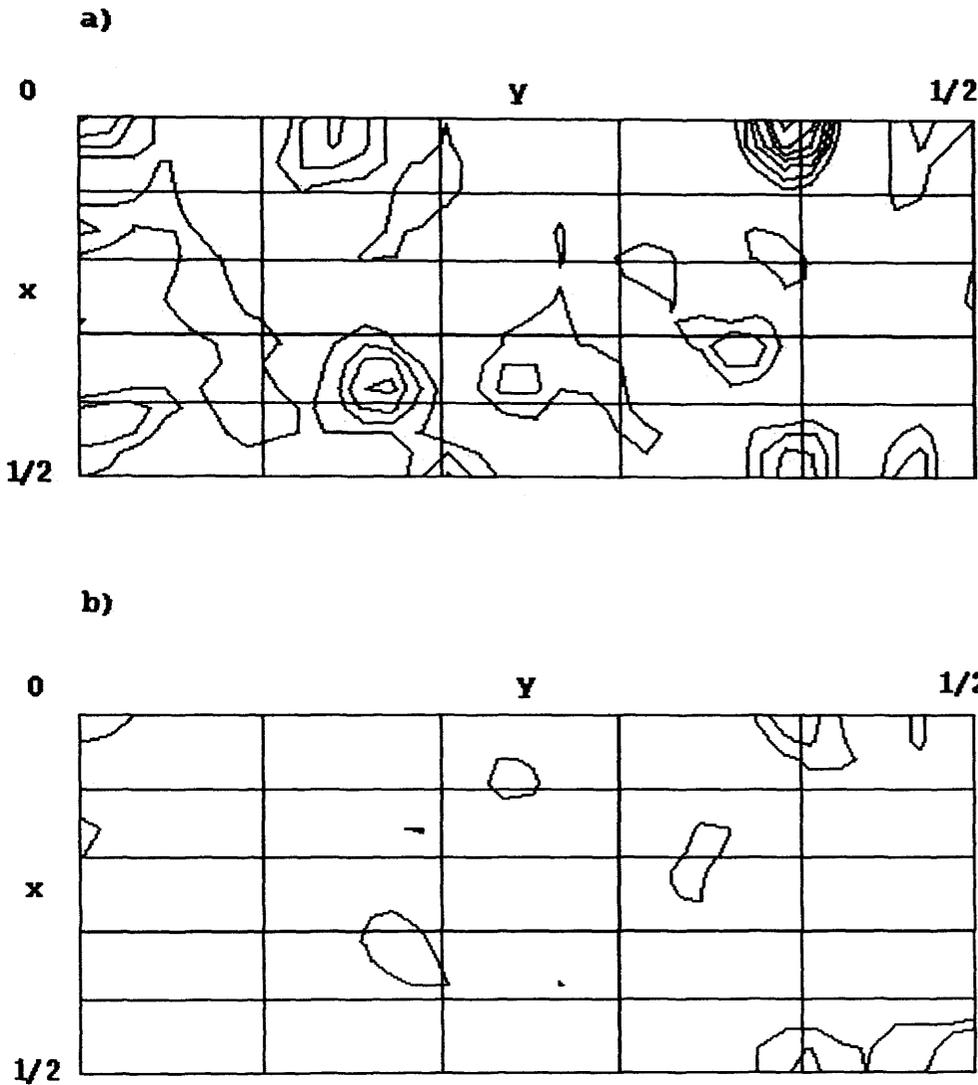


Figure 3.2. $z=1/2$ Harker sections of EMTS1 difference Pattersons. a) isomorphous difference Patterson, b) anomalous self-difference Patterson. Calculated and contoured from 30 to 4 Å. The 7σ peak shown in a) is due to a Hg atom bound to Cys 985. The difference Patterson function is defined as $\Delta P(u,v,w)=V^{-1}\sum\sum\sum_{hkl}\Delta F_{hkl}^2 e^{-2\pi i(hu+kv+lw)}$ where $(\Delta F)^2=(|F_{PH}|-|F_P|)^2$ is the amplitude contribution of the heavy atom to structure factor F_{PH} and V the volume of the unit cell of derivative crystals. $(\Delta F_{iso})^2=|F_H|^2 \cos^2(\alpha_{PH}-\alpha_H)=(1/2)|F_H|^2+(1/2)|F_H|^2 \cos 2(\alpha_{PH}-\alpha_P)$ in the case of an isomorphous difference Patterson function and $(\Delta F_{ano})^2=|F_H|^2 \sin^2(\alpha_{PH}-\alpha_H)=(1/2)|F_H|^2-(1/2)|F_H|^2 \cos 2(\alpha_{PH}-\alpha_P)$ in the case of an anomalous difference Patterson. F_H and $\alpha_{PH}-\alpha_H$ are the structure factor and phase angle of the heavy atom respectively.

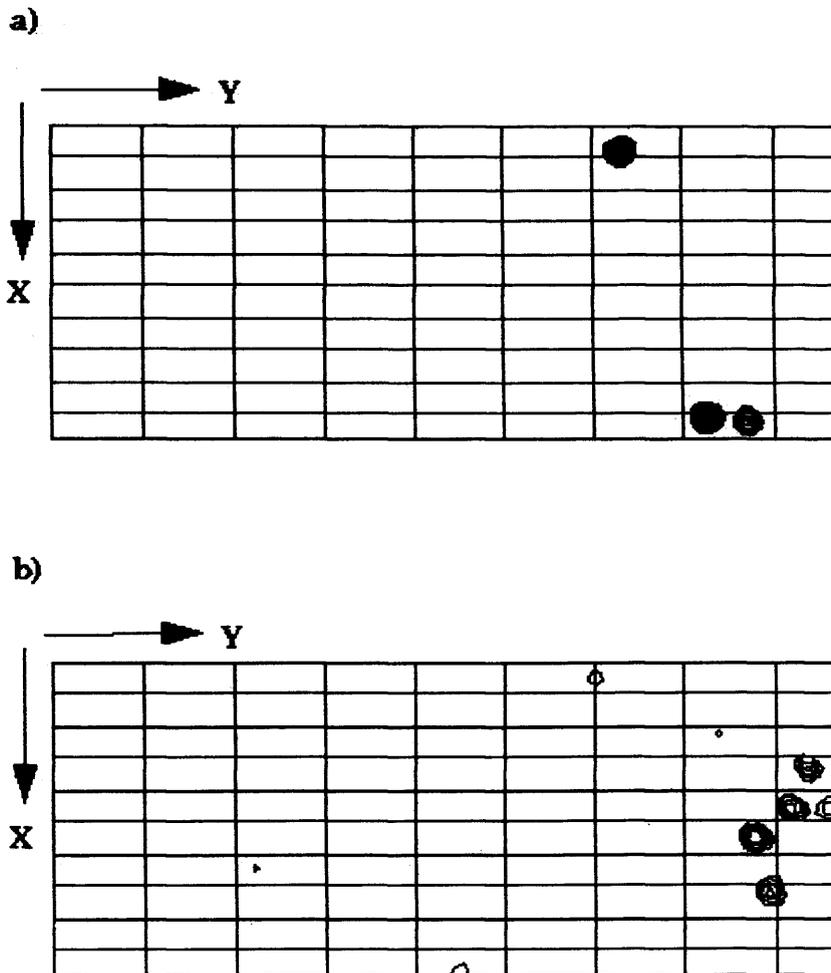


Figure 3.3. Difference Fourier of EMTS2. *a)* Hg (EMTS2) self-difference Fourier, *b)* Hg-Se (EMTS2-Se_SAD) cross-difference Fourier. The difference Fourier summation is defined as $\Delta\rho(xyz)=(1/V)\sum_{hkl}\Delta|F(hkl)|_{iso}\cos[2\pi(hk+ky+lz)-\alpha_p(hkl)]$, where α_p are the phase angles of the protein. The peaks seen in *b)* are due to Se atoms contained in the methionine cluster (see Chapter 4). See also text for details.

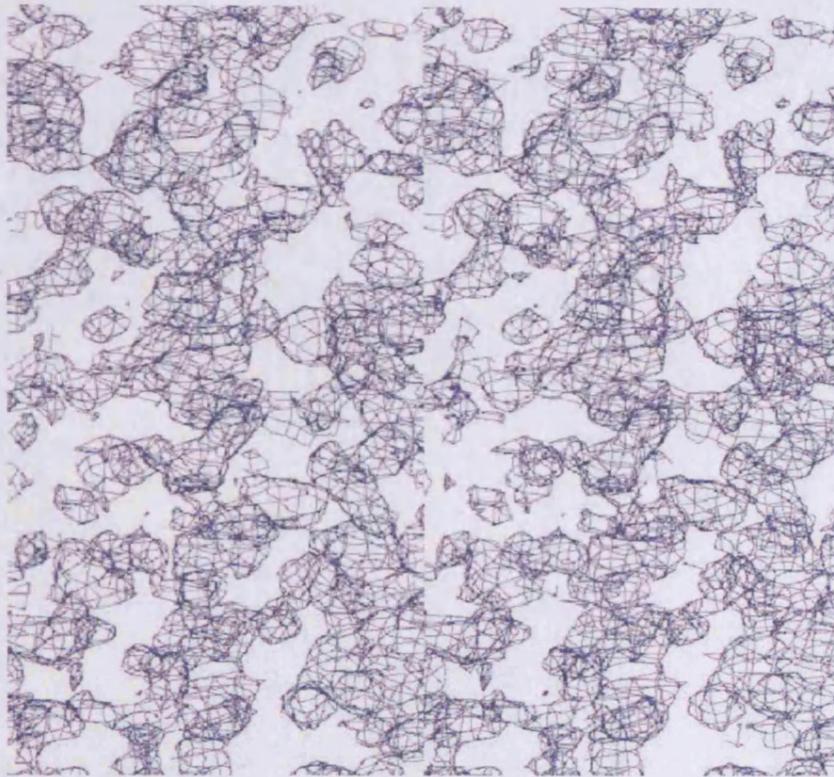


Figure 3.4. Stereo view of the first experimental electron density map of Vt. Calculated using SIRAS phases from EMTS2, the resolution is to 2.8 Å using a 1σ contour level (blue). Continuous helical density is evident. Figures 3.4 and 3.5 were both generated with XTAL VIEW (McRee, 1992) , Render (Kraulis, 1991) and Raster-3D (Merrit and Murphy, 1994).

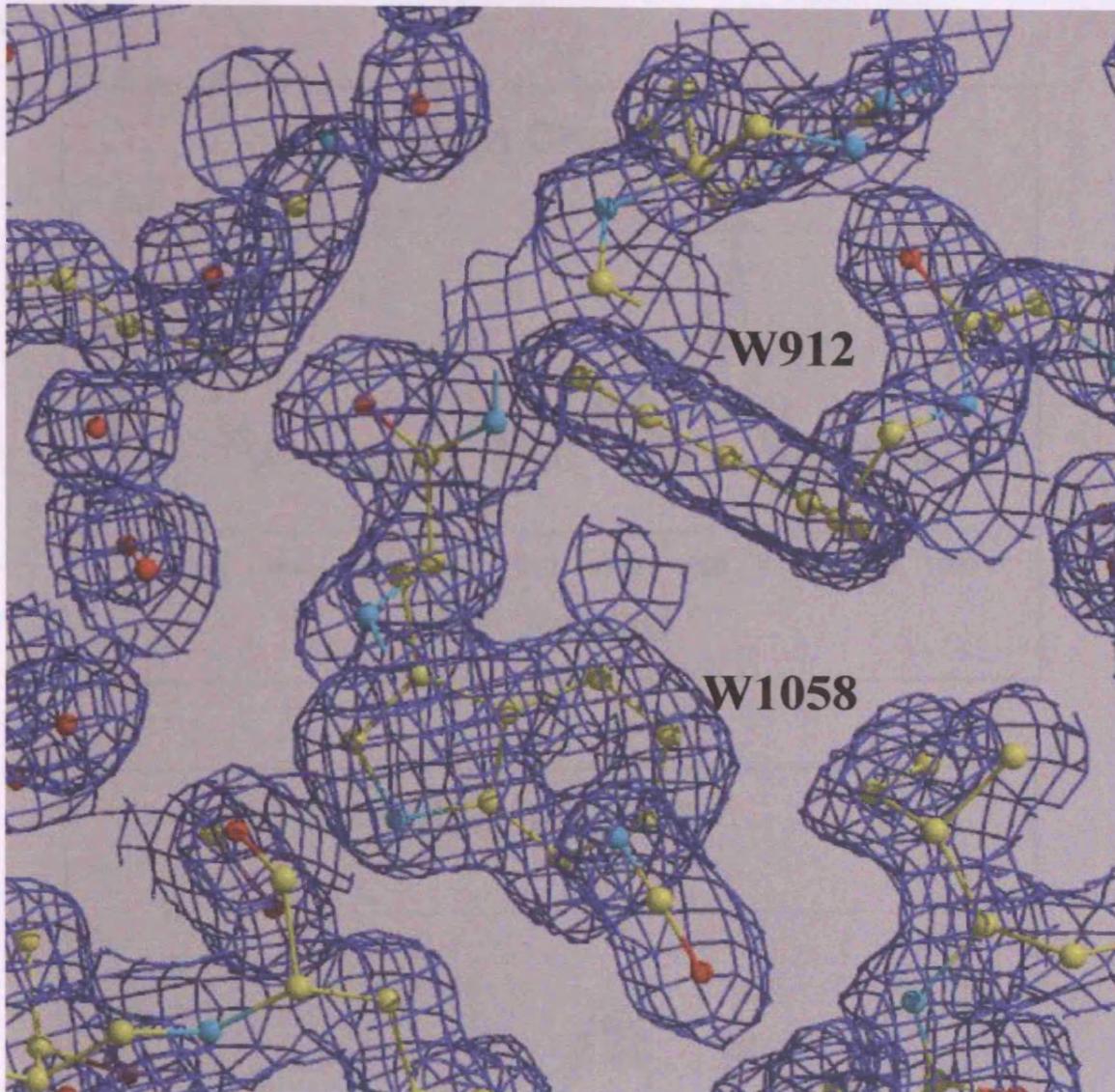
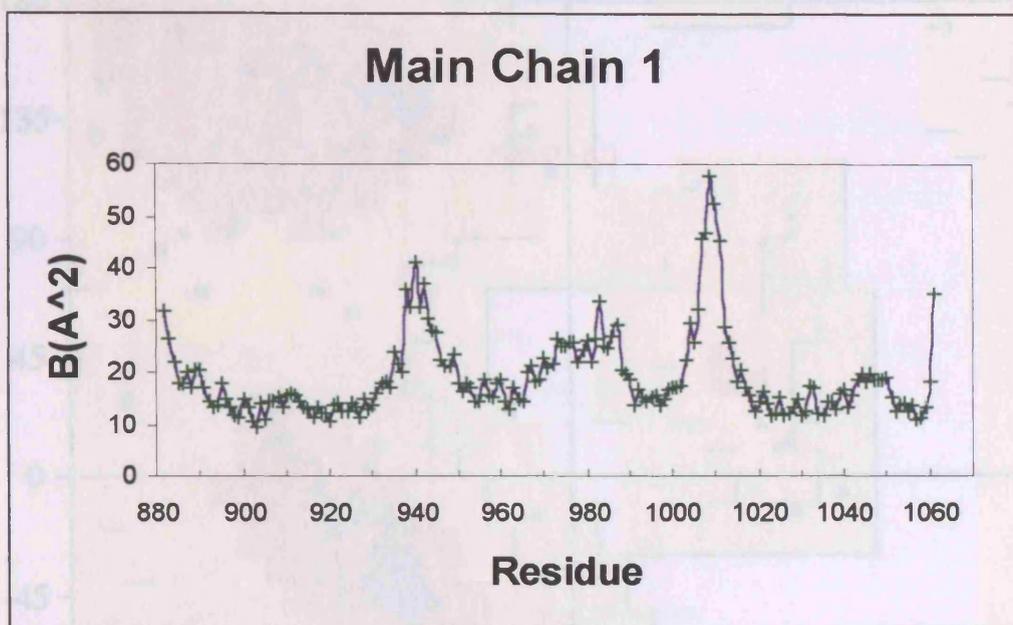


Figure 3.5. Tryptophan stacking. Detail from the final Vt model in ball-and-stick mode. Electron density is calculated to 1.8 Å and contoured at 1 sigma above the mean.

a)



b)

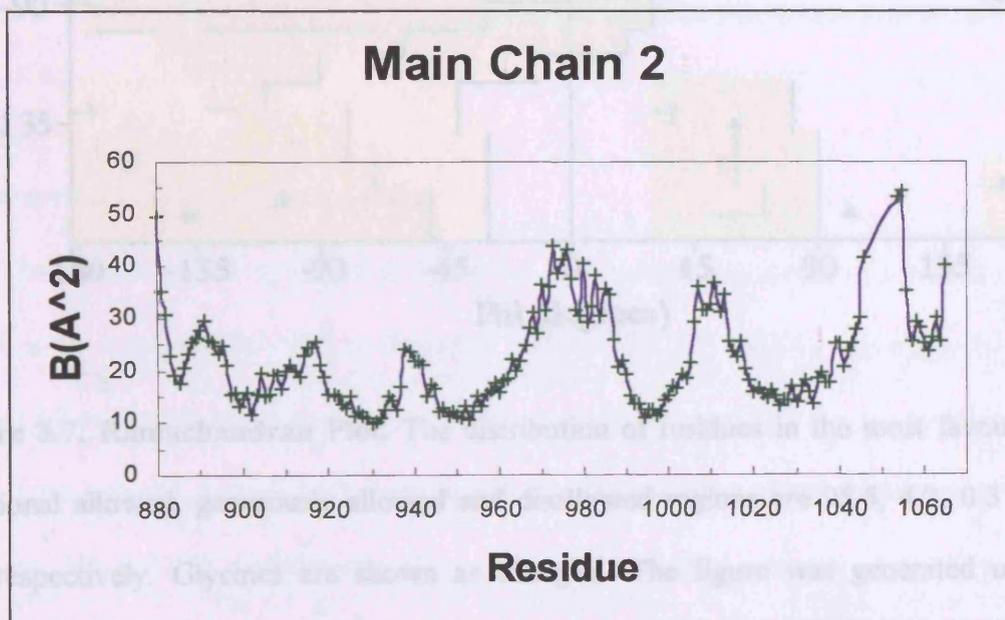


Figure 3.6. Average B factors for main chain atoms in the monoclinic model of Vt. a) molecule 1, b) molecule 2. Residues 1062-1066 in molecule 1 and 1047-1053 in molecule 2 display discontinuous or absent electron density and have therefore been omitted from the figure.

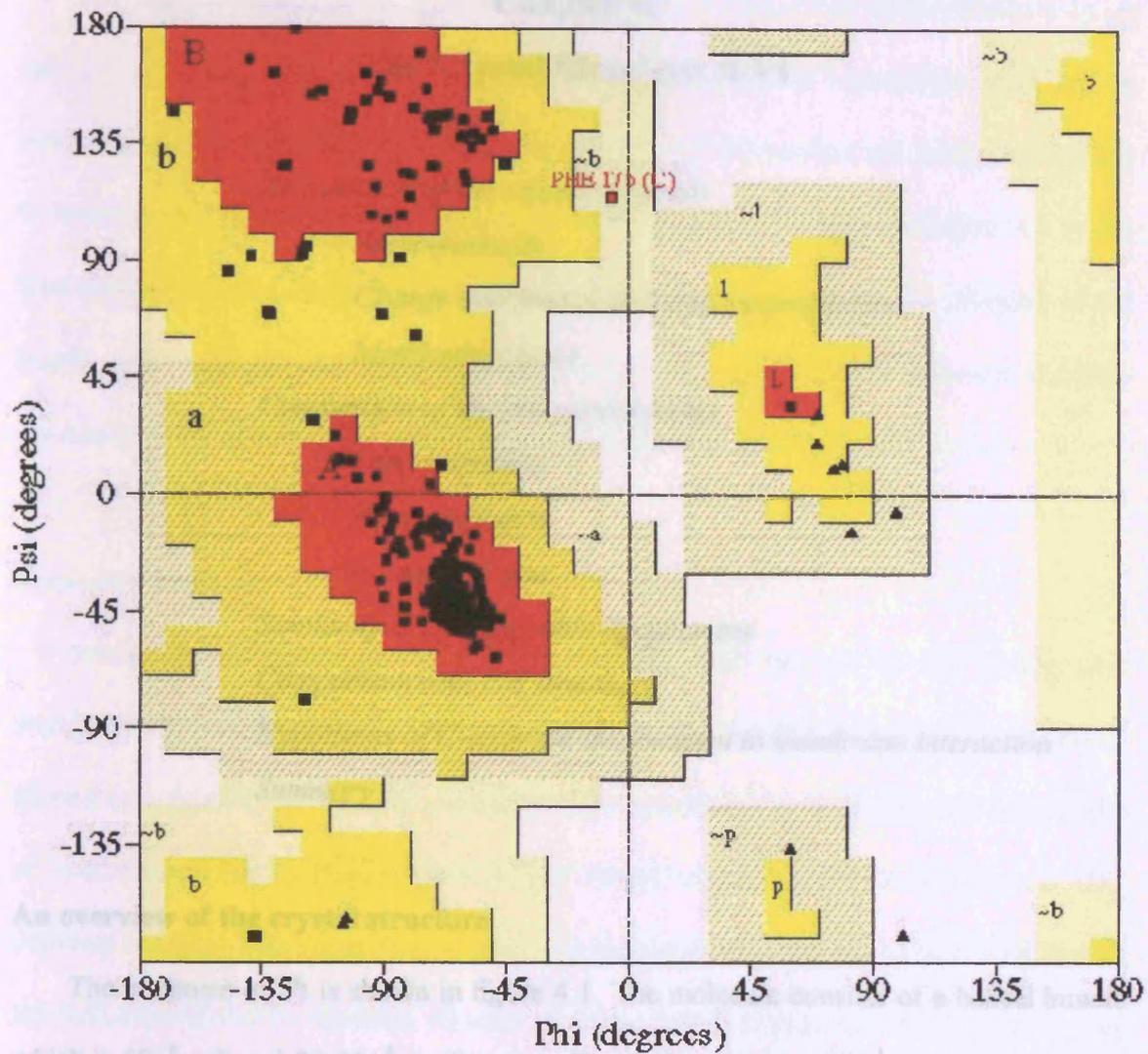


Figure 3.7. Ramachandran Plot. The distribution of residues in the most favoured, additional allowed, generously allowed and disallowed regions are 95.5, 4.2, 0.3 and 0% respectively. Glycines are shown as triangles. The figure was generated using PROCHECK (Laskowski *et al.*, 1993).

Chapter 4:

The Crystal Structure of Vt

An overview of the crystal structure

Amphipathicity

Charge distribution and surface properties

Methionine zipper

Comparison of the two crystal forms

Dimer contacts

N-terminal arm

C-terminal arm

Similarity to exchangeable lipoproteins

Comparison with EM structures

Hypothesis of C-terminal involvement in membrane interaction

Summary

An overview of the crystal structure

The structure of Vt is shown in figure 4.1. The molecule consists of a helical bundle which is 60 Å tall and 20-30 Å in diameter. Five helices (H1 to H5), connected by short loops (3-8 residues), adopt an antiparallel down-up-down-up-down circular topology. All helices are straight, except for H4 which kinks in the middle at P989. Their length varies pairwise; H1 and H2 are relatively short (16 and 20 residues respectively), H3 and H4 (29 and 31 residues) form a closely antiparallel hairpin, while H5 is the longest helix (34 residues, 52 Å long) and makes an angle of 20° with the H3-H4 hairpin which brings it nearly parallel to H1.

There are remarkably few salt bridges within or between the helices (table 4.1). A salt bridge linking H2 and H3 is seen in both copies of the monoclinic form and is conserved among vinculins of different species (figure 4.2), while a salt bridge linking the C-terminal arm (see below) to H2 and H3, is strictly conserved in vinculins and in the homologous protein, α -catenin, and could play a role in maintaining the integrity of the bundle (see below). The remaining salt bridges exhibit variation between the two molecules or do not exhibit a strong degree of conservation.

Amphipathicity

Membrane-associated helices can be classified with the aid of the hydrophobic moment plot (Eisenberg *et al.*, 1984), on which the hydrophobic moment of each helix is plotted as a function of its hydrophobicity. Hydrophobicity is defined as the vector sum of hydrophobicities as plotted on a helical wheel assuming perfect helicity and 100° between residues. The magnitude of hydrophobic moment measures the amphiphilicity of the helix (and hence its tendency to seek a surface between hydrophobic and hydrophilic phases) and the hydrophobicity measures its affinity for the membrane interior. Segments of membrane proteins in α -helices tend to fall in one of three regions of a hydrophobic moment plot: a) monomeric transmembrane anchors, which lie in the region of highest hydrophobicity and smallest hydrophobic moment; b) helices presumed to be paired (e.g. transmembrane segments of surface immunoglobulins) and helices which are bundled together in membranes (such as bacteriorhodopsin) fall in the adjacent region with higher hydrophobic moment and smaller hydrophobicity; and c) helices from surface-seeking proteins (such as melittin) which fall in the region with still higher hydrophobic moment.

All five helices of Vt are markedly amphipathic, with their hydrophobic sidechains mostly pointing into the hydrophobic core of the bundle and the hydrophilic groups pointing outwards. The amphipathicity of each helix peaks in the lower half of the molecule, as judged by Eisenberg's hydrophobic moment (table 4.2). The peak hydrophobic moment and hydrophobicity values of H3 (0.80 and -0.26 respectively) are similar to those of the bee venom peptide, melittin (sequence: GIGAVLKVLTTGLPALISWIKRKRQQ), and place it firmly in the membrane "surface-seeking" category.

Charge distribution and surface properties

The combination of short and long helices, creates a large crevice (~15 Å) at the top of the bundle, rimmed by the H2-H3 turn, the H4-H5 turn, and an acidic cluster (DEESEQATE, residues 1013-1021) at the beginning of H5 (figure 4.3). The crevice is uncharged except for a partly buried lysine at its centre (K1002) emerging from the otherwise hydrophobic surface of H4.

Positive charges cluster in two regions on the surface of the domain: a basic collar surrounding the C-terminal hairpin (see below), and a basic ladder which is centered on H3 (figure 4.3). The upper half of H4, after P989, presents the most hydrophobic exposed surface of any helix in the bundle. In both crystal forms, this surface forms part of the crystallographic dimer interface (see below).

Methionine zipper

An unusual feature of the largely aliphatic core is a methionine zipper (by analogy to a leucine zipper), which links five methionines, one in H1, three in H2 and one in H5 (figure 4.4). Three of these methionines (M926, M930 and M1031) are invariant in the known vinculin sequences and in those of α -catenin. The distances between these methionines and an adjacent cysteine (C950) vary between approximately 4 and 5 Å, although the presence of alternative methionine conformers (viewed as distinct electron density during model building) indicates a higher flexibility of these sidechains and a degree of distribution of their interactions inside the aliphatic core.

Leucine zippers are 7-residue motifs mediating homo- or hetero-dimerisation interactions in a number of cytoskeletal proteins, including actin cross-linking and intermediate filament proteins (Branden and Tooze, 1999). In an algorithm using sequences of known coiled-coil proteins (Lupas et al., 1991), Vt scores high in three regions (figure 4.5.a) involving H2, H3 and H5 (figure 4.5.b) which contain four out of the five interacting methionines (M926, M930, M933 and M1031). These same helices have been shown to dimerise when expressed as distinct peptides (see p.17) indicating that, if exposed, the hydrophobic interface of H2, H3 and H5 might have the ability to induce oligomerisation of Vt. Atypical zipper domains consisting of a heptad repeat of alanine and methionine rather than leucine residues have been previously reported (Peleg and Metzenberg, 1994).

Comparison of the two crystal forms

The four copies of Vt differ only in the conformations and degree of order of the connecting loops and chain termini. A superposition of all four C α traces of the molecule coloured according to B-values can be seen in figure 4.6. The rms deviation for helices 1 to 5 in all copies is between 0.5 and 0.7 Å. Regions of high flexibility ($>30 \text{ \AA}^2$) include the H2-H3 turn, the H4-H5 turn, the loop and hairpin region of the C-terminal arm, and the first half of H4 (residues 975-988).

The marked difference between the B-values of one of the four copies of Vt (see figure 3.6), is likely the result of a salt-bridge formed by D974 and K1047. These two residues which are strictly conserved (Figure 4.2), link the H3-H4 helices to the C-terminal arm. This linkage is only visible in the Vt copy displaying continuous electron density in the 1047-1052 C-terminal arm region (see below) and results in a stabilisation (as judged by main-chain B values) of both the H2-H3 turn and the first half of H4, prior to the kink induced by P989.

Dimer contacts

The dimer contacts observed in the two crystal forms are given in table 4.3. In both cases, contacts are established between the second half of H4 (residues 993-1001) and the first half of H5 (residues 1015-1022). The distances are essentially conserved in both crystal forms (only contacts between 996-1015 and 997-1015 are significantly different), and are primarily hydrophobic in nature with the exception of a possible salt-bridge (K996-E1015; orthorhombic form) and a hydrogen bond (Q994-Q1018; both crystal forms).

The second half of H4 presents the most hydrophobic exposed surface of any helix in the bundle. It includes I997 and V1001, which are fully exposed to solvent, a cluster of exposed threonines (T990, T993, T1000 and T1004), and a single charged residue, K996. This region has been implicated in binding paxillin (residues 978-1000), since a fusion peptide spanning residues 881-1000 (corresponding to the N-terminus of Vt through H4) binds paxillin *in vitro*, while 881-978 (N-terminus through H3) does not (Wood et al., 1994). Similarly, the first half of H5 has been implicated in focal adhesion targeting (residues 1000-1028), since a fusion peptide spanning residues 398-1028 targets to focal adhesions while 398-1000 does not (Wood et al., 1994). The interaction with paxillin appears not to be affected by vinculin conformation (Gilmore and Burridge, 1995) and Vt shows only a weak propensity towards dimerisation in solution (Johnson and Craig, 1998). The crystal dimer is therefore not likely to reflect the physiological state of the protein in solution.

N-terminal arm

The conformation of the N-terminal arm is very similar in the two molecules of the monoclinic crystal form. Upstream of H1, the chain makes an abrupt right-angled turn via a short α -helix (H0) and an N-terminal strand that packs against H1 and H2, making both salt-bridges and hydrophobic contacts. The structures deviate upstream of F885, the sidechain of which packs into a hydrophobic crevice formed by sidechains from H1 and H2. However, in both cases, the N-terminus is located close to the C-terminal hairpin.

In both molecules of the orthorhombic form, lattice contacts occur at the position that would be occupied by H0; in one molecule, H0 is replaced by a loop that takes a

different course, although F885 assumes a very similar position and orientation to that in the monoclinic crystals. In the second molecule, only very weak density is observed upstream of residue 890.

Owing to their very similar conformation, higher degree of order and lack of crystal contacts, it can be assumed that the N-terminal arms in the monoclinic crystals better reflect the structure in solution.

C-terminal arm

A C-terminal arm follows the last helix, H5. The arm contains 4 of the 6 aromatic residues in Vt, and can be divided into three segments: a flexible loop (residues 1047-1052), a β -strand (residues 1053-1061), and a hydrophobic hairpin (residues 1062-1066). Two sidechains of the β -strand, L1056 and W1058, seal the base of the helical bundle by inserting into the hydrophobic core in a cavity created by H1, the H1-H2 turn and H5. W1058 packs at right angles to W912 from the H1-H2 turn (figure 4.1*b*). These two tryptophans are the only aromatic residues in the hydrophobic core, and are conserved in the vinculins from different species and in α -catenin. The β -strand emerges beneath the turn between H1 and H2 and ends in a five-residue hairpin (TPWYQ) containing a third tryptophan (W1064). This hairpin is ordered to different extents in the two crystal forms. In the best ordered case, the C-terminal carboxylate and glutamine side chain point back towards the body of the domain, completely exposing W1064 and Y1065, which protrude 15 Å from the body of the domain into the solvent. The hairpin is surrounded by a basic collar at the base of the domain, composed of arginines and lysines from the C-terminus, the H1-H2 turn, and H5 (figure 4.3*b*).

Remarkably, in its most ordered form (orthorhombic crystal), most of the hairpin (residues 1062-1065) appears free of protein or solvent contacts (data not shown). However, crystal contacts are observed between the N-terminal arm and residues of the C-terminal β -strand as well as Q1066. The ordered state of the hairpin in this crystal form (residues TPWY) could therefore result from stereochemical constraints brought about by the clamping down of both the C-terminal β -strand and the final residue of the molecule (Q1066). This hypothesis finds some support in the other crystal form where, due to a slight tilt in the NCS-axis with respect to the one observed in the orthorhombic form, intermolecular contacts are less extensive (see table 4.3). In this case, although W1064 displays some contacts with adjacent residues (T1062, P1063) and with ordered solvent molecules, neither Y1065 or Q1066 are visible in the crystal.

Similarity to exchangeable lipoproteins

A search of the Protein Data Base for structural homologues of Vt using the DALI algorithm (Holm and Sander, 1997), revealed that although the helix bundle is a common architecture, the length of the helices and overall dimensions of Vt are unusual. The most similar structure found was apolipoprotein E (apoE), with an r.m.s. deviation of 3.5 Å over 122 residues encompassing all four helices of apoE and helices H2-H5 of Vt (DALI “Z” score = 9.7). ApoE adopts a four-helix amphipathic bundle (Wilson *et al.*, 1991) that has the same up-down-up-down circular topology as Vt. Furthermore, the lengths of the helices and shapes of the molecules are very similar.

Five-helix bundles are uncommon, and the only one in the data base with similar size and shape to Vt is apolipophorin (Breiter *et al.*, 1991). Apolipophorin is an 18 kDa

apolipoprotein that plays a similar role to apoE in insect cells. The connectivity between helices is different, however, so that it scores less well in the DALI search. The three structures are compared in Figure 4.7.

Comparison with EM structures

In electron microscopic images of intact chicken vinculin, Vt emerges as four “pearls-on-a-string” comprising a tail of approximate length 190 Å (Winkler *et al.*, 1996), whereas images of the intact pig vinculin show a globular structure with no “tail” evidence (Gimona *et al.*, 1987).

One possibility is that in the images of chicken vinculin the helical bundle is unfurled, as the total length of the helical segments in the Vt crystal structure is ~200 Å. In this case, the “pearls” could constitute helical hairpins with the chicken vinculin images being representative of an open, active conformation, and those of pig vinculin showing the inactive conformation of the protein (see also Chapter 6).

Hypothesis of C-terminal involvement in membrane interaction

Tyrosine and tryptophan residues are the most common residues in membrane-binding proteins at the boundary of the polar/non-polar interface between the hydrocarbon core and the hydrophilic head groups (Deisenhofer and Michel, 1989; Wimley and White, 1996). The three-dimensional structure of Vt creates a collar of positive charge around the C-terminal hairpin (residues TPWYQ), with contributions from several loops and helices. The structure thus suggests a model of membrane insertion whereby the positively charged collar would be attracted to the negatively

charged membrane surface causing nonspecific adsorption of the protein onto the membrane surface, and would subsequently insert the hydrophobic hairpin into the lipid bilayer in specific interactions of the aromatic residues with acidic lipids.

A similar mode of membrane association has been proposed for annexin V, an eukaryotic calcium-dependent membrane-binding protein, in which a positively charged surface of the protein binds to the acidic surface, while a tryptophan residue inserts into the membrane (Campos *et al.*, 1998).

Summary

The crystal structure of Vt is composed of five amphipathic helices forming an antiparallel bundle. Structural features are summarised in figures 4.1 and 4.3. The top of the bundle is characterised by a large crevice, while a C-terminal arm wraps across the base of the bundle and emerges as a hydrophobic hairpin surrounded by a collar of basic residues adjacent to the N-terminus. Comparison with EM structures of the intact vinculin and a resemblance to apolipoproteins, suggest an open-closed conformation of the molecule, while comparison with other membrane interacting proteins and peptides (annexin V, melittin), suggest a role of the C-terminal hairpin in interaction with acidic phospholipids.

Residue 1	Residue 2	Distance (Å)	Molecule	Position
D882	R910	2.7	2	N-terminus, H1
	K924	2.6	1	N-terminus, H2
E892	R903	2.7	1	H0, H1
		2.8	2	"
E908	K911	3.1	1	H1, H1
	R1039	3.1	2	H1, H5
R925	E960	2.9	2	H2, H3
		3.1	1	"
K966	E986	2.7	2	H3, H4
D974	K1047	3.2	1	H3-H4 turn, C-terminal arm

Table 4.1. Salt-bridges in Vt. Distances were calculated using the programme CONTACT (CCP4 suite) for both molecules of the monoclinic form. The only salt-bridge exhibiting conservation among vinculins from different species is R925-E960, while D974-K1047 is strictly conserved in both vinculins and α -catenins. The coordinate error is 0.2 Å.

Helix	Hydrophobic Moment (entire helix)	Mean Hydrophobicity (entire helix)	Peak Hydrophobic Moment	Z (overall charge)
H1	0.55	-0.33	0.74	+1(3/2)
H2	0.42	0.18	0.45	+1(3/2)
H3	0.58	-0.26	0.80	+3(7/4)
H4	0.21	-0.07	0.58	+5(6/1)
H5	0.33	-0.25	0.69	-6(2/8)
Melittin	0.34	0.10	0.62	+5(5/0)

Table 4.2. Hydrophobic moments for the helices of Vt. Values for the peak hydrophobic moment were calculated using an 11-residue window.

Molecule 1		Molecule 2		Distance (Å)	
Residue	Atom	Residue	Atom	Monoclinic	Orthorhombic
Q994	NE2	Q1018	CD	3.6	3.5
Q994	NE2	Q1018	NE2	2.7	2.7
Q994	NE2	M1022	CG	(4.0)	3.4
Q994	OE1	M1022	CE	3.4	3.4
K996	CD	E1015	OE2/OE1	(4.1)	3.2
K996	NZ	E1015	OE2/OE1	(4.7)	3.3
I997	CD1	E1015	O	(4.7)	3.5
V1001	CG2	V1001	CG2	3.3	3.3
Q1018	CG	T993	CG2	3.6	3.5
Q1018	NE2	Q994	NE2	3.0	2.9
M1022	CG	Q994	NE2	(4.0)	3.5

Table 4.3. Dimer contacts in the two crystal forms. The contacts were calculated using the programme CONTACT (CCP4 suite). Numbers in brackets indicate distances greater than the probe radius used (3.6 Å). The coordinate error is 0.2 Å.

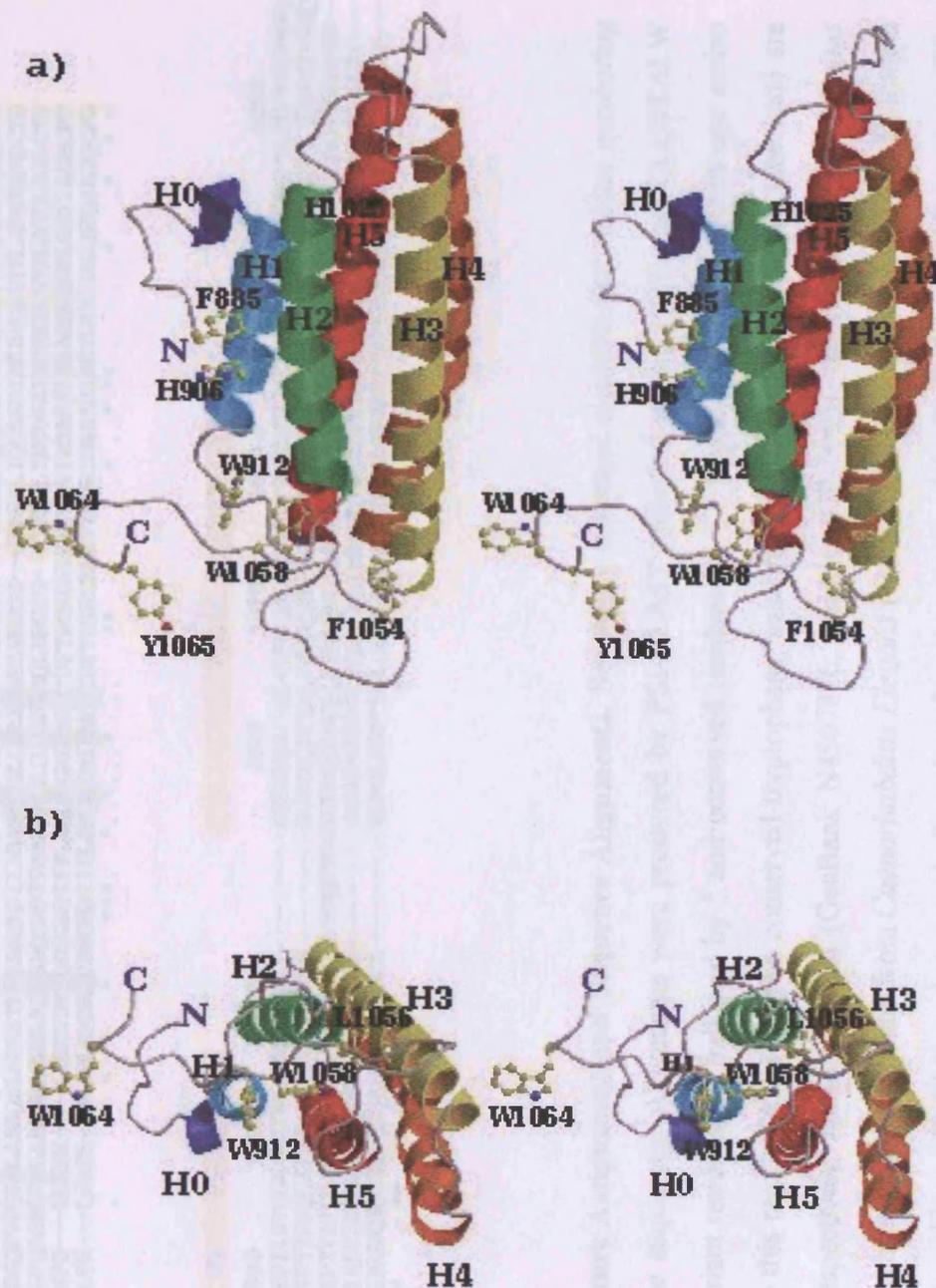


Figure 4.1. Stereo Ribbon Representation of Vt. *a)* Side view, *b)* view along the base of the bundle. Helices are shown in spectral colours from blue to red and labelled H0-H5. The N-terminal arm shown is in the consensus conformation found in the two molecules of the monoclinic crystal form. The C-terminus is that of molecule 1 in the orthorhombic form. The figure was generated by Molscript, Render and Raster-3D (Bacon and Anderson, 1988; Kraulis, 1991; Merrit and Murphy, 1994).

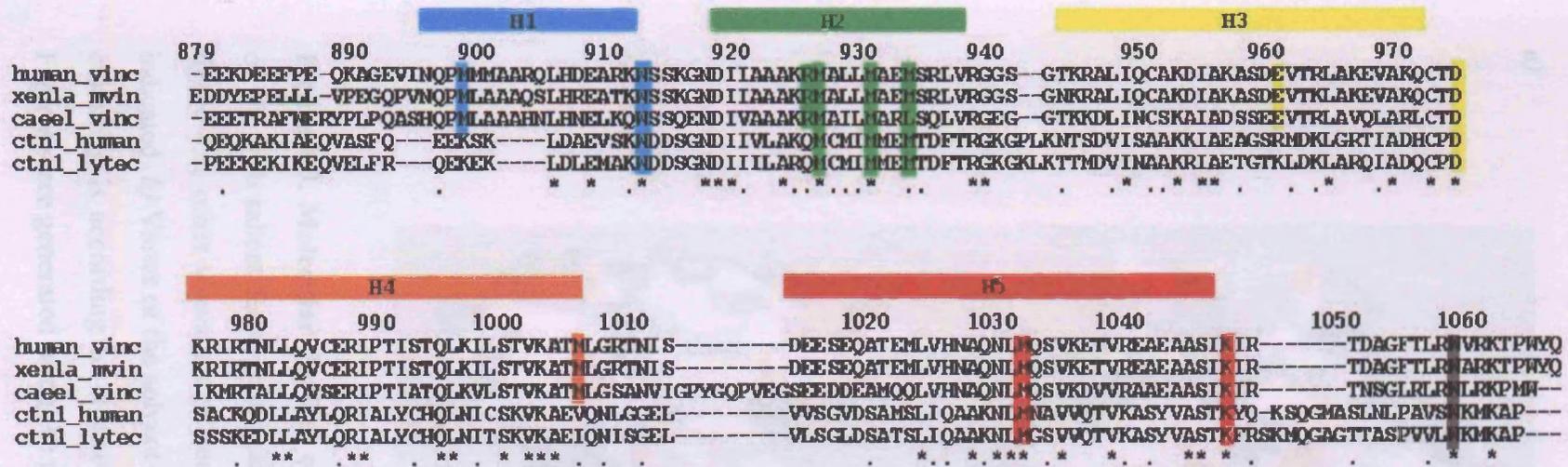


Figure 4.2. Secondary Structure Assignment and Sequence Alignments. Secondary structural elements and residue numbering for Vt are indicated above the sequences. Alignments were produced by PSI-BLAST (Altschul et al., 1997) and CLUSTALW (Thompson et al., 1994). Invariant residues are indicated by * and conserved residues by “.”. Chicken vinculin contains one amino acid difference from human in this region (A893V). The conserved tryptophans, salt-bridges and methionine zipper (see text) are highlighted. Sequences are, respectively, human vinculin (GenBank N4507876; residues 879-1066), metavinculin from *Xenopus Laevis* (GenBank Z195415; residues 47-235), vinculin from *Caenorhabditis Elegans* (GenBank J04804; residues 812-1007), human α -1 catenin (GenBank D14705; residues 679-864), and α -catenin from *Lytechinus variegatus* (GenBank AAA82613; residues 679-865). The high sequence similarities indicate that they will all adopt a similar three-dimensional structure.

a)

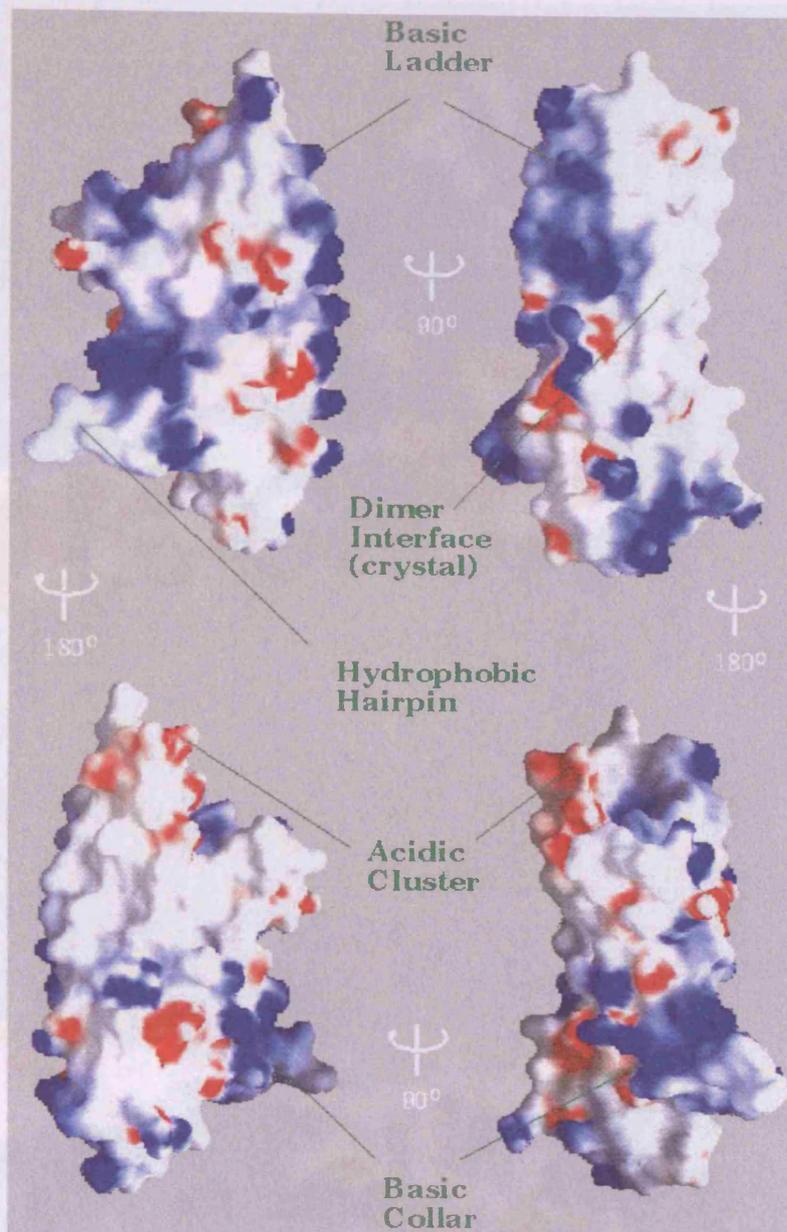
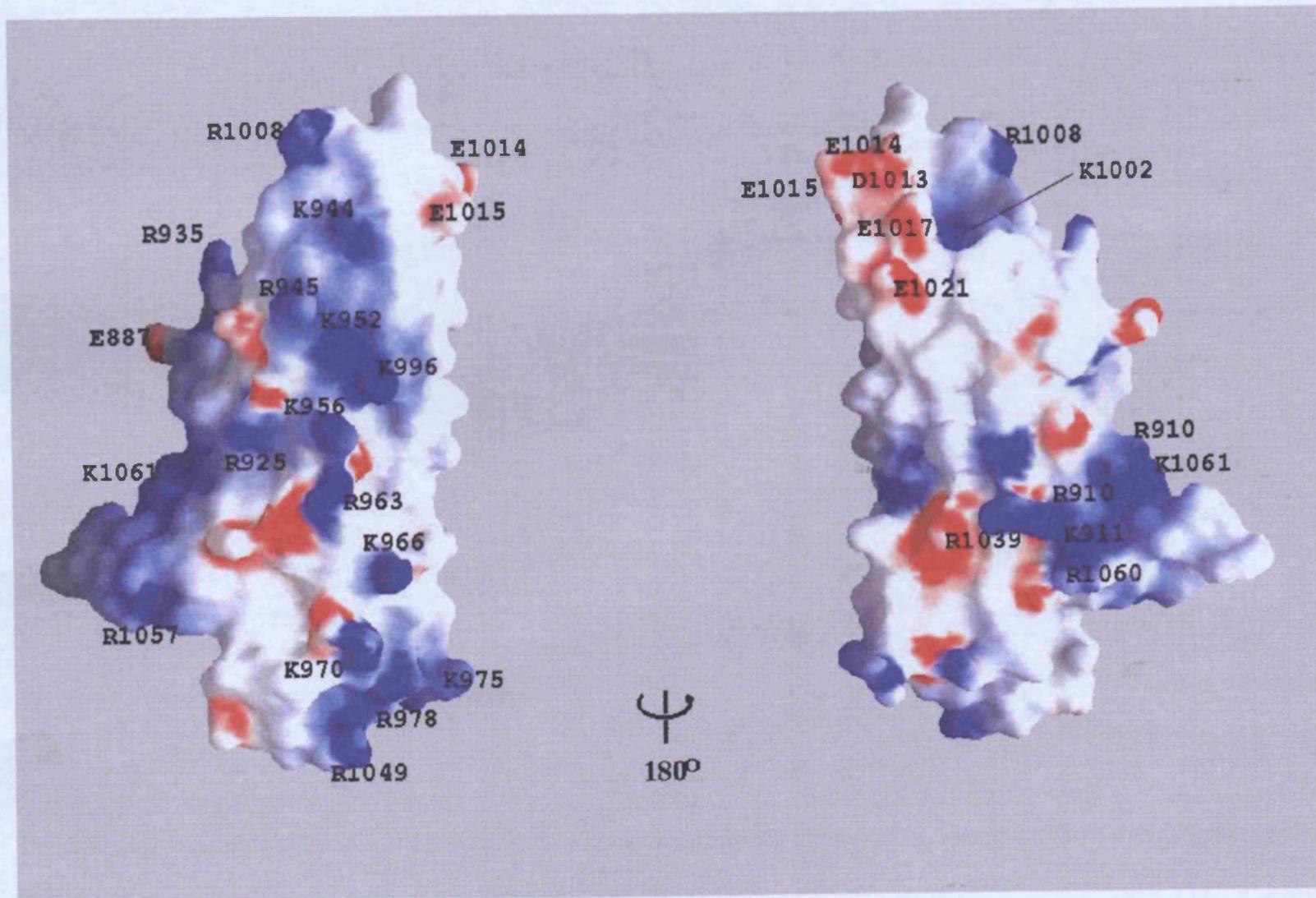


Figure 4.3. Molecular Surface of Vt. *a)* Four views of the solvent-accessible surface of Vt, with salient features indicated (see text). View at upper right is the same as in figure 4.2; other views are related by rotations of 90 or 180 about a vertical axis, as indicated. *b)* Views of the solvent-accessible surface of Vt with residue numbering. The colouring is according to electrostatic potential; blue for positive, red for negative. Figures were generated using the programme GRASP (Nicholls, 1992).

b)



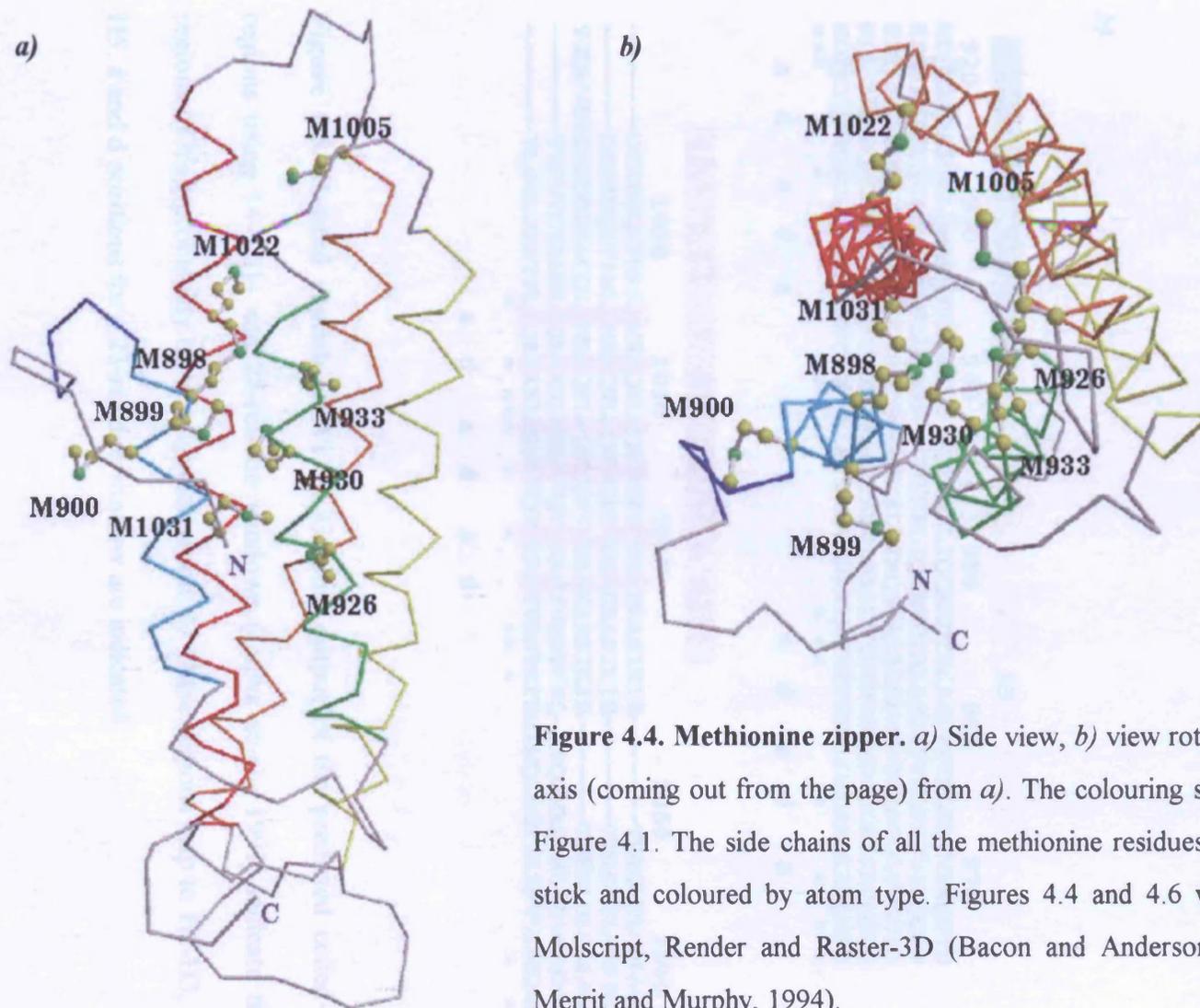
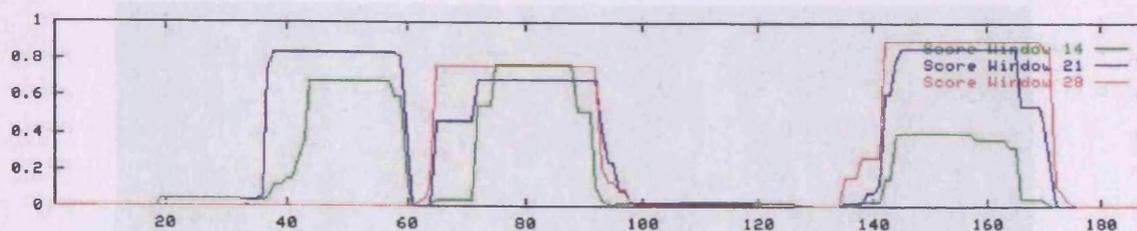


Figure 4.4. Methionine zipper. *a)* Side view, *b)* view rotated by 90° along the x-axis (coming out from the page) from *a)*. The colouring scheme is the same as in Figure 4.1. The side chains of all the methionine residues are shown as ball-and-stick and coloured by atom type. Figures 4.4 and 4.6 were both generated by Molscript, Render and Raster-3D (Bacon and Anderson, 1988; Kraulis, 1991; Merrit and Murphy, 1994).

a)



b)

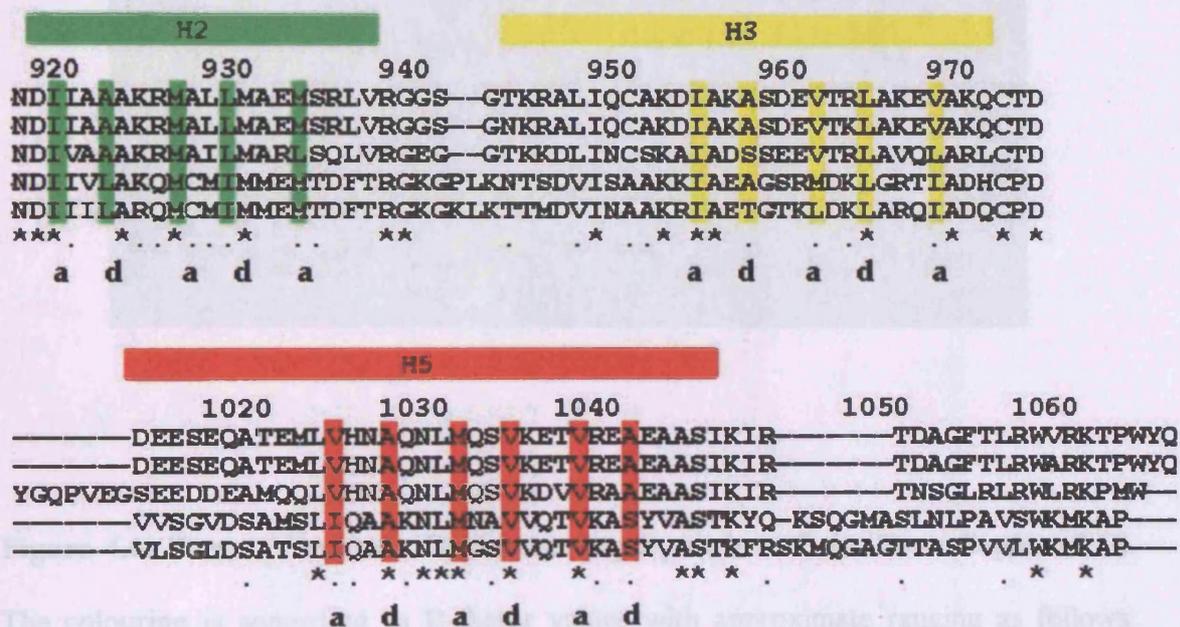


Figure 4.5. Heptad repeats in Vt. a) Graphic output of the predicted coiled-coil regions using 14-, 21- and 28-residue windows (Lupas et al., 1991) indicate three regions of high probability of Vt oligomerisation; b) These regions map to H2-H3, and H5. a and d positions for a 21-residue window are indicated.

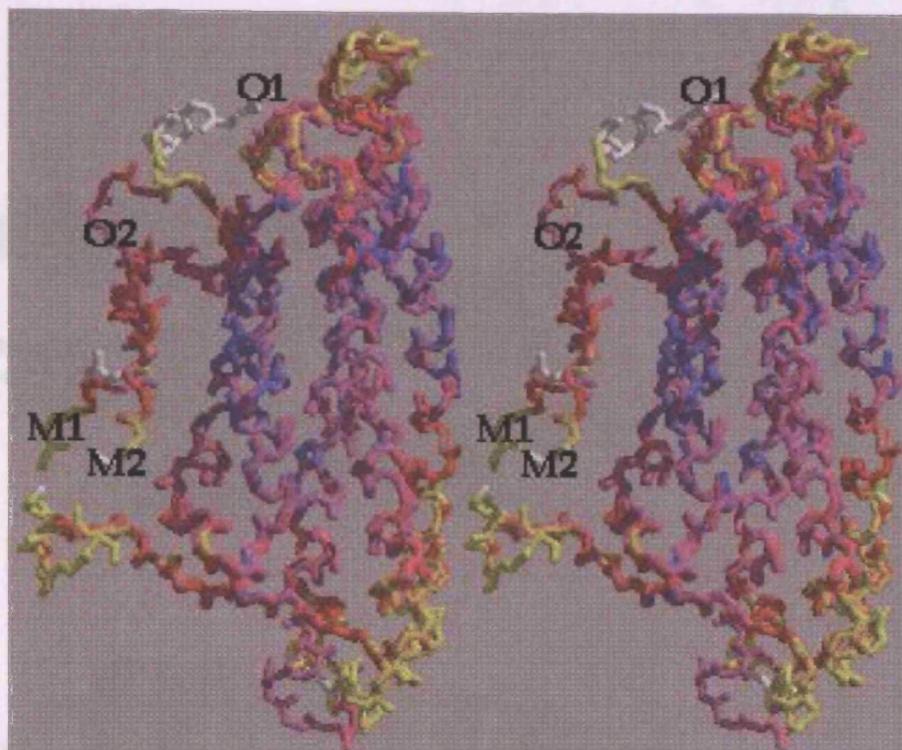


Figure 4.6. Stereo View of a Backbone Superposition of the Four Copies of Vt.

The colouring is according to B-factor values with approximate ranging as follows: blue ($9-12 \text{ \AA}^2$), magenta ($12-15 \text{ \AA}^2$), pink ($15-23 \text{ \AA}^2$), orange ($23-30 \text{ \AA}^2$), yellow ($30-47 \text{ \AA}^2$), white ($47-60 \text{ \AA}^2$). *M1*, *M2* monoclinic form; *O1*, *O2* orthorhombic form. The figure was generated using the programmes O v.6.1.2 (Jones et al., 1991), XTAL_VIEW (McRee, 1992), Render (Kraulis, 1991), and Raster3-D (Merrit and Murphy, 1994).

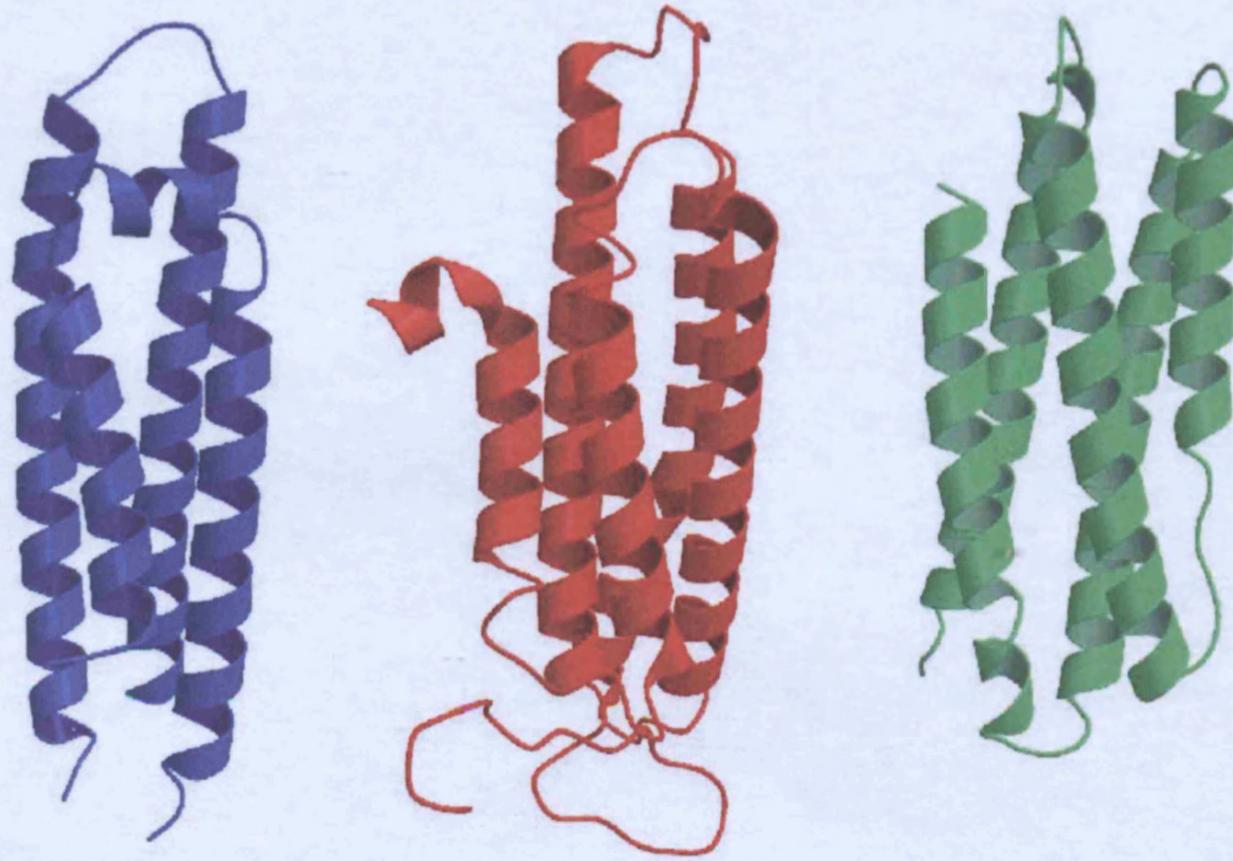


Figure 4.7. Structural Comparison of Vt with Apolipoproteins. Ribbon diagrams of apolipoprotein E (ApoE) in blue; Vt in red, and apolipophorin in green, showing the similar secondary and tertiary structures.

Chapter 5:

Structure-Based Solution Studies of Vt

Introduction

VtAC

Interactions with actin

Co-sedimentation

Stopped-flow fluorescence

Limited proteolysis

Interactions with phospholipids

Co-sedimentation

Limited proteolysis

NMR studies

¹H spectra

¹H-¹³C HSQC spectra

Materials and methods

Design, expression, and purification of VtAC

Co-sedimentation assays

Stopped-flow fluorescence

Limited proteolysis

NMR spectroscopy

Summary

Introduction

The crystal structure of Vt was obtained with the aim of providing an insight into vinculin's capacity for regulated binding to multiple ligands. Due to its static nature, the crystal structure was not capable of supplying this information on its own. However, it

provided a reliable platform from which to design and interpret a series of structure-based solution experiments. These experiments, designed to probe the conformational behaviour of Vt in the presence of ligands, could then complement the data derived from the crystal structure leading to a better overall understanding of the protein's function.

After inspection of the three-dimensional structure of Vt, Bob Liddington suggested a role for the C-terminal hairpin in membrane interactions. He proposed that the exposed tryptophan and tyrosine surrounded by a collar of basic residues, could constitute a motif for membrane association in a similar manner to that proposed for annexin V (Campos et al., 1998). On this basis, a deletion mutant (Vt Δ C) was constructed designed to maintain the 5-helix bundle architecture while lacking the β -strand and hydrophobic hairpin of the C-terminal arm (residues 1052-1066).

This chapter describes the construction of Vt Δ C and the methods used to assess its behaviour in the presence of actin, acidic and non-acidic phospholipids, and proteases; results were subsequently compared to those obtained with Vt. Binding of both Vt and Vt Δ C to phospholipids was monitored by co-sedimentation, binding to actin by co-sedimentation and stopped-flow fluorescence, and conformational changes occurring in the presence of F-actin or acidic phospholipids by limited proteolysis.

Vt Δ C

The absence of errors in the Vt Δ C construct was confirmed by DNA sequencing. Protein yield and purity was similar to that obtained for Vt (see Chapter 2). The expressed truncated protein folded normally as judged by its CD spectrum, which is

indistinguishable from wild-type Vt (data not shown) and by its limited proteolysis pattern in the presence and absence of F-actin (see p.95-96).

Interactions with actin

Co-sedimentation assays

The ability of Vt and Vt Δ C to bind actin was assessed by co-sedimentation. As shown in figure 5.1, both Vt and Vt Δ C co-sediment with actin indicating that the C-terminal arm of Vt is not required for actin binding.

Stopped-flow fluorescence

Given the clustering of the tryptophan residues in the crystal structure of Vt, actin binding was investigated by tryptophan fluorescence. The tryptophan fluorescence emission spectrum of Vt plus F-actin was quenched by 5% compared with the sum of the component spectra. This change was time-resolved using stopped-flow fluorimetry (figure 5.2). Mixing 5 μ M Vt with 5 μ M F-actin (reaction chamber concentrations) yielded an apparent first order rate constant of 12s⁻¹ when fitted to an exponential. The reaction could not be followed under strictly pseudo first order conditions because an excess of either protein gave too large a background signal. Nevertheless, doubling or halving the total concentrations of the reactants caused little change in the observed rate (i.e. binding was near complete over the concentration range used). Furthermore, the profile was better fitted to an exponential than a second order rate equation. These observations suggest that the tryptophan signal is monitoring a conformational change in Vt and/or actin rather than the binding process per se. Measurements using a

truncated vinculin indicate that tryptophans in Vt are involved (see p.95-96). Two step mechanisms in which a conformational change in Vt is induced subsequent to actin binding or actin binding stabilises a minor conformational state of Vt give different kinetic profiles, but they are difficult to distinguish over the concentration ranges available to investigation. For either mechanism, the initial second order association rate constant between Vt and F-actin is $\gg 5 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ to account for the observed rate constant at the lower concentrations used. The association process is at least an order of magnitude faster than the apparent association reported by Goldman et al. (Goldman et al., 1998) who used a Vt fusion protein and light scattering.

Similar experiments were performed on Vt Δ C, which lacks two of the three Trp residues (W1058 and W1064). Vt Δ C binds actin normally as judged by a co-sedimentation assay (see above), and by displaying a similar increase in protease sensitivity (see below), showing that the C-terminal arm is not required for actin binding. However, Vt Δ C gave no fluorescence signal change on mixing with F-actin (figure 5.2), suggesting that the fluorescence change originates from the tryptophans in Vt rather than in F-actin (which contains 4 tryptophans per monomer). Taken together, these results suggest that the fluorescence change observed for Vt is monitoring a change in the packing or flexibility of W1058 and W912 at the base of the bundle, a change which is also reflected in the increased protease sensitivity of Vt and Vt Δ C on binding F-actin (see p.95-96).

Limited proteolysis

To further explore the possibility (raised by the kinetic studies, see above) that binding to F-actin induces conformational changes in Vt, limited proteolysis of both Vt and Vt Δ C was performed in the presence or absence of F-actin.

Figure 5.3c shows that in the presence of F-actin, Vt becomes susceptible to proteolysis by elastase. Thus, under conditions in which insignificant cleavage of isolated Vt or F-actin is observed with this protease (Figures 5.3a and b), there is rapid cleavage of the Vt:F-actin complex with accumulation of a major proteolysis product at 60-120 minutes. N-terminal sequencing of the major bands shows that Vt is cleaved after residue M900, which is located within H1 and is buried by the N-terminal arm.

The rate of proteolysis by chymotrypsin is also enhanced in the presence of F-actin (Figure 5.4c), and after 120 minutes there is essentially stoichiometric conversion to an ~18 kDa fragment with an N-terminus of 913 (i.e. cleavage after W912) and its C-terminus at F1054. W912 is buried and packs against W1058 in the crystal structure.

Vt Δ C shows a similar response to Vt in the presence of these two proteases (experiments performed by Bipin Patel). In the presence of elastase, there is a rapid conversion to a ~18 kDa band (Figure 5.5b). This band has not been N-terminal sequenced, but its apparent molecular weight is consistent with cleavage at a similar position to that observed with Vt (see Figure 5.3c). Some cleavage of Vt Δ C is observed after 120 min (Figure 5.5.a) but there is no formation of a major proteolytic product, making for a distinct proteolytic pattern of Vt Δ C in the presence or absence of F-actin.

The limited number of aromatic residues in Vt Δ C, allows a more confident identification of chymotryptic digests in the absence of sequencing information. Thus, in the absence of F-actin Vt Δ C is gradually converted to a ~16 kDa product (Figure 5.6.a). This apparent molecular weight is consistent with a cleavage after W912 and is therefore the equivalent of band III in the cleavage of Vt by chymotrypsin in the absence of F-actin (Figure 5.4.b). The absence of an equivalent to band II (Figure 5.4.b) is correlated to the absence of the C-terminal arm in Vt Δ C. In the presence of F-actin, the rate of Vt Δ C cleavage by chymotrypsin is greatly enhanced with total conversion of Vt Δ C to a ~16 kDa product being observed within 30 min (Figure 5.6.b).

Interactions with phospholipids

Co-sedimentation assays

The ability of Vt and Vt Δ C to co-sediment with multilamellar vesicles composed of either phosphatidylserine (PS) or phosphatidylcholine (PC) was assessed.

Figure 5.7a shows that at physiological pH and ionic strength, Vt co-sediments with PS but not with PC, as shown previously (Johnson and Craig, 1995) whereas Vt Δ C does not co-sediment with either lipid under these conditions.

On lowering the pH, however, co-sedimentation of Vt Δ C with PS vesicles is observed, with a mid-point at pH 6.5 (figure 5.7b). Very little co-sedimentation of Vt Δ C occurs with PC vesicles, even at pH 4.5, showing that the mutant retains its specificity for acidic phospholipids. The process is reversible; increasing the pH, results in loss of binding of Vt Δ C with PS vesicles (data not shown).

Limited proteolysis

To investigate whether binding to acidic phospholipids affects the conformation or flexibility of Vt, limited proteolysis was performed using a range of different proteases. In all cases tested, a greater rate of proteolysis was observed in the presence of PS than in the presence of PC.

For elastase, under conditions in which little or no cleavage occurred in the presence of PC, rapid proteolysis is observed when bound to PS (figure 5.7a) with some accumulation of the same 18 kDa fragment that is seen in the presence of actin (see above). For chymotrypsin, the proteolysis pattern in the presence or absence of PC is identical, but in the presence of PS, the rate and extent of proteolysis greatly increases (Figures 5.7b).

NMR studies

¹H spectra

The spectra in figure 5.9.a, were obtained in order to test whether there was a pH dependence in the conformational states of Vt. The results suggest that the observed changes are due to titratable groups (carboxylates and histidines for the specified pH range) and are not indicative of conformational changes in the protein.

¹H-¹³C HSQC spectra

¹H-¹³C HSQC spectra were obtained with ¹³C-Met-substituted Vt in the absence (figure 5.9a) and presence (figure 5.9b) of a signal-broadening reagent (TEMPO) in an

attempt to obtain a probe for conformational changes of Vt upon acidic phospholipid binding.

Comparison of these spectra with the positioning of the methionine residues in the crystal structure, indicate a good correlation between the number of solvent exposed and buried methionines. Peaks 1 and 3 remain unchanged in the presence or absence of TEMPO; peaks 4, 5 and 6 remain relatively unchanged; peaks 7-10 exhibit a marked broadening in the presence of TEMPO, peak 2 is the most-solvent exposed of all methionines, as in the presence of TEMPO its signal completely disappears. The crystal structure illustrates four highly solvent-exposed methionines (M900, M1005, M1022 and M1031) which could correspond to peaks 7-10; the N-terminal methionine, which would also be expected to be fully solvent exposed, is not visible in the structure and probably corresponds to peak 2. M926, M930 and M933 are located in the most hydrophobic environment and could therefore represent peaks 1, 3, and 4. M898 and M899 are partially exposed to solvent, and could be assigned to peaks 5 and 6.

If the helical bundle were to open up in the presence of acidic phospholipids, more methionines (in particular M926, M930 and M933) would be expected to become solvent-exposed. However, subsequent attempts at obtaining a stable lipid/protein complex did not meet with success, leading for the most part to the production of protein/lipid aggregates and uninterpretable spectra.

Materials and Methods

Design, expression, and purification of VtΔC

The construct was designed, expressed, and purified following essentially the same sub-cloning and purification procedures as for Vt (see Chapter 2). The same 5' primer was used (VDU) as for the subcloning of Vt, while the 3' primer is given below:

< EcoRI > < BamHI >

5' GCC GAA TTG GGA TCC TTA ATC TGT TCT TAT CTT 3'

Successful clones were selected by monitoring for the absence of an *NcoI* site at the 3' end of the VtΔC construct.

Co-sedimentation assays

The actin co-sedimentation assay was carried out essentially as described (Hemmings et al., 1996). Briefly, equimolar amounts of rabbit muscle G-actin and Vt or VtΔC were mixed in a total volume of 100 μL of buffer containing 2 mM Tris/HCl pH 8, 0.2 mM ATP, 0.5 mM dithiothreitol, 0.2 mM CaCl₂ and 0.01% Triton X-100. Actin polymerization was initiated by the addition of KCl to 500 mM, 10 mM MgSO₄ to 10 mM, EGTA to 10 mM, and the samples were incubated at room temperature for 60 minutes. Samples were layered over a 20% sucrose gradient and centrifuged both with and without actin in a Beckman Optima Max-E ultracentrifuge at 400,000 g for 45 minutes. The supernatants were removed, the pellet washed carefully with phosphate buffered saline (PBS), and resuspended in 100 μL

PBS. Proteins present in equal volumes of both the supernatant and pellet were resolved by SDS-PAGE and the gels stained with Coomassie blue.

For the phospholipid cosedimentation assay, large, multilamellar vesicles (MLVs) were prepared essentially as described (Johnson et al., 1998). After chloroform evaporation under vacuum, the dried phospholipid films were swollen at 5 mg/mL in 150 mM NaCl, 20 mM Tris pH 7.5 for 3 h at 42°C. The vesicles were then centrifuged (20000g for 20 min at 4°C), and the pellet was resuspended in the same buffer at 0.5 mg/mL. The proteins were diluted in the same buffer to 0.1 mg/mL and incubated (30 min, 37°C) with phospholipid vesicles (0.5 mg/mL). After incubation, reaction mixtures were centrifuged (40000g for 10 min at 20°C), pellet and supernatant fractions were subjected to SDS-PAGE and proteins detected by Coomassie Brilliant Blue staining of the gels. For the pH assay, pellets were resuspended in equal volumes of 150 mM NaCl, 20 mM sodium phosphate/citrate at the required pH and protein concentrations of supernatants determined by absorbance at 595 nm using the Coomassie dye-binding protein assay (Bradford, 1976).

Limited Proteolysis

0.03 mM Vt in F-actin buffer (40 mM NaCl, 20 mM TES pH 7.0, 1 mM MgCl₂) was incubated in the presence or absence of an equimolar concentration of F-actin at RT for 10 min. Proteoliposomes of Vt and PS were obtained as described above, separated from unbound Vt by centrifugation, and resuspended in 150 mM NaCl, 20 mM Tris pH 7.5. Alternatively the same final concentration of Vt was added to purified PC vesicles. 1% (w/w) elastase or 0.2% chymotrypsin was added to each sample, incubated at 30°C, and the

reactions stopped by adding 1 mM PMSF or 10% trichloroacetic acid. Samples were analysed by SDS-PAGE and peptide bands stained with Coomassie blue or blotted onto poly(vinylidene difluoride) membranes (Millipore). N-terminal microsequencing of the blotted peptides was performed on a 476 Protein Sequencer (Applied Biosystems).

Stopped-flow Fluorescence

Stopped-flow fluorescence measurements were performed using a SF-17MV stopped-flow spectrometer (Applied PhotoPhysics, Leatherhead, UK). Tryptophan fluorescence was excited at 295 nm with the emission selected using a 340 nm cut-off filter placed in front of the photomultiplier. The temperature of the sample handling system was maintained using a water circulating temperature control system. Sample stock solutions were diluted to the desired concentrations immediately prior to loading into the syringes. A minimum of 4 traces were collected for signal averaging and subsequent non-linear regression analysis. Rabbit skeletal muscle actin, provided by R. Woolley, was purified from an acetone powder essentially as described by Pardee and Spudich (1982). The actin concentration was determined using the extinction coefficient $A_{290} (1\%) = 6.2$. In all cases, the sample buffer comprised 40 mM NaCl, 20 mM TES pH 7.0, 1 mM MgCl₂.

NMR spectroscopy

¹³C-Met-substituted Vt was grown essentially as for SeM-Vt (see Chapter 2). All spectra were obtained using a DMX 500 spectrometer at 500 MHz (¹H) and 125 MHz

(¹³C). For the proton spectra, 90% H₂O-10% D₂O buffers were used. Data were processed by Lu Yun Lian.

Summary

In an attempt to supplement the crystal structure of Vt with information about its behaviour in solution, a variety of structure-based experiments were carried out on Vt and a C-terminal truncation mutant (VtΔC) in the presence of ligands.

Both Vt and VtΔC co-sedimented with actin, while comparison of stopped-flow fluorescence data between Vt and VtΔC indicated a change in the packing and sensitivity of the Vt tryptophans at the base of the bundle. This result is further supported by limited proteolysis experiments of Vt and VtΔC in the presence of actin which demonstrated a conformational change of Vt in the N- and C-terminal regions.

VtΔC co-sedimentation with acidic MLVs was observed only upon lowering of the pH. Limited proteolysis in the presence of acidic phospholipids indicated an increased flexibility of Vt as monitored by the heightened susceptibility of Vt to proteases.

¹H NMR spectra of Vt show no conformational change of the molecule as a function of pH. HSQC NMR spectra of Vt led to the assignment of methionine groups according to their degree of exposure to solvent. However, a successful Vt/phospholipid complex that could be probed by these means, and thus yield information on the state of the helical bundle in the presence of acidic phospholipids, remains to be obtained.

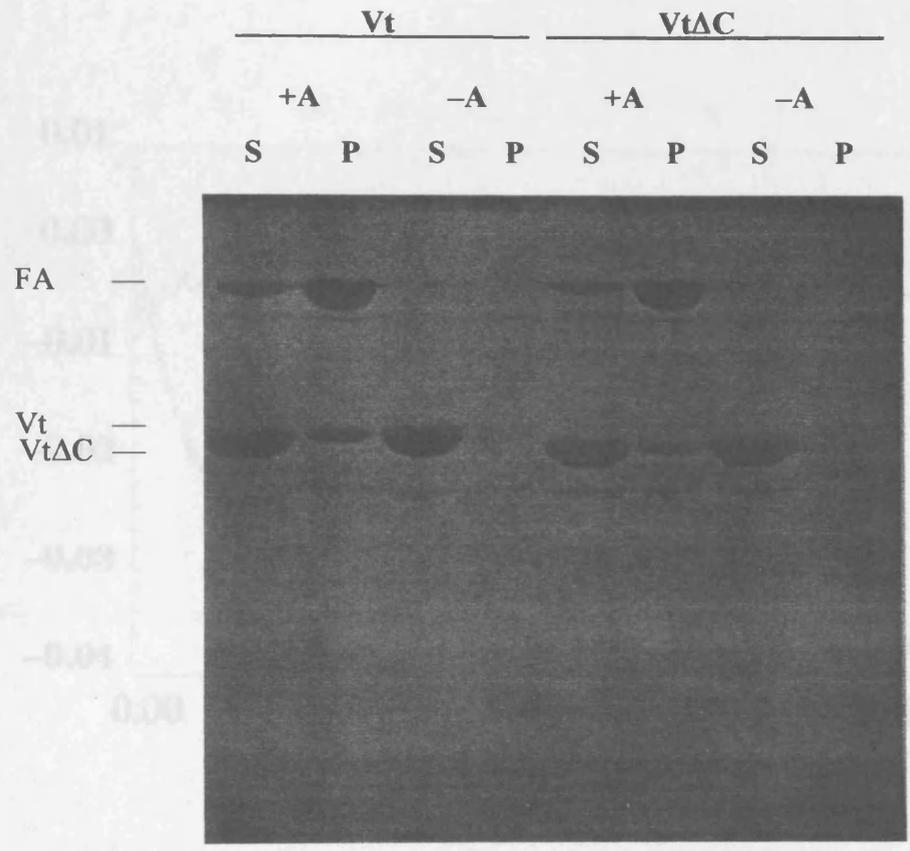


Figure 5.1. Co-sedimentation of Vt and VtΔC with actin. SDS-PAGE analysis of co-sedimentation of Vt and VtΔC in the presence (+A) or absence (-A) of F-actin. *S* = supernatant; *P* = pellet. F-actin and Vt or VtΔC were mixed in a ratio of 2:1.

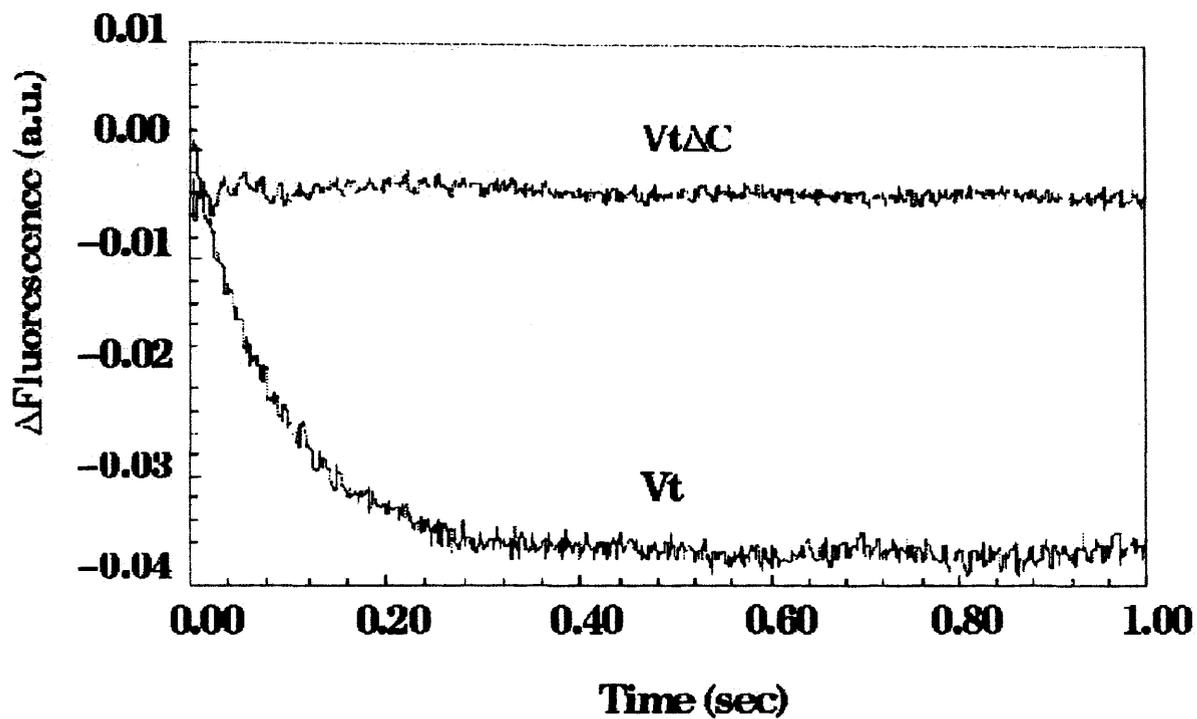


Figure 5.2. Change in tryptophan fluorescence during the interaction of Vt and $Vt\Delta C$ with F-actin as measured by stopped-flow. The change is measured as a fraction of time relative to a total signal of 1 V.

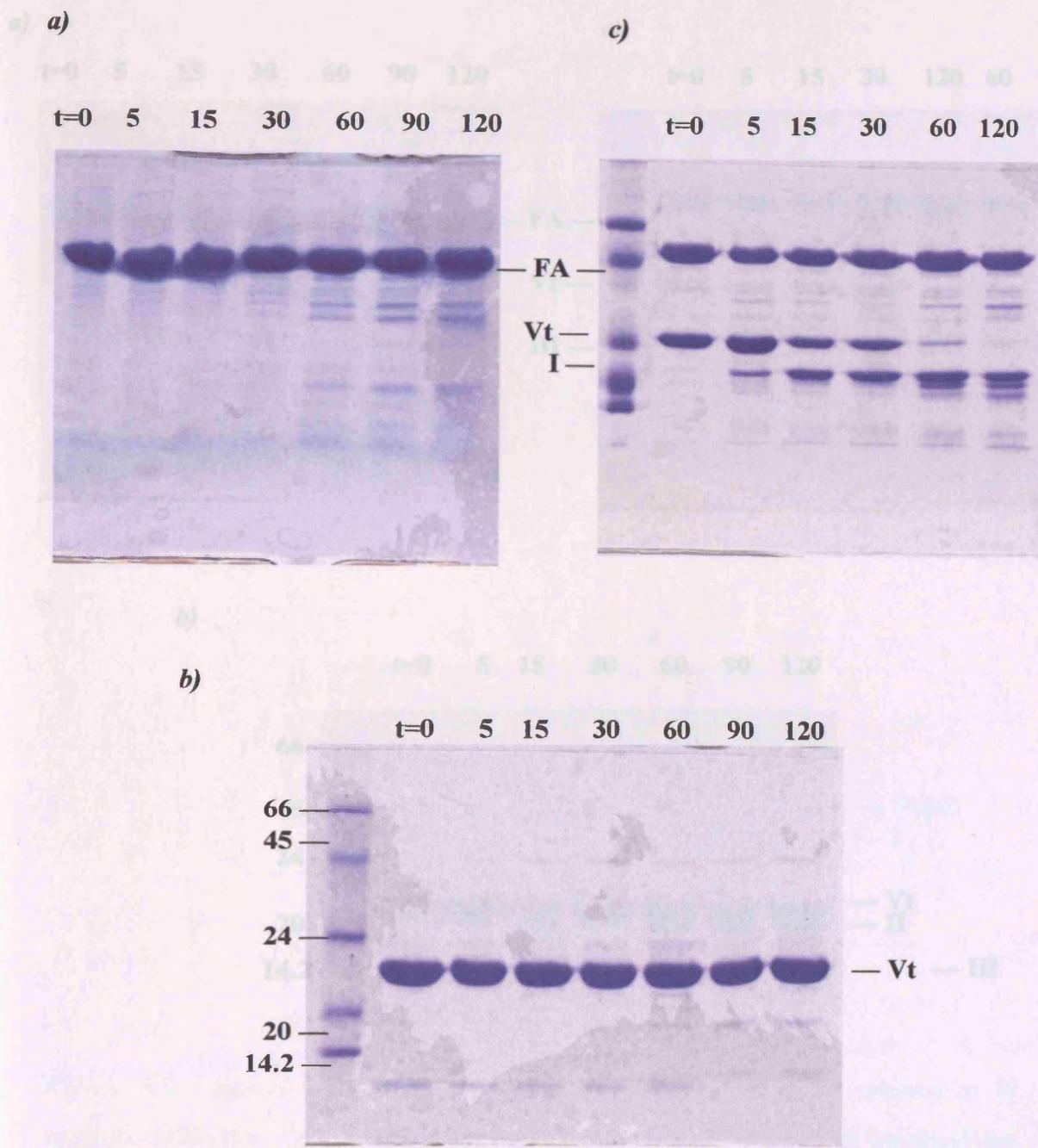


Figure 5.3. Limited proteolysis by elastase. a) F-actin (FA), b) Vt, c) Vt in the presence of F-actin. Samples were taken at time intervals indicated (minutes). Molecular weight markers indicated in c) are in kDa and identical to those in b). The proteolysis product I has its N-terminus at A901. consistent with cleavage after F1054, III has its N-terminus at 3913 and is also cleaved after F1054.

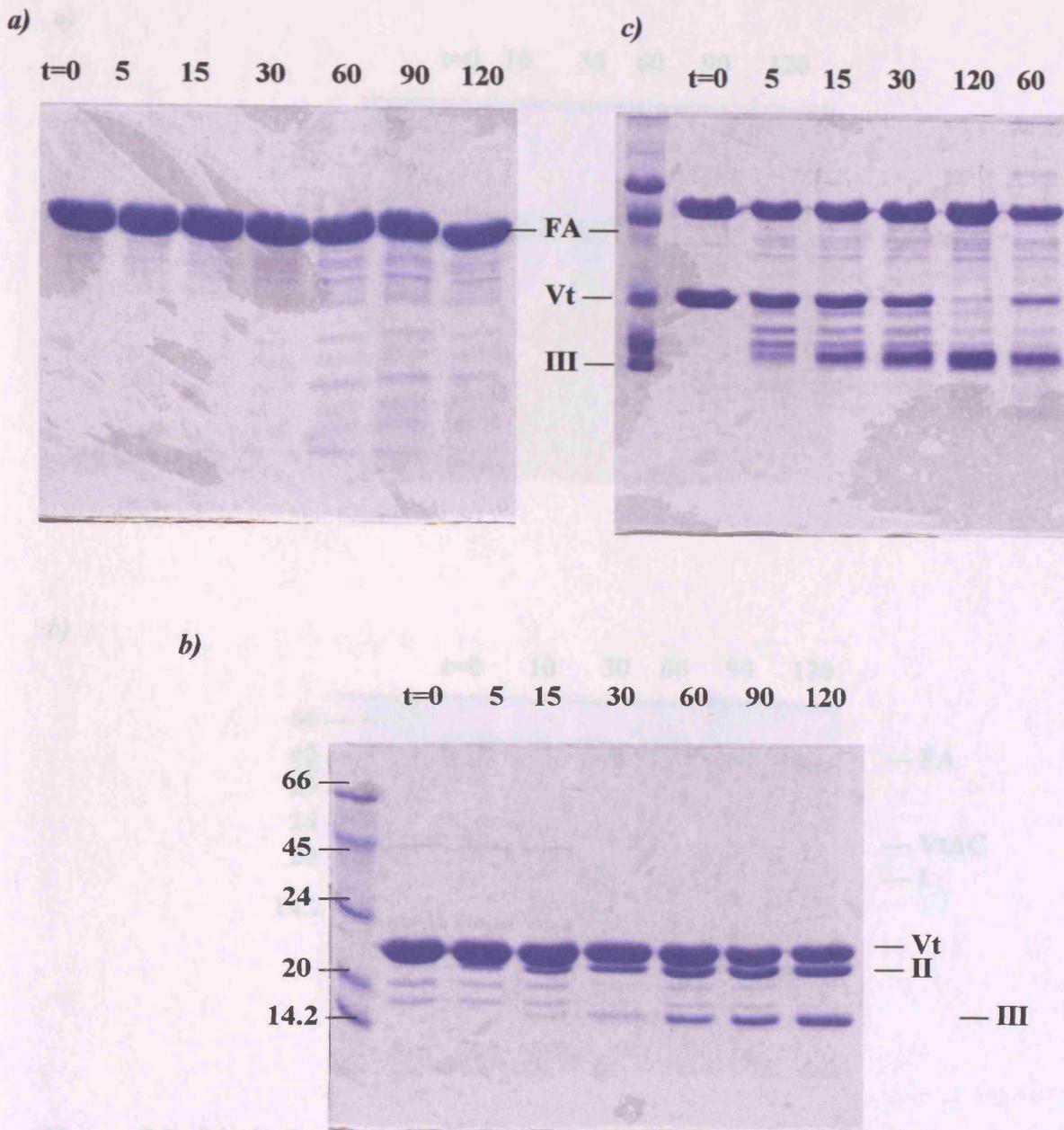
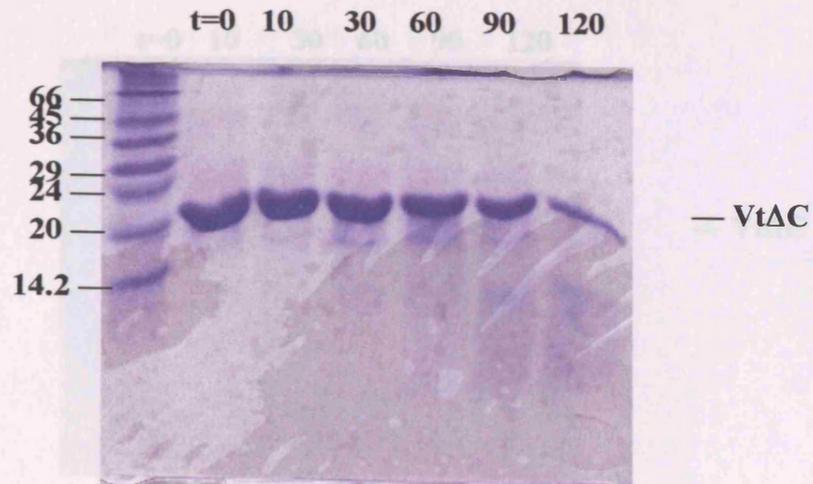


Figure 5.4. Limited proteolysis of Vt by chymotrypsin. a) F-actin (FA), b) Vt, c) Vt in the presence of F-actin (FA). Samples were taken at time intervals indicated (minutes) and

Figure 5.4. Limited proteolysis by chymotrypsin. a) F-actin (FA), b) Vt, c) Vt in the presence of F-actin. Samples were taken at time intervals indicated (minutes). Molecular weight markers indicated in c) are in kDa and identical to those in b). The proteolysis products labeled II and III are as follows: II has an authentic N-terminus and a molecular weight consistent with cleavage after F1054; III has its N-terminus at S913 and is also cleaved after F1054.

a)



b)

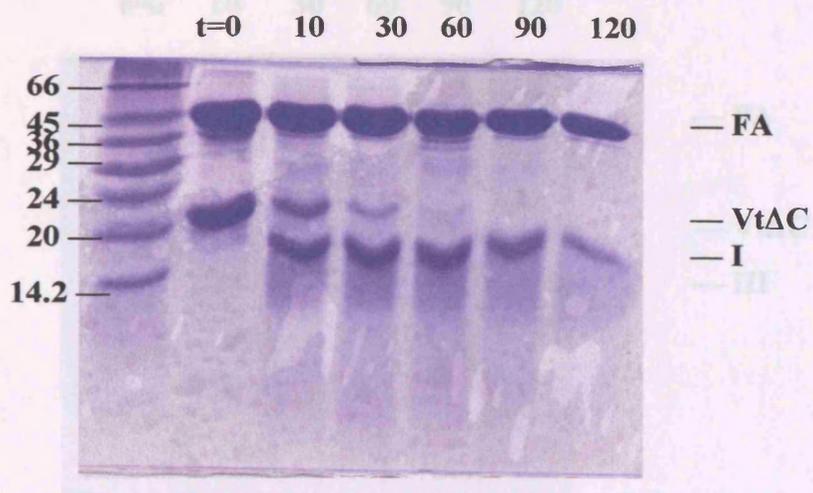
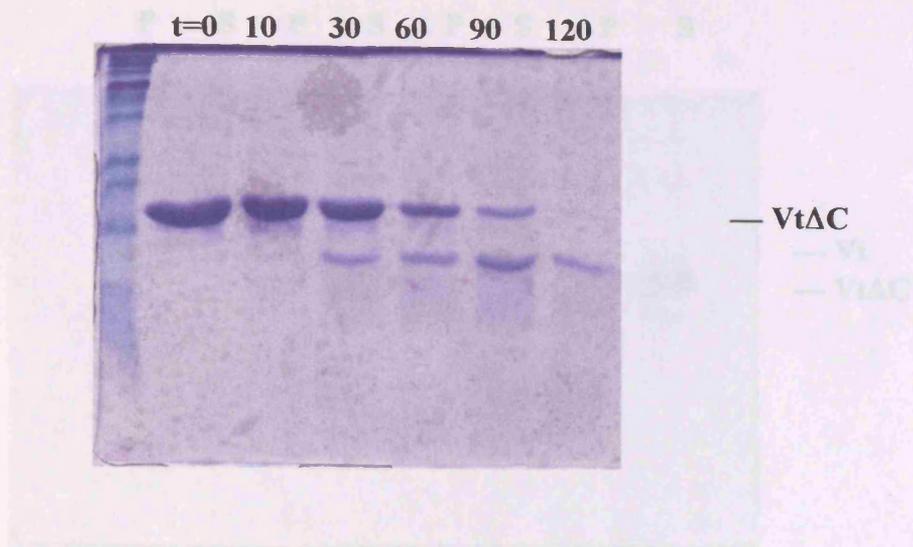


Figure 5.5. Limited proteolysis of Vt Δ C by elastase. a) In the absence or b) presence of F-actin (FA). Samples were taken at time intervals indicated (minutes) and analysed by SDS-PAGE. Molecular weight markers (in kDa) are indicated on the left. The proteolysis band observed has the same apparent molecular weight as the one defined in the legend to figure 5.3.

Note the absence of band II (defined in the legend of figure 5.4) caused by the lack of the C-terminal arm. The proteolysis band observed has an apparent molecular weight consistent with cleavage after W912.

a)



b)

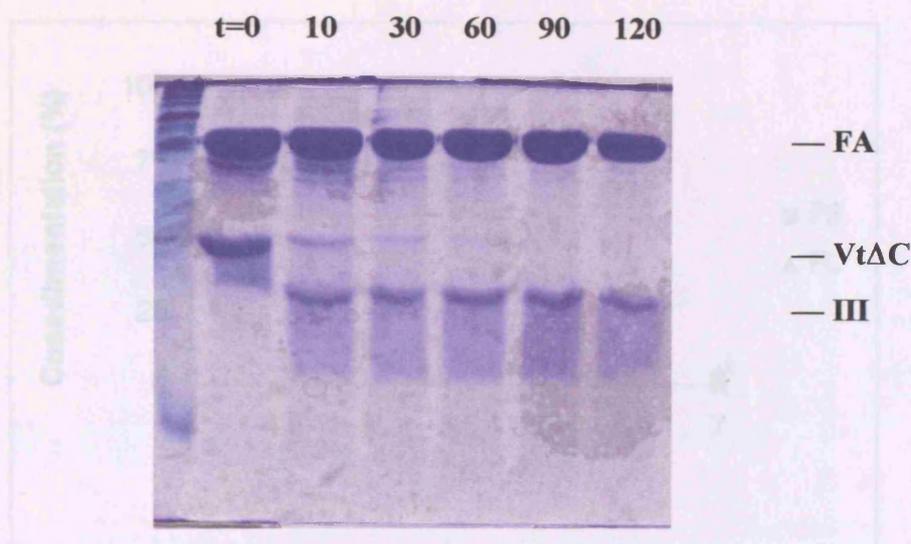


Figure 5.6. Limited proteolysis of VtΔC by chymotrypsin. *a)* In the absence or *b)* presence of F-actin (FA). Samples were taken at time intervals indicated (minutes) and analysed by SDS-PAGE. Proteolysis bands marked with roman numerals (I-III) are defined in the legends to Figure 5.4. Note the absence of band II (defined in the legend of figure 5.4) caused by the lack of the C-terminal arm. The proteolysis band observed has an apparent molecular weight consistent with cleavage after W912.

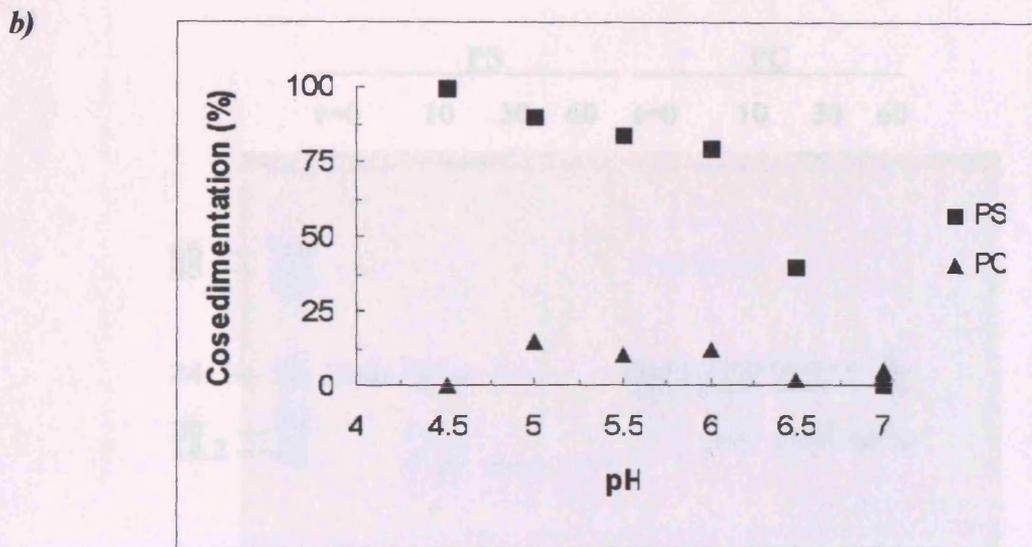
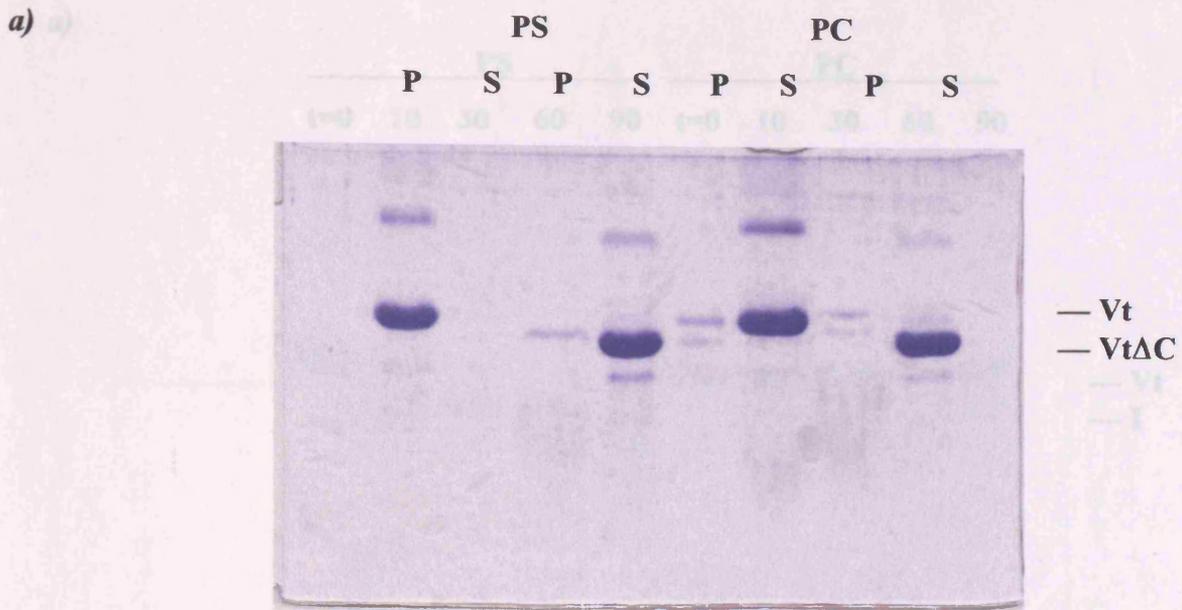
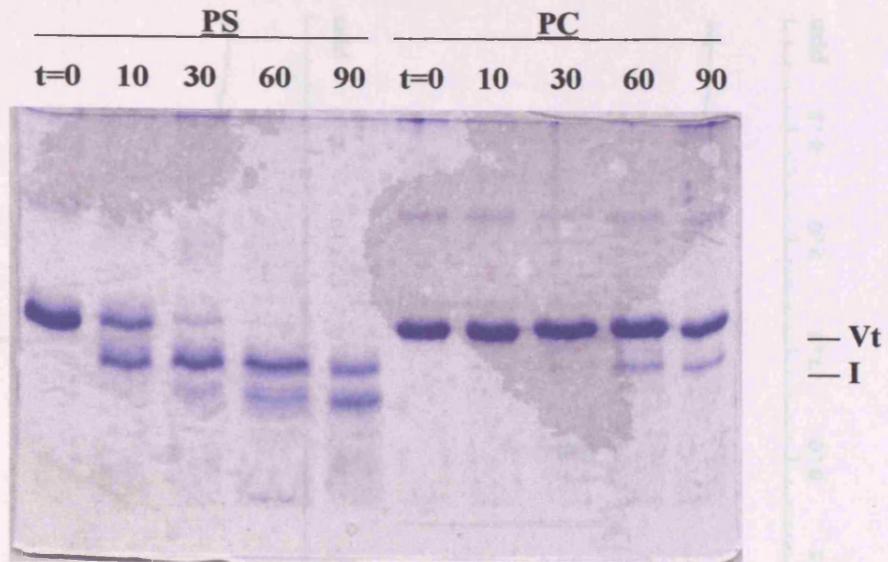


Figure 5.7. Co-sedimentation of Vt and VtΔC with phospholipids. a) SDS-PAGE analysis of co-sedimentation of Vt and VtΔC with phosphatidylserine (PS) or phosphatidylcholine (PC) MLVs at pH 7.5. P = pellet, S = supernatant. b) Co-sedimentation of VtΔC with PS or PC under the same conditions as in a), but as a function of pH. Co-sedimentation was quantified from the protein concentration of the supernatant.

a)



b)

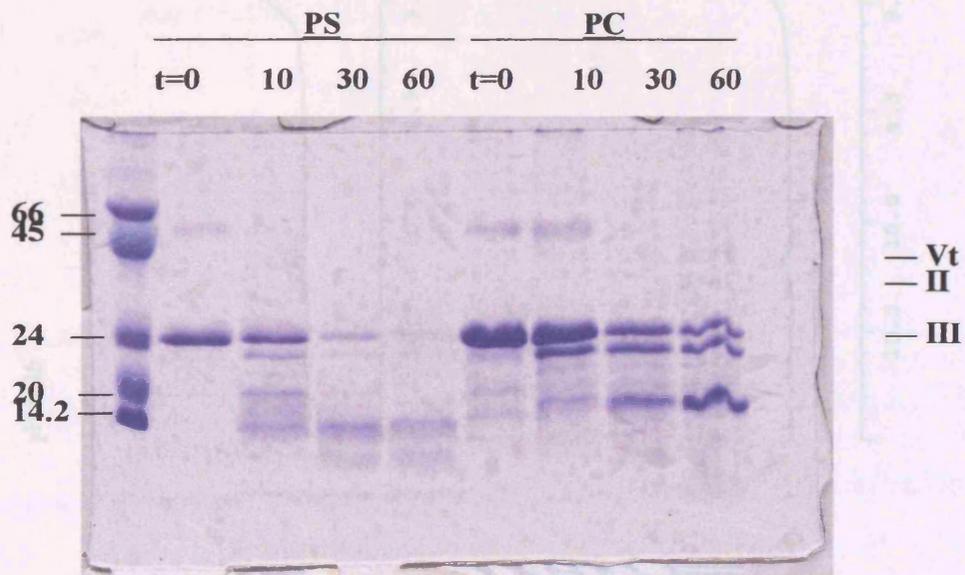


Figure 5.8. Limited proteolysis of Vt by *a)* elastase and *b)* chymotrypsin in the presence of PS or PC MLVs. Samples were taken at time intervals indicated (minutes) and analyzed by SDS-PAGE. Note that the proteolysis pattern for chymotrypsin in the presence of PC is the same as that of isolated Vt (see figure 5.5). Proteolysis bands marked with roman numerals (I-III) are defined in the legends to Figures 5.3 and 5.4.

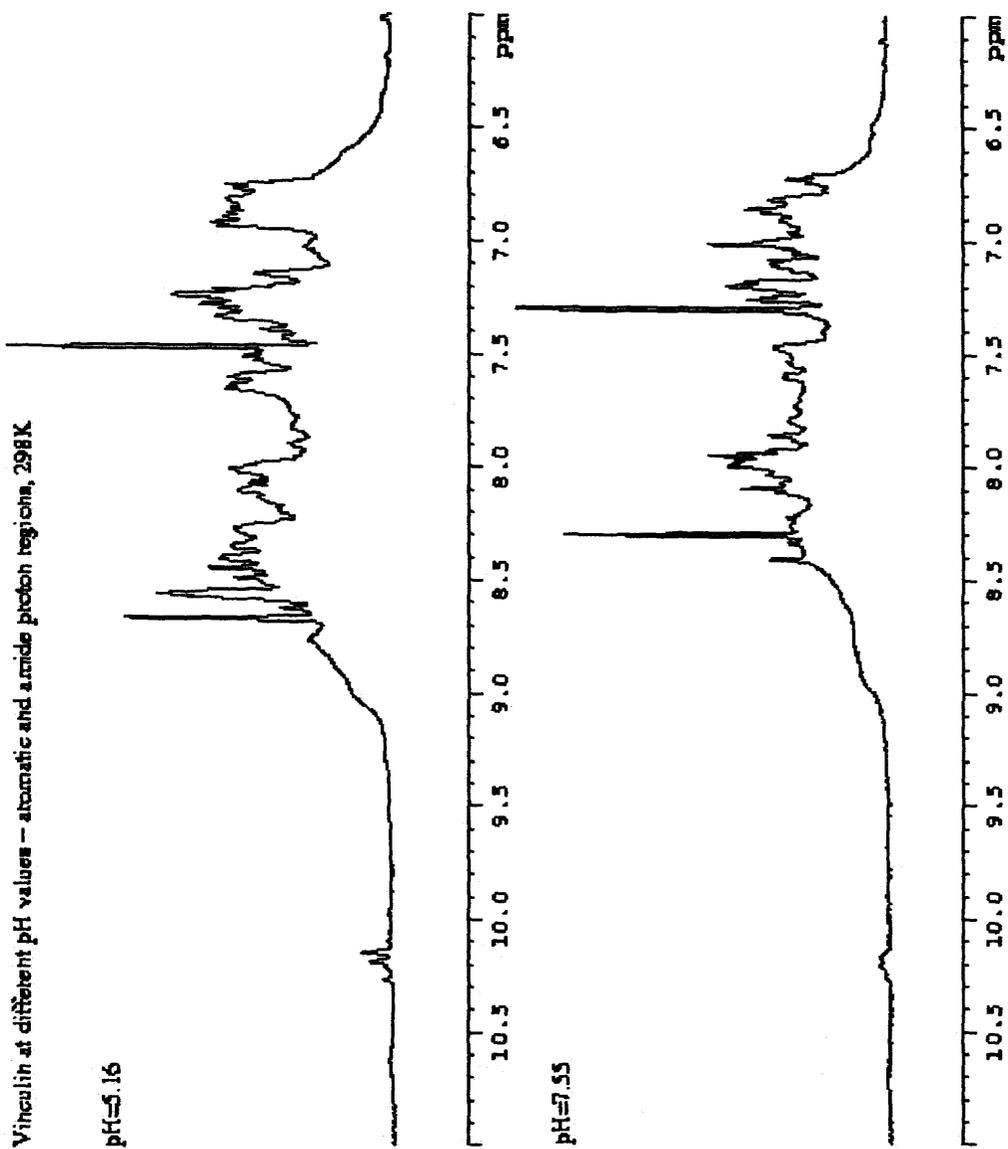


Figure 5.9. ^1H spectra of Vt at two pH values. Vt is at 0.5mM, temperature at 298K.

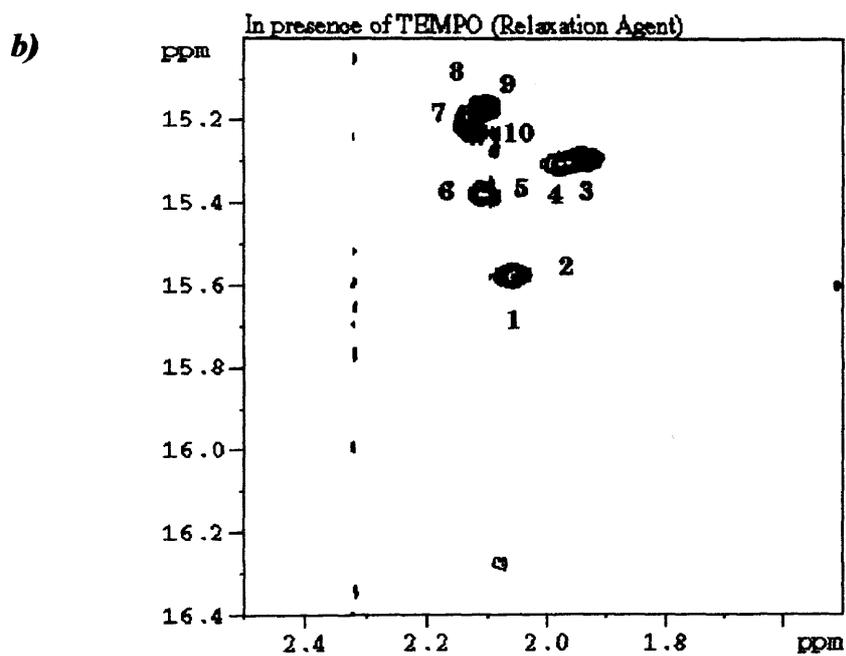
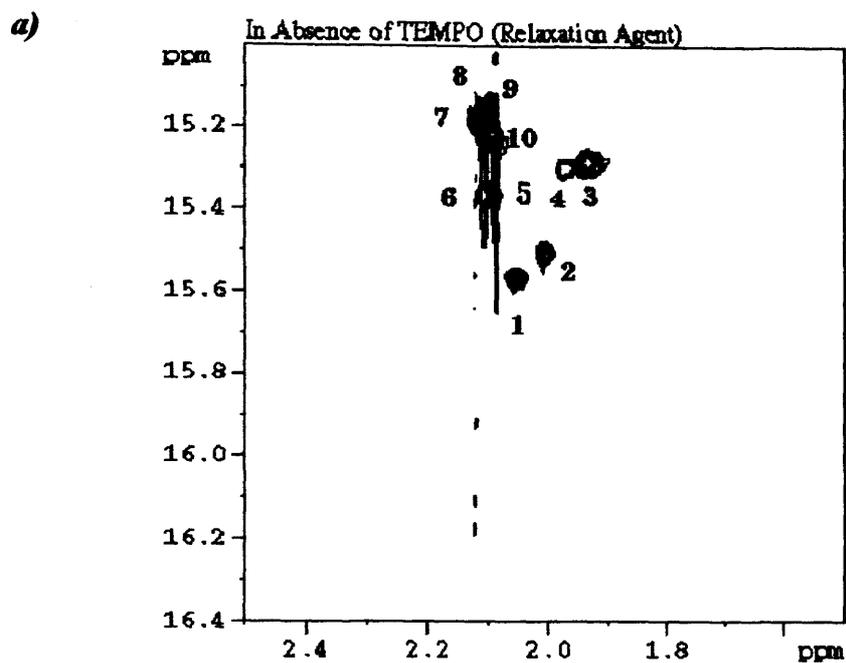


Figure 5.10. ^1H - ^{13}C HSQC spectrum of ^{13}C -labelled-Met Vt in the a) absence and b) presence of TEMPO. Vt is at 1mM and the temperature is 298K.

Chapter 6:

Discussion

Introduction

Role of C-terminal arm

Conformational changes

Relationship between actin and phospholipid binding

Other binding sites

Head-binding region

VASP, vinexin and ponsin

A pathway for activation

Perspectives

Summary

Introduction

In the previous chapter, a series of experiments demonstrated that Vt undergoes conformational changes involving its N- and C-terminal regions when in the presence of acidic phospholipids or F-actin, with an increased overall flexibility observed in the case of acidic phospholipids. This chapter re-examines the nature of these conformational changes in more detail under the light of the Vt crystal structure and, integrating current knowledge of vinculin structure-function relationships, works towards a model for vinculin activation. Preliminary results using Vt point mutants are described and a series of further experimental studies is outlined.

Role of C-terminal arm

In the presence of acidic phospholipid MLVs and under physiological pH and ionic strength, co-sedimentation of Vt but not Vt Δ C was observed (figure 5.7a). On lowering the pH however, co-sedimentation of Vt Δ C with PS vesicles was observed, with a midpoint of 6.5 (figure 5.7b). The pH midpoint suggests the titration of histidines. There are two histidines in Vt: one is contained within the basic collar (H906) while the other is on the opposite face of the domain next to the acidic cluster (H1025) (see figure 4.1a and 4.3a). H906 forms the base of the pocket that holds F885 from the N-terminal arm. Protonation of H906 would enhance the basicity of the collar, and could also weaken the bonds between H1 and the N-terminal arm, exposing the basic collar for binding to the membrane. These results show that the C-terminal arm is necessary for binding of Vt to acidic phospholipids at physiological pH, while increasing the positive charge on the mutant is sufficient to restore binding.

The C-terminal arm was shown not to be required for binding to F-actin (figure 5.1). However, Vt Δ C caused changes in bundling F-actin as monitored by light-scattering measurements at 340 nm (data not shown), suggesting that residues 1052-1066 affect the subsequent organisation of actin filaments in a manner that remains to be defined.

Conformational changes

In the presence of F-actin, Vt becomes more susceptible to proteases. N-terminal sequencing of the major proteolytic bands shows that in the presence of elastase Vt is cleaved after residue M900 (figure 5.3c), while in the case of chymotrypsin cleavage occurs after W912 and at the C-terminus after F1054 (figure 5.4c). M900 is located within H1 and

is buried under the N-terminal arm, while W912 from the H1-H2 turn packs into the hydrophobic core at the base of the bundle in the crystal structure and packs against W1058 from the C-terminal arm. Thus, cleavage of complexed Vt by either chymotrypsin or elastase occurs after residues that, based on the crystal structure, are inaccessible, leading to the conclusion that actin-bound vinculin is in a different conformation.

Access to M900 requires that the N-terminal arm peels away from H1, while access to W912 requires either that H1 peels away from the side of the bundle or that the β -strand within the C-terminal arm unplugs from the base of the bundle. The stopped-flow fluorescence data obtained with Vt and Vt Δ C in the presence of F-actin (figure 5.2) and inspection of the crystal structure suggest that the observed change in tryptophan packing is linked to the increased protease sensitivity of the N- and C-terminal arms.

To investigate whether binding to acidic phospholipids affected the conformation or flexibility of Vt, limited proteolysis was performed using a range of different proteases. In all cases tested, a greater rate of proteolysis was observed in the presence of acidic phospholipid MLVs than in the presence of PC. Thus, under conditions in which chymotrypsin or elastase caused little or no cleavage in the presence of PC, chymotrypsin causes extensive proteolysis when bound to PS MLVs vesicles (figure 5.7*b*), while elastase causes accumulation of an 18 kDa fragment (figure 5.7*a*), the same as found in the presence of actin (figure 5.3*c*). These results suggest an increase in flexibility of Vt on binding to acidic phospholipid MLVs.

The striking structural similarity of Vt with the exchangeable apolipoproteins (figure 4.6) raises the intriguing possibility that Vt functions in a similar fashion, by cycling between a soluble helical bundle and a membrane-bound unfurled form (Breiter et al., 1991).

Apolipoprotein E (apoE), apolipoprotein III and apolipoprotein Ia can all exist as helical bundles in the absence of lipid (Breiter et al., 1991; Kiss et al., 1999, Wilson et al., 1991). A truncation mutant of apolipoprotein Ia has also been crystallised, and appears to represent the unfurled form that binds lipoprotein particles (Borhani et al., 1997). In this form, the helices are intact, but they are fully extended into a large horseshoe, and the authors propose a model in which the hydrophobic interactions that stabilised the core of the helical bundle are replaced by interactions with the hydrocarbon core of the lipid.

Electron microscopy lends further support support the notion of bundle unfurling. Thus, images of chicken vinculin show a trilobar structure, representing the vinculin head, and a 200 Å long flexible “tail” described as four “pearls-on-a-string” (Molony and Burrige, 1985; Winkler et al., 1996). The “tail” is inconsistent with the crystal structure of Vt, which comprises a single domain 60 Å in length, but might be consistent with the helical hairpins of an unfurled structure.

Vt contains five helical segments. The N-terminal extension contains a further short helix and packs against H1, which could constitute the first hairpin. If H2+H3 formed a second hairpin and H4+H5 a third; 3 helical hairpins of length 30-50 Å together with extended connecting sequences of ~ 5 residues (~ 15 Å) makes a total length of ~ 150 Å. Adding to this the proline-rich region of ~ 50 residues which could constitute a fourth “pearl” (although nothing is as yet known about the structure of this region), the total length would approximately give the observed 200 Å tail. By contrast, images of pig vinculin show a globular structure with no “tail” evident (Gimona et al., 1987), which may represent the closed conformation in which the Vt helical bundle is packed against the trilobar head (see figure 6.1).

Relationship between actin and phospholipid binding

The truncation mutant (Vt Δ C) retains its specificity for acidic vesicles, albeit at a lower pH. To test the possibility that this binding could reflect the lower pI of Vt Δ C (calculated at 8.5 instead of 9.2 for Vt), co-sedimentation experiments of Vt and Vt Δ C were repeated in the presence of PIP₂, an acidic phospholipid containing more negative charges than PS. The results were identical to those obtained in the presence of PS (Bakolitsa et al., 1999) suggesting that although the C-terminus is required for the initial step of membrane attachment at physiological pH, other elements of the structure contribute to the specificity of acidic phospholipids.

A ladder of basic sidechains centered on H3 may also contribute to binding the acidic phospholipid head groups. Previous studies have shown that a peptide fragment, residues 916-970, which corresponds to the H2-H3 helical hairpin, mimics the lipid binding and inserting properties of the intact tail (Johnson et al., 1998). When isolated from the rest of the domain, the H2-H3 is highly amphipathic, and H3 in particular falls into the class of “surface-seeking” peptides (Eisenberg et al., 1994). Such an amphipathic helix could insert laterally into the vesicle surface with its hydrophobic sidechains inserting into the hydrocarbon core and the basic sidechains interacting with the acidic head groups. If this peptide does indeed mimic the full-length domain, it requires that the Vt helical bundle undergoes radical conformational changes on binding membrane.

The crystal structure bears no resemblance to other actin-binding domains (Puius et al., 1998). Several studies using recombinant peptides have attempted to localise the actin-binding site: fragments corresponding to H1-H3 or H1-H4 appear to bind actin, while

conflicting results have been reported for C-terminal fragments (Menkel et al., 1994; Hüttelmaier et al., 1997; Goldmann et al., 1998).

Given that activation of vinculin is also required for actin binding (Gilmore and Burridge, 1996; Weekes et al., 1996), it is possible that actin also binds to an (at least partly) unfurled form of the Vt bundle. The determinants of actin-binding and PIP2-binding to Vt appear to overlap substantially (Hüttelmaier et al., 1997; Menkel et al., 1994; Johnson et al., 1998), and a recent report demonstrates that binding of PIP2 to recombinant Vt actually inhibits actin binding (Steimle et al., 1999). These results suggest that Vt must be released from PIP2 prior to actin binding.

Both F-actin and acidic phospholipids have been shown to induce oligomerisation of Vt (Hüttelmaier et al., 1997; Hüttelmaier et al., 1998). In addition, expressed peptides containing residues 916-970 (H2-H3) and 1012-1066 (H5 and C-terminal arm) have been shown to migrate as dimeric species (Johnson and Craig, 1998). These same regions contain the methionine zipper (figure 4.4) with a high predicted probability of coiled-coil formation (figure 4.5). It is therefore conceivable that once the initial PIP2-induced has taken place, subsequent oligomerisation of Vt could occur. Oligomerisation of Vt has been used to propose a model of actin-binding (figure 1.8*b*) but possible mode(s) of vinculin transition from membrane-bound to actin-bound state remain to be established or indeed proposed.

Other binding sites

Head-binding region

The vinculin head domain, Vh, binds to the tail with high affinity (Johnson and Craig, 1994). In principle, Vh can inhibit the binding functions of Vt by directly masking the ligand

binding sites, by reducing the flexibility of the bundle and inhibiting conformational changes necessary for ligand binding (i.e. allosteric control), or by a combination of these mechanisms. Vh is acidic (calculated pIs: 5.18 (1-258); 5.63 (1-398); 5.86 (1-588)) whereas Vt is basic (calculated pI=9.2 for 879-1066); the head-tail intramolecular interaction is likely to be, at least in part, electrostatic in nature.

Point mutants recently produced by David Critchley's group in the acidic cluster region of H5 (figure 4.3), have shown a marked increase in Vh binding to Vt when acidic residues are mutated into arginines (manuscript in preparation). A possible explanation for this observation, would be that these residues, which are situated at the rim of the crevice formed at the top of the helical bundle, could act so as to weaken head-tail binding thereby facilitating the dynamic exchange between open and closed states. If this hypothesis is true, then it would follow that addition of negative charges in the crevice region (e.g. by phosphorylation), should further weaken the intramolecular head-tail association.

VASP, vinexin and ponsin

In the intact molecule, the residues immediately upstream of F885 are a polyproline sequence implicated in binding to VASP (Brindle et al., 1996) and ponsin (Mandai et al., 1999), and an acidic cluster (PPPRPPPPEEKDEE) implicated in vinexin-binding (Kioka et al., 1999). The crystal structure of Vt suggests that these residues are proximal to the basic collar surrounding the C-terminus where they may partly neutralise the positive charge on this region in full-length vinculin.

A pathway for activation

The above considerations lead to the following model for vinculin activation (figure 6.1). In the absence of acidic phospholipid, the trilobar vinculin head (Vh) binds and stabilises the helical bundle conformation of the vinculin tail (Vt), and masks the VASP/vinexin/ponsin binding sites in the proline-rich connecting segment and the binding site for talin on Vh. On encountering a high density of acidic phospholipid at the membrane surface, Vt is attracted by electrostatic interaction with the basic ladder and collar. Binding to PIP₂, which may be followed by insertion of the C-terminal hairpin, triggers an unfurling of the Vt helical bundle that disrupts the contacts with the head, unmasking the talin binding site on Vh, and releasing the surface-seeking potential of the Vt helices. The proline-rich region, which was sandwiched between Vh and the Vt bundle, becomes flexible and available for ligand binding.

Following PIP₂ activation and binding by talin and VASP/vinexin/ponsin, phospholipase activity causes Vt to dissociate from the membrane. Under certain conditions, the combination of head and neck sequestration may then be sufficient to prevent head-tail re-association, allowing Vt to bind F-actin. Subsequent cycles of phosphatidylinositol kinase and phospholipase activity would allow vinculin to alternate between membrane attachment and actin binding, thus contributing to the dynamic remodeling of focal adhesions sites that occurs during development, malignancy and other processes requiring cell adhesion and motility.

Perspectives

A number of issues remain to be addressed.

Mutants

Tyrosine and tryptophan residues are the most common residues in membrane-binding proteins at the boundary of the polar/non-polar interface between the hydrocarbon core and the hydrophilic head groups (Deisenhofer and Michel, 1989; Wimley and White, 1996). It would therefore be of interest to study the effect of mutants of W1064 and Y1065 on the ability of Vt to interact with acidic MLVs. Preliminary studies with recently produced point mutants of W1064 and Y1065 into arginines (David Critchley's lab) have not revealed a loss of the ability to co-sediment with acidic phospholipids under physiological conditions as judged by co-sedimentation assays and also reveal a similar increase in sensitivity in the presence of proteases (manuscript in preparation). However, co-sedimentation cannot distinguish between an electrostatic association of Vt with acidic MLVs and an insertion of Vt into the membrane, and limited proteolysis is a crude probe for conformational analysis between Vt and the YW=>RR point mutants. Other experiments, e.g. hydrophobic photolabelling or NMR (see below) need to be carried out in order to specifically analyse the role of these amino-acids in Vt-membrane interactions.

Mutational analyses on annexin V, additionally reveal the importance of conserved serine/threonine residues in interactions with membrane phospholipids. These residues, which also represent phosphorylation targets, can be studied by mutagenesis to yield information on their possible function in interactions with acidic phospholipid MLVs and in the intramolecular association of vinculin.

Other possible mutants would include residues involved in the basic ladder and mutation of H906. Cross-linking studies e.g. between H3 and H4 or H2 and H5, could additionally be attempted to further address the role of Vt unfurling in vinculin activation.

Complexes

A number of complexes with Vt can be envisaged. Vt binds paxillin, a major component of FAs that exhibits increased tyrosine phosphorylation following cell adhesion to ECM components (Burrige et al., 1992). One postulated role for paxillin is the recruitment of vinculin molecules to precursors of FAs through an interaction with Vt. The vinculin binding site for paxillin (located in H4) seems unaffected by vinculin conformation (Gilmore and Burrige, 1995). Other possible candidates for ligand formation with Vt include acidic phospholipids (or more soluble analogues) and N-terminal domains of vinculin. These complexes can be studied by crystallography or NMR methods to provide further insight into the role of vinculin at sites of adhesion.

In addition, experiments are planned to observe vinculin-binding by electron microscopy in collaboration with Dorit Hanein at the Burnham Institute, La Jolla.

NMR studies

The NMR studies of Vt so far obtained, have not supplied any new information regarding the conformational states of the molecule, with the main difficulty residing in obtaining suitable lipoprotein complexes. Similar studies obtained with apolipoprotein III (Weers et al., 1998) show clear chemical shift dispersion in both nuclei of HSQC spectra with a well defined resonance pattern. However, a number of differences exist between vinculin and apolipoprotein binding to lipids. Exchangeable apolipoproteins bind to neutral lipids like DAG whereas Vt has a well-established specificity for acidic phospholipids. In addition, vinculin has been shown to oligomerise in the presence of acidic phospholipid, a factor which, in the absence of a suitable lipid mimic, would make Vt-lipid complexes unsuitable candidates for studies by NMR. Solid-state NMR could be used to distinguish

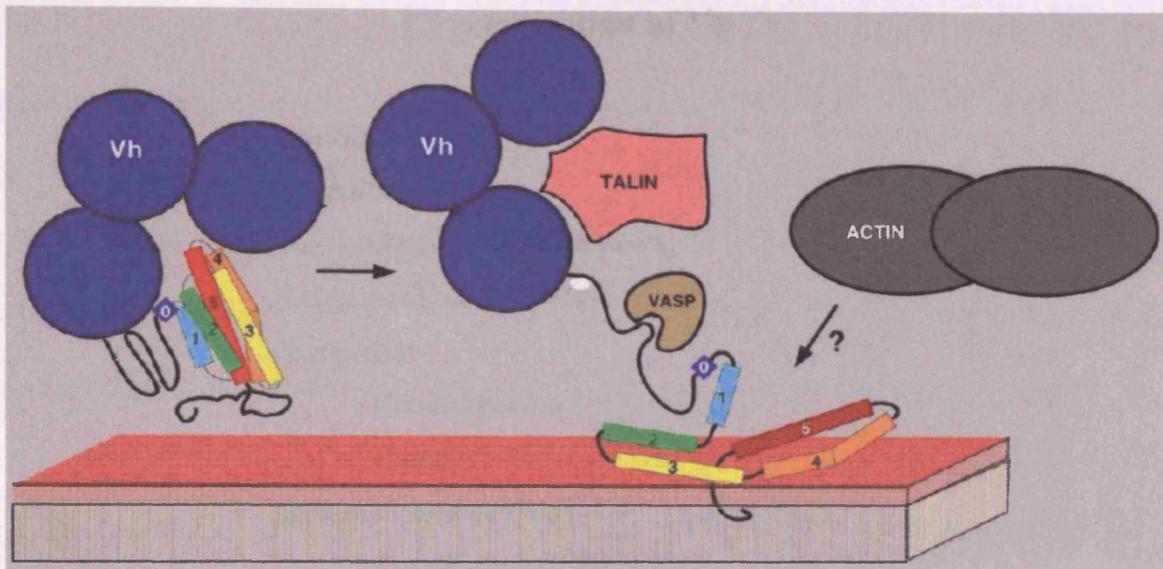
between whether Vt inserts into the membrane transversely or grazes it laterally and to ascertain the length and angle of insertion of helices into acidic MLVs. In addition, NMR studies using other ligands (e.g. paxillin peptides) could be pursued.

Crystallisation of Vg

Finally, a number of issues such as the exact nature of vinculin's intramolecular association, the function of the three 112-residue repeats and the structure of the proline-rich connecting hinge can only be addressed in the intact molecule. Cytosolic full-length vinculin is thought to exist in the closed (head-tail associated) conformation, so a crystal structure of this molecule would be capable of examining the nature of the head-tail interaction and therefore of its role in regulating cell adhesion.

Summary

The crystallographic and solution studies of Vt lead to a model for vinculin activation in which interaction with PIP2 triggers an unfurling of the Vt helical bundle that disrupts the contacts with the head, unmasking binding sites on Vh and the proline-rich region, and releasing the surface-seeking potential of the Vt helices. Subsequent phospholipase activity could cause Vt to dissociate from the membrane and, with the head and neck sequestered so as to prevent head-tail re-association, allow binding to F-actin. Cycling between membrane attachment and actin binding, vinculin could thus contribute to the dynamic remodeling of focal adhesions sites that occurs during development, malignancy and other processes requiring cell adhesion and motility. However, a number of further studies need to be carried out to further test the reliability of this model.



Summary

Figure 6.1. A model for vinculin activation. Binding of Vh to Vt allosterically inhibits unfolding of Vt. This inhibition is relieved by interaction of Vt with acidic phospholipids, which cause unfurling of the Vt bundle and sequestration of Vh with talin. The proline-rich region, previously occluded by the intramolecular interaction, becomes available for binding to VASP/vinexin/ponsin, and Vt can bind F-actin. Subsequent cycles of phosphatidylinositol kinase and phospholipase activity would allow vinculin to alternate between membrane attachment and actin binding, thus contributing to the dynamic remodelling of focal adhesions.

Chapter 7:

Crystallisation of Vg

Introduction

Crystals of Vg

Data collection and processing

Heavy metal derivatives

Discussion

Crystallisation

Phasing

Materials and methods

Purification of Vg

Summary

Introduction

In April 1997, after several months of unsuccessful attempts at growing single and sizeable Vt crystals, I undertook work on the full-length gizzard vinculin, Vg. Low-resolution diffracting crystals of Vg were rumoured to have been in existence for several years, although both crystallisation conditions and progress at solving the structure of Vg remained unknown.

This chapter describes the materials and methods used to purify Vg, the crystallisation conditions leading to the initial characterization of Vg crystals, and first attempts at crystal optimisation and phasing.

Crystals of Vg

Factorial screening of Vt at a protein concentration of ~ 4 mg/mL, resulted in the transitory appearance of a microcrystalline precipitate grown out of 30% PEG 5000 MME, 0.2 M A.S. and 0.1 M MES pH 6.5 at 22°C. Addition of magnesium sulfate to 10 mM produced fairly large ($0.3 \times 0.2 \times 0.02$ mm³) stacked plate-like crystals grown by the hanging-drop diffusion method, which when capillary-mounted on the in-house rotating anode, showed diffraction to 8 Å.

A crystal is shown in figure 7.1 and a diffraction image is shown in figure 7.2. The space group is C222₁ with unit cell parameters $a = 56.3$ Å, $b = 127.1$ Å, $c = 352.9$ Å. Crystals have one molecule per asymmetric unit and a solvent content of 52%. They diffract synchrotron radiation to better than 4 Å.

Data collection and processing

Diffraction data from native and derivative crystals were collected on a 30cm MAR plate at a wavelength of 0.87 Å. All data were processed with the programs DENZO and SCALEPACK. Data collection statistics for the native crystal are summarised in table 7.1.

Heavy metal derivatives

Three soaks of compounds exhibiting anomalous scattering at 0.87 Å (Ir, Hg, and Pt) were soaked into native Vg crystals and datasets collected essentially as for the native crystal. The mosaicity of native and putative derivative crystals varies between 0.4 and 1.1° using 20% glycerol as the cryoprotectant. A summary of the data collection statistics for these crystals is given in table 7.2.

Discussion

Crystallisation

Crystals of purified Vg have been reproducibly grown, leading to space-group characterisation and the first attempts at obtaining derivatives. However, several problems in this area have to be addressed before work on this project can be further advanced. First, in the majority of cases, nucleation takes place in two steps. Numerous, mostly stacked, plate-like Vg crystals appear within a couple of weeks of a crystallisation setup but are too thin along the c axis to be of use for data collection. Large crystals take several months to form and consistently appear out of partially dried-out/phase-separated reservoirs. The poor growth along the c axis, can be addressed by macroseeding, additives and detergents or a combination thereof, while control of the second nucleation step will require a more elaborate approach.

Second, the crystals obtained diffract synchrotron radiation to medium resolution in one direction, and show some lattice disorder along the b axis. The disorder could be the result of a number of factors ranging from suboptimal cryoprotectant conditions to microheterogeneity in the protein preparation. The quality of the cryoprotectant can be monitored by changes in the mosaicity of native crystals, while microheterogeneity of the protein sample (which could also account for the inability to obtain high resolution diffraction thus far) can be resolved either by producing Vg via an expression system, or by attempting to isolate a major isoform e.g. by preparative IEF.

Finally, the appearance of crystals equally grown in all directions but too small for data collection, indicates that basic crystallisation parameters, such as protein and precipitant concentrations, still leave room for optimisation.

Phasing

The current native dataset of Vg is two thirds complete. Before phasing by the method of MIR can be seriously attempted, the native dataset will have to be completed. Derivative data collected so far are encouraging in that isomorphism (as indicated by changes in unit cell length) does not seem to present a problem (see table 7.2). However the overall phasing quality of these data sets remains to be assessed.

Materials and Methods

Purification of Vg

The intact 117 kDa vinculin was isolated and purified from chicken gizzard smooth muscle according to Evans et al. (1984) and O'Halloran et al. (1986). All steps were carried out at 4°C unless otherwise stated. The pH adjustment of buffers was done at the temperature at which the buffer was to be used.

After trimming the chicken gizzards from any fat/connective tissue using a scalpel blade, the muscle was cut in pieces and weighed out. All subsequent steps refer to 100 g of starting material. The tissue was homogenized with a Waring blender (2x10 sec bursts at top speed) in 7-10 volumes of deionized water containing 0.5 mM PMSF. The homogenate was centrifuged (16000g, 10min) and the pellet washed again as above. After the second centrifugation step, the pellet was extracted with 1L buffer A (2 mM Tris-HCl, 1 mM EGTA, 0.5 mM EGTA, pH 9.0) for 1 hour at 37°C to extract vinculin from the myofibrils, and centrifuged (26000g, 10 min, 37°C). 10 mM MgCl₂ was added to the vinculin-containing supernatant to precipitate most of the desmin and actin stirred 15 min at RT and the pH adjusted to 7.0. After centrifugation (16000g, 10 min, RT) the supernatant was

fractionated by progressive addition of 20.5 g of $(\text{NH}_4)_2\text{SO}_4$ /100 mL of supernatant, yielding a pellet constituting the crude vinculin extract.

The ammonium sulphate pellet comprising the crude vinculin extract was resuspended in Buffer B (20 mM NaCl, 15 mM β -ME, 20 mM Tris-acetate pH 5.0) and extensively dialysed against the same buffer. Before dialysis, PMSF was added to 2 mM to control proteolysis. The dialysed protein was then centrifuged (1000g, 10 min) and the supernatant applied to a Whatman CM52 cellulose column (1.4 x 40 cm) that had been pre-equilibrated in Buffer B. After loading, the column was washed in Buffer B until the absorbance at 280 nm attained a steady base-line value. Vinculin was then eluted from the column with a 250 mL linear gradient of 200-500 mM NaCl. Column fractions were analyzed by SDS-PAGE and vinculin-containing fractions pooled and dialyzed overnight against Buffer C (20 mM NaCl, 0.1% β -ME, 0.1 mM EDTA, 20 mM sodium acetate, pH 5.0). The pooled and dialyzed fractions were loaded onto a CM-52-cellulose column previously equilibrated in Buffer C and the column eluted with a 100 mL linear gradient of 100-500 mM NaCl. The fractions were analyzed for vinculin by SDS-PAGE.

The protein at various stages of purification is shown in figure 7.3.

Summary

Initial crystallisation conditions have been determined for the intact chicken gizzard vinculin, Vg. The space group has been characterised, and medium resolution data obtained for native and derivative crystals. Future work will be focusing on optimising protein expression and purity, as well as crystallisation conditions.

Resolution Shell (Å)	total unique	% complete	Average redundancy	$\langle I \rangle / \langle \text{err} \rangle$	Rsym	
99.00	8.40	829	62.2	4.03	33.7	0.045
8.40	6.66	809	65.5	4.21	21.2	0.089
6.66	5.82	793	65.8	4.09	8.5	0.238
5.82	5.29	818	67.8	4.09	6.8	0.314
5.29	4.91	805	66.7	4.03	8.3	0.261
4.91	4.62	800	67.9	4.07	8.5	0.265
4.61	4.39	834	69.8	4.05	7.6	0.266
4.39	4.20	820	68.1	4.11	5.6	0.367
4.20	4.04	802	68.3	3.95	4.4	0.454
4.04	3.90	764	63.4	3.93	3.6	0.543
99.00	3.90	8074	66.6	4.06	12.4	0.152

Table 7.1. Data collection statistics for a native crystal of Vg. The data set was collected at SRS Daresbury, station 9.6 ($\lambda = 0.87 \text{ \AA}$) from a flash-frozen crystal at 100K.

Crystal	Cell (a, b, c) (Å)	Mosaicity	Resolution Limits (Å)	I/error (to dmin)	Compl. (%)
Native	56.3 127.1 352.9	0.82	99.00—3.90	12.4 (3.6)	66.6
IrCl ₆	56.3 126.4 351.9	0.91	99.00—4.49	20.3 (5.9)	80.9
Pt(II)	57.3 126.6 352.2	0.38	99.00—4.39	21.3 (5.3)	95.5
HgAc ₂	55.7 129.2 354.7	1.13	99.00—3.88	17.1 (3.9)	62.7

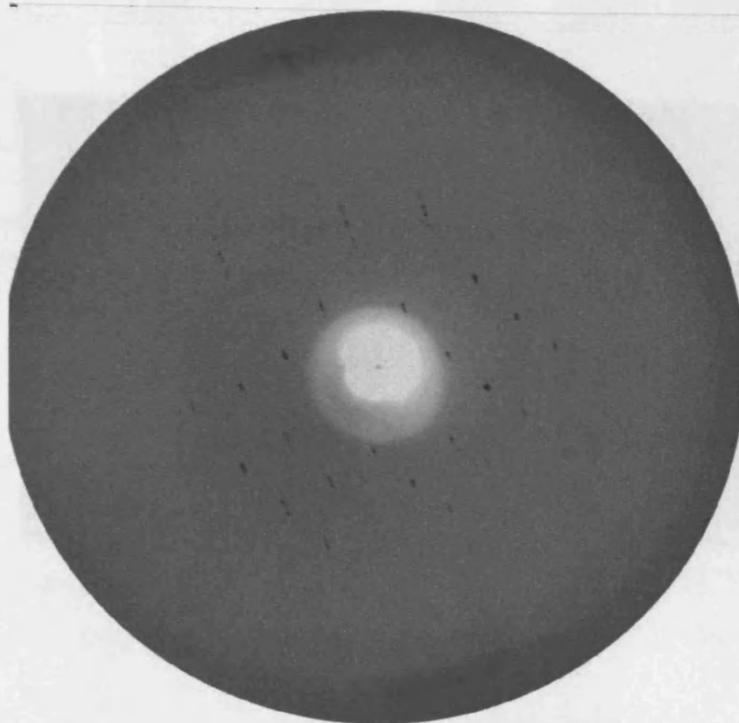
Table 7.2. Summary of data collection statistics for native and putative derivative data sets of Vg. All datasets were collected at SRS Daresbury, station 9.6 ($\lambda = 0.87 \text{ \AA}$) from flash-frozen crystals at 100K.



Figure 7.1. A crystal of Vg. Crystals grow as plates and can become quite large. The one shown has dimensions $0.3 \times 0.2 \times 0.02 \text{ mm}^3$ and is oriented with the c axis vertical.

Figure 7.2. (A) Small-angle image of a Vg crystal of View with the long axis, c^* , almost vertical. Resolution at the outer edge is 3.7 \AA . Exposure time = 100 s, crystal-to-film distance = 700 mm, $\lambda = 0.47 \text{ \AA}$. (B) Close-up view of (A).

a)



b)

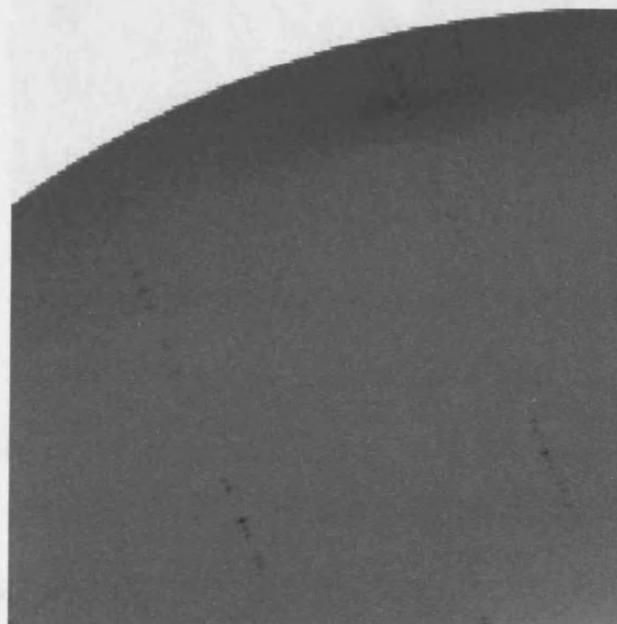
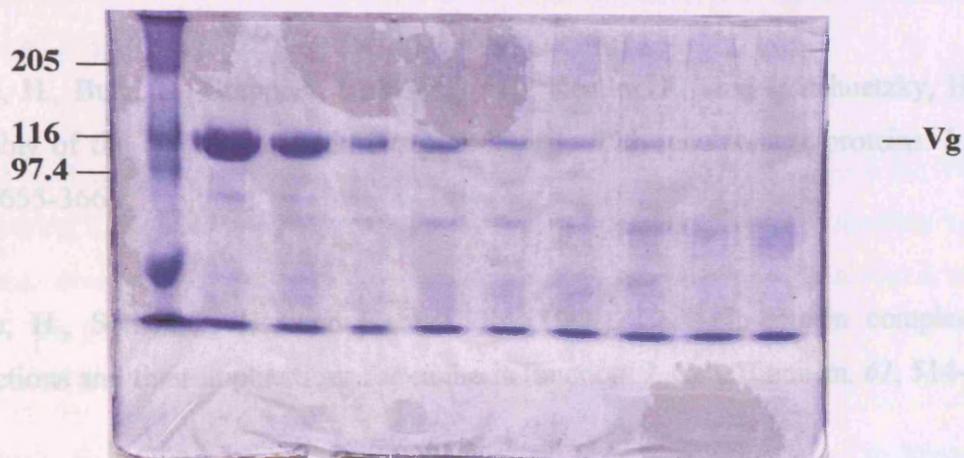


Figure 7.2. 1° oscillation images of a Vg crystal. *a)* View with the long axis, c^* , almost vertical. Resolution at the outer edge is 3.9 \AA . Exposure time = 300 s; crystal-to-film-distance = 700 mm; $\lambda = 0.87 \text{ \AA}$. *b)* Close-up view of *a)*.

a)



b)

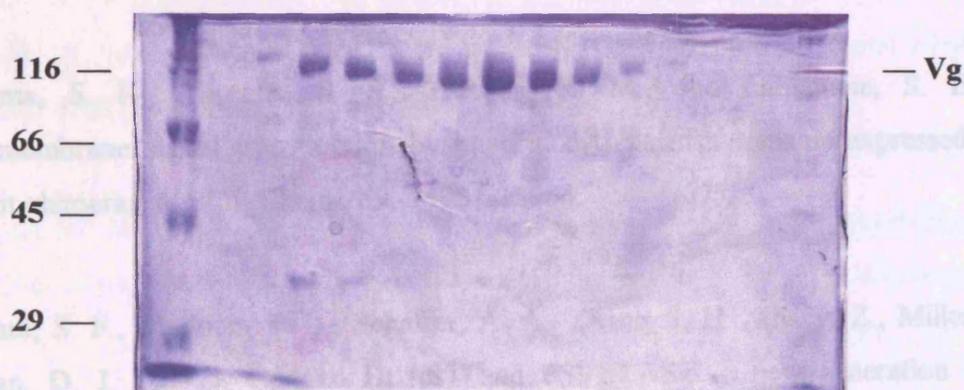


Figure 7.3. Purification of gizzard smooth-muscle vinculin (Vg) by ion-exchange. Fractions are *a*) from a DEAE column and *b*) from a CM-52 cellulose column, on a linear NaCl gradient. Molecular weight markers (in kDa) are indicated.

REFERENCES

- Aberle, H., Butz, S., Stappert, J., Weissig, H., Kemler, R., and Hoschuetzky, H. (1994). Assembly of the cadherin catenin complex in vitro with recombinant proteins. *J. Cell Sci.* *107*, 3655-3663.
- Aberle, H., Schwartz, H., and Kemler, R. (1996). Cadherin-catenin complex: protein interactions and their implications for cadherin function. *J. Cell. Biochem.* *61*, 514-523.
- Adams, P. O., Pannu, N. S., Read, R. J., and Brunger, A. T. (1999). Extending the limits of molecular replacement through combined simulated annealing and maximum-likelihood refinement. *Acta Cryst.* *D55*, 181-190.
- Akiyama, S. K., Yamada, S. S., Yamada, K. M., and LaFlamme, S. E. (1994). Transmembrane signal transduction by integrin cytoplasmic domains expressed in single subunit chimeras. *J. Biol. Chem* *269*, 15961-15964.
- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J. H., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res.* *25*, 3389-3402.
- Bacon, D. J., and Anderson, W. F. (1988). A fast algorithm for rendering space-filling molecule pictures. *J. Mol. Graphics* *6*, 219-220.
- Bakolitsa, C., de Pereda, J. M., Bagshaw, C. R., Critchley, D. R., and Liddington, R. C. (1999). Crystal structure of the vinculin tail suggests a pathway for activation. *Cell* *99*, 603-613.

Balzac, F., Belkin, A. M., Koteliansky, V. E., Balabanov, Y. V., Altruda, F., Silengo, L., and Tarone, G. (1993). Expression and functional-analysis of a cytoplasmic domain variant of the beta-1 integrin subunit. *J. Cell Biol.* *121*, 171-178.

Belkin, A., M., Zhidkova, N. I., Balzac, F., Altruda, F., Tomatis, D., Maier, A., Tarone, G., Koteliansky, V. E., and Burridge, K. (1996). beta 1D integrin displaces the beta 1A isoform in striated muscles: localization at junctional structures and signaling potential in nonmuscle cells. *J. Cell Biol.* *132*, 211-226.

Bergelson, J. M., and Hemler, M. E. (1995). Integrin-ligand binding - Do integrins use a MIDAS touch to grasp an ASP. *Curr. Biol.* *5*, 615-617.

Borhani, D. W., Rogers, D. P., Engler, J. A., and Brouillette, C. G. (1997). Crystal structure of truncated human apolipoprotein A-I suggests a lipid-bound conformation. *Proc.Natl.Acad.Sci.USA* *94*, 12291-12296.

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

Braga, V. M., Del Maschio, A., Machesky, L., and Dejana, E. (1999). Regulation of cadherin function by Rho and Rac: modulation by junction maturation and cellular context. *Mol. Biol. Cell* *10*, 9-22.

Branden, C., and Tooze, J. (1999). Introduction to protein structure. Garland Publishing, New York, 2nd edition.

Breiter, D. B., Kanost, M. R., Benning, M. M., Wesenberg, G., Law, J. H., Wells, M. A., Rayment, I., and Holden, H. M. (1991). Molecular structure of an apolipoprotein determined at 2.5 Å resolution. *Biochemistry* 30, 603-608.

Briesewitz, R., Kern, A., and Marcantonio, E. E. (1993). Ligand-dependent and -independent integrin focal contact localization: the role of the α chain cytoplasmic domain. *Mol. Biol. Cell* 4, 593-604.

Briesewitz, R., Kern, A., and Marcantonio, E. E. (1995). Assembly and function of integrin receptors is dependent on opposing α and β cytoplasmic domains. *Mol. Biol. Cell* 6, 997-1010.

Brindle, N. P. J., Holt, M. R., Davies, J. E., Price, C. J., and Critchley, D. R. (1996). The focal adhesion vasodilator-stimulated phosphoprotein (VASP) binds to the proline-rich domain in vinculin. *Biochem. J.* 318, 753-757.

Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., GrosseKunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Cryst.* D54, 905-921.

Budisa, N., Steipe, B., Demange, P., Eckerskorn, C., Kellerman, J., and Huber, R. (1995). High-level biosynthetic substitution of methionine in proteins by its analogs 2-aminohexanoic acid, selenomethionine, telluromethionine and ethionine in *Escherichia coli*. *Eur. J. Biochem.* 230, 788-796.

Burn, P., and Burger, M. M. (1987). The cytoskeletal protein vinculin contains transformation-sensitive, covalently bound lipid. *Science* 235, 476-479.

Burridge, K., and Feramisco, J. R. (1980). Microinjection and localization of a 130K protein in living fibroblasts: a relationship to actin and fibronectin. *Cell* 19, 587-595.

Burridge, K. and Chrzanowska-Wodnicka, M. (1996). Focal adhesions, contractility, and signaling. *Annu. Rev. Cell Dev. Biol.* 12, 463-519.

Burtnick, L. D., Koepf, E. K., Grimes, J., Jones, E. Y., Stuart, D. I., McLaughlin, P. J., and Robinson, R. C. (1997). The crystal structure of plasma gelsolin: implications for actin severing, capping, and nucleation. *Cell* 90, 661-670.

Campos, B., Mo, Y. D., Mealy, T. R., Li, C. W., Swairjo, M. A., Balch, C., Head, J. F., Retzinger, G., Dedman, J. R., and Seaton, B. A. (1998). Mutational and crystallographic analyses of interfacial residues in annexin V suggests direct interactions with phospholipid membrane components. *Biochemistry* 37, 8004-8010.

Chen, Y., O'Toole, E. E., Shipley, T., Forsyth, J., LaFlamme, S. E., Yamada, K. M., Shattil, S. J., and Ginsberg, M. H. (1994). Inside-out signal transduction inhibited by isolated integrin cytoplasmic domains. *J. Biol. Chem* 269, 18307-18310.

Chong, L. D., Traynorkapaln, A., Bokoch, G. M., and Schwartz, M. A. (1994). The small GTP-binding protein-rho regulates a phosphatidylinositol 4-phosphate 5-kinase in mammalian-cells. *Cell* 79, 507-513.

Christofori, G., and Semb, H. (1999). The role of the cell-adhesion molecule E-cadherin as a tumour-suppressor gene. *Trends Biochem. Sci.* 24, 73-76.

Clayton, D., Brereton, I. M., Kroon, P. A., and Smith, R. (1999). NMR studies of the low-density lipoprotein receptor-binding peptide of apolipoprotein E bound to dodecylphosphocholine micelles. *Protein Sci.* 8, 1797-1805.

Collaborative Computational Project, N. 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Cryst. D50*, 760-763.

Crick, F. H. C., and Magdoff, B. S. (1956). The theory of the method of isomorphous replacement for protein crystals. I. *Acta Cryst.* 9, 901.

Critchley, D. R., Gilmore, A., Hemmings, L., Jackson, P., McGregor, A., Ohanian, V., Patel, B., Waites, G., Wood, C. (1991). Cytoskeletal proteins in adherens-type cell matrix junctions. *Bioch. Soc. Trans.* 19, 1028-1033.

Crowther, R. A., and Blow, D. M. (1967). A method of positioning a known molecule in an unknown crystal structure. *Acta Cryst.* 23, 544-548.

Deisenhofer, J., and Michel, H. (1989). The photosynthetic reaction center from the purple bacterium, *Rhodospseudomonas viridis*. *Science* 245, 1463-1473.

DePasquale, J. A., and Izzard, C. S. (1987). Evidence for an actin-containing cytoplasmic precursor of the focal contact and the timing of incorporation of vinculin at the focal contact. *J. Cell Biol.* 105, 2803-2809.

Drenth, J. (1999). Principles of protein X-ray crystallography. Springer-Verlag, 2nd edition.

Eimer, W., Niermann, M., Eppe, M. A., and Jockusch, B. M. (1993). Molecular shape of vinculin in aqueous solution. *J. Mol. Biol.* 229, 146-152.

Eisenberg, D., Schwarz, E., Komaromy, M., and Wall, R. (1984). Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J.Mol.Biol.* 179, 125-142.

Evans, R. R., Robson, R. M., and Stromer, M. H. (1984). Properties of smooth muscle vinculin. *J. Biol. Chem.* *259*, 3916-3924.

Fernandez, J. L. R., Geiger, B., Salomon, D., and Benzeev, A. (1993). Suppression of Vinculin Expression By Antisense Transfection Confers Changes in Cell Morphology, Motility, and Anchorage-Dependent Growth of 3t3-Cells. *J. Cell. Biol.* *122*, 1285-1294.

Foty, R. A., and Steinberg, M. S. (1997). Measurement of tumor cell cohesion and suppression of invasion by E- or P- cadherin. *Cancer. Res.* *57*, 5033-5036.

Fukami, K., Furuhashi, K., Inagaki, M., Endo, T., Hatano, S., and Takenawa, T. (1992). Requirement for phosphatidylinositol 4,5-bisphosphate for alpha-actinin function. *Nature* *359*, 150-152.

Fukami, K., Endo, T., Imamura, M., and Takenawa, T. (1994). alpha-Actinin and Vinculin Are PIP(2)-Binding Proteins Involved in Signaling by Tyrosine Kinase. *J. Biol. Chem.* *269*, 1518-1522.

Gallin, W. J. (1998). Evolution of the 'classical' cadherin family of cell adhesion molecules in vertebrates. *Mol. Biol. Evol.* *15*, 1099-1107.

Geiger, B. (1979). A 130K protein from chicken gizzard: its localization at the termini of microfilament bundles in cultured chicken cells. *Cell* *18*, 193-205.

Geiger, B., Tokuyasu, K. T., Dutton, A. H., and Singer, S. J. (1980). Vinculin, an intracellular protein localized at specialized sites where microfilament bundles terminate at cell membranes. *Proc. Natl. Acad. Sci. USA* 77, 4127-4131.

Geiger, B. (1982). Microheterogeneity of Avian and Mammalian Vinculin - Distinctive Sub-Cellular Distribution of Different Isovinculins. *J. Mol. Biol.* 159, 685-701.

Giancotti, F.G., Mainiero, F. (1994). Integrin-mediated adhesion and signaling in tumorigenesis. *Biochim. Biophys. Acta*, 1198, 47-64.

Giancotti, F. G. (1997). Integrin signaling: specificity and control of cell survival and cell cycle progression. *Curr. Opin. Cell Biol.* 9, 691-700.

Gill, S. C, and von Hippel., P. H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* 182, 319-326.

Gilmore, A. P., and Burridge, K. (1996). Regulation of vinculin binding to talin and actin by phosphatidylinositol-4-5-bisphosphate. *Nature* 381, 531-535.

Gilmore, A. P., Jackson, P., Waites, G. T., and Critchley, D. R. (1992). Further characterisation of the talin binding site in the cytoskeletal protein vinculin. *J. Cell Sci.* 103, 719-731.

Gimona, M., Furst, D. O., and Small, J. V. (1987). Metavinculin and vinculin from mammalian smooth-muscle - bulk isolation and characterization. *J. Muscle Res. Cell Motil.* 8, 329-341.

Gimona, M., Small, J. V., Moeremans, M., Vandamme, J., Puype, M., and Vandekerckhove, J. (1988). Porcine vinculin and metavinculin differ by a 68-residue insert located close to the carboxy-terminal part of the molecule. *EMBO J.* 7, 2329-2334.

Goldmann, W. H., Guttenberg, Z., Tang, J. X., Kroy, K., Isenberg, G., and Ezzell, R. M. (1998). Analysis of the F-actin binding fragments of vinculin using stopped-flow and dynamic light-scattering measurements. *Eur. J. Biochem.* 254, 413-419.

Hagmann, J., and Burger, M. M. (1992). Phosphorylation of vinculin in human platelets spreading on a solid-surface. *J. Cell. Biochem.* 50, 237-244.

Hall, A. (1994). Small GTP-binding proteins and the regulation of the actin cytoskeleton. *Annu. Rev. Cell Biol.* 10, 31-54.

Harrison, C. J., Bohm, A. A., and Nelson, H. C. M. (1994). Crystal structure of the DNA binding domain of the heat shock transcription factor. *Science* 263, 224-227.

Hazan, R. B., Kang, L., Roe, S., Borgen, P. I., Rimm, D. L. (1997). Vinculin is associated with the E-cadherin adhesion complex. *J. Biol. Chem* 272, 32448-32453.

Hemmings, L., Rees, D. J. G., Ohanian, V., Bolton, S. J., Gilmore, A. P., Patel, B., Priddle, H., Trevithick, J. E., Hynes, R. O., and Critchley, D. R. (1996). Talin contains three actin-binding sites each of which is adjacent to a vinculin-binding site. *J. Cell Sci.* 109, 2715-2726.

Herrenknecht, K., Ozawa, M., Eckerskorn, C., Lottspeich, F., Lenter, M., and Kemler, R. (1991). The Uvomorulin-anchorage protein Alpha-catenin is a vinculin homolog. *Proc. Natl. Acad. Sci., USA* 88, 9156-9160.

Holm, L., and Sander, C. (1997). Dali/FSSP classification of three-dimensional protein folds. *Nucleic Acids Res.* 25, 231-234.

Hooft, R. W. W., Vriend, G., Sander, C., and Abola, E. E. (1996). Errors in protein structures. *Nature* 381, 272.

Huttelmaier, S., Bubeck, P., Rudiger, M., and Jockusch, B. M. (1997). Characterization of two F-actin-binding and oligomerization sites in the cell-contact protein vinculin. *Eur. J. Biochem.* 247, 1136-1142.

Huttelmaier, S., Mayboroda, O., Harbeck, B., Jarchau, T., Jockusch, B. M., and Rudiger, M. (1998). The interaction of the cell-contact proteins VASP and vinculin is regulated by phosphatidylinositol-4,5-bisphosphate. *Current Biology* 8, 479-488.

Hynes, R.O. (1992). Integrins: versatility, modulation and signaling in cell adhesion. *Cell* 69, 11-25.

Ito, S., Werth, D. K., Richert, N. D., and Pastan, I. (1983). Vinculin phosphorylation by the src kinase. *J. Biol. Chem.* 258, 14626-14631.

Jockusch, B. M., and Rudiger, M. (1996). Crosstalk between cell adhesion molecules: vinculin as a paradigm for regulation by conformation. *TICB* 6, 311-315.

Johnson, R. P., and Craig, S. W. (1995). The carboxy-terminal tail domain of vinculin contains a cryptic binding site for acidic phospholipids. *Biochem. Biophys. Res. Commun.* 210, 159-164.

Johnson, R. P., and Craig, S. W. (1995). F-actin binding site masked by the intramolecular association of vinculin head and tail domains. *Nature* 373, 261-264.

Johnson, R. P., and Craig, S. W. (1994). An Intramolecular Association Between the Head and Tail Domains of Vinculin Modulates Talin Binding. *J. Biol. Chem.* 269, 12611-12619.

Johnson, R. P., and Craig, S. W. (1994). Vinculin tail domain is a globular monomer with cryptic oligomerization potential. *Mol. Biol. Cell* 9, (SS) 98.

Johnson, R. P., Niggli, V., Durrer, P., and Craig, S. W. (1998). A conserved motif in the tail domain of vinculin mediates association with and insertion into acidic phospholipid bilayers. *Biochemistry* 37, 10211-10222.

Jones, P., Jackson, P., Price, G. J., Patel, B., Ohanion, V., Lear, A. L., and Critchley, D. R. (1989). Identification of a talin binding-site in the cytoskeletal protein vinculin. *J. Cell Biol.* 109, 2917-2927.

Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991). Improved methods for building protein models into electron density maps and the location of errors in these models. *Acta Crystallogr.* A47, 110-119.

Kioka, N., Sakata, S., Kawauchi, T., Amachi, T., Akiyama, S. K., Okazaki, K., Yaen, C., Yamada, K. M., and Aota, S. (1999). Vinexin: A novel vinculin-binding protein with multiple SH3 domains enhances actin cytoskeletal organisation. *J. Cell Biol.* 144, 59-69.

Kiss, R. S., Kay, C. M., and Ryan, R. O. (1999). Amphipathic α -helix bundle organization of lipid-free chicken apolipoprotein A-I. *Biochemistry* 38, 4327-4334.

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

Kreis, T., Vale R. (1999). Guidebook to the extracellular matrix, anchor, and adhesion proteins. Second edition, Oxford University Press.

Kroemker, M., Rudiger, A. H., Jockusch, B. M., and Rudiger, M. (1994). Intramolecular interactions in vinculin control α -actinin binding. *FEBS Letters* 355, 259-262.

Kuroda, S., Fukata, M., Nakagawa, M., Nakamura, T., Ookubo, T., Izawa, I., Nagase, T., Nomura, N., Tani, H. *et al.* (1998). Role of IQGAP1, a target of the small GTPases Cdc42 and Rac1, in regulation of E-cadherin-mediated cell-cell adhesion. *Science* 281, 832-835.

Laemmli, U. K. Cleavage of structural proteins during assembly of the head of bacteriophage T4. (1970). *Nature* 227, 680-685.

LaFlamme, S. E., Thomas, L. A., Yamada, S. S., and Yamada, K. M. (1994). Single subunit chimeric integrins as mimics and inhibitors of endogenous integrin functions in receptor localization, cell spreading and migration, and matrix assembly. *J. Cell Biol.* 126, 1287-1298.

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

Lee, J. O., Rieu, P., Arnaout, M. A., and Liddington, R. (1995). Crystal-structure of the A-domain from the A-subunit of integrin CR3 (CD11B/CD18). *Cell* 80, 631-638.

Lee, S. W., Reimer, C. L., Campbell, D. B., Cheresch, P., Duda, R. B., and Kocher, O. (1998). H-cadherin expression inhibits in vitro invasiveness and tumor formation in vivo. *Carcinogenesis* 19, 1157-1159.

Leslie, A. G. W. (1988). A reciprocal space algorithm for calculating molecular envelope using the algorithm of B. C. Wang. CCP4 Daresbury Study Weekend, nos. DL/SCI/R26, ISSN 0144-5677. Warrington WA4 4AD, UK: Daresbury Laboratory, for Daresbury Laboratory.

Lin, E. C. K., Ratnikov, B. I., Tsai, P. R., Gonzalez, E. R., McDonald, S., Pelletier, A. J., and Smith, J. W. (1997). Evidence that the integrin beta 3 and beta 5 subunits contain a metal ion-dependent adhesion site-like motif but lack an I domain. *J. Biol. Chem.* 272, 14236-14243.

Linask, K. K., Ludwig, C., Han, M. D., Liu X., Radice, G. L., and Knudsen K. A. (1998). N cadherin/catenin-mediated morphoregulation of somite formation. *Dev. Biol.* 202, 85-102.

Lupas, A., Van Dyke, M., and Stock, J. (1991). Predicting coiled-coils from protein sequences. *Science.* 252, 1162-1164.

Luzzati, V. (1952). Traitement statistique des erreurs dans la determination des structures cristallines. *Acta Cryst.* 5, 802-810.

Machesky, L. M., and Hall, A. (1996). Rho: a connection between membrane receptor signalling and the cytoskeleton. *Trends Cell Biol.* 6, 304-310.

Mandai, K., Nakanishi, H., Satoh, A., Takahashi, K., Satoh, K., Nishioka, H., Mizoguchi, A., and Takai, Y. (1999). Ponsin/SH3P12: An 1-afadin- and vinculin-binding protein localized at cell-cell and cell-matrix adherens junctions. *J. Cell Biol.* *144*, 1001-1017.

Marcantonio, E. E. and David, F. S. (1997). Integrin receptor signaling: the propagation of an α -helix model. *Matrix Biology.* *16*, 179-184.

Matthews, B. W. (1966). The determination of the position of anomalously scattering heavy atom groups in protein crystals. *Acta Cryst.* *20*, 230.

Matthews, B. W. (1968). Solvent content of protein crystals. *J. Mol. Biol.* *33*, 491-497.

Menkel, A. R., Kroemker, M., Bubeck, P., Ronsiek, M., Nikolai, G., and Jockusch, B. M. (1994). Characterisation of an F-actin-binding domain in the cytoskeletal protein vinculin. *J. Cell Biol.* *126*, 1287-1298.

Meredith, J., Takada, Y., Fornaro, M., Languino, L. R., and Schwartz, M. A. (1995). Inhibition of cell-cycle progression by the alternatively spliced integrin beta(1c). *Science* *269*, 1570-1572.

Merrit, E. A., and Murphy, M. E. P. (1994). Raster3D version 2.0. A program for photorealistic molecular graphics. *Acta Crystallogr.* *D50*, 869-873.

Michishita, M., Videm, V., Arnaout, M. A. (1993). A novel divalent cation-binding site in the A-domain of the beta-2-integrin-CR3 (CD11B/CD18) is essential for ligand-binding. *Cell* *72*, 857-867.

- Milam, L. M. (1985). Electron microscopy of rotary shadowed vinculin and vinculin complexes. *J. Mol. Biol.* *184*, 543-545.
- Misra, S., and Hurley, J. H. (1999). Crystal structure of a phosphatidylinositol 3-phosphate-specific membrane-targeting motif, the FYVE domain of Vps27p. *Cell* *97*, 657-666.
- Molony, L., and Burridge, K. (1985). Molecular shape and self-association of vinculin and metavinculin. *J. Cell. Biochem.* *29*, 31-36.
- Nakagawa, S., and Takeichi, M. (1998). Neural crest emigration from the neural tube depends on regulated cadherin expression. *Development* *125*, 2963-2971.
- Narayanaswami, V., Wang, J., Schieve, D., Kay, C. M., and Ryan, R. O. (1999). A molecular trigger of lipid binding-induced opening of a helix bundle exchangeable apolipoprotein. *Proc. Natl. Acad. Sci.* *96*, 4366-4371.
- Navaza, J. (1994). AMORE - an automated molecular replacement program package. *Acta Cryst.* *A50*, 157-163.
- Nicholls, A. (1992). GRASP: Graphical representation and analysis of surface properties. (New York: Columbia University).
- Nucholls, G. H., Romer, L. H., and Burridge, K. (1992). Microinjection of antibodies against talin inhibits the spreading and migration of fibroblasts. *J. Cell Sci.* *102*, 753-762.
- O'Halloran, T., Molony, L., and Burridge, K. (1986). Purification and assay of vinculin, metavinculin and talin. *Meth. Enzym.* *134*, 69-77.

Otwinowski, Z. (1993). In *Data collection and processing*, L. Sawyer, N. Isaacs and S. Bailey, eds. (Daresbury, UK: SERC), pp. 59-62.

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

Pardee, J. D., and Spudich, J. A. (1982). Purification of muscle actin. *Methods Enzymol.* 85, 164-181.

Parsons, J. T. (1996). Integrin-mediated signalling: regulation by protein tyrosine kinases and small GTP-binding proteins. *Curr. Opin. Str. Biol.* 8, 146-152.

Peleg, Y., and Metzenberg, R. L. (1994). Analysis of the DNA-binding and dimerization activities of neurospora-crassa transcription factor NUC-1. *Mol. Cell. Biol.* 14, 7816-7826.

Pertz, O., Bozic, D., Koch, A. W., Fauser, C., Brancaccio, A., and Engel, J. (1999). A new crystal structure, Ca²⁺ dependence and mutational analysis reveal molecular details of E-cadherin homoassociation. *EMBO J.* 18, 1738-1747.

Pollard, T. D., Almo, S., Quirk, S., Vinson, V., and Lattman, E. E. (1994). Structure of actin-binding proteins - insights about function at atomic resolution. *Annu. Rev. Cell Biol.* 10, 207-249.

Price, G. J., Davison, M. D., Patel, B., Eperon, I. C., and Critchley, D. R. (1987) G. J. Isolation and Characterization of a Vinculin cDNA From Chick-Embryo Fibroblasts. *Bioch. J.* 245, 595-603.

Price, G. J., Jones, P., Davison, M. D., Patel, B., Bendori, R., Geiger, B., and Critchley, D. R. (1989). Primary sequence and domain-structure of chicken vinculin. *Biochem. J.* 259, 453-461.

Puius, Y. A., Mahoney, N. M., and Almo, S. C. (1998). The modular structure of actin-regulatory proteins. *Curr. Opin. Cell Biol.* 10, 23-34.

Ramakrishnan, C., and Ramachandran, G. N. (1965). Stereochemical criteria for polypeptide and protein chain conformations. II. Allowed conformations for a pair of peptide units. *Biophys. J.* 5, 903-933.

Raussens, V., Fisher, C. A., Goormaghtigh, E., Ryan, R. O., and Ruyschaert, J. M. (1998). The low density lipoprotein receptor active conformation of apolipoprotein E - helix organization in N-terminal domain-phospholipid disc particles. *J.Biol.Chem.* 273, 25825-25830.

Read, R. J. (1986). Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Cryst.* A42, 140-149.

Read, R. J. (1990). Structure-factor probabilities for related structures. *Acta Cryst.* A46, 900-912.

Reszka, A. A., Hayashi, Y., and Horwitz, A. F. (1992). Identification of amino-acid-sequences in the integrin-beta-1 cytoplasmic domain implicated in cytoskeletal association. *J. Cell Biol.* 117, 1321-1330.

Ridley, A. J., and Hall, A. (1992). The small GTP-binding protein rho regulates the assembly of focal adhesions and actin stress fibers in response to growth-factors. *Cell* 70, 389-399.

Rozengurt, E. (1995). Convergent signaling in the action of integrins, neuropeptides, growth-factors and oncogenes. *Canc. Surv.* 24, 81-96.

Rudiger, M., Korneeva, N., Schwienbacher, C., Weiss, E. E., and Jockusch, B. M. (1998). Differential actin organization of vinculin isoforms: implications for cell type-specific microfilament anchorage. *FEBS Letts.* 431, 49-54.

Schwienbacher, C., Jockusch, B. M., and Rudiger, M. (1996). Intramolecular interactions regulate serine threonine phosphorylation of vinculin. *FEBS Letts.* 384, 71-74.

Sigal, C. T., Zhou, C. T., Buser, C. A., McLaughlin, S., and Resh, M. D. (1994). Amino-terminal basic residues of Src mediate membrane binding through electrostatic interaction with acidic phospholipids. *Proc.Natl.Acad.Sci.USA* 91, 12253-12257.

Smith, D. B., and Johnson, K. S. (1988). Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione. *Gene* 67, 31-40.

Springer, T. A. (1997). Folding of the N-terminal, ligand-binding region of integrin alpha-subunits into a beta-propeller domain. *Proc.Natl.Ac.Sci.USA* 94, 65-72.

Steimle, P. A., Hoffert, J. D., Adey, N. B., and Craig, S. W. (1999). Polyphosphoinositides inhibit the interaction of vinculin with actin filaments. *J.Biol.Chem.* 274, 18414-18420.

Steinberg, M. S. (1996). Adhesion in development: an historical overview. *Dev. Biol.* 180, 377-388.

Steinberg, M. S., and McNutt, P. M. (1999). Cadherins and their connections: adhesion junctions have broader functions. *Curr. Opin. Cell. Biol.* *11*, 554-560.

Suzuki, S. C., Inoue, T., Kimura, Y., Tanaka, T., and Takeichi, M. (1997). Neuronal circuits are subdivided by differential expression of type-II classic cadherins in postnatal mouse brains. *Mol. Cell. Neurosci.* *9*, 433-447.

Takada, Y., Ylanne, J., Mandelman, D., Puzon, W., and Ginsberg, M. H. (1992). A point mutation of integrin beta-1 subunit blocks binding of alpha-5-beta-1 to fibronectin and invasin but not recruitment to adhesion plaques. *J. Cell Biol.* *119*, 913-921.

Takeichi, M. (1995). Morphogenetic roles of classic cadherins. *Curr. Opin. Cell. Biol.* *7*, 619-627.

Tempel, B., Goldmann, W. H., Isenberg, G., and Sackmann, E. (1995). Interaction of the 47-Kda talin fragment and the 32-Kda vinculin fragment with acidic phospholipids - a computer analysis. *Bioph. J.* *69*, 228-241.

Terwilliger, T. C., and Eisenberg, D. (1987). Isomorphous replacement: effects of errors on the phase probability distribution. *Acta Cryst.* *A43*, 6-13.

Terwilliger, T. C., and Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallogr. D Biol. Crystallogr.* *55*, 849-861.

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

Turner, C. E., and Burridge, K. (1989). Detection of metavinculin in human-platelets using a modified talin overlay assay. *Eur. J. Cell. Biol.* *49*, 202-206.

VanNhiem, G. T., Krukonis, E. S., Reszka, A. A., Horwitz, A. F., and Isberg, R. R. (1996). Mutations in the cytoplasmic domain of the integrin beta(1) chain indicate a role for endocytosis factors in bacterial internalization. *J. Biol. Chem.* *271*, 7665-7672.

Vojtek, A. B., and Cooper, J. A. (1995). Rho-family members - activators of MAP kinase cascades. *Cell* *82*, 527-529.

Weekes, J., Barry, S. T., and Critchley, D. R. (1996). Acidic phospholipids inhibit the intramolecular association between the N- and C-terminal regions of vinculin, exposing actin-binding and protein kinase C phosphorylation sites. *Biochem. J.* *314*, 827-832.

Weers, P. M. M., Narayanaswami, V., Kay, C. M., and Ryan, R. O. (1999). Interaction of exchangeable apolipoprotein with phospholipid vesicles and lipoprotein particles. *J. Biol. Chem.* *274*, 21804-21810.

Weiss, E. E., Kroemker, M., Rudiger, A. H., Jockusch, B. M., and Rudiger, N. (1998). Vinculin is part of the cadherin-catenin junctional complex: Complex formation between alpha-catenin and vinculin. *J. Cell. Biol.* *141*, 755-764.

Weller, P. A., Ogryzko, E. P., Corben, E. B., Zhidkova, N. I., Patel, B., Price, G. J., Spurr, N. K., Koteliansky, V. E., and Critchley, D. R. (1990). Complete sequence of human vinculin and assignment of the gene to chromosome-10. *Proc. Natl. Acad. Sci. USA* *87*, 5667-5671.

Werth, D. K., Niedel, J. E., and Pastan, I. (1983). Vinculin, a cytoskeletal substrate of protein kinase C. *J. Biol. Chem.* 258, 11423-11426.

Wilkins J. A., Risinger, M. A., Coffey, E., and Lin, S. (1987). Purification of a vinculin binding protein from smooth muscle. *J. Cell Biol.* 104, A130.

Wilson, C., Wardell, M. R., Weisgraber, k. H., Mahley, R. W., and Agard, D. A. (1991). Three-dimensional structure of the LDL receptor-binding domain of human apolipoprotein E. *Science* 252, 1817-1822.

Wimley, W. C., and White, S. H. (1996). Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat.Struct.Biol.* 3, 842-848.

Winkler, J., Lunsdorf, H., and Jockusch, B. M. (1996). The ultrastructure of chicken gizzard vinculin as visualised by high-resolution electron microscopy. *J. Struct. Biol.* 116, 270-277.

Wood, C. K., Turner, C. E., Jackson, P., and Critchley, D. R. (1994). Characterisation of the Paxillin-Binding Site and the C-Terminal Focal Adhesion Targeting Sequence in Vinculin. *J. Cell Sci.* 107, 709-717.

Xu, W. M., Baribault, H., and Adamson, E. D. (1998). Vinculin knockout results in heart and brain defects during embryonic development. *Development* 125, 327-337.

Yamada, K. M., and Geiger, B. (1997). Molecular interactions in cell adhesion complexes. *Curr. Opin. Cell. Biol.* 9, 76-85.

Yap, A. S. (1998). The morphogenetic role of cadherin cell adhesion molecules in human cancer: a thematic review. *Cancer. Invest.* *16*, 252-261.

Ylanne, J., Chen, Y. L., O'Toole, T. E., Loftus, J. C., Takada, Y., and Ginsberg, M. H. (1993). Distinct functions of integrin α and β subunit cytoplasmic domains in cell spreading and formation of focal adhesions. *J. Cell. Biol.* *122*, 223-234.

PUBLICATIONS

Bakolitsa, C., de Pereda, J. M., Bagshaw, C. R., Critchley, D. R., and Liddington, R. C. (1999). Crystal structure of the vinculin tail suggests a pathway for activation. *Cell* 99, 603-613.