Biophysical Analysis of Tec Kinase Regulatory Regions: Implications for the Control of Kinase Activity.

by

Sharon Elizabeth Pursglove, B.Sc.(Hons)

A thesis submitted to the
University of Adelaide, South Australia
For the degree of
Doctor of Philosophy

May, 2001

Department of Biochemistry,
University of Adelaide,
South Australia

TABLE OF CONTENTS

ERRATUM

Figure 1.8 additional sentence "The orange arrows indicate the interaction surface on the protein domains."

Page 56 and 57 additional reference for the NMR experiments required: for review please see Evans 1995

Page 81 paragraph 3 line 9 should read "There is a resonance at 10.73 ppm which may be assigned to the indole of a Trp residue and is indicative of a folded protein."

Page 89 paragraph 3 line 3 should read "Strong NOEs are designated as being 2.0-2.5 Å apart."

Page 95 paragraph 2 line 3 additional sentence "Assignment of Tec SH3 domain was completed using both ^{1}H-^{1}H homonuclear and ^{1}H-^{15}N heteronuclear NMR spectra."

Additional References

Evans, J. N. S. **(1995)** Biomolecular NMR Spectroscopy. Oxford University Press, Oxford.

Kuboniwa, H., Grzesiek, S., Delaglio, F. and Bax A. **(1994)**. Measurement of HN-H alpha couplings in calcium free Calmodulin using new 2D and 3D water flip back methods. *Journal of Biomolecular NMR* **4** 871-878.

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. *Journal of Applied Crystallography* **26** 283-291.

Mulhern, T. D., Shaw, G. L., Morton, C. J., Day, A. J. and Campbell I. D. **(1997),** The SH2 domain from the tyrosine kinase Fyn in complex with a phosphotyrosyl peptide reveals insights into domain stability and binding specificity *Structure* **5** 1313-1323.

SUMMARY

The Tec family of intracellular tyrosine kinases includes Tec, Btk, Itk, Bmx, Txk and Dsrc29. There are four isofoms of Tec kinase (Tec I, Tec II, Tec III and Tec IV) generated by alternative splicing; Tec IV is the full-length transcript whereas the other major isoform, Tec III lacks 22 amino acids in the SH3 domain. These proteins contain the Src homology domains SH3, SH2 and the kinase domain (SH1) common to the Src kinases. In addition, Tec family members contain a Pleckstrin homology and Tec homology (PHTH) domain. The TH domain is a signature of this family of kinases and consists of a Btk motif followed by a proline rich region (PRR) where the Btk motif has a globular core that chelates a zinc atom by a HCCC motif and the PRR contains SH3 domain binding consensus sequences.

Many studies have shown that Tec kinase binds to and is activated by a variety of receptors in response to ligand binding. Tec kinases are crucial downstream targets of the T cell and B cell antigen receptors leading to activation and tyrosine phosphorylation of PLC-γ subunits. PH domains of Tec family proteins bind to the membrane through an interaction with $PI(3,4,5)P_3$, a product of PI3K activation. Following membrane localisation, Src family kinases phosphorylate Tec kinases within the kinase domain (Tyr 551 in Btk). Autophosphorylation of Tyr 223 in the SH3 domain follows. Activated Btk binds to a phosphotyrosine in SLP-65 resulting in the phosphorylation of PLC-γ2.

Protein tyrosine kinases are important regulators of cell differentiation and growth and comprise the largest group of known oncogenes. Src kinases are maintained in an inactive conformation through interactions between the SH2 domain and the C-terminal regulatory tyrosine and between the SH3 domain and the SH2-kinase linker. Tec kinases lack that regulatory C-terminal tyrosine and recent evidence suggests that Itk and Btk exhibit a novel mechanism of kinase activity regulation via interactions between the SH3 domain and the adjacent PRR. Itk PRRSH3 protein contains one SH3 binding consensus site (site 1) that contributes to a stable but weak intramolecular interaction. Btk, in comparison, has two consensus SH3 binding sequences within the PRR. Site 1 (^{186}KPLPPTP^{192}) can bind Src family kinases and the SH3 domain of Btk, while binding partners for site 2 (^{200}KPLPPEP^{206}) are currently unknown. Tec PRR also has two class I SH3 binding consensus sequences: site 1 ^{155}KTLPPAP^{161} and site 2 ^{165}KRRPPPPIP^{173}. There is currently no information regarding the affinity of these sites for the Tec SH3 domain.

The aim of this work was to derive NMR structures of Tec PHTH and SH3 domains and further investigate interactions between the SH3 domain and PRR of Tec kinase as

potential mechanisms of kinase activity regulation. Although a variety of different coding regions, expression systems and purification schemes were tested, structure determination of Tec PHTH domain was not successful. The solution structure of Tec SH3 domain was determined using two and three dimensional NMR techniques and the iterative approach of Nilges for NOE assignment. The Tec SH3 domain resembles a subclass of SH3 domains that includes Hck and Fyn tyrosine kinases. Each are composed of a six stranded, two β-sheet β-barrel with the second strand shared between the two sheets and a small C-terminal 3_{10} helix. Chemical shift perturbations that occurred upon ligand (SH2-kinase linker peptide like) binding were used to identify residues of the ligand binding site. This binding site is a shallow cleft on the surface of the protein and is consistent with those of other SH3 domains.

Tec PRRSH3 protein was generated as a GST fusion. Biophysical characterization, through analytical ultracentrifugation, BIAcore and NMR spectroscopy of wild-type and mutant Tec PRRSH3 proteins, was performed. Tec PRR site 1 formed an intramolecular interaction with the adjacent Tec SH3 domain. In contrast, PRR site 2 was involved in an intermolecular interaction with the SH3 domain of a neighbouring PRRSH3 molecule. Site 2 promoted the formation of PRRSH3 dimers and tetramers with a dimer dissociation constant of the order of 50 μM. In the wild-type PRRSH3 protein, more of the dimer/tetramer species existed than the intramolecularly associated monomer.

The observations described in this thesis indicate that the two SH3 domain binding consensus sequences within the PRR of Tec kinases perform distinct functions that are conserved between the different Tec family members. Tec PRR site 1 mediates intramolecular interactions and interactions with other cellular proteins whereas PRR site 2 is responsible for the dimerisation of the PRRSH3 region of Tec family proteins. Thus, Tec family members may be maintained in an inactive conformation through a mechanism involving the SH3 domain and the PRR of these proteins. This mechanism would differ depending on the presence or absence of SH3 consensus sites within the PRR.

STATEMENT

Statement of originality

This thesis contains no material which has been accepted for the award of any other degree or diploma in any University. To the best of my knowledge and belief, it contains no material that has previously been published by any other person except where due reference is made. The author consents to the thesis being made available for photocopying and loan.

.....

Sharon E. Pursglove

ACKNOWLEDGMENTS

Thankyou to Professor Peter Rathjen and Professor Barry Egan for the opportunity to complete a PhD in the Biochemistry Department, University of Adelaide.

A very BIG thankyou to Grant Booker for his support and enthusiasm when at times my own was lacking. Thanks to Grant also for the reading of this thesis and all the help he has given to get it completed, under circumstances that were not ideal. I couldn't have finished it without the help and I really appreciate it. I have enjoyed the work in the lab and I am a better scientist for being in the Booker lab. I am honoured to be the first PhD student through the Booker lab.

I'd also like to acknowledge Joel Mackay, Ray Norton and Mark Hinds for their collaborations during this PhD.

A special thanks to Terry Mulhern for everything, all the help he has given me over the years trying to introduce a total heathen to the wonders of NMR, Unix and Yowies. All the help and useful discussions were really appreciated.

Thanks to all the past members of the Booker lab that made it what it is today; Kate Stevens, Caroline Jones, Stephen Lawson, Richard Knowling, Andrew Bryce, Karen Heilbronn and Joe Wrin. It was great fun to work with you all. A special thankyou to Hayley Jackson and Joanna Doumanis for their help in cloning the SH3 protein and the linker peptide. Thanks also to Jo for her makeup tips and trendy clothes hints even when teaching me seemed hopeless.

Thankyou to the current extended Booker lab members Eric Bonython, Rebecca Bilton, Steven Inglis, Filomena Occhiodoro and Carlie Delaine. Thanks for helping out when I rang up the lab in a panic. Thankyou also to Steve for helping with the proof reading of this thesis.

Thankyou to Anne Chapman-Smith for introducing me to protein purification and Briony Forbes for her useful discussion and help with the BIAcore. Thankyou also to Keith Shearwin for the PHTH ultracentrifugation and answering my many kinetic questions. Thanks heaps to Leah Cosgrove for helping me out and giving me a job, the numerous lifts home and introducing me to the world of culturing cells en masse. Thankyou to Lynn Rogers and Tony Fratini for the student prac teaching. Thankyou to the CSU ladies and Jackie for all the media, the office ladies for all their help and BD and Serge for always being there with a helping hand.

Thanks to all the other PhD students in the Biochemistry department who made the department a great place to do a PhD. A special thanks to Steve Polyak, Melinda Lucic and Emma Parkinson who welcomed a Sydney girl into the department and made her feel at home. Thanks to Emma for those coffees, I couldn't have managed without them.

A very very special thanks to Ines Atmosukarto and Anita Merkel, who made the Booker lab a special place to work with lots of fun and lots of laughs, or cackling as Grant would say, and for agreeing to come and set up a hairdressing salon in Mildura with me. Thanks to Ines for the help throughout, introducing me to Indonesian cooking and for those Saturdays in the lab that invariably led to a 2-hour coffee break. I will miss them. Thanks heaps to Anita for always making me smile and especially for the running around she did at the end and the critical reading of this thesis. It is very much appreciated.

A big thankyou to Doreen, John, Jane, Dean Ryan, Nathan, Kev, Keita and Mark, for welcoming me into the family and making me a part of the Chapman clan. Thankyou to Doreen and John for helping us out and letting us take over their house for a few months.
A huge big thankyou to Tammie Leask, the best friend a girl can have. Thankyou for being there for me through this and for not being too cross when I don't ring or write for a while. Thanks to Rob Leighton for being the one to do the keeping in touch- I will start to return the calls now.

Thankyou heaps and heaps to Mum, Dad, Steve and Ben, for absolutely everything, especially your overwhelming support for everything I have ever wanted to do, even when you really didn't want me to do it. Thanks for always being there, even from Adelaide I knew you guys were always there for me- only a quick flight home. Thankyou for trying to understand the work when, really, it is all double-dutch. Your love and support got me through this. I couldn't have done it without you.

A huge big thanks to Gavin Chapman. Thankyou Gav for everything, for your critical reading of this thesis, for your understanding when I am grumpy, and, most of all, thankyou for being there with me through this and supporting me. I couldn't have done it without you.

DEDICATION

This thesis is dedicated to my grandparents

Raymond Pursglove,

Francis Mary and John William Bennett

And

Joseph Thomas and Emma Wake

I wish I had known them better

CHAPTER 1:

INTRODUCTION

1.1 SIGNAL TRANSDUCTION

Protein kinase cascades are an effective method for amplifying an extracellular signal that ultimately results in cell differentiation or cell proliferation. Transmembrane tyrosine kinase receptors contain extracellular regions of protein sequence that participate in ligand binding, a single span transmembrane region and intracellular regions of protein sequence. Signals are transmitted through extracellular ligands that, upon binding to the receptor, induce homodimerisation or heterodimerisation between receptor subunits and subsequent cross phosphorylation of multiple tyrosines within the intracellular catalytic portion of the protein. Various intracellular proteins then bind the phosphotyrosine residues including intracellular tyrosine kinases, phospholipase C -γ (PLC-γ) and phosphoinositol-3 kinase (PI3K) and transmit a signal through a series of protein-protein interactions to the nucleus (Karin, 1992); (Fantl *et al.*, 1993).

1.2 INTRACELLULAR PROTEIN TYROSINE KINASES

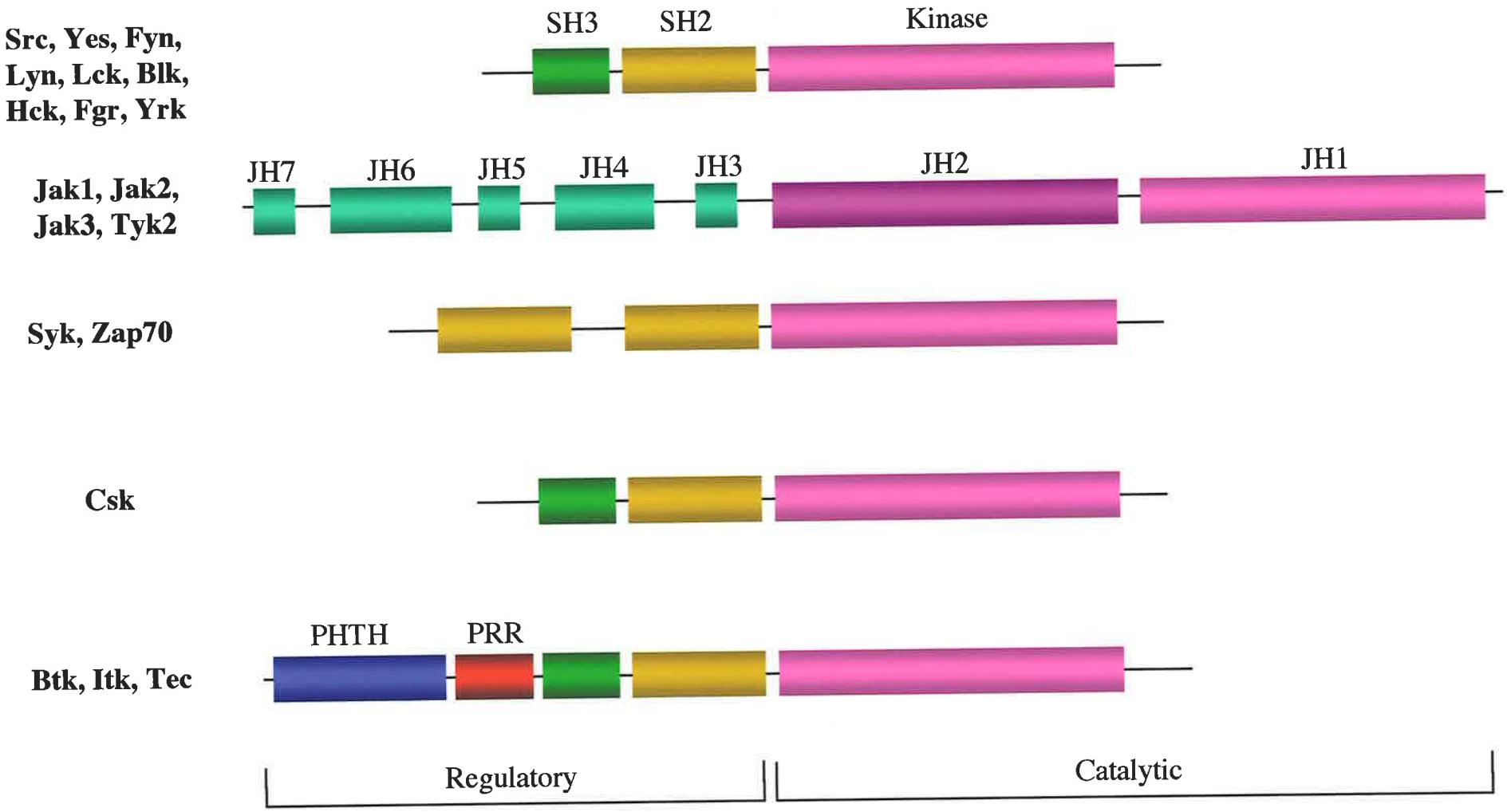
There are currently 26 members of the intracellular tyrosine kinase family separated into eight groups. All tyrosine kinases are modular proteins, consisting of a series of independently folding and independently acting domains that together modulate signal transduction pathways involved in transcription regulation, cell growth and apoptosis. All tyrosine kinases contain a hallmark catalytic or kinase domain. All, with the exception of the Jak and Fak kinases contain SH2 or SH3 domains with the majority containing both (Bolen, 1993). Figure 1.1 shows a schematic diagram of several families of tyrosine kinases with their modular domains highlighted.

1.3 THE SRC FAMILY OF INTRACELLULAR TYROSINE KINASES

The largest and most characterised group of tyrosine kinases is the Src family. Yes, Fgr, Lck and Src were isolated based on the ability to induce growth in the absence of growth factors in culture (Erpel and Courtneidge, 1995). The remaining Src family members, Fyn, Lyn, Blk, Hck and Yrk were isolated by homology screening (Erpel and Courtneidge, (1995) and references therein). Src family kinases contain a glycine at position 2 that is myristylated, and therefore anchor these kinases to the membrane. Src family kinases are predominantly responsible for the signal transduction immediately following receptor activation and membrane association is essential for their function (Toyoshima *et*

Figure 1.1

A schematic of the modular domains in Src, Jak, Syk, Csk and Tec families of intracellular tyrosine kinases. These proteins are composed of independently folding domains. The kinase domain (Src Homology 1 -SH1) is shown in pink. The regulatory domains include the SH2 domain, shown in orange, SH3 domain in green, Proline Rich Region (PRR) in red and the PHTH domain in blue. Jak kinases have several additional domains termed Jak Homology domains (JH) and these are presented in cyan. Although there are two kinase domains indicated for Jak kinase, only the most C-terminal kinase domain has catalytic activity (Heindrich, *et al.*, 1998). The non-functional kinase domain (JH2) is shown in dark pink.


Src, Yes, Fyn,
Lyn, Lck, Blk,
Hck, Fgr, Yrk
SH3
SH2
Kinase
Jak1, Jak2,
Jak3, Tyk2
JH7
JH6
JH5
JH4
JH3
JH2
JH1
Syk, Zap70
Csk
PHTH
PRR
Btk, Itk, Tec
Regulatory
Catalytic

al., 1992). The N-terminal regions are unique to each Src kinase and there is no evidence of a functional domain. Following the unique region is the Src Homology 3 (SH3) domain, Src Homology 2 (SH2) domain and the kinase domain. Src kinases contain a negative autoregulatory tyrosine at the C-terminus that is important in maintaining the kinases in an inactive conformation (section 1.6.1).

The Src family members Lck, Fgr, Hck and Blk are expressed solely in haematopoietic lineages in the adult while Src, Yes, Lyn and Fyn are expressed in a wide range of tissue types (Bolen, 1993). The functions of all Src family members have been investigated by generation of mice lacking these proteins. Lyn tyrosine kinase has been shown to physically associate with B cell antigen receptor through immunoreceptor tyrosine based activation motifs (ITAMs) and functions downstream of the receptor as mice lacking Lyn have reduced numbers of recirculating B cells (Chan *et al.*, 1997). *Src*$^{-/-}$ mice have a severely reduced survival rate following weaning, with surviving animals developing osteoporosis. This phenotype is more severe in mice that lack Src and Hck (*Src*$^{-/-}$/*Hck*$^{-/-}$) (Lowell *et al.*, 1996). Hck can therefore compensate for a lack of Src kinase during mouse development but presumably only in the haematopoietic lineages. *Src*$^{-/-}$/*Fyn*$^{-/-}$ and *Src*$^{-/-}$/*Yes*$^{-/-}$ mice die at birth (reviewed in Lowell and Soriano, (1996)). Yes kinase is expressed during neural development, however, Yes deficient mice exhibit no obvious phenotype. Fyn kinase is also expressed during neural development and mice lacking Fyn show defects in the hippocampus and consequent impaired long term memory. Lck protein has been shown to act downstream of the T cell receptor by gene targeting experiments in the mouse (reviewed in Lowell and Soriano, (1996)). Thus, there is an overall redundancy in Src family kinases that is facilitated by the presence of at least two Src family members in every cell type (Abram and Courtneidge, 2000).

1.4 THE TEC FAMILY OF INTRACELLULAR TYROSINE KINASES

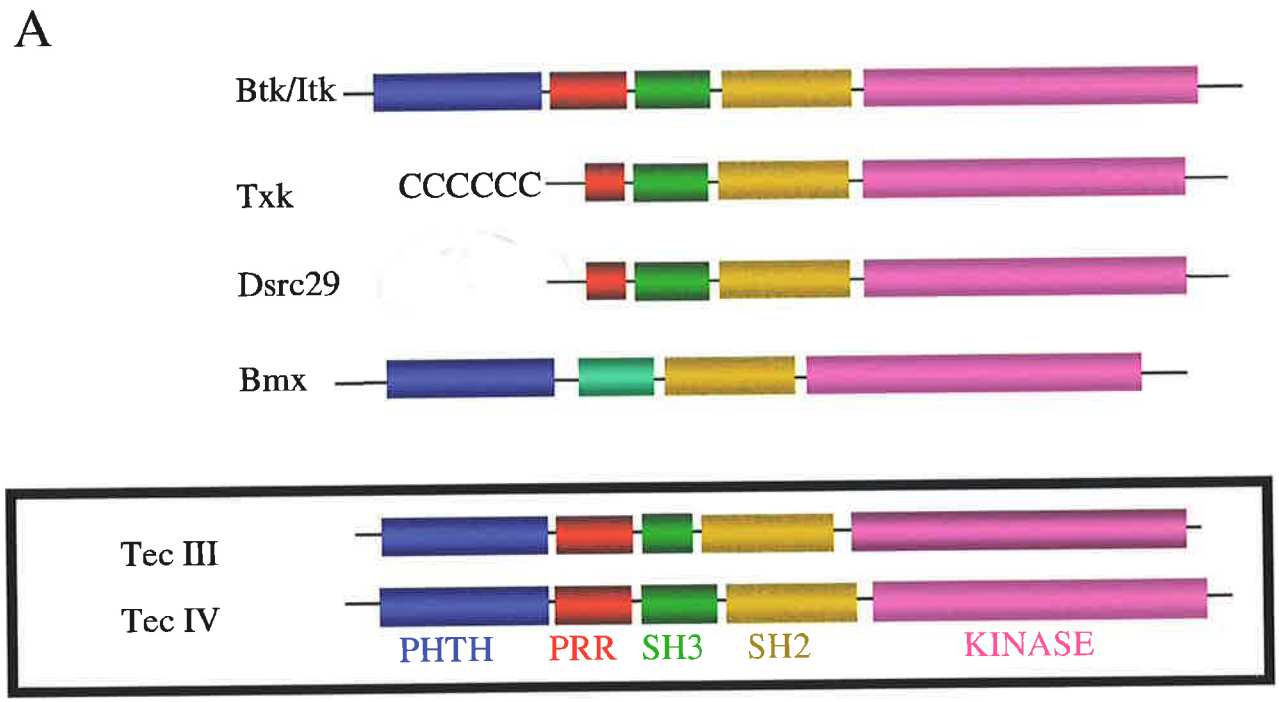
The Tec family of tyrosine kinases includes tyrosine kinase expressed in hepatocellular carcinoma (Tec), Bruton's tyrosine kinase (Btk), Interleukin-2 responsive T cell kinase (Itk; also called Emt), bone marrow tyrosine kinase gene in Chromosome X (Bmx; also called Etk), T cell expressed kinase (Txk; also called resting lymphocyte kinase Rlk) and Dsrc29. Figure 1.2A shows the modular structure of the Tec family of tyrosine kinases, which is the second largest family of tyrosine kinases. With the exception of Bmx which is also expressed in epithelial and endothelial cells, Tec kinases are predominantly expressed in haematopoietic lineages (Robinson *et al.*, 1996); (Tamagnone *et al.*, 1994);

Figure 1.2

A. Modular domains of Tec family kinases. The kinase domain is shown in pink, SH2 domain in orange, SH3 domain in green, PRR in red and PHTH domain in blue. Tec, Btk and Itk all have the same domain structure. Txk lacks the PHTH domain, which is replaced with a string of six cysteines. Bmx contains an atypical SH3 domain, which is represented in cyan. The two main isoforms of Tec kinase are shown in the boxed area.

B. Summary of currently known information about Tec family kinases. The expected size of the produced proteins, the tissue and cell lineage expression, chromosomal localisation, the activating pathways to which the proteins have been linked and phenotypes of single mutants as determined by gene targeting experiments in mice. The chromosomal locations of mouse Bmx and Itk are currently not determined (indicated with ?).
(Taken from Yang *et al.*, 2000)

A



B

Protein	Protein Size (kDa)	Tissue and Cell Lineage Expression	Chromosomal localization	Activating Pathway	Phenotype of Single Mutation
Tec29/dsrc29A/ Dsrc29C	66/55*	tissue: embryo, oocyte, and imaginal disc	polytene chromosome 28C	unknown	embryonic lethality and defective oogenesis
Bmx/Etk	80	tissue: bone marrow, lung, testis, colon, and heart cell: macrophage, neutrophil	Xp22.2 (human) ? (mouse)	IL-6R Ga12/13	unknown
Btk /Bpk/Atk/Emb	77	tissue: bone marrow, spleen, lymph node, and fetal liver cell: B, myeloid, erythroid, mast, and megakaryocyte	Xq22 (human) syntenic Xq22 (mouse)	BCR, FceRI, gp130 (IL-6R), IL-5R, IL-10R, CD19, CD28, CD38, CD40	block in B cell development, defective BCR signaling and XLA, defective degranulation
Itk/Tsk/Emt	72	tissue: thymus, spleen, and lymph node cell: T, NK, and mast cells	5q31-32 (human) ? (mouse)	TCR, CD28, CD2 and FCeRI	Reduction in CD4[1] subset and defect in TCR signaling
Rlk/Txk	52/58*	tissue: thymus, spleen, lymph node, tonsil, and testis cell: T and myeloid	4p12 (human) chromosome 5 (mouse)	TCR	subtle, some reduction of IL-2 production
Tec	72/70/66/58*	tissue: bone marrow, spleen, and thymus cell: T, B, myeloid, and hepatocarcinoma	4p12 (human) chromosome 5 (mouse)	TCR, CD28, BCR, CD38, gp130, (IL-3R, IL-6R), EpoR, G-CSFR, TpoR c-kit,	unknown

* Alternative splicing site of mRNA or alternative start site yields multiple isoforms.

(Weil *et al.*, 1997). The expression of Btk is restricted to B cells, while T cells express Itk and Txk. Tec kinase expression is more widely detected in myeloid lineages, T cells and B cells (Figure 1.2B) (reviewed in Schaeffer and Schwartzberg, (2000)).

There are four isoforms of Tec kinase (Tec I, Tec II, Tec III and Tec IV) generated by alternative splicing (Merkel *et al.*, 1999). Tec IV kinase is the full-length transcript whereas Tec III arises from the splicing out of the 66 bp exon 8 in the Tec cDNA resulting in the loss of 22 amino acids in the SH3 domain. The protein generated from the Tec III transcript is proposed to be a constitutively active isoform of the Tec protein. It is interesting to note that in some tissues the Tec III transcript is the major isoform expressed (Merkel *et al.*, 1999). Tec I protein arises from the use of an alternative exon 4 producing a 41 bp deletion in the PH domain causing a frameshift in the protein resulting in the production of a 100 amino acid protein. This protein has not been detected in a range of adult and embryonic tissues. Tec II protein is produced with alternate 21 amino acids in the C-terminus of the protein generated by the use of an alternate exon 18. The biological significance of these alternate transcripts is not clear (Merkel *et al.*, 1999).

Tec family proteins do not contain a regulatory tyrosine residue near the C-terminus nor a N-terminal myristylation sequence found in Src kinase, known to be important for regulation of the kinase activity and membrane localisation, respectively. Tec family proteins contain the Src homology domains SH3, SH2 and the kinase domain (SH1) common to the Src kinases (reviewed in Schaeffer and Schwartzberg, (2000)). Tec family SH2 and SH3 domains perform complex functions that result in both activation and repression of the proteins catalytic activity making them important regulators of these proteins. The kinase domain of Tec is structurally similar to Src and exhibits approximately 50% amino acid identity to that of other tyrosine kinases (47.8% to v-Abl and 47.4% to v-Src) (Mano *et al.*, 1990). Like other kinase domains, the Tec kinase domain contains the hallmark regions ^{383}HRDL^{386}, ^{403}DFG^{405}, ^{442}SDVWS^{446} in addition to an activating tyrosine position at 415 (equivalent to 416 v-Src) and an ATP binding site ^{273}X-G-XX-G^{278} and ^{284}Lys (Mano *et al.*, 1990). Like Src kinase, phosphorylation of tyrosine 415 (tyrosine 551 in Btk) in the loop of the Tec kinase domain results in stimulation of catalytic activity (Rawlings *et al.*, 1996).

In addition to the modular domains shared with Src kinases, Tec family members contain a Pleckstrin homology/Tec homology (PHTH) domain. PH domains are lipid-binding modules that participate in protein signalling by initiating membrane relocation of the protein. The TH domain is a signature of this family of kinases and consists of a Btk

motif followed by a proline rich region (PRR). The Btk motif is also present in the human interferon-γ binding protein (IGBP) and Ras GTP activating protein (Ras Gap) proteins and contains ten conserved amino acids of which seven amino acids are invariant (Vihinen *et al.*, 1994). The TH domain has a globular core that chelates a zinc atom by the coordination of three cysteines and a histidine (Cys 132, Cys 133 and Cys 143, and H121 in Btk) (Hyvonen and Saraste, 1997). The Btk motif is a novel protein fold unique to this family of tyrosine kinases.

1.4.1 The function of Tec family members

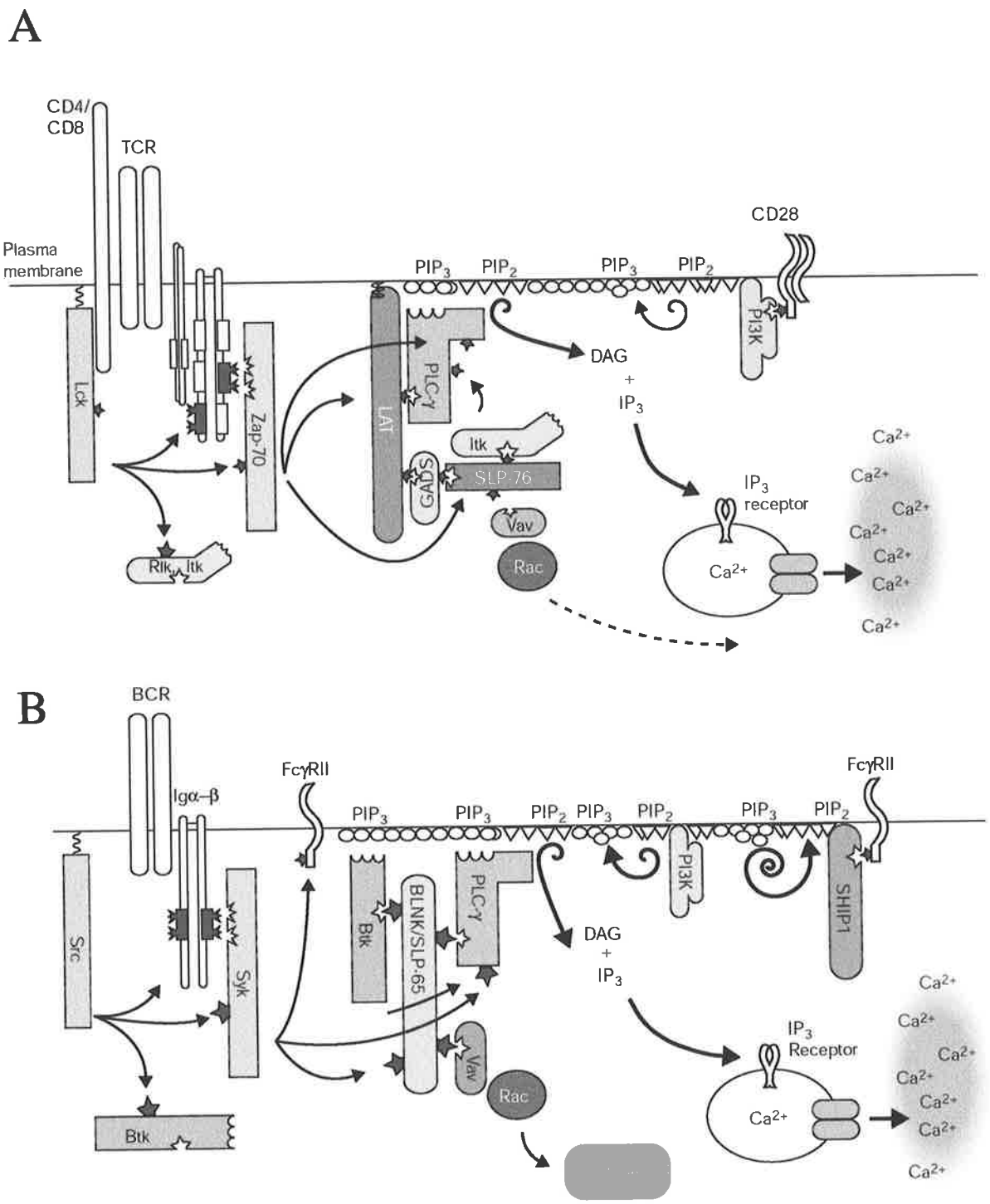
Btk, Itk/Txk and Tec are critical for downstream signalling from the B cell, T cell and Fcγ receptors, respectively (reviewed in Yang *et al.*, 2000) (Figure 1.3). X-linked agammaglobulinemia (XLA) is an inherited humoral disease that manifests as recurrent infections early in life (Rawlings and Witte, 1994). Bruton first described XLA in 1952 and in 1993 the gene was isolated and termed Bruton's tyrosine kinase or *Btk*. This was the first Src-related gene implicated in human genetic disease (Vetrie *et al.*, 1993). Patients suffering from XLA have severely reduced B cell levels and decreased levels of circulating immunoglobulins but exhibit normal myeloid and T cell function (Tsukada *et al.*, 1994); (Tsukada and Witte, 1994). Mutations that cause XLA encode Btk proteins with altered structural and functional characteristics (Vihinen *et al.*, 1999). Mutations that result in XLA have been mapped and can occur in all domains of Btk (Vihinen *et al.*, 1997). A naturally occurring X-linked immunodeficiency (XID) in mouse has also been localised to the *Btk* gene; however, the phenotype is somewhat less severe (Rawlings *et al.*, 1993). There is no evidence of similar hereditary disease caused by the other Tec family members.

Tec family members interact with many different binding partners. The unique PHTH region in Tec kinases was found to bind intracellular proteins including the guanine nucleotide exchange factor Vav, Protein kinase C and the βγ subunits of G-proteins (section 3.1.1) (Machide *et al.*, 1995); (Yao *et al.*, 1994); (Touhara *et al.*, 1994). The PH domains of Tec family members was found to bind the $PI(3,4,5)P_3$ group of phosphoinositols and it is through this interaction that these proteins become membrane bound (section 3.1.1). Binding partners for Tec family SH3 domains have also been isolated and these include Vav, Src-associated in mitosis-68 kDa (Sam68), SH3-domain binding protein that preferentially associates with Btk (Sab) and Wiskott Aldrich associated protein (WASP) (section 4.1.1) (Guinamard *et al.*, 1997); (Yamadori *et al.*, 1999); (Bunnell *et al.*, 1996). Tec kinase binds

Figure 1.3

A. A schematic representation of the function of Tec family member Itk downstream of the T cell receptors CD4/CD8, TCR and CD28. Tec family members are relocated to the membrane following activation of PI3K and the subsequent production of $PI(3,4,5)P_3$, the ligand for Tec family PH domains. Here, Tec kinases are phosphorylated (shown by stars) by Src family members (Lck in T cells and Src in B cells) within the kinase domain resulting in activation of the tyrosine activity. Subsequent phosphorylation of the tyrosine in the SH3 domain leads to complete activation of Tec family members. Following engagement of the receptors, SLP family adaptor proteins (SLP-76 in T cells and Blnk/SLP-65 in B cells) are phosphorylated by Zap 70 (T cells) or Syk (B cells) protein tyrosine kinases. A complex is formed between these adaptor proteins and Lat, Gads, Vav and Rac. Tec kinases bind to the phosphotyrosine on the SLP adaptor proteins via their SH2 domains (shown by star indents) which facilitates phosphorylation of PLC-γ by Tec family kinases. Activation of PLC-γ initiates Ca^{2+} mobilisation and activation of gene transcription.

B. A schematic representation of the function of Tec family member Btk downstream of the B cell receptors BCR and FCγRII. Rac can also facilitate Ca^{2+} production via a pathway utilising PIP_5 kinase. Ship1 phosphatase can regulate the production of calcium through hydrolysis of PIP_3.
(Adapted from Schaeffer and Schwartzberg 2000)


A
CD4/
CD8
TCR
Plasma
membrane
CD28
PIP3
PIP2
PIP3
PIP2
PI3K
Lck
Zap-70
LAT
PLC-γ
DAG
+
IP3
Itk
SLP-76
GADS
Vav
Rac
Rlk
Itk
IP3
receptor
Ca2+
B
BCR
Igα–β
FcγRII
PIP3
PIP3
PIP2
PIP3
PIP2
PIP3
PIP2
FcγRII
PI3K
SHIP1
Src
Syk
Btk
BLNK/SLP-65
PLC-γ
DAG
+
IP3
Vav
Rac
IP3
Receptor
Ca2+
Btk

proline rich regions in the CD28 protein following TCR/CD3 or CD28 receptor activation (Yang *et al.*, 1999). This links Tec kinase family members downstream of a variety of cell surface receptors. Immunoprecipitation and kinase assays have shown that Tec kinase is activated in response to the activation of a variety of receptors including the cytokine receptors granulocyte colony stimulating factor (G-CSF), erythropoietin (Epo), thrombopoietin (TPO) and interleukins-3 and 6 (IL-3 and IL-6) and the antigen receptor for T cells (Machide *et al.*, 1995); (Laffargue *et al.*, 1997); (Mano *et al.*, 1995); (Matsuda *et al.*, 1995); (Takahashi-Tezuka *et al.*, 1997); (Yamashita *et al.*, 1997); (Yang *et al.*, 1999). There is also evidence for the activation of Tec downstream of the receptor tyrosine kinase c-Kit (Tang *et al.*, 1994). Several ligands for the kinase domain of Tec kinase have been identified by yeast-2-hybrid. These include Grb10/Grb1R, suppressor of cytokine signalling–1 (SOCS-1) and BCR downstream signalling 1 (BRDG1) (Mano *et al.*, 1998); (Ohya *et al.*, 1997); (Ohya *et al.*, 1999).

Recently, ligands for the SH2 domains of Tec family members were isolated. The association of Tec family SH2 domains with the Src homology 2 domain containing Leukocyte Protein (SLP) adaptors (SLP-65/BLNK/BASH and SLP-76 in B and T cells respectively) was shown in GST pull down experiments (Su *et al.*, 1999). In T cells, SLP-76 is phosphorylated following T cell activation by Zap70 or Syk tyrosine kinases and, thus, provides the binding sites for the guanine nucleotide exchange factor for the Rho/Rac family of GTPases, Vav, the adaptor protein Nck and Itk (reviewed in Myung *et al.*, (2000). The importance of SLP-76 in T cell signalling has been shown in SLP-76 deficient jurkat cells. These cells failed to induce PLC-γ1 phosphorylation, calcium influx, extracellular signal-regulated protein kinase (ERK) activation or the upregulation of IL-2 after T cell activation. Zap70 has been shown to phosphorylate SLP-76 in jurkat cells and does so in the acidic N-terminal region at YESP/YEPP sites. This results in calcium mobilisation and initiation of transcription (Bubeck Wardenburg *et al.*, 1996).

Tec kinases are crucial downstream of the T cell and B cell antigen receptors. Activation and tyrosine phosphorylation of PLC-γ (the PLC-γ1 subunit in T cells and PLC-γ2 in B cells) is required for hydrolysis of phosphatidylinositol 4, 5 bisphosphate (PIP_2) to inositol 1, 4, 5 triphosphate ($PI(3,4,5)P_3$), a major mediator of calcium mobilisation and to diacylglycerol (DAG), an activator of protein kinase C (reviewed in Schaeffer and Schwartzberg, (2000). Calcium is required by the cell to initiate transcriptional and growth responses such as proliferation and apoptosis (reviewed in Kurosaki and Tsukada, (2000). In T cells the Tec kinase proteins involved in the response to antigen include Itk, Txk and

Tec while in B-cells Btk and Tec are important (Sommers *et al.*, 1995); (Bunnell *et al.*, 2000); (Schneider *et al.*, 2000); (Yang *et al.*, 1999); (Fluckiger *et al.*, 1998); (Kitanaka *et al.*, 1998); Rawlings, 1999).

Early experiments hinted at an important role for Tec family members in B and T cells. An increase in the tyrosine kinase activity and the phosphorylation of Btk was observed following stimulation of B cells with an anti-immunoglobulin antibody (Hinshelwood *et al.*, 1995); (Aoki *et al.*, 1994) and DT40 cells deficient in Btk showed a decrease in the phosphorylation of PLC-γ2 in response to receptor stimulation resulting in loss of phosphatidylinositol hydrolysis and no Ca^{2+} mobilisation (Takata and Kurosaki, 1996). Functional PLC-γ2 is essential in B cell function (Wang *et al.*, 2000). Btk and Tec kinase have been shown to restore Ca^{2+} mobilisation through the activation of PLC-γ2 in the absence of Syk (Venkataraman *et al.*, 1998); (Fluckiger *et al.*, 1998) and tyrosine 551 (in Btk) is absolutely required for this activation (Kurosaki and Kurosaki, 1997).

The signal transduction pathway has now been elucidated (Figure 1.3). Following receptor activation PI (4,5) P_2 is phosphorylated by PI3K to produce the high affinity ligand for the Tec family PH domains. In the case of Txk, which lacks a PH domain, there is in its place, a palmitoylated cysteine string that facilitates membrane localisation (Schneider *et al.*, 2000). Increased PI (3,4,5)P_3 results in relocation of Tec family proteins to the membrane. This is the first step in the activation of Tec family kinases and may be regulated by SHP-1, an SH2 domain containing tyrosine phosphatase which removes phosphates from PI(3,4,5)P_3 causing the release of Btk from the membrane (Bolland *et al.*, 1998).

Membrane localisation of Tec kinases facilitates activation of these kinases by Src family kinases such as Lyn, which phosphorylate Tec kinases on the tyrosine in the kinase domain (551 in Btk) (Li *et al.*, 1997), the second step and the critical step in the activation of Tec kinases. Autophosphorylation of tyrosine 223 in the SH3 domain follows, resulting in an active protein. This model differs slightly for Itk that has been shown to be constitutively attached to the membrane. In order for Itk to be activated, Zap70 must also be activated and the transmembrane adaptor, linker for activated T cells (Lat) must be phosphorylated permitting a Lat-Itk association (Yang *et al.*, 2000). During the timeframe for the activation of Tec family members, Syk kinase has phosphorylated the adaptor protein SLP-65 in B cells and Zap 70 has phosphorylated SLP-76 in T cells. The phosphotyrosine on SLP-65/SLP-76 can then bind directly to, and partially activate, PLC-γ2/PLC-γ1. However, a phosphotyrosine within SLP-65/SLP-76 also provides a ligand for binding to the SH2 domain of Tec family members. Activated Btk binds to the phosphorylated tyrosine in SLP-

65 and places it in close proximity to PLC-γ2. Btk then phosphorylates PLC-γ2 resulting in the full activation of PLC-γ2, Ca^{2+} mobilisation and gene transcription. A similar pathway involving Zap 70 tyrosine kinase, SLP-76 adaptor protein and Itk results in the activation of PLC-γ1 (described in Figure 1.3) (Shan and Wange, 1999).

Targeted gene disruption of the Tec family members has revealed a degree of redundancy in the function of Tec kinases. In T cells, the most highly expressed Tec protein is Itk. Mice null for *Itk* show a reduction in mature thymocytes and a reduction in the responses following T cell receptor cross-linking (Liao and Littman, 1995). However, mice null for *Txk* exhibit no major defects in T cell functional responses. Northern blot analysis of lymphoid organs from $Txk^{-/-}$ mice revealed a slight upregulation of *Itk* expression suggesting a compensatory role of Itk for *Txk* function (Schaeffer *et al.*, 1999). Mice lacking Txk and Itk had severely compromised responses to T cell activation due to a loss of mature T cell function. This is the same phenotype observed in the $Itk^{-/-}$ mice. $Itk^{-/-}$ / $Txk^{-/-}$ mice resemble that of the SLP-76 knockout mice again suggesting that Tec kinases, Txk and Itk and SLP-76 are a part of a complex that activates PLC-γ. This gene targeting information places Txk downstream of Src family kinases, as a second stage of activation in the initiation of transcription of IL-2 (Schaeffer *et al.*, 1999). It is interesting to note that Tec kinase has also been linked to T cell signalling and may also perform an analogous role, however, the authors suggest the function of Itk and Tec might differ in T cells as IL-2 promoter activity differed in Jurkat cells transfected with both of these proteins (Yang *et al.*, 1999). A similar redundancy of Tec kinases has been observed in B cells. The knockout of Tec kinase produced mice that were viable and fertile and had no obvious affect on the immune system. The production of *Tec* and *Btk* null mice produced a phenotype of increased severity as compared with that of $Btk^{-/-}$ mice. The numbers of peripheral B cells were reduced compared with $Btk^{-/-}$ mice and a defect during B cell development was also observed. In mice lacking both *Tec* and *Btk*, very few pre B cells can develop into immature B cells. Myeloid cell function has not yet been investigated in the $Tec^{-/-}$ mice. Thus, Btk can compensate for Tec, however, Tec can only partially compensate for Btk (Ellmeier *et al.*, 2000). Overexpression of Tec family members Bmx and Txk can reconstitute Btk signalling in human DT40 B cells (Tomlinson *et al.*, 1999). Both of these proteins can reconstitute PLC-γ calcium mobilisation and restore the ERK MAPK signalling pathways in Btk deficient DT40 cells. Bmx but not Txk rescued the function of Btk in BCR-induced apoptosis. The ability to rescue the apoptosis phenotype was isolated to the PH domain of Tec family members, consistent with the inability of Txk that lacks a PH domain

(Tomlinson *et al.*, 1999). Thus, in both B cells and T cells Tec family members can compensate for each other to differing degrees. The gene disruption of the predominant Tec family member expressed in B cells and T cells (Btk and Itk respectively) produces a severe phenotype in comparison to the gene disruption of the secondary expressed Tec family member (Tec and Txk respectively).

The atypical Tec family members, Bmx and Txk differ in their modular structure compared with other family members. Bmx lacks a PRR and a consensus SH3 domain (Tamagnone *et al.*, 1994). The function of Bmx is unclear, however, there is strong evidence for a role in prostate cancer. Bmx is required for IL-6 induced differentiation of prostate epithelial cells. In response to IL-6 receptor complex activation, Bmx protein is phosphorylated activating its kinase activity (Qiu *et al.*, 1998). Like the other Tec family members, Bmx has been shown to function downstream of PI3K and kinase activation is dependent on PI3K (Qiu *et al.*, 1998). Bmx contains a PH domain and the dependence on PI3K activation is not surprising as PH domains have been shown to associate with PIP_3 (Tomlinson *et al.*, 1999). However, Bmx activation is likely to differ compared with that of other Tec family members as an E42K mutation within the PH domain of Bmx reduces the kinase activity while the corresponding mutation in Btk (E41K) increases kinase activity (Qiu *et al.*, 1998). Bmx has also been linked to increased phosphorylation of Signal Transducer and Activator of Transcription (Stat) proteins 1, 3, and 5 and subsequent DNA binding activity (Saharinen *et al.*, 1997).

1.5 MODULAR DOMAINS CRITICAL FOR SIGNAL TRANSDUCTION

1.5.1 Src Homology 2 domains

Proteins including Src family kinases, Raf, PI3K, PLC-γ1 and GTP activating protein (GAP) interact with the stimulated Platelet-Derived Growth Factor (PDGF) or Epidermal Growth Factor (EGF) receptors (reviewed in Koch *et al.*, 1991). Although the biological activities of these proteins differ they all share a common region of homology known as the Src homology 2 (SH2) domain (reviewed in Koch *et al.*, (1991).

SH2 domains are well documented as a protein-protein interaction module that binds to phosphotyrosine residues. The SH2 domain has been shown to be indispensable for binding to activated growth factor receptors (Moran *et al.*, 1990). Gap1 and PLC-γ1 which both contain two SH2 domains, may bind activated receptors via a cooperative method

(reviewed in (Koch *et al.*, (1991). Src family members also bind activated receptors through the SH2 domain (Kypta *et al.*, 1990).

SH2 domain interactions with cellular ligands have also been reported. Following activation, STAT proteins dimerise through reciprocal SH2 domain-phosphotyrosine interactions allowing entry into the nucleus and subsequent activation of transcription (Chen *et al.*, 1998).

Deletion of the N or C-terminal regions of the Src SH2 domain increase tyrosine phosphorylation of Src by 14-30 fold. This increase in the tyrosine phosphorylation correlates with an increase in transforming potential of these proteins (Seidel-Dugan *et al.*, 1992). The phenotypes produced from SH2 domain mutations led researchers to postulate that the SH2 domain maintains Src in an inactive conformation. Liu and co-workers were able to show an intrinsic affinity of the SH2 domain for the C-terminal regulatory tyrosine (section 1.6.1.) (Liu *et al.*, 1993)

The three dimensional structure of SH2 domains have been solved using proteins in solution and in crystals. The solution structure of the p85 regulatory subunit of PI3K (P85α) was one of the first structures of an SH2 domain published (Booker *et al.*, 1992). This structure can be described as a three-stranded antiparallel β-sheet with a kink at the end of the third strand leading into a two stranded antiparallel β-sheet. An additional β-sheet is present at the C-terminus that lies along the face of the central β-sheet. Two α-helices present on opposing faces of the SH2 domain pack on either side of the central β-sheet (Booker *et al.*, 1992). SH2 domain structures in complex with phosphotyrosine peptides have also been determined and the binding surface on the SH2 domain has been described as a two pronged plug (Pascal *et al.*, 1994). The binding surface is composed of two components, one accommodating the tyrosine phosphate moity and the other the +3 amino acid residue (Hensmann *et al.*, 1994). In some instances, the second pocket is a long groove along the surface of the protein that accommodates amino acids +1 to +6 from the phosphotyrosine (Pascal *et al.*, 1994). The amino acids flanking the phosphotyrosine confer binding specificity. These sequences can be broadly separated into two classes; those that bind a motif of pTyr-hydrophilic-hydrophilic-Ile/Pro (class I) and those that bind pTyr-hydrophobic-X-hydrophobic sequences (class II). SH2 domains that bind class I phosphotyrosine moieties contain an aromatic residue in the fifth position of β-strand D (βD5) and include members from the tyrosine kinase family. The class II binders are composed of SH2 domains of p85α, pTyr-phosphatases and PLC-γ and contain an Ile or Cys at the βD5 position (Songyang *et al.*, 1995). Thus, SH2 domains interact with

phosphotyrosine residues on the receptors and facilitate the transfer of signals from the cytoplasm to the nucleus.

1.5.2 Src Homology 3 domains

SH3 domains were first identified as a conserved sequence in the N-terminal region of Src family tyrosine kinases. Another conserved region in Src family kinases, the SH2 domain, is often found adjacent to SH3 domains (Pawson *et al.*, 1993). SH3 domains have since been identified in a large number of proteins with a wide array of functions. They are present in proteins involved in cytoskeletal rearrangement (spectrin), clathrin mediated endocytosis (amphiphysins) and intracellular signalling (Src and PLC-γ). SH3 domain containing proteins involved in intracellular signalling can be further categorised into the two largest and well studied of SH3 domain groups, those with enzymatic activity and adaptor proteins that lack enzymatic activity (Pawson *et al.*, 1993). Proteins with tyrosine kinase activity or guanine nucleotide exchange activity generally only contain one SH3 domain (Pawson *et al.*, 1993). In comparison, adaptor proteins commonly have more than one SH3 domain. For example, intersectin, an adaptor protein potentially involved in endocytosis, contains five SH3 domains (Yamabhai *et al.*, 1998). SH3 domains are structurally independent and are not found in fixed topological positions within proteins, implying an independent functionality (Morton and Campbell, 1994).

SH3 domains are often referred to as molecular 'Velcro' or molecular adhesives (Morton and Campbell, 1994). They bind to proteins containing hydrophobic and proline residues, to promote the formation of specific protein-protein interactions responsible for signal transduction (Morton and Campbell, 1994); (Parrini and Mayer, 1999). Earlier work suggested that SH3 domains of proteins such as PLC-γ and growth factor receptor bound-2 (Grb2) protein target proteins to cytoskeletal microfilaments and to membrane ruffles, respectively (Bar-Sagi *et al.*, 1993). SH3 domains have now been shown to regulate the kinase activity of Src family kinases through an allosteric mechanism. The crystal structure of the kinase domain, SH3 and SH2 domain of Src and Hck kinases have revealed that the SH3 domain forms an intramolecular interaction with the polypeptide sequence connecting the SH2 and the kinase domains (SH2-kinase linker) (Xu *et al.*, 1997); (Sicheri *et al.*, 1997). This interaction, along with an interaction between the SH2 domain and the C-terminal regulatory phosphotyrosine prevent full kinase activation of these proteins. Mutations in the SH3 domain of Src can affect the stability of the inactive protein conformation and result in

a constitutively active protein (Xu *et al.*, 1997). Regulation of the Abl tyrosine kinase is also modulated through the SH3 domain, as mutation or deletion of the SH3 domain activates the transforming ability of this proto-oncogene (Gosser *et al.*, 1995) (section 1.6.3).

SH3 domains are small (~50-75 amino acids) independently folding protein domains that are amenable to structural studies. SH3 domains contain four regions of similar protein sequence. The first of these regions contains an ALYDY motif that forms a component of the SH3 domain ligand pocket (Booker *et al.*, 1993). Although this motif is not fully conserved among all known SH3 domains, the amino acid properties are conserved (Figure 1.4A). The second region is the least similar but again conservation of the amino acids with similar properties is seen at most sites. Motif three of SH3 domains contains the only invariant amino acid (Trp 215 in full length Tec). In the case of Tec and Src family members, there are often two consecutive tryptophans at this position. Tryptophan 215 is crucial for the formation of the hydrophobic ligand binding surface of SH3 domains. The most C-terminal region contains the motif GYIPSNY that is present in many SH3 domains and incorporates the final β-strands of the protein (Musacchio *et al.*, 1992). Individual amino acids that are conserved include amino acids Ala 185, Tyr 187, Leu 197, Ile 207, Leu 208 and Gly 226 in the Tec SH3 sequence (Figure 1.4A). The aspartic acid residue present in region 1 is conserved and forms a salt bridge to the ligand upon binding (Tzeng *et al.*, 2000).

There are many published structures of SH3 domains ranging from the muscle protein nebulin, through to the adaptor proteins Grb2 and Gab. Of interest to this thesis are the published structures from the tyrosine kinase family of enzymes (see an example in Figure 1.4B). Three dimensional structures of SH3 domains from the tyrosine kinase family have been reported, including those of Src, Fyn, Hck, Abl and Btk (Morton *et al.*, 1996); (Yu *et al.*, 1992); (Horita *et al.*, 1998); (Gosser *et al.*, 1995); (Tzeng *et al.*, 2000). SH3 domain structures published as an isolated domain, bound to a proline-rich ligand, as SH3SH2 complexes (Lyn) and in the context of the whole protein (Grb2) have been reported (Eck *et al.*, 1994); (Maignan *et al.*, 1995); (Yuzawa *et al.*, 2001).

SH3 domains have a five or six stranded β-barrel topology comprising two β-sheets of three strands arranged antiparallel to each other, and a single turn of 3_{10} helix. The SH3 domain of Src consists of six strands forming two sheets with the first sheet having a strong right handed twist (Yu *et al.*, 1992). Other SH3 domains that have six strands include GAP (Yang *et al.*, 1994) and PLC-γ (Kohda *et al.*, 1993). SH3 domains with five β-strands include those of the p85 subunit of PI3K (Booker *et al.*, 1993), α-Spectrin (Musacchio *et al.*,

Figure 1.4

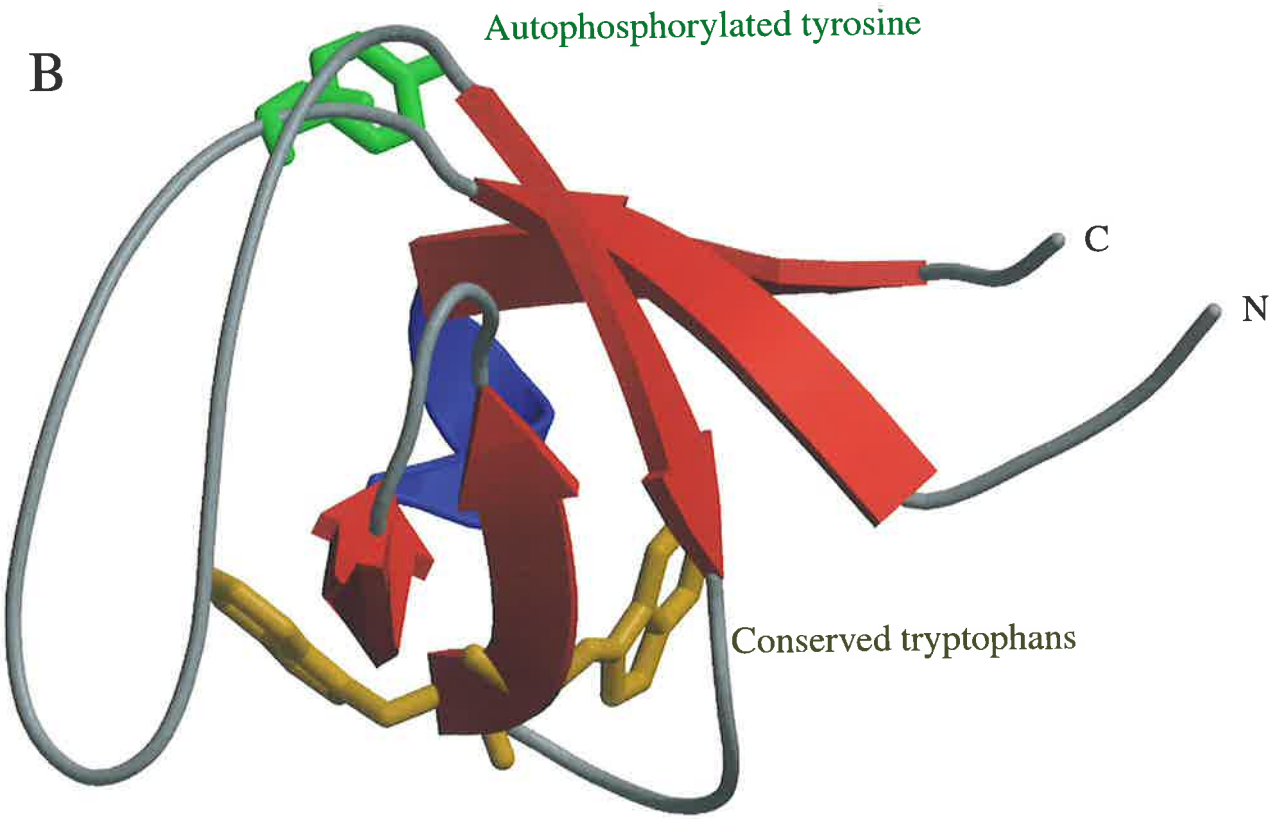
A. Alignment of SH3 domains from the Tec family members, Tec, Btk, Itk and Txk, Src family members Src, Fyn, Blk and Hck, Csk and the adaptor protein Grb2. Conserved amino acids are shown in blue (Ala 185, Leu 197 Trp 216, Gly 226 and Pro 229 in the Tec sequence), same amino acid functional group is red. The regions of similarity are boxed in black. The stars indicate conserved amino acids and : represents same amino acid functional group. The sequence alignment was conducted using a ClustalW algorithm (Thompson *et al.*, 1994).

B. Structure of the SH3 domain from Hck tyrosine kinase. Both tryptophans have been highlighted in yellow. The β-strands are shown in red and 3_{10} helix is blue (Horita *et al.*, 1998). The conserved autophosphorylated tyrosine is shown in green. The figure was produced using Molscript and Raster3D (Kraulis *et al* 1991); (Merritt *et al.*, 1997).
Hck SH3 domain PDB accession number is 5HCK.

A

```
mTec      178 NTEEIVVAMYDFQATEADHDLRLERGQEYIILEKNDLHWWRARDKY-GSEGYIPSNYVTGKKS 238
mBtk      214 TELKKVVALYDYMPMNANDLQLRKGEEYFILEESNLPWWRARDKN-GQEGYIPSNYITEAED 274
mItk      177 PEETLVIALYDYQTNDPQELALRCDEEYYLLDSSEIHWWRVQDKN-GHEGYAPSSYLVEKSP 237
mTxk       82 DERIQVKALYDFLPREPGNLALKRAEEYLILERCDPHWWKARDRF-GNEGLIPSNYVTENRL 142
mSrc       80 GGVTTFVALYDYESRTETDLSFKKGERLQIVNNTEGDWWLAHSLTTGQTGYIPSNYVAPSDS 141
mFyn       81 TGVTLFVALYDYEARTEDDLSFHKGEKFQILNSSEGDWWEARSLTTGETGYIPSNYVAPVDS 142
mBlk       51 EEERFVVALFDYAAVNDRDLQVLKGEKLQVLRST-GDWWLARSLVTGREGYVPSNFVAPVET 111
mCsk        9 PSGTECIAKYNFHGTAEQDLPFCKGDVLTIVAVTKDPNWYKAKNKVGREGIIPANYVQKREG  70
mHck       76 -EDTIVVALYDYEAIHREDLSFQKGDQMVVLEEA-GEWWKARSLATKKEGYIPSNYVARVNS 145
mGrb        1 ---MEAIAKYDFKATADDELSFKRGDILKVLNEECDQNWYKAELN-GKDGFIPKNYIEMKPH  58
consensus
```

B



1992), Fyn (Morton *et al.*, 1996), Amphiphysin (Owen *et al.*, 1998) and Lck (Hiroaki *et al.*, 1996). In several five stranded SH3 domains, including those of Fyn (Morton *et al.*, 1996) and α-Spectrin (Musacchio *et al.*, 1992), β-strand B is split between the 2 sheets. This group of SH3 domains contain a large loop between β-strands A and B, termed the RT-Src-loop, that is broken in Src by the additional β-strand (Hiroaki *et al.*, 1996). The SH3 domain of P85α contains a large loop, termed the N-Src-loop between strands B and C, which is smaller in other SH3 domains (Booker *et al.*, 1993). Most SH3 domains also include a 3_{10} helix at the C-terminal end of the SH3 domain structure, absent in the SH3 domains of Lck and PLC-γ (Hiroaki *et al.*, 1996). Despite these differences, SH3 domains retain a similar overall structure. Differences between SH3 domains may account for the specificity of binding to their distinct substrates.

The structure of the combined regulatory domains of the tyrosine kinase Lck has been determined by crystallography, and although there is very little contact between the adjacent SH3 and SH2 domains, there is an intermolecular interaction between two SH3SH2 molecules (Eck *et al.*, 1994). This structure indicates that two SH3SH2 molecules are arranged in a head to tail arrangement in which the SH3 domain binds to the SH2 domain of the other molecule. The solvent excluded surface area consists of the AB loop, the helical DE turn and the BC loop on the SH3 domain with the corresponding region on the SH2 being portions of the βD′, βE and βF strands. The authors regard this interaction as a crystal packing artefact, however, it is an interesting interaction that if it were possible could have biological implications for the binding of both the SH3 and SH2 domains (Eck *et al.*, 1994).

The structures of Src and Hck SH3 domains, solved in the intact enzyme lacking the N-terminal unique region, do not differ from those structures published as the SH3SH2 dimers in Abl and Lck, however, the relative orientation is very different (Xu *et al.*, 1997); (Sicheri *et al.*, 1997). Within the whole molecule, there are very few interactions between the SH3 and SH2 domains, except for residues 142-146 at the N-terminus of the SH2 domain of Src. Residues 142-146 form a 3_{10} helix that lies between two small clusters of hydrophobic amino acids on the surface of the Src SH3 domain. This interaction determines the relative orientation of the SH3SH2 domains and is dependent on the backbone orientation of the two amino acids adjacent to the 3_{10} helix, Asp 141 and Ser 142 (Xu *et al.*, 1997).

The adaptor molecule Grb2 mediates the Ras-signalling pathway through its interaction with Son of Sevenless (SOS) and is composed of two SH3 domains with an intervening SH2 domain. The crystal and the solution structure of Grb2 in its entirety have

been determined (Maignan *et al.*, 1995); (Yuzawa *et al.*, 2001). In contrast to the Src family kinase Lck, in the crystal structure there is no intramolecular contact between the SH3 domains and the SH2 domain although there is a shared surface between two SH3 domains. This intermolecular interaction is likely to be weak as this structure is maintained by only five hydrogen bonds and 1000Å^2 of the surface is buried (Maignan *et al.*, 1995). The ligand binding sites of both SH3 domains are exposed and predicted to still bind the proline rich peptide present in SOS. It is not known if this interaction occurs *in vivo* or if it has physiological function (Maignan *et al.*, 1995).

1.5.2.1 SH3 domain ligand interactions

Ligands capable of binding the SH3 domain of c-Abl tyrosine kinase, termed 3BP1 and 3BP2, were originally isolated by screening a cDNA expression library with the SH3 domain (Cicchetti *et al.*, 1992). Ren and co-workers, using an alanine screen, identified a ten amino acid hydrophobic PRR within 3BP1 and 3BP2 that was required for binding to the SH3 domain (Ren *et al.*, 1993). Further analysis identified three proline residues essential for SH3 domain binding; the consensus sequence derived from this was **XPXXPPP**ψ**XP** where X is any amino acid, ψ is a hydrophobic amino acid and the essential prolines are in bold. The SH3 domain of Src also binds 3BP1 and 3BP2, but at a reduced affinity, suggesting that there is another component influencing specificity of SH3 domain ligand binding (Ren *et al.*, 1993). NMR spectroscopy of these peptides in complex with the SH3 domain of Src revealed that the binding site is a hydrophobic depression on the domain's surface (Yu *et al.*, 1992). Structural experiments conducted with PI3K SH3 domain bound to a synthetic peptide (based on the GTP-hydrolysing microtubule-associated enzyme dynamin) confirmed the presence of a hydrophobic platform with charged residues at the periphery (Booker *et al.*, 1993). The authors propose that charged residues might impose specificity on the SH3-ligand interaction. The SH3 domain ligand has been minimised to a PXXP motif and the three dimensional structure of the complex of PI3K and RLP1 peptide (R_1-A_2-L_3-P_4-P_5-R_6-P_7) demonstrated that this motif forms a polyproline type 2 helix (Yu *et al.*, 1994).

There are two orientations of SH3 domain ligand binding due in part to the pseudo-symmetric nature of the polyproline type 2 helix (Saraste and Musacchio, 1994). Class I binding motifs contain a RXL motif and the established PXXP (where X is any amino acid) motif and bind in the same orientation as Src and PI3K SH3 domains with

affinities in the low micromolar range. These peptides take the form R_1-A_2-L_3-P_4-P_5-R_6-P_7 where Pro 4 to Ala and Pro 7 to Ala substitutions significantly reduce binding. Arginine in position 1 and Leu 3 substitutions were also shown to reduce binding (Yu *et al.*, 1994). Pro 4, Pro 7 and Arg 1 form one of the faces of the polyproline helix and the prolines sit in hydrophobic pockets on the surface of the SH3 domain (Feng *et al.*, 1994). Arginine 1 forms a salt bridge with the highly conserved aspartic acid (amino acid 21) in the variable loop in PI3K SH3 domain (Yu *et al.*, 1994). This salt bridge is proposed to be the major binding determinant for class I ligands (Feng *et al.*, 1994). The arginine in position 1 can be replaced with a lysine in some peptides and can still form the salt bridge (as in the case of Itk), however, in the *Caenorhabditis elegans* protein Sos the arginine to lysine substitution abolishes binding to Grb2, as does the corresponding mutagenesis of the carboxylic acid (Asp-Asn) (Andreotti *et al.*, 1997). Thus, the consensus for binding in the class I orientation has been minimised to +XXPXXP, where + is an amino acid with a positive charge (lysine or arginine), X is any amino acid and P is a proline.

Class II peptides have a consensus sequence of XXXPPLPPXR that binds in the reverse orientation where the arginine would still form a salt bridge with Asp 21 in Src kinase SH3 domain (Yu *et al.*, 1994). Amino acids required for class I binding are also critical for binding in a class II orientation. Substitution of the aspartic acid to asparagine in Src reduces the affinity of the interaction from 0.45 μM to 18 μM (Feng *et al.*, 1995). The class II binding sequence has been minimised to PXXPX+. It was initially suggested that these two ligand conformations might exist in equilibrium, however, the arginine-aspartic acid salt bridge is predicted to contribute a substantial portion of the binding energy, implying they probably do not exist in equilibrium, but rather one reaction will be favoured (Saraste and Musacchio, 1994). Table 1.1 summarises the ligand data for several SH3 domains.

The SH3 domain binding region is composed of three sections. The first is the arginine binding pocket that packs the legend's arginine residue against the indole ring of the conserved tryptophan in the SH3 domain. This interaction is strengthened by a salt bridge formed between the guanidinium group of arginine and the acidic side chain of the aspartic acid (amino acid 99 in Src) in the SH3 domain. This interaction also exists with class II ligand binding. The second and third interactions occur in binding pockets for the two essential prolines. The second pocket is composed of amino acids Tyr 92, Trp 118, Pro 133 and Tyr 136 in the Src sequence that bind proline 7, and the third pocket, made up by Tyr 90 and Tyr 136, binds proline 4 (Figure 1.5A) (Feng *et al.*, 1994). Van der Waal contacts form

Table 1.1 Overview of SH3 ligand sequences from a selection of different SH3 domains

SH3 domain protein	Binding protein	Binding class	Affinity	Sequence	Reference
Src		1	N/D	RPLPPLP	(Cheadle *et al.*, 1994)
					(Sparks *et al.*, 1995)
					(Sparks *et al.*, 1994)
		1	N/D	XXXRPLPPLPXP	(Rickles *et al.*, 1994)
	P85α	1	N/D	KPRPPRPLPV	
	Src SH2 kinase linker	2	N/D	^{250}KPQTQ^{253}	(Xu *et al.*, 1997)
	WASP*	1/2	N/D	^{307}RRQEPLPPPPPP^{322}	(Finan *et al.*, 1996)
			N/D	^{375}TGRSGPLPPPP^{388}	
			N/D	^{336}KGRSGPLPPVRL^{349}	
	3BP1	1	N/D	^{533}PPLPPLV^{538}	(Yu *et al.*, 1992)
			N/D	^{428}APTMPPPLPPGGK	
		1	8.0μM	RALPPLPRY	(Feng *et al.*, 1994)
		2	59μM	AFAPPLPRR	(Feng *et al.*, 1994)
		1	<10μm	RXLPPLPR	(Yu *et al.*, 1994)
P85α		1	N/D	RXLPPRPXX	(Yu *et al.*, 1994)
		1/2	N/D	RSLRPLPPLPPRPXX	(Rickles *et al.*, 1995)
		1	N/D	RXXRPLPPLPPP	(Rickles *et al.*, 1994)
	Dynamin	1	300μm	^{353}PAVPPARPGSRGPAP^{367}	(Booker *et al.*, 1993)
					(Gout *et al.*, 1993)
		1	<10μM	RXLPPRPXX	(Yu *et al.*, 1994)
Sem5/Grb 2	mSOS	2	N/D	PPPVPPRRR	
		2	54μM	^{1264}SPLLPKLPPKTYKRE^{1278}	(Wittekind *et al.*, 1994)
	hSOS	2	5-6μm	VPPPVPPRRR	(Goudreau *et al.*, 1994)
					(Terasawa *et al.*, 1994)
					(Lim *et al.*, 1994)
	3BP1	2	N/D	PPPXPPR	(Guruprasad *et al.*, 1995)
	Dynamin		N/D	^{353}PAVPPARPGSRGPAPG PPPAG^{372}	(Gout *et al.*, 1993)
Fyn		1	N/D	XXXRPLPP(I/L)PXX	(Rickles *et al.*, 1994)
	Btk		N/D	^{185}KKPLPPTPEE^{194}	(Cheng *et al.*, 1994)
		2	20μM	^{71}RPQVPLR^{77}	(Lee *et al.*, 1996)
					(Lee *et al.*, 1995)
Fyn R96I	PI-3K	1	50μM	PPRPLPVAPGSSKT	(Morton *et al.*, 1996)
	P85α	1	N/D	^{84}PPTPKPRPPPPLP^{96}	(Prasad *et al.*, 1993)
		2	N/D	^{303}PAPALPPKPPK^{P314}	
	3BP1	1	34μM	APTMPPPLPP	(Musacchio *et al.*, 1994)
Lyn		1/2	N/D	RXXRPLPPLPXP	(Rickles *et al.*, 1994)
	Btk	1	N/D	^{185}KKPLPPTPEE^{194}	(Cheng *et al.*, 1994))
Yes		2	N/D	GAAPPLPPRNPPR	(Rickles *et al.*, 1995)
Abl			N/D	PPPYPPPP(I/V)PXX	(Rickles *et al.*, 1994)
	3BP1	1	34μM	^{428}APTMPPPLPP^{437}	(Ren *et al.*, 1993)
					(Musacchio *et al.*, 1994)
			100X> 3BP1	APTYSPPPPP	(Pisabarro and Serrano, 1996)
Hck	Btk	1	N/D	^{185}KKPLPPTPEE^{194}	(Cheng *et al.*, 1994)
	HIV-1 NEF	2	0.3μM	^{71}TPQVPLR^{77}	(Grzesiek *et al.*, 1996)
	Ras Gap	2	10-50μM	^{129}GFPPLPPPPPQLPTLG^{144}	(Horita *et al.*, 1998)
		2	N/D	^{244}PCMSSKPQKP^{254}	(Sicheri *et al.*, 1997)
Amphiphysin-2	PRD of dynamin	2	N/D	^{400}PSRPNR^{405}	(Owen *et al.*, 1998)
	Synaptojanin	2	N/D	PSRPIP	(Owen *et al.*, 1998)

* Also been observed for binding to Fgr and PLC-γ SH3 domains. N/D affinity constant not determined.

between the prolines and the hydrophobic pocket on the SH3 domain. Class II ligands bind with less affinity due to a decrease in hydrophobic contacts and the fact that the prolines that bind each pocket are reversed (Feng *et al.*, 1994). Both class I and class II complexes have been studied by NMR and it is apparent that there are significantly fewer NOEs between the SH3 domain and the ligand in class II complexes than there are between class I complexes (Figure 1.5B) (Feng *et al.*, 1995).

There is one important exception to the observations described above. The sequence ^{249}KPQTQ^{253} found in the SH2 kinase linker of Src kinase can form an intramolecular interaction with the SH3 domain of Src. Although this sequence does not conform to either PXXP motif, within the context of the intact protein it does form a polyproline type 2 helix and interacts with the SH3 domain of Src (Xu *et al.*, 1997) and Hck tyrosine kinase (Sicheri *et al.*, 1997). Thus, while the two motifs described above fit the majority of cases, there are exceptions to the rule.

1.5.2.2 Signalling through SH3 domain

The Ras signalling pathway is one of the most studied pathways involving SH3 domain containing proteins. The SH3 domain in this pathway utilises protein-protein interactions to bring protein partners together at specific subcellular sites. The Ras signalling pathway ultimately results in the mitogenic control of cell growth, and in the case of *C.elegans*, vulval development and sex myoblast migration. The N-terminal SH3 domain present in Grb2 constitutively binds it's protein partner SOS thereby recruiting SOS to the activated EGFR complex at the membrane (Rozakis-Adcock *et al.*, 1993); (Egan *et al.*, 1993). SOS is a guanine-nucleotide releasing protein that activates guanine nucleotide exchange of Ras, promoting an interaction between Ras and the C-terminal region of the Raf1 serine/threonine kinase, facilitating downstream signalling. In this system, the SH3 domain is required to bind and localise SOS to a subcellular position that in turn facilitates an interaction between SOS and Ras and subsequent activation of Ras (Gale *et al.*, 1993).

Another well characterised SH3 domain interaction occurs between a variety of tyrosine kinases including Hck and Fyn and the HIV-1 protein, Nef. Contacts made outside the traditional SH3 ligand binding site in these interactions increase the strength of interaction. HIV-1 Nef is a non-catalytic protein critical for enhancing viral infectivity, replication in primary cells and the development of Acquired Immune Deficiency Syndrome (AIDS). HIV patients carrying virus with a defective *NEF* gene experience long term non-progressive HIV-1 infection. The interaction between the PRR of Nef and Hck is the

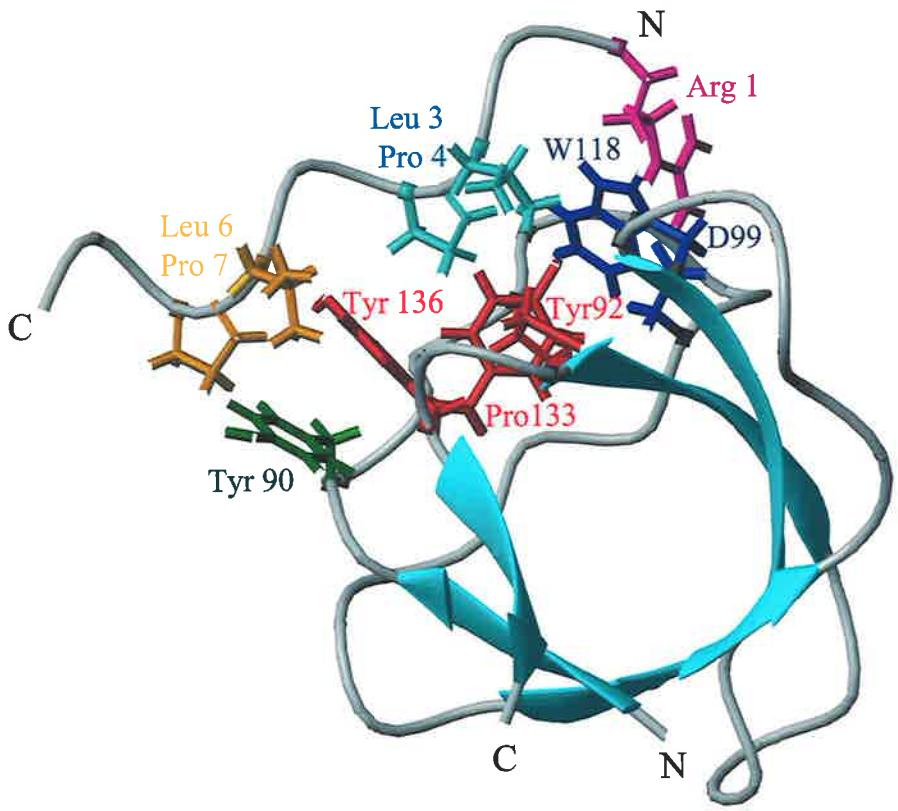
Figure 1.5

A. The structure of the Src SH3 domain bound to Class I ligand RALPPLPRY. The β-strands are shown in cyan. Arg 1 in the ligand is shown in pink and Asp 99 (shown in blue) and Trp 118 form the binding pocket. A salt bridge exists between the Asp 99 and Arg 1. Trp 118, Tyr 92, Pro 133 and Tyr 136 (shown in red) form the second pocket accommodating the first proline, Pro 4 (cyan). The second proline, Pro 7 (yellow) binding pocket is formed by Tyr 90 (shown in green) and Tyr 136 (Feng *et al.*, 1994).
PDB accession number for Src SH3 domain and a class I ligand is1RLQ.

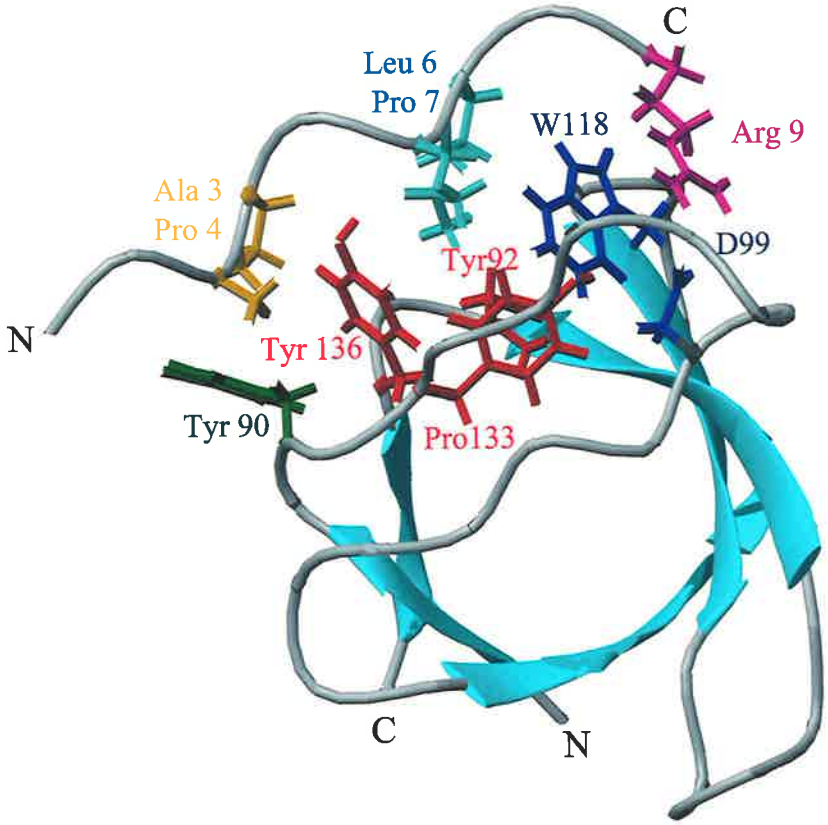
B. The structure of Src SH3 domain bound to Class II ligand AFAPPLPRR. The amino acids shown on the surface of the SH3 domain correspond to those shown in A. The amino acids participating in the interaction differ in the ligand, the two proline binding pockets bind Pro 4 and Pro 7 and Arg 9 forms a salt bridge to Asp 99 on the surface of the SH3 domain (Feng *et al.*, 1994).
PDB accession number for Src SH3 domain binding a class II ligand is 1PRM.

These figures were generated in MOLMOL (Koradi *et al.*, 1996).

A peptide RALPPLPRY



B peptide AFAPPLPRR



strongest observed for any SH3 domain with a K_D of 0.25 μM (Lee *et al.*, 1996). Hck may act downstream of Nef in HIV-infected macrophages (Briggs *et al.*, 1997). The Hck-Nef interaction increases Hck activation through an SH3 domain mediated process suggesting a novel method of regulation (Moarefi *et al.*, 1997).

1.5.3 Pleckstrin Homology domains

Pleckstrin Homology (PH) domains were first described in 1993 as conserved regions in Pleckstrin, the major Protein kinase C substrate in platelets (Haslam *et al.*, 1993); (Mayer *et al.*, 1993). This approximately 100 amino acid domain has since been found in many proteins of diverse function in vertebrates, *Drosophila melanogaster*, *C.elegans* and *Saccharomyces cerevisiae* (reviewed in Shaw, 1996). Identification of this domain was hindered by the low primary sequence identity shared between individual PH domains. New computer-based methods for identifying protein patterns or profiles were used to find previously unidentified PH domains (Musacchio *et al.*, 1993); (Shaw, 1993). PH domains have now been found in over 100 different proteins from a wide range of different subfamilies, including GTPases, serine/threonine kinases and the Tec family of tyrosine kinases (reviewed Shaw, 1996). The primary sequence is so divergent that some PH domains share as low as 4.2% identity (Musacchio *et al.*, 1993). Although there is only one invariant amino acid in all PH domains, corresponding to Trp 123 in the Btk sequence and Trp 102 in Tec, there are several regions that contain conservative amino acid substitutions and these map to β-strands of the PH domain structure (Figure 1.6A). PH domains are often present in proteins that also contain SH2, SH3 or Dbl-homology (DH) domains and these proteins can therefore form various protein-protein interactions. Such signalling proteins are capable of binding many different substrates via their different domains leading to activation of different signalling pathways.

While the primary structures of PH domains are not well conserved, their tertiary structures are very similar. Solution structures of PH domains from human Pleckstrin, mouse Spectrin, and human Dynamin (Macias *et al.*, 1994); (Yoon *et al.*, 1994); (Downing *et al.*, 1994); (Fushman *et al.*, 1995) and the crystal structure of the dynamin PH domain have been determined (Figure 1.6B) (Ferguson *et al.*, 1994); (Timm *et al.*, 1994). PH domains are modular domains and can retain their function even with large insertions. For example, the PH domain of Phospholipase C-γ (PLC-γ) contains an insertion of two SH2 domains and one SH3 domain between β strands and still retains its functionality

(Musacchio *et al.*, 1993). The structure of *Drosophila* β-Spectrin contains an α-helical insert between strands β3 and β4 (Zhang *et al.*, 1995). Three dimensional structures of PH domains determined thus far consist of a seven-stranded antiparallel β-barrel, which packs against a capping α-helix (Hyvonen and Saraste, 1997). The arrangement of the β-strands is such that strands 1, 2, 3 and 4 form one sheet and strands 5, 6 and 7 form the other sheet (Yoon *et al.*, 1994). In β-Spectrin however, β1-strand is shared between both sheets. The hydrophobic amino acids present in the capping α-helix point into the core of the protein and include the invariant tryptophan. This tryptophan appears to have a structural role as it makes contacts with many of the other hydrophobic amino acids that make up the core of the domain (Zhang *et al.*, 1995). The hydrophobic core of the PH domain from BTK kinase is composed of amino acids Leu 5, Leu 30, Leu 36, Tyr 38, Ile 45, Ile 60, Phe 101, Val 103, Leu 109, Val 111 and Trp 123 (Figure 1.6A) and interactions involving these amino acids create a strong bend in the PH domain structure (Macias *et al.*, 1994); (Downing *et al.*, 1994). In the N-terminal PH domain of Pleckstrin the hydrophobic core is capped by six lysines, four of which are conserved in most other PH domains (Yoon *et al.*, 1994). The three dimensional structure of Dynamin showed capping of the core by the presence of four tyrosine residues (Downing *et al.*, 1994), suggesting that polar or charged residues shield the hydrophobic PH domain core.

In more recent years, several structures of PH domains and adjacent regions have been determined such as the PH domain and Btk motif of Btk and the Phospho-Tyrosine Binding domain (PTB) and the PH domain of insulin receptor substrate (IRS1) (Hyvonen and Saraste, 1997); (Dhe-Paganon *et al.*, 1999). The PH domain structurally resembles the Enabled/VASP Homology (EVHI) domain and PTB domains (Prehoda *et al.*, 1999); (Eck *et al.*, 1996). EVHI domains bind to proline rich regions and are involved in targeting actin assembly machinery to sites of cytoskeletal remodelling while PTB domains bind phosphotyrosine residues and are found in adaptor proteins such as IRS-1 and Shc (Fedorov *et al.*, 1999); (Prehoda *et al.*, 1999); (Dhe-Paganon *et al.*, 1999). These three families of protein domains are now grouped into the PH domain superfamily. The ability of PH domains to bind different substrates suggests that the overall structure has some plasticity (Blomberg *et al.*, 1999).

The PH domain structure shows very strong polarisation. The more conserved face of the Dynamin PH domain, encompassing the α-helix, has a strong negative charge whereas the surface including the variable loops β1-β2, β3-β4 and β6-β7 is positively charged (Ferguson *et al.*, 1994). The structures of Pleckstrin and Spectrin PH domains show a

Figure 1.6

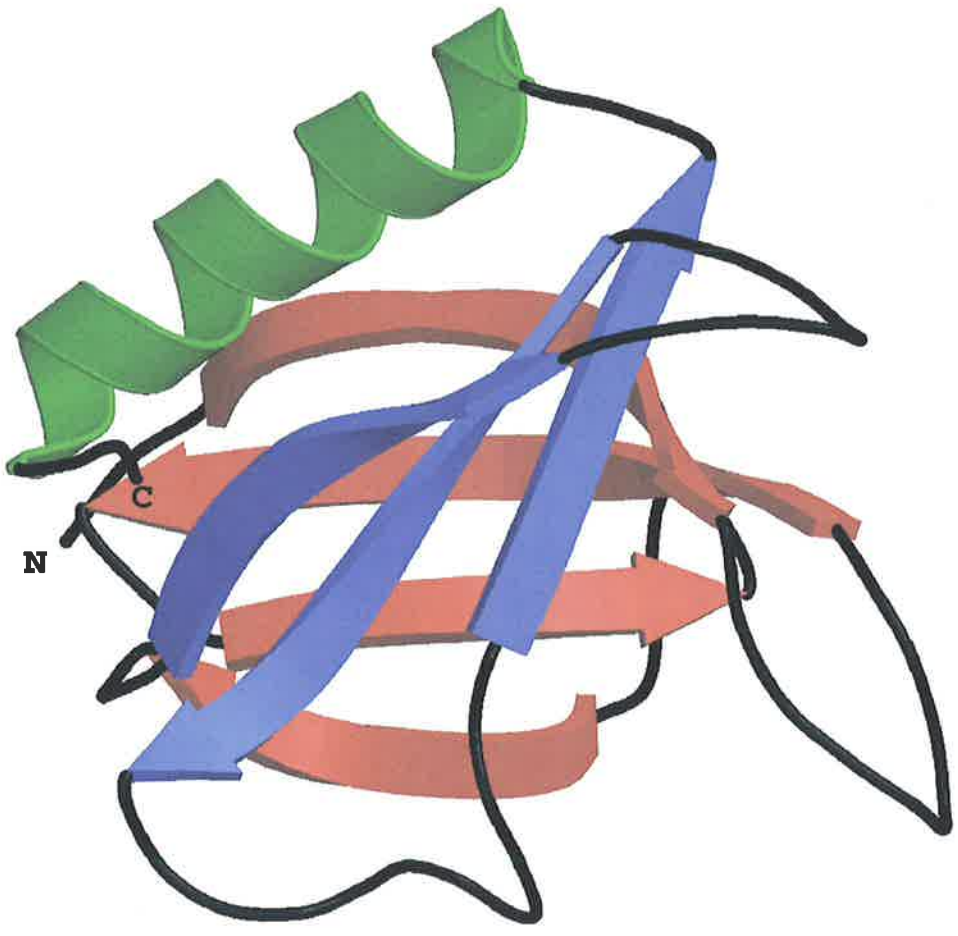
A. Alignment of PH domains from a series of proteins including the Tec family member Btk, SOS-1, PLC-δ, β-Spectrin, Dynamin and β-Ark. The only conserved amino acid in the PH domain is a conserved tryptophan represented in blue. Amino acids shown in blue represent conserved amino acids while those shown in red are members of the same amino acid group. The boxes highlight the regions of secondary structure. The sequence alignment was conducted using a ClustalW algorithm (Thompson *et al.*, 1994).

B. The structure of Dynamin PH domain. PH domains are composed of a β-barrel with a capping α-helix. The two β-sheets are shown in red (sheet 1) and blue (sheet 2) and the α-helix is shown in green. The N and C termini have been indicated (Fushman *et al.*, 1995); (Ferguson *et al.*, 1994). The figure was generated using InsightII (MSI).
Dynamin PH domain PDB accession number is 1DYN.

A

```
                    .    10        .    20        .    30        .    40        .    50
hBtk                2 ---------AAVILESIFLKRSQQKKTS----------PLNFKKRLFL  31
mSOS-1            415 GSKQLAIKKMNEIQKNIDGWEGKDIGQCCNEFIMEGTLTRVGAKHERHIF 464
rPLCδ-1             9 ------MHGLQDDPDLQALLKGSQLLKVK----------SSSWRRERFY  41
mβ-Spectrin       204 ---------MEGFLNRKHEWEAHNKKASS----------RSWHNVYCV 232
hDynamin          504 ------MKTSGNQDEILVIRKGWLTINNIG---------IMKGGSKEYWF 538
hβ-ark2           554 ..........GSHMGKDCIMHGYMSKMGN---------PFLTQWQRRYFY 584
                                        β1                        β2
                    .    60        .    70        .    80        .    90        .   100
hBtk               32 LTVH--KLSYYEYDFERGRRGSKKGSIDVEKITCVETVVPEKNPPPERQI  79
mSOS-1            465 LFDG--LMICCKSNHGQPR---L--PGASSAEYRLKEKFFMR-------- 499
rPLCδ-1            42 KLQEDCKTIWQESRKVMRSPESQ--LFSIEDIQEVRMG--HR--------  79
mβ-Spectrin       233 INNQ--EMGFYKDAKSAASG--I--PYHSEVPVSLKEAICEV-------- 268
hDynamin          539 VLTA-ENLSWYKDDEEKEKKYML--SVDNLKLRDVEKGFMSS-------- 577
hβ-ark2           585 LFPN--RLEWRGEGEAPQS------LLTMEEIQSVEET-.---------- 614
                          β3                  β4                  β5
                    .   110        .   120        .   130        .   140        .   150
hBtk               80 PRRGEESSEMEQISIIERFPYPFQVVYDEG--PLYVFSPTEELRKRWIHQ 127
mSOS-1            499 ------KVQINDKDDTSEYKHAFEIILKDG-NSVIFSAKSAEEKNNWMAA 542
rPLCδ-1            79 ------TEGLEKFARDIPEDRCFSIVFKDQRNTLDLIAPSPADAQHWVQG 123
mβ-Spectrin       268 ------ALDYKKK------KHVFKLRLSDG-NEYLFQAKDDEEMNTWIQA 305
hDynamin          577 ------KHIFALFN------TEQRNVYKDY-RQLELACETQEEVDSWKAS 614
hβ-ark2           615 -------QIKER--------KCLLLKIRGG-KQFILQCDSDPELVQWKKE 648
                                  β6                  β7                  α1
                    .   160        .   170        .   180        .   190
hBtk              128 LKNVIRYNSDLVQKYHPCFWIDGQYLCCSQTAKNAMGCQILEN 170
mSOS-1            543 LISLQYRS---------------------------------- 550
rPLCδ-1           124 LRKIIHHSGSMDQRQK------------------------------ 139
mβ-Spectrin       306 ISSA-------------------------------------- 309
hDynamin          615 FLRAGVYPERVGDK---------------------------- 628
hβ-ark2           649 LRDAYREAQQLVQRVPKMKNKPRS.----------------- 672
```

B



similar electrostatic potential (Ferguson *et al.*, 1994). The positive electrostatic potential near the variable loops led Ferguson and co-workers to propose that some PH domain ligands may be negatively charged.

The crystal structure of the PH-PTB region of IRS-1 has been determined and the PH region has a topology that is comparable to other PH domains. Juxtaposition of the PTB and PH domains does not alter the three dimensional structure of either domain compared with the isolated structures. The sheet comprising β5, β6 and β7 in the PH domain is packed against the back β1, β2, β3 and α2 of the PTB domain. There is, however, a shared surface between the two protein domains. The interaction is stabilised by hydrogen bonds between the two domains including those between PTB Phe160 - PH Asn71 and PTB Asp242 - PH Ser78/Asp77 and several potential salt bridges between PTB Glu162 - PH Arg75, and PH Lys 79, PTB Asp242 and PH Asp241 and PTB Glu200 - PH Arg89. Ligand binding to either domain does not affect the structure of the other domain (Dhe-Paganon *et al.*, 1999).

1.5.3.1 Interactions and signalling through PH domains

Proteins containing PH domains are key players in many signalling pathways particularly those that result in the rearrangement of the actin cytoskeleton but also those involved in glucose regulation. PH domains can confer their action through protein-protein interactions or protein-lipid interactions and consequently PH domains have many diverse roles. PH domains have been shown to participate in interactions between proteins and within proteins. Btk PH domain has been extensively analysed for its ability to bind to PKC, which has a major role in calcium regulation. cPKC, a mixture of α, β and γ isoforms, binds to the PH domain of Btk in a concentration dependent manner with a dissociation constant of 39 nm. The minimal binding region of the PH domain encompasses β–sheets 2 and 3 and this binds the c1 regulatory region of PKC (Yao *et al.*, 1997). The PKC binding site of BTK's PH domain can be projected onto the PH domains of Pleckstrin and PLC-γ, which shows there is overlap with the binding sites for phosphatidylinositol (4,5) bisphosphate ($PI(4,5)P_2$) and inositol (3,4,5) triphosphate ($I(3,4,5)P_3$). Furthermore, competition experiments suggest that $PI(4,5)P_2$, but not other phosphatidyl compounds such as phosphatidylserine, phosphatidylcholine, phosphatidylethanolamine, phosphatidylinositol or phosphatidylinositol 4 phosphate, compete with PKC for binding to the PH domain of Btk (Yao *et al.*, 1997).

There is evidence that PH domains can also interact intra-molecularly with other protein domains to modulate protein activity by an allosteric mechanism. Vav, a Dbl-family guanine nucleotide exchange factor and proto-oncogene, activates Rho family GTPases in response to PI3K ultimately resulting in actin cytoskeleton remodelling and gene expression changes. The PH domain of Vav is proposed to form a critical intramolecular interaction with the catalytic Dbl Homology (DH) domain that inhibits downstream signalling (Das *et al.*, 2000). This intramolecular interaction is promoted by the major PI3K substrate ($PI(4,5)P_2$) that binds the PH domain of Vav and promotes intramolecular contacts with the unphosphorylated DH domain. The PH-DH domain interaction prevents Rac, a Rho family GTPase, from binding the DH domain and, thus, prevents phosphorylation of Rac, and subsequent signalling. Conversely, products of PI3K such as $PI(3,4,5)P_3$ which bind to the PH domain of Vav with greater affinity than $PI(4,5)P_2$ prevent the inhibitory intramolecular interaction and promote the interaction of Rac and the DH domain. Thus, the PH domain of Vav can participate either in an inhibitory intramolecular interaction or in protein activation by interacting with PI3K substrates or products, respectively (Das *et al.*, 2000).

Extensive studies have also been conducted on the interaction of PH domains and βγ subunits of G-proteins. Studies have been conducted on PLC-γ, AKT (also known as protein kinase B), β-Spectrin, Insulin receptor substrate IRS-1 and β-adrenergic receptor kinase-1 (β–Ark) binding to the βγ subunit of G-proteins. The PH domain-βγ subunit interactions vary in binding affinity for the different protein partners. PLC-γ and AKT interact strongly with the βγ subunit while the interaction of the PH domain of β-Spectrin and IRS-1 with the βγ subunit is comparatively weak (Touhara *et al.*, 1994). β-Ark is involved in phosphorylating and desensitizing G-protein coupled receptors which influences G-protein coupled receptor downstream pathways. The β-Ark PH domain-βγ subunit interaction is required for relocation of β-Ark to the membrane resulting in increased phosphorylation of agonist-occupied G-protein coupled receptor substrates (Pitcher *et al.*, 1995). Localisation of β-Ark to the membrane is required for substrate phosphorylation. The C-terminal α-helix and adjacent C-terminal amino acids of the PH domain make up the G-protein βγ subunit binding surface. A mutagenesis screen revealed that the most deleterious mutants were the W643A and L647G changes (Touhara *et al.*, 1995), highlighting the importance of the conserved tryptophan in the C-terminal α-helix for the interaction with βγ subunits of G-proteins. The interaction of the PH domain of β–Ark with the βγ subunit inhibits the interaction between the βγ subunit and α subunit (Touhara *et al.*, 1994). However binding

to both PIP2 and βγ subunits of G-proteins is required to achieve effective membrane localisation and enhancement of the rate and extent of β–Ark substrate phosphorylation (Pitcher *et al.*, 1995).

A similar example of PH domain function in protein-protein and protein-lipid interactions involves the G-protein coupled receptor kinase (GRK) proteins. The GRK family regulate G-protein coupled receptor activity through phosphorylation, which in turn, promotes the activation and dissociation of heterotrimeric G-proteins and regulation of the various downstream pathways including internalisation. The PH domains of GRK2 and GRK3 can bind to negatively charged membranes and G-protein βγ-subunits and this is required for GRK2 function in cells. These two interactions in conjunction with affinity for the receptor itself may work cooperatively to recruit GRK2 to the membrane. Although inhibitions of either interaction results in dramatic loss of receptor phosphorylation, binding of Gβγ subunits is the limiting step for GRK2-dependent receptor phosphorylation (Carman *et al.*, 2000). Thus, the PH domain of GRK2 functions to recruit the kinase to the membrane via two cooperative interactions.

PH domain-lipid interactions lead to membrane localisation via the PH domain. Adaptor proteins such as B cell adaptor molecule of 32kDa (Bam32) and Grb2 associated binder 1 (Gab1) are relocated to the membrane upon protein activation due to PH domain interactions. Gab1 is a scaffolding protein that acts downstream of the Met receptor. The interaction of the PH domain of Gab1 and $PI(3,4,5)P_3$ in epithelial cells is crucial for subcellular localisation of Gab1 and efficient morphogenesis downstream of Met (Maroun *et al.*, 1999). Bam32 acts downstream of B cell receptor ligation where it is recruited to the membrane by binding 3′-phosphoinositols (Marshall *et al.*, 2000).

Several PH domains show clear preference for binding to particular phospholipid while others do not. PH domains exhibiting preference for $PI(3,4,5)P_3$ include those of Btk, T lymphoma invasion and metastasis protein (Tiam) and Son of Sevenless (SOS) (Rameh *et al.*, 1997). The PH domain of Dynamin exhibits preference for PI(4)P and β–Ark, PLC-γ and Pleckstrin PH domains bind $PI(4,5)P_2$ (Zheng *et al.*, 1996). The binding site also varies among different PH domains. The PI(4)P binding site of the Dynamin PH domain is a depression bounded by the first β-sheet (β strands 1-4) and the loop between β3 and β4 (Zheng *et al.*, 1996) while the binding site of $PI(4,5)P_2$ to the PH domain of Pleckstrin is in the N-terminal region of the protein near the lip of the β–barrel. Positively charged lysines may contribute to the binding of the Pleckstrin PH domain to negatively charged phosphoinositols (Harlan *et al.*, 1994). PLC-δ interacts with $Ins(1,4,5)P_3$ (the $PI(4,5)P_2$ head

group) through the positively charged surface made up of the β1β2 and β3β4 loops. Binding of PLC-γ to Ins(1,4,5)P_3 and Btk to Ins(1,3,4,5)P_4 involves similar molecular interactions suggesting that, although the ligands differ, the mechanism by which the PH domain binds is very similar. PH domains that show specificity for 3′-phosphorylated inositol derivatives may also be downstream targets of PI3K (Blomberg *et al.*, 1999).

1.6 REGULATION OF TYROSINE KINASE ACTIVITY

Protein tyrosine kinases are important regulators of cell differentiation and growth and comprise the largest group of known oncogenes (Rodrigues and Park, 1994). Overexpression of receptor tyrosine kinases can result in neoplastic transformation. The receptor is otherwise normal and may or may not associate with ligand indicating that overexpression is sufficient to increase the transforming ability of these proteins. This has been seen for epidermal growth factor receptor and colony stimulating factor 1 receptor. Increased kinase activation of Src family members has been associated with many different cancers, including those originating from Rous sarcoma virus (v-Src) and Moloney Murine Leukemia Virus (Lck) (Punt, 1992). Viral forms of Src family kinases possess deregulated tyrosine kinase activity that results in constitutive activation of the cellular pathways controlling cell growth and differentiation. A fusion protein of BCR and Abl kinase formed through translocation of the Abl sequence from its normal location on chromosome 9 to the breakpoint cluster region on chromosome 22 is oncogenic. This protein has intrinsic protein tyrosine kinase (PTK) activity and has a higher transforming ability than the wild-type protein (reviewed in Punt, 1992). Thus, tight regulation of the enzymatic activity of tyrosine kinases is critical for maintaining normal cellular function and preventing oncogenic activation.

1.6.1 Src family tyrosine kinases

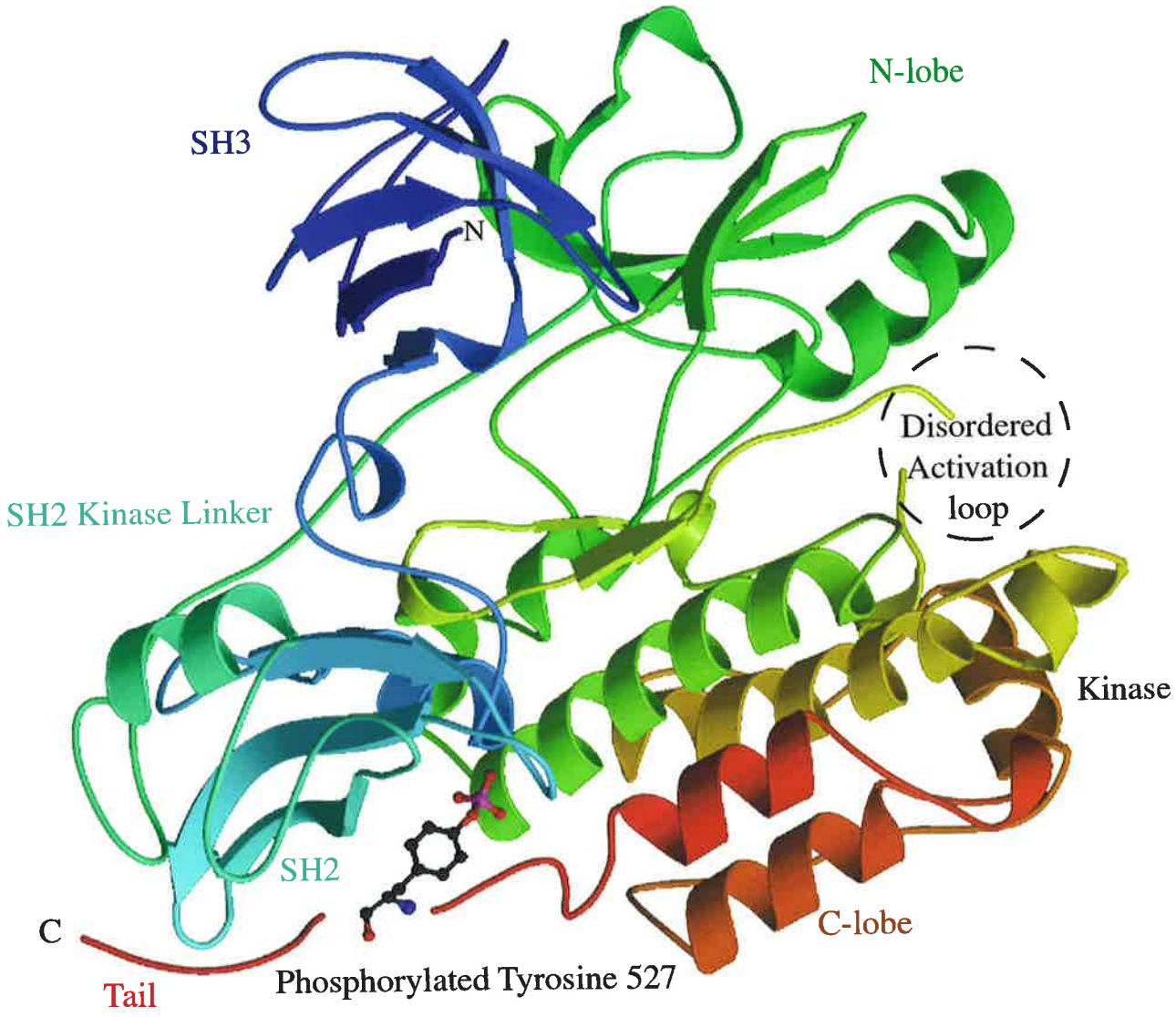
Src family tyrosine kinases are tightly regulated proteins. v-Src encoded by the Rous Sarcoma Virus (RSV) was first discovered by it's ability to augment cellular transformation. The discovery of an almost identical cellular counterpart that was unable to initiate cellular transformation prompted research into Src kinase. The major difference between the v-Src and cellular Src (c-Src) proteins is that the C-terminal 19 amino acids of c-Src have been replaced by 12 amino acids unique to v-Src (Takeya and Hanafusa, 1983). Early analysis determined that these proteins were tyrosine phosphorylated on different tyrosines. c-Src

was predominantly phosphorylated on Tyr 527 in the C-terminal tail whereas v-Src is phosphorylated at Tyr 416 in the kinase domain (Smart *et al.*, 1981); (Cooper *et al.*, 1986). The phosphorylation status of the C-terminus of c-Src was discovered to modulate the tyrosine kinase activity and transforming activity of Src kinase. Three papers published in 1987 showed that phosphorylated tyrosine 527 at the C-terminus of Src kinase suppresses the tyrosine kinase activity. In comparison, tyrosine 416 in the kinase domain of Src kinase enhances the transforming ability of Src kinase (Kmiecik and Shalloway, 1987); (Piwnica-Worms *et al.*, 1987); (Cartwright *et al.*, 1987). The SH3 and SH2 domains of c-Src were implicated in the maintenance of the inactive state of the kinase by deletion and mutagenesis studies. Mutations within the SH2 domain in the conserved FLVRES sequence resulted in an increase in transformation ability in both c-Src and also mutated Y527F Src suggesting a positive role for the SH2 domain in regulation of the inactive conformation (Hirai and Varmus, 1990). The discovery that the SH2 domain of Src kinase could in fact bind the phosphorylated tyrosine in the C-terminal tail confirmed the inhibitory role of the SH2 domain (Liu *et al.*, 1993). A deletion of the SH3 domain also results in constitutive activation of Src kinase, however, until recently, the mechanism by which this occurred was unknown (Seidel-Dugan *et al.*, 1992). The crystal structures of the SH3, SH2 and catalytic domains of the Src kinases Hck and Src shed light on the intramolecular interactions in the inactive kinases (Xu *et al.*, 1997); (Sicheri *et al.*, 1997). These papers confirmed an interaction between the SH2 domain and the regulatory tyrosine and also revealed an interaction between the SH3 domain and the polypeptide sequence between the SH2 domain and the kinase domain termed the SH2-kinase linker, that would participate in increasing the stability of this conformation (Figure 1.7). This compact domain organisation pushes the two lobes of the catalytic domain close together and promotes a conformational change in the small lobe that disables the active site. The SH2-kinase linker contains only one proline and yet forms a polyproline type 2 helix that interacts with SH3 domains (Xu *et al.*, 1997). The amino acids present in this region (^{246}TSKPQTQGLAKDA^{259}) do not match the consensus SH3 domain ligand (Xu *et al.*, 1997).

Following the publication of the structures of Src and Hck, the role of the SH2-kinase linker in this closed conformation has been further established. Interactions between the linker and the kinase domain have been shown to be crucial to maintain the inactive state. Leu 255 forms an interaction with the catalytic domain and mutation of this amino acid to valine, alanine or proline increased Src kinase activity whereas mutation to methionine repressed Src kinase activity. Thus, a hydrophobic interaction of a large amino

Figure 1.7

The crystal structure of the SH3, SH2 and the kinase domains of chicken Src kinase. The SH3 domain is represented in blue, the SH2 domain is shown in cyan. The kinase domain is shown in green and red. The intramolecular interactions between the SH3 domain and the SH2-kinase linker and the SH2 domain and the C-terminal tyrosine are shown. The structure surrounding the activation loop could not be determined and is highlighted by the dashed circle (Xu *et al.*, 1997). The regulatory tyrosine Tyr 527 is shown as a ball and stick model. This figure was adapted from the NCBI structure file. PDB accession number for the crystal structure of Src kinase is 1FMK.


N-lobe
SH3
N
Disordered
Activation
loop
SH2 Kinase Linker
Kinase
SH2
C
C-lobe
Tail
Phosphorylated Tyrosine 527

acid is required for the interaction with the kinase domain and maintenance of the closed conformation (Gonfloni *et al.*, 1999).

The activation of the Src family member Hck by SH3 domain displacement holds intriguing possibilities for the activation of all Src kinases. Moarefi and co-workers have shown a marked increase in the activation of Hck when the high affinity ligand Nef binds to the SH3 domain. The activation of Hck is more pronounced through this SH3 domain interaction than an interaction of the SH2 domain. Thus, the authors propose that activation of Src family kinases by high affinity ligands maybe a general mechanism for cellular activation of these kinases (Moarefi *et al.*, 1997).

There are other factors at play that affect the activity of the Src protein. The C-terminal tyrosine must be phosphorylated for the closed conformation to form. c-Src kinase (Csk) a tyrosine kinase that selectively phosphorylates tyrosine 527 is a negative regulator of Src activity. Mouse embryos lacking Csk arrest after 9-12 days post coitum and were shown to have increased tyrosine kinase activity of Src family kinases Src, Fyn and Lyn (Nada *et al.*, 1993). Therefore, in the absence of Csk, Src family kinases become derepressed and their tyrosine kinase activity is increased (Nada *et al.*, 1993). The SH3 and SH2 domains of Src were shown to be critical for the repression of kinase activation through a Csk dependent mechanism. c-Src SH2 domain mutants were insensitive to Csk action and could not rescue a fatal phenotype in *Saccharomyces-Pombe* (Superti-Furga *et al.*, 1993). Csk was, however, able to partially rescue the phenotype from SH3 domain mutants. Thus, this work confirms the importance of the SH2 and the SH3 domains in the maintenance of Src in an inactive conformation. The authors proposed that the SH2 domain interacts with the C-terminal tyrosine that is phosphorylated by Csk. The SH3 domain would stabilise this interaction (Superti-Furga *et al.*, 1993).

Other effectors for the tyrosine kinase activity of Src family members include the interplay between the phosphorylation and dephosphorylation at the C-terminal tyrosine. The phosphate is removed from Tyr 527 by Protein Tyrosine Phosphatase PTPα, which therefore controls Src kinase activity. PTPα, like Src, is widely expressed with the most abundant site of expression being the brain. Overexpression of PTPα has been shown to activate Src kinase and lead to neoplastic transformation through the dephosphorylation of Tyr 527 (Zheng *et al.*, 1992); (den Hertog *et al.*, 1993). There is evidence to suggest that Src may phosphorylate Tyr 789 in the C-terminus of PTPα, which in turn affects the binding of Grb2 to PTPα (Zheng *et al.*, 2000). A mutation within PTPα of Tyr 789 to Phe blocks transformation through an inability to dephosphorylate and activate Src kinase (Zheng *et al.*,

2000). Following phosphorylation of Tyr 789 within PTPα, by Src kinase, the SH2 domain of Src forms an interaction with this phosphotyrosine. The authors propose that partially activated Src phosphorylates Tyr 789 of PTPα, which promotes an interaction between Src's SH2 domain and PTPα phosphotyrosine 789 therefore removing the intramolecular interaction between Src SH2 domain and the regulatory tyrosine at the C-terminus. This positions Tyr 527 of Src in a location where the D1 catalytic subunit of PTPα can bind and dephosphorylate it leading to complete activation of Src (Zheng *et al.*, 2000).

Another possible method of regulation of Src family kinases includes allosteric effectors of tyrosine kinase activity. One such protein has been isolated and termed Src interacting (SIN) protein. This protein binds to the SH3 domain of Src and becomes tyrosine-phosphorylated. Phosphorylated SIN can now interact with the SH2 domain of Src rendering Src kinase active and suggests that SIN is both an activator and a substrate of Src kinase (Alexandropoulos and Baltimore, 1996).

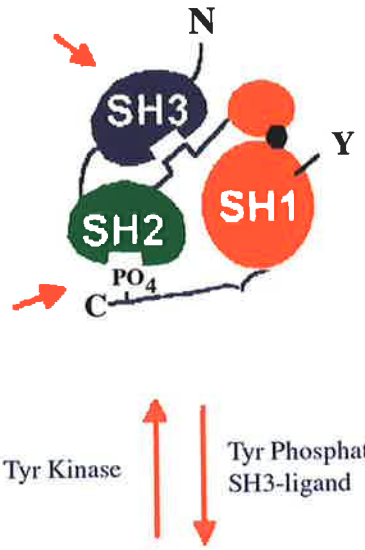
1.6.2 Tec family tyrosine kinases

Recent evidence suggests that Itk and Btk exhibit a novel mechanism of kinase activity regulation via intramolecular interactions between the SH3 domain and the adjacent PRR. Tec kinases do not contain a C-terminal regulatory tyrosine or a myristylation sequence at the N-terminus and it is therefore likely that their regulation is different to that of Src family members (Figure 1.8). Oncogenic activation of Tec family members occurs through several different mechanisms. Deletion of the SH3 domain (186-233) of Tec family kinases results in constitutive kinase activation (Yamashita *et al.*, 1996); (Zhu *et al.*, 1994) as determined by transient expression in 293 cells and anti phosphotyrosine immunoblotting.

Activation of wild-type Tec proteins requires phosphorylation of Tyr 551 in the activation loop of the kinase domain by a Src kinase family member and membrane association of the protein via the PH domain (Mahajan *et al.*, 1995); (Afar *et al.*, 1996); (Rawlings *et al.*, 1996). PI3K-γ increases the kinase activity of Btk and relocates Btk to the membrane by an interaction between the Btk PH domain and lipid products of PI3K activity (Li *et al.*, 1997). A constitutively active form of Btk has been isolated from a retroviral random mutagenesis screen, which contains a point mutation in the PH domain E41K (Li *et al.*, 1995). Over expression of Btk with the E41K mutation in the PH domain *in vivo* results in a severe phenotype in which the development of follicular recirculating B cells is blocked (Dingjan *et al.*, 1998). The increase in transforming ability and phosphorylation of the

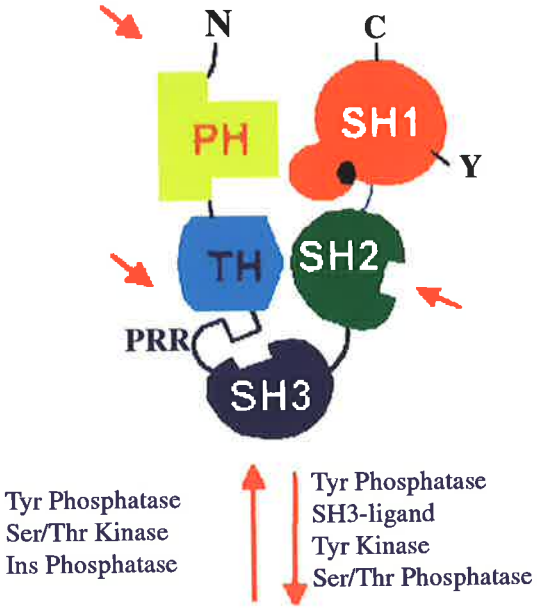
Figure 1.8

A schematic diagram of the possible regulatory mechanism of Tec family kinases. One method represents the mechanism analogous to that of Src family members. The other method could be an interaction between the SH3 domain and the PRR of Tec that would maintain these kinases in a closed conformation. The kinase domain is shown in orange with both lobes highlighted, the SH2 domain in green, the SH3 domain in dark blue, TH in light blue and the PH domain in yellow.
Taken from Yang *et al* (2000).



ACTIVATION

Src family



ACTIVATION

Tec family

E41K form of Btk has been shown to correlate to an increase in the membrane localisation of the Btk protein (Varnai *et al.*, 1999). Structurally, the phenotype observed with the E41K mutant can be explained due to the ability to bind two molecules of $PI(3,4,5)P_3$ suggesting a possible mechanism for the enhanced membrane localisation (Baraldi *et al.*, 1999). The substituted lysine residue contacts the second molecule of $Ins(1,3,4,5)P_4$ through an exposed surface and makes minimal contacts with the sidechains of Lys 26, Lys 41 and Lys 53 (Baraldi *et al.*, 1999). The E41K mutated Btk PH domain binds to $Ins(1,3,4,5)P_4$ at a two fold lower affinity compared with the wild-type.

Coexpression of Btk with Lyn, Fyn, Blk, and Hck tyrosine kinases was shown to increase the tyrosine phosphorylation of Btk 5-10 fold, a level equivalent to that observed in response to cross-linking of FcεRI on mast cells, membrane IgM on B cells and IL-5 receptor on pro B cells (Mahajan *et al.*, 1995); (Rawlings *et al.*, 1996). Btk has been shown to directly associate with Fyn, Lyn and Hck through an interaction modulated by the TH region of Btk and the SH3 domain of the Src family proteins (Cheng *et al.*, 1994). No single protein domain in Lyn or Blk has been shown to be required for phosphorylation of Btk, although kinase activity, an intact ATP binding site and the myristylation site were required for Lyn activation of Btk (Mahajan *et al.*, 1995); (Rawlings *et al.*, 1996). The region in the kinase domain of Tec kinases surrounding Tyr 551 is similar to the activating tyrosine in Src's kinase domain. The SH3 and SH2 domains, but not the PH domain, are required for Btk activation by Lyn (Mahajan *et al.*, 1995). Rawlings and co-workers found that the SH3 and SH2 domains of Btk were not required for the activation by Src family members (Rawlings *et al.*, 1996). Overexpression of the Csk protein indirectly diminishes Src dependent activation of Btk through inhibition of Src family kinase activation (Afar *et al.*, 1996). Syk kinase was not shown to effect tyrosine kinas activation of Btk, indicating other tyrosine kinases could not substitute for Src family kinases (Mahajan *et al.*, 1995); (Rawlings *et al.*, 1996). Rawlings and co-workers suggested that, in response to Lyn mediated phosphorylation of Btk, a second autophosphorylation event occurs. Subsequently, tyrosine 223 within the SH3 domain was found to be autophosphorylated (Park *et al.*, 1996). Mutation of tyrosine 223 to a phenylalanine potentiates the transforming ability of the protein, however, this does not correlate to an increase in kinase activity. This implies an alteration in signal transduction mediated by the SH3 domain (Park *et al.*, 1996).

Recent evidence has suggested a novel mechanism for maintenance of the Tec family kinases in an inactive form incorporating the SH3 domain. Andreotti and co-workers showed a weak but stable intramolecular interaction between the PRR and the SH3 domain

of Itk (Andreotti *et al.*, 1997). The PRRSH3 region of Itk is monomeric and following structure determination was shown to form an intramolecular interaction (Figure 1.9). The internal ^{155}KPLPPTP^{161} ligand of Itk forms a left handed polyproline type II helix where the dipeptides ^{157}LP^{158} and ^{160}TP^{161} contact the SH3 domain pocket. NOEs from the ligand can be clearly seen and show the ligand interacts in a class I orientation. The prolines are critical for this interaction because if they are mutated to alanine a polyproline structure does not form and there is no interaction with the SH3 domain. A panel of Itk-GST fusion proteins were tested *in vitro* for their ability to bind cellular ligands of Itk's SH3 domain. Sam68 is a protein that interacts with the Itk SH3 domain (Bunnell *et al.*, 1996). The SH3 domain of Itk preferentially binds Sam68 over the adjacent intramolecular PRRSH3 interaction. The interaction between Sam68 and Itk's SH3 domain is reduced in Itk PRRSH3 proteins also containing the SH2 domain and the N-terminal region. Inhibition of the Sam68-Itk SH3 domain interaction was not observed when only the SH2 domain was present, however, was observed in the presence of the TH domain alone. This suggests the N-terminal region and the SH2 domain may stabilise the PRR-SH3 domain interaction. These data are consistent with the weak nature of the PRR SH3 intramolecular interaction and its inability to compete with high affinity ligands for binding to the SH3 domain (Andreotti *et al.*, 1997).

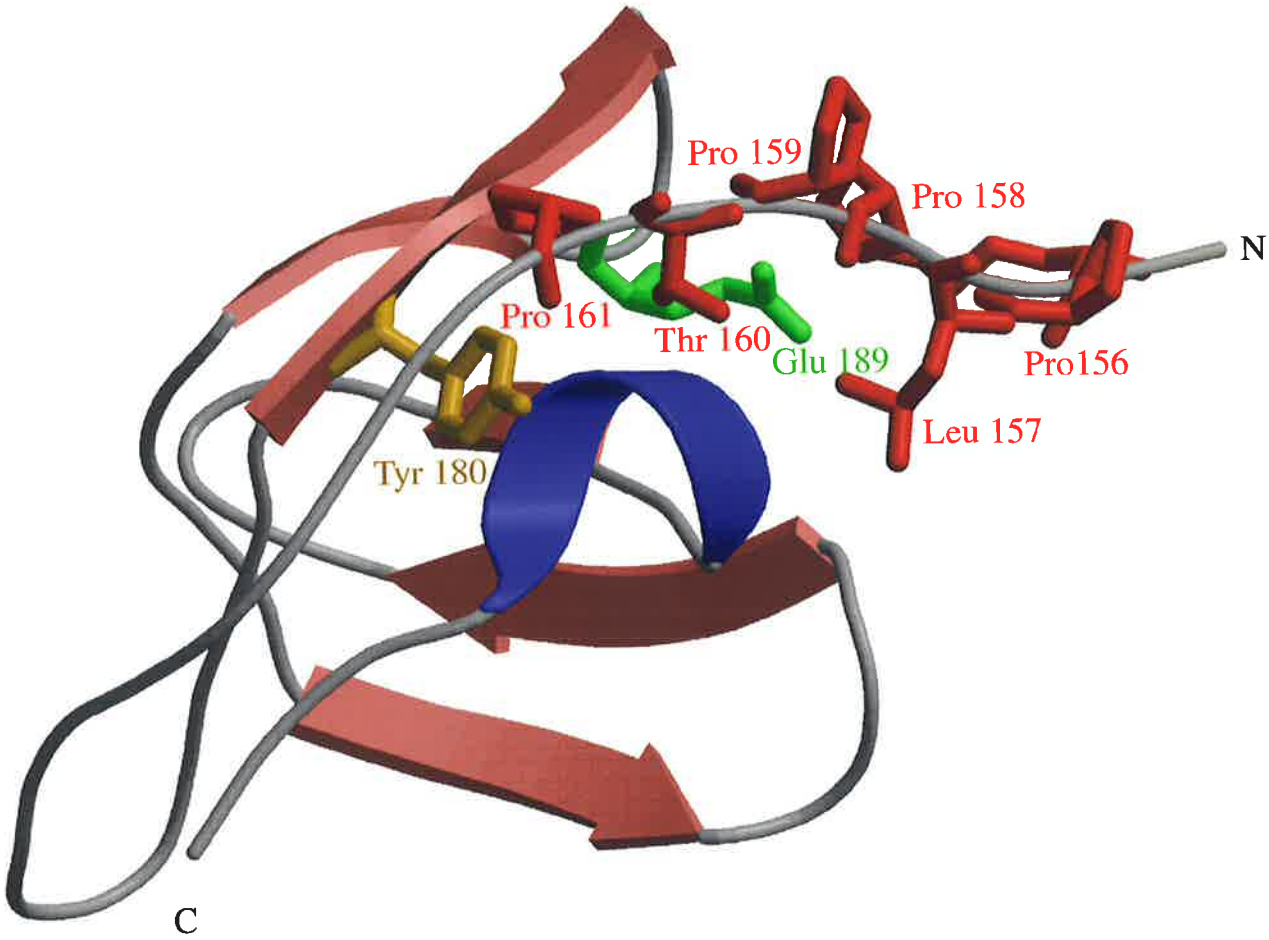
Btk has two consensus SH3 binding sequences within the PRR. The SH3 domain of Btk has been shown to bind a PRR of c-Cbl and the TH region (composed of the Btk motif and the PRR) can bind Lyn, Fyn and Hck SH3 domains (Cheng *et al.*, 1994); (Yang *et al.*, 1995). The first consensus sequences (^{186}KPLPPTP^{192}) can bind Src family kinases while binding partners for the second site (^{200}KPLPPEP^{206}) are currently unknown. ^{186}KPLPPTP^{192} has been shown by fluorescence experiments to bind the SH3 domain of Btk with a dissociation constant of 54.8 μM, which is comparable to the interaction of c-Cbl PRR and Btk SH3 domain (34.5 μM). Thus, the PRR of Btk could compete with the PRR of c-Cbl to bind to Btk's SH3 domain. It should be noted that this is an intermolecular interaction rather than an intramolecular interaction (Patel *et al.*, 1997).

A further interaction for Btk PRRSH3 domain has been observed. Hansson and co-workers (2001) observed the PRRSH3 region of Btk forms dimers with an equilibrium dissociation constant of 60 μM. A peptide corresponding to amino acids 181-192 of the Btk TH region was titrated into the PRRSH3 dimer solution and changes in the HSQC spectrum indicated that peptide addition caused dimer dissociation. The environment of the amino acids in the ligand binding site remain the same (Hansson *et al.*, 2001). Thus, the dimer interface is equivalent to the SH3 domain ligand binding site.

Figure 1.9

Structure of the Itk PRRSH3 domain. The PRR is shown in red, the β-strands are shown in pink and the 3_{10} helix is shown in blue. The side chains for amino acids Pro 156 to Pro161 in the PRR are shown in red. The autophosphorylated tyrosine (Tyr 180) is shown in yellow and conserved acidic amino acid Glu 189 is shown in green (Andreotti *et al.*, 1997). The figure was generated using Molscript and Raster3D (Kraulis, *et al* 1991); (Merritt *et al.*, 1997).

The PDB accession number for Itk PRRSH3 is 1AWJ.


Pro 159
Pro 158
N
Pro 161
Thr 160
Glu 189
Pro156
Leu 157
Tyr 180
C

The linker between the SH2 domain and the kinase domain of Tec contains the same characteristics as the equivalent region in Src kinase (Xu *et al.*, 1997); (Sicheri *et al.*, 1997). It is therefore possible that equivalent interactions could form between Tec SH3 domain and SH2-kinase linker as observed in Src. It is not known whether an SH2 kinase linker-SH3 domain interaction can maintain Tec in an inactive conformation.

The SH3 domain of Tec family members plays a critical role in both the activation and repression of the tyrosine kinase activity of this family of enzymes. When activated, the SH3 domain can maximise the tyrosine kinase activity. Evidence suggests that the SH3 domain can maintain the tyrosine kinase in the inactive conformation through an interaction with the adjacent PRR. The autophosphorylated tyrosine present in the SH3 domain is located within the ligand binding site of the SH3 domain making it possible that binding of the PRR to the SH3 domain would render the phosphorylated tyrosine inaccessible (Morrogh *et al.*, 1999). Btk activation is proposed to involve a series of events. Initially, Btk is present in the cytoplasm in an inactive conformation with possible intramolecular interactions maintaining Tec kinases in an inactive conformation. PI3K activation by Src kinases generates $PI(3,4,5)P_3$ and an interaction between Btk's PH domain and the lipid relocates Btk to the membrane where Src family kinases phosphorylate Btk on Tyr 551 (Nore *et al.*, 2000). This in turn precipitates the dimerisation of Btk proteins. Autophosphorylation then occurs on Tyr 223 in the SH3 domain resulting in an active enzyme. Btk protein would be released from the membrane by the dephosphorylation of $PI(3,4,5)P_3$ (Baraldi *et al.*, 1999).

1.6.3 Abl and Csk family

c-Abl is the cellular homologue of the v-Abl oncogene of the Abelson murine leukemia virus (A-MuLV) responsible for the transforming ability of the virus. cDNAs of c-Abl have been cloned from human, mouse, fruit fly and nematode and to date, one other family member, Arg, has been isolated from the human genome. There are two alternatively spliced transcripts of c-Abl both of which encode SH3, SH2 and kinase domains but different N-termini, one of which contains a myristylation site (Wang., 1993). Oncogenic forms of c-Abl are found in A-MuLV, Hardy-Zuckerman-2 feline sarcoma virus (HZ2-FeSV) and in humans have been linked to chronic myelogenous leukemia (CML) (Franz *et al.*, 1989).

The c-Abl protein activity is tightly regulated *in vivo* and the mechanism of its regulation differs from that of Src family kinases, as there are no internal phosphorylation

sites in the protein (Nam *et al.*, 1996). Overexpression of the c-Abl does not result in cellular transformation (Nam *et al.*, 1996), however, transformation does occur with Abl proteins that lack the SH3 domain, contain a relocated SH3 domain or harbour activating mutations in the kinase domain (Nam *et al.*, 1996). Furthermore, deletion of the SH2 domain prevents transformation in the presence or absence of the SH3 domain (Mayer and Baltimore, 1994), indicating that the SH3 domain is required for suppression of the intrinsic transforming ability and this function is position dependent. In contrast, the SH2 domain is required for transforming activity and this function is independent of its position in the protein (Mayer and Baltimore, 1994).

SH3 domain dependent transforming ability and negative regulation of the kinase activity of the Abl tyrosine kinase is under investigation, although there is some indication that this regulation is through a cis-acting mechanism. Relocation of the SH3 domain within the protein activates c-Abl approximately 20-fold, although structurally these are subtle changes they result in almost complete derepression of c-Abl. The physical relationship between the SH3 domain and other regions of the protein is important for repression of activity indicating the repression is likely to be due to a cis-acting mechanism. This is not through an intramolecular interaction with the C-terminal PRR as the affinity of c-Abl SH3 domain for its PRR is low. The model proposed is a trans-acting mechanism whereby a cellular protein would interact precisely but weakly with the SH3 domain and other parts of the protein (Mayer and Baltimore, 1994). Several candidate inhibitory proteins have been identified as possible inhibitors of c-Abl. One such candidate retinoblastoma protein (RB) acts as an inhibitor of Abl through an interaction between the C-terminal region of RB and the ATP binding lobe of the kinase domain of c-Abl. Hyperphosphorylation of RB results in the release of c-Abl and activation of Abl tyrosine kinase activity during S-phase of the cell cycle. RB inhibits the kinase activity of c-Abl through a mechanism that is independent of the SH3 domain (Welch and Wang, 1993).

Abl associated protein (AAP1) binds to the SH3 domain of c-Abl at amino acids ^{239}LPPGPPPPVP^{248} of the c-Abl PRR and this association is enhanced by the presence of the c-Abl SH2 domain. AAP1 cannot bind the SH3 domain of BCR-Abl fusion protein possibly due to an interaction between the SH2 domain of c-Abl and BCR-Abl or oligomerisation of BCR-Abl protein. AAP1 down-regulates the kinase activity of c-Abl detected using an *in vitro* kinase assay and does so through a direct interaction with the SH3 domain. Thus, AAP1 is a direct inhibitor, but not a substrate of c-Abl (Zhu and Shore, 1996).

Abl interactor protein 2 (Abi2), isolated using a yeast-2-hybrid screen of a human lymphocyte library with Abl SH3 domain as a bait, contains a proposed homeodomain, three PEST domains, two PRRs, a polyproline region and a C-terminal SH3 domain. The presence of the homeodomain suggests that Abi2 may bind DNA. Abi2 regulates c-Abl through two interactions; one between the N-terminal PRR of Abi2 and the SH3 domain of c-Abl and the other through the C-terminal SH3 domain of Abi2 and the PRR region (amino acids 544-637) of Abl. Abi2 may function as a tumour suppressor as removal of amino acids 1-157 from Abi2 increases Abl tyrosine kinase activity in NIH3T3 cells (Dai and Pendergast, 1995). A related protein Abi1 interacts with the kinase domain of Abl kinase and is a substrate for v-Abl tyrosine kinase. Abi1 regulates the activity of v-Abl by forming an interaction with the kinase domain of Abl (Shi *et al.*, 1995).

Pag (also known as macrophage 23kDa stress protein (MSP23) was isolated by yeast-2-hybrid with the Abl SH3 from a HELA cDNA library. Pag protein suppresses the tyrosine kinase activity of c-Abl in a dose dependent manner and requires both the SH3 domain and the ATP lobe of the kinase domain of Abl to form a stable complex. Pag can inhibit phosphorylation of c-Abl in a kinase assay suggesting that Pag is an inhibitor of Abl tyrosine kinase (Wen and Van Etten, 1997). Thus, regulation of Abl tyrosine kinase is likely to be through an interaction with cellular ligands that maintain Abl in an inactive conformation.

The Csk family of tyrosine kinases do not contain the myristylation signal at the N-terminus, an autophosphorylated tyrosine residue in the catalytic domain and the regulatory tyrosine near the C-terminus observed in Src kinases. c-Src kinase (Csk) was the first member identified due to its ability to phosphorylate tyrosine 527 of Src family kinases (section 1.6.1). Csk kinase is a cytoplasmic protein that is ubiquitously expressed (Sabe *et al.*, 1992). Csk contains a high proportion of hydrophilic residues and it is these that are predicted to direct the specificity of Csk (reviewed in Superti-Furga and Courtneidge, (1995)). The regulation of Csk activity is currently unclear. Removal of the SH2 or SH3 domains of Csk has no effect on its kinase activity although such proteins can no longer inhibit Src kinase (Sabe *et al.*, 1994). The SH3 domains of Csk family members are atypical as they lack both ALYDY and PSNY motifs. Thus, the SH3 domain of Csk is unlikely to have a functional role in the regulation, however, it might participate in protein-protein interactions (reviewed in Superti-Furga and Courtneidge, 1995). Thus, the mechanism employed would be predicted to be unique based on the protein domain structure of the protein.

Regulation of tyrosine kinases is a very tightly controlled process. SH3 domains of tyrosine kinases have been shown to participate in the regulation of a variety of proteins including Src, Abl and Tec kinases. In the case of Abl tyrosine kinas, relocation of the SH3 domain within the protein can also activate the protein. The mechanism by which the SH3 domain can repress kinase activity differs between different families. Abl kinase appears to use an allosteric mechanism of repression whereby a protein binds to the SH3 domain and inactivates Abl kinase. Both Src and Tec family members appear to utilise intramolecular interactions to maintain the proteins in an inactive conformation.

The importance of Tec family kinases downstream of the B and T cell antigen receptors has been confirmed by gene targeting studies. The phenotype of mice lacking the Tec family members Btk and Itk are severe with an almost total lack of signalling from the B cells and T cells receptors respectively.

1.7 AIMS AND APPROACH

The N-terminal region of Tec kinase is predicted to perform critical functions in the regulation of this tyrosine kinase family (section 1.6.2). The PH domain of Tec family members is absolutely required for membrane localisation of the protein and subsequent activation while the SH3 domain is required for the maintenance and regulation of an inactive conformation. It is likely that an intramolecular interaction involving the SH3 domain and the PRR of Tec kinase maintains this kinase in a closed inactive state (section 1.6.2). A structural investigation of these regions of Tec kinase should reveal information as to a possible regulatory mechanism of Tec family kinases.

One aim of this work was to determine the three dimensional structure of Tec PHTH domain. The PHTH region of the Tec family member Btk, has been shown to bind to many different signalling molecules, however, whether these interactions are observed for Tec kinase is currently unknown. The solution structure of the Tec PHTH domain will provide the basis for further investigation of these interactions and facilitate analysis of the interactions of other signalling molecules with the PHTH region of Tec from a biophysical perspective.

The second aim of this work was to investigate the potential interactions of the SH3 domain of Tec kinase. The solution structure of Tec SH3 domain was to be determined by Nuclear Magnetic Resonance spectroscopy and analysis of the binding surface undertaken. An investigation of potential interactions between the SH3 domain and the PRR of Tec kinase and whether these interactions are mediated through an intra or intermolecular mechanism was also undertaken.

CHAPTER 2:

MATERIALS AND METHODS

2.1 ABBREVIATIONS

APS	ammonium persulphate
ATP	adenosine triphosphate
BCIG	5-bromo-4-chloro-3-indolyl-β-D-galactoside
bisacrylamide	N, N'-methylene-bisacrylamide
βME	beta-mercaptoethanol
bp	base pair
BSA	bovine serum albumin
CIP	calf intestinal phosphatase
Da	Dalton
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DTE	dithioerythritol
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetra-acetic acid
EtBr	ethidium bromide
GLB	gel loading buffer
GST	glutathione s-transferase
HEPES	N-2-hydroxyethyl piperazine-N-ethane sulphonic acid
HPLC	high performance liquid chromatography
HBS	HEPES buffered saline
IPTG	Isopropyl-β-D-thiogalactopyranoside
kb	kilobase pair
kDa	kilodalton
K_D	equilibrium dissociation constant
LB	luria broth
mA	milli-amperes
MOPS	3-(N-morpholino)propane sulfonic acid
MQ	milli-Q filtered
mwco	molecular weight cut off
NiNTA	nickel nitrilotriacetic acid
NMR	nuclear magnetic resonance

OAc	acetate
OD_n	optical density at a wavelength of n nm
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
PMSF	phenylmethylsulphonyl fluoride
RNA	ribonucleic acid
RNase	ribonuclease
RU	response unit
SDS	sodium dodecyl sulfate
TAE	Tris acetate EDTA
TBS	Tris buffered saline
TFA	Tri-fluoro acetic acid
TTBS	Tris buffered saline, 0.1% Triton X-100
TEMED	N,N,N',N'-tetramethyl-ethylenediamine
Tween 20	polyethylene-sorbitan monolaurate
U	units
UV	ultra violet
V	volts
v	volume
w	weight
µF	micro-farads
2D	two-dimensional
3D	three-dimensional
4D	four-dimensional

2.2 MATERIALS

2.2.1 Chemicals and Reagents

Sigma Chemical Co. supplied the following chemicals: acrylamide, bisacrylamide, agarose (Type 1), ampicillin, ATP, βME, BSA, EtBr, EDTA, glycerol, SDS, thrombin, Tris base, TEMED, triton X-100, reduced glutathione, bromophenol blue, coomassie brilliant blue, H_2NaPO_4, HNa_2PO_4, $NiSO_4$ and imidazole.

BDH Chemicals supplied: chloroform, ethanol, isopropanol, KOAc, MgOAc, $MgCl_2$, NaCl, NaOH, Phenol, TFA.

Sources for other reagents:

Scimar	IPTG and DTT
Pharmacia	Sephadex G75, glutathione sepharose and Resource Q column.
Diagnostic Chemicals Ltd.	DTE
Oxoid	Bacto-agar, Bacto-tryptone, yeast extract and Bacto Nutrient Broth
Bioserve	glutathione agarose resin
Qiagen	Ni NTA resin
Aldrich	$^{15}NH_4Cl$
Novachem	D_2O
BioRad	Bradford reagent

2.2.2 Radiochemicals

[α-^{32}P] dATP was supplied by Geneworks.

2.2.3 Kits

Pharmacia	T7 Sequenase kit
Geneworks	BRESAclean Nucleic acid purification kit
Stratagene	Quickchange™ Site directed mutagenesis kit

2.2.4 Enzymes

Restriction endonucleases were supplied by Pharmacia and New England Biolabs. Other enzymes were obtained from the following sources:

Roche	calf intestinal phosphatase
Sigma	RNase A, thrombin and enterokinase
Pharmacia	T7 DNA polymerase and T4 DNA Ligase.

2.2.5 Buffers

10 x GLB	50% glycerol, 0.1% SDS, 500 μg/μl bromophenol blue, 500 μg/μl xylene cyanol
10 x SD	330 mM Tris-HOAc pH 7.8, 625 mM KOAc 100 mM MgOAc (filter sterilised), 40 mM Spermidine, 5 mM DTE
1 x TAE	40 mM Tris-base, pH 8.2, 20 mM NaOAc, 10 mM EDTA
1 x TE	10 mM Tris-base, pH 7.5, 1 mM EDTA
10 x CIP	500 mM Tris-base, pH 8.5, 1 mM EDTA
10 x Ligase	500 mM Tris-base pH 7.4, 100 mM $MgCl_2$, 100 mM DTT 10 mM ATP
Megadeath	0.1 M NaOH, 0.5% SDS, 10 mM Tris-base, pH 8.0, 1 mM EDTA
Solution I	50 mM glucose, 25 mM Tris-base pH 8.0, 10 mM EDTA pH 8.0
Lysis/Solution II	0.2 M NaOH, 1% SDS
Solution III	3 M KOAc, 2 M HOAc
$CaCl_2$ solution	62 mM $CaCl_2$, 15.5% (w/v) glycerol, made up to 145 ml with MQ water
PCR Buffer	1-3 mM $MgCl_2$, 5 x reaction buffer (provided with enzyme), MQ H_2O
SDS load buffer:	250 mM Tris-base pH 6.8, 10% SDS, 0.5% bromophenol blue, 50% glycerol, 5% βME
HPLC Buffer A	0.1% TFA (v/v), 0.2 μM filtered
HPLC Buffer B	0.088% (v/v) TFA, 80% (v/v) Acetonitrile, 0.2 μM filtered

PBS	25 mM $NaHPO_4$, 250 mM NaCl pH 7 unless otherwise stated, 0.5 μM filtered
HBS	10 mM Hepes, 150 mM NaCl, 3 mM EDTA, 0.005% P20 surfactant, 0.2 μM filtered
TBS	25 mM Tris base, 250 mM NaCl, 0.5 μM filtered
TTBS	1 x TBS, 0.1% triton X-100, 0.5 μM filtered
NiNTA binding buffer	500 mM NaCl, 5 mM imidazole, 20 mM Tris base, pH 7.9
NiNTA wash buffer	500 mM NaCl, 60 mM imidazole, 20 mM Tris base, pH 7.9
NiNTA elute buffer	500 mM NaCl, 1 M imidazole, 20 mM Tris base, pH 7.9
NiNTA strip buffer	100 mM EDTA, 250 mM NaCl, 20 mM Tris base pH 7.9
NiNTA charge buffer	50 mM $NiSO_4$

2.2.6 Plasmid vectors

pGEX-2T	Pharmacia
pGEX-4T-2:	Stratagene
pGEMT-Easy:	Promega
pYEX-2T	Amrad Biotech
pET	Novagen

2.2.7 Oligonucleotides

Synthetic DNA primers were synthesised by Geneworks using a 380B Applied Biosystems DNA Synthesiser. Shown 5′ to 3′.

Sequencing Primers

T7	TAATACGACTCACTATAGGGAGA
T3	ATTAACCCTCACTAAAGGGA
RSP	AAACAGCTATGACCATG
USP	GTAAAACGACGGCCAGT
pGEX-5′	GGGCTGGCAAGCCACGTTTGGTG
pGEX-3′	CCGGGAGCTGCATGTGTCAGAG
Trx-5′	CGAACGCCAGCACATGGACA

PCR primers

Primers used for the amplification of the region encoding theSH3 domain

SH3JD5′	CGCGGATCCAAGAAATCGTTGTAG
SH3JD3′	GCGCTCGAGTCACTCATATTGATCTA

Primers used for amplification of the PRR and mutagenesis

PRR5′	ATGGATCCAGTATAAGAAAGACC
SH3JD3′	GCGCTCGAGTCACTCATATTGATCTA
PRRΔ1-5′	GTATAAGAAAGACCCTGGCGCCCGCGCCAGAAATAAAG
PRRΔ1-3′	CTTTATTTGTGGCGCGGGCGCCAGGGTCTTTCTTATAC
PRRΔ2-5′	CGCCAGAAATAAAGAAGAGAAGGGCGGCCGCGGCCATTCC CCCAGAGGAAGAAAATA
PRRΔ2-3′	ATATTTTCTTCCTCTGGGGGAATGGCCGCGGCCGCCCTTCT CTTCTTTATTTCTGGCG

Primers for amplification of the regions spaning thePHTH domain

PHTHext3′	ACAACGAATTCATTCAGTATT
HTH5′	GCAGGATCCCCACAAAGCAGGGACCGATG
TH5′	CAATAGGATCCATGATTAAATAC
TH3′	GGCAGGGGAATTCTTAACTACTC
PH5′	CAACCAGGGATCCGAGATGAATTTC
TH5′	GTTCTTGAATTCTTATTTTAAC

Primers for the production of the Linker peptide

5′ Linker	GATCCGGCAAAAACGCGCCGACCGTGCGGGGCAAAAGCTATGATA AATAAG
3′ Linker	AATTCTTATTTATCATAGCTTTTGCCCGGCACGGTCGGCGCGTTT TTGCCG

2.2.8 Peptides purchased

The PRR site 1 peptide was purchased from AUSPEP to a purity of greater than 70%. The peptide spans amino acids ^{152}SIRKTLPPAPEI^{163} in the Tec sequence. Mass spectrometric data confirmed the correct weight of the peptide and the purity.

2.2.9 Bacterial strains

DH5α strain *E. coli* were used for chemical heat shock and electroporation transformations constituting routine subcloning. Protein expression and extraction was conducted in *E. coli* strain BL-21 and BL-21 (λDE3)

E. coli strain phenotypes were as follows:

DH5α: *supE*44 Δ*lac* U169 (phi80 *lacZ*ΔM15) *hsdR*17 *recA*1 *endA*1*gyrA*96 *thi*-1*relA*1

BL-21 B F dcm ompT hsdS(rb^-mb^-) gal

BL-21 (λDE3) BL-21 carrying a λDE3 insertion

Strain stocks were stored at -80°C in 40% glycerol.

2.2.10 Yeast strains

DY150 MATa, *ura* 3-52, *leu* 2-3, 112 *trp*1-1, *ade* 2-1, *his* 3-11, *can* 1-100

2.2.11 Bacterial growth media

Luria broth: 1% (w/v) Bacto-tryptone, 0.5% (w/v) yeast extract
1% (w/v) NaCl, adjusted to pH 7.0 with NaOH

2YT 1.6%(w/v) Bacto-tryptone (Difco), 1% (w/v) yeast extract (Difco), 0.5% (w/v) NaCl, adjusted to pH 7.0 with NaOH.

YENB medium: 0.75% Bacto yeast extract, 0.8% Bacto Nutrient Broth.

SOC Medium: 2% Bacto-tryptone, 0.5% Bacto yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose.

MinA media salt solution (final concentration) 60 mM K_2HPO_4, 33 mM KH_2PO_4, 1.7 mM Na_3Citrate, 15 mM NH_4Cl. Following autoclaving 0.005% (w/v) thiamine, 0.2% (w/v) glucose, 0.8 mM $MgSO_4$ was added and made to 1 L with sterile MQ water

MOPS media — MOPS solution 40 mM MOPS pH 7.4, 4 mM tricine pH 7.4, 0.1 mM $FeSO_4$, 2.7 mM K_2SO_4, 5 mM $CaCl_2$, 5.25 mM $MgCl_2$, and 50 mM NaCl. Micronutrient solution 0.3 μM $(NH_4)_6 Mo_7O_{24}$, 40 μM H_2BO_4, 3 μM $CoCl_2$, 1 μM $CuSO_4$, 8 μM $MnCl_2$ and 1 mM $ZnSO_4$. Add to this 0.35% (w/v) glucose, 0.4% glycerol and 1.32 mM K_2HPO_4. Made to 1 L with sterile MQ water

The MOPS solution with micronutrients was made as a 10 x stock.

M9 media — M9 minimal salts per 1l 0.6 g Na2HPO4, 0.3 g KH2PO4, 0.05 g NaCl and 0.1 g NH4Cl. Added to 1 ml 1 M MgSO4 (autoclaved), 10 ml 20% glucose (filter sterilised), 1 ml 1% thiamine (filter sterilised), 10 ml 100 mM CaCl2 (autoclaved). Made to 1L with sterile MQ water.

Growth media were prepared in MQ water and sterilised by autoclaving except where indicated. LB agar plates were made by adding 1.5% (w/v) Bacto-agar (Difco) to the LB. Minimal media was prepared by making a 10 x salt solution and autoclaving. This was then added to 900 ml autoclaved water.

Ampicillin (100 μg/ml) was added after the medium cooled to 55°C to maintain selective pressure for recombinant plasmids in transformed bacteria.

2.2.12 Yeast growth media

YNB: 0.67% yeast nitrogen base without amino acids, 2% glucose

YNB selective media: YNB media, plus 20 μg/ml histidine, 20 μg/ml adenine and 20 μg/ml tryptophan

2.2.13 Molecular weight markers

DNA

*Hpa*II digested pUC19 markers were purchased from Geneworks.

Band sizes (bp): 501, 489, 404, 331, 242, 190, 147, 111, 110, 67, 34, 26.

*Eco*RI digested SPP-1 bacteriophage DNA Markers were purchased from Geneworks.

Band sizes (kb): 8.51, 7.35, 6.11, 4.84, 3.59, 2.81, 1.95, 1.86, 1.51, 1.39, 1.16, 0.98, 0.72, 0.48, 0.36.

During cloning of recombinant DNA, fragment sizes and approximate concentrations were estimated by comparison with loading agarose mini-gels with 500 ng of marker DNA.

Protein

SDS-7 Markers were purchased from Sigma.

Protein constituents: Albumin (Bovine) 66 kDa, Albumin (Egg) 45 kDa, Glyceraldehyde-3-P-dehydrogenase 36 kDa, Carbonic anhydrase (Bovine) 29 kDa, Trypsinogen (Bovine Pancreas) 24 kDa, Trypsin inhibitor (Soybean) 20 kDa, Lactalbumin (Bovine milk) 14.2 kDa.

Minigels were loaded with 10 µl unless otherwise stated.

2.2.14 Miscellaneous materials

X-ray film:	Konica
Centricon column	Millipore
NMR tubes	Wilmad Glass co
BIAcore chips	Pharmacia

2.3 MOLECULAR METHODS

2.3.1 Mini-preparation of plasmid DNA

1.5 ml of LB containing 100 μg/ml of ampicillin was inoculated with a single bacterial colony and grown overnight at 37°C in a rotating drum. Each culture was poured into a 1.5 ml Eppendorf tube and centrifuged at maximum speed for 20 seconds. The majority of the medium was removed leaving around 100 μl. Bacterial pellets were resuspended in the remaining medium by vortexing and lysed by the addition of 300 μl of Megadeath solution (0.1 M NaOH, 0.5% SDS, 10 mM Tris-base pH 8.0 and 1 mM EDTA). Cell debris was precipitated by mixing 155 μl of NaOAc pH 5.2 with the mixture and centrifugation at maximum speed for 4 minutes. Nucleic acids were precipitated by mixing 1 ml of 100% ethanol with the supernatant. The sample was briefly vortexed and centrifuged at maximum speed for 4 minutes prior to removal of the supernatant. The nucleic acid pellet was washed by the addition of 400 μl of 70% ethanol followed by vortexing and a brief centrifugation. The remaining liquid was removed and the pellet was dried for 5-10 minutes at 37 °C. DNA was resuspended in 20 μl of MQ H_2O containing 10 μg/μl RNase A.

2.3.2 Midi-preparation of Plasmid DNA

50 ml of LB containing 100 μg/ml ampicillin in a 250 ml flask was inoculated with a single bacterial colony and grown overnight at 37°C with shaking. The culture was transferred to a 40 ml oakridge tube and centrifuged in a RC-5 Sorvall centrifuge or SS-34 rotor (Dupont) at 6000 rpm for 10 minutes. The supernatant was removed and the bacterial pellet was resuspended in 3 ml of Solution I (50 mM glucose, 25 mM Tris-base pH 8.0, 10 mM EDTA). The suspension was gently mixed with 6 ml of fresh Solution II (0.2 M NaOH, 1% SDS) to lyse the cells. Following 5 minutes incubation on ice, cell debris was precipitated by the addition of 4.5 ml of Solution III (3 M KOAc, 2 M HOAc). The solution was gently mixed by inversion, left on ice for 5 minutes, then mixed vigorously, left on ice for a further 15 - 20 minutes, and centrifuged at 14,000 rpm for 15 minutes at 4°C. The supernatant was mixed with 8 ml of isopropanol in a clean oakridge tube, and nucleic acids were precipitated by centrifugation at 12,000 rpm for 5 minutes at 4°C. The supernatant was drained and the pellet was dissolved in 400 μl of MQ H_2O. RNA was removed by

incubation at 37°C for 30 minutes with 2 µl of RNase A (10 mg/ml). The sample was extracted with an equal volume of phenol/chloroform and then with an equal volume of chloroform alone. The aqueous phase was precipitated by addition of 100 µl of 7 M NH_4Ac and 1 ml of 100% ethanol. DNA was precipitated for 20 minutes at -20°C then pelleted at 14,000 rpm in a bench-top centrifuge for 15 minutes. The DNA was washed in 400 µl of 70% ethanol, dried, and resuspended in 200 µl of MQ H_2O.

2.3.3 Restriction endonuclease digestion of DNA

Plasmid DNA was digested with 4 units of enzyme per microgram of DNA and incubated at the appropriate temperature for 1-6 hours. All restriction digestions were carried out in SD buffer (33 mM Tris-HOAc pH 7.8, 62.5 mM KOAc, 10 mM MgOAc, 4 mM Spermidine, 0.5 mM DTE). Plasmid DNA was assayed for complete digestion by TAE agarose gel electrophoresis.

2.3.4 Agarose gel electrophoresis

Agarose gel electrophoresis was carried out using horizontal mini-gels prepared by pouring 10 ml of gel solution (1% to 3% w/v) agarose in 1 x TAE onto a 7.5 cm x 5.0 cm glass microscope slide. Agarose mini-gels were submerged in 1 x TAE and samples containing 1 x GLB (5% glycerol, 0.01% SDS, 50 µg/µl bromophenol blue, 50 µg/µl xylene cyanol) were typically electrophoresed at 100 V. DNA was visualised by EtBr staining and photographed upon exposure to medium wavelength UV light using a tracktel thermal imager.

2.3.5 Purification of linear DNA fragments

Linear DNA fragments were run on appropriate percentage TAE agarose gels and visualised under long wavelength UV light. Bands were removed from preparative gels using a clean scalpel blade and purified from agarose using the Bresaclean kit (Geneworks) according to the manufacturers instructions.

2.3.6 Removal of 5´ phosphate groups from vector DNA fragments

To inhibit intramolecular ligation of the vector backbone 5´ phosphate groups were removed by treatment with Calf Intestinal Phosphatase (CIP) prior to gel purification.

Restriction digested vector DNA was reacted with 1 µl CIP for 30 minutes at 37°C. Dephosphorylated vector DNA was electrophoresed on 1% TAE agarose gels (section 2.3.2) and purified (section 2.3.3).

2.3.7 Ligation reactions

Complementary end ligation reactions were carried out with 25 ng purified vector, 50-100 ng DNA insert in the presence of ligation buffer (50 mM Tris-base pH 7.4, 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP) and 2 units T_4 DNA Ligase. Reactions were incubated at room temperature for 1 hour.

2.3.8 Annealing of complementary oligonucleotides

Complementary DNA oligonucleotides were annealed by heating to 100°C and slow cooling. The annealed DNA was restriction enzyme digested with appropriate restriction enzymes and electrophoresed on a 2% TAE agarose gel. The annealed DNA was gel purified and ligated into appropriately restricted pGEX-4T-2.

2.3.9 Preparation of calcium chloride competent bacterial cells

5 ml of LB was inoculated with a single colony of DH5α strain bacteria and grown overnight at 37°C with shaking. 5 ml of overnight culture was used to inoculate 500 ml of LB in bevelled shakeflasks. The culture was grown at 37°C to an OD_{600} of 0.4-0.5 for DH5α and 0.6 for BL-21. Cells were poured into two 250 ml oakridge tubes and chilled on ice for 5 minutes prior to centrifugation at 5000 rpm for 15 minutes at 4°C. The supernatant was removed and the cell pellets were each resuspended in 50 ml of $CaCl_2$ solution (62 mM $CaCl_2$, 15.5% (w/v) glycerol) and centrifuged at 5000 rpm for 15 minutes at 4°C. The supernatant was aspirated and the pellets were resuspended in 10 mls of ice-cold $CaCl_2$ solution. 200 µl aliquots were snap frozen in a dry ice/ethanol bath and stored at -80°C.

2.3.10 Bacterial heat shock transformation

$CaCl_2$ competent DH5α cells were thawed on ice for 5-10 minutes. 200 µl aliquots were mixed with DNA (approximately 10 ng of plasmid DNA; half of a ligation reaction) and left on ice for 30 minutes. The cell/DNA mixture was heat shocked for 2 minutes at

42°C and mixed with 1 ml of LB. Cells were allowed to recover by incubation at 37°C for 45 minutes and were pelleted by brief centrifugation in a microfuge at maximum speed. The majority of the LB was removed, leaving around 100 µl, and cells were resuspended and plated on LB plates containing 100 µg/ml ampicillin. 20 ml of 50 µg/ml BCIG (dissolved in dimethyl formamide) and 50 ml of 50 µg/ml IPTG were spread onto plates prior to plating bacteria to colour select for recombinant bacteria.

2.3.11 Preparation of electrocompetent cells for site directed mutagenesis

An isolated DH5α strain colony was picked into 10 ml of YENB medium in a 50 ml flask and incubated overnight at 37°C with shaking. Two baffled 2 L flasks each with 500 ml of YENB, were inoculated with 5 mls of the overnight culture. The cells were allowed to grow to an OD_{600} of 0.8 before centrifugation at 2,600 x g for 10 minutes at 4°C in a pre-cooled rotor. The supernatant was carefully removed and the cells were resuspended in 100 ml of ice-cold 10% glycerol. Following centrifugation at 2,600 x g the supernatant was removed and the pellet resuspended in 100 ml cold 10% glycerol and centrifuged again. The supernatant was removed and the cells were resuspended 2 ml of 10% glycerol. 120 µl aliquots were snap frozen in a dry ice/ethanol bath and stored at -80°C.

2.3.12 Transformation of bacteria by electroporation

Electro-competent cells were allowed to thaw on ice for 10 minutes and 40 µl aliquots were mixed with 1-2 µl of DNA direct from the PCR site directed mutagenesis reaction mix. The cell/DNA mix was transferred to a chilled 0.2 cm gap electroporation cuvette (BIO-RAD) and electroporated at 25 mF, 2,500 V, and 200 ohms in a BIORAD Gene Pulser. Immediately after electric shock cells were suspended in 1 ml of SOC medium and incubated at 37°C for 1 hour. The bacteria were plated onto LB plates containing 100 µg/ml ampicillin and grown overnight at 37°C.

2.3.13 Manual sequencing of plasmid DNA

15 µl of mini-prep DNA (section 2.3.1) was incubated at 37°C for 15 minutes with 1.5 µl RNase A (10 mg/ml) and denatured with the addition of 3.5 µl 1 M NaOH/1 mM EDTA for 15 minutes at 37°C. Denatured plasmid was purified by centrifugation at 1800

rpm for 3 minutes through a Sepharose CL-6B spin column. 1 µl of primer (4.5 ng/ml) was added to 7 µl of purified plasmid DNA in the presence of annealing buffer (20 mM Tris-base pH 7.5, 10 mM $MgCl_2$, 25 mM NaCl) in a total of 10 µl and annealed at 65°C for 2 minutes, then at 37°C for 5 minutes, and room temperature for 15 minutes, and finally kept on ice until required. Briefly 6 µl of labelling mix, 4U modified T7 DNA polymerase, and 5µl µCi/ml [α^{32}P]-dATP) was added to the annealed primer/ DNA mix and incubated at room temperature for 5 minutes. 3.5 µl of the mixture was mixed with 2.5 µl of each termination mix in a separate tube and incubated at 37°C for 5 minutes. Reactions were halted by the addition of 4 µl of FLB (95% (w/v) deionised formamide, 20 mM EDTA, 500 mg/ml bromophenol blue, 500 mg/ml xylene cyanol) and run on a 7M urea / 6% polyacrylamide (2:34 bis-acrylamide to acrylamide) denaturing gel.

6% polyacrylamide sequencing gels were prepared using Sequenca gel solution (National Diagnostics). 400 µl of 10% APS and 20 µl of TEMED were added to 40 mls of stock gel solution prior to pouring between clean glass plates (20 x 40 cm) separated by 0.4 mm spacers. After 30 minutes, the comb was removed and the well flushed with water. Gels were pre-electrophoresed for 1 hour at 2000 V/50 mA and the wells were flushed with 1 x TAE before loading samples. Gels were electrophoresed at 50 mA at a temperature of 50 - 55°C. After electrophoresis, the glass plates were prised apart and the gel transferred to 3 mm paper. The gel was dried down at 70°C on a vacuum gel drier and exposed to fast X-ray film.

2.3.14 Automated sequencing of plasmid DNA

DNA was prepared using the midi preparation protocol (section 2.3.2) or using the BRESAspin minipreparation kit. 0.5 mg of miniprep DNA was subjected to cycle sequencing in the presence of 100 ng of primer and dye terminator mix (Perkin Elmer) in a total volume of 20 µl. The reaction was cycled through the following steps 25 times.

Step 1: 96°C for 30 seconds

Step 2: 50°C for 15 seconds

Step 3: 60°C for 4 minutes

Step 4: repeat steps 1-3 25 times

Step 5: 4°C for 18 hours

Completed reactions were added to 50 µl of 95% ethanol, mixed and allowed to precipitate for 15 minutes at -80°C. DNA was pelleted for 10 minutes at 14,000 rpm,

washed in 250 µl of 70% ethanol, centrifuged again for 5 minutes and air dried. Reactions were analysed at the Institute for Medical and Veterinary Science Sequencing Centre, Adelaide, Australia and viewed on the Editview program (PE Biosystems).

2.3.15 PCR with Taq polymerase

RT-PCR reactions were carried out using Taq polymerase (Geneworks) according to the manufacturer's instructions. Briefly, 25 µl volume PCR reactions containing 20 ng of DNA, 100 ng of each primer, 2 µl of 10 mM dNTPs, 5 ml 5 x PCR buffer containing 1-3 mM MgSO4, and 1 U of Taq polymerase were overlaid with mineral oil, heated to 96°C for 5 minutes, and cycled as follows in a PTC-100 Thermal cycler.

Step 1	96°C for 5 minutes
Step 2	46°C for 1 minute
Step 3	72°C for 1minute
Step 4	repeat steps 1-3, 2 times
Step 5	96°C for 1 minute
Step 6	55°C for 1 minute
Step 7	72°C for 1 minute
Step 8	repeat steps 5-7, 29 times

The reactions were removed from mineral oil following and placed into a clean Eppendorf.

2.3.16 Site directed Mutagenesis by PCR

Site directed mutagenesis was performed essentially as per the Stratagene Quickchange™ Site directed mutagenesis kit. Briefly 5-50 ng template plasmid DNA, 125 ng of each primer, 1 x reaction buffer (provided), 0.05 mM each NTP and 2.5 U of Pfu DNA polymerase made up to 50 µl with MQ H20 and overlayed with mineral oil. The cycling parameters are as follows

Step 1	95°C for 30 seconds
Step 2	95°C for 30 seconds
Step 3	55°C for 1minute
Step 4	68°C for 2minutes/kb of plasmid length
Step 5	repeat steps 1-4 12-18 times

Following the PCR reaction the reaction mix was removed and digested with 20 U DpnI restriction enzyme at 37°C. DNA was transformed by electroporation and plated onto LB plates containing ampicillin.

2.4 PROTEIN METHODS

2.4.1 pGEX Fusion protein induction and extraction from *E. coli*

Protein induction in shakeflasks

Overnight cultures of BL-21λDE3 *E. coli* containing pGEX-4T-2 constructs were diluted 1/100 into 500 ml cultures of LB containing 100 μg/ml ampicillin. Cultures were grown at 30°C or 37°C until they reached an OD_{600} of 0.6, at which time they were induced to produce the GST fusion protein by addition of IPTG to a final concentration of 0.2 mM. Induction was allowed to proceed from 4 hours to a maximum of overnight and bacteria were pelleted at 5,000 rpm, 4°C for 15 minutes in a Beckman JA10 rotor and frozen. The pellet was resuspended in 1 x TTBS at approximately 50 ml/L of culture medium. The solution was sonicated with three 30 second bursts or french pressed, with the incremental addition of PMSF to a final concentration of 1 mM and the solution spun at 10,000 rpm, 4°C for 15 minutes.

Small-scale unlabelled protein preparation of Tec PHTH were grown in MOPS or M9 minimal media supplemented with 100 μg/ml ampicillin (Neidhardt *et al.*, 1974). Bacteria were grown at 30°C, induced with IPTG to a final concentration of 0.2 mM for 4-16 hours, pelleted and stored at -20°C.

Preparations of THSH3, TH and PHTH protein were grown in LB supplemented with 100 μg/ml ampicillin. Bacteria were grown at 37°C and induced with IPTG to a final concentration of 0.2 mM for a minimum of 4 hours in the presence of $ZnSO_4$ to a final concentration of 30 μM. The bacteria was then pelleted and stored at –20°C

Large-scale protein preparations (2 L) of Tec SH3 domain were grown in LB or Min A (Miller, 1972) minimal media supplemented with 0.81 g/l $^{15}NH_4Cl$. Bacteria were grown at 37°C, induced with IPTG to a final concentration of 0.2 mM for 4 hours, pelleted and stored at -20°C.

Protein production in fermenters

500 ml fermentations were conducted using 2YT culture media supplemented with 100 μg/ml ampillicin. Fermenter inoculations were initiated by the addition of 200 μl of an overnight culture of OD_{600} 1.4, per 500 ml fermenter run. Antifoam was added to minimise foaming due to stirring. Glucose was added incrementally to a final concentration of 2% glucose. The growth conditions in the fermenter were regulated by a pH meter and air

monitors, which control the acidity of the media at approximately pH 7.0 and the air flow within the fermenter, respectively. Induction was commenced when the culture reached an OD_{600} of approximately 3.5 by the addition of IPTG to a final concentration of 0.2 mM. The bacteria were grown for a further 4 hours before harvesting.

2.4.2 pET fusion protein induction and extraction from *E .coli*

pET vectors were transformed in BL-21λDE3 *E. coli* and grown in LB medium supplemented with 100 μg/ml ampicillin. Bacteria were grown to an OD_{600} of 0.6, induced to a final IPTG concentration of 1.0 mM, grown a further 4 hours and harvested. Bacterial pellets were resuspended in 1 x binding buffer containing 5 mM imidazole, 500 mM NaCl and 20 mM Tris-base pH 7.9, lysed using three 30 second sonication bursts and centrifuged at 10,000 rpm. The soluble fraction was 0.5 μM filtered and applied to a NiNTA column equilibrated in lysis solution.

Protein was isolated from the insoluble pellet by resuspension in 1 x binding buffer containing 5 mM imidazole and 6 M urea and sonicated as above. Following centrifugation at 10,000 rpm the supernatent was applied to a NiNTA column.

2.4.3 pYex fusion protein Induction and extraction from *S.cerriveiase*

Yeast pYex constructs were transformed in yeast and grown in the yeast media YNB plus histidine, adenosine, tryptophan and glucose (Mitchell *et al.*, 1993). Yeast was grown overnight in 10 ml volumes, pelleted, resuspended in the same media and then induced by the addition of copper sulfate to a final concentration of 0.5 mM. Yeast were then grown for a further 4 hours. Following lysis by French pressing the soluble fraction was purified as per the bacterial GST fusion proteins.

2.4.4. Affinity purification using glutathione agarose/Sepharose

Column based glutathione agarose/sepharose purification

GST-fusion proteins were purified from bacterial protein extracts using a glutathione agarose (Bioserve) or glutathione Sepharose (Pharmacia) column. The column was equilibrated in 1 x TBS with at least three column volumes and then with 1 x TBS/0.1% (v/v) triton X-100 (TTBS). The bacterial extracts were 0.45 μm filtered and loaded on the column at a flow rate of 4 ml/min. The column was then washed extensively with TTBS

and TBS. Bound protein was eluted by the addition of 10 mM reduced glutathione in 1 x TBS pH 8.0. The GST fusion proteins were stored at -80°C.

Prior to affinity purification of ^{15}N-labelled GST fusion proteins the GST affinity column was washed with two column volumes of 70% ethanol, six column volumes of MQ water and two column volumes of 6 M guanidiniun-HCL. The column was extensively equilibrated in TBS following the cleaning procedure.

Batch method purification of GST fusion proteins

Glutathione agarose beads equilibrated in TTBS were mixed with the lysed bacteria for 30 minutes. The mix was placed in an empty PD10 column and washed extensively with TTBS and TBS. Elution was facilitated using reduced glutathione in 1x TBS pH 8.0 and 1 ml fractions were collected.

2.4.5 Nickel affinity purification

Nickel affinity resins (Qiagen –NiNTA) were used to purify pET fusion proteins. Purification was conducted under both non-denaturing and denaturing conditions, the latter achieved by the addition of 6 M urea. Bacteria were resuspended in 1 x binding buffer (5 mM imidazole, 500 mM NaCl and 20 mM Tris base pH 7.9) at 100 ml binding buffer/1L of LB media. Three 30 second sonication bursts were used to lyse the *E. coli*, which was supplemented with PMSF to a final concentration of 1 mM. Following lysis, the lysates were centrifuged and the supernatent applied to a NiNTA column. The column was washed extensively with binding buffer followed by a wash buffer (500 mM NaCl, 60 mM imidazole, 20 mM Tris base, pH 7.9). The protein sample was eluted from the column using elute buffer (1 M imidazole, 500 mM NaCl and 20 mM Tris base pH 7.9).

Purification of pET-PHTH bacterial lysates was also conducted using a binding/lysis buffer containing no imidazole (20 mM Tris base pH 7.9 and 100 mM NaCl). The column was washed with lysis buffer and elution conducted using firstly 50 mM imidazole (50 mM imidazole, 500 mM NaCl and 20 mM Tris base pH 7.9) followed by a 100 mM imidazole elution (100 mM imidazole, 500 mM NaCl and 20 mM Tris base pH 7.9). The protein samples were collected in 1 ml aliquots and stored at 4°C until required.

2.4.6 Protein refolding on the NiNTA column

Denatured protein was loaded on a NiNTA column pre-equilibrated in 1 x binding buffer (5 mM imidazole, 500 mM NaCl and 20 mM Tris base pH 7.9) with 6 M Urea. The column was extensively washed with the binding buffer. The column was washed with two column volumes of 4 M urea buffer (5 mM imidazole, 500 mM NaCl, 20 mM Tris base pH 7.9 and 4 M urea) then 2 M urea and 1 M urea. The sample was eluted in 250 mM imidazole containing 125 mM NaCl and 20 mM Tris base pH 7.9.

Denatured protein was also eluted from the NiNTA column as above, however, there was no incremental decrease in the urea content. The protein was loaded on the column in 1 x binding buffer (5 mM imidazole, 500 mM NaCl and 20mM Tris base PH 7.9) containing 6 M urea, washed with urea binding buffer and eluted in 250 mM imidazole containing 125 mM NaCl and 20 mM Tris base pH 7.9.

2.4.7 Protein refolding by dialysis

Denatured protein was eluted from a NiNTA affinity column in elute buffer (500 mM NaCl, 1 M imidazole, 20 mM Tris base, pH 7.9) containing 6 M urea. Dialysis was conducted using pre boiled dialysis tubing with a 10 kDa mwco. Two different methods were used.

Stepwise refolding

Stepwise dialysis was carried out by successive incubations with 4 M, 2 M, and 1 M urea solutions. The dialysis buffer was changed four times and a 1/500 dilution was used. Decreasing urea concentrations were used 4 M, 2 M and 1 M urea and the final conditions were 250 mM NaCl, 25 mM Tris base and 15 mM β-mercaptoethanol. The concentration of NaCl and Tris base were adjusted when the buffer was changed at 4 M urea to the final conditions of 250 mM NaCl, 25 mM Tris-base and 15 mM β-mercaptoethanol. Dialysis buffer was changed every 12 hours.

Direct dialysis

The direct dialysis started in elute buffer (500 mM NaCl, 1 M imidazole, 20 mM Tris base, pH 7.9) containing 6 M Urea and was dialysed directly into of 150 mM NaCl, 50 mM Tris-base and 7.5 mM DTT using a 1/500 dilution overnight.

A PHTH sample was dialysed against 250 mM NaCl, 25 mM Tris base for 24 hours in a solution containing Peg 3360 and 1.5 mM DTT to assist in the refolding. The buffer contained a 13.5:1 PEG 3650: protein ratio as determined in Cleland *et al* (1992) based on a hydrophobicity index of –63.25 for Trx-PHTH (Cleland *et al.*, 1992).

2.4.8 Determination of protein concentration by Bradford assay

10 µl of BSA standards (0 - 1 mg/ml) or samples were mixed with 200 µl of 1 in 4 diluted Bradford reagent (BIO-RAD) in a 96-well tray. Absorbance at 600 nm wavelength was measured in an Emax plate reader (Molecular Dynamics). Samples and standards were performed in triplicate and averages were used in further calculations. Protein concentrations of samples were determined by calculation from the line of best fit of the standard curve (Bradford, 1976).

2.4.9 Thrombin cleavage

Purified fusion protein was cleaved with bovine thrombin to remove the GST fusion partner. Thrombin digests were performed in the TBS reduced glutathione elute buffer supplemented with 2.5 mM $CaCl_2$. Reactions were incubated at 37°C for a minimum of 1 hour to a maximum of 24 hours with ratios of enzyme (units) to total protein (mg) ranging from 1:1 to 4:1.

Digestion on the column

A solution containing 200 units of thrombin in TBS supplemented with final concentrations of 2.5 mM $CaCl_2$, 10 mM DTT and 0.2% NaN_3 was added to the column and incubated overnight at room temperature. TBS was applied to the column at 2 ml/min to elute the protein sample. The GST fraction was eluted by the addition of 10 mM reduced glutathione solution in TBS pH 8.0.

2.4.10 Enterokinase cleavage

pET purified protein was proteolytically digested with enterokinase. Enterokinase diluted 1/20, 1/50, 1/100 and 1/200 (v/v) were used to digest samples of concentration 0.7 mg/ml protein. 10 µl samples were removed at different timepoints, the reaction stopped by the addition of PMSF to a final concentration of 1 mM and the samples prepared for SDS-PAGE analysis.

2.4.11 Size exclusion chromatography

Prior to ^{15}N sample purification using this Superdex G75 column (Pharmacia) it was washed for 2 hours with 0.5 M NaOH at a flow rate of 1 ml/minute and then equilibrated extensively in buffer. The column was standardised against a solution containing the proteins lysozyme (14 kDa), chymotrypsin (24 kDa), ovalbumin (45 kDa) and BSA (66 kDa) to generate a standard curve.

Thrombin digested GST-fusion proteins were loaded onto a 300 ml Superdex G75 column (Pharmacia) equilibrated overnight in 1 x HBS, 1 x TBS 0.01% NaN_3, 1 x PBS, 0.01% NaN_3 or 1 x TBS containing 10 mM DTT and 30 mM $ZnSO_4$. The sample was loaded and run at 2 ml/minute. Protein was detected at A_{280} and 5 ml fractions collected. Samples corresponding to the protein of interest were confirmed by SDS-PAGE, pooled and stored at 4°C.

Pet-PHTH digested samples were run on a Superdex G75 PC 3.2/30 (Pharmacia) column at a flow rate of 50 µl/minute equilibrated in TBS. Protein elution was detected at A_{280}. 100 µl samples were collected for further analysis and the appropriate fractions retained. The column was standardised using a solution of lysozyme (14 kDa), chymotrypsin (24 kDa), ovalbumin (45 kDa) and BSA (66 kDa). The relative elutions are shown on the profiles.

2.4.12 Ion exchange chromatography of THSH3 protein

THSH3 was purified using a 1 ml Resource Q ion exchange column equilibrated in 25 mM Tris base pH 8.0 and 250 mM NaCl. Elution was conducted using a 0-50% gradient over 10 minutes with a final buffer of 25 mM Tris pH 8.0 and 1 M NaCl and a flow rate of 1 ml/minute. Samples were detected at A_{280} and the fractions containing the protein of interest retained and further analysed by SDS-PAGE.

2.4.13 PD10 buffer exchange and sample generation.

PD10 columns were equilibrated in 20 ml of the buffer relevant to the designed experiment, HBS, 10 mM Phosphate 0.01%NaN3, MQ H_2O or 1 x PBS for BIAcore, NMR or ultracentrifugation experiments. Samples were concentrated to 2 ml using an amicon stirred cell, loaded on the PD10 column and 1 ml fractions collected until 15 ml had been eluted. Sample elution was monitored by Bradford reaction and fractions containing the

protein of interest were pooled and further concentrated using a centricon spin column of 3 kDa mwco.

2.4.14 Ultrafiltration of the synthesised peptide

Preparations of the high affinity peptide were separated from the GST fusion partner using an amicon stirred cell. A 10,000 mwco membrane was used. The GST fusion protein was retained in the stirred cell while the synthesised peptide passed though the membrane. The linker peptide was acidified before applying to HPLC column for further purification.

2.4.15 Reverse phase high performance liquid chromatography analysis

Reverse phase HPLC was used to separate and purify peptides and prepare samples for mass spectrometric analysis. A C4 analytical column (Brownlee Aquapore BU300, 7 μm particle size, 300-A pore size 4.6 mm x 100 mm) and a C18 analytical column (Brownlee analytical RP-18, Spheri-5, 5 μm particle size, 4.6 x 100mm) were used for testing the purification of the synthesised peptide and preparation of peptides to be analysed by mass spectroscopy. A C4 prep10 column (Analytical Prep10 Butyl Aquapore C4, 7 μm particle size, 10 x 100mm) was used for purification of the peptides. Samples were acidified with TFA prior to loading on the column. Columns were equilibrated with 0.1%TFA in H_2O between runs. Various linear gradients using 0.1%TFA in H_2O and 0.088%TFA/80% acetonitrile were used. A flow rate of 0.5 ml/minute and 2.0 ml/minute was used for the analytical and Prep10 columns respectively. The eluted peptides were detected by absorbance at 215 nm and 280 nm using a Flow Through Ultraviolet Detector (Waters, Model 490). Fractions were collected in Eppendorf tubes. Reverse phase HPLC was controlled using Waters Expert™, and Milenium™ chromatography software.

The synthesised peptide (Chapter 4) was acidified and a gradient of 0.1% TFA in water and 0.1%TFA in 80% acetonitrile was used with a C8 column. Gradients of 50-100% over 20 minutes were sufficient to get good separation of the peptide and GST and any residual GST fusion still present. The samples were lyophilised twice to remove any residual TFA.

2.4.16 Determination of molecular mass

Molecular masses were determined by electrospray mass spectrometry conducted at the Waite Institute, Adelaide, Australia. Samples were lyophilised from HPLC buffer. Aliquots (1-2 ul) of the SH3 domain NMR samples in 10 mM phosphate were sent for mass spectrometry.

2.4.17 SDS-PAGE Gels

Denaturing 12.5% acrylamide Tris/Tricine SDS gels were prepared using a Hoefer gel pouring apparatus. The resolving gel, containing 6.35 ml acrylamide (40%), 5.3 ml 50% glycerol, 8.325 ml 3x gel buffer (3.0 M Tris base, 0.3% SDS), 5 ml MQ H_2O, 35 µl Temed and 100 µl 10% APS was allowed to polymerise. The stacking gel was prepared by combining 0.712 ml acrylamide, 1.95 ml 3 x gel buffer, 5.16 ml MQ H_2O, 7.5 µl Temed and 62.5 µl APS. Samples were mixed with 5 x SDS load buffer to a 2 x final concentration and heated to 100°C for 3 minutes before being loaded onto the gel. Gels were electrophoresed at 50 mA using a continuous buffering system, with the anode buffer (0.2 M Tris-base pH 8.9) and the cathode buffer (0.1 M Tris-base/1%SDS and 1.0 M Tricine). The gels were run until the ion front reached the bottom of the gel.

Protein content was stained for using coomassie stain (1 g Coomassie, 30% methanol, 10% acetic acid). Stain was removed by incubating overnight in Destain solution (50% methanol, 5% glacial acetic acid). SDS-PAGE gels were dried using a vacuum gel drier.

2.5 NMR SPECTROSCOPY

2.5.1 NMR conducted on the PHTH proteins

NMR tubes were cleaned for NMR experiments by soaking overnight in a 10 M Nitric acid solution. Tubes where then washed ten times in distilled water and then ten times in MQ water. The tubes were dried in a 60°C oven overnight, sealed and stored until required.

NMR experiments were collected on a 1 mM PHTH sample at 25°C. One dimensional presaturation experiments were conducted initially to test the purity and the signal strength of the sample. The acquisition of these experiments was conducted using the program UXNMR and the parameters were as follows.

	DQF-COSY	TOCSY	NOESY
Pulse program	Cosydfprtp	Dipsiprtp	noesyprtp
Temperature	25°C	25°C	25°C
Mixing time	-	43 ms	140 ms
P1 (90° pulse)	6.8 µsec	6.8 µsec	6.8 µsec
ds	72 µs	72 µs	72 µs
Spectral frequency	500.139 MHz	500.139 MHz	500.139 MHz
SWH	6994 Hz	6994Hz	6994Hz
(Spectral width)	13.885 ppm	13.885ppm	13.885ppm
F2 TD (points)	4096	4096	4096
F1TD (points)	400	400	400
NS (scans)	80	64	128

Mixing times of 100 ms, 150 ms and 200 ms were used for the NOESY experiments. These data were Fourier transformed using States/TPPI with a 90° sinebell function in dimension 1 and using complex processing with a 45° sine squared function with zero filling in D2, total points were 2048 and 1024 in D1 and D2 respectively. Data manipulation was conducted using FELIX (MSI). Spectra were referenced to water at 1024 points in D1 and 512 points in D2 with a chemical shift of 4.7 ppm. The data was transferred into XEASY for printing. One dimensional pre-saturation spectra were recorded on a 0.5 mM TH domain protein sample at 25°C. A spectral width of 8000Hz was used with 8192 data points and 64 transients collected.

2.5.2 NMR conducted on the SH3 protein

Initial 3D-NOESY HSQC and 3D-TOCSY HMQC data were recorded and processed at the Biomolecular Research Institute Melbourne with the assistance of Dr Mark G. Hinds. Tec SH3 domain was buffer exchanged 10 mM phosphate, 0.01% NaN_3, pH 6.0 using PD10 columns (Pharmacia). Sample concentrations of 2.0 mM and 1.25 mM were used in this study for the unlabelled and ^{15}N labelled samples, respectively. NMR experiments were performed on Bruker AMX-500 and DRX-600. All data sets were recorded at 25°C using 5 mm inverse triple resonance $^{1}H/^{13}C/^{15}N$ pfg probes. These data were recorded using 1024 x 32 x 72 complex points and sweep widths of 1824 Hz in F1 and 7507 Hz in F2 and F3. NOESY and TOCSY mixing times of 200 ms and 50 ms, respectively, were used for both the 3D and 2D experiments. Cosine-squared functions were applied to the data prior to Fourier transformation. The final matrix size was 896 x 64 x 256. 2D ^{1}H-^{15}N HSQC data was recorded with spectral widths of 2500 Hz in F2 and 6250 Hz in F1, using 400 t1 increments. The centre of the spectra were referenced to 4.85 ppm on the proton axis and 120 ppm on the nitrogen-15 axis.

2D-NOESY, TOCSY and DQF-COSY experiments were recorded on a Varian Inova 600 spectrometer with a sweep width of 8000 Hz in both dimensions and a minimum of 2048 x 512 complex points. A cosine-squared function was applied to the data prior to Fourier transformation. This data was also zero-filled to a final matrix size of 2048 x 2048. Amide exchange experiments 2D [^{1}H-^{15}N]-HSQC and 2D[^{1}H-^{15}N]-HMQC-J were conducted to clarify slow and fast exchanging amino acids. Spectra were calibrated on the centre of the spectra at 4.85 ppm and 120 ppm. Processed data was analysed using XEASY (Bartels *et al.*, 1995).

2.5.3 Ligand titration experiments

The site 1 peptide was purchased from AUSPEP. A NMR sample of Tec SH3 domain was made to a concentration of 0.75 mM buffered with 10 mM phosphate pH 6.0 with an addition of 10% D_2O. Synthesised ligand was dissolved in water to a concentration of 41.2 mM and added in molar amounts ranging from 0.5 to 6 (0.375 mM-4.5 mM) molar excess. PRR site 1 peptide was also dissolved to concentration of 41.2 mM and was added to a new sample of Tec SH3 domain in increasing amounts from 0.5 to 8 molar excess (0.375 mM-6 mM). The pH of the sample was checked between experiments and maintained at pH 6.0. HSQC experiments were conducted using the pulse program ghsqcse

with 2048 complex points in the proton dimension and 64 t_1 increments at 25°C. The spectra were processed using a sine squared function and linear prediction used. The spectra were all calibrated on 512 complex points and 4.85 ppm in the proton dimension and 64 complex points and 120 ppm in the nitrogen dimension. Chemical shift values were recorded using XEASY software (Bartels *et al* , 1995).

Calculation of affinity of interaction

Plots of difference in chemical shift V's concentration of peptide were generated in excel and the kinetics of the reaction investigated. The affinity of reaction was determined as in Tsukabe *et al.*, (1996). This method is a Scatchard based method and does not require the reaction to be at equilibrium, however, does require the reaction to be a slow exchanging reaction. Briefly,

R + S = RS where R and S are free components and RS is the complex. The rate constant for the reaction is K= [RS]/[R][S] where [RS], [R] and [S] are the equilibrium concentrations of the complex [RS] and the substrates [R], [S] respectively.

Assuming the concentration of the ligand > receptor a Scatchard plot is generated

$\frac{\Delta}{[L]} = -K\Delta + \Delta oK$ where

K can be substituted with $1/K_D$ giving the equation

$$\frac{\Delta}{[L]} = \frac{-1}{K_D}\Delta + \Delta o \frac{1}{K_D}$$

where

$\Delta = [\delta_{\text{free receptor}} - \delta_{\text{RL @ given [L]}}]$

$\Delta_0 = \delta_{\text{free}} - \delta_{\text{RL@saturation}}$

[L] = concentration of ligand (mM)

K = equilibrium constant (association)

K_D = dissociation constant

Plot $\Delta/[L]$ against Δ and the slope of the line is $1/K_D$

2.6 STRUCTURE DETERMINATION

2.6.1 Structural restraints

Distance restraints were derived from a 2D [^{1}H-^{1}H]-NOESY spectrum (recorded in D_2O) and a 3D [^{1}H-^{15}N]-NOESY-HSQC spectrum, both with mixing times of 200 ms. Distance restraint upper bounds were calibrated using the method of (Xu *et al.*, 2000). so that upper bounds ranged from 6.0 to 2.2 Å. Lower bounds were set to 1.8 Å. Hydrogen bond restraints were imposed for amide groups detectable in [^{1}H-^{15}N]-HSQC spectra recorded more than 1 hour after protein, lyophilised from H_2O, was dissolved in D_2O. The hydrogen-acceptor distance was restrained between 1.7 and 2.2 Å and the donor-acceptor distance was restrained between 2.7 and 3.2 Å.

Values of $^{3}J_{HN\text{-}H\alpha}$ were measured from F1 and F2 cross peak line widths in a [^{1}H-^{15}N]-HMQC-J experiment using the method of (Wishart and Wang, 1998). Torsion angle restraints for ϕ-angles of $-120 \pm 40°$ were imposed for $^{3}J_{HN\text{-}H\alpha} \bullet 8$ Hz and $-60 \pm 30°$ for $\bullet 5$ Hz.

2.6.2 Structure calculations

Structures were calculated in X-PLOR (version 3.851) (Brunger, 1991) using the ARIA method of Nilges *et al.*, 1997. A raw set of distances restraints was extracted from the NOESY peak list by interrogating the chemical shift assignments using windows of ±0.04 ppm for F1 and ±0.02 ppm for F2 in the 2D NOESY and windows of ±0.06 ppm for F1 (1H), ±0.25 ppm for F2 (15N) and ±0.03 ppm for F3 (1H) in the 3D NOESY-HSQC. An initial ensemble of 40 structures was calculated using the raw restraints, which were composed of 49 unambiguous restraints (22 intraresidue, 6 sequential, 2 medium range and 19 long range restraints), 1566 ambiguous restraints and 31 ϕ-angle restraints. The 10 structures with the lowest overall energy were retained. Ten ARIA iterations followed during which the assignment parameter (N_p) was reduced from 0.9999 to 0.75 and the violation tolerance (V_{tol}) was reduced from 2.0 to 0 Å. Restraints were excluded if they were violated in more than 50% of the retained structures from the previous iteration. Hydrogen bond restraints were included for the slowly exchanging amide groups when a single acceptor was found with suitable hydrogen bonding geometry in at least 50 % of the retained structures. Restraint lists were filtered to remove redundancy with the restraint with the

lowest upper bound being retained. A total of 100 structures were calculated using the final restraint list (generated with N_p of 0.75 and V_{tol} of 0 Å). These final restraints included 1228 NOE-derived distance restraints including 956 restraints assigned unambiguously (468 intraresidue, 173 sequential, 55 medium range and 260 long range) and 272 restraints for which the assignment remained ambiguous. The remaining restraints were 34 hydrogen bond restraints (two restraints per hydrogen bond) and 31 ϕ-angle restraints. A final round of refinement of the 100 structures from the last ARIA iteration was carried out in explicit solvent using the OPLSX non-bonded parameter set (Linge and Nilges, 1999). The 20 structures with the lowest overall energies were selected as the final ensemble.

2.6.3 Structural analyses

Hydrogen bonding and secondary structure were analysed using MOLMOL (Koradi *et al.*, 1996) and Ramachandran properties and angular order parameters were measured using the in-house program ANGORDER and PROCHECK.

The well-defined residues were identified by iterative fitting of the C^{α} atoms to define the subset with the best-defined C^{α} positions. Following each iteration the C^{α} atom with the highest RMSD was excluded prior to the next fit. At each iteration the C^{α} RMSD, for the retained subset, was divided by the number of residues in the subset. The set of best-defined residues was the subset for which this ratio was the minimum.

2.7 SURFACE PLASMON RESONANCE –BIAcore

Surface plasmon resonance of PRRSH3 proteins was carried out using a BIAcore 2000 instrument. Covalent attachment of the Tec SH3 domain to the CM5 biosensor chip was achieved by the amine coupling method as per manufacturers instructions. Briefly, Tec SH3 domain was diluted in 10 mM NaOAc pH 4.6 and injected onto the activated CM5 surface at 5 µl/minute with HBS (10 mM Hepes, 150 mM NaCl 3 mM EDTA, 0.005% P20) as the running buffer. Residual binding sites were quenched with ethanolamine. Two different amounts of Tec SH3 domain were coupled to the chip surface 1500-3000 RU.

The PRRSH3 proteins were buffer exchanged into HBS prior to the experiments. In the binding analysis PRRSH3 proteins at different concentrations were injected for 5

minutes at a flow rate of 10 µl/minute with HBS as the running buffer. The dissociation phase was initiated by switching the injection stream from the PRRSH3 sample to HBS and was conducted for 68 minutes. Duplicate injections were conducted in random order.

Regeneration was facilitated by an extended wash of the chip surface with HBS buffer at 30 µl/minute for 45 minutes. The response of the SH3 domain surface was consistent within one experiment consisting of 12 injections. Each experiment was conducted on a new BIAcore surface on the chip. Report points were taken at pre injection, post injection and after the regeneration run. BIAcore analysis was conducted using the BIAevaluation software.

Kinetic model used

Equilibrium affinity constants of the PRRSH3 interaction with the Tec SH3 domain were determined using the BIAevaluation program as per manufacturers instructions. The analysis of Tec PRRSH3 protein binding to the SH3 domain on the chip was conducted using the Steady State model provided in the BIAevaluation software. Briefly,

A + B= AB at equilibrium.

$$Req = \frac{KCR_{max}}{KC_{n+1}}$$

Req = response at equilibrium

K = equilibrium affinity constant

R_{max} = maximum surface response.

C = concentration

Erratum

$$Req = \frac{KCR_{max}}{1+KCn}$$

n = steric interference factor specifying how many binding sites on average are blocked by binding one analyte molecule (default=1, Tec SH3:Tec PRR =1)

The affinity constant is determined from a plot of Req against concentration (C).

The closeness of the fit is described using χ^2 values where χ^2 is

$$\chi^2 = \frac{1}{n-p} \sum^{n} (r_f - r_x)^2$$

where r_f is the fitted value

r_x is the experimental value

n is the number of data points

p is the number of fitted parameters

As n>>p then

n-p =n

and χ^2 reduces to the average squared residual per data point.

2.8 FLUORESCENCE EXPERIMENTS

Purified Trx-PHTH protein were diluted to a final concentration of 0.25 mM in varying urea concentrations in a buffer containing 250 mM NaCl, 25 mM Tris-base, 1.5 mM DTT and a 13.5:1 PEG 3650: protein ratio as determined in (Cleland *et al.*, 1992). A 0.4 mg/ml sample that had been dialysed against 250 mM NaCl, 25 mM Tris-base and 1.5 mM DTT for 24 hours in a solution containing Peg 3360 to assist in the refolding was also tested.

GST-PHTH (0.37 mg/ml) in 1 x TBS and 10 mM reduced glutathione and PHTH (0.1 mg/ml) in 1 x PBS were also tested by fluorescence. Background buffer fluorescence scans were subtracted from all the protein fluorescence scans. The purity and size of the protein samples were confirmed by SDS-PAGE.

Fluorescence spectroscopy was conducted on an SLM-Aminco Bowman Series 2 Fluorescence spectrometer polarimeter with the excitation wavelength of 295 nm an emission wavelength of 350 nm and spectra were collected with the range of 310-400 nm.

2.9 ANALYTICAL ULTRACENTRIFUGATION EXPERIMENTS

Sedimentation equilibrium experiments to characterise the self association of Tec PRRSH3 wild-type were carried out on a Beckman Optima XL-A analytical ultracentrifuge equipped with an An-60ti rotor. Data was recorded and processed by Dr Joel Mackay, Biochemistry, University of Sydney. Protein samples were made up in phosphate buffered saline (pH 7.0) at a number of loading concentrations in the range 8–100 μM, and data were recorded at both 4 and 20 °C at speeds of 20,000, 30,000, 42,000, 48,000 and 56,000 rpm. Data were collected in six-sector cells as absorbance (248 and 280 nm) versus radius scans (0.001 cm increments). Scans were collected at 4-hour intervals and compared in order to determine when the sample reached equilibrium. Analysis of the data was carried out using

the NONLIN software (Johnson *et al.*, 1981), and the best model and final parameters were determined by examination of the residuals derived from fits to monomer, monomer ↔ dimer, monomer ↔ trimer, and monomer ↔ dimer ↔ tetramer models (all ideal species models). The partial specific volumes of each domain were determined from the amino acid sequences (Perkins, 1986), and the solvent density was taken to be 1.0066 g ml^{-1} at 20°C.

Purified PHTH protein was buffer exchanged into 10 mM phosphate, 150 mM NaCl, 0.01 mM NaN_3 and 1 mM DTT. Two concentrations of PHTH protein 0.025 mM and 0.0125 mM were conducted at three different rotor speeds, 12 K, 18 K and 28 K. All experiments were conducted at 5°C.

2.10 CONTAINMENT FACILITIES

Manipulations that involved organisms containing recombinant DNA were carried out in accordance with the regulations and approval of the Australian Academy of Science Committee on Recombinant DNA and the University Council of the University of Adelaide.

CHAPTER 3:

CHARACTERISATION OF TEC PHTH PROTEINS

3.1 INTRODUCTION

3.1.1 PHTH domains of Tec family members

PH domains are present in all Tec family members except Txk. In Tec kinases, the PH domain is extended C-terminally to include the TH region incorporating the Btk motif and a PRR. Besides Tec family members, the TH domain exhibits only limited similarity with other proteins. The Btk motif has been identified adjacent to a PH domain in *Drosophila* Gap protein and in a human interferon-γ binding protein. The PRR region has similarity to a region in p85 subunit of PI3K and SLK1/SSP31 in yeast (Vihinen *et al.*, 1994). Tec kinases are the only known family of intracellular tyrosine kinases that contain a PH domain. PH domains of Tec family members show higher conservation than generally seen for non-paralogous PH domains with over 50% identity compared with less than 10% observed between some PH domains (Okoh and Vihinen, 1999) (Figure 3.1A).

Btk has been the most studied Tec family protein because mutations in Btk, within the PH domain, result in X-linked agammaglobulinemia in humans (Figure 3.2A). The three dimensional structure of the Btk PH domain and Btk motif has been determined with and without the inositol (1,3,4,5) tetrakisphosphate ligand (Hyvonen and Saraste, 1997); (Baraldi *et al.*, 1999). The structure of the Btk PH Btk motif region is shown in Figure 3.1B. The PH domain of Btk has a fold common to other PH domains, that of a seven stranded β-barrel with a capping α-helix. A large loop between β-strands 5 and 6 includes a short stretch of α-helix. The Btk motif is globular and packs against strands 5 and 7 of the Btk PH domain. It coordinates a single zinc atom in a distorted tetrahedral geometry with the sidechains of amino acids His 143, Cys 154, Cys 155 and Cys 165. This zinc coordination has similarities to Lim repeats and GATA-1 zinc fingers, which are also coordinated through a three cysteine and one histidine pattern (Hyvonen and Saraste, 1997). The Btk motif is essential for the structural stability of the Btk PH domain (Baraldi *et al.*, 1999). The PH domains of the other Tec family members have been modelled and are predicted to be similar to that of Btk (Okoh and Vihinen, 1999) (Figure 3.1B).

A considerable amount of work has been conducted on possible interactions of the Btk PH domain. The PH domain of Btk has been shown to bind the βγ subunits of G-proteins (Gβγ) with high affinity (Touhara *et al.*, 1994). It is possible that the Btk motif will improve the interaction of Btk's PH domain with Gβγ subunits as an increase in binding is

Figure 3.1

A. Sequence alignment of Tec family PHTH regions. Conserved amino acids are represented in blue and similar amino acids are shown in red. The amino acid regions boxed in blue indicate the β-stands in the structure of the PH domain. Red boxes indicate the α-helices and green boxes indicate residues involved in zinc binding. The tryptophan, conserved in all PH domains is boxed in black. The alignment was created using a ClustalW algorithm (Thompson *et al.*, 1994) and coloured using the Mac Boxshade program.

B. Three-dimensional structure of the PH domain, Btk motif of human Btk that contains a R28C amino acid substitution. The β-strands are shown in blue and are numbered 1-7 from the N terminus. The two α-helices are shown in red. The zinc atom is coloured magenta. Taken from Hyvonen and Saraste (1997)(PDB accession 1BTK).

A

```
                .   10        .   20        .   30        .   40        .   50        .   60        .   70        .   80
mTec     1 MNFNTILEEILIKRSQQKKKTSLLNYKERLCVLPKSVLSYYEGR---AEKKYRKGVIDISKIKCVEIVKNDDGVIPCQN-   76
mBmx     1 MESKSILEELLLKKSQQKKKMSPNNYKERLFVLTKTSLSYYEYD--KMKRGSRKGSIEIKKIRCVEKVNLEEQTPVERQ-   77
mBtk     1 -MAAVILESIFLKRSQQKKKTSPLNFKKRLFLLTVHKLSYYEYDFERGRRGSKKGSIDVEKITCVETVIPEKNPPPERQI   79
mItk     1 MNNFILLEEQLIKKSQQKRRTSPSNFKVRFFVLTKASLAYFEDR--HGKKRTLKGSIELSRIKCVEIVKSDIS-IPCHY-   76
mTxk       ------------------------------------------------------------------------------

                .   90        .  100        .  110        .  120        .  130        .  140        .  150        .  160
mTec    76 -----------KFPFQ-----VVHDANTLYIFAPSPQSRDRWVKKLKEEIKNNNNIMIKYRPKFWADGSYQCCRQTE  137
mBmx    77 -----------------YPFQIVYKDGLLYVYASNEESRCQWLKALQKEIRGNPHLLIKYRSGFFVDGKFLCCQQSC  137
mBtk    80 PRRGEESSEMEQISIIERFPYPFQVVYDEGPLYVFSPTEELRKRWIHQLKNVIRYNSDLVQKYRPCFWIDGQYLCCSQTA  159
mItk    76 -----------KYPFQTLVYLQVVHDNYLLYVFAPDCESRQRWVLTLKEETRNNNSLVSKYRPNFWMDGRWRCCSQLE  143
mTxk     1 -----------------------------------------------------------MILSSYSSFQSVLCCCCRCSV   22

                .  170        .  180        .  190        .  200        .  210        .  220        .  230        .  240
mTec   138 KLAPGCEKYN---------LFESSIRKTLPPA-----PEI-----KKRRPPPPIPPEEENTEEIVVAMYD-----FQATE  193
mBmx   138 KAAPGCTLWE-AYADLHIAISDEKHRAPTFPERLLKIPRAVPVLKMDASSSGAILPQYDSYSKKSCGSQPTSNIRYIPRE  216
mBtk   160 KNAMGCQILENRNGSLKPGSSHRKTKKPLPPT-----PEEDQILKKPLPPEPTAAPISTTELKKVVALYD-----YMPMN  229
mItk   144 KPAVGCAPYD---------PSKNASKKPLPPT-----PED-----NRRS-------FQEPEETLVIALYD-----YQTND  192
mTxk    23 QKRQVRTQISLSREEELSEKHSQRQRPWFAKLMGKTQSNRGGVQPSKRKPLPPLPQEPPDERIQVKALYD-----FLPRE   97

                .  250        .  260
mTec   194 AH----DLRLERGQEYIILE  209
mBmx   217 DCPDWWQVRKLKSEEDIACS  236
mBtk   230 AN----DLQLRKGEEYFILE  245
mItk   193 PQ----ELALRCDEEYYLLD  208
mTxk    98 PG----NLALKRAEEYLILE  113
```

B

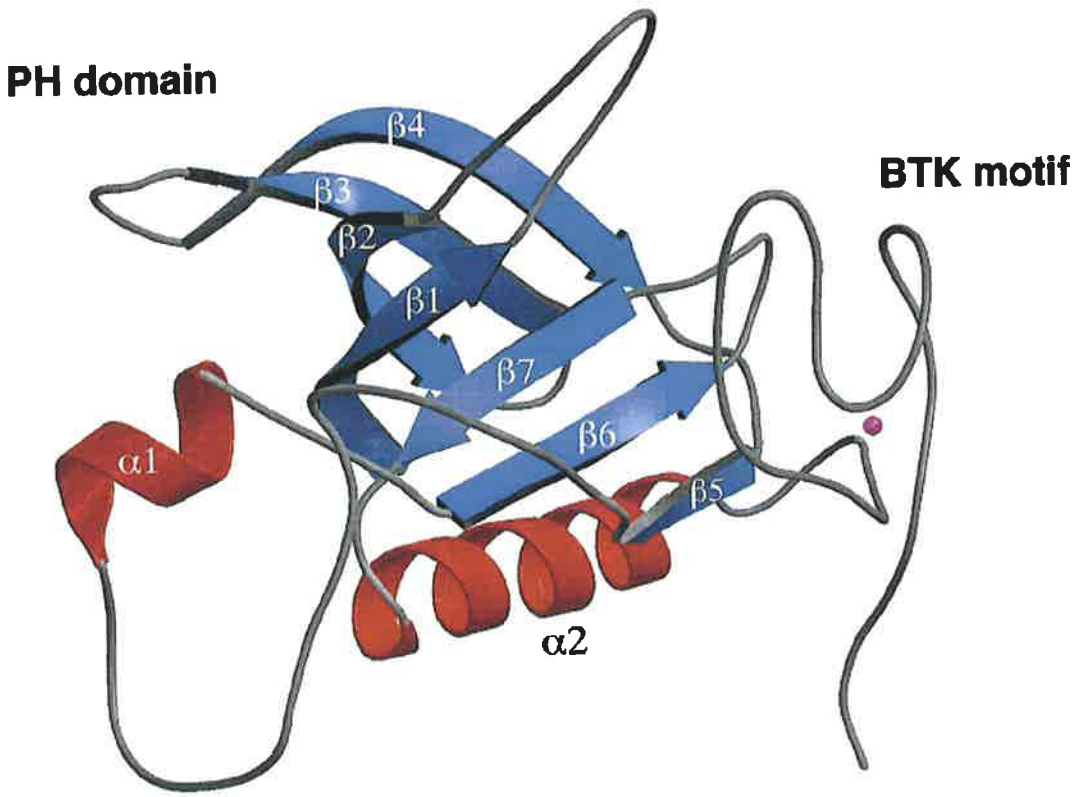


Figure 3.2

A. Sequence of the Btk PHTH region and mutations found to map to this region in the X-linked disease X-linked agammaglobulanemia. The amino acids indicated above the sequence are the mutations detected in XLA patients. Amino acids indicated with # are sites of deletions, @ indicates an insertion site and X is a nonsense mutation. Obtained from BTKbase (www.uta.fi/imt/bioinfo/BTKbase/).

B. Ribbon diagram of the PH-Btk domain from Btk (I) A ribbon representation of the Btk PH domain and Btk motif showing residues mutated in XLA patients as ball-and-stick models. The mutated amino acids are labelled. β-Strands are shown in blue and the α-helices are shown in red. (II) Same view of the molecule as in (I), showing the residues that form the I(1,3,4,5)P_4 binding site as ball-and-stick models. Taken from Hyvonen and Saraste (1997).

A

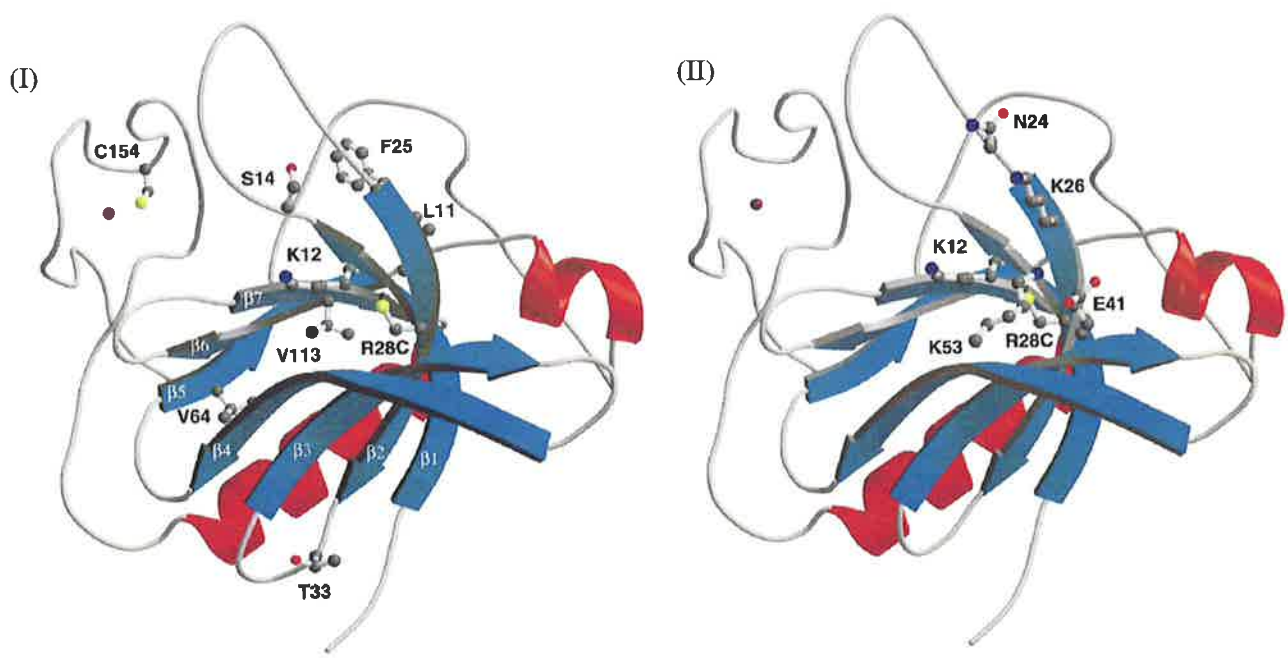
PH Domain

```
                                       P                N
     X                           X   X   H                C                               F
     # @ ##   @ PRXFX   E   #    S RC      P     #S#       ##       @ #   # @   N# D   #   @
1    MAAVILESIF LKRSQQKKKT SPLNFKKRLF LLTVHKLSYY EYDFERGRRG SKKGSIDVEK ITCVETVVPE
     X                                   H

               X                                                      X
       #    ##@ P #            X                @#   #       XD#F P        X   H   #  #  #X
71   KNPPPERQIP RRGEESSEME QISIIERFPY PFQVVYDEGP LYVFSPTEEL RKRWIHQLKN VIRYNSDL
```

TH Domain

```
                X        R                                         P            @
          X     #     @   SG @  @      X     @#  @        ## @#   #      X#
139 VQ KYHPCFWIDG QYLCCSQTAK NAMGCQILEN RNGSLKPGSS HRKTKKPLPP TPEEDQILKK

                #       ##
201     PLPPEPAAAP VSTSE
```

B



observed when the TH domain is present (Touhara *et al.*, 1994). Thus, the Btk motif is also likely to play a role in the interaction of Tec family members with Gβγ. Another Tec family member, Itk, has also been shown to be activated in response to Gβγ suggesting a more general role for Gβγ in the activation of Tec family members (Langhans-Rajasekaran *et al.*, 1995).

The PH domain of Btk interacts with filamentous actin (F-actin), Protein kinase Cμ, BAP35 and phosphoinositols. Btk binds to F-actin *in vitro* and *in vivo* through an interaction that requires basic amino acids towards the N-terminus of the Btk PH domain (Yao *et al.*, 1999). Amino acids 11-20 of the PH domain represent the minimum actin binding site. Competition studies indicate that F-actin can compete with PIP_2 for binding to the PH domain but F-actin has no effect on PKC or Gβγ binding. These data suggest that F-actin may play a role in localisation of Btk to the membrane (Yao *et al.*, 1999). The Btk PH domain has been shown to bind protein kinase C with an affinity of 39 nM, through a region incorporating the β2-β3 strands of the PH domain. This region of the PH domain also binds PIP_2 and competes for binding with PKC (Yao *et al.*, 1997). BAP35 is a binding partner of Btk and is phosphorylated by Btk following B-cell receptor (BCR) activation, suggesting that BAP35 is downstream of Btk in BCR signalling (Yang and Desiderio, 1997). BAP35 binds to the PHTH region of Btk, yet the PH domain alone retains some binding capacity for BAP35 (Yang and Desiderio, 1997). Yeast-2-hybrid screening of a liver cDNA library has revealed an interaction of Tec PHTH with the cytoskeletal crosslinking protein α actinin-4. This further consolidates the cytoskeletal link of Tec kinases through the PHTH domain (Merkel and Booker (2000), unpublished results).

Several PH domains have the ability to bind $I(1,3,4,5)P_4$. The interaction between Btk PH domain and $I(1,3,4,5)P_4$ occurs with an affinity of 40 nM as shown by isothermal titration calorimetry (Baraldi *et al.*, 1999). Although the overall structure of the PH domain does not alter significantly with binding to the lipid headgroup, the position of the β1-β2 loop is slightly shifted as it is directly involved in binding to the ligand. In the complex this loop can be modelled but, in the absence of legend, there was partial disorder in the electron density map suggesting that it is more stable in the complex. Interactions also occur with the β3-β4 loop and key residues Lys 12 and Arg 28 coordinate the ligand (Baraldi *et al.*, 1999). These residues interact with the 3′- and 4′- phosphates whereas amino acids on the β1-β2 loop are involved in coordinating the 5′- phosphate. Arginine 28 contacts the 3′-phosphate through two hydrogen bonds, which bury the 3′- and 4′ phosphates in the cavity formed by loops β1-β2 and β3-β4 (Figure 3.2B). PI3K is a key regulator of Btk function through the

production of PI(3,4,5)P_3 (Baraldi *et al.*, 1999). An activating mutation of Btk (E41K) results in constitutive membrane association possibly because this protein interacts with two PI(1,3,4,5)P_3 molecules in the crystal lattice. This suggests that activation is a result of membrane association. This interaction serves to recruit Btk to the membrane where activation of Btk promotes growth and differentiation of B cells.

The PH domains of other Tec family members have also been studied, albeit in less detail than Btk. Itk, a T cell specific Tec family member has also been shown to bind and be phosphorylated by PKC, although the mechanism by which this occurs is not yet clear (Kawakami *et al.*, 1995). Laser scanning confocal microscopy demonstrated a direct interaction between T cell receptor/ Complement determinant 3 (TCR/CD3) and the PH domain of Itk. A deletion mutant of Itk lacking the PH domain no longer binds to the receptor, indicating that the PH domain of Itk is indispensable for membrane localisation and subsequent tyrosine activation of the kinase (Ching *et al.*, 1999). Membrane localisation of Itk is also likely to be mediated by products of PI3K. August and co-workers provided evidence that the PH domain of Itk binds inositol phosphates phosphorylated at the 3′ position (August *et al.*, 1997). The interaction with 3′ inositol phosphates is crucial for membrane association and kinase activation.

Immunoprecipitation studies and subsequent kinase assays have shown that Tec tyrosine kinase binds several proteins via its PH domain and that these interactions result in Tec kinase activation. Tec kinase associates with c-Kit and is tyrosine phosphorylated following stem cell factor binding. This association takes place via a region incorporating the PHTH domain of Tec kinase (Tang *et al.*, 1994). Tec also interacts with the SH3 domain of the Src family kinase Lyn (Mano *et al.*, 1994) and with Vav through the TH region (Machide *et al.*, 1995). This places Tec kinase downstream of a variety of receptors including the erythropoietin receptor and the interleukin 3 receptor (Machide *et al.*, 1995); (Mano *et al.*, 1995). Like other Tec family members, Tec has been shown to bind PI(3,4,5)P_3, which implies that Tec kinase is also downstream of PI3K (Yang *et al.*, 2001). There is also evidence that p110, the 110 kDa subunit of PI3K, and Tec family members act in synergy to increase calcium flux in the cell. This requires the PH domain of Tec family members and may suggest that Tec activates PLC-γ to induce calcium flux (Scharenberg *et al.*, 1998). By analogy with Btk, it appears reasonable that Tec kinase would also bind to protein kinase C.

3.1.2 XLA and other disease states associated with mutations in PH domains

X-linked agammaglobulinemia (XLA), a disease state in which patients lack circulating B cells and therefore have impaired immune defence systems, is due to a lack of Btk function. Btk inactivation can be attributed to, among other things, an inability of Btk to relocate to the membrane. Binding to phospholipids plays an important role in this relocation. XLA mutations that map to the PHTH region of Btk can be categorised into structural mutants that disrupt the overall fold of the domain, and functional mutants that alter ligand binding sites (Figure 3.2A). Figure 3.2A shows the sequence of Btk PHTH domain and highlights the different amino acids involved in XLA. Significantly, several functional mutations map to the phospholipid-binding site. For example, substitution of Arg 28 accounts for a large number of XLA causing mutations, manifest by an inability to bind the 3′-phosphate of $PI(3,4,5)P_3$. All the mutations remove the positive charge that would bind the 3′-phosphate. Other mutations in the binding pocket that substitute larger amino acids, such as K12R and S14F, are likely to hinder ligand binding by steric interference. Thus, the binding of Btk to phospholipids is crucial for the function of Btk. The mutations have been mapped onto the structure of Btk PHTH and are shown in Figure 3.2B.

Wiskott Aldrich Syndrome (WAS), a human heritable disease, is caused by mutation of the PH domain in the Wiskott Aldrich Syndrome protein (WASP). This protein induces formation of actin clusters and, thus, regulates actin cytoskeleton (Imai *et al.*, 1999). Introduction of a missense mutation in the WASP PH domain results in a loss of actin cluster formation. This mutant also lacks the ability to bind to phosphatidylinositol 4,5-bisphosphate (PIP_2), a key regulator in actin re-arrangement and the authors suggest that the loss of binding to PIP_2 results in a decrease in protein activity. High proportions of WASP mutations map to the PH domain suggesting that PIP_2 binding may be the critical step in activation of actin re-arrangement (Imai *et al.*, 1999). Reduced PH domain-PIP_2 binding efficiency is also observed with a construct containing a C38W mutation in WASP. These results highlight the importance of the WASP PH domain's ability to bind PIP_2 and actin for membrane retention (Miki *et al.*, 1996).

30-45 fold overexpression of Akt (PKB) has been reported in some human ovarian cancers. Several PH domain-containing proteins including Vav, Dbl, Ost and Ect-2 have been shown to have oncogenic transforming ability. The significance of the PH domain in cellular transformation remains unclear as these proteins also contain adjacent Dbl domains. It is not known if one or both of these domains are responsible for oncogenic transformation (Shaw, 1996). Several human diseases have been linked to potential HIKE (HIK(X8)E)

motifs including acute mast cell and myeloid leukemia's and Hirschsprung disease. These diseases have mutations in the HIKE motif, which has been suggested to bind PH domains. The inability of the PH domain to bind to the HIKE motif may perturb the downstream signalling pathways from the proteins containing the HIKE motif (Ciccarelli *et al.*, 2000).

3.2 AIMS

The PHTH region is likely to be important for the proper function of Tec kinase given that mutations in the Btk PH domain have been linked to Btk function and XLA. Since it was not possible to investigate the structure of the whole protein, a reductionist approach was taken. The PHTH region of Tec kinase was investigated in isolation with the aim of relating this data to the whole enzyme.

Millimolar protein concentrations are required for NMR experiments and, thus, high-level protein production is required. Several different combinations of Tec kinase N-terminal domains were investigated to achieve the best induction and expression. Figure 3.3 summarises the different PHTH proteins investigated and the relative number of amino acids in each. At the onset of this project, there was very little structural information about the TH domain so the expression and further investigation of the TH domain was the primary aim. Figure 3.4 shows a flow chart and the experiments conducted on the different PHTH containing proteins of Tec kinase. Although structural determination of the PHTH domain was the ultimate aim of this work, this chapter will present only the expression, purification and preliminary biophysical characterisation of different Tec PHTH proteins. Low level soluble expression, limited refolding success and overall low solubility of the proteins limited the progress of this aim.

Figure 3.3

A schematic representation of the different proteins used in this study. The sequence of mouse Tec containing the PHTH, PRR and SH3 domain is shown. The PH domain and the TH region are coloured blue and pink, respectively. The PRR and the SH3 domain are coloured red and green, respectively. The location within Tec kinase, the start and finish of the sequence, the number of amino acids and the calculated molecular weights of each protein tested are shown.

50

MNFNTILEEI LIKRSQQKKK TSLLNYKERL CVLPKSVLSY YEGRAEKKYR

100

KGVIDISKIK CVEIVKNDDG VIPCQNKFPF QVVHDANTLY IFAPSPQSRD

150

RWVKKLKEEI KNNNNIMIKY HPKFWADGSY QCCRQTEKLA PGCEKYNLFE

200

SSIRKTLPPA PEIKKRRPPP PIPPEEENTE EIVVAMYDFQ ATEAHDLRLE

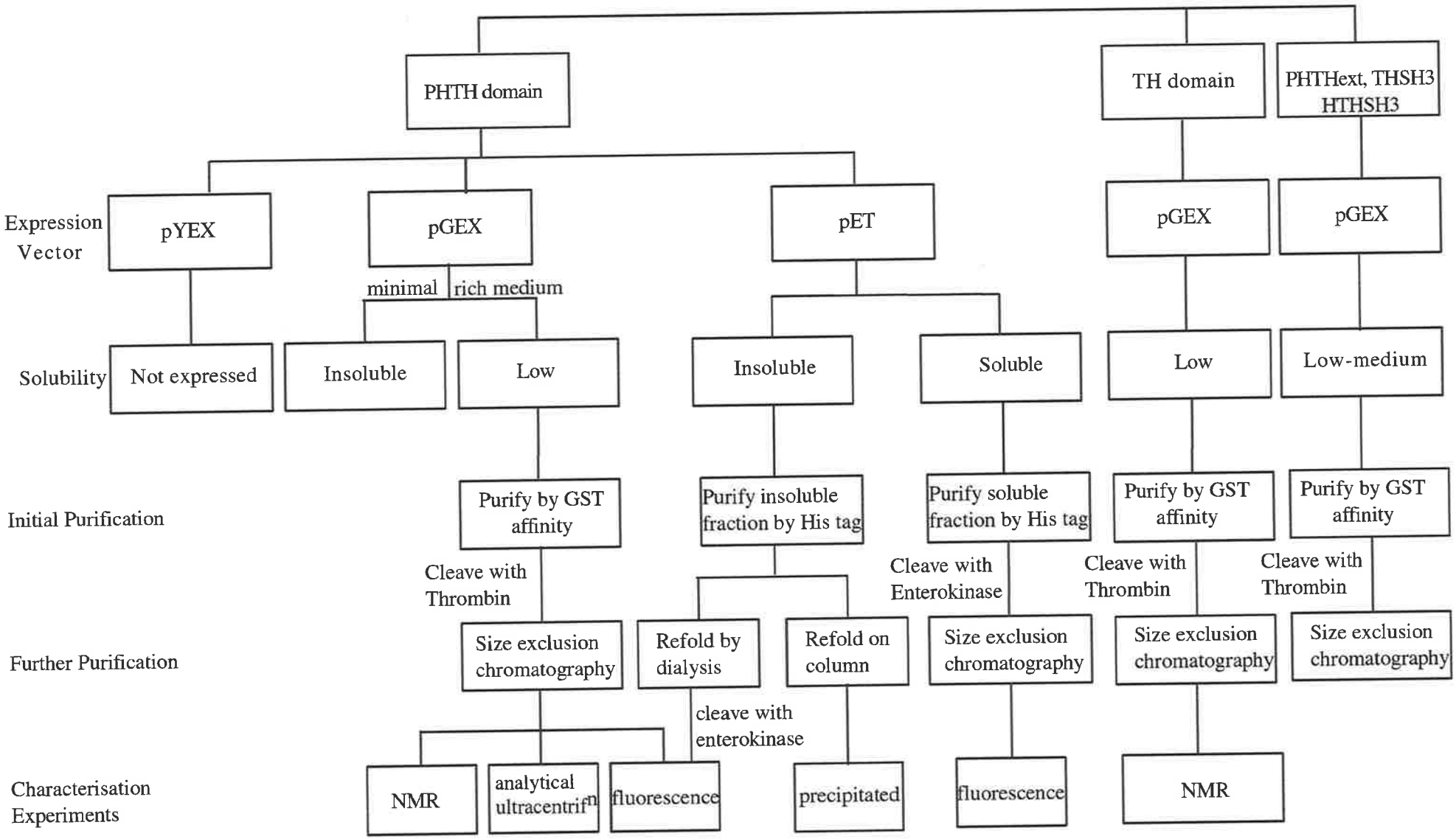
244

RGQEYIILEK NDLHWWRARD KYGSEGYIPS NYVTGKKSNN LDQYE

Protein	Sequence	No of amino acids	Molecular weight (Da)
PHTH 	^{1}MNFN.....FESS^{152}	152	181774
PHTHext 	^{1}MNFN.....ENTE^{180}	180	21170
PH 	^{1}MNFN.....IKNN^{111}	113	13289
TH 	^{117}MIKY.....FESS^{152}	35	4259
THSH3 	^{111}NNNN.....DQYE^{244}	133	15792
HTHSH3 	^{96}PQRS.....DQYE^{244}	148	17815

Figure 3.4

Flow chart showing the experiments conducted on the different PHTH proteins. The proteins are shown at the top of the chart and the different experimental stages are highlighted through the chart. The first level represents the vector system used for protein expression. The second level down is the induction and solubility assessment followed by the purification procedure in two stages and finally other biochemical analyses conducted on these proteins.


PHTH domain
TH domain
PHTHext, THSH3
HTHSH3
Expression Vector
pYEX
pGEX
pET
pGEX
pGEX
minimal
rich medium
Solubility
Not expressed
Insoluble
Low
Insoluble
Soluble
Low
Low-medium
Initial Purification
Purify by GST affinity
Purify insoluble fraction by His tag
Purify soluble fraction by His tag
Purify by GST affinity
Purify by GST affinity
Cleave with Thrombin
Cleave with Enterokinase
Cleave with Thrombin
Cleave with Thrombin
Further Purification
Size exclusion chromatography
Refold by dialysis
Refold on column
Size exclusion chromatography
Size exclusion chromatography
Size exclusion chromatography
cleave with enterokinase
Characterisation Experiments
NMR
analytical ultracentrifn
fluorescence
precipitated
fluorescence
NMR

3.3 EXPRESSION AND PURIFICATION OF THE PHTH REGION

3.3.1 PHTH fusion protein expression

Protein structure determination by NMR spectroscopy is limited largely by spectral resolution. Overlaps of peaks in NMR spectra make assignment of each peak to a certain pair of protons difficult (section 4.2). Isotopic labelling of the protein of interest facilitates the use of three-dimensional and even four-dimensional NMR experiments capable of providing enhanced resolution thereby separating otherwise overlapping peaks. Currently, the structure of proteins of less than 100-120 amino acids can be determined without the need for isotopic labelling (Roberts, 1993). Tec PHTH domain is 152 amino acids and therefore requires labelling to minimise spectral overlap.

A cDNA encoding amino acids 1-152 of the mouse Tec IV sequence was PCR amplified and cloned into pGEX-2T, pET and pYEX expression vectors using 5′ *Bam*HI and 3′ *Eco*RI restriction sites (Figure 3.5). This region corresponds to the PH domain and the adjacent Btk motif but does not include the PRR. This protein will be referred to as the Tec PHTH domain. Recombinant plasmids were manually sequenced using appropriate sequencing primers to ensure expression of the correct protein (section 2.2.7, 2.3.13) (not shown). pGEX-2T-PHTH and pET-PHTH plasmids were transformed into BL-21λDE3 *E. coli*. The BL-21 strain of *E. coli* is commonly used for recombinant protein expression because it does not contain the *Ion* protease or the *ompT* outer membrane protease making protein degradation less likely (Grodberg and Dunn, 1988). The DE3 host contains a chromosomal copy of the gene for T7 polymerase for better regulation of protein induction. During the work, various expression systems were analysed in an attempt to increase recombinant protein yields (Figure 3.4). Figure 3.4 shows a flow chart of the different expression systems and constructs used in this work.

The glutathione-S-transferase (GST) fusion partner used in pGEX plasmids aids high protein expression, increased yields of soluble protein and simplifies purification of the protein. Glutathione agarose chromatography allows for one step purification of GST-fusions from the bulk of the other *E. coli* proteins.

pGEX-PHTH transformed bacteria induced to express the GST-fusion protein were analysed by SDS-PAGE to determine expression and solubility of GST-PHTH protein. Production of GST-PHTH was compared between bacteria grown in shakeflasks and

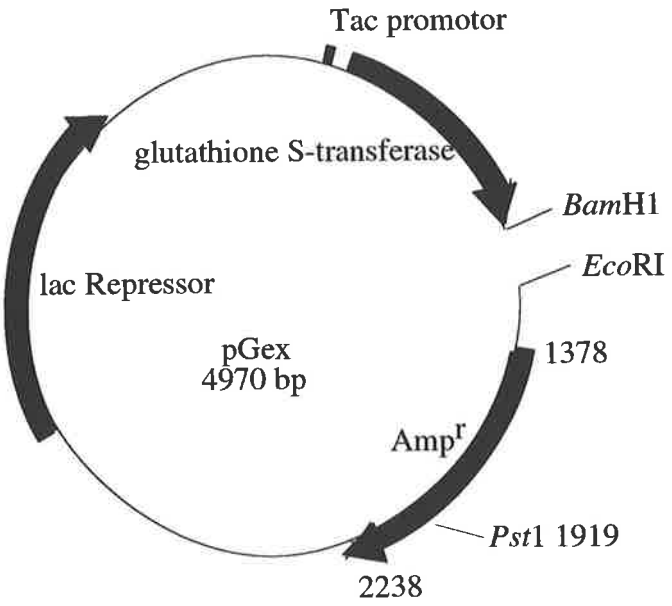
Figure 3.5

Vector maps of the pGEX-2T and pET-32a plasmids used to express PHTH proteins.

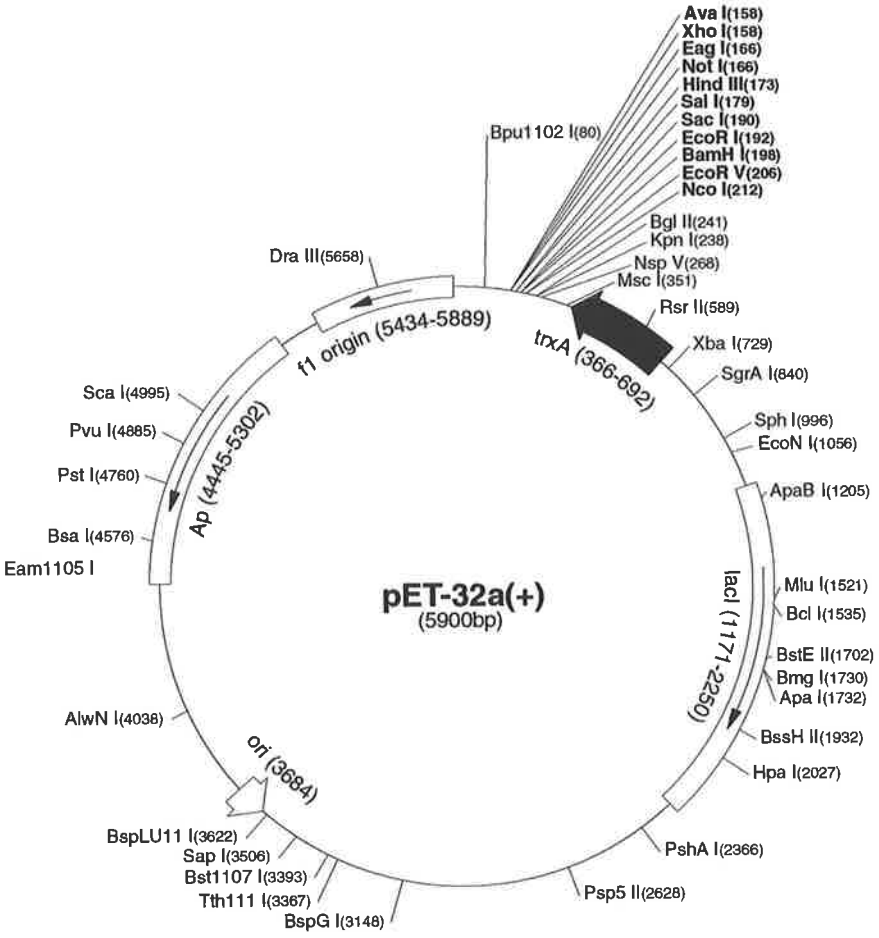
A. pGEX-2T contains an ampicillin resistance gene, a lac repressor cassette and the Tac promoter. The multiple cloning site is located 3′ to the GST coding region to facilitate in-frame production of the PHTH fusion proteins.

B. The pET plasmid contains the ampicillin resistance gene, the lacI cassette and the T7 promoter. There are N and C terminal histidine tags in the pET vector, however a 3′ stop codon in the PHTH insert prevents translation of the C-terminal tag. pET-32a vector also contains the gene for Trx protein production 5′ to the multiple cloning site. Both plasmids were digested with *Bam*HI and *Eco*RI to facilitate cloning of the PHTH PCR products generated.

A


Tac promotor
glutathione S-transferase
BamH1
EcoRI
lac Repressor
pGex
4970 bp
1378
Amp^r
Pst1 1919
2238

B


Ava I(158)
Xho I(158)
Eag I(166)
Not I(166)
Hind III(173)
Sal I(179)
Sac I(190)
EcoR I(192)
BamH I(198)
EcoR V(206)
Nco I(212)
Bpu1102 I(80)
Bgl II(241)
Kpn I(238)
Nsp V(268)
Msc I(351)
Rsr II(589)
Dra III(5658)
f1 origin (5434-5889)
trxA (366-692)
Xba I(729)
SgrA I(840)
Sca I(4995)
Pvu I(4885)
Pst I(4760)
Ap (4445-5302)
Sph I(996)
EcoN I(1056)
ApaB I(1205)
Bsa I(4576)
Eam1105 I
pET-32a(+)
(5900bp)
lacI (1171-2250)
Mlu I(1521)
Bcl I(1535)
BstE II(1702)
Bmg I(1730)
Apa I(1732)
AlwN I(4038)
ori (3684)
BssH II(1932)
Hpa I(2027)
BspLU11 I(3622)
Sap I(3506)
Bst1107 I(3393)
Tth111 I(3367)
BspG I(3148)
PshA I(2366)
Psp5 II(2628)

fermenters, 2YT and LB, 37°C and 30°C and IPTG to a final concentration of 0.1 mM and 0.2 mM.

Bacteria were inoculated into fermenters or flasks and recombinant protein expression was induced by addition of IPTG to a final concentration of 0.2 mM once fermenter cultures had reached OD_{600} of 4.0 and shakeflasks had reached OD_{600} of 0.6 (section 2.4.1). Whole cell lysates from bacteria grown in 0.2 mM IPTG for a minimum of 4 hours contained a band of approximately 45 kDa that was not observed in the absence of IPTG (Figure 3.6A (I)). This size is consistent with the expected size for GST-PHTH fusion protein. Following lysis of the *E. coli* extracts from shakeflasks and fermenters, SDS-PAGE of the soluble and insoluble fractions indicated the majority of GST-PHTH protein produced was in the insoluble fraction (Figure 3.6A (II)). These bacteria contained inclusion bodies as observed by microscopy (data not shown).

Soluble protein expression in shakeflasks was compared between 2YT and LB media. Bacteria grown in 2YT produced more recombinant protein, however, a greater proportion of GST-PHTH protein was insoluble (P) compared with the LB cultured bacteria (Figure 3.6B). This is possibly because the 2YT media is richer in nutrients than the LB media and the rate of protein production is too great for the folding machinery of the cell. Cultures were grown at 30°C and 37°C with addition of IPTG to a final concentration of 0.1 mM and 0.2 mM in order to establish conditions that resulted in greater proportions of soluble protein production. No improvement in soluble protein production was observed in bacteria grown at either lower temperatures or at different IPTG concentrations (Figure 3.6C (I) and (II)). Optimisation experiments did not produce significantly higher yields of soluble GST-PHTH fusion protein. As a result, larger culture volumes (5 L) were required to produce enough soluble protein for NMR spectroscopy experiments. Expression of GST-PHTH protein was also assessed in minimal media, however, soluble expression was negligible as visualised by SDS-PAGE and it was not pursued (Figure 3.7A).

Like pGEX, the pET expression vector contains a fusion partner to promote the production of soluble protein. In the case of pET, this fusion protein is Thioredoxin (Trx). Both N- and C-terminal 6-histidine tags are also present to facilitate purification using nickel affinity purification, however, only the N-terminal histidine tag is present in the PHTH protein as there is a stop codon at the C-terminus of the PHTH insert preventing translation of the histidine tag sequence. The pET bacterial expression system has the advantage of allowing purification of the histidine tagged protein in a denatured state. Thus, yields could

Figure 3.6

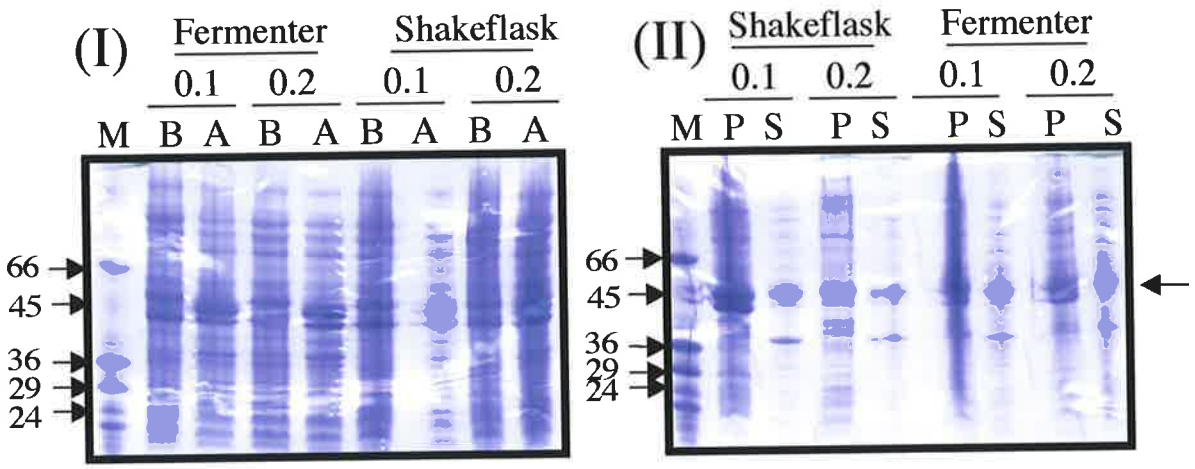
A. (I) Comparison of GST-PHTH protein expression in shakeflasks and fermenters. Shakeflasks were inoculated at 1/100 dilution from an overnight culture and grown to an absorbance of A_{600} 0.6. 1 ml samples were removed at both pre-induction (B) and post-induction stages (A). Two IPTG concentrations were tested, 0.1 mM and 0.2 mM. Fermenters were inoculated from an overnight culture with 200 µl and the culture grown to an OD_{600} of 4 at which it was induced with IPTG to 0.2 mM. Following electrophoresis on 12.5% SDS-PAGE gels, proteins were visualised by Coomassie staining. SDS-7 markers were used for protein size estimation. Sizes of the markers are indicated.

(II) Solubility of expressed GST-PHTH protein generated from bacteria cultured in shakeflasks and fermenters. Following induction the bacterial pellet was resuspended in TTBS and sonicated in the presence of 1 mM PMSF. Samples were run on a 12.5% SDS-PAGE gel and visualised by Coomassie blue staining. The soluble and the insoluble fractions are indicated by S and P respectively.

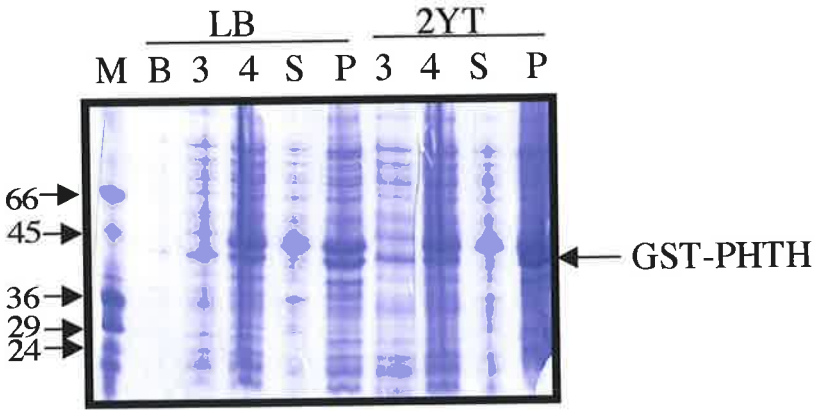
B. Solubility of GST-PHTH protein expressed in LB medium and 2YT medium as indicated. Samples were removed just prior to induction (3) and following induction at 4 hours (4) and run on a 12.5% SDS-PAGE gel. A sample was also removed from the LB culture 1 hour before induction (B). The soluble fraction is labelled S and the insoluble fraction is labelled P. Visualisation was by Coomassie staining. Sizes of the markers are indicated.

C. Comparison of GST-PHTH protein expression at 30°C and 37°C and with induction using 0.1 mM and 0.2 mM IPTG. Samples were electrophoresed on a 12.5% SDS-PAGE gel and visualised by Coomassie staining. A is after induction, B is before induction shown on gel (I), P is the insoluble fraction and S is soluble fraction shown on gel (II).

A


(I)
Fermenter
Shakeflask
0.1
0.2
0.1
0.2
M B A B A B A B A
66
45
36
29
24
(II)
Shakeflask
Fermenter
0.1
0.2
0.1
0.2
M P S P S P S P S
66
45
36
29
24

B


LB
2YT
M B 3 4 S P 3 4 S P
66
45
36
29
24
GST-PHTH

C

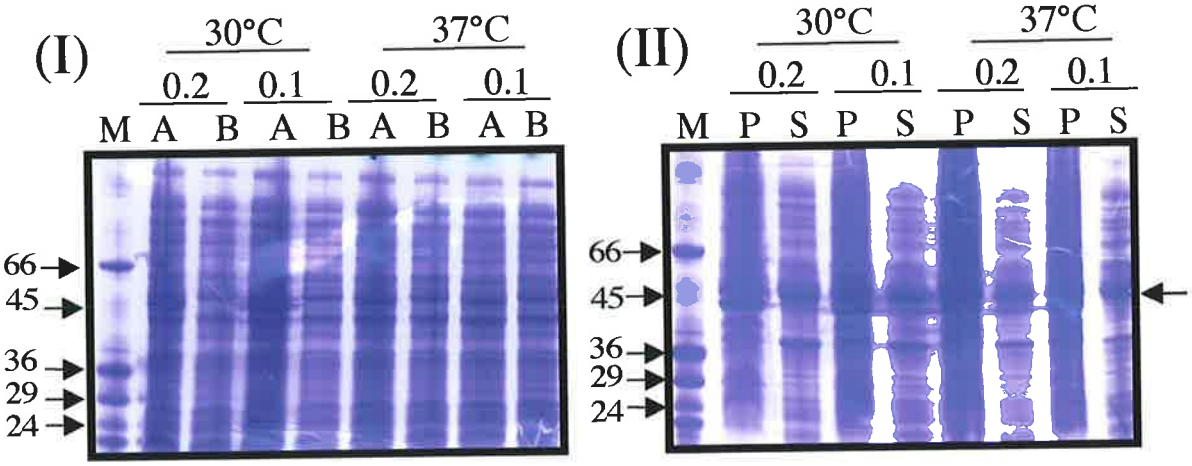
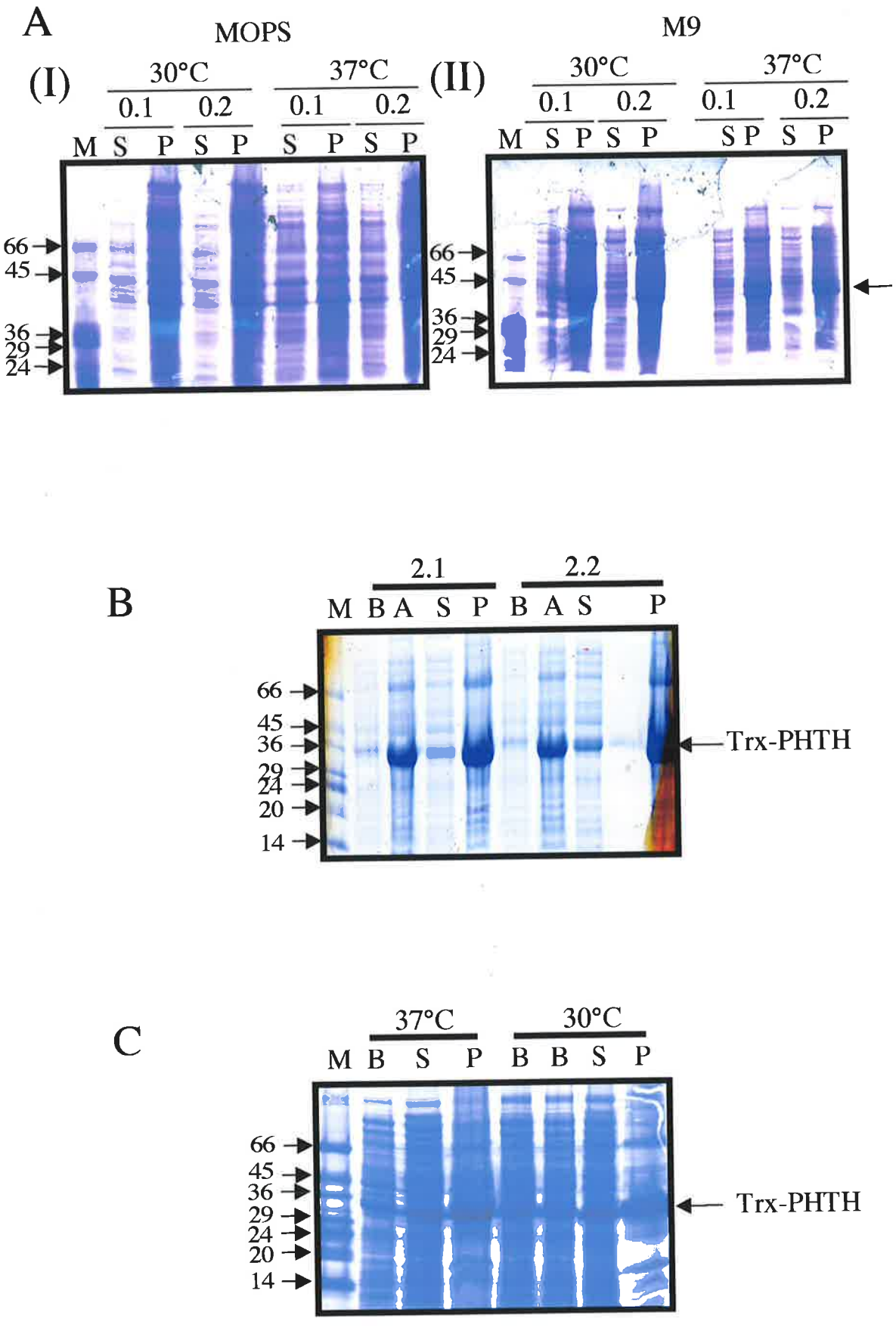

(I)
30°C
37°C
0.2
0.1
0.2
0.1
M A B A B A B A B
66
45
36
29
24
(II)
30°C
37°C
0.2
0.1
0.2
0.1
M P S P S P S P S
66
45
36
29
24

Figure 3.7

A. 12.5% SDS-PAGE analysis of the expression and solubility of GST-PHTH grown under minimal media culture conditions. Two different media preparations were investigated, MOPS (I) and M9 (II). Both cultures were induced for 4 hours after induction with IPTG. Both preparations were optimised for temperature (30°C or 37°C) and IPTG concentration (0.1 mM or 0.2 mM). The soluble fraction (S) and the insoluble fraction (P) were assessed in all cases. The samples were compared to SDS-7 protein markers for estimation of protein size. The induced band is indicated with an arrow.

B. Expression of pET-PHTH. Cultures (colonies 2.1 and 2.2) were grown at 37°C in LB media and induced with 1 mM IPTG. Samples were removed before (B) and after induction (A) pelleted and resuspended in 1x load buffer. The soluble (S) and insoluble fractions (P) following lysis by sonication are also shown on the gel. Samples were loaded on a 12.5% SDS-PAGE gel, electrophoresed and stained with Coomassie blue. For size comparison SDS-7 markers were also loaded and are indicated on the gel.

C. SDS-PAGE analysis of soluble Trx-PHTH protein production following lysis by french pressing. Two different temperatures were analysed as indicated above the gels and before induction samples analysed (B) and the after induction sample analysed for the 30°C sample (A). Also shown on the gel is the supernatant (S) and pellet following lysis (P). Samples were loaded on a 12.5% SDS-PAGE gel, electrophoresed and stained with Coomassie blue.


A
MOPS
M9
(I)
(II)
30°C
37°C
0.1
0.2
M S P S P S P S P
M S P S P S P S P
66
45
36
29
24
B
2.1
2.2
M B A S P B A S P
66
45
36
29
24
20
14
Trx-PHTH
C
37°C
30°C
M B S P B B S P
66
45
36
29
24
20
14
Trx-PHTH

be increased if the PHTH protein could be purified from inclusion bodies using the pET bacterial expression system.

Expression levels and solubility were assessed as for pGEX protein expression with the exception that cultures were induced with a final concentration of 1 mM IPTG (section 2.4.2). Figure 3.7B shows the expression and solubility of two independent Trx-PHTH clones, 2.1 and 2.2. Trx-PHTH fusion protein expression was induced in bacteria grown in 1 mM IPTG; however, the majority of the protein was present in the insoluble fraction after lysis (Figure 3.7C). Yields of soluble protein were equivalent to that achieved using the pGEX system. Growth of bacteria at 30°C did not improve soluble protein yields compared with cells grown at 37°C (Figure 3.7C).

Expression of recombinant PHTH protein using a yeast expression system was attempted to circumvent problems of insolubility and inclusion body formation. Yeast use different groups of chaperones and fold proteins co-translationally rather than post-translationally (Hartl, 1996). Correct post-translational modification has also been reported in yeast in contrast to bacteria (Mitchell *et al.*, 1993).

SDS-PAGE conducted on yeast lysates (harbouring pYEX-PHTH plasmid) induced with copper sulfate to produce GST-PHTH, showed no significant induced band of the expected size (not shown). Variation of the induction time and copper sulfate concentration did not improve induction. This system was not pursued.

The solution structure of the Btk PH domain and Btk motif, published during these studies, indicated that the Btk motif chelates a Zn atom through a motif containing three cysteines and a histidine (Hyvonen and Saraste, 1997). Soluble yields of GST-PHTH protein were analysed with the addition of zinc. GST-PHTH protein was generated under the same conditions, with the exception that 30 μM $ZnSO_4$ was added in conjunction with the chemical inducer IPTG. Polyacrylamide gels containing supernatant and pellet fractions from the lysed *E. coli* showed a slight increase in yield of soluble recombinant protein (not shown). However, this increase was only small and further purification of this protein was not attempted. Addition of zinc may increase the stability of the Btk motif and, thus, the overall stability of the protein. This is consistent with an increase in yields of THSH3 and TH domain fusion proteins grown in the presence of zinc (section 3.5.2).

3.3.2 Purification PHTH proteins from pGEX and pET

Two methods were investigated for the preparation of GST-PHTH fusion proteins; 1) the GST-PHTH protein was eluted from the column with reduced glutathione and then digested with thrombin and 2) the GST-PHTH protein was cleaved with thrombin while still attached to the column resin (section 2.4.9). Thrombin digestion of column-purified fusion protein resulted in precipitation of the PHTH protein such that very little soluble protein was observed by SDS-PAGE. Affinity chromatography and digestion while still attached to the column improved yields with 11 mg PHTH protein per litre LB culture produced. Figure 3.8A shows a profile of the purification of Tec PHTH domain. Elution of the PHTH was facilitated by application of a thrombin mixture overnight, followed by elution 16 hours later with TBS (section 2.4.9). The GST component was eluted with 10 mM reduced glutathione in 1 x TBS pH 8.0. The gel shown in Figure 3.8B shows the purification of the PHTH protein. There is no obvious fusion protein in the flow through suggesting all the GST-PHTH bound the column. The PHTH sample was then concentrated using an Amicon stirred cell (Amicon) and buffer exchanged using a PD10 fast desalting column (Pharmacia) equilibrated in 10 mM phosphate, 0.2 mM DTT and 0.02% NaN_3 pH 5.5 (section 2.4.13). Purified PHTH protein for further downstream analysis was assessed for purity using polyacrylamide gel electrophoresis and the size determined by mass spectrometric analysis shown in Figure 3.8C. The peak of molecular weight 18178 Da is consistent with the expected mass of the PHTH protein. The peak at 18252 Da may reflect a small degree of post-translational modification, however, the exact modification is not known. Column-based proteolytic cleavage provided the highest observed yields of the purified PHTH protein and although not ideal, sufficient material for further experiments was generated using this method.

The pET expression system was also used in an attempt to produce large quantities of native soluble protein from bacteria induced to express Trx-PHTH protein. The protein sample was eluted from the column using a Tris-based elution buffer containing 1 M imidazole, which competes with the 6-histidine tag for binding to the nickel (Ni) (section 2.4.5). The yields returned were low as visualised by SDS-PAGE, which also showed that the majority of the protein was eluting at the wash stage (60 mM imidazole) not at the elution stage (1 M imidazole) (not shown).

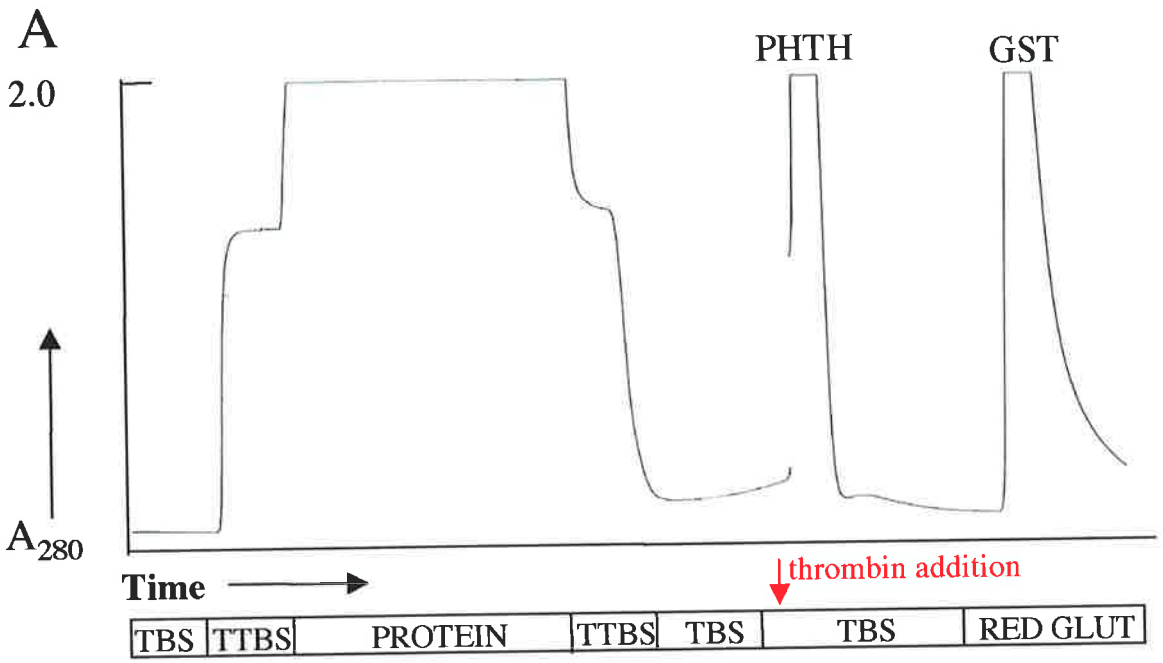
The pET expression system also allows for insoluble protein purification. To determine if folded Trx-PHTH protein could be recovered from inclusion bodies Trx-PHTH protein was purified under denaturing conditions. Inclusion bodies were purified using a Ni

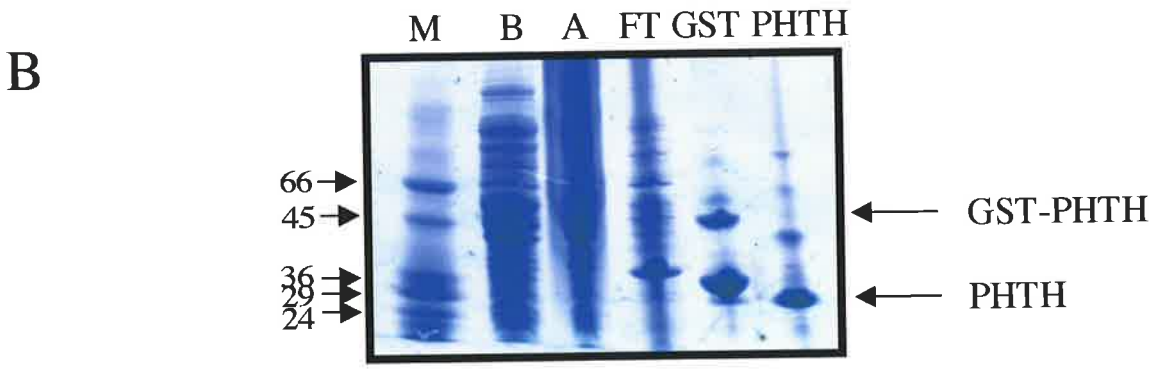
Figure 3.8

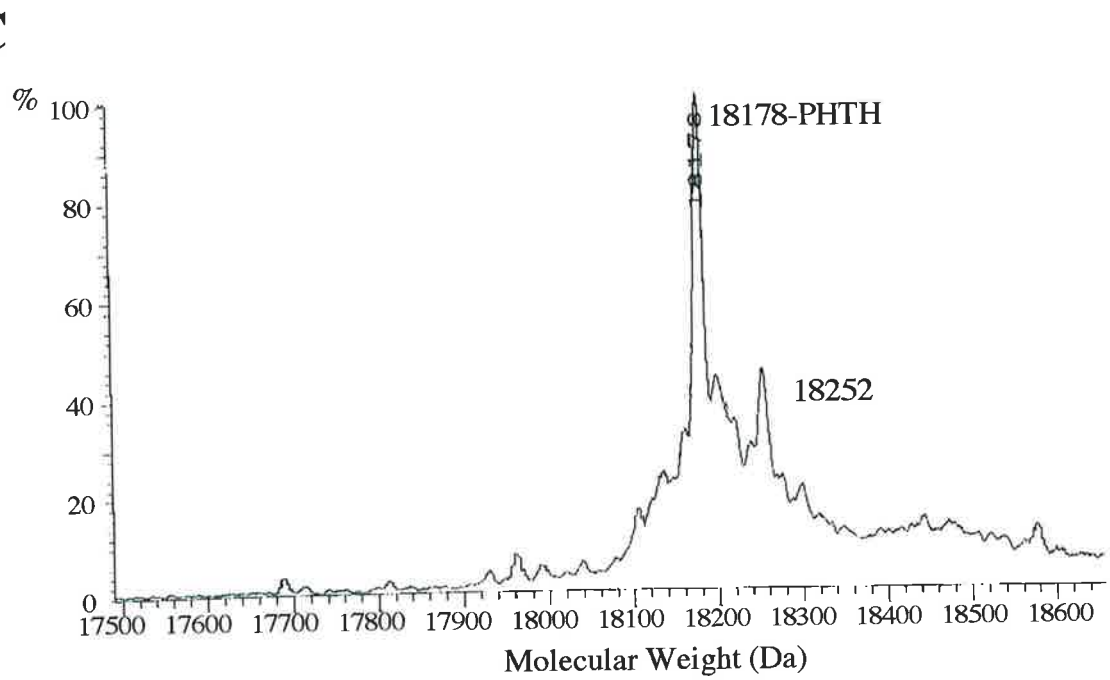
A. Glutathione Sepharose column chromatography of GST-PHTH protein. 5 L LB cultures of pGEX-PHTH were grown in shakeflasks induced with IPTG to a final concentration of 0.2 mM and pelleted. These pellets were resuspended in TTBS and sonicated. A glutathione Sepharose column was equilibrated in TBS to establish a baseline following which it was equilibrated in TTBS. Filtered supernatant was loaded on the column at flow rates of between 3-4 ml/min and then the column washed with TTBS and then TBS. The GST-PHTH protein was cleaved overnight and eluted the following morning by addition of TBS to the column. The GST was eluted from the column by TBS and 10 mM reduced glutathione pH 8. The proteins eluted at each stage are indicated above the profile.

B. SDS-PAGE analysis of GST-PHTH purification. Pre induction (B) and post induction (A) samples were generated from the 5 L pGEX-PHTH culture preparation. The flow through (FT), TBS elution (PHTH) and reduced glutathione wash (GST) were diluted with equal amounts of 5 x SDS page load buffer of which 10 µl was loaded on a 12.5% SDS-PAGE gel. SDS-7 markers were loaded for size estimation. The gel was visualised by Coomassie staining.

C. Mass spectrometry profile of purified PHTH protein. The molecular masses (Da) of the peaks are indicated above the peak. PHTH protein was purified by HPLC, lyophilised and sent for analysis at the Chemistry Department, Wollongong University, Wollongong.


A
2.0
PHTH
GST
A_{280}
Time
thrombin addition
TBS
TTBS
PROTEIN
TTBS
TBS
TBS
RED GLUT


B
M
B
A
FT
GST
PHTH
66
45
36
29
24
GST-PHTH
PHTH


C
%
100
80
60
40
20
0
18178-PHTH
18252
17500
17600
17700
17800
17900
18000
18100
18200
18300
18400
18500
18600
Molecular Weight (Da)

affinity column in the presence of 6 M urea. Elution of the denatured Trx-PHTH from the Ni affinity column also occurred at the wash stage when the buffer contained 60 mM imidazole (section 2.4.5). Figure 3.9A shows the SDS-PAGE monitoring the purification of Trx-PHTH. A significant amount of Trx-PHTH protein was recovered in the first wash stage and the protein also eluted in fractions 14 and 15. Yields of the denatured fusion protein were 15 mg/L of culture broth.

A series of experiments were conducted to refold the Trx-PHTH protein while still bound to the Ni column. This has the advantage of minimising incorrect associations since the protein is anchored at one end (Holzinger *et al.*, 1996); (Stempfer *et al.*, 1996). The binding/lysis buffer contained 20 mM Tris-base pH 7.9, 100 mM NaCl and 6 M urea but lacked imidazole. The Trx-PHTH fusion protein was eluted in two stages; the first in 50 mM imidazole, 20 mM Tris-base pH 7.9, 100 mM NaCl and the second 100 mM imidazole, 20 mM Tris-base pH 7.9, 100 mM NaCl (section 2.4.5). Again, the yields were very low, indicating that the protein had precipitated on the column. Stepwise refolding, using 6 M, 4 M, 2 M and 1 M urea increments, increased the yields to 6 mg fusion protein /L of culture broth (section 2.4.6). Polyacrylamide gels were performed to analyse the purity of the samples eluted from the column at different urea concentrations as well as at the final elution conditions (Figure 3.9B). The fusion protein eluted in fractions 4, 5, 6 and 7 although some breakdown was also observed. Further analysis of the breakdown products was not undertaken. It appeared that refolding on the column was successful, however, following short term storage, these samples precipitated from solution.

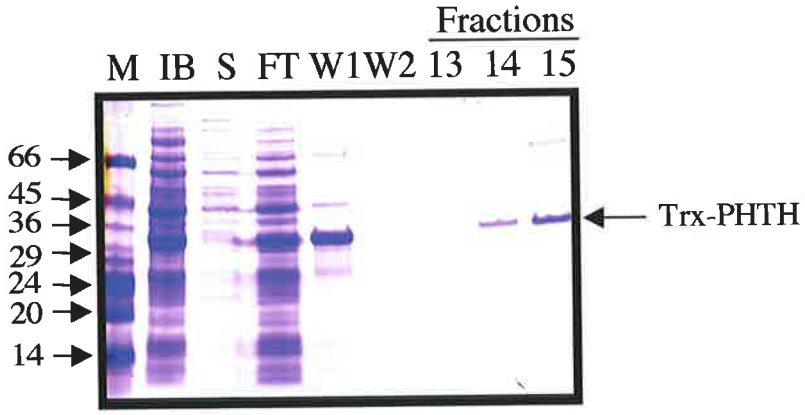
Dialysis was also investigated as a refolding method using purified denatured Trx-PHTH protein. The denatured fusion protein was purified on the Ni affinity column and eluted under denaturing conditions. Two different methods of dialysis refolding were investigated: stepwise dialysis (using several different urea concentrations) and direct dialysis (one step dialysis from 6 M urea to no urea). Both methods resulted in approximately 50% protein loss. The stepwise dialysis was conducted on two samples of different concentrations, 0.5 mg/ml and 1 mg/ml. The starting buffer was 6 M urea, 1 M imidazole, 500 mM NaCl and 20 mM Tris base pH 7.9 dialysed into a final buffer of 250 mM NaCl, 25 mM Tris base pH 7.9 and 15 mM β-mercaptoethanol with steps of decreasing concentrations of urea (section 2.4.7). Some precipitation was evident and some protein bound the membrane. Stepwise dialysis resulted in 40% and 55% protein loss giving final concentrations of 0.3 mg/ml and 0.45 mg/ml, respectively. By direct dialysis, the protein was dialysed into 150 mM NaCl, 50 mM Tris base, 1 mM EDTA and 7.5 mM DTT using

Figure 3.9

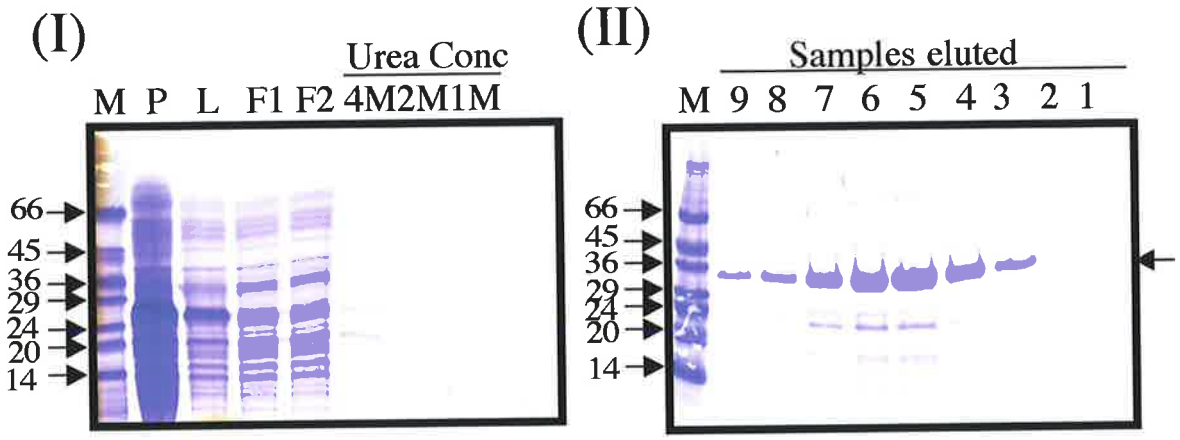
A. Purification of Trx-PHTH monitored by 12.5% SDS-PAGE. The bacteria were grown in LB medium and induced with 1 mM IPTG. Bacteria were lysed using french press and the soluble fraction (S) loaded onto a nickel affinity column equilibrated in TBS and 0.5 mM imidazole. The insoluble fraction (P) was also investigated on the gel. The flow through (FT) from the column was monitored and the two wash steps, one in 5 mM imidazole (W1) and the other with 60 mM imidazole (W2). The sample was eluted with 1 M imidazole and 1 ml fractions collected. Trx-PHTH eluted in fractions 14 and 15 as shown on the gel. Size markers were loaded for estimation of the size of the Trx-PHTH protein.

B. Purification of Trx-PHTH protein from inclusion bodies monitored by 12.5% SDS-PAGE. The sample (L) was loaded onto a nickel affinity column in 500 mM NaCl, 5 mM imidazole, 20 mM Tris base and 6 M urea. The pellet was also shown on the gel (P). The column was washed with 5 mM and 60 mM imidazole and samples taken (F1 and F2 respectively). The urea concentration was reduced while the protein was still bound to the column. 4 M, 2 M and 1 M urea concentrations were applied to the column and samples monitored (gel I). The protein was then eluted in 1 ml fractions with 250 mM imidazole, 125 mM NaCl and 20 mM Tris base and visualised on a 12.5% SDS-PAGE gel. Trx-PHTH eluted from fraction 3-9 as shown on gel II. The arrow indicates Trx-PHTH protein.

A


Fractions
M IB S FT W1 W2 13 14 15
66
45
36
29
24
20
14
Trx-PHTH

B


(I)
Urea Conc
M P L F1 F2 4M 2M 1M
66
45
36
29
24
20
14
(II)
Samples eluted
M 9 8 7 6 5 4 3 2 1
66
45
36
29
24
20
14

500 (1 ml/500 ml) fold equivalents overnight (section 2.4.7). Some protein precipitation was evident in the dialysis reaction. An initial protein concentration of 0.5 mg/ml resulted in 60% of the protein lost giving a final concentration of 0.2 mg/ml. Refolding was monitored by fluorescence spectroscopy (3.4.1). Refolding experiments were also conducted with the addition of polyethylene glycol (PEG) to assist in refolding. Purification of Trx-PHTH under denaturing conditions did not improve the yield of soluble protein due to precipitation during the refolding procedures.

The pET vectors contain two different protease digestion sites, thrombin and enterokinase. Various enterokinase concentrations were assessed to cleave the fusion protein from the PHTH domain (section 2.4.10). As expected the rate of digestion was dependent on the amount of enterokinase added but even at a mass ratio of 1/20 the enzyme failed to completely digest the Trx-PHTH protein within 5 hours (Figure 3.10A). The protein bands evident on the polyacrylamide gel represent sizes of approximately 36 kDa, 24 kDa and 17 kDa. The sizes expected from enterokinase digestion were 17 kDa and 18 kDa representing the Trx fusion partner and the PHTH protein respectively. To determine the identity of the fragments produced, samples of Trx-PHTH were digested with both thrombin and enterokinase and Ni affinity resin was added (section 2.4.9, 2.4.10). The protein fragment containing the 6-histidine tag will still bind to the resin. Trx-PHTH protein will contain only the N-terminal 6-histidine tag as a stop codon present in the PHTH insert will prevent translation of the C-terminal 6-histidine tag. The fusion partner (the Trx protein) will retain the 6-histidine tag and pellet with the addition of the resin (labelled Dp). The size of the fragment that remains in solution (labelled Ds) is approximately 24 kDa when compared with Sigma size markers. This is bigger than the expected mass for Tec PHTH and the remaining fragment is smaller than the fusion protein, thus, it appears the enzymes were cleaving the protein at a different site within the fusion protein (Figure 3.10B). This discrepancy was further analysed using small-scale size exclusion chromatography. Undigested Trx-PHTH samples and digested Trx-PHTH samples were loaded on a size exclusion Sephadex G75 column equilibrated in TBS (section 2.4.11). The profiles of both undigested and digested samples are represented in Figure 3.10C and Figure 3.10D, respectively. The column was calibrated using a set of size markers including lysozyme (14 kDa), chymotrypsin (24 kDa), ovalbumin (45 kDa) and bovine serum albumin (BSA) (66 kDa) and the peaks eluted from the Trx-PHTH digest were compared with these standards. The peaks from the Trx-PHTH digest appeared to elute at sizes consistent with those expected: 17074 and 18136 Da representing the Trx fusion partner and the PHTH protein,

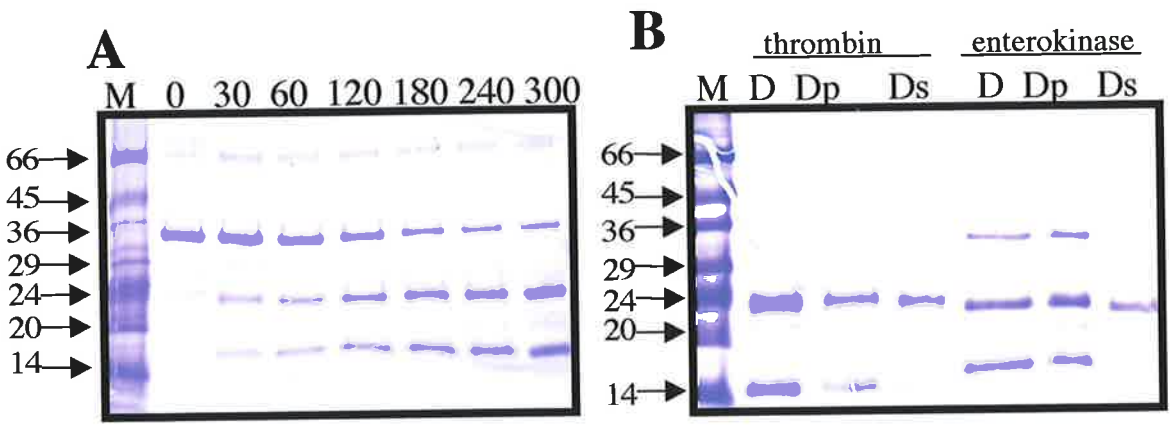
Figure 3.10

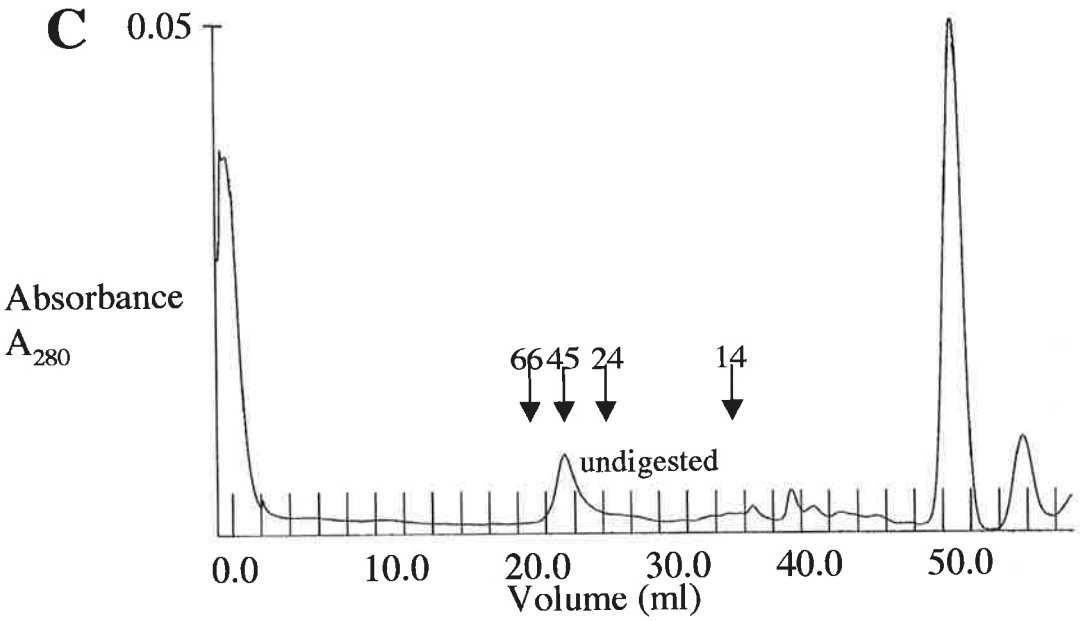
A. 12.5% SDS-PAGE analysis of time points taken during enterokinase digestion of Trx-PHTH. 1/20 mass ratio dilution of enterokinase was used and time points removed and digestion stopped by the addition of PMSF. 10 µl of these samples was diluted with equal volume of 5 x SDS-PAGE loading buffer. Samples were removed up to 5 hours. Visualisation was by Coomassie staining.

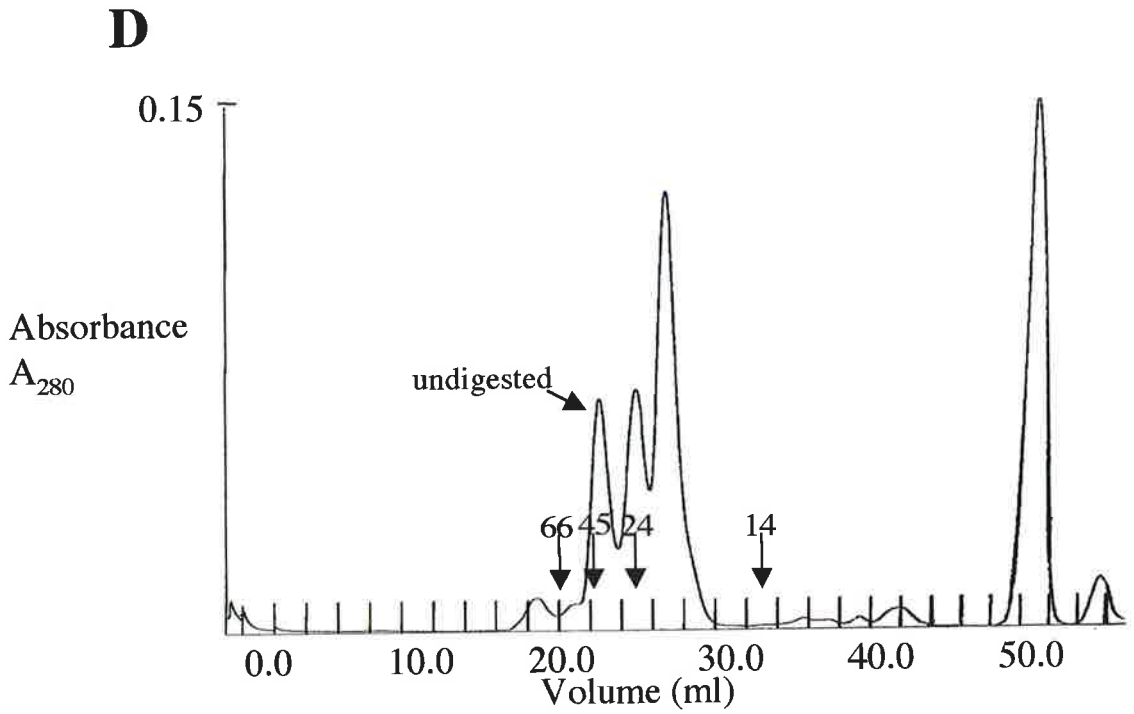
B. 12.5% SDS-PAGE of digested Trx-PHTH protein. Trx-PHTH protein was either digested with thrombin or enterokinase and mixed with nickel resin. These samples were shaken and then centrifuged and the two phases separated and retained. D alone is the digestion mixture used in the experiment. Dp is the pellet after centrifugation whereas the Ds sample is the supernatant.

C. Small scale size exclusion chromatography of Trx-PHTH protein. A Superdex 75 column was equilibrated in 1 x TBS, Trx-PHTH protein was loaded and run at 50 µl/min. 100 µl fractions were collected. The sizes at which standard elute is highlighted on the profile.

D. Size exclusion chromatography of enterokinase digested Trx-PHTH protein. Chromatography was carried out as for C.


A
M 0 30 60 120 180 240 300
66
45
36
29
24
20
14
B
thrombin
enterokinase
M D Dp Ds D Dp Ds
66
45
36
29
24
20
14


C
0.05
Absorbance
A280
66 45 24
14
undigested
0.0
10.0
20.0
30.0
40.0
50.0
Volume (ml)


D
0.15
Absorbance
A280
undigested
66 45 24
14
0.0
10.0
20.0
30.0
40.0
50.0
Volume (ml)

respectively. There is a large contaminating peak of low molecular weight present in both column preparations. The identity of this species is unknown.

Mass spectroscopy was conducted on undigested and digested proteins and indicated that the PHTH protein was 22,257 Da and the Trx-PHTH fusion protein was 35,602 Da (not shown). The size of Trx-PHTH agrees with the expected value while the size observed for the PHTH domain is larger than expected. This data confirmed that the enterokinase enzyme cut the Trx-PHTH fusion protein at an erroneous site. Analysis of the sequence did not reveal where the enterokinase was digesting the protein and the digestion by enterokinase was not pursued.

Using the pET system for purification of the PHTH domain did not improve the solubility or stability when compared with the pGEX fusion system. Although purifying the protein under denatured conditions can generate more PHTH fusion protein, it was lost during the subsequent refolding and dialysis steps. The yields obtained were not sufficient to pursue for NMR experiments and, thus, the GST expression system was used for further expression.

3.4 ANALYSIS OF TEC PHTH PROTEIN

3.4.1 Fluorescence experiments of TRX-PHTH and GST-PHTH

Unfolded protein can be recovered from the inclusion bodies using the pET expression system (section 3.3.2). However, an assay was required to determine if purified and refolded protein was correctly folded. Although the PHTH region of Tec kinase has been shown to interact with several different proteins, there was no biological assay for the PHTH region of Tec kinase at the time of this study. Circular dichroism is an ideal technique to determine the fold of a protein, however, this technique was not available. Fluorescence spectroscopy was used to compare the folding state of refolded Trx-PHTH protein and GST-PHTH protein that was purified in the native conformation. This method relies on probing the chemical environment around specific amino acids, including tryptophans and, to a lesser extent, tyrosines and phenylalanines (Brand and Witholt, 1967). These amino acids contain delocalised aromatic systems that act as intrinsic fluorophor groups with tryptophan accounting for 90% of fluorescence of native proteins. The protein sample is excited at a wavelength of 275 nm and the emission spectra recorded at

wavelengths spanning 310-400 nm. The maxima expected for a pure sample of tryptophamide is at 350nm.

There are two tryptophans in the PHTH region of the protein: Trp 102 in the PH region and Trp 125 in the Btk motif. Trp 102 corresponds to the conserved tryptophan in all PH domains and is predicted to be present in the core of the protein based on the 3D structures of other PH domains (Downing *et al.*, 1994). This buried tryptophan would therefore be a good indicator of whether or not the PH domain is folded. The tryptophan in the Btk motif is expected to be solvent exposed, and as such may not be a sensitive probe of folding.

A series of five samples of Trx-PHTH buffered in 1.5 mM DTT, 13.5:1 PEG: protein, 250 mM NaCl, 25 mM Tris-base and different urea concentrations to simulate refolding stages (section 2.4.7), were analysed by fluorescence spectroscopy. Figure 3.11A shows the different spectra of Trx-PHTH obtained with differing concentrations of urea in the buffer. The emission observed in 8 M urea was greater than that observed in 3 M, 2 M and 1 M urea. There was a distinct shift of the maxima from approximately 350 nm to 333 nm. Trx-PHTH protein in 8 M urea produced spectra with an emission maximum of 348 nm, close to the 350 nm expected for tryptophamide. This indicates that the tryptophans in the 8 M urea sample of Trx-PHTH were more solvent exposed and therefore the protein is likely to be unfolded. These spectra only had one peak indicating that all tryptophans were in identical environments. Trx-PHTH protein samples in 3 M, 2 M and 1 M urea have fluorescence maxima of 340 nm, 337 nm, and 335 nm respectively shown in Figure 3.11A. The maxima is shifting away from the 348 nm seen in the 8 M urea sample indicating the environment surrounding the tryptophans is changed with different urea concentrations. These spectra also show marked solvent quenching of the tryptophans in the urea-containing buffer as the absorbance units were considerably lower in the urea containing buffers. Thus, it appears that the urea buffer has quenched the tryptophans, or in the unfolded state, the tryptophans are in close proximity to a quenching group within the protein. Lowering urea concentration caused a shift in the wavelength maxima such that in the 1 M urea spectrum, two peaks were present similar to the GST purified PHTH protein. This implies that the protein was shifted back to the same fold as seen for the GST-PHTH sample and the size exclusion purified sample.

The fluorescence spectra obtained from GST-PHTH and size exclusion purified PHTH, derived from the GST-PHTH protein were very similar, indicating that the presence of GST in the fusion protein does not change the maximum absorbance for the tryptophans

Figure 3.11

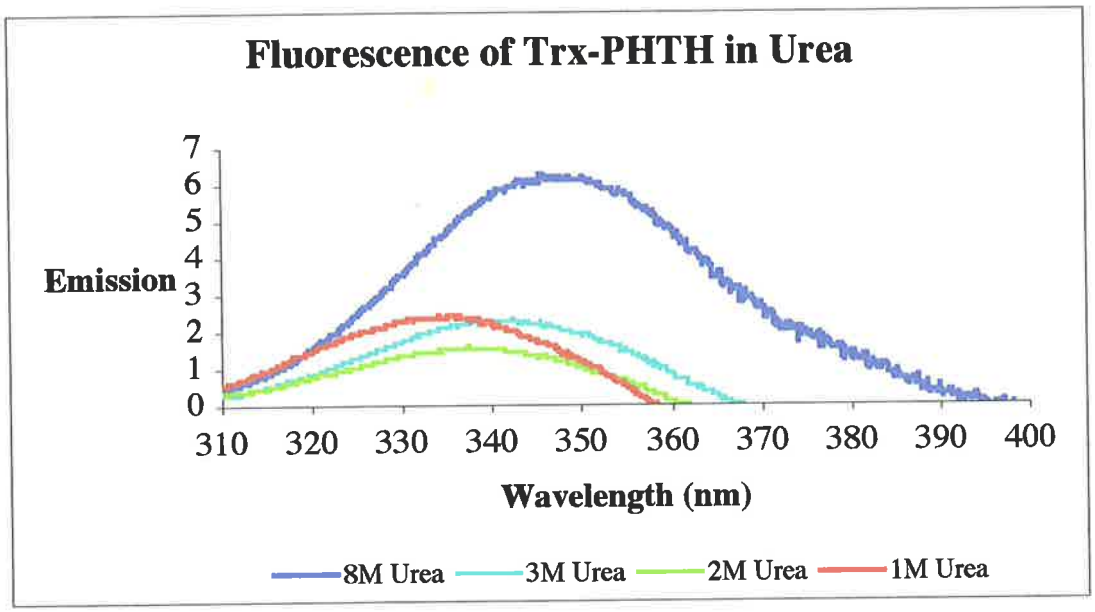
A. Fluorescence spectra of Trx-PHTH protein in different concentrations of urea. Samples of Trx-PHTH were diluted to 0.25 mg/ml with varying amounts of 6 M urea. The protein sample was excited at wavelength of 295 nm and emission spectra recorded at 310-400 nm. Spectra were recorded on a SLM-Aminco Bowman Series 2 Fluorescence spectrometer polarimeter. The spectrum recorded on the sample in 8 M urea is shown in blue, the 3 M in cyan, 2 M in green and the 1 M sample in red.

B. The emission spectra of three different PHTH proteins. The concentrations of the proteins were:

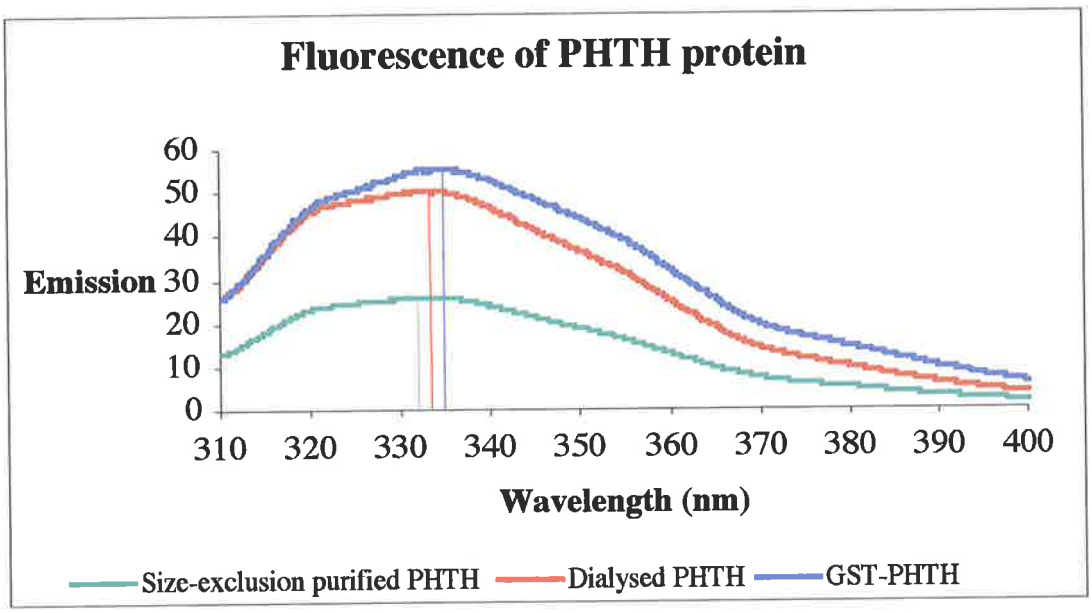
dialysed Trx-PHTH sample	0.423 mg/ml shown in pink
GST-PHTH sample	0.375 mg/ml shown in blue
Size exclusion purified PHTH	0.13 mg/ml shown in cyan

The fluorescence experiments were conducted as for A.

A


Fluorescence of Trx-PHTH in Urea
Emission
0
1
2
3
4
5
6
7
310
320
330
340
350
360
370
380
390
400
Wavelength (nm)
8M Urea
3M Urea
2M Urea
1M Urea

B


Fluorescence of PHTH protein
Emission
0
10
20
30
40
50
60
310
320
330
340
350
360
370
380
390
400
Wavelength (nm)
Size-exclusion purified PHTH
Dialysed PHTH
GST-PHTH

in the PHTH domain (section 2.8). GST-PHTH fusion protein has local maxima at 321 and 333 nm, as does the size exclusion purified sample (Figure 3.11B). The emission units are greater for the GST-PHTH sample because the sample is approximately double the concentration of size exclusion purified sample. The shift from the expected 350 nm suggests the tryptophans in the protein are protected from the solvent possibly buried in the core of the protein.

Trx-PHTH protein dialysed against 250 mM NaCl, 25 mM Tris-base, 1.5 mM DTT and PEG had two maxima, one at 320 nm and one at 333 nm. These maxima are consistent with those observed in the GST purified PHTH protein. This implies the Trx-PHTH was refolded into the same conformation as GST-PHTH protein.

Fluorescence spectroscopy is a useful tool for monitoring the environments around tryptophans and subsequently determining if a protein is folded correctly provided the correct controls are used (Brand and Witholt, 1967). The fluorescence experiments indicated that Trx-PHTH refolded by dialysis, contained tryptophans in the same environment as the GST-PHTH purified protein. The yields from the dialysis procedure did not improve sufficiently to make this a viable option for PHTH protein generation.

3.4.2 PHTH NMR spectroscopic analysis

The aim of this work was to determine the three dimensional structure of the PHTH domain of Tec kinase to gain some insight into the function and regulation of Tec kinase. Although expression yields were not optimal, PHTH protein was expressed and purified as a GST fusion protein and the purity determined both by mass spectrometry and PAGE. The protein for further experiments was generated as a GST fusion protein as the yields were comparable to those of the pET system and refolding of the protein was not required. Preliminary NMR experiments were conducted at the Biomolecular Research Institute, Melbourne in collaboration with Assoc Professor Raymond S. Norton. Nuclear Overhauser Effect SpectroscopY (^{1}H-^{1}H NOESY), TOtal Correlation SpectroscopY (^{1}H-^{1}H TOCSY) and Double Quantum Filtered-Correlation SpectroscopY (^{1}H-^{1}H DQF-COSY) experiments were collected on a Bruker AMX 500 spectrometer equilibrated at 25°C (section 2.5.1). The concentration of the PHTH domain was 1 mM.

These experiments had very little signal. 2D spectra from these experiments showing the fingerprint region as well as the aromatic region are represented in Figure 3.12. The lack of signal maybe due to the PHTH protein aggregating. Support for this conclusion was the presence of a small amount of precipitate at the conclusion of the experiments. The

fingerprint region of the ^{1}H-^{1}H COSY spectrum should contain approximately 150 cross peaks representing correlations between the backbone amide and the CαH of each residue (except proline). However, it can be seen from Figure 3.12A that only approximately 50 cross peaks were visible in this spectrum. The lack of cross peaks in this finger print region is indicative of an aggregated protein complex.

In order to optimise these experiments, a series of different mixing time ^{1}H-^{1}H NOESY experiments were conducted to determine the mixing time giving the maximum signal. Three NOESY spectra were recorded with mixing times of 100 ms, 150 ms and 200 ms on a sample of concentration of 1 mM. These experiments were conducted essentially as above, however, the temperature was at 30°C. The best signal for the ^{1}H-^{1}H NOESY experiments was the 200 ms mixing time. More NOE peaks were present with better signal to noise. In the three different mixing times tested for ^{1}H-^{1}H NOESY experiments, the same regions are highlighted, as in the previous Figure and are shown in Figure 3.13. A series of different ^{1}H-^{1}H TOCSY mixing times, 58 ms, 42 ms and 26 ms were also tested to optimise the ^{1}H-^{1}H TOCSY experiment. The signal to noise was not improved compared with the ^{1}H-^{1}H TOCSY data shown in Figure 3.12. The NMR experiments conducted showed that the PHTH domain was probably aggregating and the presence of a precipitate at the completion of the experiments indicated that the PHTH domain is unstable at temperatures greater than 25°C over the time scale of the NMR experiment. Further investigation was required to confirm the aggregation of the PHTH protein.

3.4.3 Determination of dimerisation by analytical ultracentrifugation

The lack of cross-peaks and broad peak shape in the ^{1}H-^{1}H TOCSY and ^{1}H-^{1}H NOESY spectra (section 3.4.2) when compared with a smaller protein of 67 residues such as Tec SH3 domain (Chapter 4) suggested that at NMR concentrations the PHTH domain protein was present as a multimer. To determine if PHTH domain protein was aggregating, a sample was analysed by ultracentrifugation. This work was conducted in collaboration with Dr Keith Shearwin, Department of Biochemistry, University of Adelaide. Equilibrium sedimentation can be performed to characterise macromolecules and their interactions in solution (Schuster and Toedt, 1996). Equilibrium sedimentation ultracentrifugation generates a value for the molecular weight of the molecule in the solution. This value can be compared with the calculated molecular weight to determine the multimeric status of the protein in the solution.

Figure 3.12

^{1}H-^{1}H Homonuclear NMR experiments of Tec PHTH domain buffered in 10 mM phosphate, 0.2 mM DTT, 0.2 mM NaN_3 pH 5.5. The NMR data collected on Tec PHTH was acquired at the BRI Melbourne on a 1.0 mM sample. All spectra were recorded at 25°C with 6994 Hz in both dimensions. Mixing times of 43 ms and 140 ms were used for the ^{1}H-^{1}H TOCSY and ^{1}H-^{1}H NOESY, respectively. Final spectral sizes were 2048 and 1024 in D1 and D2, respectively. The data was manipulated in FELIX (MSI) and transferred to XEASY (Bartels *et al*, 1995) for printing. The experiment type is indicated above each spectrum.

A. ^{1}H-1H DQF-COSY spectrum

B. ^{1}H-1H TOCSY spectrum

C. ^{1}H-1H NOESY spectrum

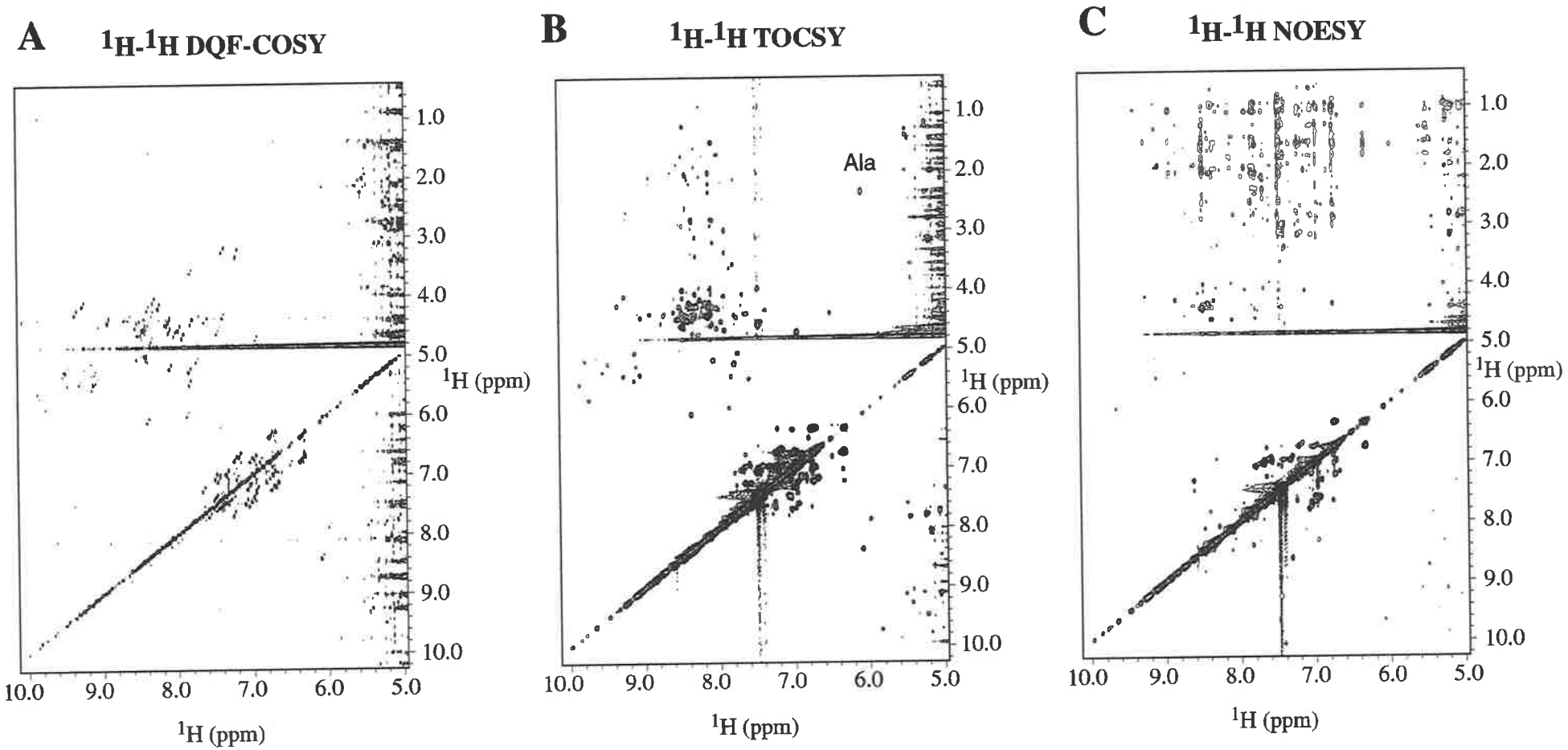
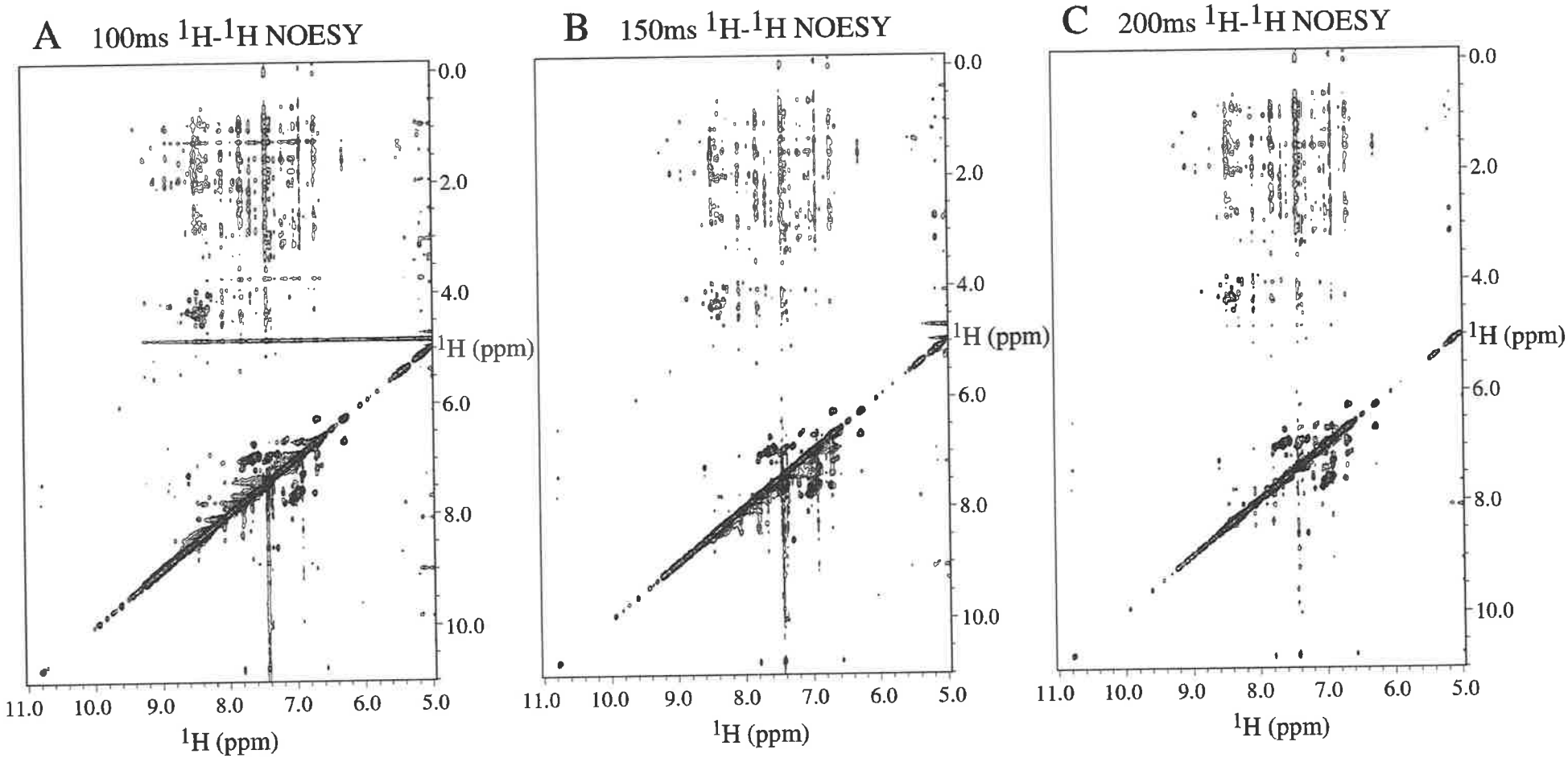

A
1H-1H DQF-COSY
B
1H-1H TOCSY
Ala
C
1H-1H NOESY
1H (ppm)
10.0
9.0
8.0
7.0
6.0
5.0
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
10.0

Figure 3.13

Optimisation of the NOESY mixing time for Tec PHTH domain. The NMR data collected on Tec PHTH was acquired at the BRI Melbourne on a 1.0 mM sample. These spectra were recorded at 30°C with 6994 Hz in both dimensions. Mixing times for the NOESY experiments were (A) 100 ms, (B) 150 ms and (C) 200ms, and are indicated above the spectrum. Final spectral sizes were 2048 and 1024 in D1 and D2, respectively. The data was manipulated in FELIX (MSI) and transferred to XEASY (Bartels *et al*, 1995) for printing.


A 100ms ^{1}H-^{1}H NOESY
B 150ms ^{1}H-^{1}H NOESY
C 200ms ^{1}H-^{1}H NOESY
0.0
2.0
4.0
6.0
8.0
10.0
^{1}H (ppm)
11.0
10.0
9.0
8.0
7.0
6.0
5.0
^{1}H (ppm)

Tec PHTH protein solutions of 0.025 mM and 0.0125 mM were centrifuged at rotor speeds of 12K, 18K and 28K until equilibrium was reached and absorbance at 280 nm collected. The data were fitted to an ideal model of a monomer-dimer interaction and a plot of the absorbance versus radial distance was produced. The residuals from the model are also shown (Figure 3.14B). A weak monomer-dimer transition was occurring between different PHTH molecules with a Ka of $1.6x10^3$ M^{-1} corresponding to an equilibrium dissociation constant of 625 μM. Calculation of the percentages of different multimers versus molarity/concentration indicated that at 1 mM 60% of the PHTH domain in the sample was present as a dimer. Importantly, at the concentration used for NMR analysis, the majority of the sample would have been in a dimeric form (Figure 3.14C). Analytical ultracentrifugation has showed the PHTH domain was present in a monomer to dimer equilibrium, consistent with the NMR spectroscopy. Given the difficulties experienced in producing sufficient material and its propensity to aggregate at the high concentrations necessary for structural analysis it was decided not to pursue structure determination further.

3.5 N-TERMINAL REGION OF TEC PROTEIN

3.5.1 Expression tests of TH domain, THSH3 domain, PHTHext, HTHSH3 and the PH domain

A series of GST fusion proteins incorporating different regions of the PHTH domain of Tec were expressed to optimise the production of soluble protein of this N-terminal regulatory region. All the following proteins were generated as for the PHTH domain. cDNAs corresponding to different regions within the N-terminus of Tec kinase were PCR amplified, digested and ligated into the *Bam*HI and *Eco*RI sites in the pGEX-2T plasmid. All constructs, verified by sequencing using the GST sequencing primers (section 2.3.13), were transformed into BL-21λDE3 bacteria for expression studies. A summary of the proteins investigated is shown in Figure 3.4.

The proposed TH region lacking the PRR (Tec amino acids 117 to 152) was expressed to determine whether this region could form a stable fold independent of the rest of the molecule. A band at the expected size was observed in the induced lane following electrophoresis when bacteria harbouring the pGEX-TH plasmid were induced by the addition of IPTG to a final concentration of 0.2 mM (not shown). Following lysis, there was evidence of soluble protein production. Expression of soluble protein was also observed for

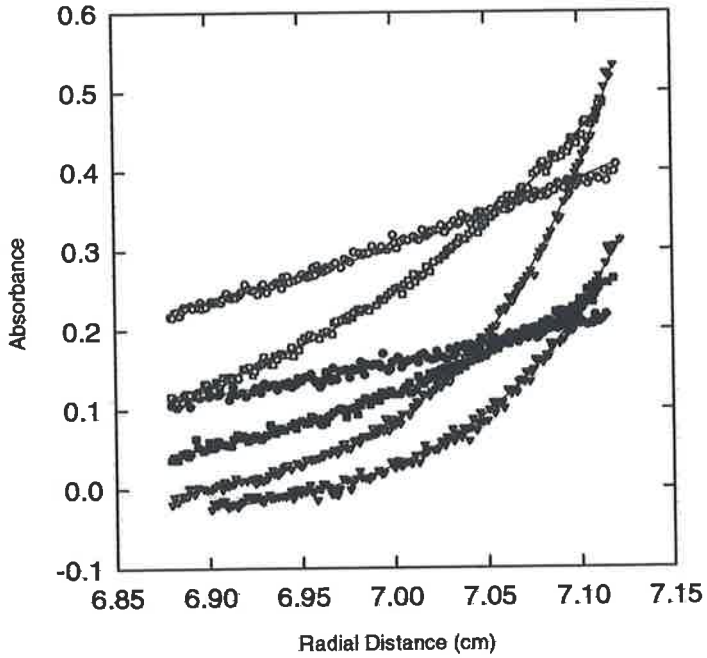
Figure 3.14

A. Ultracentrifugation of Tec PHTH protein. Plot of the absorbance against radial distance for all six runs, two different concentrations (0.025 mM and 0.0125 mM) at three different speeds (12K, 18K and 28K) of PHTH protein purified from a GST fusion in 10 mM phosphate, 150 mM NaCl, 0.01 mM NaN_3 and 1 mM DTT.

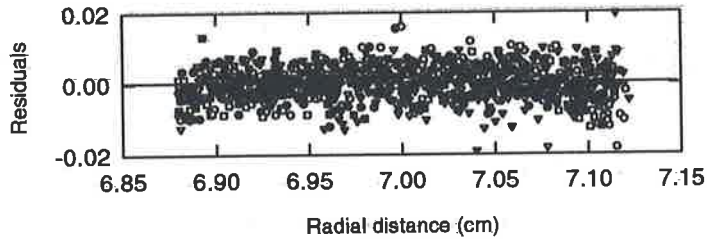
B. Residuals of the fitting of PHTH protein to a monomer-dimer interaction.

C. Plots showing the monomer and dimer ratios of PHTH protein at different concentrations of total protein. (I) the x-axis represented in mg/ml, (II) the x-axis shown as factor of 10 molarity. The dimer form is shown in blue and the monomer in red.

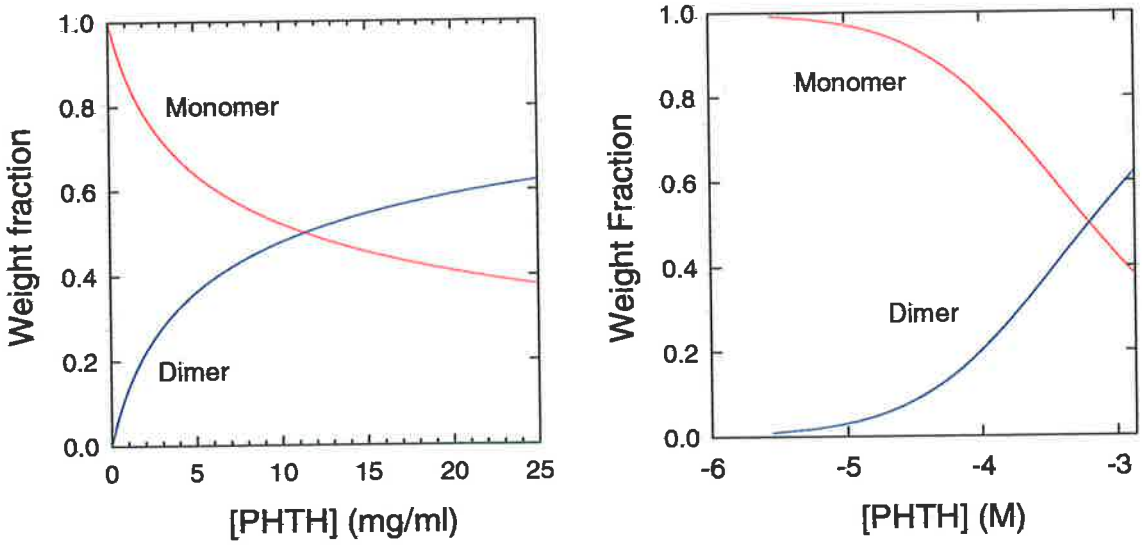
A


0.6
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
Absorbance
6.85
6.90
6.95
7.00
7.05
7.10
7.15
Radial Distance (cm)

B


0.02
0.00
-0.02
Residuals
6.85
6.90
6.95
7.00
7.05
7.10
7.15
Radial distance (cm)

C


1.0
0.8
0.6
0.4
0.2
0.0
Weight fraction
Monomer
Dimer
0
5
10
15
20
25
[PHTH] (mg/ml)
Weight Fraction
Monomer
Dimer
-6
-5
-4
-3
[PHTH] (M)

two other constructs including the THSH3 domain (GST-THSH3) amino acids N^{111}-E^{244} of the Tec protein sequence and the GST-HTHSH3 (THSH3 with the additional N-terminal helix from the PH domain) (Figure 3.15A and B). The SH3 domain is a well characterised stable fold that is amenable to structural studies (section 1.5.2) and expression of the TH domain linked to this domain may have improved the overall solubility of these proteins (Figure 3.15A). The PHTH protein extended to incorporate the PRR, (termed PHTHext) also produced some soluble protein (Figure 3.15C). These proteins were produced in larger quantities and investigated further.

The expression level of the PH domain alone was tested. Although this protein is expressed in bacteria, the protein produced was insoluble and no further analysis was undertaken.

3.5.2 Purification of the GST-TH. GST-THSH3 and GST-PHTHext proteins

The GST-TH domain lacking the PRR was expressed and purified by GST affinity chromatography. Purification resulted in a band of approximately 30 kDa representing the GST fusion protein (26 kDa) and the TH domain of Tec protein (4 kDa) (Figure 3.16A) (section 2.4.1). Gel 3.16B shows the polyacrylamide electrophoresis gel monitoring the purification of the GST-TH protein. Yields of 60 mg fusion protein /L of rich media were produced for this protein. GST-TH protein from the glutathione column was judged to be approximately 95% pure by SDS-PAGE gel electrophoresis (Figure 3.16B). After glutathione sepharose affinity chromatography the material eluting with 10 mM reduced glutathione was pooled and subjected to thrombin digestion. The fusion protein band had been completely digested within 1 hour revealing the 26 kDa GST fusion partner (Figure 3.16BII). At 4 kDa, the TH domain is too small to be visualised in this electrophoresis system although the appearance of a band smaller (GST protein) than the GST-TH fusion is suggestive of complete thrombin digestion. The TH domain (eluted at 235 ml) was separated from the GST fusion partner (eluted at 150 ml) by gel filtration (Figure 3.16C) (section 2.4.11) and the purified material was concentrated. The TH domain was buffer exchanged into water supplemented with 10% D_2O, the pH adjusted to 5.1 and the sample readied for NMR analysis.

The THSH3 was expressed and affinity purified using the GST fusion protein. The yields were 40 mg fusion protein/L of LB media however, when these inductions were performed with the addition of $ZnSO_4$ to a final concentration of 30 μM, the protein yields

Figure 3.15

A. (I) 12.5% SDS-PAGE analysis of four pGEX-THSH3 bacterial clones. The bacteria were grown in LB medium and induced by the addition of 0.2 mM IPTG. Pre induction samples (B) were removed and compared to post induction samples (A). The arrow highlights the expressed band.

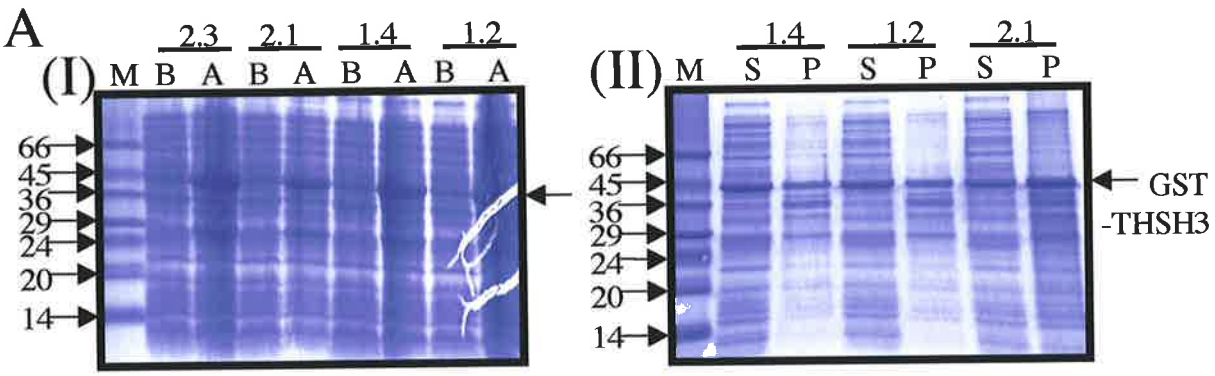
(II) Solubility GST-THSH3 protein expressed by pGEX-THSH3 bacterial clones. Following lysis supernatent samples (S) and pellet samples (P) were electrophoresed. Samples from three of the four clones were run on a 12.5% SDS-PAGE gel. SDS-7 markers were loaded to estimate the size of THSH3 protein (indicated with an arrow). Proteins were visualised by Coomassie staining.

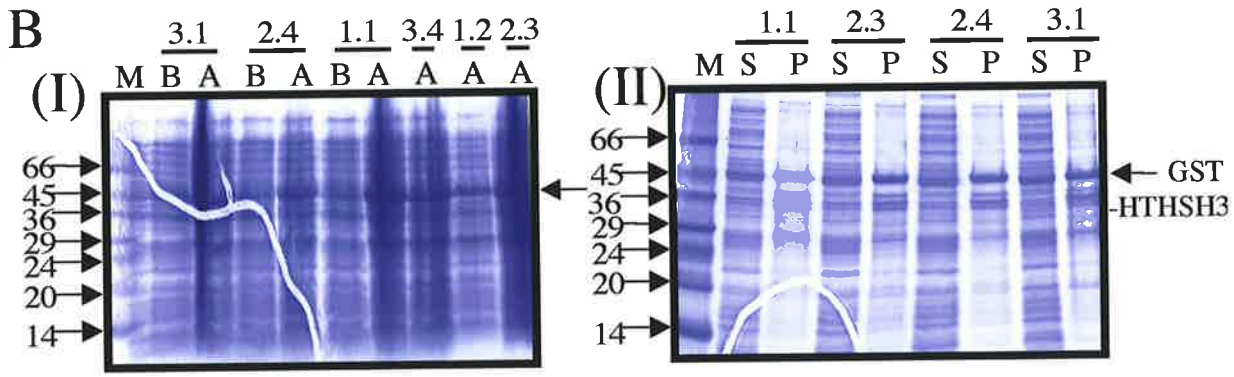
B. Expression and solubility of GST-HTHSH3 protein. Six different pGEX-HTHSH3 bacterial clones were grown in LB and induced with the addition of 0.2 mM IPTG. Samples were loaded on a 12.5% SDS-PAGE gel and proteins visualised by Coomassie staining.

(I) Pre (B) and Post (A) sample were removed and pelleted and resuspended in 5 x loading buffer. Pre and post samples are shown for clones 3.1, 2.4, 1.1 and post induction samples are shown for 3.4, 1.2, and 2.3.

(II) Solubility of sonicated lysates from pGEX-HTHSH3 clones 1.1, 2.3, 2.4 and 3.1. Following lysis by sonication the supernatent (S) and pellet (P) were assessed by SDS-PAGE gel.

C. 12.5% SDS-PAGE analysis of pGEX-PHTHext. Bacteria were grown in LB and induced with 0.2 mM IPTG. Pre and post induction samples were removed and electrophoresed. Samples were compared to SDS-7 size markers for size estimation.


A
(I)
2.3 2.1 1.4 1.2
M B A B A B A B A
66
45
36
29
24
20
14
(II)
1.4 1.2 2.1
M S P S P S P
GST
-THSH3


B
(I)
3.1 2.4 1.1 3.4 1.2 2.3
M B A B A B A A A A
66
45
36
29
24
20
14
(II)
1.1 2.3 2.4 3.1
M S P S P S P S P
GST
-HTHSH3

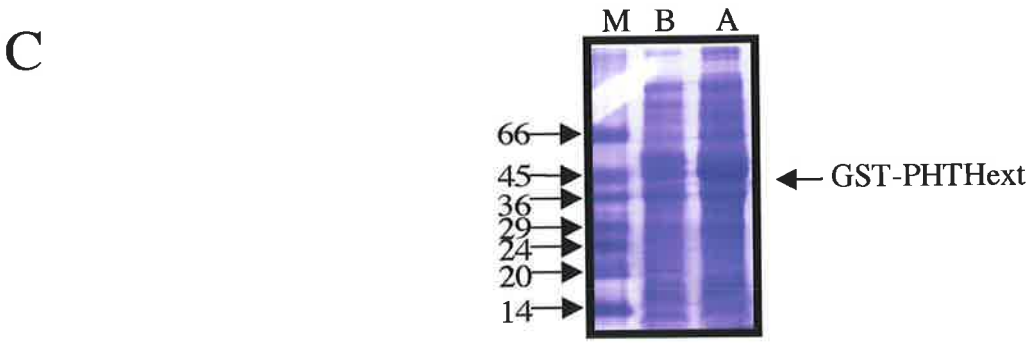

C
M B A
66
45
36
29
24
20
14
GST-PHTHext

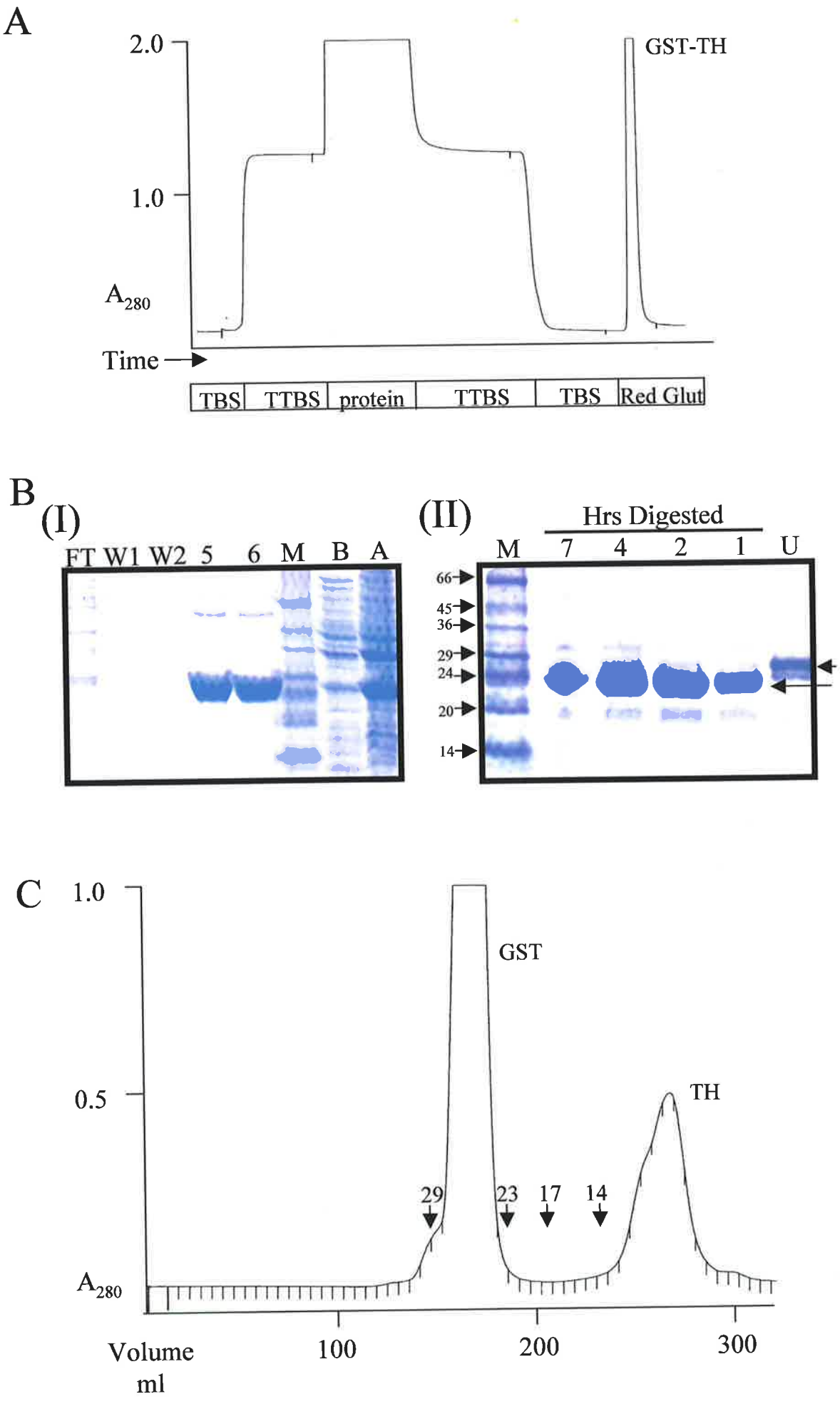
Figure 3.16

A Glutathione agarose chromatography of GST-TH protein. LB cultures were grown in shakeflasks, induced to express GST-TH domain with a final IPTG concentration of 0.2 mM and pelleted. The pellets were resuspended in TTBS and sonicated. A glutathione-agarose column was equilibrated in TBS to establish a baseline following which it was equilibrated in TTBS. Filtered supernatant was loaded on the column (protein) at flow rates of between 3-4 ml/min and then the column washed with TTBS and then TBS. The GST-TH was eluted from the column by TBS supplemented with 10 mM reduced glutathione pH 8. Proteins eluted at each stage are indicated below the profile.

B. (I) SDS-PAGE analysis of the GST-TH purification. Samples before (B) and after induction (A) were collected as well as the flow through (FT) that did not bind to the column. The TTBS (W1) and TBS wash (W2) were also analysed.

(II) Digestion of GST-TH protein monitored by SDS-PAGE. The fusion protein is shown in the undigested lane (U). Digested samples were tested at 1, 2, 4 and 7 hours. Digestion was visualised by Coomassie staining and SDS-7 markers loaded for size estimation.

C. Profile of TH domain purification from the GST fusion partner. 80 mg of GST-TH was digested with 100 U thrombin for 6 hours at 37°C and then applied to the column equilibrated in TBS. Size calibration (kDa) of the column is indicated in black.


A
2.0
1.0
A_{280}
Time
GST-TH
TBS
TTBS
protein
TTBS
TBS
Red Glut
B
(I)
FT W1 W2 5 6 M B A
(II)
Hrs Digested
M 7 4 2 1 U
66
45
36
29
24
20
14
C
1.0
0.5
A_{280}
GST
TH
29
23
17
14
Volume
ml
100
200
300

were significantly greater than previous preparations at approximately 100 mg fusion protein/L of LB medium. Higher yields of soluble protein by addition of zinc suggest that zinc stabilised the TH domain of Tec kinase. Following thrombin digestion, THSH3 and GST were separated by size exclusion chromatography (section 2.4.11). Baseline separation was not achieved by size exclusion chromatography possibly due to the presence of dimers of the THSH3 protein (data not shown)(section 5.3). The THSH3 dimer size would be 32 kDa close to the 26 kDa of the GST fusion partner and would not be separated from the GST fusion protein by size exclusion chromatography. Ion exchange chromatography achieved some separation of the THSH3 from the GST fusion partner; however, there was still a trace of GST. Although purification by the GST affinity resin was not attempted, this protein was not pursued as the size exclusion chromatography suggested the formation of THSH3 dimers that are not suitable for NMR analysis.

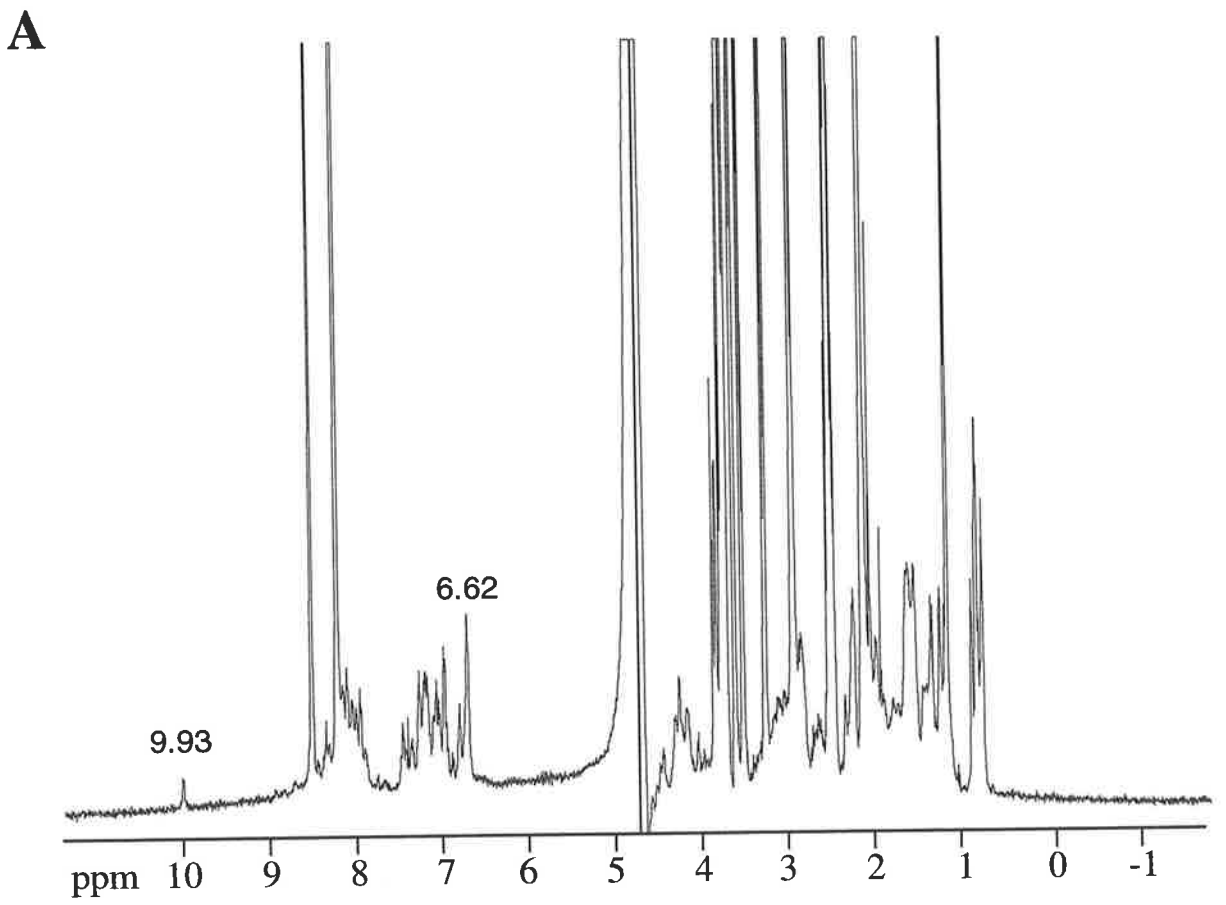
The GST-PHTHext fusion protein was expressed and purified to homogeneity using affinity chromatography followed by a gel filtration chromatography step. This protein expressed well and produced soluble yields consistent with those seen for the PHTH domain at 15 mg/L LB medium but the addition of the PRR did not improve the soluble yields to a significant level for use in NMR experiments and the protein was not pursued.

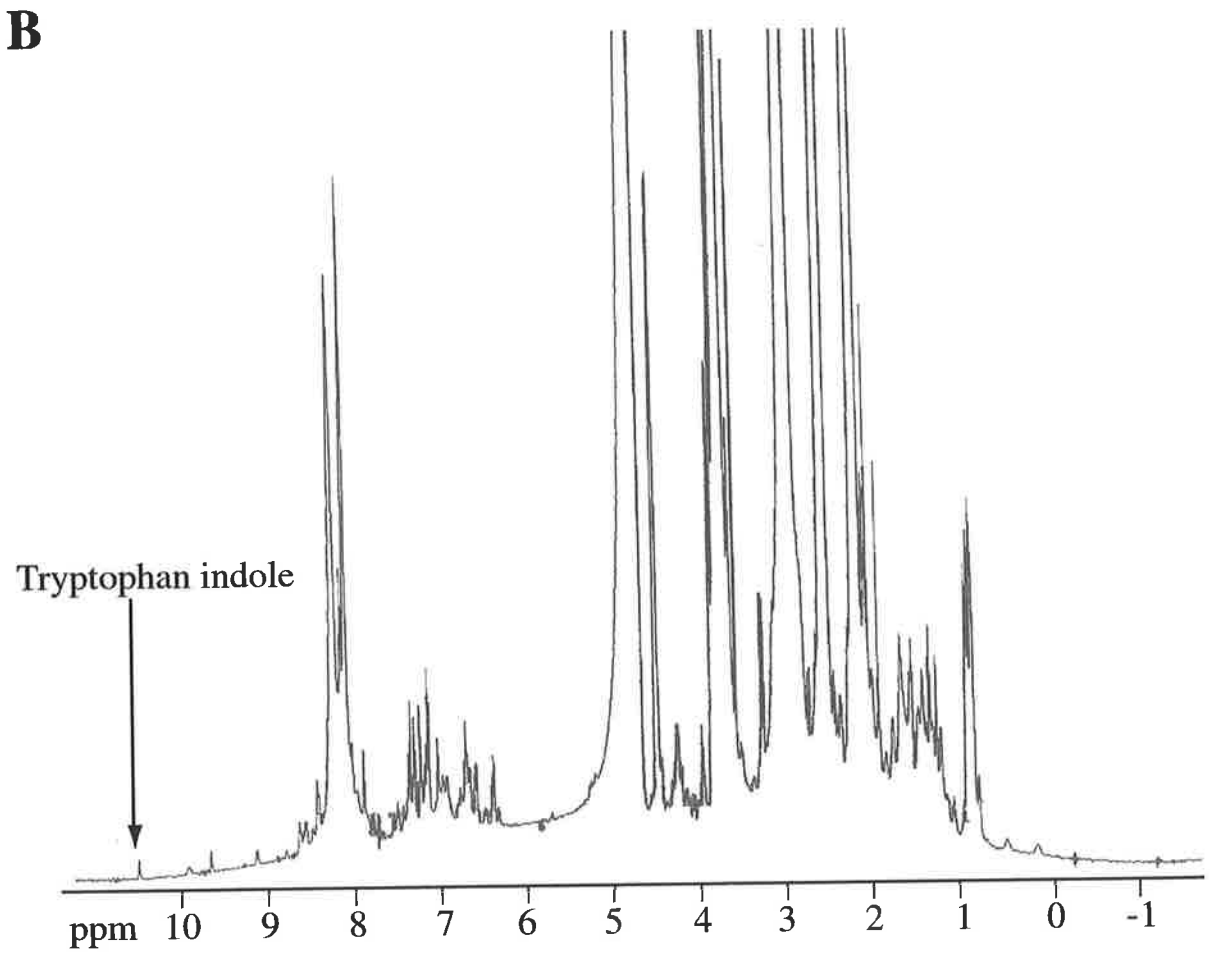
3.5.3 NMR experiments on the TH domain of Tec kinase

The TH domain was further analysed by NMR. The 1D spectrum of the purified TH domain at 0.5 mM shows relatively sharp peaks clustered into small ranges of chemical shift consistent with the values expected of the random coil positions for each amino acid. This lack of dispersion together with an absence of the Trp NεH shift near δ=10 ppm observed in the PHTH 1D spectrum strongly suggests this peptide is not structured in solution. The addition of zinc following purification of the TH domain did not improve protein folding as determined by NMR. However, one dimensional spectra of TH domain protein purified in the presence of 30 μM $ZnSO_4$, 10 mM DTT and TBS buffer showed subtle changes to the previous spectra (Figure 3.17A compared with B). This spectra has suggestion of a small proportion of folded protein. There is a resonance peak at 10.73, suggestive of a folded tryptophan peak. This peak has also been observed in the PHTH NMR spectra (section 3.4.2). Purification of the TH domain in the presence of zinc showed some suggestion of a folded protein, however, the protein was still predominantly unfolded and was not pursued for structural studies.

Figure 3.17

One dimensional spectra of Tec TH domain. The NMR data was collected on a Varian Inova 600 spectrometer. One dimensional spectra were recorded using a spectral width of 8000 Hz and 8192 points and 64 transients on 0.5 mM sample of Tec TH domain at 25°C. The data were processed and printed in VNMR.


A
6.62
9.93
ppm 10 9 8 7 6 5 4 3 2 1 0 -1


B
Tryptophan indole
ppm 10 9 8 7 6 5 4 3 2 1 0 -1

3.6 DISCUSSION

Mutations present in XLA patients highlight the importance of the PH domain of Btk in downstream signalling pathways originating from the B cell receptor. The PH domain has been shown to bind to a variety of proteins including PKC and β-Ark. The interaction of the PH domain with phosphatidylinositol phosphates relocates Btk to the membrane in close proximity with its substrates. This then promotes the downstream signalling originating from the B cell receptor. Recent evidence from our laboratory suggests that Tec plays an equivalent function downstream of the Fcγ receptor during Fcγ-mediated phagocytosis (Atmosukarto *et al* manuscript in preparation (2001)). Thus, it is predicted that the PH domain of Tec would play an equivalent role and localising the Tec protein to the membrane near its substrates.

The structure of the PH domain and the Btk motif from Btk indicates that these two domains are mutually inclusive as they both aid the other in maintaining structural integrity. The PH domain forms a seven stranded β-barrel with a capping α-helix with the Btk motif including a bound zinc atom clearly associated.

This project aimed to determine the solution structure of the PHTH region of Tec kinase to relate it to a possible function. Several structures of PH domains have been published allowing predictions of possible functional mechanisms. Whether these possible mechanisms can also be extended to Tec kinase is unknown and determination of the PHTH domain structure of Tec would provide insight into these functions. This work would also supplement functional investigations of Tec kinase PHTH domain.

Structural studies require millimolar concentrations of protein generally in a 500 µl volume and the main focus of this chapter was to improve expression and purification yields of a series of proteins incorporating different regions of the N-terminal regulatory domains of Tec kinase to provide enough protein for NMR experiments. The PHTH domain of Tec kinase is a 17 kDa protein and, thus, protein concentrations of 17 mg/ml are required for a 1 mM protein sample for NMR studies. The main protein tested was the PHTH domain, lacking the PRR (PHTH). This protein was expressed using the bacterial expression systems pGEX and pET and the yeast system pYEX and affinity purified. Both the bacterial systems expressed well, however, yields of soluble protein were low. Yields of 11 mg/L of PHTH protein were obtained following glutathione affinity purification using the pGEX system compared with 16 mg/L Trx-PHTH fusion protein when purified under denaturing

conditions and 6 mg/L Trx-PHTH fusion protein when refolding on the column. The advantage gained from purifying under denatured conditions was lost during refolding experiments as precipitation of the protein was evident. Fluorescence experiments indicated that refolding of denatured PHTH was possible, however, the refolding yields were low. NMR and analytical ultracentrifugation experiments were conducted on PHTH protein generated from 5 L LB cultures of pGEX-PHTH. These experiments indicated the presence of a dimeric species at the protein concentrations necessary for NMR spectroscopy. Analysis of the PHTH protein was not continued as dimerisation of the protein and its instability was not ideal for NMR experiments.

Several other proteins were investigated for soluble protein production with an aim to determine the NMR structure of the TH component of the PHTH region in conjunction with other regions of Tec kinase, including the SH3 domain. These include the PHTH with an extension of the PRR called PHTHext, the PH domain alone, the TH alone, THSH3 and an extension of the THSH3 including the C-terminal α-helix of the PH domain (HTHSH3). Good yields were observed for GST-THSH3 and GST-TH with maximum yields of 40 mg/L and 60 mg/L in rich medium. The addition of zinc to the culture medium improved soluble yields of the THSH3 protein to 100 mg/L fusion protein. The THSH3 protein expressed well, however, size exclusion chromatography showed possible dimer formation, which is not suitable for NMR analysis.

Purification and analysis of the TH domain of Tec in the presence of zinc did result in some evidence of a folded protein, although this was a small proportion of the whole protein. Saraste and co-workers suggest the PHTH domain of Btk is a mutually inclusive domain, thus, it is interesting that in the presence of zinc there is some evidence of a folded protein (Hyvonen and Saraste, 1997). The other constructs were not pursued as the solubility tests suggested that the proteins were not expressed in the soluble fraction. A final summary of the proteins investigated is shown in Table 3.1

The inability to express Tec PHTH in a variety of bacterial systems and one yeast system implies that expression and purification of Tec PHTH may require expression in a mammalian system. This requires the formation of a mammalian expression construct encoding the PHTH region and tissue culture techniques to yield some protein for downstream analysis (Archer *et al.*, 1993). This will minimise the likelihood of incorrect folding of the protein, however, expression of sufficient protein for NMR is unlikely in a mammalian system. There is evidence that other workers have encountered difficulties in expressing and purifying PH domain proteins possibly due to the membrane binding

Table 3.1 Summary of the expression, solubility, purification and further experiments for PHTH proteins tested

Protein	Expression	Solubility	Affinity Purification	Further Experiments
pGEX-PHTH	++++	++	✔	✔
pET-PHTH	+++++	+	✔	✔
pYEX-PHTH	+	-	-	✔
pGEX-PHTHext	++++	++	✔	-
pGEX-PH	+++++	+	-	-
pGEX-TH	+++++	+++	✔	✔
PGEX-THSH3	+++++	+++	✔	✔
PGEX-HTHSH3	+++++	+	-	-

+ indicates range from good (+++++) to bad (+)

✔ indicates that experiments were conducted

- indicates experiments were not conducted

function of PH domains. Initial studies on Btk PH domain were conducted on the R28C mutant form of the protein as this was the only protein that produced high quality crystals. The authors had also tried expression of the PH domain alone, however, this protein was also insoluble (Hyvonen and Saraste, 1997).

This work has provided an in depth characterisation of the biophysical properties of the PHTH region of Tec kinase. Although the aim of structural determination was not reached the data collected provides important information regarding the solubility and protein expression of the N-terminal region of Tec kinase.

CHAPTER 4:

STRUCTURE DETERMINATION OF TEC SH3 DOMAIN

4.1 TEC FAMILY SH3 DOMAINS

The SH3 domain of Tec family kinases is required to regulate the enzyme's activity. Deletions within the SH3 domain of the Tec kinases, Btk and Tec, result in deleterious effects on downstream signalling. A 21 amino acid deletion in Btk's SH3 domain has been identified in patients suffering X-linked agammaglobulinemia (XLA) (Zhu *et al.*, 1994). XLA patients lack circulating B cells, and as such are susceptible to infection. Many mutations that cause XLA map within the PHTH domain of Btk while relatively few are present in the SH3 domain. The 21 amino acid deletion is due to mutation of the 5′splice site for intron 8, a T to G change that results in splicing out of the 63 bp exon 8. The loss of 21 amino acids removes a section of the SH3 domain binding pocket as well as the 3_{10} helix at the C-terminal end, and in addition, β-strand E from β-sheet 1 and β-strand F from β-sheet 2. The authors propose that the deletion of 21 amino acids within the SH3 domain alters the structure, which then disturbs the interaction with crucial downstream signalling proteins (Zhu *et al.*, 1994). Tec III is an equivalent naturally occurring isoform observed in mice although biological effects of the missing 21 amino acids have not been determined (Mano *et al.*, 1993); (Merkel *et al.*, 1999). A major auto-phosphorylation site (Tyr 223) was found within the Btk SH3 domain following activation by Src family members (Rawlings *et al.*, 1996); (Park *et al.*, 1996). The ligand binding properties of Btk's SH3 domain are altered when this tyrosine is phosphorylated. Tyrosine 223 is located at the edge of the binding pocket and when phosphorylated may block access for cellular ligands. In the unphosphorylated form the SH3 domain of Btk binds to Wiskott Aldrich associated protein (WASP) and c-Cbl (Morrogh *et al.*, 1999). Immunoprecipitation and Western experiments have shown that when phosphorylated, Btk can no longer bind to WASP but retains the ability to bind to c-Cbl, suggesting the SH3 domain can bind these two proteins at different sites (Morrogh *et al.*, 1999). The authors also show that the phosphorylated form of the SH3 domain binds the activated form of Syk through its C-terminal region. This is the first suggestion that the phosphorylation status of Btk may regulate SH3 domain ligand binding (Morrogh *et al.*, 1999). Biophysical experiments have shown that Btk SH3 domain can also bind a series of proteins including Fyn, Lyn and Hck through the SH3 domain (Cheng *et al.*, 1994). Sab (SH3 domain binding protein that preferentially associates with Btk) was isolated by far western expression cloning by its ability to bind the SH3 domain of Btk and has since been shown to regulate Btk function (Matsushita *et al.*, 1998). Sab has been shown by immunoprecipitation experiments to inhibit the auto- and transphosphorylation of

Btk (Yamadori *et al.*, 1999). The SH3 domain of Itk has also been investigated for its potential ligands. SH3 domain ligands isolated for Itk include Sam68, WASP and hnRNP-K and to a lesser extent c-Cbl and Fyn (Bunnell *et al.*, 1996). Itk SH3 domain has also been shown to bind to CD28 (Marengere *et al.*, 1997). By analogy, it would be expected that Tec kinase would also bind to WASP through its SH3 domain.

The SH3 domain of Tec family members is important for regulation of these kinases. Tec protein lacking the SH3 domain (TecΔSH3) is constitutively active. This mutation does not completely activate Tec because transformed foci are not observed in 3T3 fibroblast cells expressing TecΔSH3. TecΔSH3 is, however, hyperphosphorylated and has elevated kinase activity in these cells, suggesting a role for the SH3 domain in down regulation of enzyme activity (Yamashita *et al.*, 1996).

The solution structure of Btk SH3 domain has been determined and is similar to that of other SH3 domains (Hansson *et al.*, 1998). It has five β-strands that form two β-sheets in a β-barrel structure. The RT loop of the Btk SH3 domain is long and has a tight turn at the C-terminal end. The highly conserved Tyr 223 is present at the edge of the potential binding site. The sections Asp 224-Pro 227 and Asp 232-Leu 235 of the RT loop run antiparallel to each other and are well defined with hydrogen bonds between Leu 235 NH to Tyr 225 CO and Leu 233 CO. The hydrophobic core of the SH3 domain consists of residues Val 219, Leu 233, Leu 235, Tyr 241, Ile 243, Ala 254, Ile 264 and Val 269 and 56% of the surface area is polar. The potential binding site involves the conserved residues Asp 224, Tyr 225, Asp 232, Trp 251, Pro 265 and Tyr 268 bounded by the autophosphorylated tyrosine Tyr 223 and the RT loop. The authors noted the presence of two conformational states for Btk SH3 domain; however, the biological relevance of this is unknown (Hansson *et al.*, 1998). The Btk SH3 domain binds to a proline rich ligand from the tyrosine kinase c-Cbl in a class I orientation. Analysis of chemical shift changes and NOEs indicate that the RT loop, n-Src loop and the loop between β-strands β4 and β5 bind to the ligand. The form of the Btk SH3 domain lacking 21 amino acids exhibits reduced binding to the PRR of c-Cbl, possibly due to improper folding, and this may be the basis of XLA in patients carrying this mutation (Tzeng *et al.*, 2000); (Patel *et al.*, 1997).

4.2 STRUCTURAL DETERMINATION BY NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY

Nuclear magnetic resonance spectroscopy (NMR) has recently emerged from behind the shadow of X-ray crystallography and is now a valuable technique for the determination of solution three dimensional (3D) structures. NMR and X-ray crystallography exist in a symbiotic relationship that allows both fields to gain from the other. NMR can supplement the information generated from X-ray crystallography by providing structural information in a more biological context, that is, in solution. X-ray crystallography aided NMR during its earlier years by allowing a means to compare independently derived structures (Wuthrich, 1995).

NMR uses the magnetic spin properties of atomic nuclei to identify atoms that are close together in space (Howard, 1998). ^{1}H, ^{15}N, ^{13}C and ^{31}P nuclei orientated by a magnetic field absorb radiation at characteristic frequencies, which manifest as NMR peaks (Howard, 1998). The frequency at which the nuclei resonate is dependent on the local molecular environment of the atom of interest and these are generally expressed as chemical shift values in parts per million relative to a reference signal (Roberts, 1993). This allows for the variation in the field strengths of different magnets. In an extended polypeptide chain, where all amino acid side chains are exposed equally to solvent, the resonances derived from each different functional group on the chain will resonate together. When placed in a protein environment, protons will be shifted upfield (lower ppm value) or downfield depending on their local environment. The resonances observed for an unstructured polypeptide chain include the methyl groups (CH_3) at ~1 ppm, the methylene group (CH_2) at ~2-3 ppm, α-protons at 4-5 ppm, aromatic groups at 6-7.5 ppm and finally amide groups at 7-11 ppm (Wuthrich, 1995). These resonances are called random coil chemical shift values. Ideally, in a one dimensional (1D) ^{1}H spectrum where the observed signal intensity is plotted against chemical shift, different protons should resonate at discrete positions. However, in practice the spectral overlap for a small protein is such that structure determination from 1D spectra is not possible (Howard, 1998).

Splitting the 1D spectra into two dimensions was the first step to determining full protein structures by NMR spectroscopy. Adaptation of key 2D experiments was required to minimise overlap before the structure of Bull seminal protease inhibitor IIA (BUSI IIA) (Williamson *et al.*, 1985) could be solved by NMR. Since then protein structural determination by NMR has undergone major advancements through production of higher field magnets, routine application of 2D, 3D and 4D NMR experiments, availability of

isotopically labelled protein of interest and development of efficient and reliable software for calculation of the protein structures (Roberts, 1993). A dimeric protein of molecular weight 142 kDa has now been solved by NMR spectroscopy (McEvoy *et al.*, 1997).

NMR data collection can be broadly separated into the Homonuclear (same nuclei eg ^{1}H-^{1}H) and Heteronuclear (different nuclei eg ^{1}H-^{15}N) NMR experiments. Homonuclear experiments generally include ^{1}H-^{1}H 2D COSY (COrrelation SpectroscopY), ^{1}H-^{1}H 2D TOCSY (TOtal Correlation SpectroscopY or HOHAHA HOmonuclear HArtmann HAhn spectroscopy) and ^{1}H-^{1}H 2D NOESY (Nuclear Overhauser Effect SpectroscopY). Both ^{1}H-^{1}H COSY and ^{1}H-^{1}H TOCSY experiments reveal information about through bond linkages between different protons. In 2D contour plots, each proton is represented as a peak on the diagonal and an interaction between two particular protons is represented as a peak off the diagonal. Extrapolation to the diagonal would establish which two protons this cross peak is emanating from. The spin-lock time of the ^{1}H-^{1}H TOCSY experiment (usually called mixing time) can be varied to allow further magnetisation transfer through a particular amino acid. ^{1}H-^{1}H TOCSY can, in all cases, reveal cross peaks for all the protons in an amino acid side chain, however, the distance of transfer from the amide proton through the side chain can differ. ^{1}H-^{1}H NOESY experiments reveal through space information through an effect called the Nuclear Overhauser Effect, a nuclear relaxation phenomenon, the intensity of which is correlated with the inverse sixth power of the inter-nuclear distance (Wuthrich, 1995); (Wuthrich *et al.*, 1991). An NOE cross peak is observed when two protons are within approximately 5Å of each other and this can be extrapolated to the diagonal to identify the interacting protons (Clowes *et al.*, 1995). The intensity of the NOE cross peak relates to the distance between the two protons. The NOE peaks are classified into strong, medium or weak and it is this that reflects the closeness in space (Hinds and Norton, 1994).

Heteronuclear 2D experiments require isotopically labelled proteins to detect a correlation between two different nuclei. ^{1}H-^{15}N HMQC (Heteronuclear Multiple Quantum Coherence) experiments present cross peaks that are from H and N bonds, that is, a peak on the plot would represent every NH in the protein. This reduces the spectral overlap by applying another variable, that of specific nitrogen resonances. Proteins greater than 10 kDa require isotope labelling, however, proteins of smaller sizes can also be isotopically labelled to aid spectral assignment. Standard ^{1}H-^{1}H 2D TOCSY or ^{1}H-^{1}H NOESY experiments can be split by applying a third ^{15}N dimension (Wuthrich, 1990). This yields the same

information as determined by a regular 2D ^{1}H-^{1}H NOESY or ^{1}H-^{1}H TOCSY experiment, however the additional nitrogen dimension significantly lowers overlap.

The second stage of structure determination by NMR requires sequence specific resonance assignments for each amino acid in the protein. This procedure, developed by Wuthrich and co-workers, requires only the amino acid sequence of the protein and the recorded NMR data. Sequence specific assignments are achieved in two stages (Wuthrich, 1986). Firstly, individual amino acid spin systems are identified. A spin system is defined as a group of resonances arising from the same amino acid (Wuthrich, 1986). This is achieved using the ^{1}H-^{15}N COSY and ^{1}H-^{15}N TOCSY data and although this can be completed using 2D homonuclear spectra, it is considerably easier using 3D data as a 10 kDa protein has in excess of 500 NMR observable protons (Primrose, 1993). The spin system can be determined through analysing cross peaks to identify those that are linked. By comparing these with tables of random coil chemical shifts, different amino acids can be determined. Distinguishing amino acids with similar side chains, i.e. those belonging to AMX spin systems (Asp, Asn, Cys and Ser), is more difficult (Gronenborn and Clore, 1990). In the event of only one of a particular amino acid in the chain being present, this would determine where in the polypeptide chain it resides. Generally, there is more than one of a particular amino acid in a protein of 10 kDa and the position within the chain is now determined by attempting to link groups of two or more amino acids together using ^{1}H-^{15}N NOESY NMR data. NOE cross peaks indicate that two protons are closer than 5Å in space (Clowes *et al.*, 1995). The pattern of NOE links between amino acids differs depending on the secondary structural elements present in the amino acid. If the amino acid region is present in β-sheet conformation, strong NOEs exist between the $C\alpha H_i$ and NH_{i+1} due their close proximity to each other. In the case of α-helix, a strong NOE peak occurs between NH_i and NH_{i+1}. Inspection of the NOE data for the strong NOEs can link more than one amino acid together and inspection with the primary sequence can generally allocate the two spin systems to a position within the protein sequence (Wuthrich, 1995).

The maximum number of NOE conformational constraints possible must be collected for the production of a high-resolution structure (Wuthrich, 1995). Distance restraints are determined by determining the strength of the NOE peak. Strong NOE peaks are designated as being 2.0-2.5Å apart if it is an intra-residue NOE but 2.0-5.0Å apart if it is a long range NOE. Some of the NOE peaks are designated a distance constraint and then preliminary structure calculations are undertaken. The first structure calculations were conducted using metric matrix distance geometry (reviewed in Gronenborn and Clore, (1990)), however,

since then, many different methods have been developed. One such method is dynamical stimulated annealing which results in the structures produced having lower folding energy levels (reviewed in Gronenborn and Clore, (1990)).

4.2.1 Structure determination of Tec SH3 domain

The method used for the initial assignment of distance restraints of Tec SH3 domain, to be used in further structure calculations, was Ambiguous Restraints for Iterative Assignment (ARIA). This method required the complete proton assignment of the chemical shifts and a list of partially assigned NOEs. ARIA then generated a list of restraints using a combination of ambiguous distance restraints and an iterative assignment process. ARIA allows this process to be largely automated as it performs a series of routines including partial assignment, calibration, violation analysis and merging and the organization of the iterative procedure (Nilges *et al.*, 1997). A very rudimentary list of NOEs, mainly sequential and secondary structure NOEs, are used to generate the preliminary structures. Further iterations compare possible restraints and remove those that are violated in all structures and makes possibilities from the remainder. Each iteration produces more unambiguous restraints and results in reduced ambiguous restraints until a final family of structures is produced when the structures and data sets converge (Nilges *et al.*, 1997).

The resulting structures are further analysed and more NOEs assigned unambiguously. Each iteration inputs more distance constraints and the resulting structure has better defined regions and more constraints per amino acid. Convergence is judged by the residual constraint violations and all satisfactory solutions are included as a group of conformers (Wuthrich, 1995). Greater number of restraints per amino acid has been shown to lead to the most favourable energy conformations (MacArthur *et al.*, 1994).

NMR spectroscopy can yield information other than just purely 3D structural information. NMR can detect conformation transitions, denaturation, internal mobility, weak ligand interactions and other dynamical information (Wuthrich, 1989).

4.3 AIMS

The SH3 domain is crucial for the regulation and activity of the Src and Tec families of tyrosine kinases. The function of the SH3 domain of Tec family members in regulation is currently under investigation, however, Tec family members lacking part of the SH3 domain or the entire SH3 domain are constitutively active. This study of the SH3 domain from Tec

kinase has been conducted from a structural perspective, the results of which will complement other research into the biological function of Tec kinase.

The SH3 domain of Tec has been cloned into the pGEX-4T-2 vector for high level expression and purification from a bacterial source. The protein is produced C-terminal to the bacterial Glutathione S Transferase (GST) protein to facilitate downstream purification and aid in the high level expression of the protein. A two step purification protocol is required, the first an affinity glutathione agarose column to remove the bulk of the bacterial proteins and after protease treatment, a size exclusion gel filtration column to separate the target protein from the fusion partner. Tec SH3 domain protein has been produced as both an unlabelled protein and a ^{15}N labelled protein. The ^{15}N labelling of the protein was required to minimise overlap in the NMR spectra and facilitate the assignment of the spectra.

This chapter presents the three dimensional structure of Tec SH3 domain and compares the structure with that of other SH3 domains already determined. An investigation of the binding site of Tec SH3 domain has also been undertaken and will be presented in this chapter. The initial 3 dimensional ^{15}N-^{1}H TOCSY and ^{15}N-^{1}H NOESY NMR experiments were kindly collected and processed at the Biomolecular Research Institute, Melbourne by Dr Mark Hinds. The assignment and data collection of the 2D experiments was collected and processed with the assistance of Dr Terry Mulhern. Dr Terry Mulhern kindly conducted ARIA calculations and optimisation.

4.4 EXPRESSION AND PURIFICATION OF TEC SH3 DOMAIN

4.4.1 Tec SH3 protein expression and analysis

A region corresponding to the SH3 domain of Tec IV kinase, amino acids 181-245, was cloned into pGEX-4T-2. Following sequencing pGEX-4T-2-SH3 plasmid DNA was transformed into the *Escherichia coli* bacterial strain BL-21λDE3. Preliminary induction and solubility experiments were undertaken to ensure the viability of this protein for structural studies. Bacteria transformed with pGEX-4T-2-SH3 were grown at 37°C to an OD_{600} of 0.6 and protein expression was induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to 0.2 mM for 4 hours. Following lysis by French pressing, soluble and insoluble material from the culture and pre- and post- IPTG induction samples were electrophoresed on a 12.5% polyacrylamide gel. Figure 4.1A shows an induced band at the expected size for labelled and unlabelled GST-SH3 protein. The induced band observed at approximately 34 kDa was consistent with the size expected for GST-SH3 (26 kDa GST and 8 kDa SH3). GST-SH3 protein produced in bacteria contains a high proportion of soluble protein (lane S Figure 4.1C).

Tec SH3 domain was purified initially using glutathione agarose (Bioserve) that binds specifically to the GST component of the fusion protein. *E. coli* induced to express GST-SH3 fusion protein were pelleted, resuspended in TTBS and lysed by French press (section 2.4.4). The fusion protein was applied to the glutathione agarose column equilibrated in TTBS and washed extensively in both TTBS and TBS. The protein was eluted by addition of 10 mM reduced glutathione in TBS and yields of GST-SH3 domain returned were routinely of the order of 300 mg/L for rich medium cultures. A profile of the purification of GST-Tec SH3 domain is represented in Figure 4.1B.

The purified fusion protein was then cleaved using 4 U bovine thrombin/mg protein for a minimum of 18 hours. The reaction had not gone to completion when shorter digestion times were used, resulting in the extended digestion times for future preparations (Figure 4.2A). The absence of a linker peptide between the globular SH3 domain and the large globular GST fusion protein may sterically hinder access to the cleavage site by the thrombin protease resulting in long digestion times.

The Tec SH3 domain was separated from the GST fusion partner by gel filtration using Superdex G75 beads (Pharmacia) equilibrated in 25 mM NaH_2PO_4, 150 mM NaCl, pH

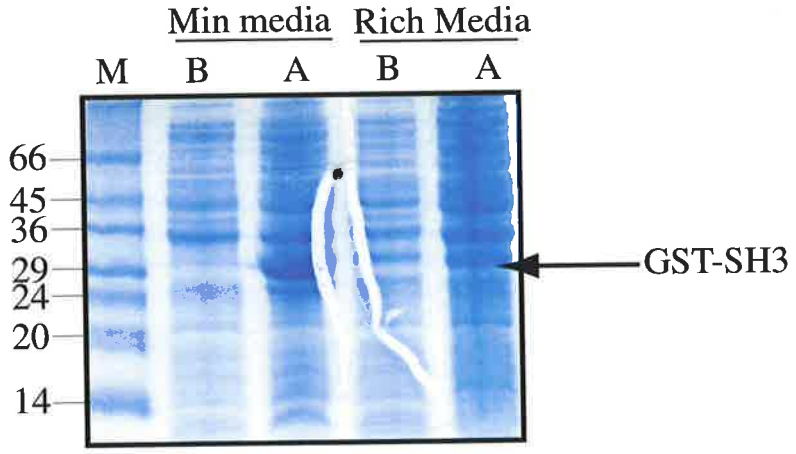
Figure 4.1

A. SDS-PAGE analysis of GST-SH3 protein induction. BL-21λDE3 bacteria transformed with pGEX-SH3 domain were grown either in min A medium (min) or LB medium (rich) and a 1 ml sample removed prior to induction (B) or following 4 hours treatment with 0.1 mM IPTG (A). These were centrifuged and the pellet resuspended in 200 µl 2 x SDS-PAGE buffer. 10 µl was loaded on a 12.5% SDS-PAGE gel, electrophoresed and visualised after Coomassie staining. Markers used are SDS-7 markers (Sigma). The induced band representing GST-SH3 domain is indicated with an arrow.

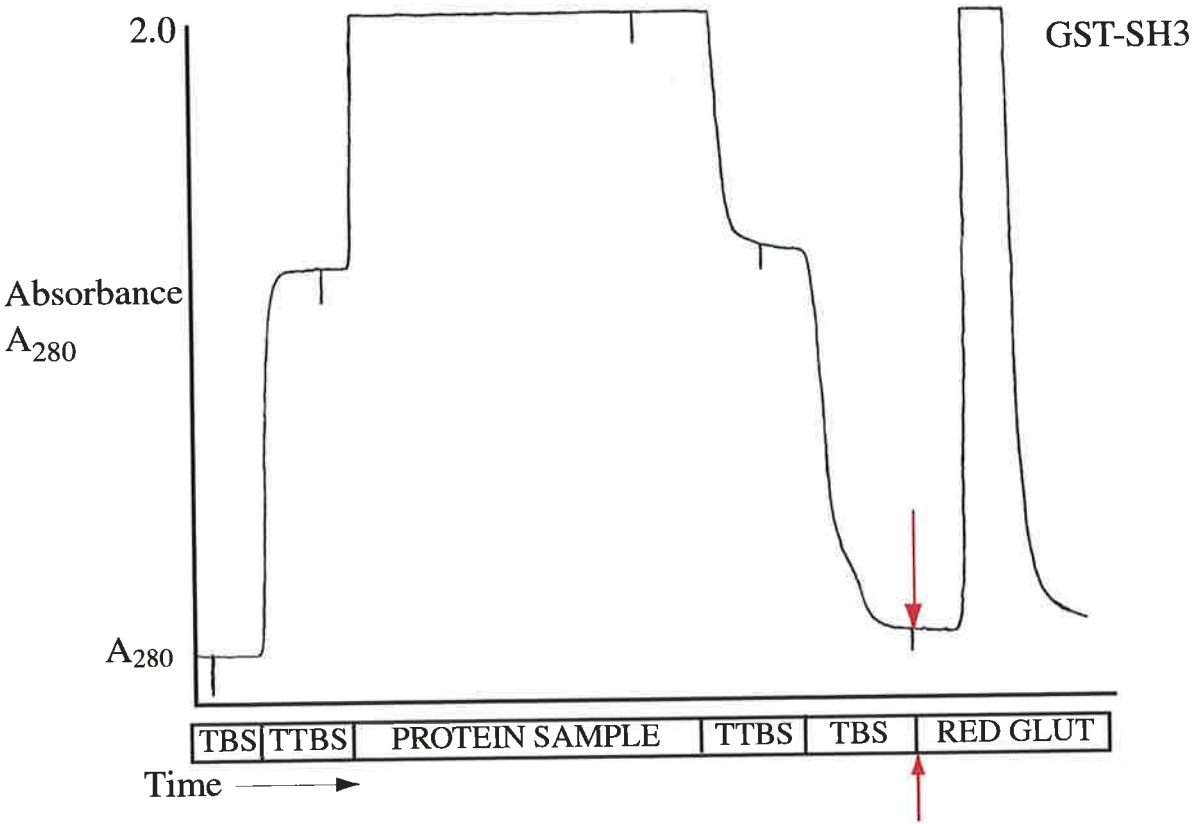
B. Glutathione Sepharose chromatography of GST-Tec SH3 domain. The bacteria were lysed and prepared as described in 2.4.4. The column was equilibrated firstly in TBS and then TTBS and the sample loaded. Following extensive washing the GST-SH3 fusion was eluted by the addition of 10 mM reduced glutathione in TBS pH 8.0 and frozen at -80°C. The reduced glutathione was added where indicated (red arrow).

C. 12.5% SDS-PAGE analysis of affinity purified GST-SH3 protein. (I) Pre (B) and post (A) induction samples were prepared as for A from bacteria grown in LB medium. Following lysis the insoluble fraction (P) and the soluble fraction (S) were collected for analysis. The flow through (FT) of the affinity column was collected and 10 µl added to 10 µl 5 x load buffer. TF, TTBS (1) and TBS (2) wash and the GST-SH3 (SH3) protein eluted by the addition of reduced glutathione. Gel (II) shows the soluble fraction loaded on the column (L), the SH3 (SH3) eluted from the column and the flow through (FT) containing proteins that do not bind the glutathione Sepharose affinity resin. SDS-7 makers were loaded for size comparison. Both gels were stained for visualisation with Coomassie blue.

A

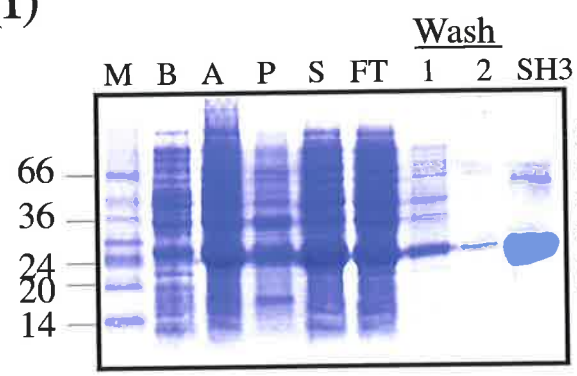

Min media
Rich Media
M
B
A
B
A
66
45
36
29
24
20
14
GST-SH3

B


2.0
GST-SH3
Absorbance
A280
A280
TBS
TTBS
PROTEIN SAMPLE
TTBS
TBS
RED GLUT
Time

C

(I)


Wash
M B A P S FT 1 2 SH3
66
36
24
20
14

(II)

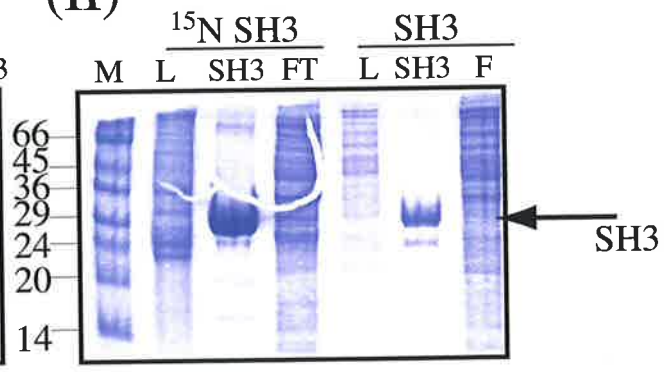

15N SH3
SH3
M L SH3 FT L SH3 F
66
45
36
29
24
20
14
SH3

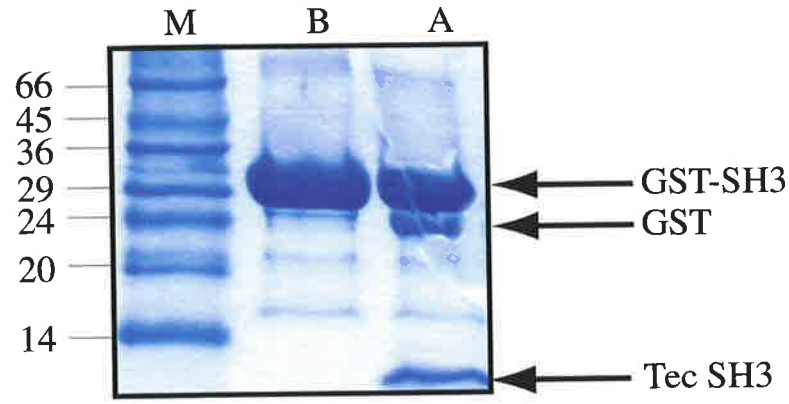
Figure 4.2

A. 12.5% SDS-PAGE analysis of thrombin digested GST-SH3 domain. Samples were removed prior to digestion (U) and then a sample removed after digestion (D) had progressed for 3 hours. 10 µl of each sample was diluted in 5 x SDS-PAGE load buffer and electrophoresed. The positions of the different bands are shown with arrows. Visualisation was by Coomassie blue staining. SDS-7 markers were loaded for size estimation.

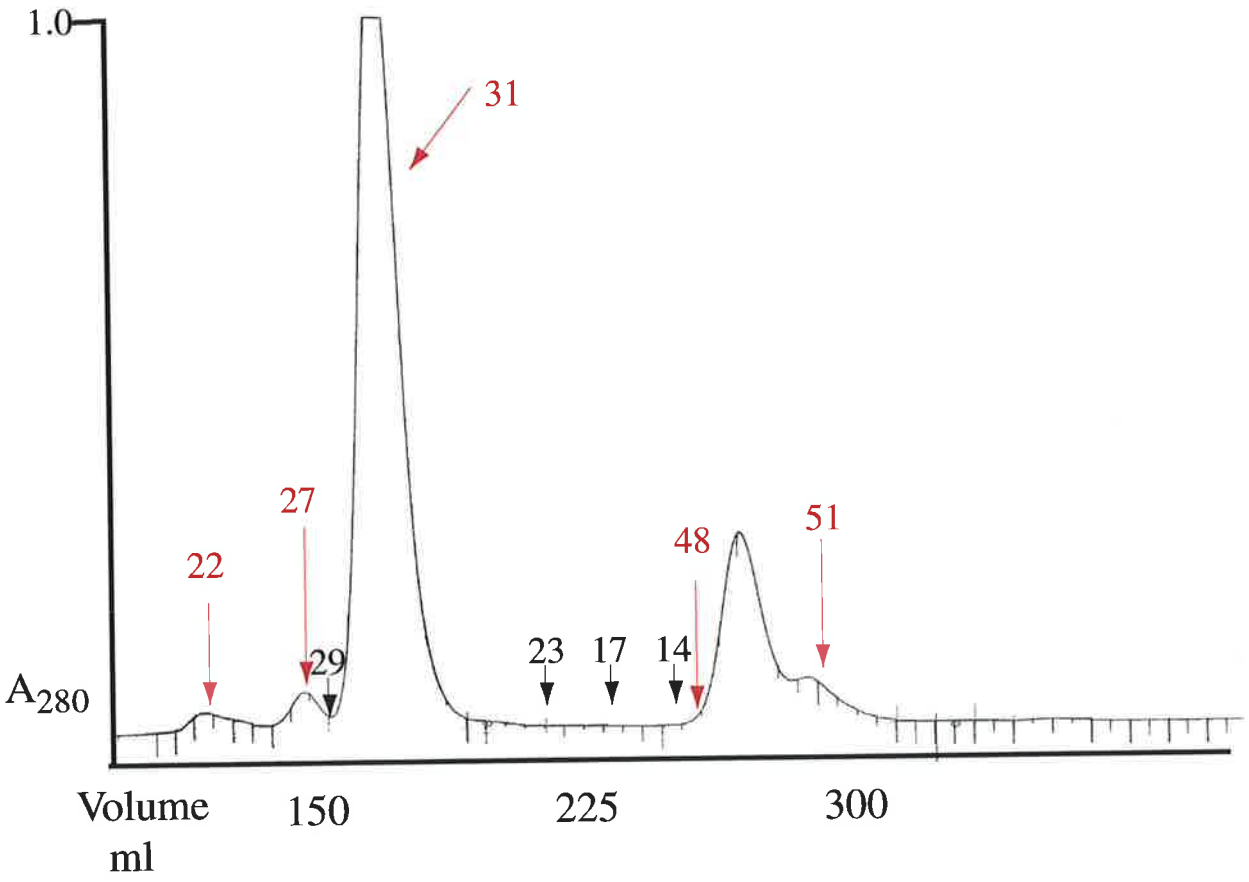
B. Size exclusion chromatography profile of Tec SH3 domain following digestion with thrombin. A Superdex G75 column was equilibrated in 1 x PBS pH 8 with a flow rate of 2 ml/minute. 4 ml of 6 mg/ml GST-SH3 protein was digested with thrombin and loaded onto the Superdex G75 column and 5 ml fractions collected. The protein elution was detected at A_{280} with a full scale absorbance of 1. The fractions from the column are shown in blue. The sizes from the calibration are indicated in black.

C. 12.5% SDS-PAGE of the size exclusion profile for digested GST-SH3 domain. The samples were diluted 1:1 into load buffer and electrophoresed. The sample loaded onto the size exclusion column was analysed to determine the extent of digestion (L) and fractions thoughout the purification tested. Fraction 22 was found to contain no detectable protein. Fraction 27 and 31 contained GST and some undigested GST-SH3. Samples were compared to SDS-7 markers for molecular weight determination.

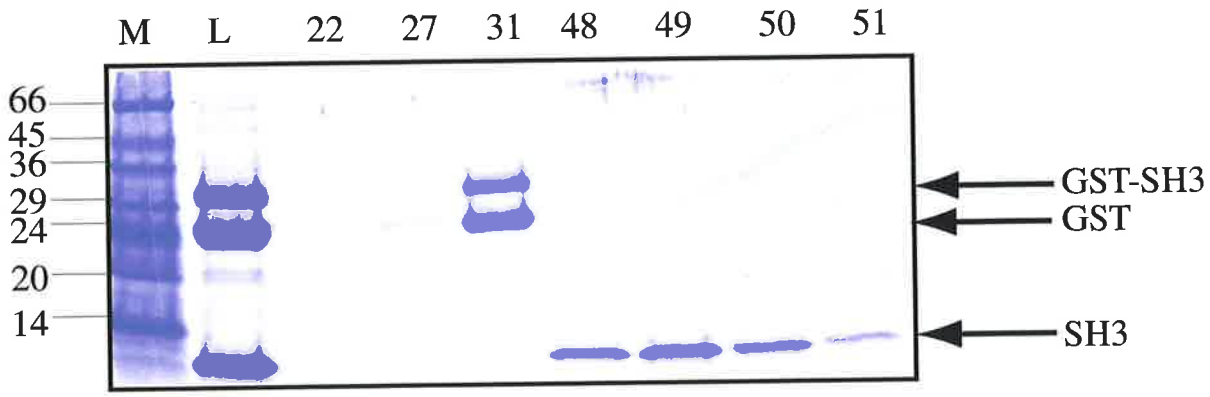
A


M
B
A
66
45
36
29
24
20
14
GST-SH3
GST
Tec SH3

B


1.0
31
27
22
48
51
29
23
17
14
A280
Volume
ml
150
225
300

C


M
L
22
27
31
48
49
50
51
66
45
36
29
24
20
14
GST-SH3
GST
SH3

8.0 (section 2.4.11). 24 mg of GST-SH3 protein was digested with thrombin overnight and loaded on the Superdex G75 column. The GST-SH3 protein was fully digested and 3 mg of purified Tec SH3 domain was recovered. The Superdex beads efficiently separated the 26 kDa GST fusion partner from the 8 kDa SH3 domain (Figure 4.2B). Fractions 48-52 corresponding to the SH3 domain protein were pooled (Figure 4.2C) and concentrated using an Amicon stirred cell to a final volume of 2 ml. Fast desalting PD10 columns (Pharmacia) were used to buffer exchange the protein into 10 mM phosphate, 0.01% (w/v) NaN_3, pH 6.0 for NMR analysis. The protein was further concentrated using Centricon spin columns (3 kDa cut off) to a final volume of 500 µl. Figure 4.3A shows Tec SH3 domain sequence. Although denaturing polyacrylamide gel electrophoresis was used to monitor the protein throughout the purification procedure, mass spectrometric analysis of all the final samples was undertaken. The experimental molecular mass of Tec SH3 domain of 7899 agreed with expected values computed from the amino acid sequence shown in Figures 4.3 A and B. The purity of Tec SH3 domain is greater than 95% as observed by mass spectrometry.

Structure determination by NMR is hampered by overlap in the 2D spectra. Heteronuclear single quantum coherence (HSQC) based experiments used for structure determination can reduce the spectral overlap and this is facilitated by using a ^{15}N enriched sample. Tests were conducted on GST-SH3 protein grown in Min A medium, a salt based system lacking the yeast extract present in LB medium. Proteins can be isotopically labelled by growing bacteria in a defined minimal medium, such as Min A (Miller, 1972), containing ^{15}N-labelled NH_4Cl as the only nitrogen source.

To determine if NMR samples of SH3 protein can be generated from bacteria grown in Min A medium, GST-SH3 fusion protein was expressed and analysed by polyacrylamide gel electrophoresis. Expression of GST-SH3 protein in minimal media was observed by SDS-PAGE (Figure 4.1A). 150 mg/L of GST-SH3 fusion protein was recovered following affinity purification on glutathione agarose columns from protein grown in minimal medium. Protein yields were reduced compared with that of the rich medium possibly due to the increased stress on the bacteria to produce the protein with limited nutrients in the broth. This protein was cleaved with thrombin and the SH3 protein was purified from the GST fusion partner using Superdex G75 gel filtration (section 2.4.11). Analysis by mass spectroscopy indicated that purified labelled SH3 domain protein had a molecular weight of M_r 7994, consistent with the expected value and confirmed uniform ^{15}N incorporation (Figure 4.3). The purity of the sample is greater than 95% as observed by mass spectrometry. The cleavage of the ^{15}N labelled GST-fusion protein sometimes resulted in the production of a

Figure 4.3

A. Primary sequence of the Tec SH3 domain representing amino acids 181-245 of the Tec IV sequence. The N-terminal residues Gly^{-2} and Ser^{-1} are derived from the fusion partner after thrombin digestion.

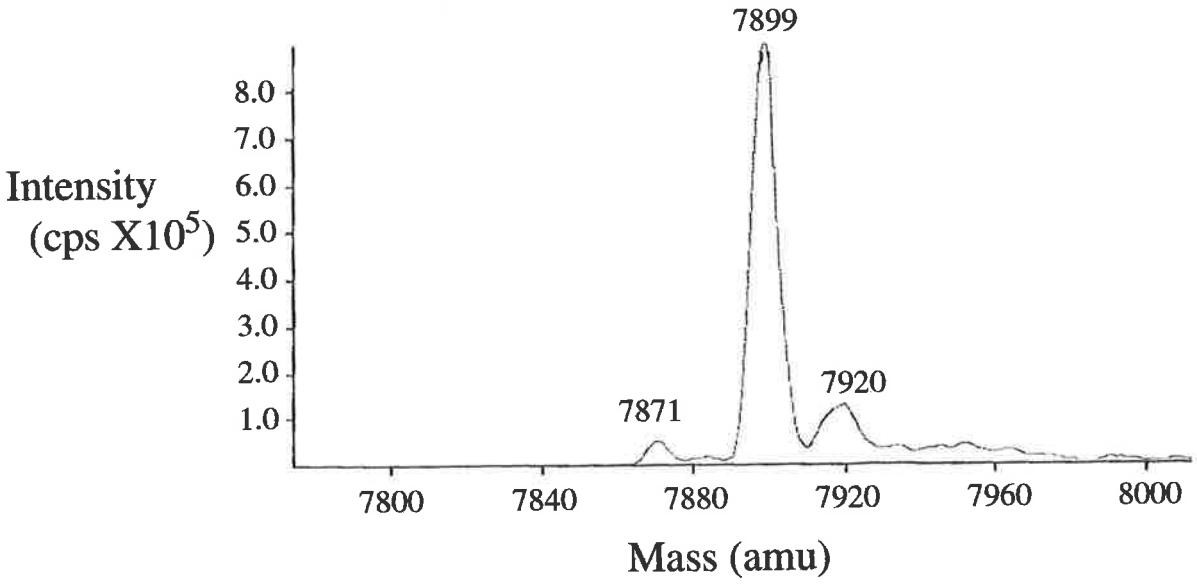
B. Mass spectrometry profile of Tec SH3 domain produced in LB medium. 2 µl of Tec SH3 domain was diluted into methanol/acetonitrile and analysed at the Waite mass spectrometry unit, Adelaide. The purity of Tec SH3 domain protein was greater than 95% as observed by mass spectrometry. A size of 7899 Da is consistent with the expected mass for Tec SH3 domain.

C. Mass spectrometry profile of Tec SH3 domain produced in min A medium supplemented with $^{15}NH_4Cl$. The size of 7994 Da is consistent with the expected mass. The purity of ^{15}N SH3 domain is greater than 95%. The labelled SH3 domain sample is fully labelled as there is no evidence of an unlabelled contaminating peak.

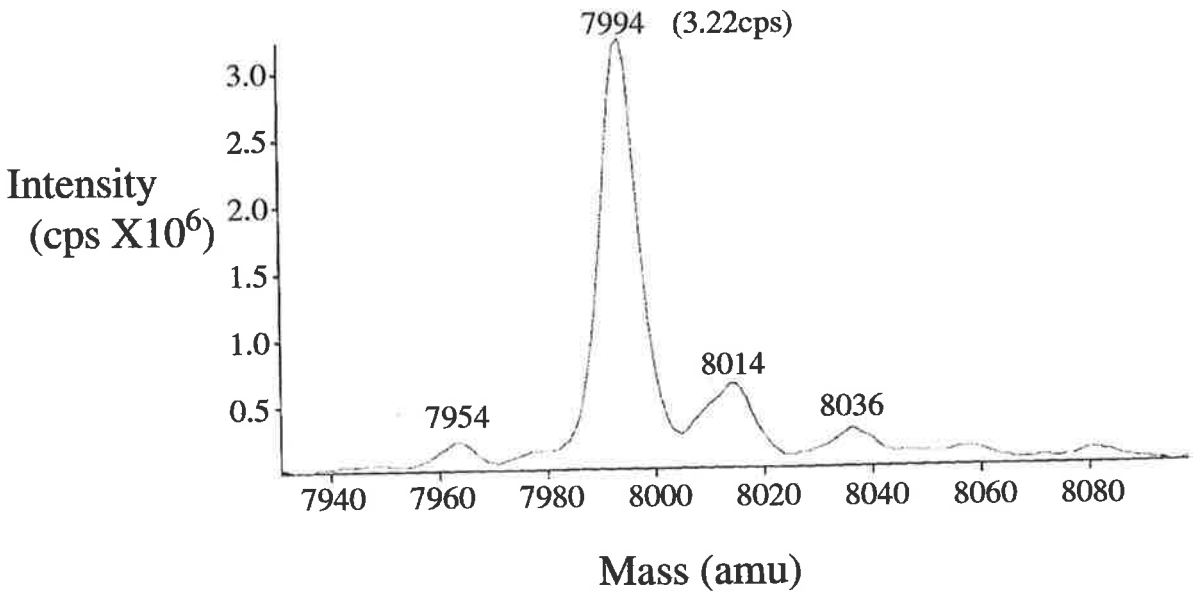
A

```
-2 -1 181                                  190                                  200                                  210
G S E I V V A M Y D F Q A T E A H D L R L E R G Q E Y I I L E K N D L

                  220                                  230                                  240
H W W R A R D K Y G S E G Y I P S N Y V T G K K S N N L D Q Y D
```

B



C



second cleavage product of lower molecular mass than that expected for the SH3 domain. This fraction was not further analysed and was not used for the production of Tec SH3 NMR samples.

4.5 NUCLEAR MAGNETIC RESONANCE DATA COLLECTION AND SEQUENTIAL ASSIGNMENT OF TEC SH3 DOMAIN

SH3 protein sample concentrations of 2.0 mM and 1.25 mM, as determined by Bradford analysis, were used in this study for the unlabelled and ^{15}N labelled samples, respectively. 500 µl of these samples in 10 mM NaH_2PO_4, 0.01% (w/v) NaN_3, pH 6.0 were supplemented with 10% (v/v) D_2O and placed in a 5 mm high-resolution thin-walled glass NMR tube. Samples for D_2O experiments were freeze dried twice before resuspending in 100% D_2O. NMR tubes were prepared for NMR as described in 2.5.1.

NMR experiments were performed on Bruker AMX-500 and DRX-600 and Varian Inova 600 spectrometers. All data sets were recorded at 25°C using 5 mm inverse triple resonance $^1H/^{13}C/^{15}N$ pfg probes. In most cases, the carrier frequency was set on the D_2O frequency and solvent presaturation conducted except for the D_2O experiments for which no presaturation was required.

Initial two dimensional experiments were conducted. 2D ^1H-^1H-TOCSY and ^1H-^1H NOESY experiments are shown in Figure 4.4. The NMR data was recorded and processed with the assistance of Dr T. Mulhern, University of Adelaide, Adelaide.

4.5.1 Sequential Assignment

Sequential assignment was conducted as per Wuthrich (1986). The 3D ^{15}N-^1H HSQC-TOCSY and 3D ^{15}N-^1H HSQC-NOESY spectra were used to begin the assignment process. The 3D ^{15}N-^1H HSQC-TOCSY was used to identify spin systems of the different amino acids by comparing the chemical shifts to those expected if the amino acid was in a random coil (Wuthrich, 1986). Identification of the different amino acid types was significantly easier using 3D data sets rather than relying on the 2D data as the spectra were ordered into columns (referred to as strips) based on ^{15}N chemical shift such that ideally one amino acid was represented per strip. One amino acid will have all its protons represented in one strip. The strip was selected based on the nitrogen frequency and the amide proton chemical shift. The other protons are then presented in a column with decreasing proton chemical shift. This is generally the case, however, some amino acids with very close ^{15}N

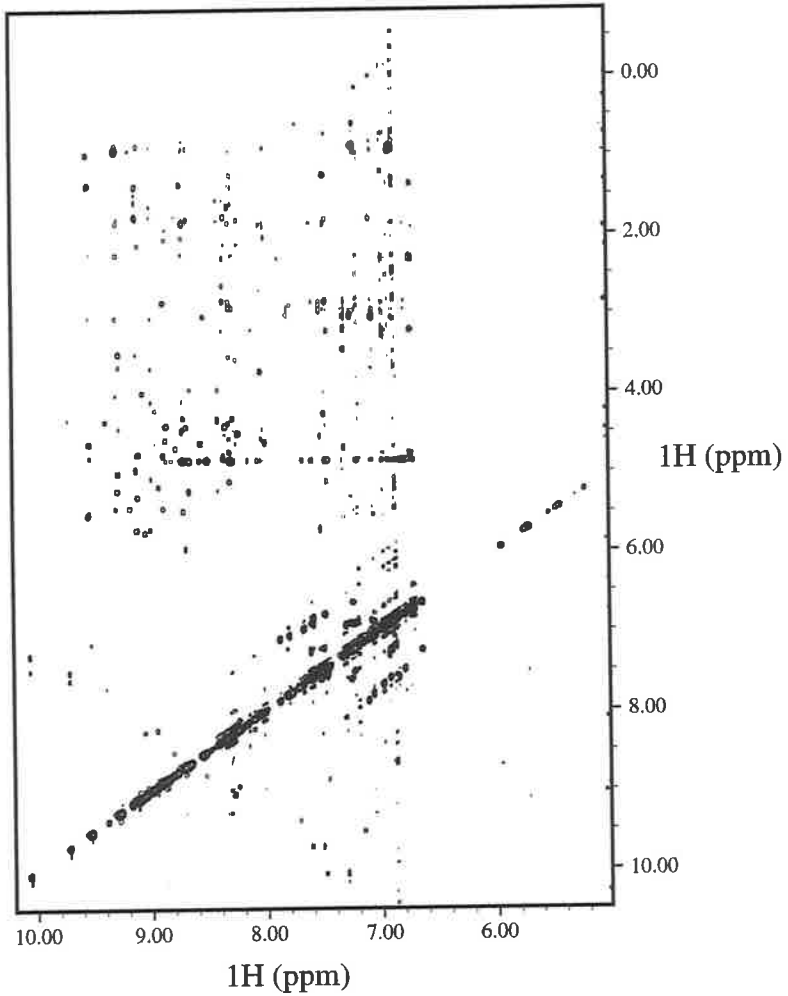
Figure 4.4

Spectra of Tec SH3 domain. The NMR spectra were collected at the BRI, Melbourne on a 2.75 mM (21 mg/ml) sample of Tec SH3 domain in 10 mM Na_2HPO_4, 0.01% w/v NaN_3, pH 6.0 supplemented with 10% D_2O at 25°C. ^1H-^1H NOESY and ^1H-^1H TOCSY of mixing times 200 ms and 50 ms respectively were recorded with a sweep width of 8000 Hz in both dimensions and a minimum of 2048 x 512 complex points. A cosine-squared function was applied to the data prior to Fourier transformation. This data was also zero-filled to a final matrix size of 2048 x 2048. The data was transferred into XEASY (Bartels *et al* 1995) for further analysis. These spectra were plotted from XEASY showing the region from 6 ppm to 10 ppm in both dimensions.

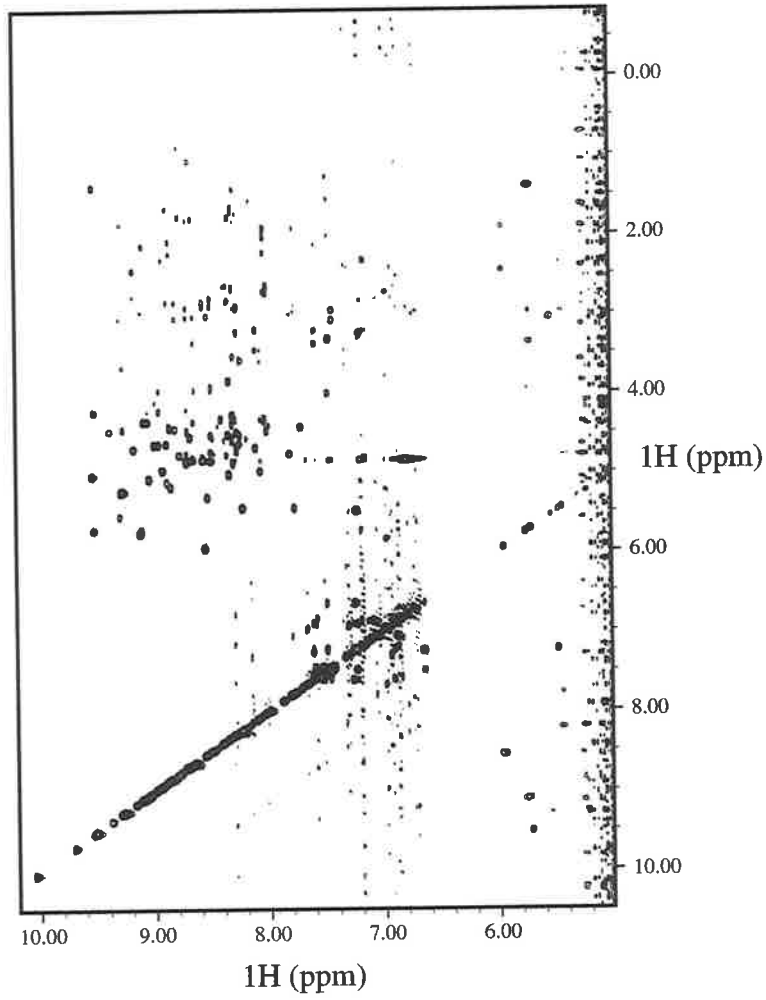
A. ^1H-^1H NOESY spectra of Tec SH3 domain.

B. ^1H-^1H TOCSY spectra of Tec SH3 domain.

A. ^{1}H-^{1}H Noesy Tec SH3 domain



B. ^{1}H-^{1}H TOCSY Tec SH3 domain



chemical shifts did reside in the same strip due to the poor digital resolution in the nitrogen dimension; 2000 Hz for 64 data points or 0.5 Hz per data point. This was seen for amino acids Asp 242 and Leu 213 with nitrogen chemical shifts of 120.30 ppm and 120.25 ppm, respectively. The identification of some of the amino acids was determined based on cross peaks from the 3D data alone, others were identified to their amino acid type only. Sequential amino acids were determined by the presence of strong cross peaks present in the 3D ^{15}N-^{1}H HSQC-NOESY spectra that were not present in the 3D ^{15}N-^{1}H HSQC-TOCSY spectra. β-Sheet structure in proteins results in strong C^{α}H-NH NOE cross peaks and, thus, these cross peaks were investigated initially to link some of the amino acids together. The strips associated with amino acids Glu 203 to Ser 224 are represented in Figure 4.5 with arrows representing the sequential NOE peaks used to link the amino acids together. Both TOCSY and NOESY cross peaks are represented in this figure. Once the C^{α}H-NH NOEs were identified, other NOEs were assessed to determine whether they represented NOEs between other protons in the side chains of these amino acids. The strips corresponding to the 3_{10} helix at the C-terminal of the SH3 domain are represented in Figure 4.6, the presence of strong NH_i to NH_{i+1} being indicative of helical properties.

Complete backbone ^{1}H and ^{15}N assignments were obtained for 65 of 67 residues, the exceptions being Gly -2, for which no resonances could be assigned, and Ser -1 for which no backbone NH resonance could be assigned. A contour plot of a ^{15}N-^{1}H HSQC experiment conducted on uniformly ^{15}N labelled Tec SH3 domain is shown in Figure 4.7. This HSQC shows good dispersion in both the ^{15}N and ^{1}H dimensions. Although not evident in the high resolution spectrum, one section (~118-122 ppm ^{15}N) displays some overlap in the three dimensional experiments. All backbone amide resonances except those of the N-terminal Gly -2 and Ser -1 NH resonances are labelled. Arg N^{ε}H, Gln $N^{\varepsilon 2}$H and Asp $N^{\delta 2}$H and Trp $N^{\varepsilon 1}$H resonances have been highlighted. Complete side chain assignments of non-exchangeable protons were obtained with the exception of the $H^{\varepsilon 1}$ protons of His 195 and His 214. Exchangeable side chain groups including all Trp $H^{\varepsilon 1}$, Asn $H^{\delta 2}$, Gln $H^{\varepsilon 2}$ and Arg H^{ε} were assigned, but no Arg H^{η} or Ser and Thr hydroxyl protons could be assigned (table 4.1).

The absence of NOE contacts involving residues beyond 238 and the presence of sharp resonances corresponding to these residues suggest that this region of the protein is flexible. Assignment of Ile 228, Pro 229 and Ser 230 was complicated by the lack of TOCSY transfer beyond the C^{α}H proton in the 3D ^{15}N-^{1}H TOCSY HSQC experiment for Ile 228 and Ser 230 and the significant shift from random coil values of the C^{α}H protons for Pro 229 (3.76 ppm) and Ser 230 (2.74 ppm). These values suggest that the C^{α}H protons of

Figure 4.5

Strip plot of residues 203-224 of Tec SH3 domain. 3D ^{15}N-^{1}H NOESY HSQC and 3D ^{15}N-^{1}H TOCSY HMQC data were recorded using 1024 x 32 x 72 complex points and sweep widths of 1824 Hz in F1 and 7507 Hz in F2 and F3. NOESY and TOCSY mixing times of 200 ms and 50 ms, respectively, were used for these experiments. Cosine-squared functions were applied to the data prior to Fourier transformation. The final matrix size was 896 x 64 x 256. The protein sample was 1.25 mM pH 6.0 buffered in 10 mM NaH_2PO_4 0.01 % w/v NaN_3. All data were collected at 25°C. Processed data was analysed using XEASY and used for generation of the strip plots (Bartels *et al* 1995). Glu 203-Ser 224 strips are shown with the NOESY cross peaks shown in blue and the corresponding TOCSY peaks overlaid in red. The amino acids shown in each strip are indicated at the bottom of the strip along with their nitrogen chemical shifts. The arrows indicate the strong $C\alpha H_i$-$C\alpha H_{i+1}$ cross peaks observed for a protein with β-strand secondary conformation.

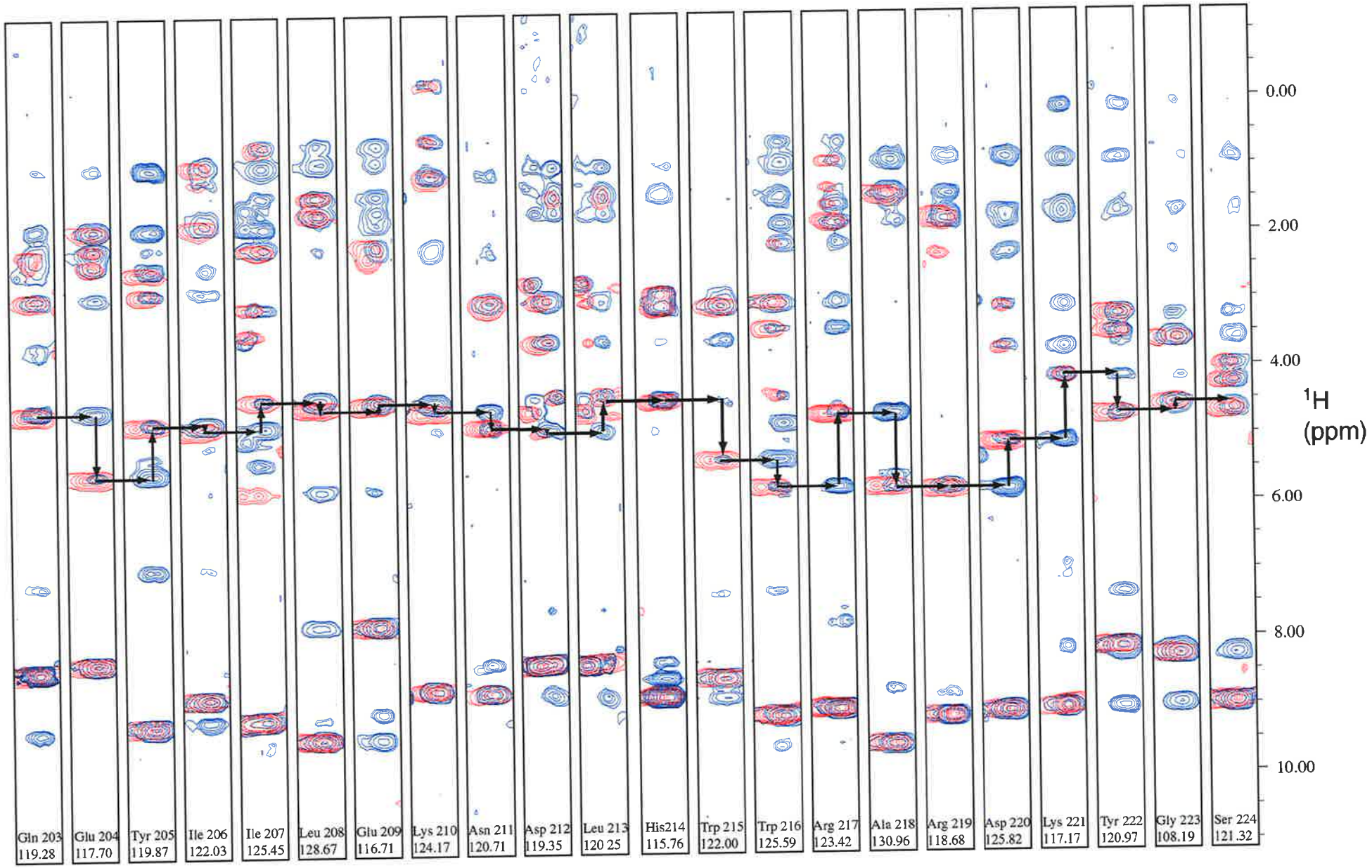

Gln 203
119.28
Glu 204
117.70
Tyr 205
119.87
Ile 206
122.03
Ile 207
125.45
Leu 208
128.67
Glu 209
116.71
Lys 210
124.17
Asn 211
120.71
Asp 212
119.35
Leu 213
120 25
His214
115.76
Trp 215
122.00
Trp 216
125.59
Arg 217
123.42
Ala 218
130.96
Arg 219
118.68
Asp 220
125.82
Lys 221
117.17
Tyr 222
120.97
Gly 223
108.19
Ser 224
121.32
0.00
2.00
4.00
6.00
8.00
10.00
^{1}H
(ppm)

Figure 4.6

Strip plot of residues 228-237 of Tec SH3 domain. The contour strip plots were generated as for figure 4.5 and show the strips for amino acids Ile 228-Lys 237 incorporating the 3_{10} helix of Tec SH3 domain. The amino acids shown in the strips are labelled on each strip along with their nitrogen resonance. Arrows indicate both the strong $C\alpha H_i$-$C\alpha H_{i+1}$ for the β-sheet but also observed in this set of strips is a set of peaks showing strong $N\alpha H_i$-$N\alpha H_{i+1}$ indicative of helical structure. The NOESY cross peaks are shown in blue and the TOCSY cross peaks are shown in red.

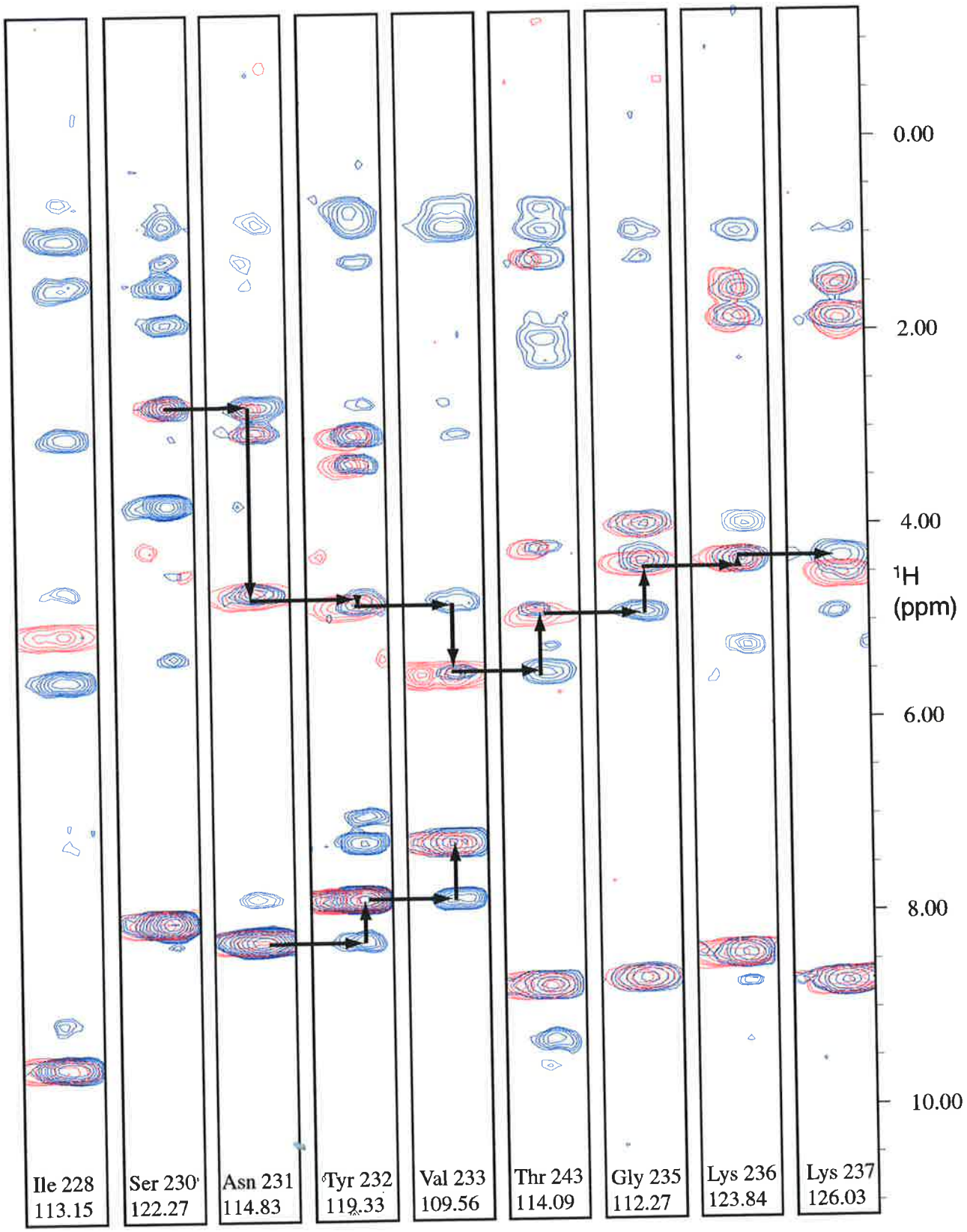

0.00
2.00
4.00
^{1}H
(ppm)
6.00
8.00
10.00
Ile 228
113.15
Ser 230
122.27
Asn 231
114.83
Tyr 232
119.33
Val 233
109.56
Thr 243
114.09
Gly 235
112.27
Lys 236
123.84
Lys 237
126.03

Figure 4.7

Contour plot of a fully assigned 600 MHz 2D ^{1}H-^{15}N HSQC spectrum. Spectra were recorded on a 1.25 mM solution of uniformly ^{15}N-labelled Tec SH3 domain in 90% H_2O/10% D_2O at pH 6.0 and 25 °C. Spectral widths of 2500 Hz in F2 and 6250 Hz in F1 were used, along with 400 t_1 increments. All residues are shown, except for Pro^{229} and the N-terminal dipeptide. All Arg, Asn and Gln side-chain NH resonances have been highlighted and connecting lines drawn between the two NH resonances for Asn and Gln side chains. This figure was generated in XEASY (Bartels *et al* 1995).

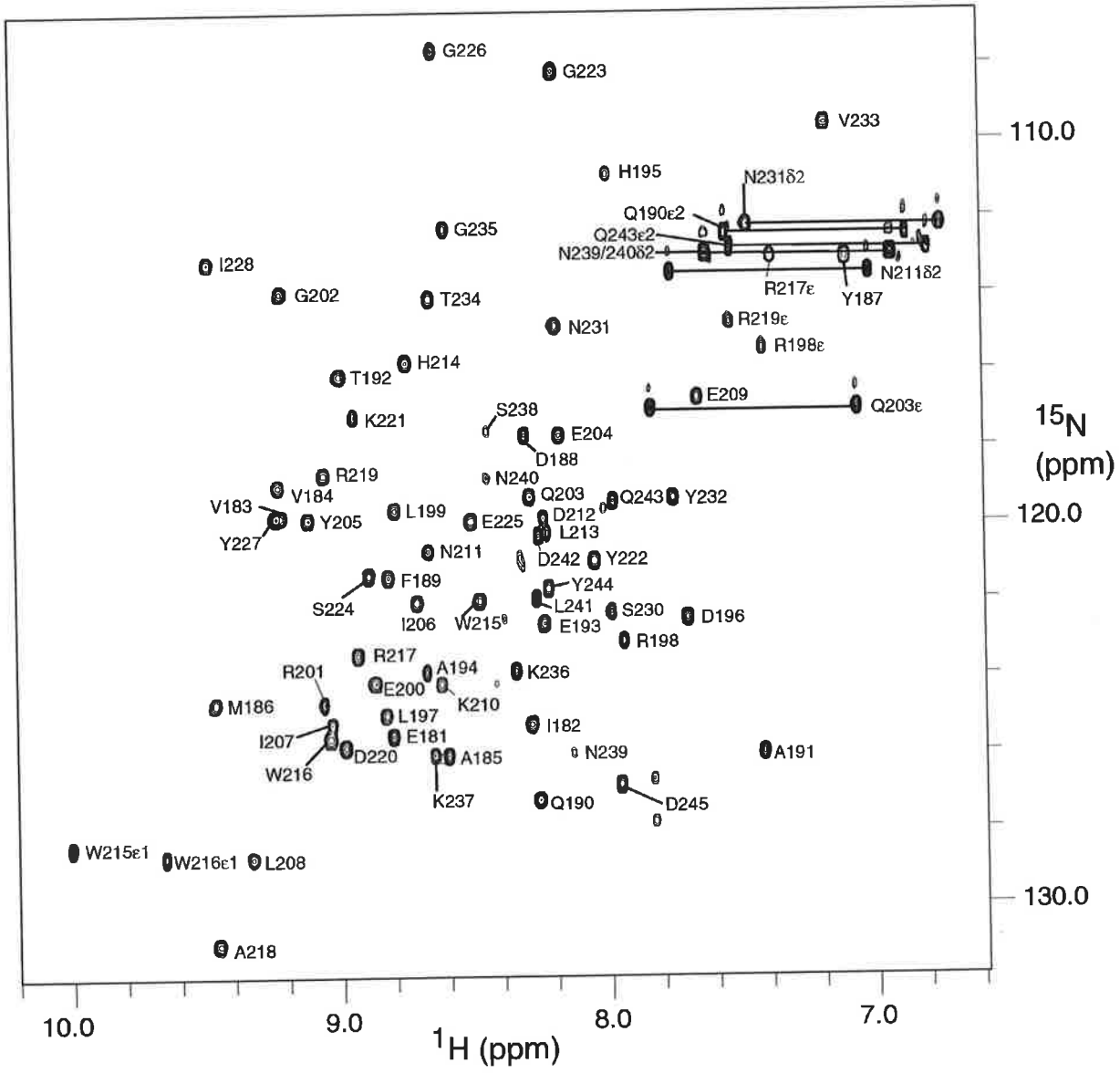

G226
G223
V233
H195
N231δ2
Q190ε2
Q243ε2
N239/240δ2
G235
I228
G202
T234
R217ε
Y187
N211δ2
N231
R219ε
R198ε
H214
T192
E209
Q203ε
K221
S238
E204
D188
R219
N240
V184
Q203
Q243
Y232
V183
L199
E225
D212
Y205
L213
Y227
N211
D242
Y222
F189
Y244
S224
L241
S230
D196
I206
W215
E193
R198
R217
R201
A194
K236
E200
K210
M186
L197
I182
I207
E181
N239
A191
D220
A185
W216
K237
Q190
D245
W215ε1
W216ε1
L208
A218
110.0
120.0
130.0
15N (ppm)
10.0
9.0
8.0
7.0
1H (ppm)

Table 4.1 Complete chemical shift assignments for Tec SH3 domain.

Residue	Residue name	^{15}N	HN	Hα	Hβ	Hγ	Other
-2	Gly						
-1	Ser			4.622	3.926 3.792		
181	Glu	125.52	8.865	4.456	1.978	2.104	
182	Ile	125.225	8.346	5.052	1.824	1.608 1.201 γ2 0.982	δ 1 0.89
183	Val	119.35	9.261	5.266	2.293	1.029 0.986	
184	Val	118.985	9.309	5.27	1.893	0.979 0.935	
185	Ala	126.035	8.687	4.859	1.84		
168	Met	124.637	9.534	4.262	1.403	2.274	ε 1.842
187	Tyr	113.068	7.205	4.851	3.241 2.333		δ 6.77 ε 6.748
188	Asp	117.697	8.39	4.813	2.882 2.671		
189	Phe	121.361	8.911	5.14	3.264 2.892		δ 7.496 ε 7.558 ζ 7.614
190	Gln	127.221	8.34	4.298	1.941 1.887	2.334	ε21 7.623 ε22 6.884 Nε2 112.404
191	Ala	125.971	7.516	4.016	1.293		
192	Thr	116.103	9.087	4.377	4.242	1.224	
193	Glu	122.635	8.312	4.507	1.776 1.695	1.93 1.87	
194	Ala	123.856	8.771	4.141	1.431		
195	His	110.873	8.1	5.001*	3.619* 3.401*		δ 2 7.124*
196	Asp	122.465	7.808	5.46	2.947 2.901		
197	Leu	124.975	8.91	4.661	1.856 1.716	1.707	δ 0.76
198	Arg	123.053	8.036	4.448	2.082 1.942	2.036 1.585	ε 7.506 Nε 115.135 δ 3.358
199	Leu	119.616	8.893	5.206	2.273 1.289	1.842	δ1 0.778 δ2 0.056
200	Glu	124.136	8.953	4.997	2.144 2.011	2.33	
201	Arg	124.666	9.108	3.549	1.819 1.619	1.538 1.477	δ2 3.319 δ3 3.272 ε 7.236
202	Gly	113.936	9.3	4.474 3.704			
203	Gln	119.278	8.378	4.546	2.864 2.181	2.527 2.334	ε21 7.123 ε22 7.897 Nε116.998
204	Glu	117.707	8.257	5.481	1.871	2.37 2.125	

...Cont'd

* assigned using D_20 spectra

^1H-^1H Homonuclear and ^1H-^{15}N heteronuclear experiments were used for the assignment of Tec SH3 domain. Chemical shifts were referenced to 4.85 ppm and 120 ppm for the proton and nitrogen dimensions respectively.
*assigned using D_20 spectra

Table 4.1 cont'd *Complete chemical shift assignments for Tec SH3 domain.*

Residue	Residue name	^{15}N	HN	Hα	Hβ	Hγ	Other
205	Tyr	119.866	9.198	4.73	2.814 2.483		δ6.914 ε7.12
206	Ile	122.035	8.803	4.803	1.82	1.634 1.146 γ2 0.913	δ 0.886
207	Ile	125.447	9.128	4.39	2.173	1.847 1.375 γ2 0.655	δ 0.637
208	Leu	128.686	9.407	4.504	1.399	1.654	δ1 0.87 δ2 0.837
209	Glu	116.709	7.744	4.456	2.19 2.085	2.319	
210	Lys	124.17	8.674	4.584	1.125 1.066	0.561 -0.264	δ2 1.514 δ3 0.882 ε 2.191
211	Asn	120.705	8.739	4.789	3.071 2.956		δ21 7.805 δ22 7.053 Nδ 113.447
212	Asp	119.353	8.33	4.883	3.554 2.964		
213	Leu	120.249	8.285	4.347	1.648 1.429	1.28	δ1 0.994 δ2 0.963
214	His	115.762	8.836	4.421*	3.044* 2.824*		δ 6.732
315	Trp	121.997	8.57	5.341	3.071		δ1 7.345 ε1 10.082 ε3 6.681 Hh2 7.261 ζ2 7.52 ζ3 7.296 Nε 128.26
216	Trp	125.592	9.133	5.76	3.408 3.017		δ1 7.547 ε1 9.748 ε3 6.934 Hh2 7.288 ζ2 7.637 ζ3 7.255 Nε 128.648
217	Arg	123.424	9.022	4.687	1.831 1.593	1.325 0.942	δ2 3.095 δ3 3.005 ε 7.473 Nε 112.791
218	Ala	130.962	9.548	5.753	1.434		
319	Arg	118.677	9.134	5.8	2.311 1.778	1.873 1.675	δ2 3.412 δ3 3.255 ε 7.624 Nε 114.74
220	Asp	125.823	9.067	5.103	3.711 3.095		
221	Lys	117.167	9.013	4.139	1.794 1.686	0.946 0.168	δ2 1.565 δ3 1.504 ε 2.816

...Cont'd

* assigned using D_20 spectra

Table 4.1 cont'd Complete chemical shift assignments for Tec SH3 domain.

Residue	Residue name	^{15}N	HN	Hα	Hβ	Hγ	Other
222	Tyr	120.972	8.136	4.713	3.486 3.251		δ 6.972 ε 7.344
223	Gly	108.188	8.276	4.555 3.617			
224	Ser	121.322	8.979	4.672	4.254 3.996		
225	Glu	119.905	8.587	5.993	1.964 1.929	2.487 2.186	
226	Gly	107.627	8.717	4.139 4.047			
227	Tyr	119.798	9.315	5.582	3.14 3.098		δ 6.921 ε 7.122 H?13.087
228	Ile	113.151	9.558	5.067	1.85	1.531 1.034 γ2 1.467	δ 0.651
229	Pro	-		3.759	1.245 0.887	0.677 0.307	δ 2 2.633 δ 3 2.529
230	Ser	122.273	8.069	2.739	1.879 1.511		
231	Asn	114.832	8.261	4.668	3.007 2.763		δ21 6.813 δ22 7.539 Nδ 112.2
232	Tyr	119.333	7.839	4.803	3.346 3.063		Qδ7.015 Qε6.993
233	Val	109.564	7.258	5.522	2.086	0.929 0.74	
234	Thr	114.089	8.713	4.896	4.232	1.256	
235	Gly	112.272	8.674	4.387 3.994			
236	Lys	123.837	8.395	4.349	1.857	1.449	Qδ1.607 Qε2.98
237	Lys	126.026	8.735	4.51	1.937 1.835	1.543 1.5	Qδ1.775 Qε3.088
238	Ser	117.478	8.525	4.546	3.982 3.93		
239	Asn	125.916	8.222	4.592	2.919 2.799		Hδ217.68 Hδ226.988 Nδ2112.963
240	Asn	118.885	8.529	4.765	2.913 2.814		Hδ217.682 Hδ226.989 Nδ2112.963
241	Leu	121.697	8.319	4.422	1.75	1.692	Qδ11.025 Qδ20.962 H?41.026
242	Asp	120.407	8.326	4.609	2.717		
243	Gln	119.412	8.022	4.359	2.065 1.962	2.264	Nε2112.806 Hε21 6.839 Hε227.589
244	Tyr	121.674	8.305	4.743	3.277 2.955		Qδ7.226 Qε6.908
245	Asp	126.808	8.041	4.485	2.761 2.67		

* assigned using D_20 spectra

residues 229 and 230 are in close proximity to an aromatic ring, most likely that of the Trp 216.

The chemical shift deviation of $C^{\alpha}H$, NH and N from random coil, plotted against residue number is presented in Figure 4.8. The $C^{\alpha}H$ chemical shift deviation from random coil values is a useful prediction of secondary structure and the plot shown in Figure 4.8D indicates the presence of seven β-strands (Wishart *et al.*, 1995); (Wishart and Sykes, 1994); (Wishart *et al.*, 1992). The secondary structural elements are not as obvious for the HN and N chemical shifts as with the $C^{\alpha}H$ chemical shifts due to a decrease in digital resolution along the nitrogen frequency (Figure 4.8A B and C). The plots indicate that the secondary structure of Tec SH3 domain is predominantly β-strand with seven β-strands and a small helix (α1) present between amino acids 229-231. This pattern of secondary structure is similar to that of other published SH3 domains' suggesting that α1 is likely to be a 3_{10} helix.

4.6 STRUCTURE DETERMINATION OF TEC SH3 DOMAIN

Distance restraints were derived from a 2D ^{1}H-^{1}H-NOESY spectrum (recorded in D_2O) and a 3D ^{1}H-^{15}N-NOESY-HSQC spectrum, with mixing times of 200 ms. Distance restraint upper bounds were calibrated using the method of Xu *et al* (2000) so that upper bounds ranged from 6.0 to 2.2 Å. Lower bounds were set to 1.8 Å. Hydrogen bond restraints were imposed for amide groups that had not exchanged ^{1}H for ^{2}D within 1 hour at 20°C after resuspension of a lyophilised sample (from H_2O) to D_2O. The hydrogen-acceptor distance was restrained between 1.7 and 2.2 Å and the donor-acceptor distance was restrained between 2.7 and 3.2 Å. A complete list of distance restraints is shown in Appendix 1.

Values of $^{3}J_{HN\text{-}H\alpha}$ were measured from F1 and F2 cross peak line widths in a ^{1}H-^{15}N-HMQC-J experiment using the method of Wishart and Wang (1998). Torsion angle restraints for ϕ-angles of -120 ± 40° were imposed for $^{3}J_{HN\text{-}H\alpha}$ • 8 Hz and –60 ± 30° for • 5 Hz. Dr T. Mulhern, Biochemistry Department, University of Adelaide, Adelaide kindly carried out the structure calculations.

Structures were calculated in X-PLOR (version 3.851) (Brunger, 1991) using the ARIA method of Nilges *et al.*, (1997). A raw set of distance restraints was extracted from the NOESY peak list by interrogating the chemical shift assignments using windows of ±0.04 ppm for F1 and ±0.02 ppm for F2 in the 2D ^{1}H-^{1}H NOESY and windows of ±0.06 ppm for F1 (^{1}H), ±0.25 ppm for F2 (^{15}N) and ±0.03 ppm for F3 (^{1}H) in the 3D ^{15}N-^{1}H

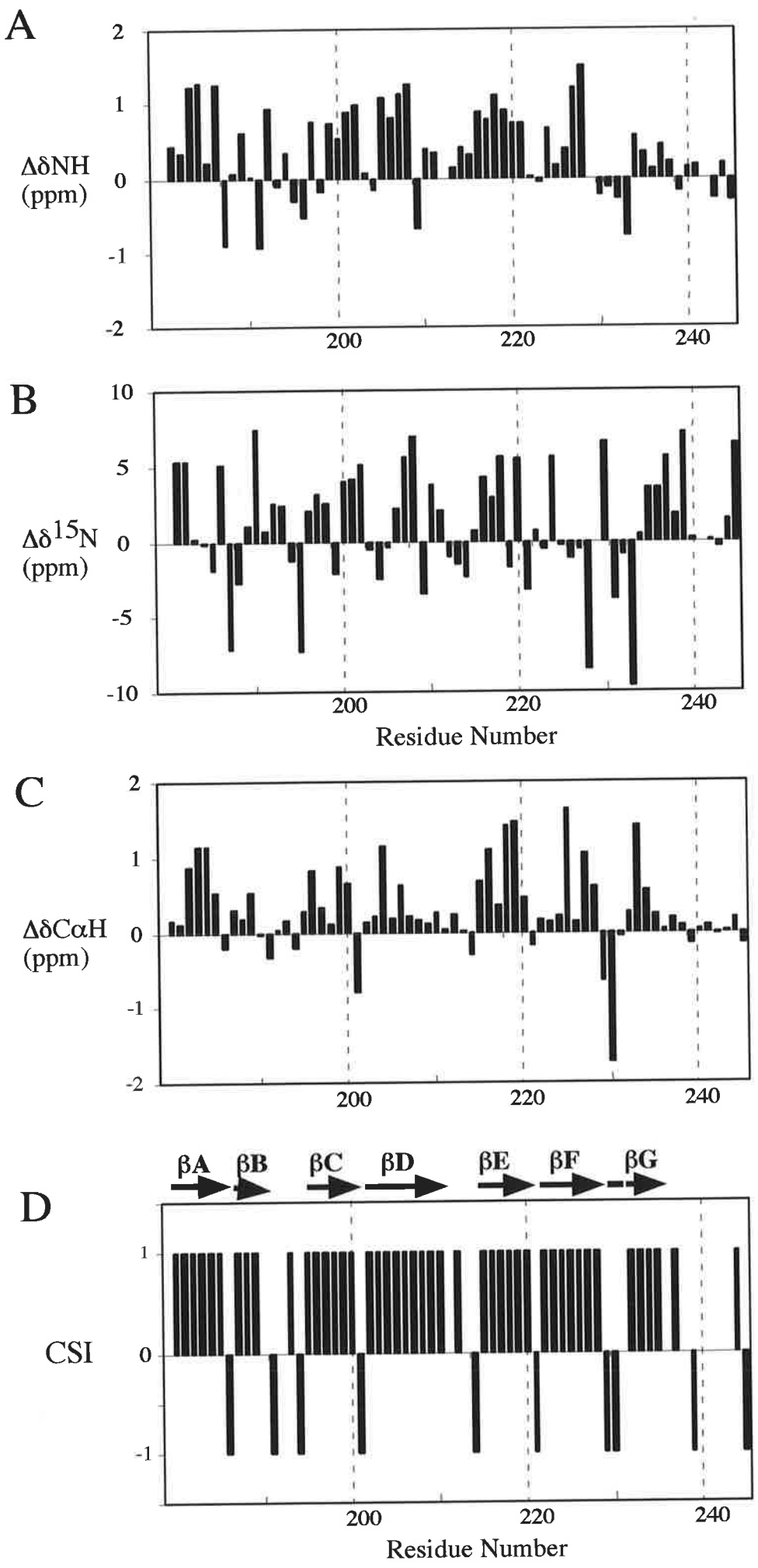
Figure 4.8

A. Deviation of NH chemical shifts from random coil values (Wishart *et al*, 1995), plotted against residue number. Sections of three or more with values greater than 0.1 ppm indicate β-sheet, whereas values of less than -0.1 ppm indicate α-helix. NH chemical shifts are limited by limited spectral resolution.

B. Deviation of ^{15}N chemical shifts from random coil values (Wishart *et al.*, 1995), plotted against residue number.

C. Deviation of CαH chemical shifts from random coil values (Wishart *et al*, 1995), plotted against residue number. Sections of three or more with values greater than 0.1 ppm indicate β-sheet, whereas values of less than -0.1 ppm indicate α-helix.

D. Chemical shift index plotted against residue number (Wishart *et al*, 1992). Deviations in chemical shift greater than 0.1 are given a value of 1 and those lower than -0.1 are given a value of -1. Chemical shift index plotted against residue number highlights regions of β-strand and α-helix. Regions of β-strands predicted from this analysis are shown in black boxes and are labelled βA-βG. A region of helix is predicted and is shown as a line labelled α1.


A
2
1
0
-1
-2
ΔδNH
(ppm)
200
220
240
B
10
5
0
-5
-10
Δδ15N
(ppm)
200
220
240
Residue Number
C
2
1
0
-1
-2
ΔδCαH
(ppm)
200
220
240
βA
βB
βC
βD
βE
βF
βG
D
1
0
-1
CSI
200
220
240
Residue Number

NOESY HSQC. An initial ensemble of 40 structures was calculated using the raw restraints, which were composed of 49 unambiguous restraints (22 intraresidue, 6 sequential, 2 medium range and 19 long-range restraints), 1566 ambiguous restraints and 31 ϕ-angle restraints. The ten structures with the lowest overall energy were retained. Ten ARIA iterations followed during which the assignment parameter (N_p) was reduced from 0.9999 to 0.75 and the violation tolerance (V_{tol}) was reduced from 2.0 to 0 Å. Restraints were excluded if they were violated in more than 50% of the retained structures from the previous iteration. Hydrogen bond restraints were included for the slowly exchanging amide groups when a single acceptor was found with suitable hydrogen bonding geometry in at least 50 % of the retained structures. Restraint lists were filtered to remove redundancy with the restraint with the lowest upper bound being retained. A total of 100 structures were calculated using the final restraint list (generated with N_p of 0.75 and V_{tol} of 0 Å). These final restraints included 1228 NOE-derived distance restraints. Of these, 956 were assigned unambiguously (468 intraresidue, 173 sequential, 55 medium range and 260 long range) and 272 restraints for which the assignment remained ambiguous. The remaining restraints were 34 hydrogen bond restraints (two restraints per hydrogen bond) and 31 ϕ-angle restraints. A final round of refinement of the 100 structures from the last ARIA iteration was carried out in explicit solvent using the OPLSX non-bonded parameter set (Linge and Nilges, 1999). The 20 structures with the lowest overall energies were selected as the final ensemble. A list of the structural statistics of Tec SH3 domain is shown in Table 4.2.

Hydrogen bonding and secondary structure were analysed using MOLMOL (Koradi *et al.*, 1996) and Ramachandran properties and angular order parameters were measured using the in-house program ANGORDER and PROCHECK.

The well-defined residues were identified by iterative fitting of the C^{α} atoms to define the subset with the best-defined C^{α} positions. Following each iteration the C^{α} atom with the highest RMSD was excluded before the next fit. At each iteration, the C^{α} RMSD, for the retained subset, was divided by the number of residues in the subset. The set of best-defined residues was the subset for which this ratio was the minimum.

4.6.1 Tec SH3 domain

The final 20 structures are shown in Figure 4.9. The number of NOEs observed per residue is represented in Figure 4.10B. Residues $^{-2}$Gly and $^{-1}$Ser, which remain following proteolytic cleavage from the fusion partner, have very few medium and long range NOEs:

Table 4.2 Structural statistics for Tec SH3 domain

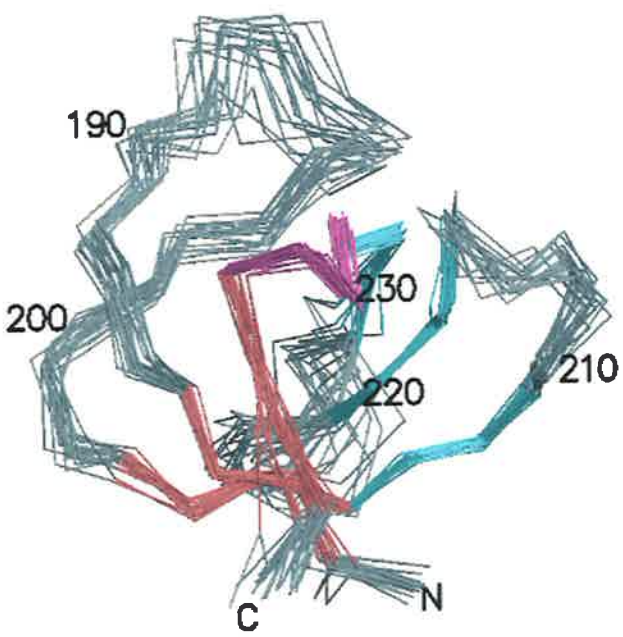
RMSD from experimental distance restraints	
956 Unambiguous (Å)	$2.16 \times 10^{-2} \pm 2.1 \times 10^{-3}$
272 Ambiguous (Å)	$2.07 \times 10^{-2} \pm 4.5 \times 10^{-3}$
17 Hydrogen bond (Å)[a]	$2.33 \times 10^{-2} \pm 7.0 \times 10^{-3}$
31 Dihedral angle (°)	$4.81 \times 10^{-1} \pm 2.73 \times 10^{-1}$
r.m.s. deviation *versus* average structures (Å)[b]	
Backbone (N, C',C')	3.21 ± 0.90, 0.68 ± 0.13, 0.48 ± 0.06
All heavy atoms	3.67 ± 0.90, 1.26 ± 0.15, 1.07 ± 0.12
Deviation from idealized geometry	
Bond (Å)	$3.93 \times 10^{-3} \pm 1.8 \times 10^{-4}$
Angle (°)	$5.86 \times 10^{-1} \pm 3.7 \times 10^{-2}$
Improper (°)	$5.18 \times 10^{-1} \pm 3.1 \times 10^{-2}$
Non-bonded energies in CSDX/OPLS force field (kcal.mol^{-1})[c]	
vdW	-257.7 ± 18.7
Elec	-1959.1 ± 107.6
Residues in allowed ϕ/ψ regions of the Ramachandran plot (%)[d] using the program ANGORDER	
Most favored region	66.7 (81.9)
Additionally allowed region	27.6 (17.8)
Generously allowed region	27.6 (17.8)
Disallowed regions	2.4 (0.1)

[a]Two restraints per hydrogen bond giving 34 restraints in total. [b]The three values shown represent all 67 residues, 53 non-terminal residues (excluding: Gly^{-2}, Ser^{-1}, Glu^{181} & Gly^{235}-Asp^{245}) and the 43 well defined (Ile182-Ala191, His195-Lys210, His214-Asp220, Glu225-Thr234) residues, respectively. See text for definition of well defined. [c]Lennard-Jones and electrostatic energies evaluated using the default OLPSX parameters for protein only. [d]Values for the 43 well defined residues are in parentheses.

Energy constants used were 500 kcal mol^{-1} rad^{-2} for the bond angles and improper dihedral angles and 1000 kcal mol^{-1}Å^{-2} (Linge and Nilges, 1999).

Figure 4.9

An overlay of the final Tec SH3 domain ensemble. Structures were generated using the X-PLOR program using dynamical simulated annealing protocol (Nigles *et al* 1991). Ambiguous distance restraints were generated for every peak on the NOESY spectrum based purely on chemical shift considerations. These were fitted iteratively against preliminary structures to unambiguously define interacting groups using the Ambiguous Restraints for Iterative Assignment (ARIA) protocol (Nilges *et al* 1997). The ensemble shown represents the best 20 structures superimposed on a single structure. Represented in red is β-sheet 1 composed of β-strands A, B and E whereas β-sheet 2, composed of β-strands B´, C and D, is shown in blue. The 3_{10} helix is shown in magenta spanning amino acids 230-231 in Tec kinase. The numbering on the structures is consistent with the amino acid number of full-length Tec kinase. The figure was generated in Molscript and Raster3D (Kraulis *et al* 1991; Merritt *et al.*, 1997).


190
200
230
210
220
C
N

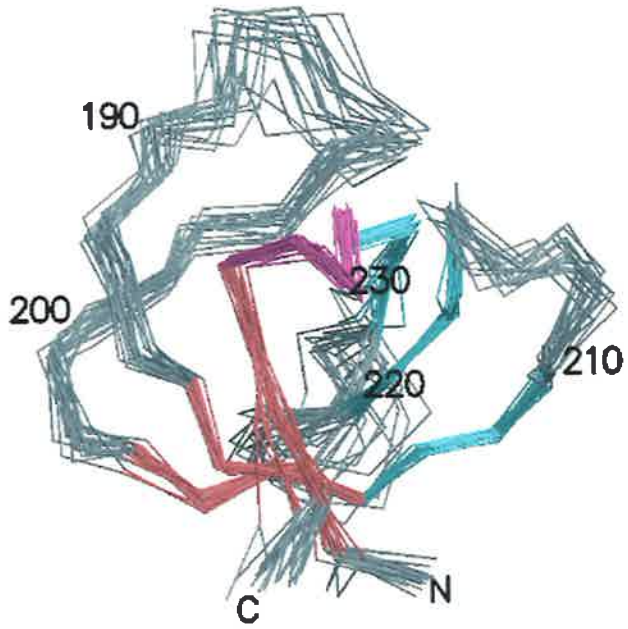
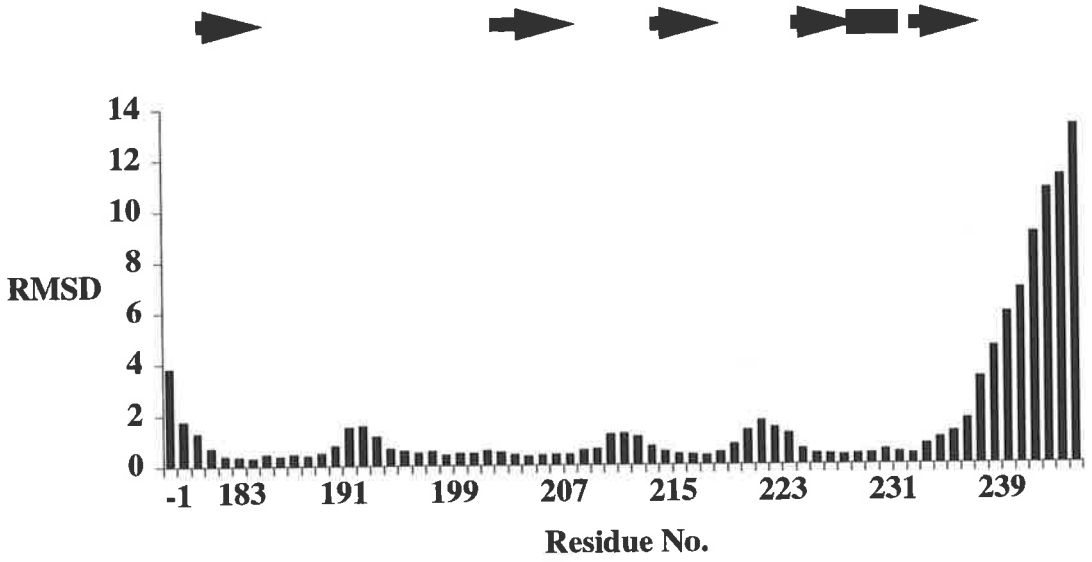

190
200
230
210
220
C
N

Figure 4.10

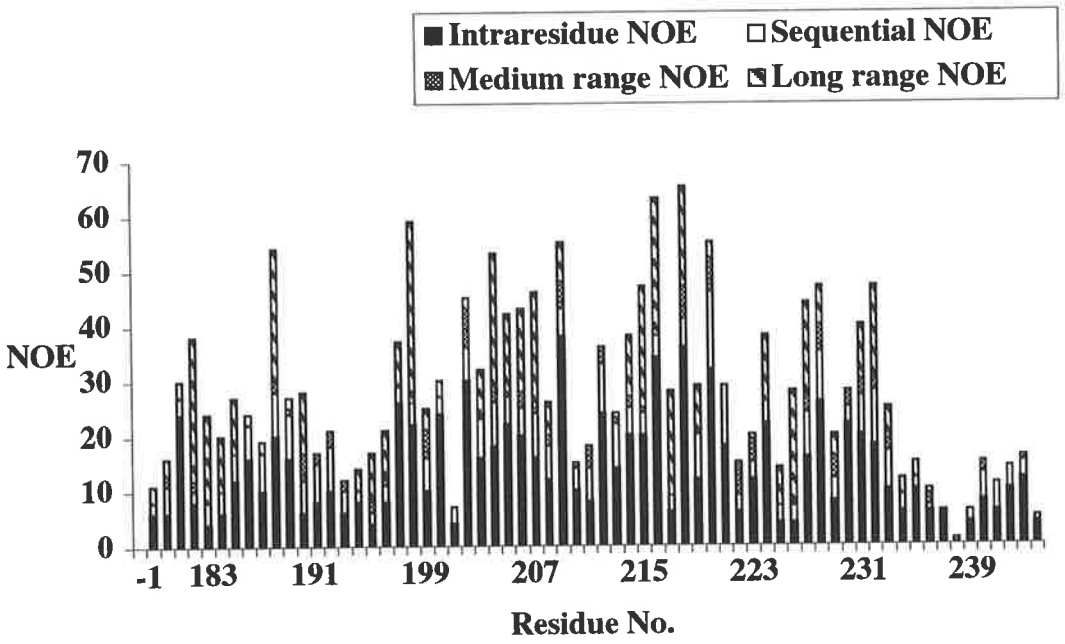
A. The RMSD of Cα within the ensemble of Tec SH3 domain. The RMSD is plotted against residue number and shows areas of high motility. Areas of low motility is shown by low RMSD values and is indicated by the arrows above the graph. The rectangle represents the 3_{10} helix present in the C-terminal end of Tec SH3 domain..

B. Graph of the total NOEs used in the Tec SH3 domain structure calculations. The NOEs used in the calculations are plotted against residue number highlighting well-defined regions and more flexible regions.

A


RMSD
14
12
10
8
6
4
2
0
-1
183
191
199
207
215
223
231
239
Residue No.

B


Intraresidue NOE
Sequential NOE
Medium range NOE
Long range NOE
NOE
70
60
50
40
30
20
10
0
-1
183
191
199
207
215
223
231
239
Residue No.

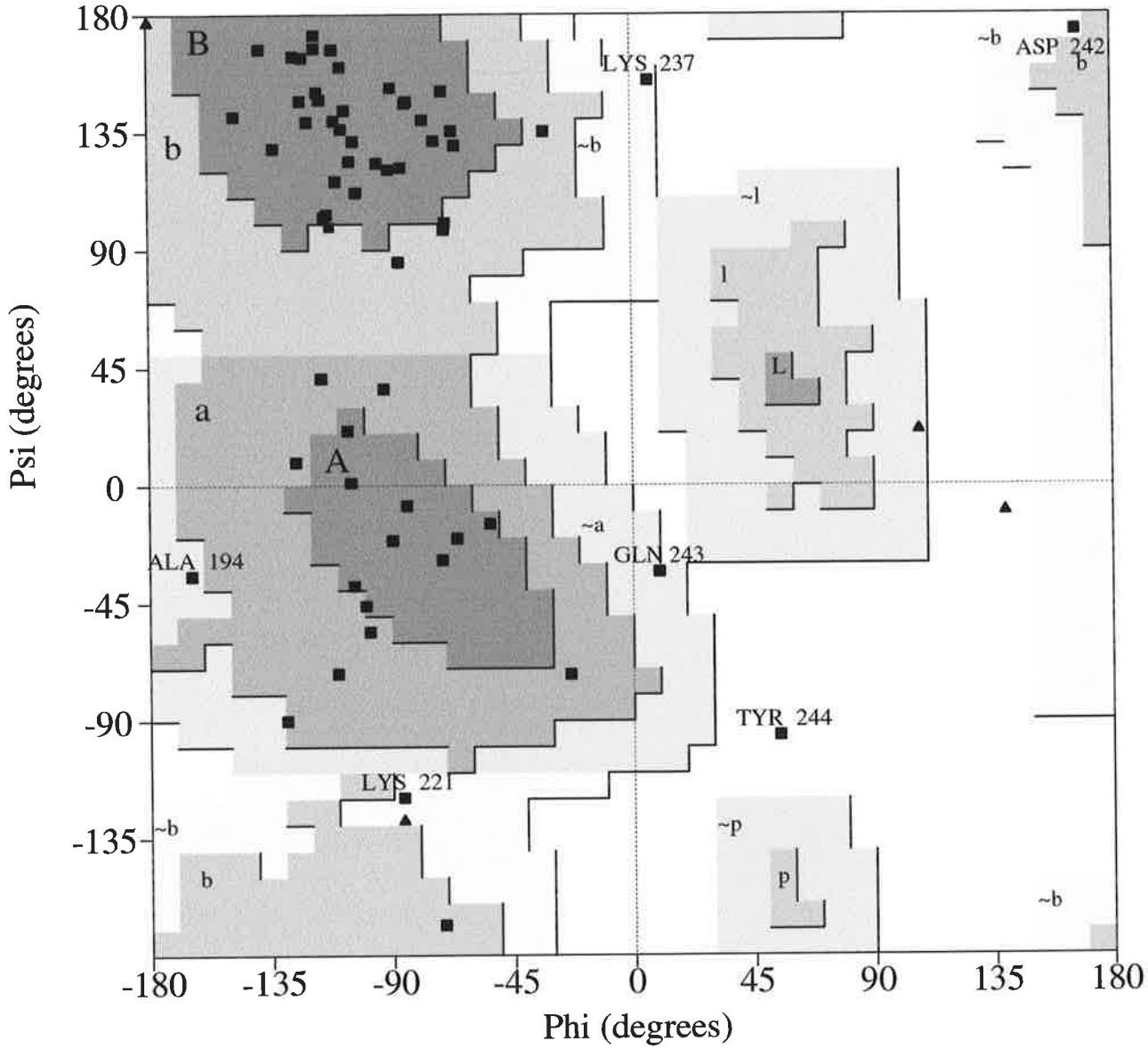
Gly^{-2} has 5 and Ser^{-1} has none. Glu 181 has only three long range NOEs suggesting the N-terminal region is disordered. The C-terminal region, amino acids Ser 236-Asp 245, is also poorly defined with very few medium and long range NOEs suggesting a disordered structure. Sections of the AB or RT loop, amino acids Phe 189-Ala 191 and His 195-Leu 199 of Tec SH3 domain are well defined with many NOEs and possible reciprocal hydrogen bonds between Leu 199 and Phe 189 that generate a well-defined structure. The β-strands of Tec SH3 domain are well defined with many medium and long-range contacts. RMSD values plotted against residue number indicate well-defined and poorly defined regions (Figure 4.10A) and confirms that the strand regions are well defined. The RMSD to the mean co-ordinates of the well-defined residues of Tec SH3 domain is 0.48 ± 0.06 Å for an ensemble of 20 final structures. The Ramachandran plot generated in PROCHECK of Tec SH3 domain is presented in Figure 4.11. The Phi and Psi angles of the amino acids in Tec SH3 domain are plotted against each other and compared with regions of allowed Phi and Psi angles based on steric hindrance and known angles for different amino acids. The only amino acid that falls in the disallowed region is Tyr 244, which is located on the C-terminal flexible tail of the SH3 domain. Other amino acids found in the generously allowed regions include Lys 221, Lys 237, Asp 242 and Gln 243 of which three are present in the C-terminal extension. The Ramachandran properties when all Gly, Pro and the C-terminal amino acids were removed were 76% most favoured with all amino acids falling into the generously allowed regions (Figure 4.11).

Like other SH3 domains, the Tec SH3 domain has two anti-parallel β-sheets arranged in a β-barrel with the two sheets at right angles to each other. Both sheets have three strands and one strand is shared between the two sheets. The stereo ribbon diagram of Tec SH3 domain is shown in Figure 4.12. The chemical shift data presented (Figure 4.8) predicts that residues 187-190 and 195-200 form a pair of β-strands (labelled βB and βC in Figure 4.8), however, the second strand is shared between the two sheets (designated as B and B´). Strand B´ contains a β-bulge due to hydrogen bonding of both HN Leu 208 and HN Glu 209 to the carbonyl O of Arg 217 similar to the extended β-hairpin loop present in this position in other SH3 domains (Horita *et al.*, 1998). A single turn of 3_{10} helix links strands D and E and incorporates Ser 230-Tyr 232. The first sheet is made up of β-strand A, Ile 182-Met 186, β-strand B, Gly 202-Tyr 205 and β-Strand E, Val 233-Thr 234. The second sheet is composed of β-strand B´ Tyr 205-Glu 209 (β-bulge Leu 208-Glu 209), β-strand C Leu 213-Arg 219 and β-strand D Glu 225-Pro 229 (Figure 4.13). While hydrogen bonds can only be inferred from indirect evidence from the exchange data, analysis of the

Figure 4.11

A. Ramachandran properties of the minimised SH3 domain of Tec kinase. The amino acids that do not conform to the correct Psi and Phi angles have been highlighted on the Ramachandran plot. All amino acids except Pro 229 are represented on the Ramachandran plot. The plot was generated using PROCHECK. The percentage of residues fitting the allowed regions are shown in the table (B). The flexible C-terminus (amino acids 241-245) has been removed for the $^{-2}$Gly-^{240}Asn calculations.

A



B

	$^{-2}$Gly-^{245}Asp		$^{-2}$Gly-^{240}Asn	
Residues in most favoured regions [A,B,L]	42	70.0%	42	76.4%
Residues in additional allowed regions [a,b,l,p]	12	20.0%	10	18.2%
Residues in generously allowed regions [~a,~b,~l,~p]	5	8.3%	3	5.5%
Residues in disallowed regions	1	1.7%	0	0.0%
Number of non-glycine and non-proline residues	60	100.0%	55	100.0%
Number of end-residues (excl. Gly and Pro)	1		1	
Number of glycine residues (shown as triangles)	5		5	
Number of proline residues	1		1	
Total number of residues	67		62	

Figure 4.12

Stereo ribbon diagram of Tec SH3 domain. The β-strands, depicted as arrows, are numbered from the N-terminus. The 3_{10} helix is shown by a rectangle. The colour scheme is identical to figure 4.9 such that β-sheet 1 is coloured red, β-sheet 2 is coloured cyan and the 3_{10} helix is coloured magenta. Amino acid numbers shown correspond to the mouse full-length Tec IV protein. The figure was generated in Molscript and Raster3D (Kraulis *et al* 1991; Merritt *et al.*, 1997).

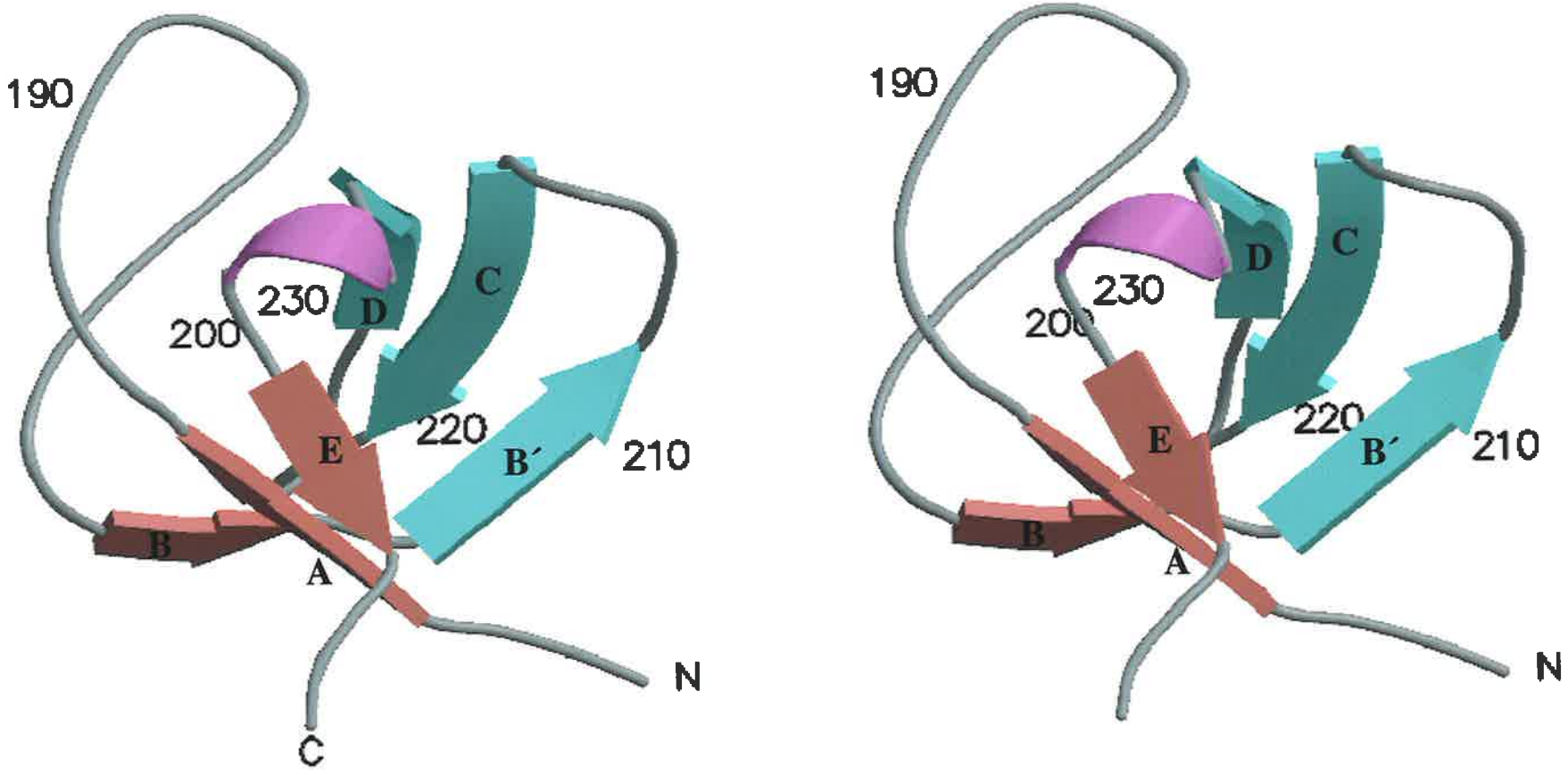
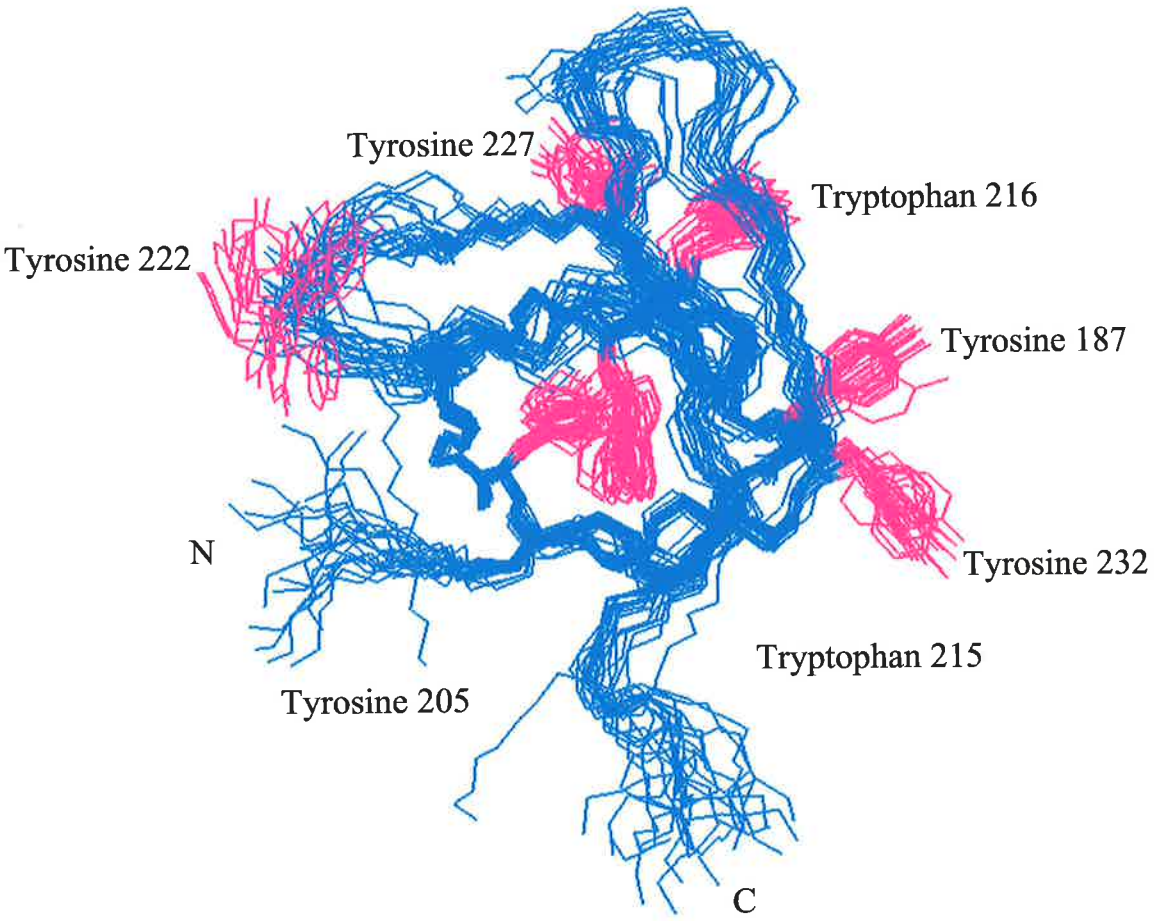

190
200
230
D
C
E
220
B´
210
B
A
N
C
190
D
C
230
E
220
B´
210
B
A
N

Figure 4.13

A trace plot of Tec SH3 domain ensemble. The ensemble shown represents the best 20 structures superimposed on a single average structure. The backbone of Tec SH3 domain is shown in blue and the sidechains of the aromatic tyrosine and tryptophan residues are shown in pink. The structure highlights well-defined core aromatic residues and the less defined surface aromatic residues. The numbering is consistent with the full-length Tec IV protein. The N- and C-termini are indicated. The figure was generated in InsightII.


Tyrosine 227
Tryptophan 216
Tyrosine 222
Tyrosine 187
N
Tyrosine 232
Tryptophan 215
Tyrosine 205
C

final structures indicates the closeness of the atoms involved are consistent with the sections of protein structure forming β-sheets. There is significant hydrogen bonding between the two sheets and between the 3_{10} helix and RT loop as implied by both hydrogen exchange experiments and proximity in the final structures. The 3_{10} helix forms hydrogen bonds with amino acids from both sheet structures, for example, there is a hydrogen bond predicted between Leu 197 and Tyr 229. The RT loop between strands A and B is 19 amino acids and shows no propensity to form any additional secondary structural elements despite the possibility of hydrogen bonds between Phe 189 and Leu 199. A hydrogen bond between Phe 189 and Leu 199 could stabilise the loop. Hydrogen bonds between Leu 197 and Tyr 227 in strand two positions the loop along the top of the barrel. This possible hydrogen bond could place the RT loop in a fixed position in the unliganded form. Similar hydrogen bond propensity has also been observed in the SH3 domain of Btk (Hansson *et al.*, 1998). In comparison, the n-Src loop between strands B´ and C spanning amino acids Asn 211 and Leu 213 does not contain hydrogen bonds to any other part of the protein and is therefore less well defined. The autophosphorylated tyrosine, Tyr 187, is conserved among a variety of SH3 domains including Tec family members and Src family members. Tyrosine 187 is found in the RT loop, and in the 3D structure, is on the surface of the protein domain.

The hydrophobic core of the SH3 domain protein is composed of Val 183, Ala 185, Ala 181, Leu 197, Leu 199, Tyr 205, Trp 216, Ala 218, Ile 228, Ser 230 and Val 233. Hydrophobic amino acids in the core of the protein are a good measure of the quality of the structure. The internal hydrophobic amino acids exhibit a high level of rigidity due to the numerous contacts made in the core of the protein. Figure 4.13 shows the final structures with the hydrophobic tyrosine and tryptophan amino acids shown. The contacts made by the different amino acids are represented in the contact map generated using MOLMOL (Figure 4.14B compared with the sequence presented in 4.14A) (Koradi, *et al* 1996). The contacts within one sheet are clear from the contact map as different amino acids within one strand have made contact with amino acids within a different strand on the same sheet. The amino acids involved in the core form contacts between many other amino acids within the protein core. For example, Trp 216 makes contacts with amino acids Val 183, Ile 207 and Leu 208 that are present in the core of the protein. Solvent accessibility has been analysed using the program MOLMOL and is plotted as accessibility versus residue number. The β-sheet regions have amino acid backbones less accessible to the solvent than those of the C-terminal region (Figure 4.14C). Thus, a decrease in solvent accessibility is consistent with the locations of the β-strands in Tec SH3 domain.

Figure 4.14

A. Primary sequence of the Tec SH3 domain representing amino acids 181-245 of the Tec IV sequence. The N-terminal residues Gly^{-2} and Ser^{-1} are derived from the fusion partner after thrombin digestion.

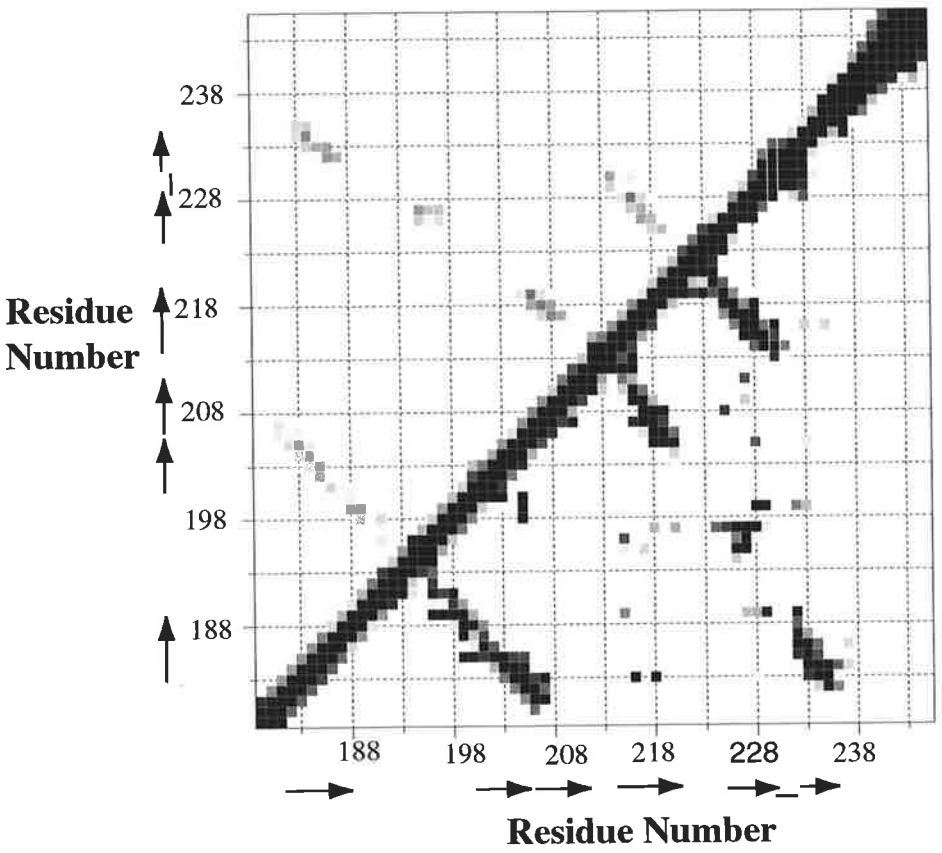
B. Contact plot of Tec SH3 domain generated in the program MOLMOL (Koradi *et al* 1996). This plot indicates the contacts between the different residues in Tec SH3 domain. The arrows represent the locations of the β-strands in Tec SH3 domain.

C. Solvent accessibility of residues within Tec SH3 domain. Solvent accessibility can be a reflection of the presence of secondary structure elements. The arrows above the graph reflect the positions of β-strands in Tec SH3 domain, which is consistent with solvent accessibility. The graph is generated in MOLMOL (Koradi, *et al* 1996). Amino acid numbers shown correspond to the mouse full-length Tec IV protein.

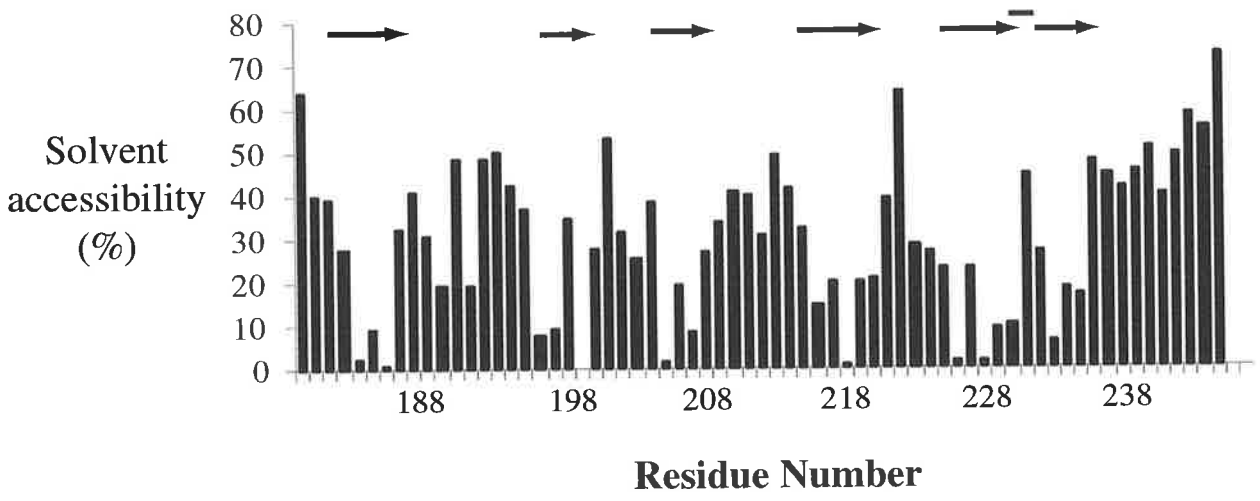
A

```
-2 -1 181                                190                                200                                210
G  S  E  I  V  V  A  M  Y  D  F  Q  A  T  E  A  H  D  L  R  L  E  R  G  Q  E  Y  I  I  L  E  K  N  D  L

H  W  W  R  A  R  D  K  Y  G  S  E  G  Y  I  P  S  N  Y  V  T  G  K  K  S  N  N  L  D  Q  Y  D
                220                                230                                240
```

B



C



4.6.2 Comparison to other SH3 domains

A comparison of the three dimensional structure of SH3 domains of Fyn, Hck, Abl and PI3K are shown in Figure 4.15. The three dimensional fold of Tec SH3 domain shows strong similarities with the SH3 domains of GAP (Yang *et al.*, 1994), Spectrin (Zhang *et al.*, 1995), Hck (Horita *et al.*, 1998) and Fyn (Morton *et al.*, 1996) all of which have long loops between β-strand 1 and β-strand 2.

Another similarity between these SH3 domains is that the second strand participates in the formation of both sheet structures. This direct linkage between β-strands (labelled B and B´ in Tec SH3 domain) includes a β-bulge that changes the direction of the β-strand enabling it to form hydrogen bonds with both sheets. These SH3 domains also contain a 3_{10} helix between β-strands 4 and 5 and the overall positioning of the strands is similar (Yang *et al.*, 1994). The SH3 domain of Lck is also similar to Tec in that it contains five β-strands in similar locations and does not contain any secondary structure in the RT loop. Lack of structure in the RT loop of Lck is due to the presence of a proline that increases flexibility in this region (Hiroaki *et al.*, 1996). The SH3 domain of Grb2 is similar to that of Fyn and spectrin as the β-strand locations are in similar positions within the SH3 domain (Goudreau *et al.*, 1994).

Tec SH3 domain does not contain any secondary structure inserts within different sections of the protein. Src SH3 domain secondary structure is different from that of Spectrin, Tec and PLC-γ (Kohda *et al.*, 1993). It contains a shorter RT loop plus additional fragments of secondary structure in the n-Src loop (Yang *et al.*, 1994). The SH3 domain of PLC-γ contains eight β-strands forming three antiparallel β-sheets. An additional β-strand in the extended RT loop aids in stabilisation of the protein (Yang *et al.*, 1994); (Kohda *et al.*, 1993). P85α SH3 domain, although composed of five strands has a large insert of 15 amino acids which forms two helices within the RT loop (Booker *et al.*, 1993); (Yang *et al.*, 1994). Amphiphysin SH3 domain is very close in structure to p85α, it contains a 13 amino acid insert between β-strands 3 and 4 which forms two helices and can interact with a region N-terminal to the SH3 domain of Amphiphysin (Owen *et al.*, 1998). It is interesting to note that the residues within the RT loop of Abl SH3 domain are less acidic when compared with other RT loops and, thus, the ligands they bind do not have a positive amino acid (Gosser *et al.*, 1995). Topological differences in the SH3 domain structure do not correlate with distinct ligand binding orientations.

Figure 4.15

Structures of a variety of SH3 domains. The β-strands are shown in cyan and the 3_{10} helix in red. The structures are all in the same orientation. The figures were generated in MOLMOL (Koradi *et al* 1996).

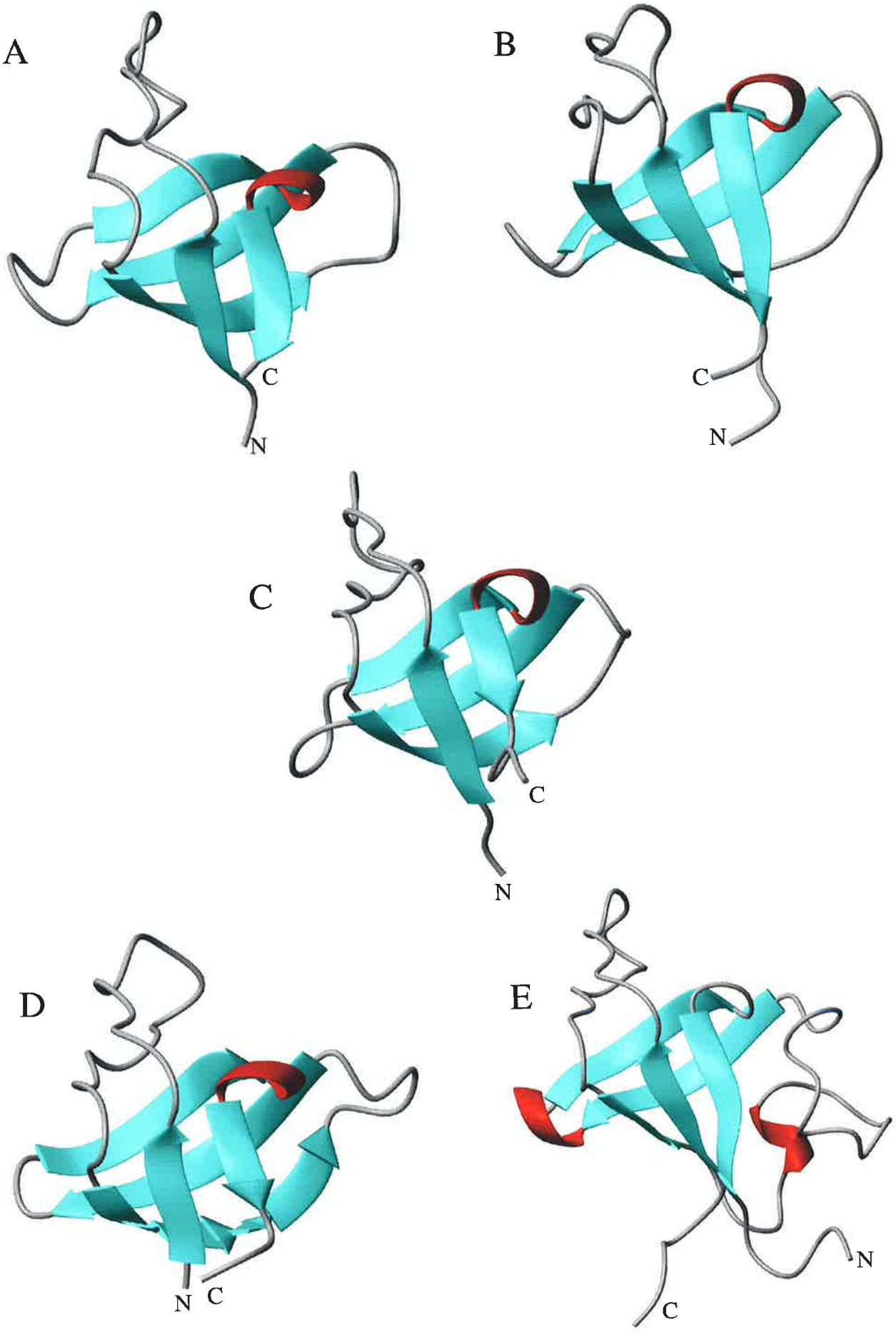
A. Fyn SH3 domain, amino acids 84-141 in the full sequence (PDB accession no. 1NYK).

B. Hck SH3 domain, amino acids 78-138 in the full sequence (PDB accession no. 5HCK).

C. Tec SH3 domain, amino acid 181-245 in the full sequence.

D. Abl SH3 domain, amino acids 65-120 in the full sequence (PDB accession no. 1ABQ).

E. PI3K SH3 domain, amino acids 2-84 in the full sequence (PDB accession no. 1PNJ).


A
C
N
B
C
N
C
C
N
D
N
C
E
N
C

4.7 TEC SH3 DOMAIN BINDS PROLINE RICH LIGANDS

The structure of Tec SH3 domain is very similar to other SH3 domains and, thus, it is predicted that the binding site would be equivalent. A proposed ligand was produced by annealing two complementary oligonucleotides together which when transcribed and translated produced the protein sequence GSGKNAPTVPGKSYDK (section 2.3.8)(Figure 4.16A). The binding orientation for Tec SH3 domain was unknown and, thus, the peptide was synthesised with both class I and class II ligand consensus sequences and was loosely based on the SH2 kinase linker in Tec and will be referred to as linker peptide. These oligonucleotides were annealed and cloned into pGEX-2T, expressed in bacteria and purified as a GST fusion protein. Cloning of this linker peptide into pGEX-2T allows both large scale production and isotope labelling of this ligand if required. The GST-linker fusion protein was expressed at high levels and found to be soluble with very little protein present in the insoluble fraction. Production of GST-linker protein was scaled up to 1 L. Cultures were grown, induced for 4 hours with a final concentration of 0.2 mM IPTG and affinity purified as per the SH3 domain. Yields achieved for the purification of GST-linker protein were in the order of 265 mg fusion protein/1L of LB medium. The fusion protein was cleaved with thrombin and following ultrafiltration was acidified and purified by high performance liquid chromatography (HPLC). Figure 4.16A shows the DNA sequence required to produce the protein along with the HPLC purification (Figure 4.16B). Mass spectrometry indicated the mass determined for this protein was consistent with the expected mass of 1606 Da (Figure 4.16C).

Nuclear magnetic resonance spectroscopy is a useful tool for the investigation of protein ligand interactions. NMR can provide detailed information regarding conformational changes both in the protein and the ligand upon complex formation (Craik and Wilce, 1997). More recently, NMR has been used for the mapping and rapid determination of protein interaction surfaces for drug design (Hajduk *et al.*, 1999). In the absence of more detailed three dimensional structure information, NMR can be used for determination of ligand binding surfaces that can help to model the binding surface (Craik and Wilce, 1997). Relative affinities of interactions can also be readily extracted from NMR data collected. Chemical shifts are valuable markers for interactions of proteins and ligands and can be used to trace changes in the environment of the nuclei that occur with complex formation (Craik and Wilce, 1997).

Figure 4.16

A. Sequence of DNA encoding the linker peptide. Complementary oligonucleotides were synthesised with a *Bam*HI overhang at the 5′ end and an *Eco*RI overhang at the 3′ end to facilitate cloning into pGEX-2T. The two strands were annealed and cloned into pGEX-2T.

B. HPLC profile of the purification of the high affinity peptide using a 50-100% buffer B (0.088% TFA in 80% acetonitrile) gradient over 20 minutes. The fusion protein was produced in bacteria and preliminary purification was conducted using a GST-affinity column. Following digestion with thrombin the sample was separated from the GST component by ultracentrifugation and acidified. The eluted peak at 18.288 minutes was collected and lyophilised twice.

C. Mass spectrometry profile of the high affinity peptide. The sample was analysed at the Waite mass spectrometry unit, Adelaide.

A

BamH1

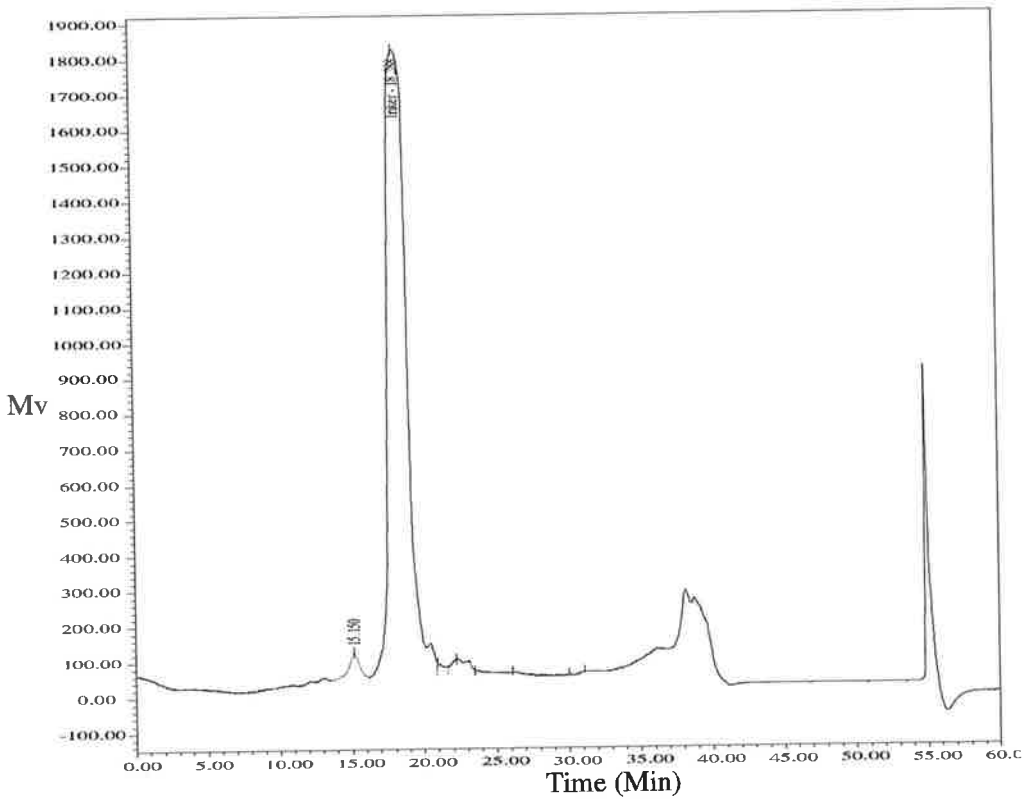
5' GA TCC GGC AAA AAC GCG CCG ACC GTG CCG GGC AAA AGC TAT GAT AAA UAA G

G CCG TTT TTG CGC GGC TGG CAC GGC CCG TTT TCG ATA CTA TTT ATT CTT AA 3'

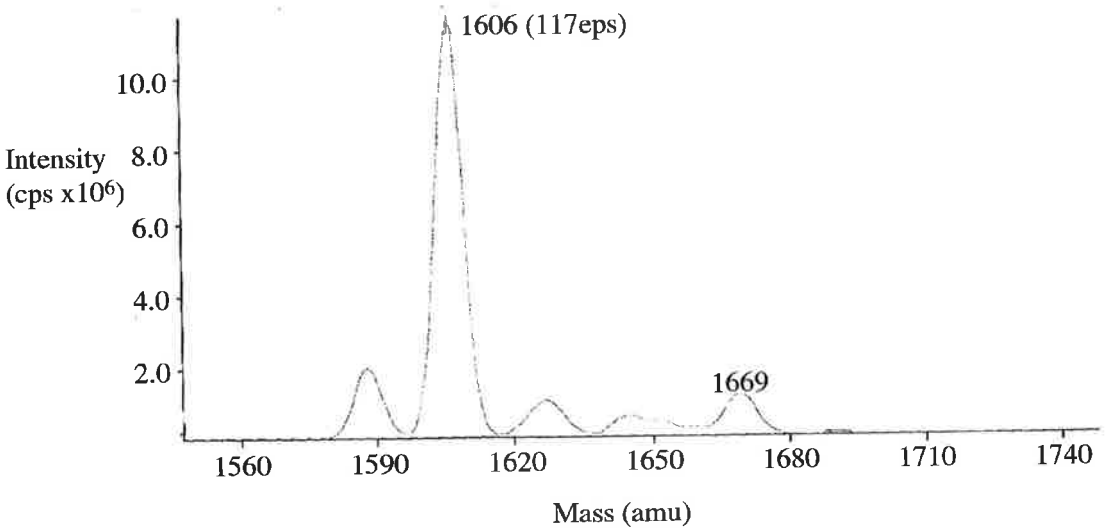
EcoR1

GSGKNAPTVPGKSYDK

B



C

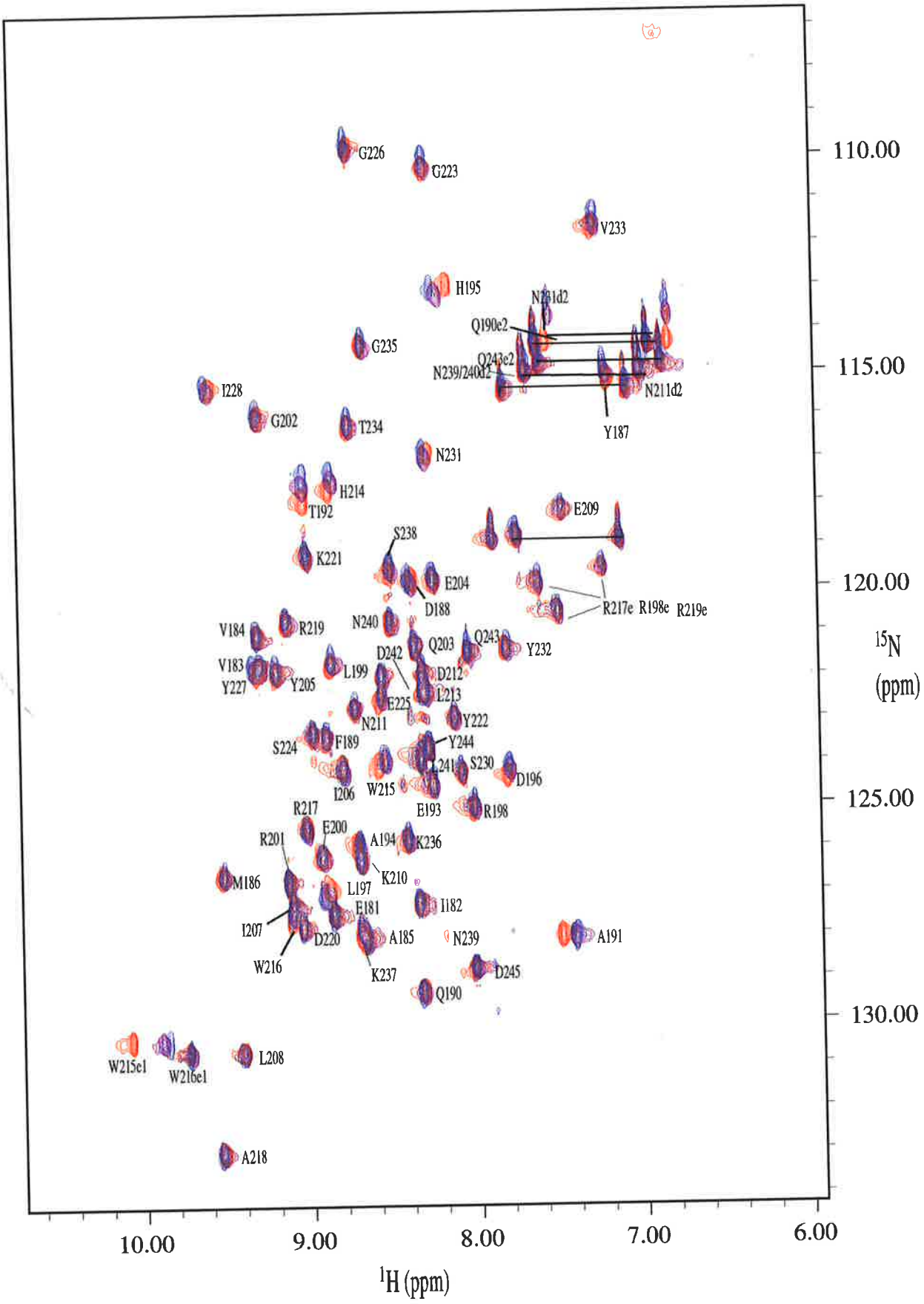


In order to map the binding site on the SH3 domain, a series of 2D ^{15}N-^{1}H HSQC NMR spectra were recorded. The HSQC experiments were recorded on a 0.75 mM ^{15}N labelled SH3 domain sample buffered with 10 mM sodium phosphate pH 6.0 and 10%(v/v) D_2O, 0.01% w/v NaN_3. The signal to noise in these NMR experiments must be optimal to allow relatively fast experimental acquisition times for multiple additions of ligand in one day. The protein sample is therefore required to be sufficiently concentrated to provide adequate signal but not too concentrated that multiple fold excess of ligand cannot be achieved. Increasing molar amounts of 41.5 mM linker peptide (dissolved in water) were added to the SH3 domain sample and successive HSQC experiments were run. The pH of the sample was measured following each addition of linker peptide and maintained at pH 6.0. The NMR experiments were conducted at 25°C with a final matrix size of 2048 and 256 complex points in the ^{1}H and ^{15}N dimensions respectively. Individual peaks in the spectrum moved incrementally during the ligand titration series indicating that the peak position represents the average of the bound and free chemical shifts (Figure 4.17). This is consistent with the fast exchange kinetics for the complex formation. Chemical exchange refers to a process in which the environment of any nucleus changes between two or more conformations and its NMR parameters differ (eg chemical shift, scalar coupling and relaxation) (Lian and Roberts, 1993). Fast exchange interactions results in a single resonance peak, which reflects the weighted average of the chemical shifts of the two individual states- that is the bound, compared with the unbound form. Slow exchange in comparison would produce two resonances for each state (Lian and Roberts, 1993). When the HSQC profiles for different ligand concentrations are overlaid, amino acids involved in ligand binding become apparent (Figure 4.17).

The amino acids involved in ligand binding were observed as chemical shift changes in both the nitrogen and the proton dimension on the HSQC spectra due to changes in their local environment as they interact with the ligand. The amino acids involved in binding include Tyr 187, Asp 188, Ala 191, Thr 192, His 195, Leu 197, His 214, Trp 215, Gly 226, Ile 228 and Asn 231 (Figure 4.19A). Pro 229 shown in blue (Figure 4.19A) is in the middle of the binding pocket, however, it lacks an amide group and therefore is not detected in a HSQC experiment. It is highly likely that the side chain of Proline 229 has a change in chemical shift due to ligand binding which could be further detected by a ^{1}H-^{1}H 2D TOCSY or ^{1}H-^{1}H 2D NOESY in D_2O. Chemical shift analysis of ligand protein binding events indicates the amino acids whose environments change with ligand binding. These amino acids need not be in the binding pocket but may map to other regions of the protein that

Figure 4.17

^{1}H-^{15}N HSQC plots of ^{15}N labelled Tec SH3 domain with increasing concentration of linker peptide added. The spectra were recorded on a Varian Inova 600 spectrometer with spectral widths of 8000Hz in the ^{1}H dimension and 2000Hz in the ^{15}N dimension. The number of increments collected was 64 and the number of points was 2048 with an acquisition time of 0.128 seconds. 16 transients were collected. The spectra were collected at 25°C using the ghsqcse experiment from Varian. Both dimensions in the spectra were processed with a sinebell function and linear prediction was used in the nitrogen dimension using the VNMR software. The final spectral size was 2048 x 256 and this was visualised and further analysed using XEASY (Bartels *et al.,* 1995). The protein sample was at 0.75mM in 10 mM NaH_2PO_4 supplemented with 10% D_2O pH 6.0. The linker peptide was dissolved in water to a final concentration of 41.5 mM and this was added incrementally to the SH3 domain protein. The pH was checked following addition of the peptide and maintained at pH 6.0. The spectra in the SH3 domain with no site 1 peptide is represented in red, when 2.75 mM peptide 3 fold excess (purple) is present and when 5.25 mM-7 fold excess (blue) peptide is present. The numbering is consistent with mouse Tec protein assigned in section 4.5.1.


G226
G223
V233
H195
N211d2
Q190e2
G235
Q243e2
N239/240d2
N211d2
I228
G202
T234
Y187
N231
H214
T192
E209
S238
K221
E204
D188
R217e R198e R219e
N240
R219
V184
Q203
Q243
Y232
V183
Y227
Y205
L199
D242
D212
L213
E225
N211
Y222
S224
F189
Y244
L241
S230
D196
I206
W215
E193
R198
R217
E200
R201
A194
K236
K210
M186
L197
E181
I182
I207
D220
A185
N239
A191
W216
K237
D245
Q190
W215e1
W216e1
L208
A218
110.00
115.00
120.00
125.00
130.00
^{15}N (ppm)
10.00
9.00
8.00
7.00
6.00
^{1}H (ppm)

undergo a conformational change with ligand binding (Foster *et al.*, 1998). This can lead to over estimation of the extent of the interface (Ramos *et al.*, 2000). Mapping the observed chemical shift changes on to the structure of Tec SH3 showed the binding pocket is a shallow hydrophobic patch on the surface of the protein (Figure 4.19B). The autophosphorylated tyrosine is located at the edge of the binding pocket and the NH group is shifted with ligand binding to Tec SH3 domain as is also observed for Btk (Morrogh *et al.*, 1999). This region in Tec SH3 domain is consistent with other SH3 domain ligand interfaces. The structure of Btk SH3 domain has been determined in a complex with a peptide from c-Cbl (Tzeng *et al.*, 2000). The amino acids in Btk showing the greatest chemical shift changes were Trp 251 and Ser 266 with Tyr 223, Tyr 268 and Tyr 225 also showing chemical shift changes. The amino acids involved in ligand binding in Tec SH3 domain show a reduced content of tyrosine residues compared with that observed in the binding site of other SH3 domains (for example Btk). These tyrosines may still interact with the ligand; however, the backbone of these amino acids does not alter significantly with ligand binding and was not detected in these HSQC experiments. Aspartic acid 188 undergoes a chemical shift change with ligand binding which is consistent with this amino acid interacting with the ligand and forming a salt bridge. Salt bridge formation is the main determinant of class I or class II binding. A further degree of complexity arises in these experiments with the observation that this peptide contains the consensus sequences required to bind in both a class I and class II orientation. These experiments do not allow differentiation as to which orientation the peptide is binding. The chemical shift changes for each amino acid in the protein can be graphed to highlight which amino acids undergo changes upon ligand binding (Figure 4.18A). The chemical shift values of the Tec SH3 domain backbone amide protons are shown in Table 4.3 for both the complex and the uncomplexed forms.

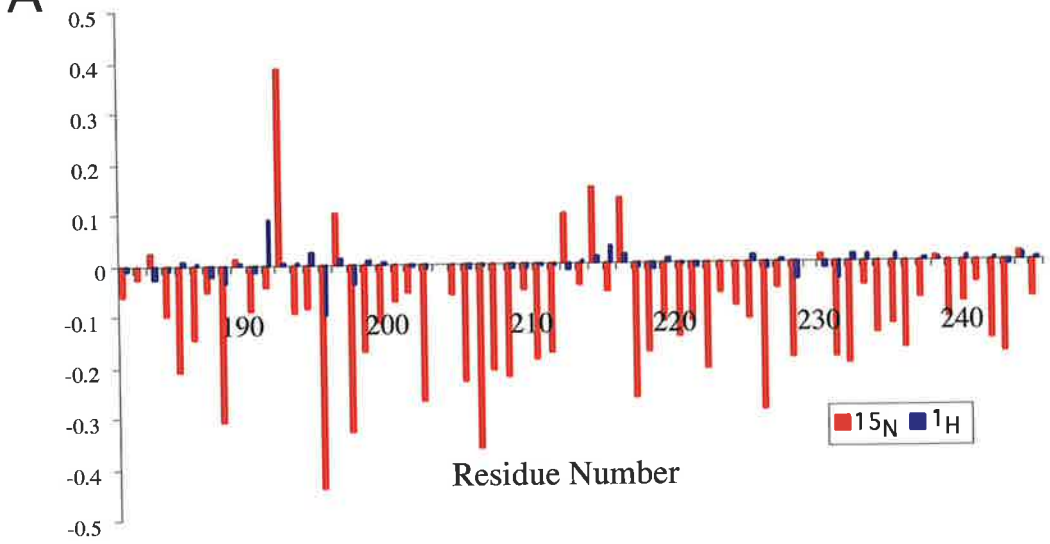
Equilibrium dissociation constants can be derived from the chemical shift data (Craik and Wilce, 1997). The ^{1}H chemical shift data is used for the calculation of the dissociation constant as the resolution is better in the proton dimension compared with the nitrogen dimension. The change in chemical shift relative to the chemical shift of the unbound SH3 domain can be plotted against concentration of ligand to determine if saturation of the SH3 domain with the ligand has been reached. The relationship between the chemical shift change and excess ligand (mM) has been plotted for the amino hydrogen for amino acids Asn 231, Trp 215ε, His 195 and Asp 188 in Figure 4.18B. These plots show that with the addition of 5.25 mM ligand this interaction has not reached saturation. Equilibrium

Figure 4.18

A. Graph of chemical shift changes of Tec SH3 domain. Blue columns represent the hydrogen frequencies and those in red represent the nitrogen frequencies. The residue numbers are indicated under the graph and reflect those of the full-length Tec IV protein.

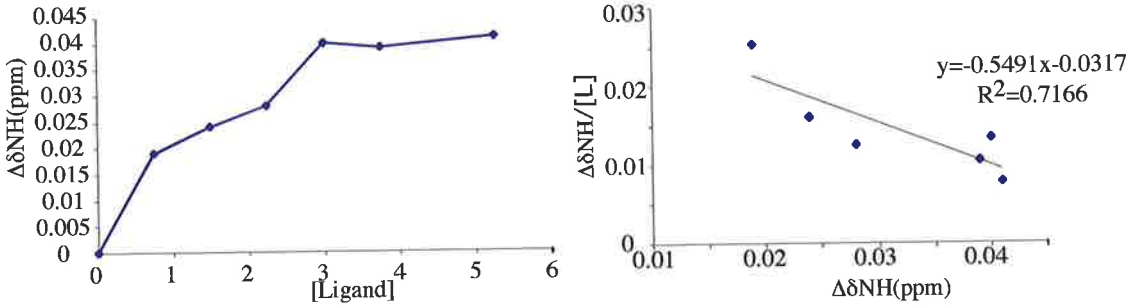
B. Titration curves highlighting the effect of increasing peptide concentrations on amino acids Tryptophan 215 NεH, Asparagine 231 NH, and Histidine 195 HN and Asp 188 NH as measured by NMR. Change in the chemical shift of the hydrogen frequency is plotted against the ligand concentration in mM. Scatchard based analysis was conducted to determine the dissociation constant. Chemical shift change relative to ligand concentration was plotted against chemical shift change where the slope of the curve is $1/K_D$. The R^2 values indicate the goodness of the fit to the slope of the curves.

A

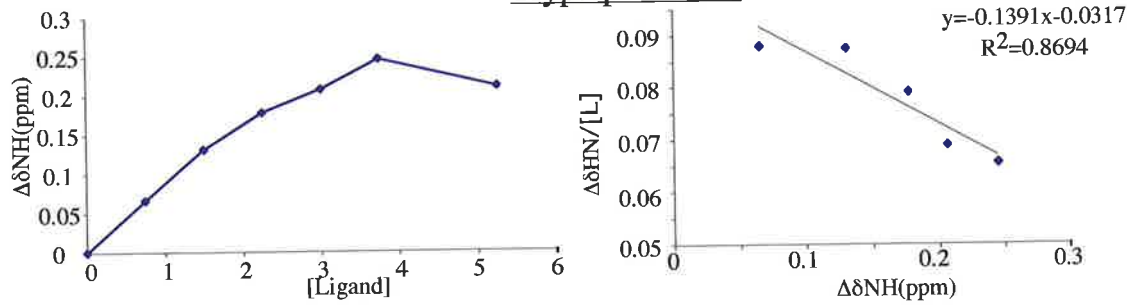


B

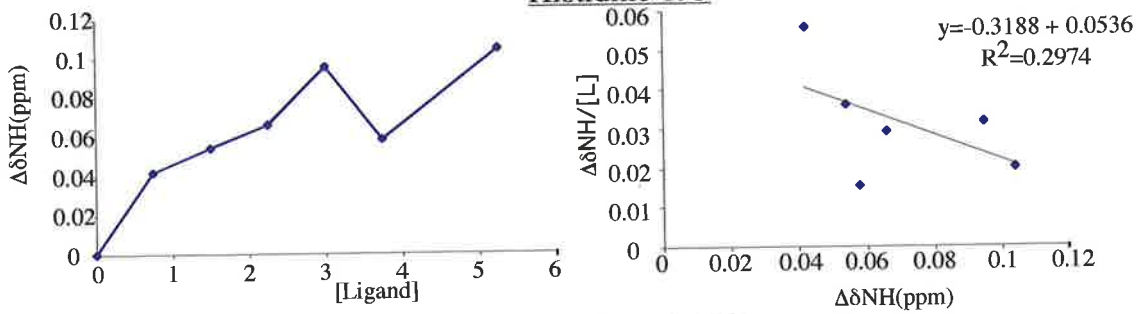
Asparagine 231



Tryptophan 215ε



Histidine 195



Aspartic Acid 188

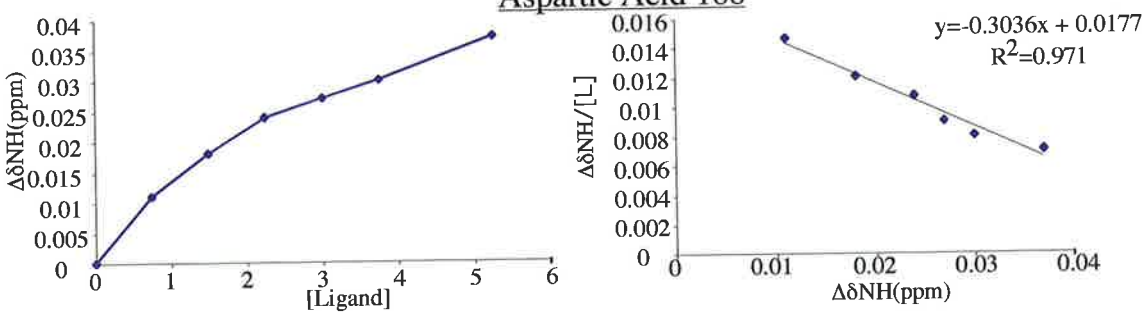
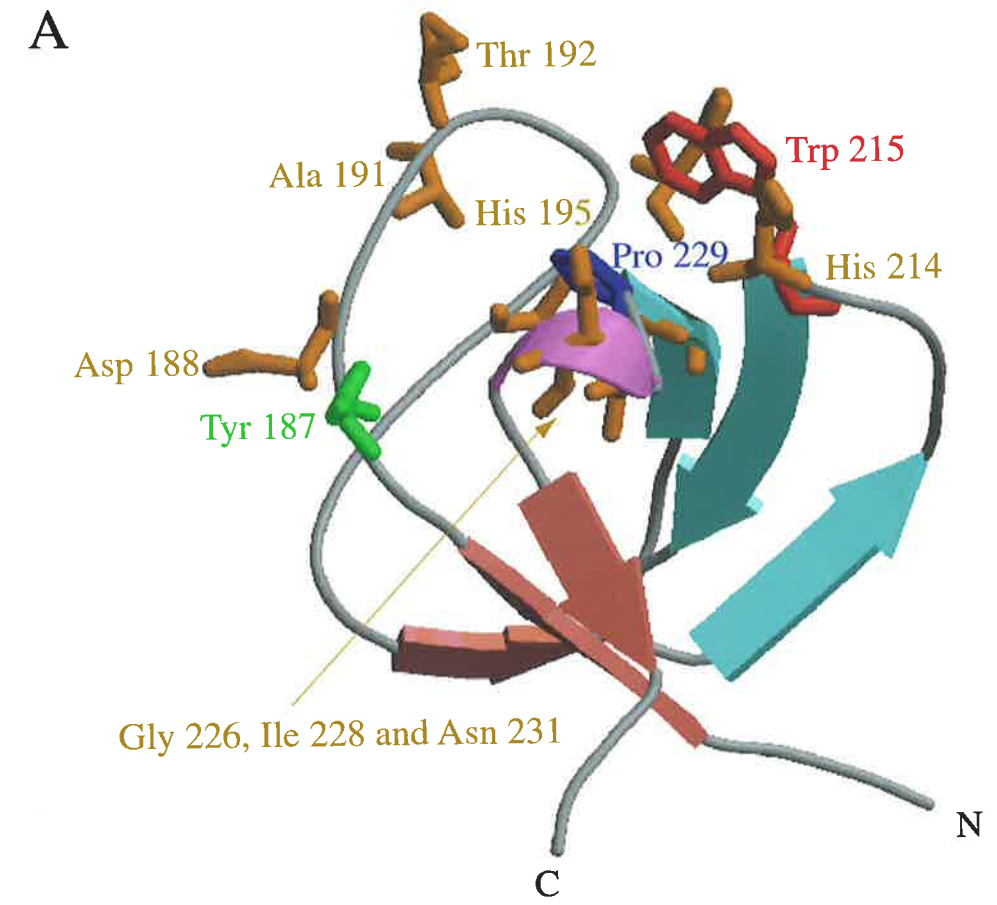


Figure 4.19

A. A ribbon diagram highlighting the amino acids involved in ligand binding. β-sheet 1 is shown in red and β-sheet 2 is shown in light blue. The 3_{10} helix is shown in magenta. Trp 215 is shown in red, Tyr 187 is shown in green and Pro 229 is shown in blue. The other amino acids involved in ligand binding are shown in brown and are named. The N- and C-termini are indicated. The numbering is consistent with the number for full-length Tec kinase. The figure was generated using Molscript and Raster3D (Kraulis *et al* 1991); (Merritt *et al.*, 1997).

B. Surface plot of Tec SH3 domain with the ligand-binding site generated in Insight II. The plot has been rotated 90° along the Z axis relative to A. The binding site (magenta) was determined using NMR spectroscopy. Proline 51 is highlighted in yellow, as although it is not detected in an HSQC experiment it is likely to be involved in any interactions involving ligand. Tyrosine 187 is shown in white.


A
Thr 192
Ala 191
His 195
Trp 215
Pro 229
His 214
Asp 188
Tyr 187
Gly 226, Ile 228 and Asn 231
N
C

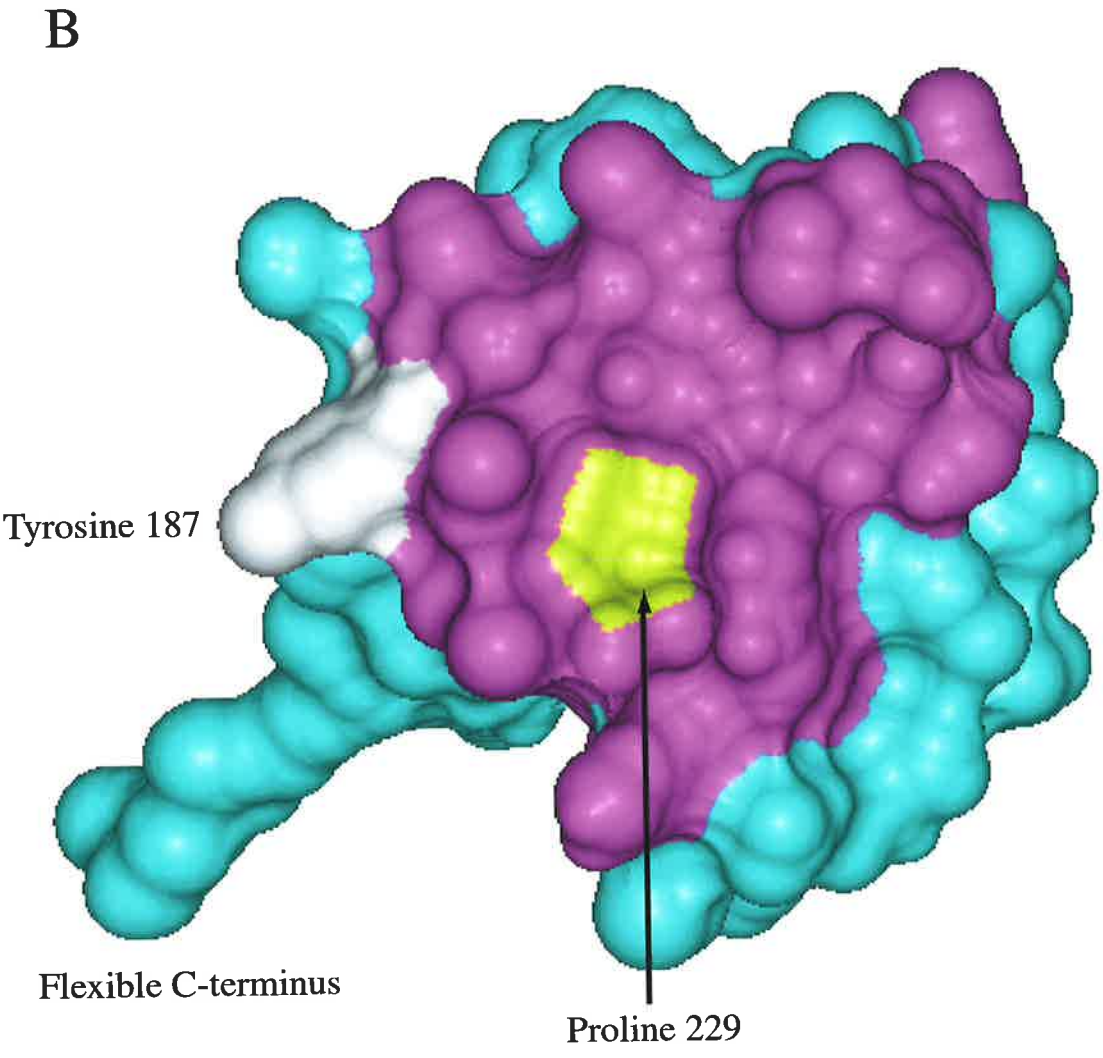

B
Tyrosine 187
Flexible C-terminus
Proline 229

Table 4.3 Chemical Shifts of Tec SH3 domain with the addition of the Linker peptide

	No peptide added		4.5mM peptide added	
Residue No.	**^{1}H Chemical Shift**	**^{15}N Chemical Shift**	**^{1}H Chemical Shift**	**^{15}N Chemical Shift**
E181	8.854	127.485	8.864	127.545
I182	8.348	127.252	8.351	127.28
V183*	9.248	121.773	9.277	121.751
V184	9.287	121.031	9.298	121.132
A185	8.672	128.098	8.667	128.31
M186	9.518	126.619	9.515	126.766
Y187	7.174	115.149	7.199	115.202
D188	8.359	119.775	8.395	120.085
F189	8.893	123.437	8.889	123.428
Q190	8.324	129.263	8.338	129.354
A191	7.496	128.01	7.409	128.054
T192	9.01	117.912	9.008	117.527
E193	8.255	124.465	8.252	124.561
A194	8.717	125.915	8.694	126.002
H195	8.125	112.965	8.225	113.405
D196	7.81	124.3	7.801	124.199
L197	8.876	126.947	8.918	127.275
R198	8.021	125.043	8.016	125.215
L199	8.854	121.767	8.853	121.886
E200	8.92	126.179	8.921	126.254
R201	9.128	126.714	9.135	126.772
G202	9.271	115.961	9.28	116.23
Q203	8.36	121.348	8.36	121.348
E204	8.242	119.775	8.244	119.837
Y205	9.18	121.895	9.189	122.127
I206	8.789	124.08	8.796	124.444
I207	9.1	127.332	9.101	127.542
L208	9.399	130.659	9.41	130.881
E209	7.734	118.747	7.743	118.802
K210	8.692	126.238	8.698	126.429
N211	8.711	122.731	8.719	122.908
D212	8.307	121.96	8.322	121.863
L213	8.32	122.345	8.317	122.388
H214	8.854	117.655	8.845	117.508
W215	8.573	124.05	8.543	124.108
W216	9.115	127.671	9.102	127.546
R217	9.008	125.472	9.018	125.738
A218	9.531	133.01	9.546	133.188
R219	9.128	120.739	9.123	120.856
D220	9.041	127.797	9.048	127.943
K221	8.984	119.197	8.995	119.313
Y222	8.124	122.956	8.124	123.167
G223	8.256	110.254	8.258	110.315
S224	8.963	123.344	8.967	123.431

...Cont'd

* the signal from these residues was not detected in all the HSQC experiments

Table 4.3 cont'd *Chemical Shifts of Tec SH3 domain with the addition of the Linker peptide*

	No peptide added		4.5mM peptide added	
Residue No.	**^{1}H Chemical Shift**	**^{15}N Chemical Shift**	**^{1}H Chemical Shift**	**^{15}N Chemical Shift**
E225	8.571	122.541	8.562	122.652
G226	8.707	109.771	8.721	110.062
Y227	9.284	121.831	9.281	121.884
I228	9.544	115.278	9.579	115.466
P229#	-	-	-	-
S230	8.073	124.243	8.089	124.234
N231	8.255	116.82	8.291	117.009
Y232	7.818	121.357	7.809	121.558
V233	7.251	111.651	7.241	111.702
T234	8.735	116.209	8.737	116.352
G235	8.652	114.337	8.643	114.463
K236	8.411	125.814	8.411	125.987
K237	8.698	127.97	8.695	128.046
S238	8.503	119.582	8.5	119.577
N239*	8.193	128.008	8.195	128.121
N240	8.506	120.773	8.5	120.855
L241	8.327	123.911	8.327	123.954
D242	8.56	122.003	8.557	122.157
Q243	8.034	121.446	8.044	121.625
Y244	8.287	123.699	8.278	123.684
D245	8.021	128.75	8.017	128.824
190ε2	7.604	114.378	7.613	114.436
190ε2	6.927	114.378	6.938	114.432
203ε	7.122	118.876	7.118	119.058
203ε	7.878	118.811	7.886	119.05
211δ2	7.07	115.406	7.067	115.463
211δ2	7.812	115.406	7.808	115.462
215ε1	10.065	130.376	9.851	130.621
216ε1	9.727	130.657	9.723	130.88
231δ2	7.539	114.314	7.524	113.668
231δ2	6.81	114.314	6.819	113.661
239/240δ2	7.682	114.957	7.677	115.084
239/240δ2	6.992	114.957	6.989	115.064
243ε2	7.591	114.764	7.587	114.942
243ε2	6.849	114.764	6.857	114.949

Proline residues are not detected by HSQC experiment
* the signal from these residues was not detected in all the HSQC experiments

constants are best determined from an interaction that has reached saturation. Scatchard analysis can also be used to determine the equilibrium dissociation constant of a reaction not yet at equilibrium (Tsukube, *et al* 1996). A plot of chemical shift change relative to the concentration of the ligand is plotted against chemical shift change as shown in Figure 4.18B where the reciprocal of the slope is the dissociation constant (Tsukube, *et al* 1996). The dissociation constant was determined using the hydrogen chemical shifts of amino acids Asn 231, His 214, Asp 188, Leu 197 and Trp 215ε1 as they exhibited the greatest chemical shift changes upon binding. The dissociation constant for binding of the linker peptide to the SH3 domain is 4 mM ± 2 mM. This interaction is relatively weak as is also clear from the ratio of the peptide to the SH3 domain required to produce chemical shift changes. As the ligand orientation preference of Tec SH3 domain was unknown the peptide used in these experiments contains both a class I and class II ligand. The orientation of binding of SH3 domain ligands to SH3 domains was originally thought to exist in equilibrium, however, that has since been shown not to be the case. Thus, the presence of both ligand orientations should have allowed the reaction with the higher affinity to prevail. Thus, one ligand orientation must prevail but these experiments did not determine whether Tec SH3 domain binds in the class I or class II orientation.

The overall binding region of the published SH3 domain structure are consistent and the main determinants of binding orientation are amino acids present in the charged loop or the RT loop.

4.8 CONCLUSIONS

SH3 domains are crucial for the control of kinase activity of the protein tyrosine kinase family of proteins. The crystal structure of the tyrosine kinases Src and Hck have shown that two relatively weak interactions are required to hold Src kinase in an inactive conformation (Xu *et al.*, 1997); (Sicheri *et al.*, 1997). These interactions are broken upon binding of intracellular ligands for the SH3 and SH2 domains, facilitating activation of the protein. The importance of the SH3 domain of Tec family members is highlighted by the fact that when the SH3 domain is incomplete or removed, Tec tyrosine kinase is constitutively active. Current research has indicated the possibility that the SH3 domain of Tec kinases participates in an intramolecular interaction also to maintain Tec kinases in an inactive conformation. Thus, the SH3 domain of Tec kinase may form a critical interaction required for maintenance of the inactive form of the protein.

Structural studies can provide valuable information to supplement cell-based experiments conducted to investigate the regulation of the Tec family of kinases. This chapter presents the solution structure of the SH3 domain of the intracellular tyrosine kinase Tec. The SH3 domain was expressed in bacteria and purified to homogeneity using a combination of affinity chromatography and size exclusion chromatography. A series of NMR experiments including homonuclear ^{1}H-^{1}H NOESY and ^{1}H-^{1}H TOCSY experiments and heteronuclear 3D ^{15}N-^{1}H NOESY and ^{15}N-^{1}H TOCSY experiments were used to generate a complete list of chemical shifts and distance restraints. An iterative approach has been used to unambiguously assign the restraints and produce a high quality three-dimensional structure.

The Tec SH3 domain is composed of two β-sheets, each of which has three β-strands. The second strand is shared between the two sheets by the formation of a β-bulge changing the direction of the β-strand. The ligand binding site on the surface of the Tec SH3 domain was also determined. The ligand surface is consistent with that observed for other SH3 domains as a hydrophobic patch on the side opposite the N-and C-termini of the protein. The autophosphorylated tyrosine is located on the edge of the binding pocket and would be expected to alter the ligand binding properties of Tec SH3 domain when it is present in the phosphorylated form compared with the unphosphorylated form. The ligand binding properties of Btk SH3 domain have been observed to change with the phosphorylation of the equivalent tyrosine (Morrogh *et al.*, 1999).

The SH3 domain is a crucial player in the regulation of tyrosine kinase activity as the SH3 domain participates in an intramolecular interaction that maintains these kinases in an inactive conformation. The Tec SH3 domain structure can now allow rapid NMR analysis of potential proline rich ligands and determination of equilibrium binding dissociation affinities. Potential intramolecular interactions in Tec kinase will be investigated (Chapter 5) and the three dimensional structure of the Tec SH3 domain provides the first stage of these NMR based investigations.

CHAPTER 5:

INTERACTIONS OF TEC PRRSH3 PROTEINS

5.1 INTRODUCTION

Transduction of signals from the cell membrane to the nucleus involves a series of protein-protein interactions. One interaction that is common to signal transduction pathways is that of PRR's interacting with a variety of modular domains.

PRR's facilitate binding to hydrophobic patches on the surface of a variety of proteins. The formation of a polyproline helix also called a left handed 3_{10} helix presents a hydrophobic patch on one side of the helix while the carbonyl groups provide good hydrogen bonding sites on the opposite side. The orientation of these amino acids is restricted due to the conformation of the adjacent prolines thereby reducing the entropy cost upon binding the partner protein. These regions in proteins, normally near the N or C-termini, form 'sticky arms' for interactions with other proteins (Kay *et al.*, 2000).

5.1.1 Protein domains binding PRR sequences

Three domain types are known to bind PRR sequences, the SH3 domain already introduced in the previous chapter, WW domains and EVHI domains. WW domains, so called because of two tryptophan residues spaced between 20 and 22 amino acids apart, form a small globular structure composed of a three stranded antiparallel β-sheet. Proline sequences PPxY or PPLP bind to WW domains and are often flanked by additional prolines. The binding pocket of the WW domain includes three hydrophobic amino acids: leucine, tyrosine, one of the conserved tryptophans and a histidine. The prolines in the ligand form Van der Waal contacts with the tryptophan on the WW domain and the tyrosine on the ligand is coordinated by the histidine (Kay *et al.*, 2000). WW domain interactions have been associated with several disease states, including muscular dystrophy, Liddles syndrome, Huntington's and Alzheimer's diseases. Liddles syndrome causes hypertension because of a 12 amino acid mutation in the β and γ subunits of the amiloride-sensitive epithelial sodium channel protein that includes the PPxY motif. Mutated sodium channel subunits no longer bind to Nedd-4 protein, leading to decrease of ubiquitin-mediated breakdown of the sodium channels resulting in a longer half-life. This increases the number of sodium channels and the corresponding sodium imbalance causes high blood pressure (Kay *et al.*, 2000).

More recently, the Enabled, VASP, Homology 1 (EVHI) domain has been shown to interact with proline ligands. This domain is present in Ena, vasodilator-stimulated phosphoprotein and WASP family of proteins that regulate actin cytoskeleton remodelling. This family of proteins bind ligands with the sequence (D/E)FPPPP through a V-shaped

binding site formed by aromatic residues (Fedorov *et al.*, 1999). Profilin, a regulator of F-actin distribution, has also been shown to bind a PRR that forms a weak polyproline type 2 helix. This interaction requires a minimum of 6-8 proline residues, with increased proline content resulting in an increase in binding affinity. Although this interaction occurs through a hydrophobic patch on the surface of the protein similar to that found in other PRR interacting domains, Profilin contains no structural homology with WW or SH3 domains (Mahoney *et al.*, 1999).

5.1.2 Allosteric regulation of Src

The association and dissociation rates of PRR-SH3 interactions can be very fast and are often found in situations requiring the rapid recruitment or interchange of several proteins (Kay *et al.*, 2000). PRR-SH3 domain interactions are thought to participate in regulating the kinase activity of Src family kinases through an intramolecular interaction. These proline rich interactions are allosteric regulators of tyrosine kinase activity; they bind at regions other than the active site and regulate protein kinase activity.

The PRR-SH3 domain interaction has been seen in the structure of Src and Hck kinases (Xu *et al.*, 1997); (Sicheri *et al.*, 1997) and probably by extension the other family members and the Tec family of tyrosine kinases. The Src family of tyrosine kinases are maintained in an inactive conformation by an interaction between the SH3 domain and the amino acid region between the SH2 domain and the kinase domain (SH2-kinase linker) and the SH2 domain and the regulatory tyrosine near the C-terminus. The SH2-kinase linker does not conform to the recognised SH3 domain ligand consensus sequence, yet, it forms a polyproline type 2 helix that can interact with the Src SH3 domain. These two relatively weak interactions sequester the SH3 domain and the SH2 domain. Thus, kinase activation is facilitated through the binding of a specific cellular ligand to either of these domains resulting in relaxation of the regulatory interactions (section 1.6.1).

5.1.3 Allosteric regulation of Tec family kinases

An intramolecular interaction between an SH3 domain and a PRR within the same protein was first described in the Tec family member Itk (Andreotti *et al.*, 1997). This interaction has implications for the regulation of the protein's tyrosine kinase activity. Multidimensional NMR was conducted on a fragment representing amino acids 154-232 (Itk PRRSH3) of the Itk protein. This work showed that the PRR (KPLPPTP), located just N

terminal to the SH3 domain, is able to fold back and bind the SH3 domain in a class I orientation (section 1.6.2). NOE contacts between the SH3 binding pocket and KPLPPTP were identified, however, there was no long/medium range NOEs observed for the loop region connecting the two domains suggesting the interaction is restricted to the PRR and is specific for the PRR and the SH3 domain. Mutation of the critical proline residues to alanine (KPLAPTA) resulted in no intramolecular interaction, confirming the requirement of the proline residues in this interaction.

Biological studies investigating the accessibility of Itk SH3 domain in an *in vitro* context show that the high affinity Itk SH3 domain ligand, Sam 68, will inhibit the intramolecular interaction between the PRR and the SH3 domain (Andreotti, *et al* 1997). However, the PRR-SH3 interaction is stabilised by the presence of the Itk SH2 domain and TH domain and to a lessor degree by the presence of the TH domain alone (Andreotti *et al.*, 1997). Thus, the SH2 domain may stabilise the interaction of the SH3 domain and the PRR through an indirect interaction with the amino terminal region. The same is true for Itk PRR binding proteins. Grb2 has been shown to interact with the PRR of Itk; however, this interaction is not sufficiently strong to interfere with the intramolecular interaction between the Itk PRR and the SH3 domain. Thus, although the intramolecular interaction of the SH3 and the PRR can not compete with high affinity ligands for the SH3 domain of Itk, in the context of the whole protein the intramolecular interaction is more stable and may successfully compete with low to intermediate affinity ligands for the Itk SH3 domain and the PRR. In cell lysates activated by pervanadate, the amount of Grb2 associated with the PRR increases with increased tyrosine phosphorylation of Itk suggesting that cellular ligands can bind Itk with increasing success as the protein becomes phosphorylated (Andreotti *et al.*, 1997).

Intramolecular PRR-SH3 interactions have also been demonstrated for another Tec family member, Btk. Fluorescence experiments showed that the SH3 domain of Btk could interact with the PRR in an intermolecular interaction. Btk has two potential PRR sites and one, ^{186}KPLPPTP^{192}, has been showed to bind to the SH3 domains of Lyn, Fyn and Hck (Cheng *et al.*, 1994). A peptide representing this site was used by the authors to show an ability to bind intermolecularly to the Btk SH3 domain with an apparent affinity of 55 μM, an affinity comparable to those seen for the SH3 domains of Lyn, Fyn and Hck binding to the Btk PRR. Proline 189 of Btk PRR was shown to be crucial for this interaction and the affinity increased when Lys 186 was substituted for an arginine. A comparison between class I and class II peptides highlighted that the SH3 domain of Btk will bind to class I

ligands with a 10 fold higher affinity than class II ligands. A cellular ligand of the Btk SH3 domain is a PRR within c-Cbl that binds to Btk SH3 domain with a dissociation constant of 35 μM. Thus, the intermolecular interaction between Btk SH3 domain and the PRR is of a similar order to a known intracellular ligand suggesting that at any one time Btk SH3 domain could be bound to either c-Cbl or through an intermolecular PRR interaction (Patel *et al.*, 1997).

A dimerisation reaction between Btk SH3 domain and the PRR has been recently reported (Hansson *et al.*, 2001). Hansson and co-workers used gel permeation chromatography profiles, titrations with proline rich ligands, intermolecular NMR cross relaxation measurements and ^{15}N NMR relaxation measurements to show a monomer-dimer interaction in the order of 60 μM. The authors suggest this interaction provides the opportunity for autophosphorylation of Btk to occur through a *trans* rather than *cis* mechanism. Alternatively, autophosphorylation of the tyrosine in the SH3 domain may destabilise a preformed dimer. Both cases have implications for the regulation of Btk activation (Hansson *et al.*, 2001).

Tec kinase contains four SH3 binding consensus sequences within the PRR, complicating possible interactions. Three class I SH3 binding ligands (+XXPXXP) ^{155}KTLPPAP, ^{165}KRRPPPP, ^{167}RPPPPIP and a fourth incomplete ligand (PXXP) ^{170}PPIP exist in an arrangement where three of the four ligands overlap (Merkel *et al.*, 1999). There is currently no information regarding the affinity of these sites for the Tec SH3 domain.

The possibility of both intra- and intermolecular interactions in the Tec family members Itk and Btk has been seen *in vitro*, however, the situation may be more complicated in Tec kinase. The relative contributions of the PRR sites to any dimerisation or intramolecular interaction is not known for Tec kinase and together they may form an intricate mechanism for interactions both within Tec and with other cellular ligands. The biological significance of such an interaction for the Tec family of proteins is not known, however, the PRR of these proteins may regulate the protein kinase activity in an allosteric manner by an interaction with the SH3 domain. The protein is maintained in an inactive conformation until active competition occurs between either the PRR and a cellular substrate containing a high affinity ligand or an interaction with the SH3 domain and a high affinity ligand (Andreotti *et al.*, 1997). This could provide a very tightly regulated and highly specific mechanism of activation for Tec kinase.

5.2 AIMS

The regulation of tyrosine kinase activity is crucial to the biological signalling pathways these proteins control. To understand the biological function of Tec kinas, a reductionist approach has been taken. PRR-SH3 domain interactions have been shown to allosterically regulate tyrosine kinase activity and the potential for similar regulation has been demonstrated for Tec family members Itk and Btk. To determine if a similar interaction can occur in Tec kinase, a protein incorporating the PRR and the SH3 domain (PRRSH3) of Tec kinase was produced. This protein was analysed by biophysical experiments including surface plasmon resonance (BIAcore), analytical ultracentrifugation and NMR spectroscopy. The contribution of the different PRR sites within Tec kinase to form intra- or intermolecular interactions was determined. The PRR was broadly split into two main sites, site 1 and site 2, in which site 1 incorporated ^{155}KTLPPAP and site 2 ^{165}KRRPPPP, ^{167}RPPPPIP and ^{170}PPIP. Analysis of the contributions of each of these sites to any interactions observed was assessed using a site directed mutagenesis approach. The central proline required for polyproline type 2 helix formation was mutated to an alanine in all the ligands and the resulting proteins generated as GST-fusion proteins. Biophysical characterisation of the resulting mutant proteins was undertaken. Joel Mackay kindly carried out the ultracentrifugation experiments at the University of Sydney.

5.3 TEC PRRSH3 PROTEIN FORMS DIMERS

The DNA encoding the Tec PRRSH3 region was amplified by the polymerase chain reaction (PCR) from the complete Tec IV cDNA (James Ihle, St Judes) using primers PRR5′ and SH3JD3′as shown in Figure 5.1A (section 2.3.14). PCR was carried out with Taq polymerase for 30 cycles at an annealing temperature of 65°C (section 2.3.14) and electrophoresed on a 2% TAE agarose gel. Amplified products were observed in reactions containing 2 mM and 2.5 mM $MgCl_2$ at the expected size of 280 bp. Products from PCR reactions conducted in 2.0 mM $MgCl_2$ (Figure 5.1B) were agarose gel purified and ligated into pGemT. Figure 5.1C shows the plasmid map of pGemT and the location of the restriction enzymes used for diagnostic analysis. DNA minipreparations from resulting colonies were screened by digestion with *Pst*I and *Bam*HI endonucleases and a plasmid that contains a 317 bp fragment (Figure 5.1C), referred to as pGemPRRSH3, was grown in a 1.5 ml culture for plasmid preparation (section 2.3.3). The agarose gel shown in Figure 5.1C shows the restriction digestion of 11 putative positive colonies. Seven of the 11 had the correct insert although the band of 317 bp is not apparent on this figure. The remaining four colonies possibly contained a cloning artefact that was not further characterised. The PRR-SH3 domain coding region was excised from pGemPRRSH3 as a 280 bp DNA fragment using *Bam*HI/*Xho*I sites derived from the PCR primers. This fragment was then ligated into appropriately prepared pGEX-4T-2 and clones selected by *Bam*HI/*Pst*I and *Bam*HI/*Xho*I double restriction digestions (section 2.3.3). The pGEX-4T-2 vector contains a GST fusion partner for ease of purification and cloning at the *Bam*HI site ensures in frame production of the protein of interest. The plasmid map of pGEX-4T-2 is shown in Figure 5.2A and the position of the multiple cloning site is highlighted. As shown in Figure 5.2B *Bam*HI/*Pst*I and *Bam*HI/*Xho*I digestion results in DNA fragments of 1250 bp and 280 bp consistent with the sizes expected from the map shown in Figure 5.2B. Sequencing using pGEX 5′ and pGEX 3′ primers confirmed that Tec PRRSH3 domain coding sequence was intact and in the correct reading frame with respect to GST and the translated sequence is shown in Figure 5.2C (section 2.3.13). pGEX-4T-2-PRRSH3 plasmid DNA was transformed into BL-21λDE3 *E. coli* to facilitate protein expression.

E. coli containing pGEX-4T-2-PRRSH3 was induced to express the GST-PRRSH3 fusion protein using conditions similar to that of GST-SH3 domain protein production. The affinity purification and corresponding SDS-PAGE are shown in Figure 5.3A. The yields obtained from this system were 150 mg of fusion protein per litre of LB medium and 120

Figure 5.1

A. cDNA sequence of the PRRSH3 region from Tec kinase. The linker between the TH domain and the PRR is coloured black, the PRR is coloured blue and the SH3 domain is coloured red. The numbering is consistent with the mouse Tec IV cDNA. The primers designed to amplify Tec PRRSH3 region are shown, as is the region of complementarity where the primers will bind. The direction of amplification of each primer is shown by the arrows.

B. PCR experiment of Tec IV cDNA visualised on a 2% TAE agarose gel. 20 ng of Tec IV cDNA was amplified using the primers described in A. The PCR was conducted with 30 cycles of amplification. 5 µl of PCR product was added to 1 µl of glycerol load buffer and electrophoresed at 80 V for 40 minutes. A band of 280 bp (lane labelled PRRSH3) was observed when compared to the *Hpa*II digested pUC19 DNA markers.

C. (I) The plasmid map of pGEM-PRRSH3 is shown. Indicated on the map is the location were the PRRSH3 was cloned and the diagnostic restriction enzyme sites are indicated. pGEM-T contains an ampicillin resistance gene and a lac Z gene.

(II) The PCR product shown in B was cloned into pGEM-T DNA. The ligations were plated onto plates with IPTG and BCIG and 11 of the resulting white colonies were picked and cultured overnight. DNA preparation were conducted and 4 µl of each was digested with *Bam*HI and *Pst*I and run on a 1% TAE agarose gel. Lanes 1-11 are the *Bam*HI/*Pst*I digests of the DNA. The first lane is *Hpa*II digested pUC19 DNA markers (M).

A

460 AACAATAATA TCATGATTAA ATACCATCCT AAATTCTGGG CAGATGGGAG

510 TTACCAGTGT TGTAGACAAA CAGAAAAACT AGCACCCGGA TGTGAGAAGT

*Bam*HI
ATGGATCCAGT ATAAGAAAGA CC ➤

560 ACAATCTTTT TGAGAGTAGT ATAAGAAAGA CCCTGCCTCC CGCGCCAGAA

610 ATAAAGAAGA GAAGGCCTCC TCCACCAATT CCCCCAGAGG AAGAAAATAC

660 TGAAGAAATC GTTGTAGCGA TGTATGACTT CCAAGCGACG GAAGCACATG

710 ACCTCAGGTT AGAGAGAGGC CAAGAGTATA TCATCCTGGA AAAGAATGAC

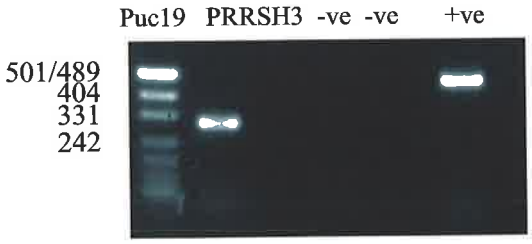
760 CTCCATTGGT GGAGAGCGAG AGATAAGTAT GGGAGTGAAG GATATATCCC

810 AAGTAATTAC GTCACAGGGA AGAAATCCAA CAACTTAGAT CAATATGAGT

Xho I
➤ **ATCTA GTTATACTCACTGAGCTCGCG**
Stop

860 GGTACTGCAG AAATACCAAC AGAAGCAAAG CAGAACAGCT CCTCAGAACG

B



C

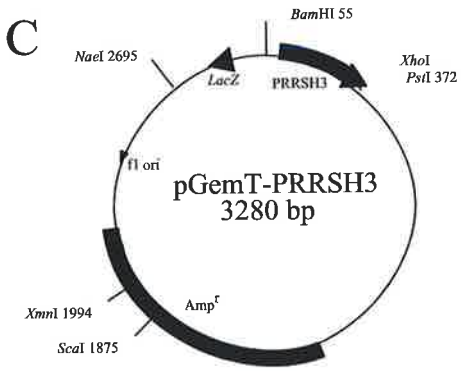


Figure 5.2

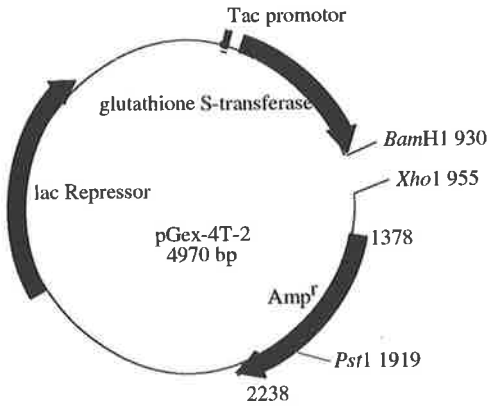
A. Plasmid map of pGEX-4T-2 used for cloning Tec PRRSH3. pGEX-4T-2 contains an ampicillin resistance cassette, a lac represser cassette and the Tac promoter. pGEX-4T-2 plasmid was digested with *Bam*HI and *Xho*I to facilitate cloning of the PCR product generated. PRRSH3 contains *Bam*HI/*Xho*I sites suitable for directional cloning.

B. (I) Plasmid map of pGEX-4T-2-PRRSH3.

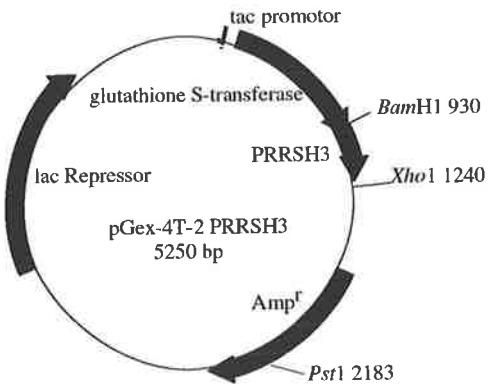
(II) The insert in pGEM-T was digested with *Bam*HI and *Xho*I and ligated into pGEX-4T-2 also digested with *Bam*HI and *Xho*I. Correct colonies were screened by digestion and visualised on a 1% TAE agarose gel. DNA with correct inserts produced a 280 bp fragments with *Bam*HI/*Xho*I digestion and a 1250 bp fragment upon *Bam*HI/*Pst*I digestion.

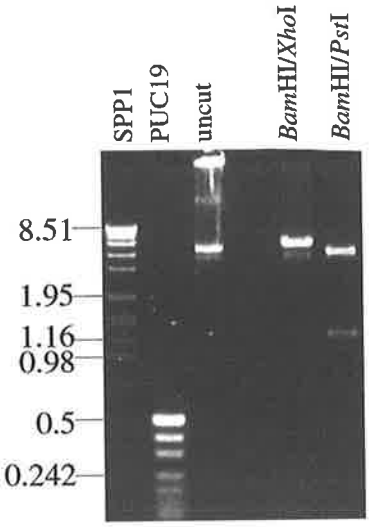
C. Translation of the PRRSH3 protein encoded by pGEX-4T-2-PRRSH3 plasmid. The DNA numbers reflect that of mouse Tec IV and the numbers in the protein start from 1 at the beginning of the cloned region. A Gly Ser dipeptide will be added to the N-terminus because of restriction enzyme sites used during cloning.

A



B





C

```
  1  G  S  S  I  R  K  T  L  A  P  A  P  E  I  K  K  R  R  P  P      20
930  GGATCCAGTATAAGAAAGACCCTGGCGCCCGCGCCAGAAATAAAGAAGAGAAGGCCTCCT   989

 21  P  P  I  P  P  E  E  E  N  T  E  E  I  V  V  A  M  Y  D  F      40
990  CCACCAATTCCCCCAGAGGAAGAAAATACTGAAGAAATCGTTGTAGCGATGTATGACTTC   1049

 41  Q  A  T  E  A  H  D  L  R  L  E  R  G  Q  E  Y  I  I  L  E      60
1050  CAAGCGACGGAAGCACATGACCTCAGGTTAGAGAGAGGCCAAGAGTATATCATCCTGGAA   1109

 61  K  N  D  L  H  W  W  R  A  R  D  K  Y  G  S  E  G  Y  I  P      80
1110  AAGAATGACCTCCATTGGTGGAGAGCGAGAGATAAGTATGGGAGTGAAGGATATATCCCA   1169

 81  S  N  Y  V  T  G  K  K  S  N  N  L  D  Q  Y  E  *  L  E  G     100
1170  AGTAATTACGTCACAGGGAAGAAATCCAACAACTTAGATCAATATGAGTGACTCGAGGGA   1229
```

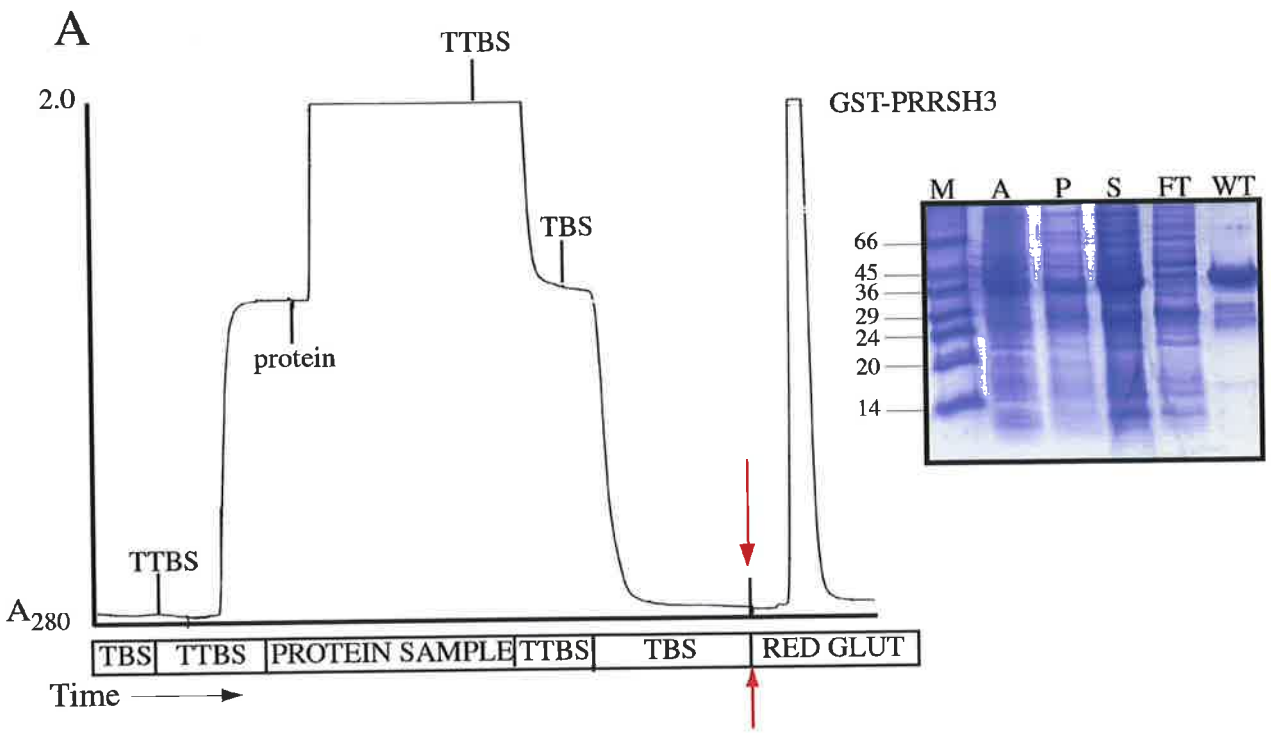
Figure 5.3

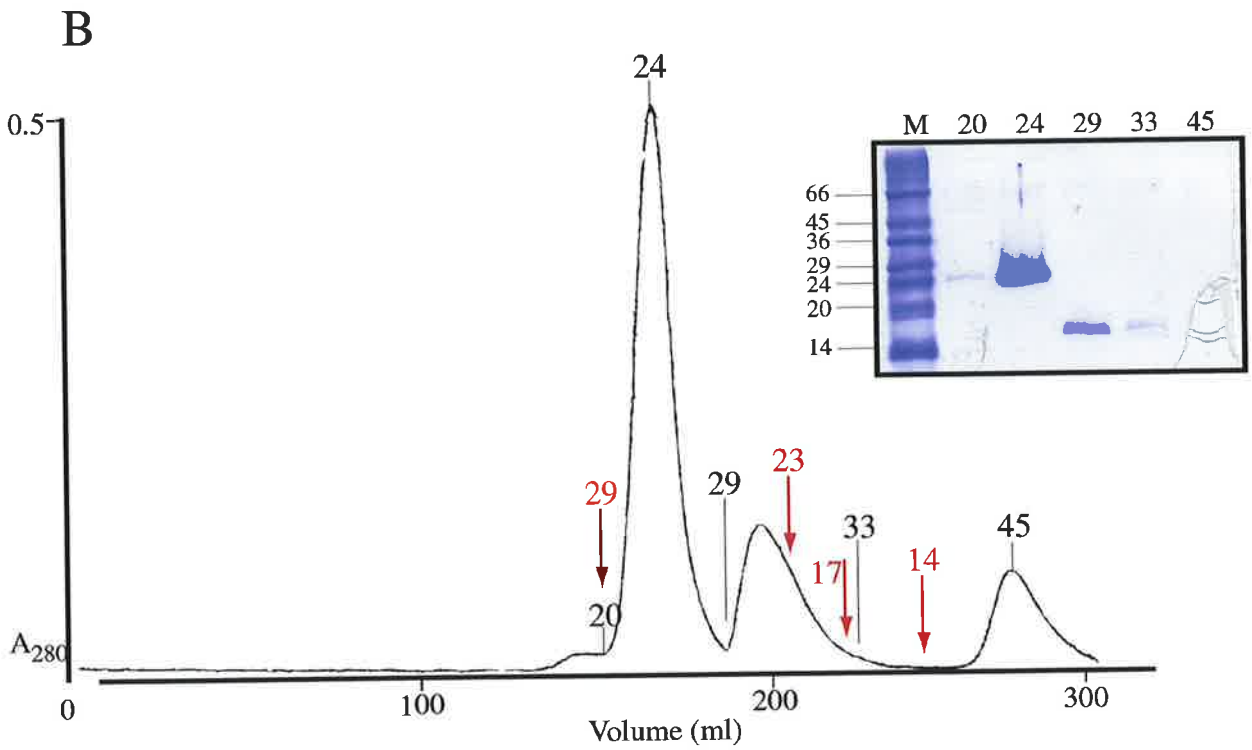
A. Profile of the affinity purification of 1.5 L of GST-PRRSH3 produced in LB medium. The bacteria were grown to an OD_{600}= 0.6 and induced with a final concentration of IPTG of 0.2 mM. French pressing was conducted to lyse the bacteria in the presence of 1 mM PMSF to minimise proteolytic digestion. The bacterial supernatant was filtered and applied to a pre-equilibrated glutathione agarose column at 4 ml/minute. Equilibration of the column was conducted by washing with TBS followed by washing with TTBS. Elution was facilitated with 10 mM reduced glutathione in TBS pH 8.0. The arrow indicates addition of reduced glutathione to the column.

12.5% SDS-PAGE conducted on samples from GST-PRRSH3 affinity purification. Bacterial cultures were grown, 1 ml was pelleted and resuspended in 200 µl 2x SDS-PAGE load buffer of which 10 µl was loaded on the gel. The induced band is present in lane A. Following lysis the two fractions, insoluble fraction (labelled p) and soluble fraction (s) were visualised by SDS-PAGE. The flow through (FT) from application to the column was also electrophoresed as was the eluted protein (WT).

B. Profile of the size exclusion purification of GST-PRRSH3. 280 mg of GST-PRRSH3 protein eluted from the affinity column was digested with 59 U of thrombin and separated by Superdex G75 size exclusion chromatography (absorbance full scale 0.5). The column was equilibrated in 1 x PBS, and 0.02% NaN_3. 5 ml fractions from the different peaks were collected and pooled. The red arrows indicate where calibration proteins eluted. Sizes are indicated above the elution profile.

12.5% SDS-PAGE conducted on fractions eluted from the size exclusion column. SDS-7 markers were loaded in lane 1 as a molecular weight comparison. 10 µl of each fraction was added to 10 µl of 5 x SDS-PAGE gel loading buffer and heated to 100°C for 3 minutes before loading 10 µl on the gel.


A
TTBS
2.0
GST-PRRSH3
TBS
protein
M A P S FT WT
66
45
36
29
24
20
14
TTBS
A280
TBS TTBS PROTEIN SAMPLE TTBS TBS RED GLUT
Time


B
24
0.5
M 20 24 29 33 45
66
45
36
29
24
20
14
29
29
23
33
17
14
45
20
A280
0
100
200
300
Volume (ml)

mg fusion protein per litre of Min A minimal medium (Miller 1972) (Figure 5.3A). The PRRSH3 protein was cleaved from the GST fusion protein by thrombin digestion at 37°C for a period of 4 hours. Complete thrombin digestion was achieved in this time. The profile obtained from this size exclusion purification step (Figure 5.3B) shows three peaks, the bulk of the sample was present in the GST (fraction 24) peak eluted at 170 ml. The PRRSH3 protein eluted at 200 ml (fractions 29-33, lane 29 in the gel). An additional peak eluted at 290 ml (fraction 45) may represent reduced glutathione or a buffer salt peak as no protein band was observed on a protein SDS-PAGE gel. No further analysis of the peak was conducted. The PRRSH3 protein eluted from the size exclusion column was subjected to mass spectrometric analysis that showed a molecular weight 11214 Da, consistent with the expected mass of the PRRSH3 wild-type protein. There are several higher molecular mass peaks, which may reflect some degree of post translational modification (Figure 5.9A). Although size exclusion chromatography was used to separate the GST fusion partner from the PRRSH3 protein, the separation was suboptimal as there was no baseline separation between these proteins. The PRRSH3 protein eluted from the column earlier than predicted based on the mass calibration of the column (Figure 5.3B). This was the first indication that this protein may be forming multimeric complexes. To further understand this interaction, and determine the degree of multimerisation a series of biophysical experiments were undertaken.

One dimensional NMR spectra recorded on the purified PRRSH3 protein appeared to contain several sharp peaks but also broad peaks. Upon comparison with similar spectra for the SH3 domain samples (Chapter 4), it is clear that the signal to noise is lower and the peaks are significantly broadened, consistent with a protein of greater than 12 kDa. This data suggests that the PRRSH3 forms multimeric species in solution. This was confirmed by comparison of 2-D ^{1}H-^{1}H NOESY and ^{1}H-^{1}H TOCSY spectra of Tec PRRSH3 and Tec SH3 domain. The number of NOEs observed in the NOESY spectrum was lower than seen for the smaller SH3 domain indicating the formation of aggregated protein in the NMR sample (Figure 5.4).

Since aggregation is a concentration dependent phenomenon, a dilution series was undertaken in order to test for conditions that minimise the aggregation but maximise the signal to noise. Figure 5.5 shows a series of HSQC spectra of ^{15}N labelled PRRSH3 protein ranging from 0.0625 mM to 0.5 mM concentrations. On average, 60 resonances were observed in these spectra for an 11 kDa protein compared with the 65 resonances observed for the 8 kDa SH3 domain (Figure 4.7). The intensity of the peaks did not decrease with the

Figure 5.4

A. Homonuclear ^{1}H-^{1}H TOCSY spectrum conducted on Tec PRRSH3 protein. Tec PRRSH3 was made up to a concentration of 0.35 mM in a buffer containing 10 mM NaH_2PO_4, 150 mM NaCl pH 6.0 supplemented with 10% (v/v) D_2O. ^{1}H-^{1}H TOCSY data were recorded on a Varian Inova 600 spectrometer with spectral widths of 8000 Hz in both dimensions. The spectra were processed in XEASY (Bartels, *et al* 1997).

B. Homonuclear ^{1}H-^{1}H NOESY spectrum conducted on Tec SH3 protein. Tec SH3 protein was made up to a concentration of 2.75 mM in a buffer containing 10 mM NaH_2PO_4, pH 6.0 supplemented with 10% (v/v) D_2O. NOESY data were recorded at BRI, Melbourne with spectral widths of 8000 Hz in both dimensions. The spectra were processed in XEASY (Bartels, *et al* 1997).

C. Homonuclear ^{1}H-^{1}H NOESY spectrum conducted on Tec PRRSH3 protein. Tec PRRSH3 protein was made up to a concentration of 0.35 mM in a buffer containing 10 mM NaH_2PO_4, 150 mM NaCl pH 6.0 supplemented with 10% (v/v) D_2O. 1H-1H NOESY data were recorded on a Varian Inova 600 spectrometer with spectral widths of 8000 Hz in both dimensions. The spectra were processed in XEASY (Bartels, *et al* 1997).

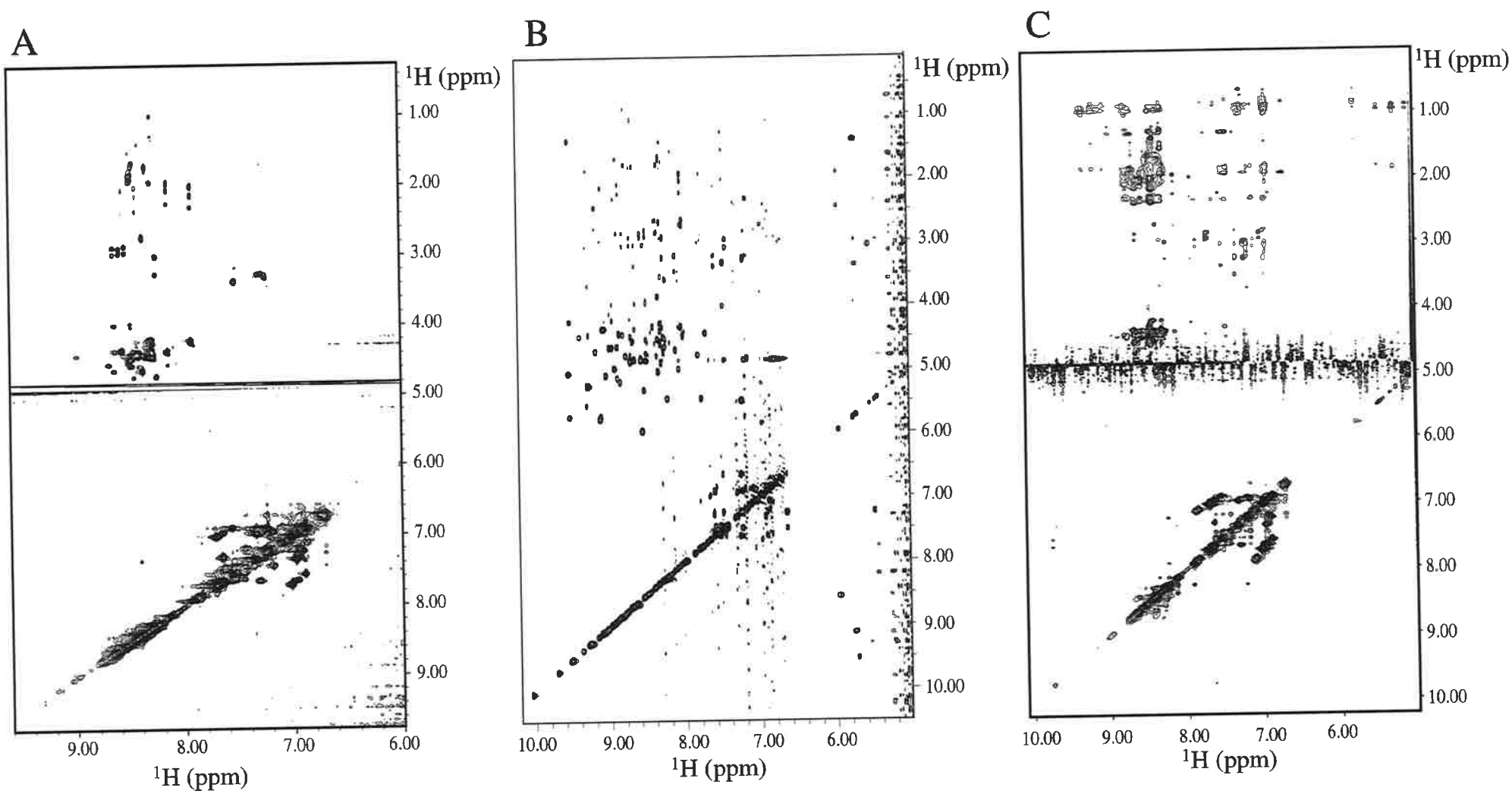

A
B
C
^{1}H (ppm)
1.00
2.00
3.00
4.00
5.00
6.00
7.00
8.00
9.00
10.00
^{1}H (ppm)

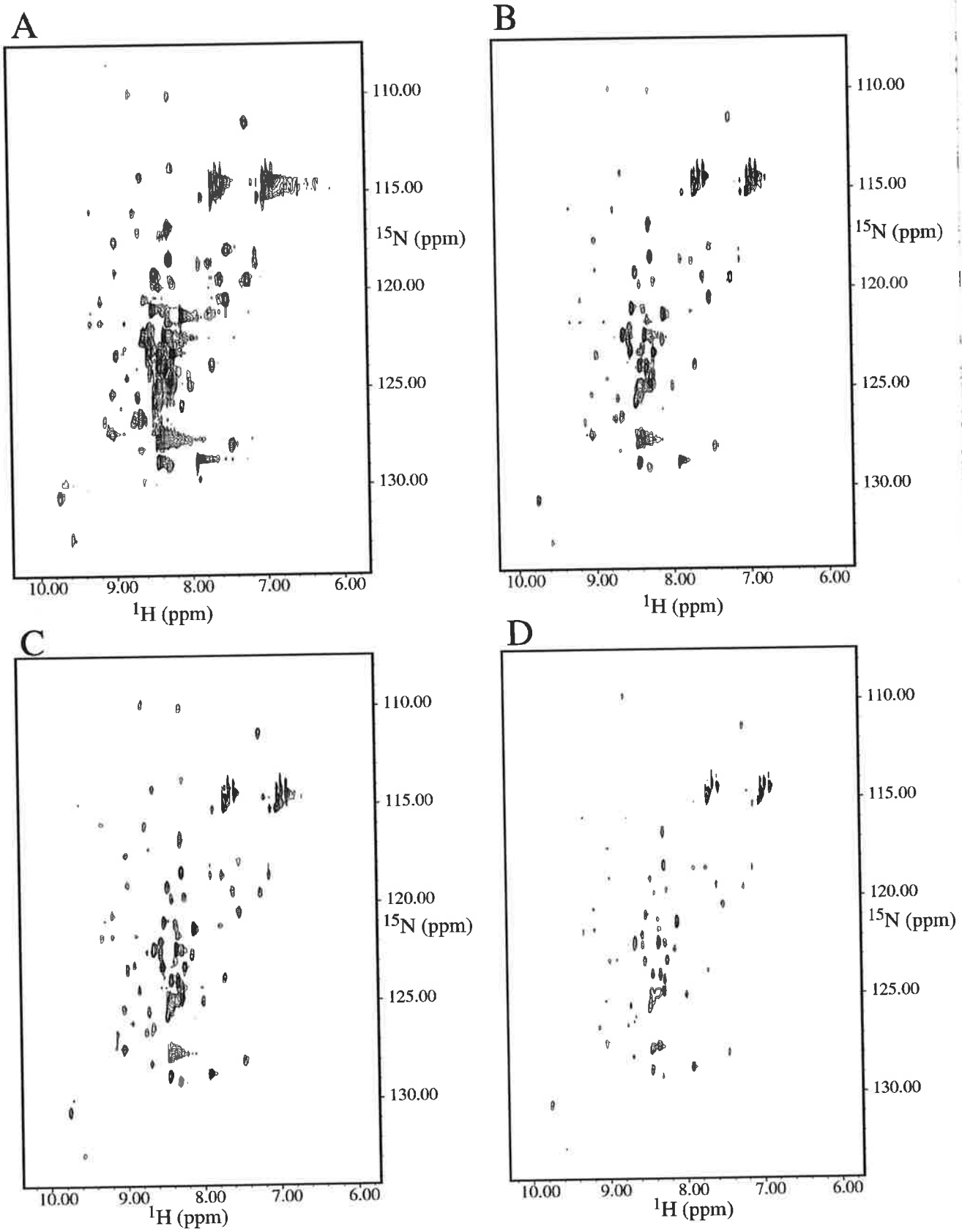
Figure 5.5

A. ^{15}N-^{1}H 2D HSQC experiment conducted on uniformly labelled ^{15}N PRRSH3 protein. ^{15}N PRRSH3 was buffered in 1 x PBS pH 6.0 at a final concentration of 0.5 mM. The data were collected on a Varian Inova 600 spectrometer equilibrated at 25°C. The 2048 points, 64 t_1 increments and 16 scans were collected with an acquisition time of 0.128 seconds producing a final size of 2048 x 256 points. Spectral widths of 8000 Hz and 2000 Hz were employed for the hydrogen and nitrogen frequencies, respectively. The data were processed with a square sinebell function on the VNMR software package. Data was transferred into XEASY for further analysis (Bartels *et al* 1997).

B. ^{15}N-^{1}H 2D HSQC collected as for A however the PRRSH3 protein was diluted 1:1 with 1 x PBS producing a protein concentration of 0.25 mM.

C. ^{15}N-^{1}H 2D HSQC of Tec PRRSH3 protein diluted to 0.125 mM. The data was collected as for A with the exception that 64 scans were collected to compensate for the decreased protein concentration.

D. ^{15}N-^{1}H 2D HSQC of Tec PRRSH3 protein diluted to 0.0625 mM. 128 t_1 increments and 128 scans were collected for improved signal to noise.


A
B
C
D
110.00
115.00
120.00
125.00
130.00
15N (ppm)
10.00
9.00
8.00
7.00
6.00
1H (ppm)

same gradient as the decrease in the sample concentration consistent with the formation of protein aggregates.

Sedimentation equilibrium is an ideal method to determine the degree of aggregation of the PRRSH3 protein in solution. Experiments to characterise the self association of Tec PRRSH3 wild-type were carried out on a Beckman Optima XL-A analytical ultracentrifuge equipped with an An-60ti rotor. Data were recorded and processed by Dr Joel Mackay, University of Sydney. Protein samples were made up in phosphate buffered saline (pH 7.0) at a number of loading concentrations in the range 8–100 μM, and data were recorded at both 4°C and 20 °C at speeds of 20,000, 30,000, 42,000, 48,000 and 56,000 rpm. Data were collected in six-sector cells as absorbance (248 and 280 nm) versus radius scans (0.001 cm increments). Scans were collected at 4-hour intervals and compared in order to determine when the sample reached equilibrium. Analysis of the data was carried out using the NONLIN software (Johnson *et al.*, 1981), and the best model and final parameters were determined by examination of the residuals derived from fits to monomer, monomer ↔ dimer, monomer ↔ trimer, and monomer ↔ dimer ↔ tetramer models (all ideal species models). The partial specific volumes of each domain were determined from the amino acid sequences (Perkins, 1986), and the solvent density was taken to be 1.0066 g ml^{-1} at 20°C.

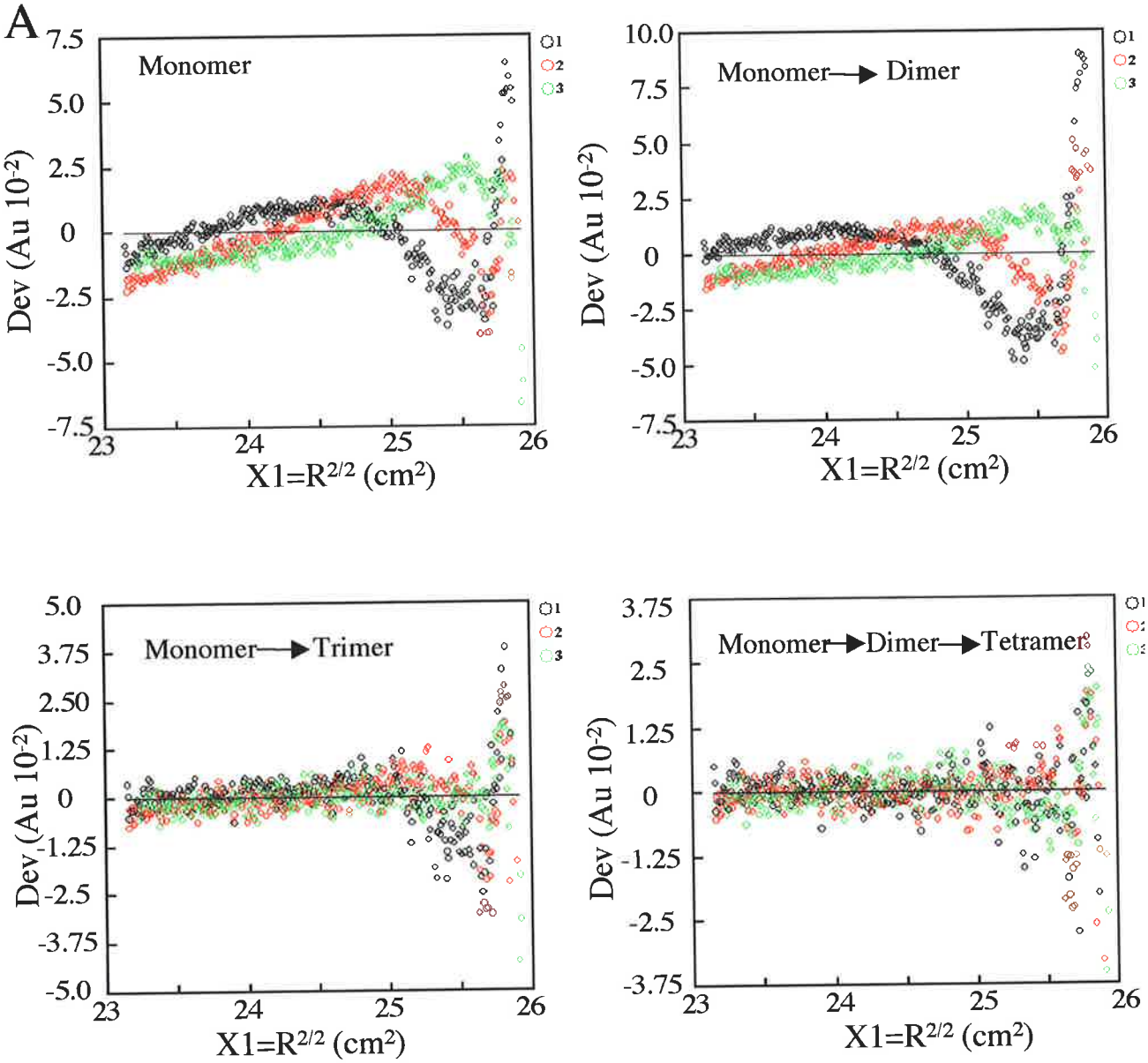
PRRSH3 protein was purified and concentrated in 1 x PBS. Three different concentrations were used with OD_{280} of 0.8, 0.27 and 0.08 the most concentrated sample being equivalent to 0.1 mM. The data obtained for Tec PRRSH3 domain protein was fitted to several models. Figure 5.6A shows the residuals of the different models with the deviation from the ideal model shown for all the models tested. Examination of the residual fits showed that the fit to the monomer model of association was poor with large deviations from zero. The same was observed for the monomer- dimer fit. The residual plots observed for the monomer-trimer and monomer-dimer-tetramer interactions were similar but the residuals fitted better to the monomer-dimer-tetramer association model. Thus, PRRSH3 protein was associating in a monomer to dimer to tetramer association pattern. A plot of the fraction of protein against molarity (Figure 5.6B) highlighted that at the 1 mM concentrations used for NMR spectroscopy approximately 70% of this protein will be tetrameric, 20% will be a dimer and only 10% will be present as a monomer (Figure 5.6). The dimer production increased sharply and then plateaued whereas the tetrameric species consistently increased, suggesting the requirement of a dimeric species for tetramer production. Thermodynamic association constants obtained from this data are representative of two associations: firstly, the monomer to dimer transition and, secondly, the monomer to

Figure 5.6

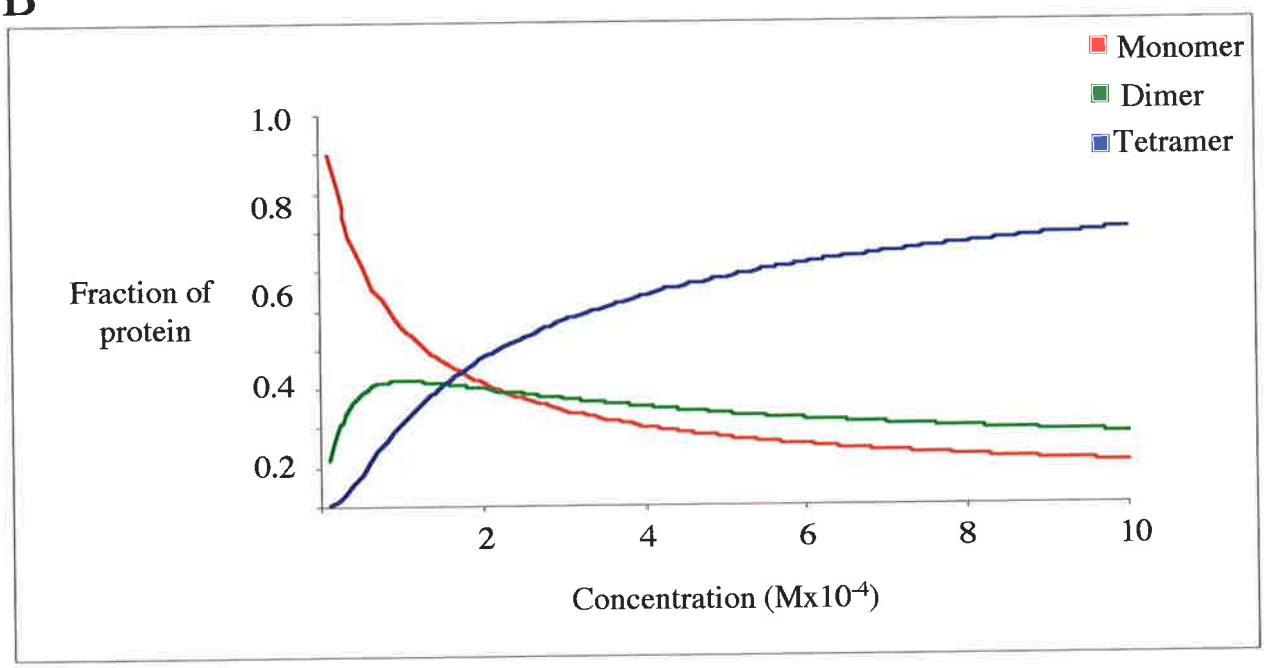
A. Ultracentifugation data collected on PRRSH3 protein. Residuals observed when sedimentation equilibrium data for the PRRSH3 wild-type protein was fitted to a series of different models; monomer, monomer-dimer, monomer-trimer and monomer-dimer-tetramer. The sedimentation equilibrium experiments were conducted on a Beckman Optima XL-A analytical ultracentrifuge with an An-60ti rotor. The samples were collected from the size exclusion preparation in 1 x PBS and further concentrated using a Centricon 3 spin column until a concentration of 0.1 mM was reached or diluted appropriately. Three different PRRSH3 protein concentrations were tested, 0.1 mM, 0.03 mM and 0.01 mM at wavelengths 248 nm and 280 nm.

B. A plot of ultracentrifugation data as a fraction of protein against concentration for monomer dimer and tetramer species. The monomer species (red line) decreased rapidly. Dimer formation (green) increased sharply and then plateaued at 1 mM protein concentration and remained almost level through the concentrations plotted. The tetramer species (blue) present in a solution of PRRSH3 increased over the concentrations plotted in this graph however the rate decreased as protein concentration increased.

A


Monomer
Monomer→ Dimer
Monomer→Trimer
Monomer→Dimer→Tetramer
Dev (Au 10-2)
X1=R2/2 (cm2)
1
2
3

B


Monomer
Dimer
Tetramer
Fraction of protein
Concentration (Mx10-4)

tetramer transition. The monomer to dimer transition can also be represented by the dissociation constant K_D. PRRSH3 protein was determined to have a dissociation constant of K_D of 125 μM (range 200-91 μM) for the monomer to dimer transition, which corresponds to an association constant of 8000 M^{-1} (range 5000-11000 M^{-1}). The association constant for the monomer to tetramer association is $1.2x10^{12} M^{-3}$ (range $0.8\text{-}1.6x10^{12} M^{-3}$).

Various biophysical methods have shown that the PRRSH3 protein is aggregating and ultracentrifugation has shown the presence of tetramer with a small percentage dimer. This data highlights that in the wild-type protein the isolated PRRSH3 domain will preferentially form aggregates rather than bind intramolecularly. This is in contrast to the evidence seen for the PRRSH3 region of Itk. In Itk, the SH3 and the PRR intramolecularly associate and form a weak although stable complex observed in NMR experiments. The PRR region of Itk contains only one SH3 consensus-binding site (Andreotti *et al.*, 1997) whereas the PRR region of Tec contains two sites with one containing three overlapping consensus binding sites. The complexity of the PRR of Tec kinase could explain the different results observed compared with Itk, however, the dimer/tetramer formation observed is consistent with a dimerisation interaction observed for a similar region in Btk. Site-directed mutagenesis was used to further characterise the dimerisation interaction observed and determine the contributions of the individual sites to this PRRSH3 dimerisation event.

5.4 MUTAGENESIS AND EXPRESSION OF TEC PRRSH3 PROTEINS

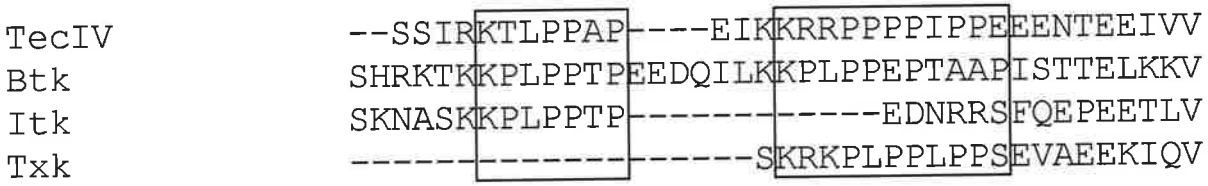
In order to elucidate the contribution of each SH3 consensus-binding site present in the PRR to the dimer/tetramer interaction observed for the PRRSH3 wild-type protein, site directed mutagenesis was used to mutate these sites individually and also in concert. Figure 5.7 shows alignment of PRR's from Tec family members. Also shown are the amino acid sequences to which the PRR of Tec kinase was mutated. SH3 domain ligands form a polyproline type 2 helix and disruption of this helix would be expected to prevent ligand binding (Andreotti *et al.*, 1997). pGEX-4T-2-PRRSH3Δ1, encoding the PRRSH3 protein that lacks PRR SH3 consensus binding site 1 was formed using pGEX-4T-2-PRRSH3 plasmid DNA as the template for site-directed mutagenesis by the Quickchange™ mutagenesis kit (section 2.2.3). Site 1 primers incorporate base changes that alter Pro 158 to Ala and would be expected to disrupt the polyproline type 2 helix. Following PCR amplification of the mutants and subsequent *Dpn*I digestion the DNA was electroporated into DH5α competent cells and minipreparations from transformants were screened by DNA

Figure 5.7

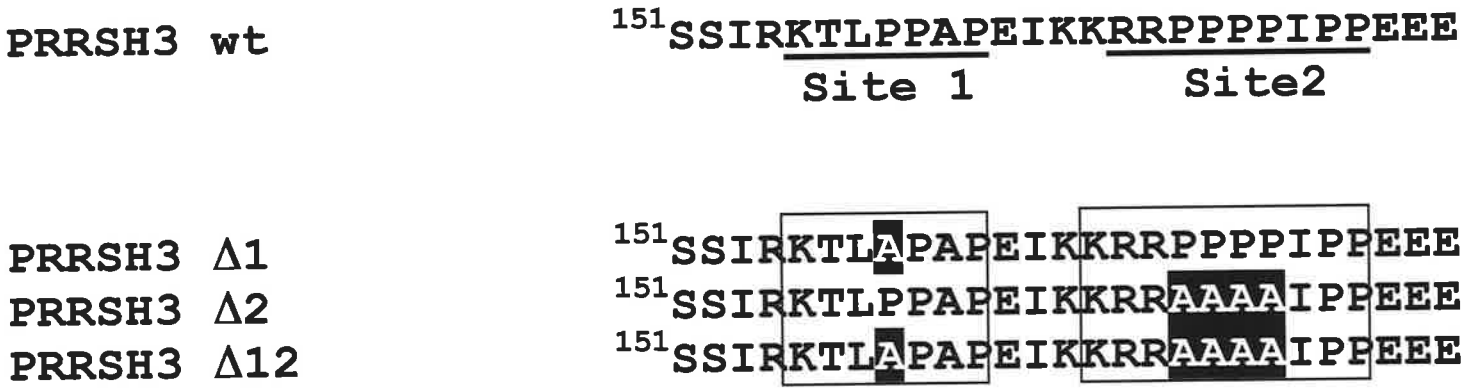
A. Alignment of the proline rich regions of Tec family members Tec, Btk, Itk and Txk. The two regions referred to as site 1 and site 2 are boxed in the alignment. Tec and Btk have two SH3 binding consensus regions, while Itk and Txk only have one. Itk contains PRR site 1 and Txk contains the second region of proline rich residues (PRR site 2). All sites conform to the SH3 ligand class 1 sequence of +XXPXXP. The alignment was conducted using a ClustalW algorithm (Thompson *et al.*, 1994).

B. Wild-type Tec PRR aligned with PRR mutants generated. PRRSH3Δ1 mutant contains a Pro-Ala mutation at residue number 158. PRRSH3Δ2 mutant contains Pro-Ala mutations at residues 168-171. The numbers reflect the position in the mouse Tec IV kinase. The mutants were generated using the Stratagene Quikchange PCR mutagenesis kit using the primers in section 2.2.7.

A


TecIV
Btk
Itk
Txk
--SSIRKTLPPAP----EIKKRRPPPPIPPEEENTEEIVV
SHRKTKKPLPPTPEEDQILKKPLPPEPTAAPISTTELKKV
SKNASKKPLPPTP----------EDNRRSFQEPEETLV
----------SKRKPLPPLPPSEVAEEKIQV

B


PRRSH3 wt
151SSIRKTLPPAPEIKKRRPPPPIPPEEE
Site 1
Site2
PRRSH3 Δ1
PRRSH3 Δ2
PRRSH3 Δ12
151SSIRKTLAPAPEIKKRRPPPPIPPEEE
151SSIRKTLPPAPEIKKRRAAAAIPPEEE
151SSIRKTLAPAPEIKKRRAAAAIPPEEE

sequencing (section 2.3.13). One out of 11 plasmids sequenced contained the desired mutation. This plasmid carries the correct mutation and is termed pGEX-4T-2-PRRSH3Δ1.

pGEX-4T-2-PRRSH3Δ2 was also generated using Quickchange™ mutagenesis using pGEX-4T-2-PRRSH3 as a template. The PCR mutagenesis of PRR site 2 required 20 cycles compared with the 15 required for the site 1 mutagenesis, possibly because the number of bases being changed was greater, resulting in lower efficiency. This mutagenesis reaction mutated prolines 168-171 each to alanine. pGEX-4T-2-PRRSH3Δ12 double mutant was generated using pGEX-4T-2-PRRSH3Δ2 as the template and the site 1 primers in a Quickchange™ mutagenesis reaction. Products of the mutagenesis were screened using DNA sequencing and correct sequence was confirmed by sequencing in the opposite direction (section 2.3.13). Figure 5.8 shows the sequencing chromatograms over the two PRR sites for all the proteins produced. pGEX-4T-2-PRRSH3Δ1, pGEX-4T-2-PRRSH3Δ2 and pGEX-4T-2-PRRSH3Δ12 were transformed into BL-21λDE3 *E. coli* for expression studies.

PRRSH3Δ1, PRRSH3Δ2 and PRRSH3Δ12 proteins were expressed and purified (section 2.4.1, 2.4.4, 2.4.9, 2.4.11). The GST fusion proteins were affinity purified on glutathione agarose, cleaved with thrombin and then separated by size exclusion chromatography on a Superdex G75 column equilibrated in 1 x PBS or 1 x HBS. The profiles from these purifications suggested that these proteins might also aggregate (data not shown). The purifications were monitored by SDS-PAGE. The yields of these proteins differed with 75 mg/L-fusion protein for PRRSH3Δ1 and 113 mg/L and 89 mg/L for PRRSH3Δ2 and PRRSH3Δ12, respectively and the mass of these proteins was confirmed by mass spectrometry (Figure 5.9). Analysis by mass spectrometry of the PRRSH3Δ1 protein shows a degree of breakdown. The PRRSH3 wild-type protein has a mass of 11214 Da consistent with the calculated mass. The PRRSH3Δ1 protein has a mass of 11188 Da as shown by the larger of the peaks. The size of the PRRSH3Δ1 protein is slightly smaller than that of PRRSH3 wild-type reflecting the P-A substitution. The relative positions of the breakdown are indicated by the peak. All digestions are occurring from the N-terminus and are cleaving at lysine residues. The peak at 10687 Da has been cleaved at residue Lys 156 producing the fragment Ser 151-Lys 156 not seen by mass spectrometry in this size region. The peak at 9638 Da has been cleaved at residue Lys 166 producing two fragments, one of Ser 151-Lys 156 and the other Thr 157-Lys 166 (Figure 5.9). PRRSH3Δ2 protein has a full-length mass of 11110 Da consistent with P-A mutation of residues 168-171. There is a

Figure 5.8

Sequence data for wild-type and mutant PRRSH3 proteins. Sequencing was conducted by PCR using the sequencing primers for pGEX as per the pGEX manual. The PRR regions have been boxed in red (Site 1) or blue (Site 2) reflecting the two different sites. The sites are labelled above indicating whether they are mutant (MUTANT) or wild-type sequences (WT). The coding region of each of these constructs was sequenced in both directions to confirm that no additional mutations had been incorporated.

A. wild-type PRRSH3 coding sequence with the two PRR sites boxed. The first site, Site 1 is boxed in red and the second site, Site 2 is boxed in blue.

B. wild-type PRRSH3 coding sequence electropherogram

C. PRRSH3Δ1 coding sequence electropherogram

D. PRRSH3Δ2 non-coding sequence electropherogram

E. PRRSH3Δ12 non-coding sequence electropherogram

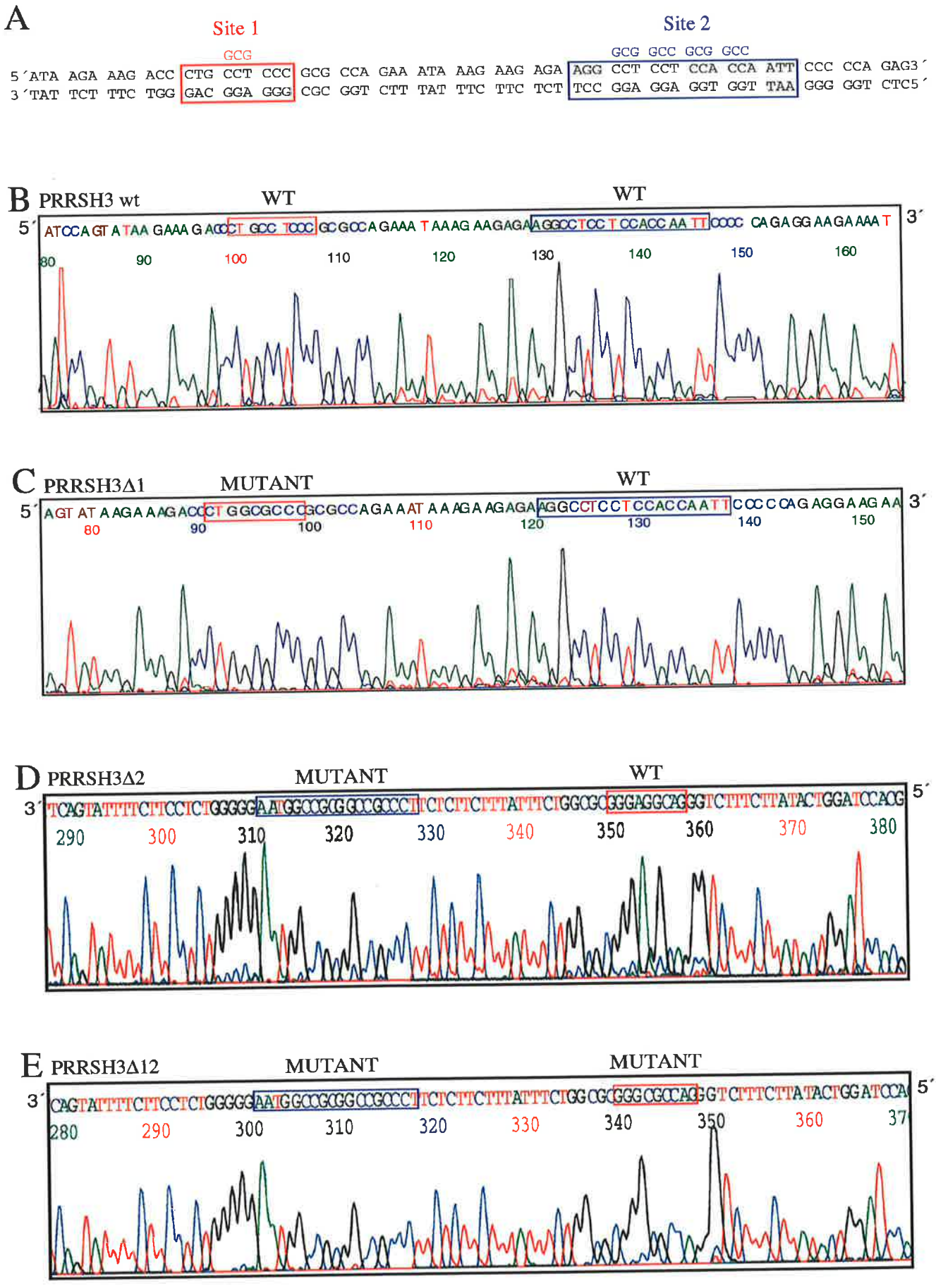

A
Site 1
GCG
Site 2
GCG GCC GCG GCC
5´ATA AGA AAG ACC CTG CCT CCC GCG CCA GAA ATA AAG AAG AGA AGG CCT CCT CCA CCA ATT CCC CCA GAG3´
3´TAT TCT TTC TGG GAC GGA GGG CGC GGT CTT TAT TTC TTC TCT TCC GGA GGA GGT GGT TAA GGG GGT CTC5´
B PRRSH3 wt
WT
WT
5´
3´
80 90 100 110 120 130 140 150 160
C PRRSH3Δ1
MUTANT
WT
5´
3´
80 90 100 110 120 130 140 150
D PRRSH3Δ2
MUTANT
WT
3´
5´
290 300 310 320 330 340 350 360 370 380
E PRRSH3Δ12
MUTANT
MUTANT
3´
5´
280 290 300 310 320 330 340 350 360

Figure 5.9

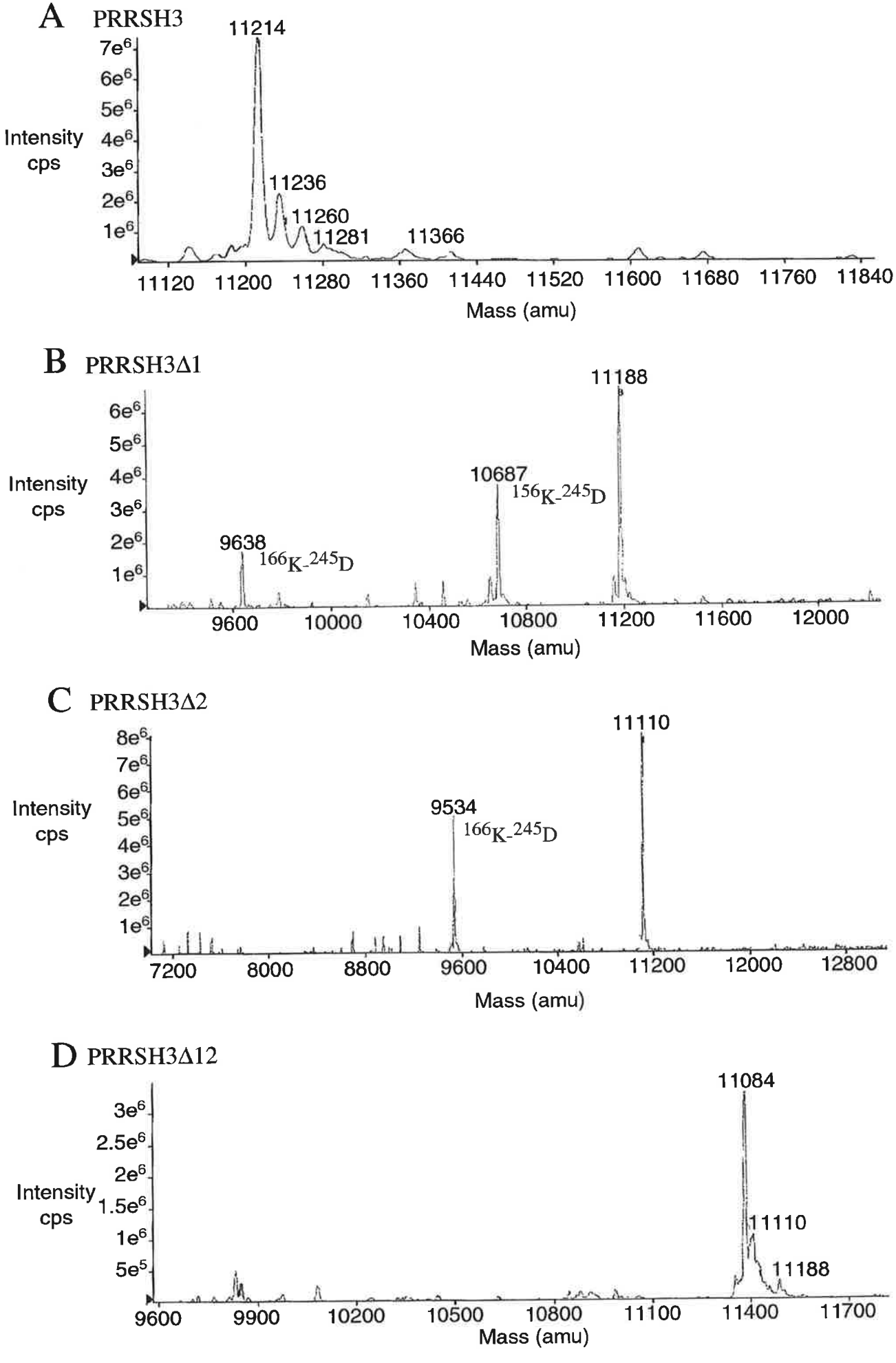
Analysis of the PRRSH3 proteins by mass spectrometry. Samples were purified by HPLC, lyophilised and sent for analysis to the Waite mass spectrometry unit, University of Adelaide. The molecular mass (amu) of the peaks are indicated above the peak. Molecular masses of 11214, 11188, 11110 and 11084 Da are consistent with calculated molecular masses for:

A. PRRSH3 wild-type,

B. PRRSH3Δ1

C. PRRSH3Δ2 and

D. PRRSH3Δ12, respectively.


A PRRSH3
11214
11236
11260
11281
11366
Intensity cps
7e6
6e6
5e6
4e6
3e6
2e6
1e6
11120 11200 11280 11360 11440 11520 11600 11680 11760 11840
Mass (amu)
B PRRSH3Δ1
11188
10687
156K-245D
9638
166K-245D
Intensity cps
6e6
5e6
4e6
3e6
2e6
1e6
9600 10000 10400 10800 11200 11600 12000
Mass (amu)
C PRRSH3Δ2
11110
9534
166K-245D
Intensity cps
8e6
7e6
6e6
5e6
4e6
3e6
2e6
1e6
7200 8000 8800 9600 10400 11200 12000 12800
Mass (amu)
D PRRSH3Δ12
11084
11110
11188
Intensity cps
3e6
2.5e6
2e6
1.5e6
1e6
5e5
9600 9900 10200 10500 10800 11100 11400 11700
Mass (amu)

breakdown peak at mass 9534 Da consistent with cleavage at residue Lys 166. This is the same digestion observed for the PRRSH3Δ1 protein. PRRSH3Δ12 protein has a mass of 11084 Da as observed by mass spectrometry; this is consistent with the P-A mutation of residues 158 and 168-171. Further preparations were generated with shorter digestion times in order to minimise the breakdown products observed.

5.5 PRRSH3 MUTANT PROTEINS FORM DIMERS TO DIFFERING DEGREES

5.5.1 Ultracentrifugation

Analytical ultracentrifugation was used to determine the contribution of each of the PRR sites to the dimerisation/tetramerisation reaction already observed. These ultracentrifugation experiments are a solution based assay that provide a relative molecular mass and, thus, allow extrapolation to the multimeric state of the different PRRSH3 proteins. Thus, the relative percentage of different forms of the protein can be determined by ultracentrifugation. The data were collected and processed by Dr Joel Mackay, Department of Biochemistry, University of Sydney.

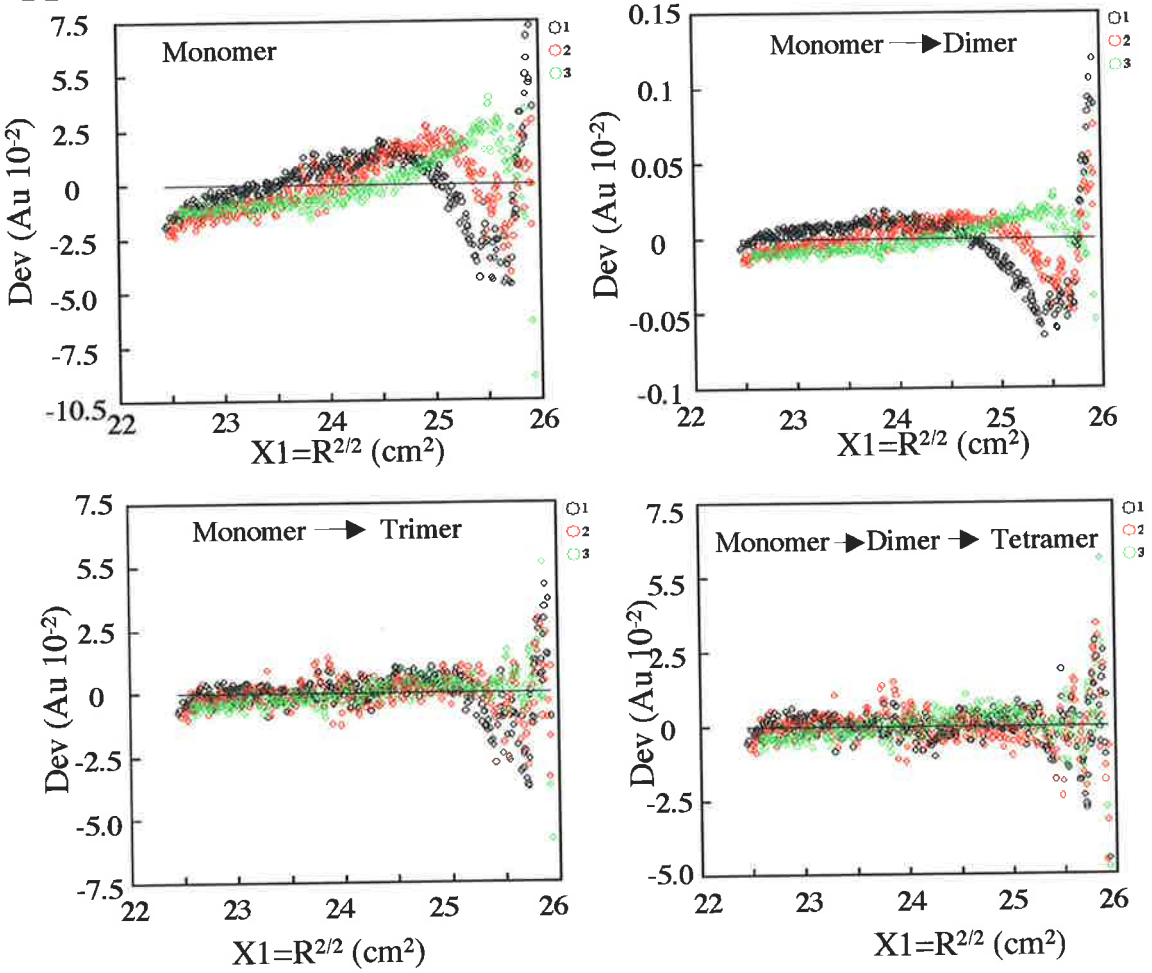
The data and processing of the PRRSH3 mutant proteins was conducted in an analogous manner to the wild-type PRRSH3 protein. These data were fitted to a series of ideal models and residuals to these fits analysed. Similar to the wild-type protein, all the mutants fitted best to a monomer-dimer-tetramer model. Figure 5.10A to 5.12A shows the residual plots from the fitted data for PRRSH3Δ1, PRRSH3Δ2 and PRRSH3Δ12 respectively. The residuals are minimised in the monomer dimer tetramer association model. Ultracentrifugation data can be visualised following analysis by the NONLIN analysis software by plotting the relative amounts (fractions of the protein solution) of each different multimeric state against protein concentration. Figure 5.6B shows the results for the PRRSH3 wild-type protein indicating the presence of a high proportion of tetramer, which is also observed for PRRSH3Δ1 at the same concentration (Figure 5.10B). At 1 mM concentration there will be approximately 7% monomer, 25% dimer and 68% tetramer in a solution of PRRSH3Δ1. This implies that PRRSH3Δ1 equilibrium is favoured to dimer/tetramer production similar to PRRSH3 wild-type protein. PRRSH3Δ2 and PRRSH3Δ12 have a lower proportion of the protein present as a dimer or tetramer compared with PRRSH3 wild-type at all the concentrations tested. PRRSH3Δ2 contains 25%

Figure 5.10

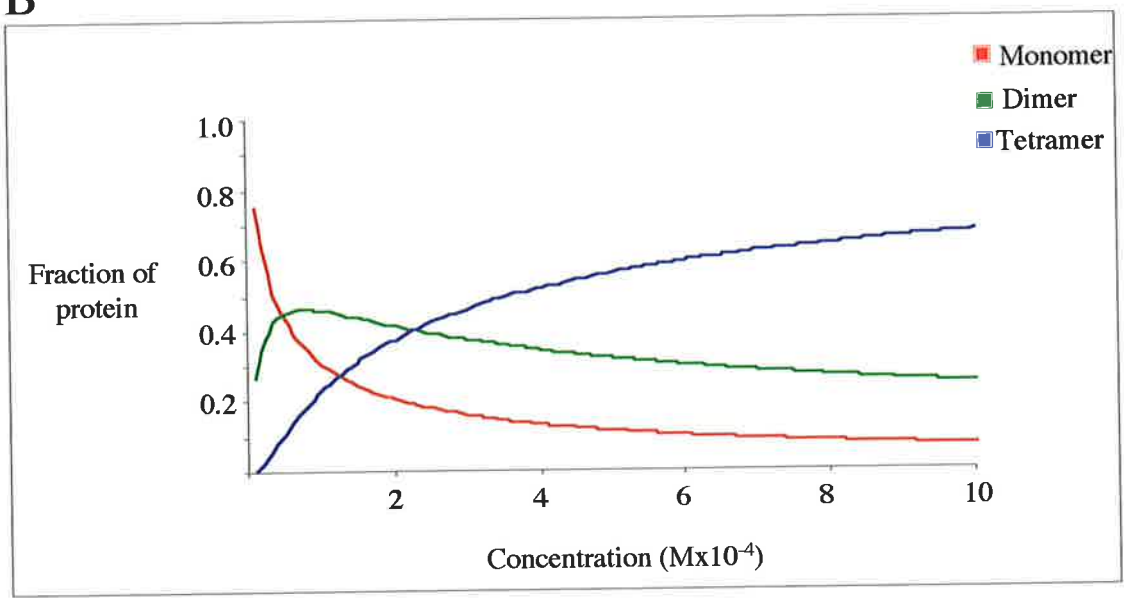
A. Ultracentrifugation data for PRRSH3Δ1 protein. Residuals observed when sedimentation PRRSH3Δ1 protein equilibrium data was fitted to a series of different models; monomer, monomer-dimer, monomer-trimer and monomer-dimer-tetramer. The sedimentation equilibrium experiments were conducted on a Beckman Optima XL-A analytical ultracentrifuge with an An-60ti rotor. The samples were collected from the size exclusion preparation in 1 x PBS further concentrated using Centricon 3 until a concentration of 0.1 mM was reached or diluted appropriately. Three different PRRSH3 protein concentrations were tested, 0.1 mM, 0.03 mM and 0.01 mM at two different wavelengths 248 nm and 280 nm. The residuals were observed to best fit the monomer-dimer-tetramer equilibrium model.

B. Plot of ultracentrifugation data for PRRSH3Δ1 as a fraction of protein against concentration of monomer, dimer and tetramer species. The monomer species (red) decreased exponentially. The tetramer species (blue) present in the solution of PRRSH3 increased over the concentrations plotted in this graph however the rate decreased as protein concentration increased. Dimer formation (green) increased sharply and then plateaued at less than 0.1 mM protein concentration and then slowly decreased.

A


Monomer
Monomer → Dimer
Monomer → Trimer
Monomer → Dimer → Tetramer
Dev (Au 10-2)
X1=R2/2 (cm2)
1
2
3

B


Monomer
Dimer
Tetramer
Fraction of protein
Concentration (Mx10-4)

monomer, 31% dimer and 42% tetramer at a 1 mM concentration whereas PRRSH3Δ12 contains 30% monomer, 20% dimer and 50% tetramer. Although PRRSH3Δ2 and PRRSH3Δ12 samples do form the same associations as the PRRSH3 wild-type and PRRSH3Δ1 proteins, they do so at a lower affinity and under the conditions of these and the following experiments were predominantly monomer (concentrations below 100 μM) (Figures 5.11B and 5.12B respectively). Thermodynamic association constants obtained from this data are shown in Table 5.1 and indicate that the equilibrium monomer to dimer transition favours dimer formation for both the wild-type and PRRSH3Δ1 proteins with $K_{Dmonomer\text{-}dimer}$ of 125 μM and 50 μM respectively. The $K_{Dmonomer\text{-}dimer}$ values for PRRSH3Δ2 and PRRSH3Δ12 are 417 μM and 909 μM, 4 and 8 fold weaker than PRRSH3 wild-type, respectively. PRRSH3Δ2 and PRRSH3Δ12 dimerise and tetramerise to a lesser extent than PRRSH3 wild-type and PRRSH3Δ1. The thermodynamic association constants for the monomer to tetramer transition for the PRRSH3 proteins are shown in Table 5.1 and these indicate that the equilibrium for tetramer production is favoured for the wild-type PRRSH3 protein ($1.2x10^{12}\ M^{-3}$) and PRRSH3Δ1 ($6.3x10^{12}\ M^{-3}$) whereas for PRRSH3Δ2 ($2.5x10^{10}\ M^{-3}$) and PRRSH3Δ12 ($1.5x10^{10} M^{-3}$) tetramer production is not favoured. This ultracentrifugation data indicates that site 2 in the PRR promotes an intermolecular interaction with the SH3 domain. The proteins with an intact PRR site 2 (PRRSH3 wild-type and PRRSH3Δ1) dimerise and tetramerise under the conditions of these ultracentrifugation experiments (section 5.3). The lower association constants for the PRRSH3Δ2 and PRRSH3Δ12 proteins suggest that they do not form dimers and tetramers; hence, PRR site 1 cannot promote dimerisation. By extension, proteins containing PRR site 1 are potentially forming an intramolecular interaction. A schematic representing the observed results of the ultracentrifugation is shown in Figure 5.13. The PRRSH3 protein is represented and those interactions that are favoured based on the ultracentrifugation data are shown in pink boxes. BIAcore experiments were conducted to further characterise the association of Tec PRRSH3.

5.5.2 BIAsensor experiments indicate that PRRSH3 mutants bind immobilised SH3 protein differentially

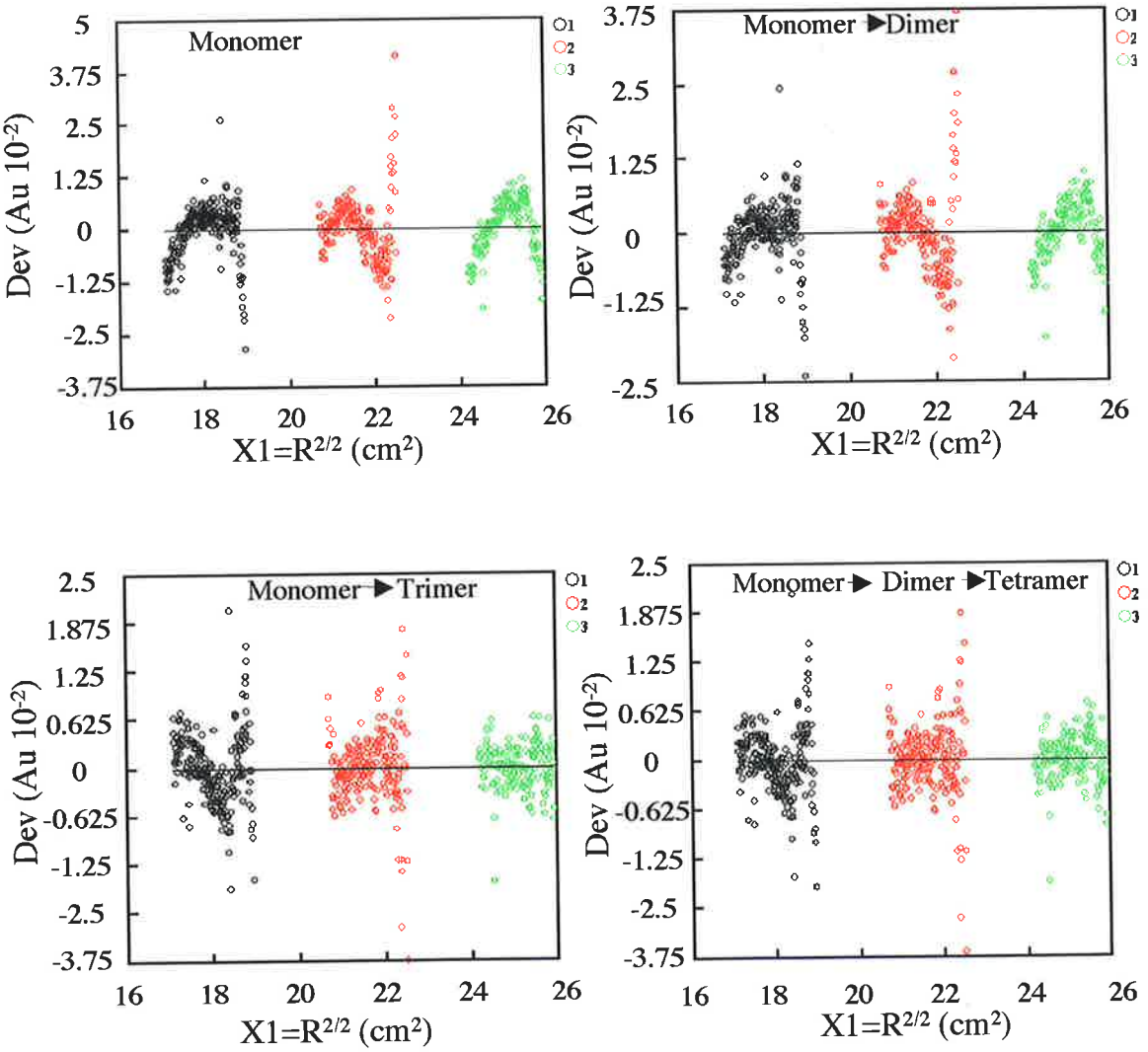
Analytical ultracentrifugation has shown that the SH3 consensus sites within the PRR have different dimerisation affinities and, thus, may perform different functions in the context of the whole protein. PRRSH3 wild-type protein dimerises and tetramerises as does

Figure 5.11

A. Ultracentrifugation data for PRRSH3Δ2 protein. Residuals observed when sedimentation equilibrium data for the PRRSH3Δ2 protein was fitted to a series of different models; monomer, monomer-dimer, monomer-trimer and monomer-dimer-tetramer. The sedimentation equilibrium experiments were conducted on a Beckman Optima XL-A analytical ultracentrifuge with an An-60ti rotor. The samples were collected from the size exclusion preparation in 1 x PBS further concentrated using Centricon 3 until a concentration of 0.1 mM was reached or diluted appropriately. Three different PRRSH3 protein concentrations were tested, 0.1 mM, 0.03 mM and 0.01 mM at wavelengths of 248 nm and 280 nm. The residuals were observed to best fit the monomer-dimer-tetramer equilibrium.

B. Plot of ultracentrifugation data for PRRSH3Δ2 protein as a fraction of protein against concentration for monomer dimer and tetramer species. The monomer species (red) decreased while the tetramer species (blue) increased linearly over the concentrations plotted in this graph. Dimer formation (green) increased sharply and then plateaued at 0.2 mM protein concentration and remained level through the concentrations plotted.

A


Monomer
Monomer ▶Dimer
Monomer ▶Trimer
Monomer ▶ Dimer ▶Tetramer
Dev (Au 10-2)
X1=R2/2 (cm2)
1
2
3

B

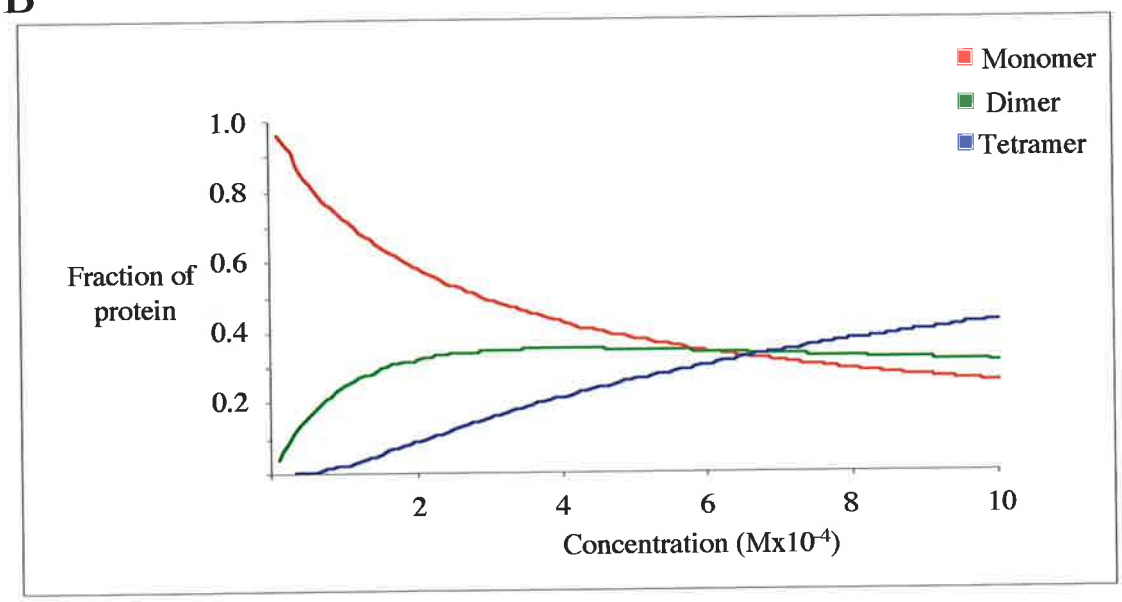
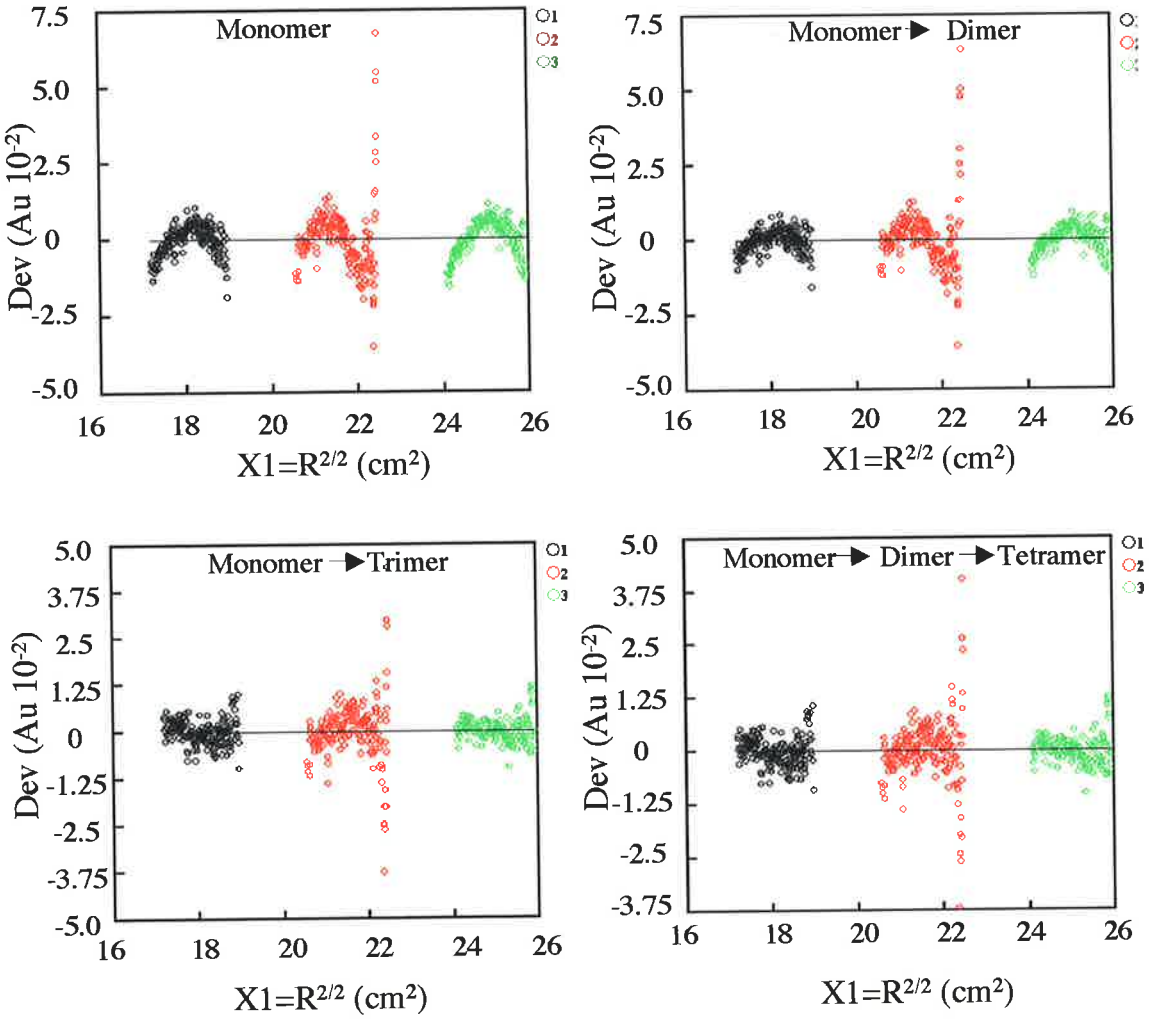

Monomer
Dimer
Tetramer
Fraction of protein
Concentration (Mx10-4)

Figure 5.12

A. Ultracentrifugation data for PRRSH3Δ12 protein. Residuals observed when sedimentation equilibrium data for the PRRSH3Δ12 protein was fitted to a series of different models; monomer, monomer-dimer, monomer-trimer and monomer-dimer-tetramer. The sedimentation equilibrium experiments were conducted on a Beckman Optima XL-A analytical ultracentrifuge with an An-60ti rotor. The samples were collected from the size exclusion preparation in 1 x PBS further concentrated using Centricon 3 until a concentration of 0.1 mM was reached or diluted appropriately. Three different PRRSH3 protein concentrations were tested, 0.1 mM, 0.03 mM and 0.01 mM at wavelengths of 248 nm and 280 nm. The residuals were observed to best fit the monomer-dimer-tetramer equilibrium.

B. Plot of ultracentrifugation data for PRRSH3Δ12 protein as a fraction of protein against concentration for monomer dimer and tetramer species. The monomer species (red) decreased while the tetramer species (blue) increased linearly over the concentrations plotted. Dimer formation (green) increased sharply and then plateaued at 0.2 mM protein concentration and remained level through the concentrations plotted.

A


Monomer
Monomer ▶ Dimer
Monomer ▶ Trimer
Monomer ▶ Dimer ▶ Tetramer
Dev (Au 10-2)
X1=R2/2 (cm2)
1
2
3

B

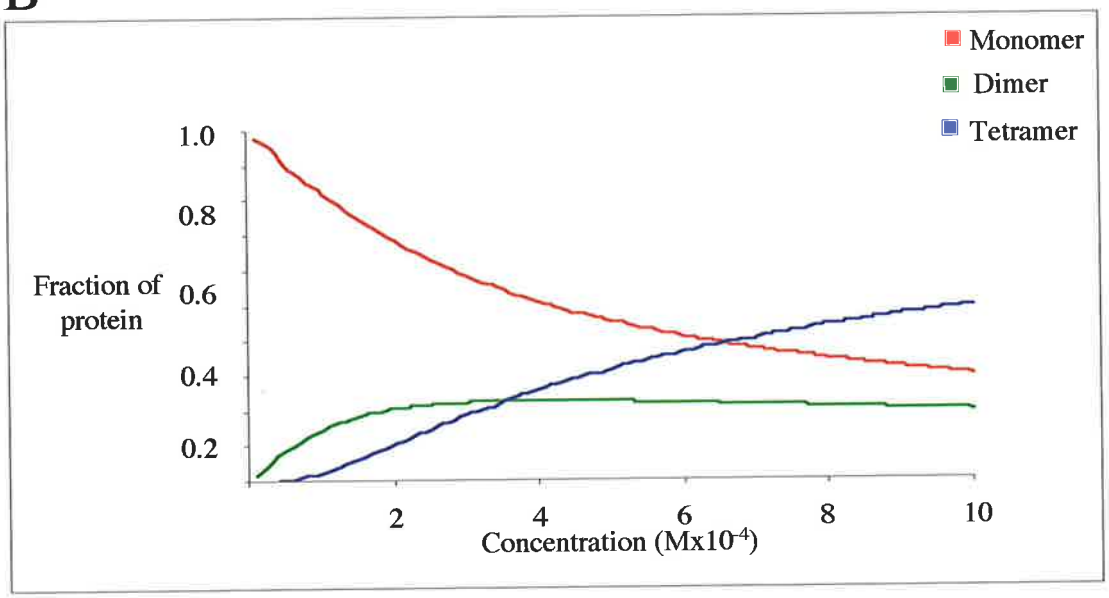

Monomer
Dimer
Tetramer
Fraction of protein
1.0
0.8
0.6
0.4
0.2
2
4
6
8
10
Concentration (Mx10-4)

Figure 5.13

Schematic representation of possible PRRSH3 protein interactions. The SH3 domain component of the protein is shown by the grey oval with the ligand binding site highlighted by the indent. PRR site 1 (red) and is furthest away from the SH3 domain. PRR site 2 is indicated in green. Length of arrows indicates the probable shift of the equilibrium reaction. The three options highlighted by pink boxes are the conclusions determined from the sedimentation equilibrium experiments conducted on the mutant proteins. Question marks indicate interactions where no information is known.

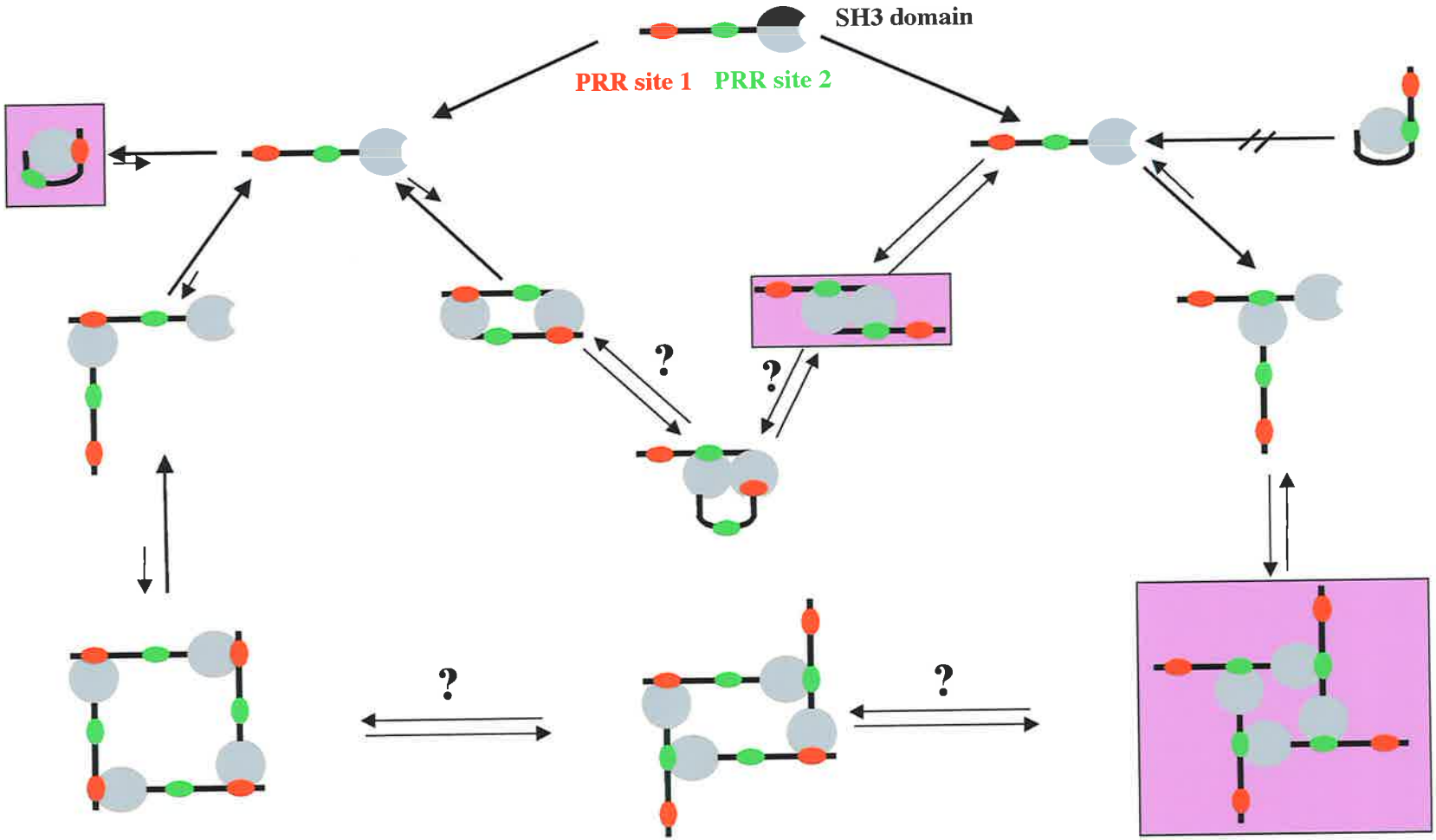

SH3 domain
PRR site 1
PRR site 2
?
?
?
?

Table 5.1 Thermodynamic Equilibrium Constants determined by Analytical Ultracentrifugation for Tec PRRSH3 proteins

	$K_{2\ monomer\text{-}dimer}$ M^{-1}	K_2 Range M^{-1}	$K_{D\ monomer\text{-}dimer}$ μM	K_D Range μM	$K_{4\ dimer\text{-}tetramer}$ M^{-3}	K_4 Range M^{-3}
PRRSH3 Wild-type	8000	5000-11000	125	200-91	$1.2x10^{12}$	$0.8\text{-}1.6x10^{12}$
PRRSH3Δ1	20000	14000-30000	50	71-33	$6.3x10^{12}$	$4.3\text{-}9.3x10^{12}$
PRRSH3Δ2	2400	1800-3200	417	555-312	$2.5x10^{10}$	$2.1\text{-}3.1x10^{10}$
PRRSH3Δ12	1100	830-1600	909	1205-625	$1.5x10^{10}$	$1.3\text{-}1.8x10^{10}$

PRRSH3Δ1 (at 170 μM there are approximately equal amounts of monomer, dimer and tetramer), however, PRRSH3Δ2 and PRRSH3Δ12 have reduced affinity of dimerisation (Δ2; 700 μM for equal amounts of monomer, dimer and tetramer). To gain a better understanding of the specificity and strength of the different interactions occurring between the SH3 domain and the different components of the PRR of Tec kinase, these proteins were analysed by surface plasmon resonance technology -BIAcore.

Surface plasmon resonance technology is a highly reproducible method of determining kinetics of a variety of macromolecular interactions. This kinetic information can relate the structure to the molecular function of a variety of macromolecules (Myszka, 1997). BIAcore experiments consist of one component attached to a gold surface termed the ligand and an analyte in solution that is flowed across the surface of the chip. A light source monitors the reaction and following an interaction between ligand and analyte at the surface, the light is diffracted at different angles to produce a response in arbitrary units (Myszka, 1997). The binding of the analyte to the ligand is monitored in real time. The dissociation phase can be monitored by replacement of the sample by buffer and the loss of bound analyte is followed (Morton *et al.*, 1995). BIAcore is a solid phase assay that can detect the interaction of the PRRSH3 protein for the immobilised SH3 domain on the chip surface. The complex interactions occurring in solution are not detected by BIAcore experiments.

BIAcore experiments were conducted with the SH3 domain immobilised on the chip as a ligand. Different PRRSH3 proteins as analytes were flowed over the chip and the binding reaction was monitored in real time. Tec SH3 domain protein was immobilised to a CM5 chip as per the manufacturer's instructions. This chip binds proteins via primary amine groups and can thus couple through either lysine sidechains or the N-terminus of the protein. As both positions are capable of coupling it is difficult to determine the exact orientation of the protein, and by extension, the accessibility of the SH3 domain-binding pocket to the analytes in solution. Other methods for immobilisation such as metal chelates, steptavidin and antibody coupling were not available. The SH3 domain sample was diluted in NaOAc pH 4.6 to facilitate immobilisation. The SH3 domain was unstable at pH 4.6 and, thus, the sample was diluted just before running the immobilisation procedure. The SH3 sample used was tested for structural integrity by NMR and mass spectrometry prior to use (Chapter 4). A profile obtained for the immobilisation of the SH3 domain is shown in Figure 5.14B, which can be compared, to a blank lane (Figure 5.14A). The PRRSH3 kinetic experiments were conducted with 1500-3000 RU of SH3 domain immobilised (Figure 5.14B).

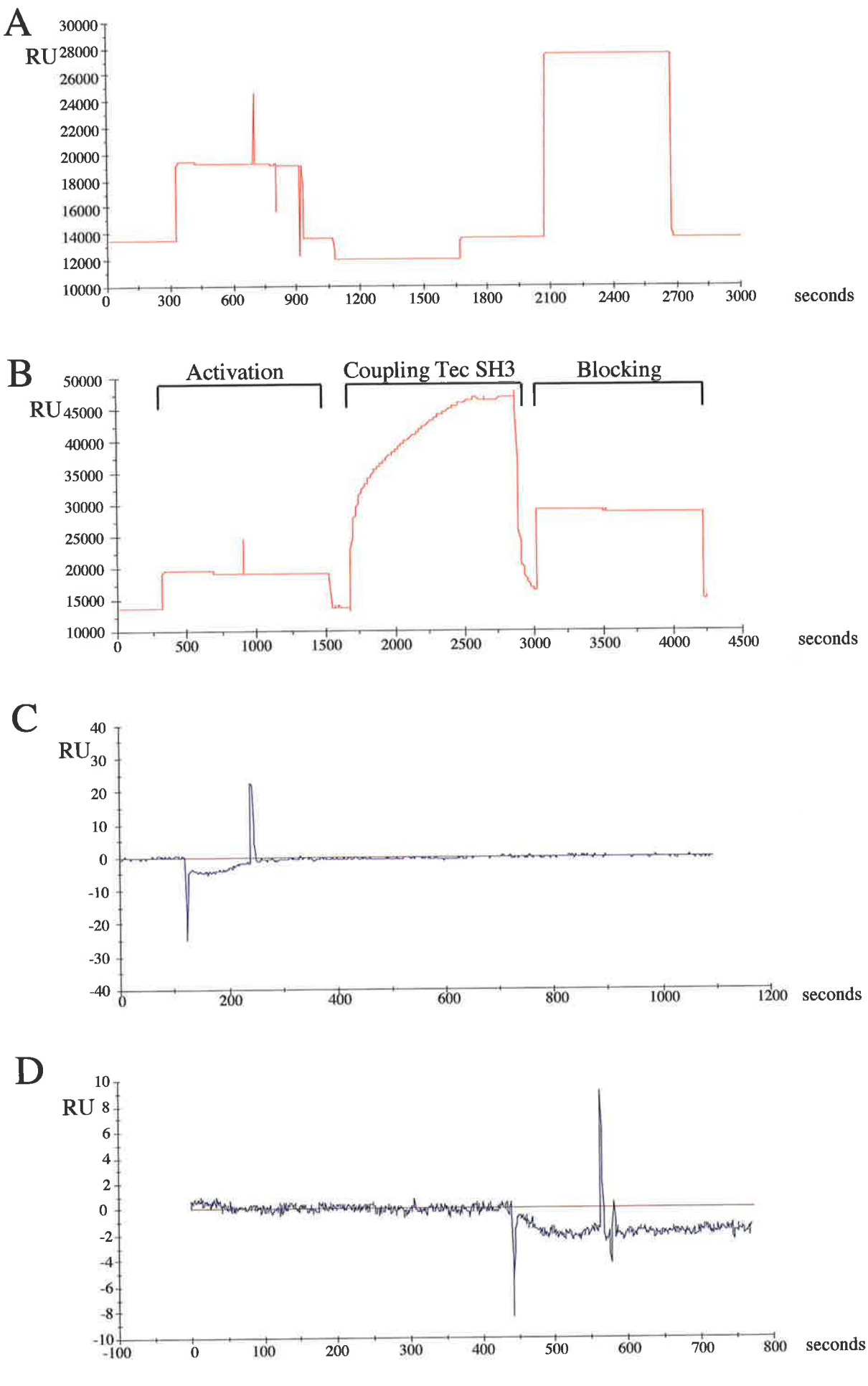
Figure 5.14

A. Sensorgram representing the immobilisation of a blank lane on a CM5 BIAcore chip. CM5 chips couple through amide chemistry to form a covalent linkage. The chip is activated by 100 µl of a 1:1 NHS/EDC mixture and all unbound amide groups are blocked by the injection of 100 µl ethanolamine. A buffer sample (HBS, 10 mM HEPES, 150 mM NaCl 3 mM EDTA, 0.005% P20 surfactant) was injected over the chip surface at 5 µl/minute. One blank lane was activated per chip. In all experiments, the response from the background lane was subtracted from the sensorgram of interest.

B. Sensorgram showing immobilisation of Tec SH3 domain onto a CM5 BIAcore chip. The different stages of immobilisation have been highlighted. The SH3 domain protein was diluted to 1 mg/ml in 10 mM NaOAc pH 4.6 to facilitate amide bond coupling. The chip was activated as for A (activation) and the 100 µl of 1 mg/ml Tec SH3 domain injected on the chip (coupling SH3). All residual binding sites were quenched with ethanolamine (blocking). Immobilisation was conducted at a flow rate of 5 µl/minute. This resulted in approximately 1500 RU SH3 domain immobilised on the surface of the chip.

C. Sensorgram highlighting the response from injecting 50 mM purified GST over a Tec SH3 domain coupled chip. Background has been subtracted from this sensorgram. 30 µl of 50 µM GST was injected at a flow rate of 10 µl/minute. Dissociation was initiated by replacement of GST protein with the running buffer alone.

D. Sensorgram highlighting the response from injecting 8.8 µM purified SH3 domain over a Tec SH3 domain coupled chip. Background has been subtracted from this sensorgram. 20 µl of 8.8 µM SH3 domain was injected at a flow rate of 10 µl/minute and dissociation was initiated by replacement of SH3 domain protein with the running buffer alone.


A
RU
30000
28000
26000
24000
22000
20000
18000
16000
14000
12000
10000
0
300
600
900
1200
1500
1800
2100
2400
2700
3000
seconds
B
Activation
Coupling Tec SH3
Blocking
RU
50000
45000
40000
35000
30000
25000
20000
15000
10000
0
500
1000
1500
2000
2500
3000
3500
4000
4500
seconds
C
RU
40
30
20
10
0
-10
-20
-30
-40
0
200
400
600
800
1000
1200
seconds
D
RU
10
8
6
4
2
0
-2
-4
-6
-8
-10
-100
0
100
200
300
400
500
600
700
800
seconds

Preliminary binding experiments indicated that the interaction between the Tec PRRSH3 and the Tec SH3 domain was characterised by very fast association and dissociation rates as well as low RU signal in this system. Therefore, the level of 1500 RU of coupled ligand was chosen in order to obtain sufficient sensitivity (Panayotou *et al.*, 1993). GST has been shown to dimerise in the presence of reduced glutathione and since these proteins were purified as GST fusions, the effect of trace amounts of GST was investigated. Injection of 50 μM GST over the immobilised SH3 domain surface gave a response of lower than 5 RU (Figure 5.14C). Likewise, injection of Tec SH3 domain also showed a response of 5 RU when passed over the immobilised SH3 domain (Figure 5.14D). Thus, neither the GST nor the SH3 domain alone bind to the immobilised SH3 domain.

Mass transport, a phenomenon where the analyte 'hops' along the surface from one ligand molecule to another due to excess ligand bound on the chip surface can have significant influence on BIAcore experiments and must therefore be minimised. To ensure that the protein-protein interaction is not limited by mass transport effects a series of injections were conducted at different flow rates. 10 μl or 20 μl of 10 mM PRRSH3 was injected over the chip at three different flow rates, 5, 10 and 20 μl/minute. The association reaction is mass transport limited if the initial gradient of association differs with the different flow rates, however, if mass transport effects are not observed the initial association rates will be unaffected. A plot of the three different flow rates overlaid highlighted that the initial rates of association were essentially the same irrespective of the flow rate (Figure 5.15A). Thus, the PRR-SH3 interaction at 10 μl/minute was not mass transport limited and this flow rate was chosen as the flow rate for the following experiments.

Chip surface regeneration removes analyte bound to the ligand but does not remove or adversely affect the ligand coupled to the chip. Regeneration allows multiple injections on the same surface while ensuring constant ligand amounts to bind to the analyte. Initial experiments showed that the dissociation of Tec PRRSH3 proteins from the immobilised SH3 domain was through a rapid dissociation reaction, however, a small amount of the PRRSH3 protein remained bound to the immobilised SH3 domain. A regeneration protocol for the Tec SH3 domain chip was investigated. Tec SH3 domain is not very stable under acidic conditions and, thus, it was unlikely the protein would tolerate regeneration using an acid-based system. Regeneration by washing in NaCl, NaOH or urea was tested. Pulses with NaOH or urea irreversibly unfolded the SH3 domain and the surface no longer bound the PRRSH3 protein. NaCl washing failed to remove the PRRSH3 protein bound to the SH3

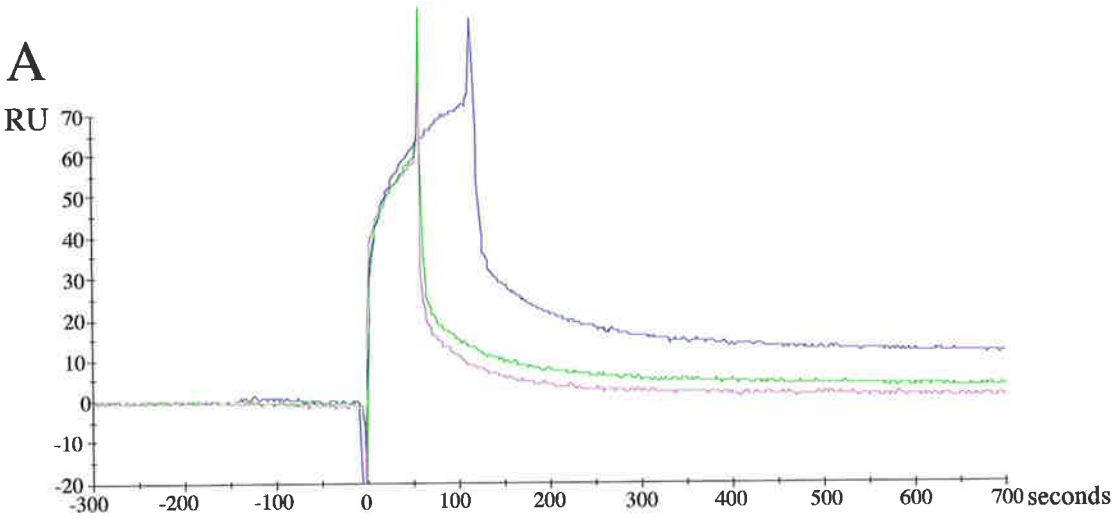
Figure 5.15

A. Sensorgrams of 10 μM PRRSH3 protein injected over a BIAcore chip coupled to 1500 RU Tec SH3 domain. 10 μl or 20 μl of 10 μM PRRSH3 was injected over the chip at three different flow rates, 5 μl/minute (purple), 10 μl/minute (green) and 20 μl/minute (pink). The initial slope of the curve was the same with all three flow rates indicating that at these concentrations the interaction was not mass transport limited.

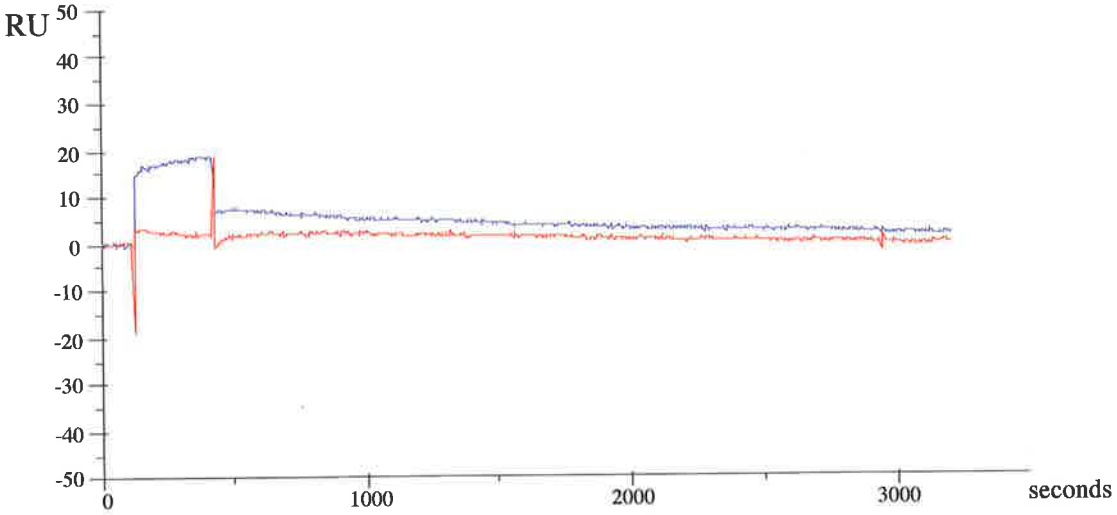
B. Sensorgram of a regeneration protocol derived from Manninen *et al.*, (1998). The regeneration used three 5 μl pulses of glycine pH 2.2, 0.05% SDS and 4 M urea to regenerate the SH3 domain on the chip. Extensive washing with HBS followed. 50 μM PRRSH3 protein was injected over the surface to determine the binding efficiency. The sensorgram in blue is the response prior to regeneration and the sensorgram in red is the response following regeneration.

C. Sensorgrams of the wash regeneration procedure used for all experiments that follow. Buffer was flowed over the surface at 30 μl/minute for 45 minutes following the dissociation phase. The sensorgram represented in navy blue represents the response of 50 μl of 50 μM wild-type PRRSH3 protein injected at a flow rate of 10 μl/minute. The sensorgram in red is the response following a series of 12 injections over the lane on the chip. The sensorgram shown in light blue is the binding observed following two series of 12 injections on the chip.

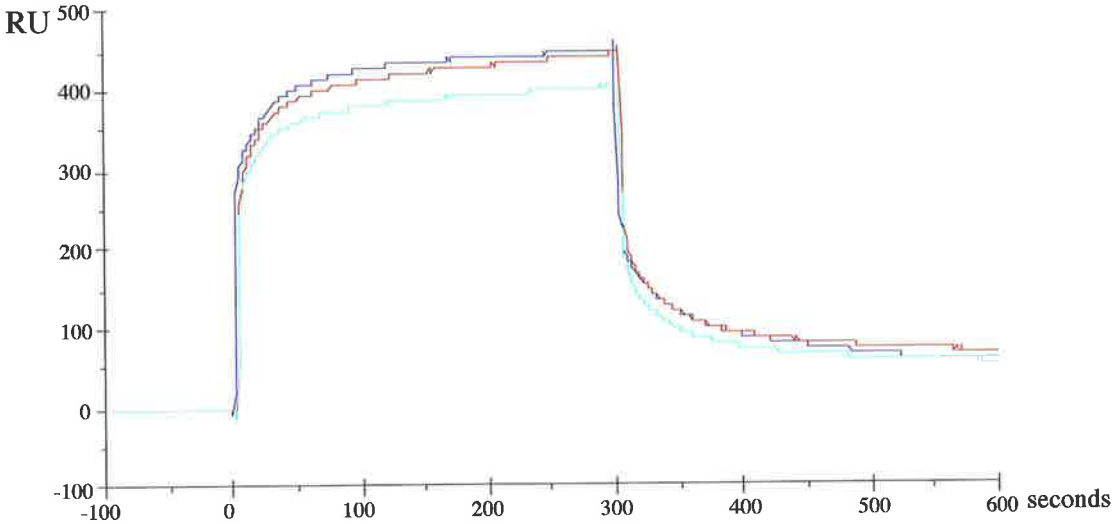
A


RU
70
60
50
40
30
20
10
0
-10
-20
-300
-200
-100
0
100
200
300
400
500
600
700
seconds

B


RU
50
40
30
20
10
0
-10
-20
-30
-40
-50
0
1000
2000
3000
seconds

C


RU
500
400
300
200
100
0
-100
-100
0
100
200
300
400
500
600
seconds

domain. Manninen and co-workers published a regeneration procedure for Hck SH3 domain that used 5 µl (at a flow rate of 5 µl/minute) pulses of glycine buffer pH 2.2, 0.05% SDS and 4 M urea (Manninen *et al.*, 1998). This method was investigated for the regeneration of the Tec SH3 domain chip but also resulted in degradation of the SH3 domain on the chip surface (Figure 5.15B). Regeneration of Hck SH3 domain was performed by a prolonged wash step following injection (Lee *et al.*, 1995). This was investigated for Tec SH3 domain surface regeneration. Prolonged washing of the Tec SH3 surface between injections meant that over the period of a single kinetic experiment (consisting of a series of 12 injections) the SH3 coated chip bound the PRRSH3 protein consistently (Figure 5.15C). Performing more than one kinetic experiment per SH3 domain surface lowered the binding capabilities of the SH3 domain on the chip. Thus, for each kinetic experiment a new chip surface was generated.

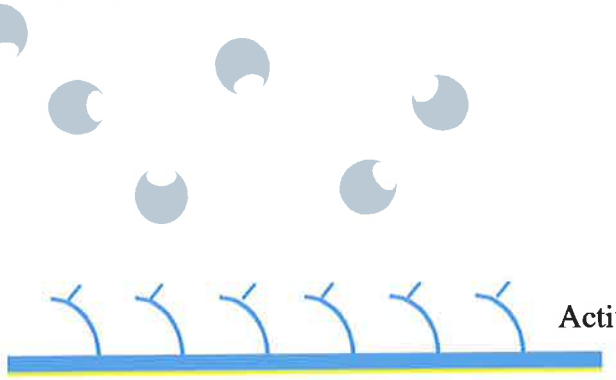
Six lanes were immobilised with 1000-1500 RU of Tec SH3 domain and a series of 12 injections (six different dilutions) conducted for each protein on one lane per Tec PRRSH3 protein. Experiments for PRRSH3Δ1 and PRRSH3Δ12 were conducted on two surfaces with approximately 3500 RU SH3 domain immobilised. Six different concentrations of the PRRSH3 protein, both as a GST fusion and as an isolated species, were injected over a single lane per protein. To minimise the buffer effects on the sensorgram, these proteins were buffer exchanged into HBS using PD10 columns, size exclusion chromatography or, if the sample was concentrated enough, it was diluted to the appropriate concentration with HBS (Pharmacia). Following injections, the surface was washed for 45 minutes at 30 µl/minute with HBS to return the signal to baseline. The injections were conducted randomly and in duplicate to ensure there were no additive errors from an increase in the protein concentration. Over the course of the experiment, duplicate injections of the same protein concentration produced very similar profiles, indicating that available binding sites on the chip did not reduce significantly between injections. Figure 5.16 shows a representation of the BIAcore experiments conducted on Tec PRRSH3 proteins. Step 3 highlights that there are multiple interactions taking place in solution, but only direct binding to the surface is monitored. The interactions occurring in solution are not detected by BIAcore analysis.

Six different concentrations of the mutant and wild-type PRRSH3 proteins with or without the GST fusion partner, ranging from concentrations of 1 µM to 100 µM were injected over the chip surface. Ultracentrifugation data has shown that at concentrations lower than 100 µM all the PRRSH3 proteins are predominantly present in the monomeric

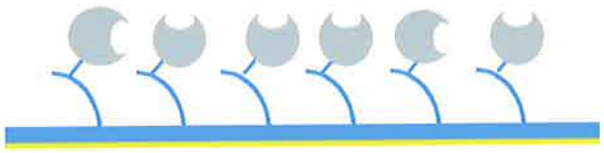
Figure 5.16

Schematic representation of the BIAcore experiments undertaken in this study. The BIAcore surface is shown in light blue, the SH3 domain as grey circles with the SH3 binding surface represented by the indent. The PRRSH3 proteins are shown at step 3. There is only one PRR site shown in pink to simplify the diagram. Step 1 is the activation of the surface followed by Step 2 where the SH3 domain is covalently attached to the chip surface. Step 3 shows the complex interactions taking place in these BIAcore experiments. The PRRSH3 proteins can interact intramolecularly as well as with the SH3 domain on the chip. Only those interactions of PRRSH3 domain with the SH3 domain on the chip will be detected by the instrument. The intramolecular interactions will not be detected.

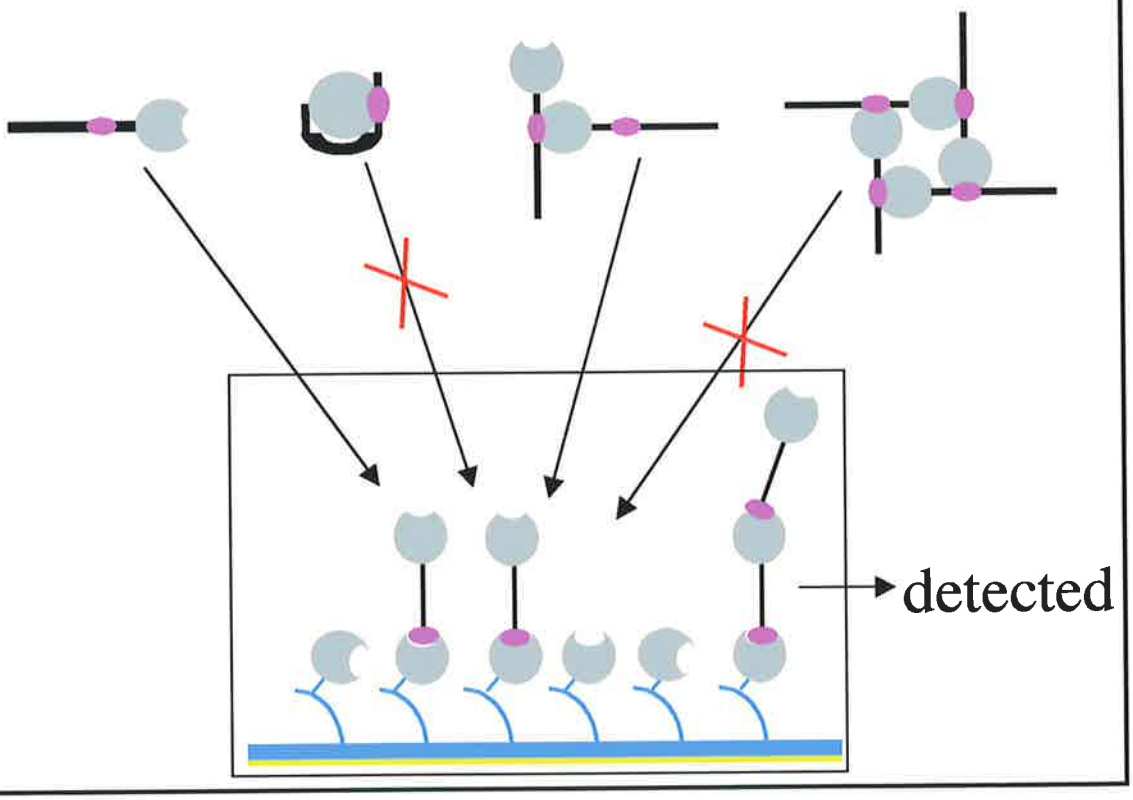
Step 1: Coupling Tec SH3 domain through amide coupling



Step 2: Tec SH3 Bound to the chip surface



Step 3: Possible interaction of PRRSH3 proteins in BIAcore Experiments



form. Comparing Figure 5.17 with Figure 5.18, clear differences were observed in the total RU bound at equilibrium for the GST-PRRSH3Δ2 and GST-PRRSH3Δ12 compared with GST-PRRSH3 and GST-PRRSH3Δ1. All sensorgrams showed a rapid association and dissociation interaction. Approximately 90% of the protein dissociated very fast from the surface whereas the remaining 10% dissociated slowly, suggesting a biphasic association and dissociation mechanism. The GST-PRRSH3 wild-type protein at a concentration of 100 μM produced binding of approximately 860 RU, while at the same concentration the GST-PRRSH3Δ1 mutant produced a response of 820 RU (Figure 5.17A, C), the GST-PRRSH3Δ2 and GST-PRRSH3Δ12 of 161 and 250 RU, respectively (Figure 5.18A, C). Thus, GST-PRRSH3Δ1 produced a response of 95% that of the wild-type GST-PRRSH3 indicating that the PRR site 2 can bind intermolecularly to the Tec SH3 domain on the chip. The response seen for GST-PRRSH3Δ2 was about 20% the response of wild-type PRRSH3 indicating that PRR site 1 does not interact significantly with the SH3 domain on the chip. PRRSH3Δ12 bound the SH3 domain on the chip to a greater extent than PRRSH3Δ2 possibly because of non-specific interactions.

Purified PRRSH3 proteins cleaved from the GST fusion partner were also tested for binding to the SH3-coupled surface. The PRRSH3 proteins were passed over the surface at a slightly lower concentration compared with the GST-fusion proteins, as the isolated proteins did not concentrate to 100 μM without precipitation. Isolated Tec PRRSH3 proteins exhibited decreased stability compared with GST fusions. Although a direct comparison of equilibrium RU reached between the isolated protein and the GST fusion protein is not useful due to differences in the mass of the analyte, the binding pattern between the PRRSH3 proteins is similar. Running wild-type PRRSH3 over different lanes with 1500 units SH3 domain immobilised gave an equilibrium response of approximately 300 RU at a concentration of 85 μM compared with PRRSH3Δ2 which gave a response of approximately 20 RU at 25 μM (Figure 5.19 A and C). The PRRSH3Δ2 protein therefore has less than 30% the binding capability of the wild-type PRRSH3 protein. PRRSH3Δ1 and PRRSH3Δ12 experiments had approximately twice the SH3 domain immobilised to the chip (3000 RU) and, thus, have two times more SH3 domain available to bind the PRRSH3 protein so that comparisons of equilibrium RU between these and the wild-type and PRRSH3Δ2 protein cannot be done. However, a comparison between PRRSH3Δ1 and PRRSH3Δ12 shows that at equilibrium PRRSH3Δ1 has a binding of 600 RU (at 45 μM) to the SH3 domain on the chip surface and PRRSH3Δ12 has 200 RU equilibrium binding at a

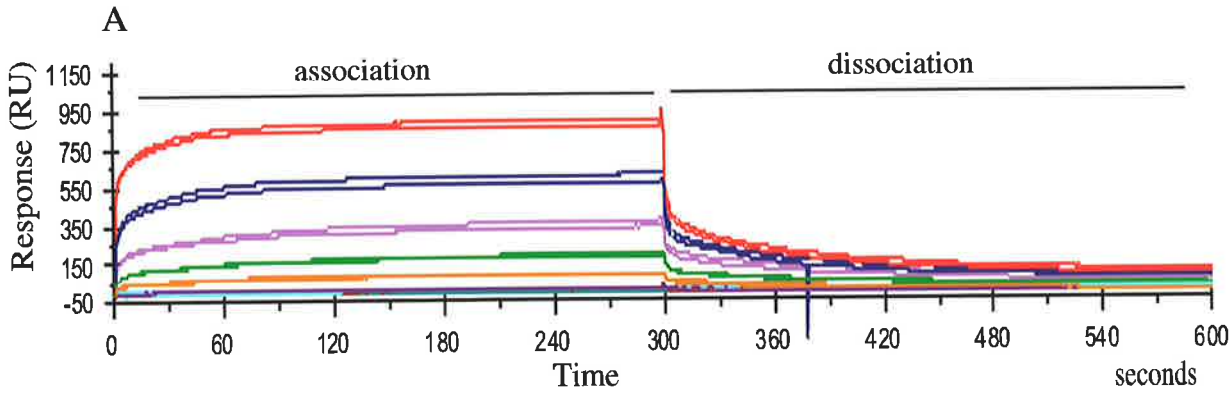
Figure 5.17

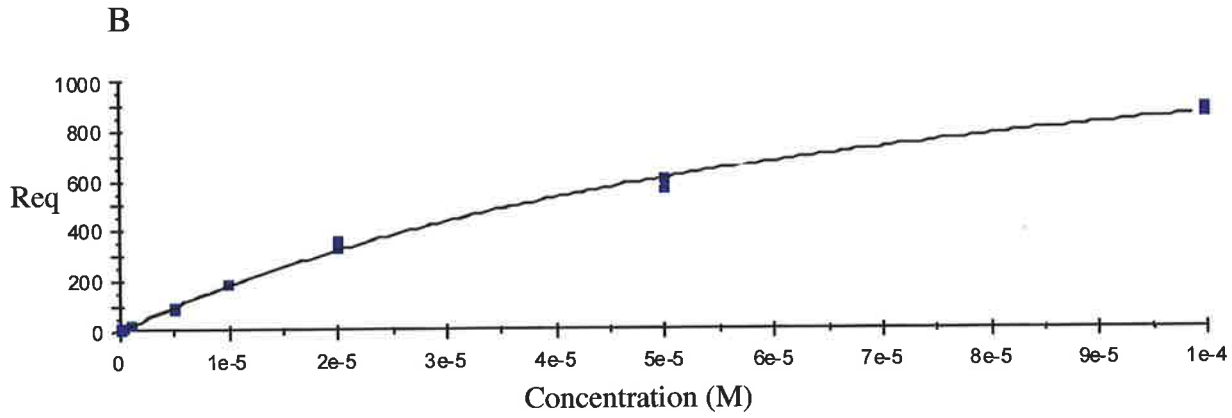
A. Sensorgrams for different concentrations of wild-type GST-PRRSH3 protein binding 1500 RU Tec SH3 domain on the chip surface. Data was collected on a BIAcore 2000 machine. The GST-PRRSH3 protein samples were buffer exchanged into HBS to minimise buffer spikes observed during association and dissociation phases. 50 µl Samples were passed over the surface at 10 µl/minute for the duration of the association phase and dissociation was initiated by replacing the sample flow with HBS buffer. The surface was washed at 10 µl/minute for 1600 seconds followed by a further wash at 30 µl/minute for 2500 seconds to ensure a return to baseline values. The sensorgram in the control lane was subtracted from these sensorgrams and they were overlayed. Six different concentrations of GST-PRRSH3Δ2 protein (100 µM, red; 50 µM, blue; 20 µM, pink; 10 µM, green; 5 µM, orange; 1 µM, purple; 0.5 µM, light blue; 0.1 µM, green were injected over the surface in duplicate (represented in the same colour).

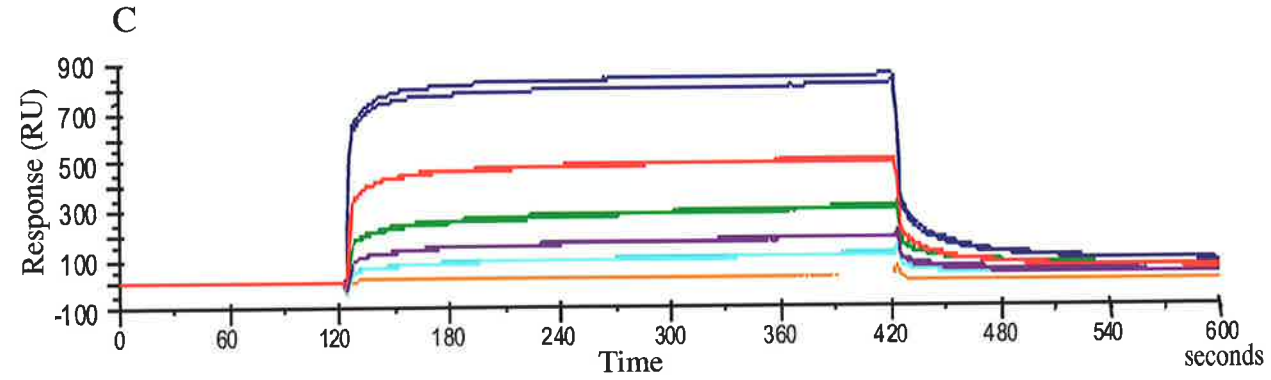
B. Kinetic analysis of the interaction between GST-PRRSH3 (wild-type) protein and the SH3 domain. Steady state equilibrium analysis was undertaken at the region of the sensorgrams just prior to dissociation using the BIAevaluation software package. A region of approximately 50 seconds at 240-300 seconds was chosen as the region where equilibrium was reached. A plot of response at equilibrium (Req) versus concentration was generated and equilibrium dissociation constants determined.

C. The sensorgrams of GST-PRRSH3Δ1 injected over a chip surface with 1500 RU Tec SH3 domain immobilised on the chip. The experiments were completed as described in A. The concentrations of PRRSH3Δ12 injected over the surface were 100 µM (blue), 50 µM (pink), 20 µM (teal), 10 µM (purple), 5 µM (light blue) and 1 µM (orange).

D. Kinetic analysis of the GST-PRRSH3Δ1 SH3 domain interaction. Constants were derived as described in B.


A
association
dissociation
Response (RU)
1150
950
750
550
350
150
-50
0
60
120
180
240
300
360
420
480
540
600
Time
seconds


B
Req
1000
800
600
400
200
0
0
1e-5
2e-5
3e-5
4e-5
5e-5
6e-5
7e-5
8e-5
9e-5
1e-4
Concentration (M)


C
Response (RU)
900
700
500
300
100
-100
0
60
120
180
240
300
360
420
480
540
600
Time
seconds

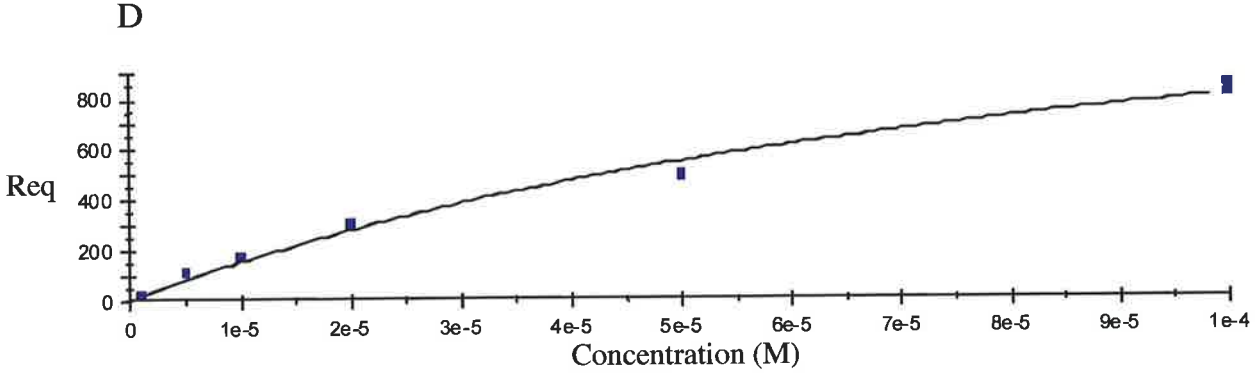

D
Req
800
600
400
200
0
0
1e-5
2e-5
3e-5
4e-5
5e-5
6e-5
7e-5
8e-5
9e-5
1e-4
Concentration (M)

Figure 5.18

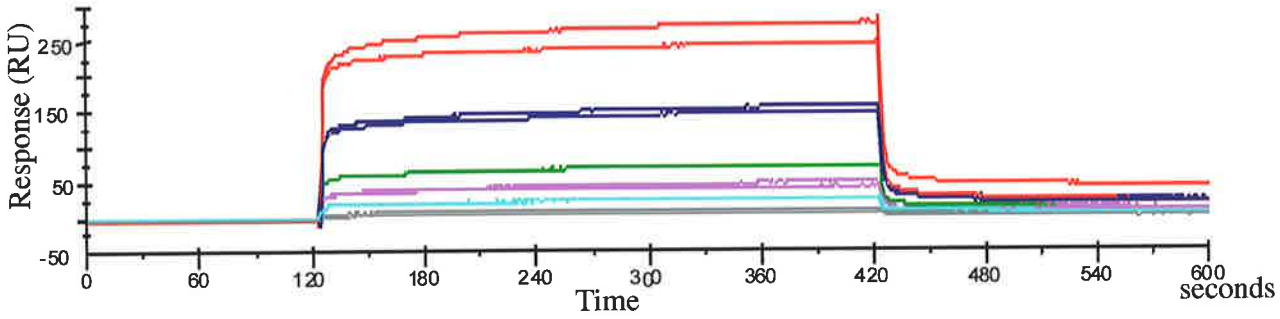
A. Sensorgrams for different concentrations of GST-PRRSH3Δ2 binding to 1500 RU Tec SH3 domain on the surface. The experiments were completed as described in Figure 5.17A. Six different concentrations of GST PRRSH3Δ2 protein (100 μM, red; 50 μM, blue; 20 μM, green; 10 μM, pink; 5 μM, teal; 1 μM, grey) were injected over the surface in duplicate (presented in the same colour). Sensorgrams with air spikes were removed.

B. Equilibrium dissociation constants for the GST-PRRSH3Δ2 SH3 domain interaction. Constants were derived as described in Figure 5.17B with the exception that approximately 50 seconds at 350-400 seconds was chosen as the region where equilibrium was reached.

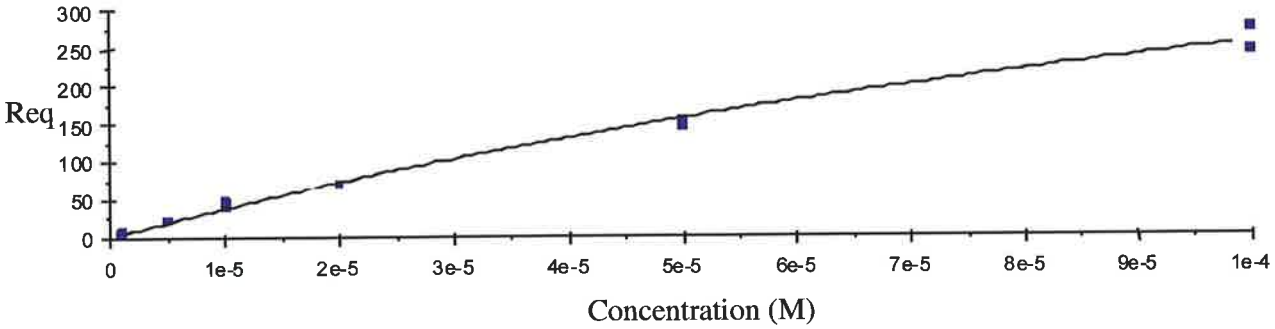
C. Sensorgrams of GST-PRRSH3Δ12 injected over a chip immobilised with 1500 RU Tec SH3 domain. The experiments were completed as described in Figure 5.17A. The concentrations of PRRSH3Δ12 injected over the surface were 100 μM (teal), 50 μM (red), 20 μM (green), 10 μM (orange), 5 μM (pink) and 1 μM (blue). Duplicate sensorgrams were removed in the 100 μM, 50 μM and 1 μM concentrations due to air spikes. Sensorgrams of the same concentration are presented in the same colour.

D. Equilibrium dissociation constants for the GST-PRRSH3Δ12 SH3 domain interaction. Constants were derived as described in Figure 5.17B with the exception that approximately 50 seconds at 350-400 seconds was chosen as the region where equilibrium was reached.

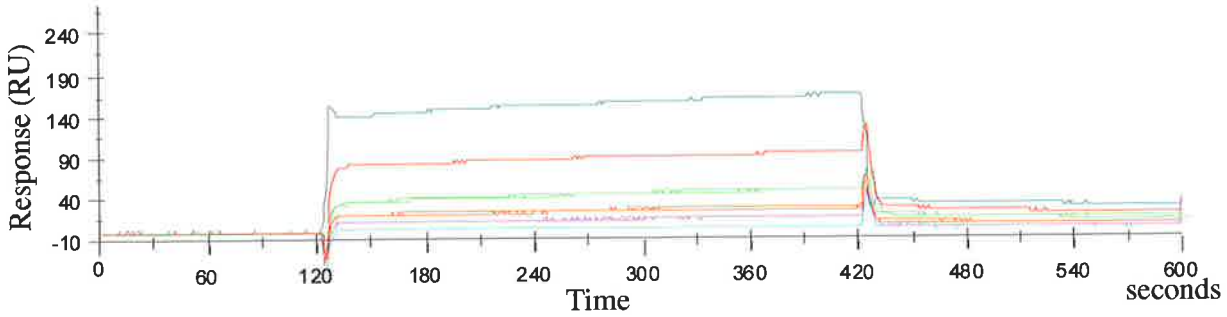
A


Response (RU)
250
150
50
-50
0
60
120
180
240
300
360
420
480
540
600
Time
seconds

B


Req
300
250
200
150
100
50
0
1e-5
2e-5
3e-5
4e-5
5e-5
6e-5
7e-5
8e-5
9e-5
1e-4
Concentration (M)

C


Response (RU)
240
190
140
90
40
-10
0
60
120
180
240
300
360
420
480
540
600
Time
seconds

D

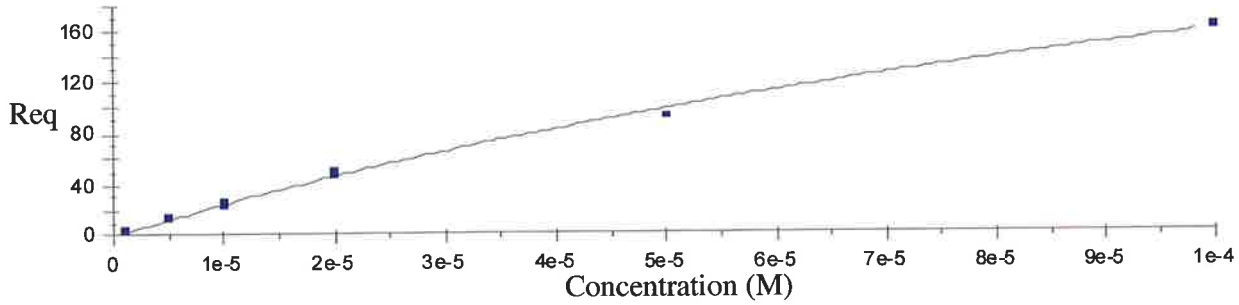

Req
160
120
80
40
0
1e-5
2e-5
3e-5
4e-5
5e-5
6e-5
7e-5
8e-5
9e-5
1e-4
Concentration (M)

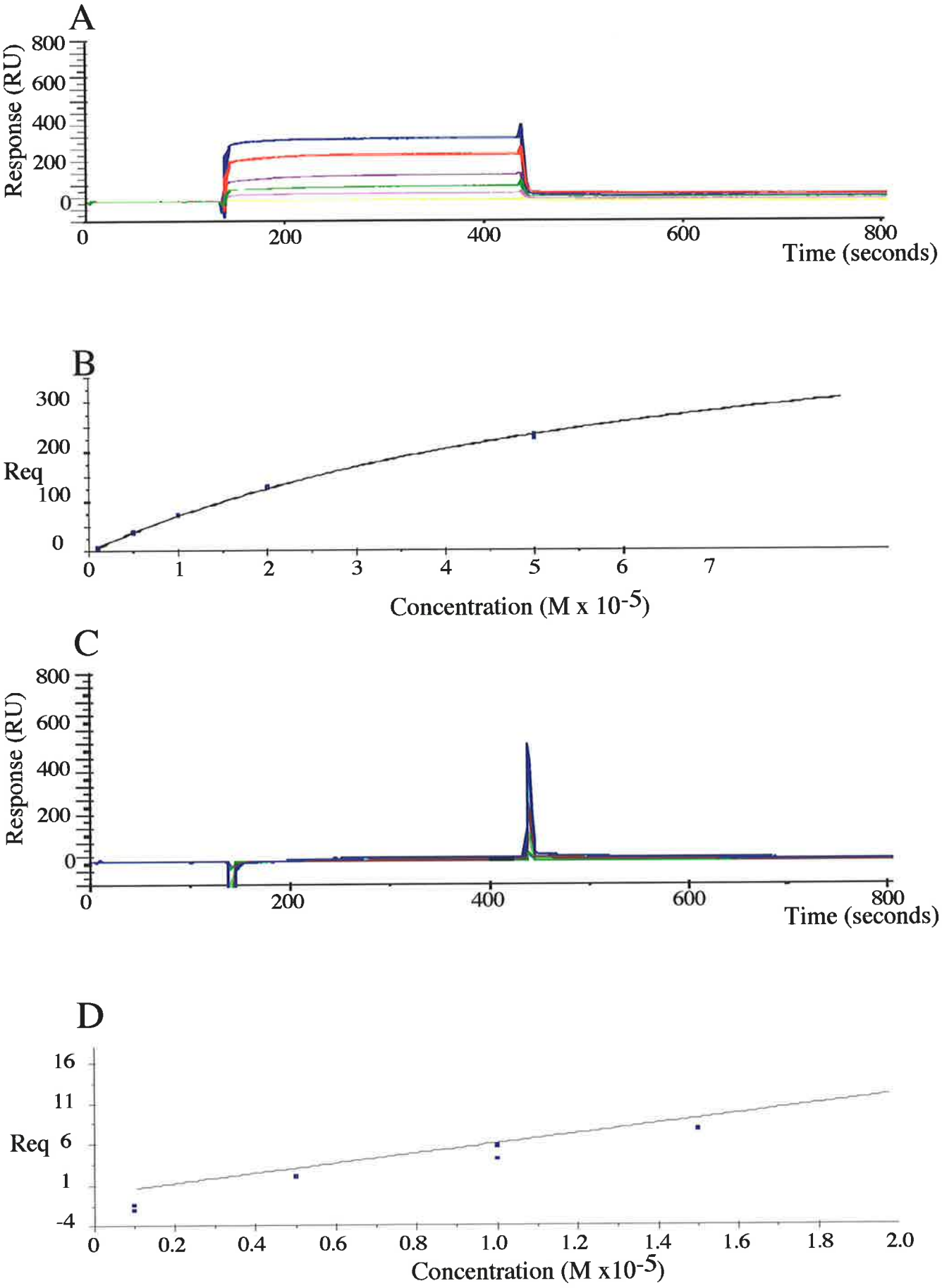
Figure 5.19

A. Sensorgrams for different concentrations of wild-type PRRSH3 binding to Tec 1500 RU SH3 domain on the surface. The experiments were completed as described in Figure 5.17A. Six different concentrations of wild-type PRRSH3 protein (85 μM, blue; 50 μM, pink; 20 μM, purple; 10 μM, teal; 5 μM, pink; 1 μM, yellow) were injected over the surface in duplicate (represented in the same colour). Low solubility at high concentrations limited the concentrations tested.

B. Equilibrium dissociation constants for the wild-type PRRSH3 SH3 domain interaction. Constants were derived as described in Figure 5.17B with the exception that approximately 50 seconds at 350-400 seconds was chosen as the region where equilibrium was reached.

C. Sensorgrams for different concentrations of PRRSH3Δ2 binding to 1500 RU Tec SH3 domain on the surface. The experiments were completed as described in Figure 5.17A. Six different concentrations of PRRSH3Δ2 protein (25μM,blue; 20μM, light blue; 15μM, brown; 10μM, light green; 5μM, dark green; 1μM, purple) were injected over the surface in duplicate (represented in the same colour). Low solubility at high concentrations limited the concentrations tested.

D. Equilibrium dissociation constants for the PRRSH3Δ2 SH3 domain interaction. Constants were derived as described in Figure 5.17B with the exception that approximately 50 seconds at 350-400 seconds was chosen as the region where equilibrium was reached.


A
Response (RU)
800
600
400
200
0
0
200
400
600
800
Time (seconds)
B
300
200
Req
100
0
0
1
2
3
4
5
6
7
Concentration (M x 10-5)
C
Response (RU)
800
600
400
200
0
0
200
400
600
800
Time (seconds)
D
16
11
Req 6
1
-4
0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
Concentration (M x10-5)

concentration of 60μM (Figure 5.20A, C). The fast association/fast dissociation observed for all PRRSH3 proteins binding to the SH3 domain suggests a weak interaction between the Tec SH3 domain and the Tec PRRSH3 proteins. Sensorgrams for the PRRSH3 proteins are consistent with those for GST fusion proteins. PRRSH3 wild-type and PRRSH3Δ1 bind the immobilised SH3 domain at similar levels and PRRSH3Δ2 and PRRSH3Δ12 bind the SH3 domain at similar levels. The relative responses for the different PRRSH3 proteins are maintained with different immobilisation levels of the Tec SH3 domain.

Numerous potential interactions that may occur between the PRRSH3 protein and the SH3 domain in solution necessitate caution interpreting kinetic constants derived from this BIAcore data. There is no BIAcore based method to detect the interactions occurring in the solution, and the association with the SH3 domain surface is the only interaction detected by the instrument. The rapid association/dissociation observed is not amenable for determination of individual on and off rates; thus, steady state equilibrium kinetic analysis is the most appropriate method of analysis. Steady-state equilibrium analysis has the added advantage that it is independent of the amount of ligand immobilised on the chip. The kinetic analysis was conducted using the BIAevaluation program as per the manufacturers instructions. A section of approximately 50 seconds prior to dissociation was chosen for equilibrium analysis. Equilibrium dissociation constants (K_D) were derived from plots of response units at equilibrium (Req) versus molar concentration using at least eight values (Figure 5.17-5.20 B and D for the respective analysis of the different PRRSH3 proteins). The equilibrium association and dissociation constants derived for the different PRRSH3 constructs are displayed in Table 5.2. The χ^2 value represents the mean square of the signal deviation from the fitted curve and is a reflection of the "goodness of fit" of the curve to the data points. Ideally, values for χ^2 of below 10 are required (BIAcore manual). While the χ^2 values for these curves fitted for PRRSH3 proteins binding to the SH3 domain are frequently higher than 10, this is using a simple 1:1 interaction model which may not represent the true interaction given that an intramolecular interaction can also occur. High χ^2 values may reflect the complexity of the interactions and suggests a simple dimeric association does not ideally fit the data consistent with tetramerisation already observed. The equilibrium constants derived will reflect the 1:1 intermolecular interaction of the Tec SH3 domain and the Tec PRR and will give an indication of the relative affinities for this interaction.

The equilibrium dissociation constants derived for wild-type PRRSH3 and PRRSH3Δ1 proteins were similar, both as isolated proteins and as GST fusion proteins. The association was slightly stronger for the wild-type GST-PRRSH3 protein with a K_D of 70.6

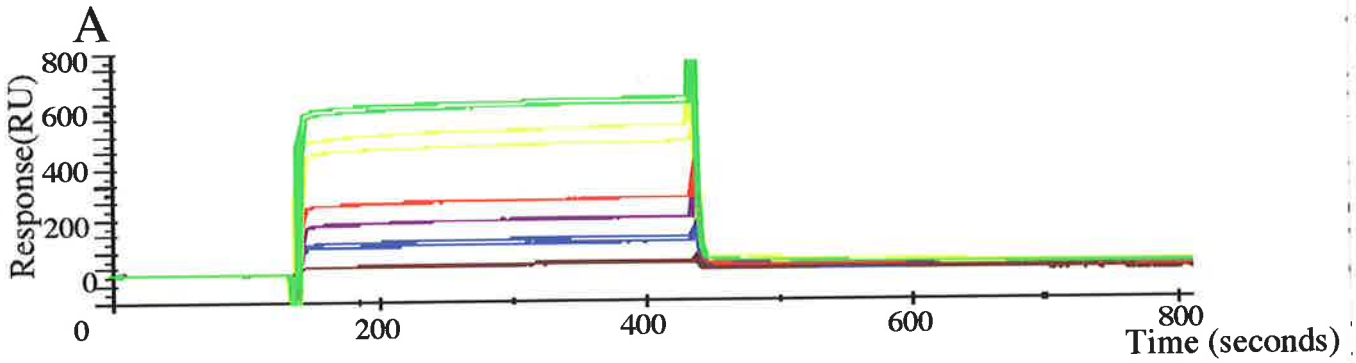
Figure 5.20

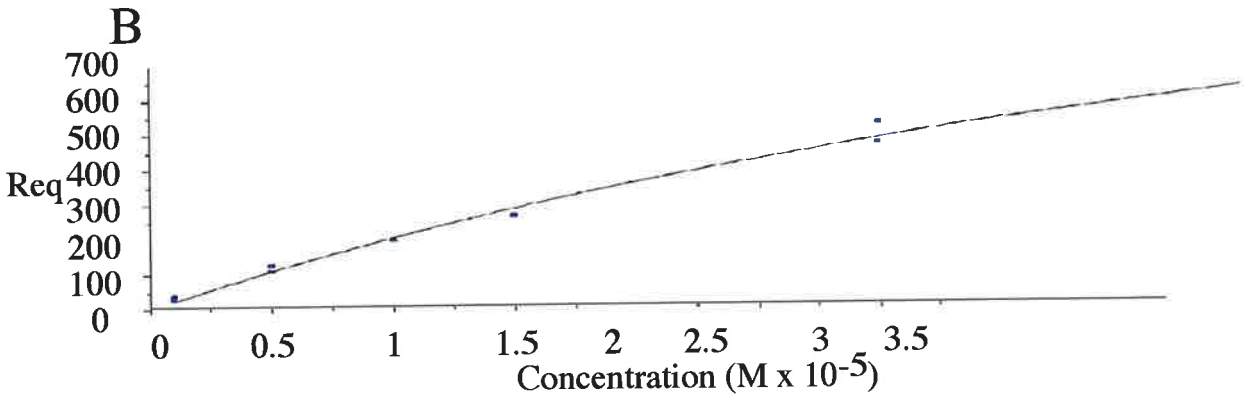
A. Sensorgrams for different concentrations of PRRSH3Δ1 binding to 3000 RU Tec SH3 domain on the surface. The experiments were completed as described in Figure 5.17A. Six different concentrations of PRRSH3Δ1 protein (45 μM, green; 30 μM, yellow; 15 μM, pink; 10 μM, purple; 5 μM, blue; 1 μM, dark red) were injected over the surface in duplicate (represented in the same colour). Low solubility at high concentrations limited the concentrations tested.

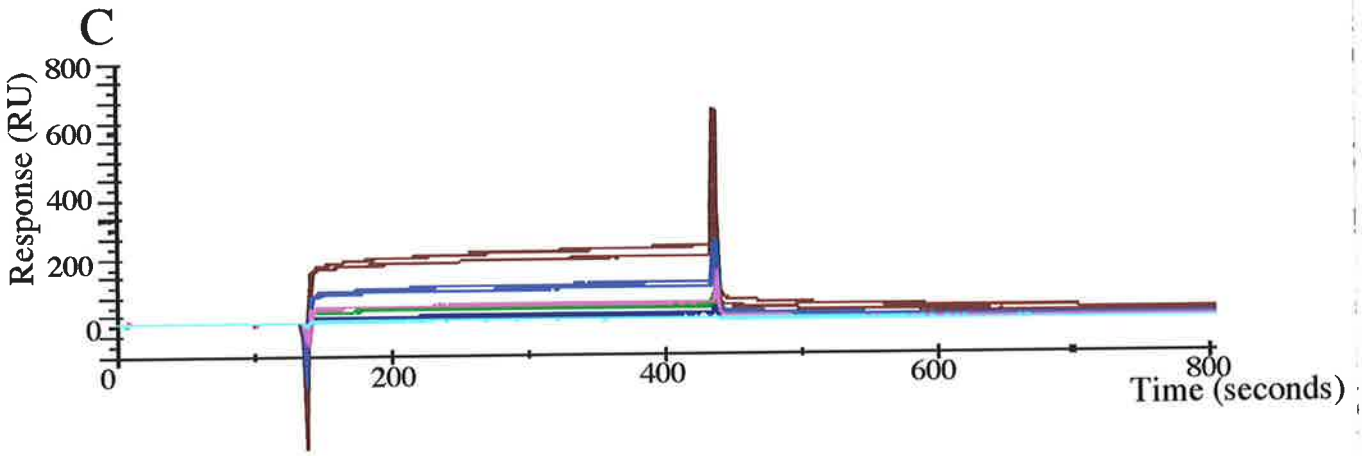
B. Equilibrium dissociation constants for the PRRSH3Δ1 SH3 domain interaction. Constants were derived as described in Figure 5.17B with the exception that approximately 50 seconds at 350-400 seconds was chosen as the region where equilibrium was reached.

C. The sensorgrams of PRRSH3Δ12 injected over a chip surface containing 3000 RU Tec SH3 domain. The experiments were completed as described in Figure 5.17A. Concentrations of PRRSH3Δ12 protein (60 μM, purple; 30 μM, dark blue; 15 μM, pink; 10 μM, light blue; 5 μM, blue; 1 μM, aqua) were injected over the chip surface in duplicate (presented in the same colour).

D. Equilibrium dissociation constants for the PRRSH3Δ12 SH3 domain interaction. Constants were derived as described in Figure 5.17B with the exception that approximately 50 seconds at 350-400 seconds was chosen as the region where equilibrium was reached.


A
Response(RU)
800
600
400
200
0
0
200
400
600
800
Time (seconds)


B
Req
700
600
500
400
300
200
100
0
0
0.5
1
1.5
2
2.5
3
3.5
Concentration (M x 10-5)


C
Response (RU)
800
600
400
200
0
0
200
400
600
800
Time (seconds)

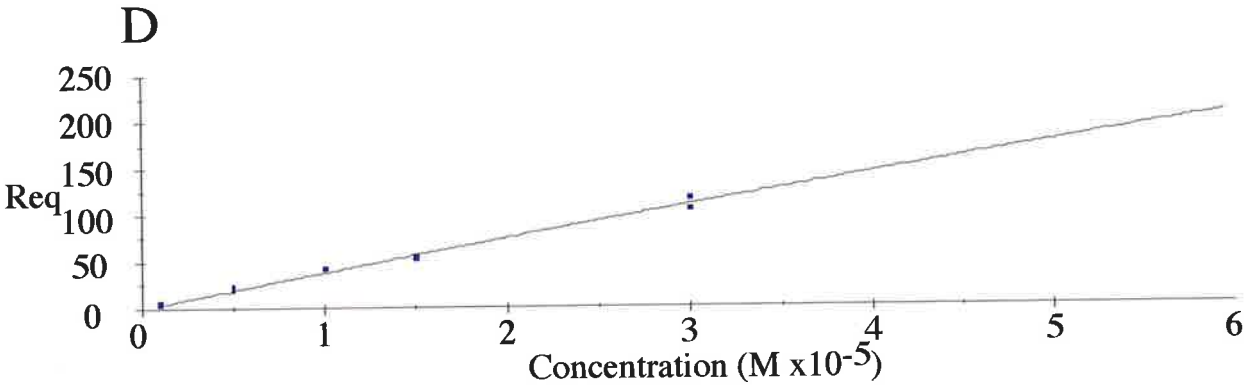

D
Req
250
200
150
100
50
0
0
1
2
3
4
5
6
Concentration (M x10-5)

Table 5.2 Kinetic Constants derived by BIAcore for PRRSH3 proteins

	K_A M^{-1}	K_D μM	R_{max} RU	χ^2
Wildtype				
+GST	14200	70.6	1460	353
-GST	14600	68.3	555	10.7
PRRSH3Δ1				
+GST	11500	86.9	1500	1160
-GST	15500	64.3	1520	372
PRRSH3Δ2				
+GST	6320	158	144	8.09
-GST	442	2200	1380	4.72
PRRSH3Δ12				
+GST	5920	169	684	76.6
-GST	1950	513	2000	57.4

μM compared with GST-PRRSH3Δ1 at 86.9 μM, however, when the GST is removed the dissociation constants are very close, 68.3 μM and 64.3 μM for PRRSH3 wild-type and PRRSH3Δ1, respectively. This data indicates that both these proteins have the potential to interact with the SH3 domain on the solid phase at comparable binding affinities. Affinity constants are similar to those reported for other SH3 PXXP interactions (Lee *et al.*, 1995). Therefore, the binding affinity of wild-type PRRSH3 to the SH3 domain is due solely to the second binding site in the PRR (PRR site 2).

Dissociation constants were derived for the GST-PRRSH3Δ2 and GST-PRRSH3Δ12 interaction with the SH3 domain surface. The constants determined were 158 μM and 169 μM for GST-PRRSH3Δ2 and GST-PRRSH3Δ12, respectively. These values indicate a 2-fold reduction in affinity for the SH3 domain on the chip compared with the wild-type GST-PRRSH3 and GST-PRRSH3Δ1 proteins. The affinity was lower for these proteins when they lacked the GST fusion partner at dissociation constants of 2200 μM and 513 μM for PRRSH3Δ2 and PRRSH3Δ12, respectively. The signal to noise for the PRRSH3Δ2 protein binding to the SH3 domain by BIAcore analysis was very low and the affinity constant may therefore be affected by the low response. This data suggests that Tec PRRSH3Δ2 protein has reduced binding to the Tec SH3 domain on the chip. The PRRSH3Δ12 dissociation constant also differed from the affinity constant determined for the GST-PRRSH3Δ12. It is possible that additional interactions between GST molecules are occurring. These may be preformed dimers and, thus, the mass binding to the chip is greater than expected or when the PRRSH3Δ12 binds the chip the GST may form a secondary dimerisation surface. This secondary interaction was only significant when the PRRSH3 proteins bound with low affinity.

The BIAcore assay can only detect the association of the PRRSH3 proteins with the SH3 domain surface. The analysis of the dimer interaction between the SH3 domain surface and the PRRSH3 protein confirms that PRR site 2 is responsible for the majority of the binding to the SH3 domain. In contrast, the decreased binding of PRRSH3Δ2 and PRRSH3Δ12 to the SH3 domain surface indicates that PRR site 1 has reduced intermolecular affinity for the SH3 domain. Thus, the presence of the PRR site 1 decreases the affinity of the dimerisation reaction, perhaps by intramolecular competition. It is possible that PRR site 1 is already bound to the SH3 domain in an intramolecular manner and, thus, is inaccessible for binding the SH3 domain surface.

Another method of investigating the affinity of binding between the SH3 domain and the different sites within the proline rich region is to synthesise the peptides and again conduct BIAcore analysis. A peptide corresponding to PRR site 1 ^{152}SIRKTLPPAPEI^{163} was produced to measure interactions involving this region in isolation. A peptide corresponding to PRR site 2 could not be synthesised, most likely due to technical issues stemming from the four adjacent prolines. Thus, the site 2 peptide could not be further investigated.

Both the PRR site 1 ligand and the linker peptide (section 4.7) were investigated for binding to the SH3 domain surface. The detection limits of the BIAcore system can make detecting small molecules difficult. It was predicted that neither of these peptides would produce a significant response due to their size of M_r 1321 Da and M_r 1606 Da, respectively and, thus, the interaction of site 1 peptide and the linker peptide with the SH3 domain surface was determined using an indirect method. A competition experiment was conducted in which a constant amount of PRRSH3 protein (20 μM) was mixed with increasing amounts of either peptide ligands to a final concentration of 500 μM peptide. A schematic is described in Figure 5.21A highlighting the possible interactions in this experiment. Complex equilibria exist as the small peptide can interact with both the SH3 domain on the surface and the SH3 domain in solution. This BIAcore experiment detects the PRRSH3 protein interaction with the SH3 domain on the chip surface. Increasing concentrations of the low molecular weight ligand will block the available binding surfaces on the immobilised SH3 domain and, thus, reduce the binding of PRRSH3 proteins, which will produce a lower response. The extent of small molecular weight ligand binding the PRRSH3 molecules in solution cannot be determined (Figure 5.21A).

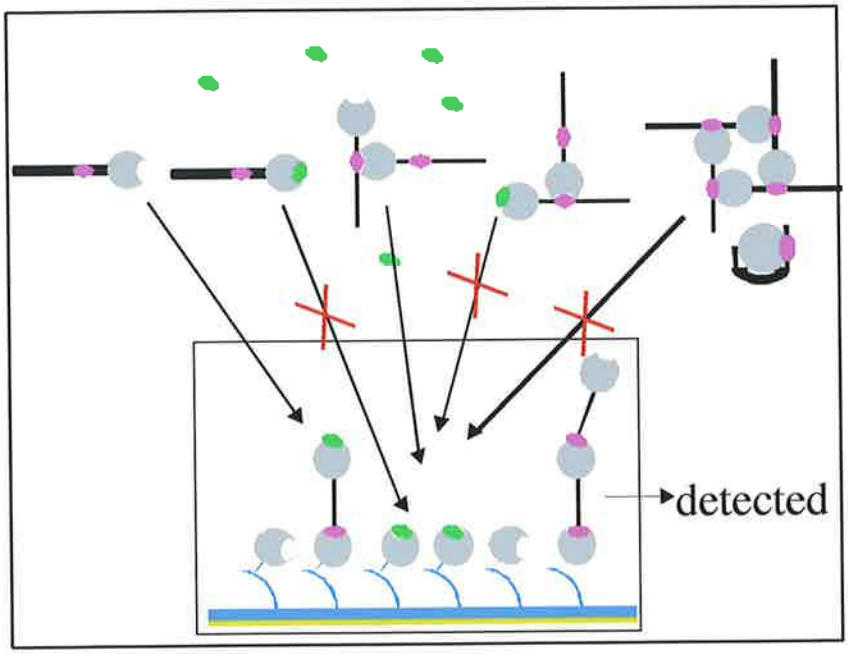
The site 1 peptide and linker peptide were initially passed over the chip at concentrations ranging from 10 μM-500 μM peptide with no PRRSH3 protein present to determine if peptide binding could be measured directly. The site 1 peptide did show a slight response (approx 25 RU) at 500 μM but was insufficient for rigorous analysis to be undertaken (shown in dark blue on the sensorgram). The peptide was mixed with a fixed concentration of PRRSH3 of 20 μM. Figure 5.21B shows the sensorgram obtained for the competition experiments where the PRRSH3 is in competition with the site 1 peptide. When the concentration of site 1 peptide is increased there is a corresponding decrease in the equilibrium response for the PRRSH3 protein indicating that the site 1 peptide and the PRRSH3 domain are competing for the same binding surface on the immobilised SH3 domain. The sensorgrams derived for the linker peptide were the same as observed for the PRR site 1 peptide (data not shown). While this data was not quantitatively assessed,

Figure 5.21

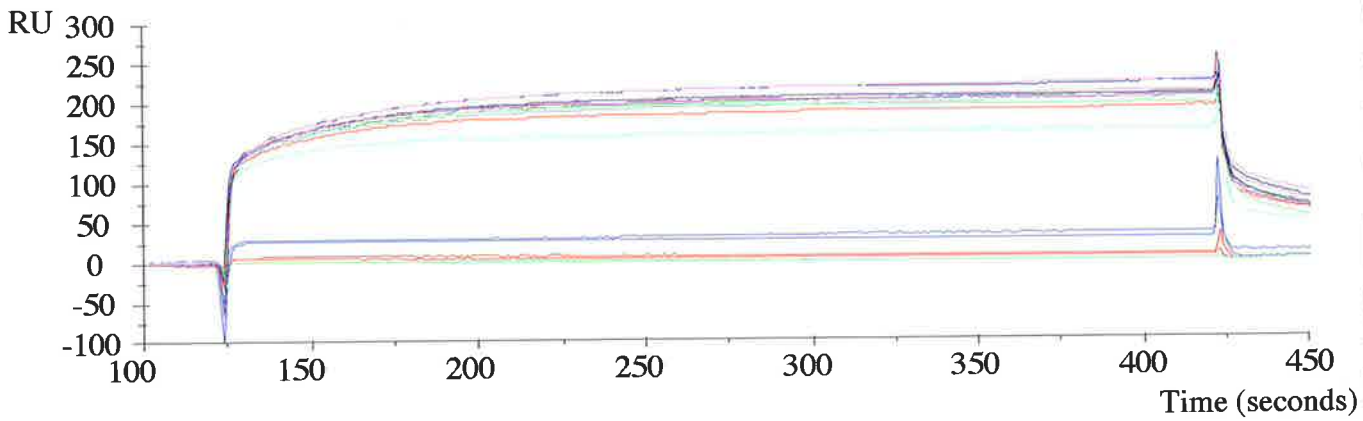
A. Schematic of the BIAcore competition experiments. The colour scheme matches that for Figure 5.16. The SH3 domain is represented in grey and the PRR ligand is shown in pink. Small ligands (green), not detectable directly by BIAcore, are in competition with wild-type PRRSH3 protein in solution for binding to the immobilised SH3 domain surface. With increased amount of small ligand, decreased binding of the larger ligand will be observed.

B. Sensorgrams for different concentrations of site 1 peptide binding to 1500 RU Tec SH3 domain on the chip surface in competition with 20 μM wild-type PRRSH3 protein. Data was collected on a BIAcore 2000 machine. Samples were mixed just prior to the beginning of the experiment. 50 μl Samples were injected on to the surface at 10 μl/minute for the association phase and dissociation was initiated by replacing the sample flow with HBS buffer. The surface was washed with HBS at 10 μl/minute for 2100 seconds to ensure a return to baseline. The sensorgram in the background lane was subtracted from these sensorgrams and they were overlayed. Eight different concentrations of PRR site 1 peptide (0 μM, magenta; 10 μM, blue; 20 μM, brown; 50 μM, dark green; 100 μM, green; 150 μM, grey; 250 μM, dark red; 500 μM, aqua) in competition with 20 μM wild-type PRRSH3 protein were injected over the surface in duplicate (same colour). PRR site 1 peptide alone was also tested at three concentrations (10 μM, green; 100 μM, red; 500 μM, blue). Approximately 25 RU was observed for the 500 μM sample.

A


detected

B


RU
300
250
200
150
100
50
0
-50
-100
100
150
200
250
300
350
400
450
Time (seconds)

competition between the PRRSH3 protein and the site 1 peptide for surface binding provides evidence that the PRRSH3 protein and the peptide ligands are binding the same site on the SH3 domain.

The BIAcore data has confirmed the potential for an intermolecular interaction between the PRR and the adjacent SH3 domain. The mutant protein analysis has shown that PRRSH3Δ1 can bind intermolecularly with the SH3 domain surface with the same affinity as PRRSH3 wild-type. Thus, PRR site 2 can interact with the immobilised SH3 domain. This is consistent with the ultracentrifugation data that indicated that PRRSH3Δ1 formed dimers and tetramers at similar rates to wild-type PRRSH3. The reverse is true for PRRSH3Δ2. This protein could not bind the SH3 domain surface and it did not multimerise by ultracentrifugation to the same extent as the wild-type PRRSH3 protein and PRRSH3Δ1. This suggests that PRR site 1 is not available to either interact with the SH3 domain surface or dimerise with other PRRSH3 molecules. By extension, it is possible that PRR site 1 is facilitating an intramolecular interaction within the PRRSH3 protein.

5.5.3 Chemical shift perturbation analysis of PRRSH3 protein

Changes in chemical shift upon ligand addition can be a powerful method to establish residues of a protein that interact with the ligand. Using ^{15}N-labelled protein samples and HSQC provides a convenient, solution-based, rapid method to follow chemical shift changes. To identify the amino acids participating in the PRR interaction with the SH3 domain seen in the BIAcore experiments, a HSQC titration experiment series was conducted to determine the site of interaction between PRR site 1 peptide and the SH3 domain. The PRR site 1 peptide spans amino acids ^{152}SIRKTLPPAPEI^{163} in the Tec amino acid sequence. A 0.75 mM ^{15}N SH3 domain sample, buffered with 10 mM phosphate pH 6.0 was supplemented with 10% D_2O. A 41.2 mM solution of PRR site 1 peptide was added in increasing molar amounts to the SH3 sample. The pH of the solution was maintained at pH 6.0 throughout the experiment. A series of ^{15}N-^{1}H HSQC experiments were conducted at 25°C with increasing site 1 peptide concentrations to monitor the change in chemical shift of the HSQC cross peaks. Changes in chemical shift indicate a change in the environment of the amino acid indicative of an interaction (Craik and Wilce, 1997). Amino acids involved in the PRR site 1 interaction with Tec SH3 domain can be identified from changes in cross peak chemical shifts.

Spectra were overlaid and Figure 5.22 shows the transition of the HSQC spectra through additions of increasing amounts of ligand. A single set of peaks were observed in the HSQC spectra indicating that the interaction is a fast exchanging interaction and the observed peaks in the HSQC are the weighted average of the chemical shifts of the two individual states- that is the bound compared with the unbound forms. The interaction observed for the linker peptide (Chapter 4) also underwent fast exchange on a NMR timescale. The largest change in chemical shift was observed for amino acids His 195 and Trp 215ε and these shifts are observed in the presence of a small amount of ligand. This indicates the environment surrounding the backbone of His 195 and Trp 215 changes with increasing ligand concentration. The HSQC experiment detects only backbone chemical shift changes through the N-H bond. Thus, amino acids whose sidechains interact with the ligand to a greater extent than the backbone may not be detected. Importantly, the amino acids involved in binding the site 1 peptide are also affected upon binding the linker ligand (section 4.7) implicating the same shallow cleft on the surface of the protein in ligand binding. The amino acids involved include the tyrosines Tyr 187, Tyr 205 and Tyr 232 and His 195, Asn 231, Trp 215, Thr 192, Ser 230, Trp 216, His 214, Glu 193, Leu 197 and Asp 188 (Figure 5.22 see also Figure 5.24). These sites are present as a hydrophobic pocket on the surface of the protein and represent the same binding region observed for other SH3 domains binding to PRR (Tzeng *et al.*, 2000). Table 5.3 lists the chemical shifts of Tec SH3 domain both in isolation and as a complex with PRR site 1.

The binding affinity of the PRR site 1 synthetic peptide to the SH3 domain was weak as determined by NMR analysis. An eight-fold (6mM) excess of ligand to protein was not sufficient for the entire SH3 domain present to be saturated with ligand. Additions of more ligand would be required to saturate the SH3 domain. The graphs presented in Figure 5.23 show the chemical shift changes against peptide concentrations for the amino acids Asn 231, His 195 and Trp 215ε1. Affinity constants were derived using Scatchard analysis using at least six points (section 2.5.7) as shown in Figure 5.23B. Change in chemical shift was plotted as a function of ligand concentration and the dissociation constant K_D calculated (Tsukube, *et al* 1996). All chemical shift changes were analysed and the dissociation constant determined from four amino acids with significant chemical shift changes His 214, Trp 215ε1, Asn 231 and His 195. The K_D value determined from His 195 deviated by greater than one standard deviation and was not included in the calculation. The K_D of the interaction between Tec SH3 domain and the PRR site 1 was 3 mM ± 1 mM. The affinity constant for the binding of site 1 peptide to the SH3 domain was in the low millimolar range

Figure 5.22

^{1}H-^{15}N HSQC plots of ^{15}N labelled Tec SH3 domain in the presence of increasing concentrations of Site 1 peptide. The spectra were recorded on a Varian Inova spectrometer with spectral widths of 8000 Hz in the ^{1}H dimension and 2000 Hz in the ^{15}N dimension. The number of t_1 increments collected was 64 and the number of points was 2048 with an acquisition time of 0.128 seconds. 16 transients were collected. The spectra were collected at 25°C using the ghsqcse experiment from Varian. Both dimensions in the spectra were processed with a sinebell function and linear prediction was used in the nitrogen dimension using the VNMR software. The final spectral size was 2048 and 256 and this was visualised and further analysed using XEASY (Bartels *et al* 1997). The protein sample was at 0.75 mM in 10 mM $NaHPO_4$ supplemented with 10% D_2O pH 6.0. The site 1 peptide was dissolved in water to a final concentration of 41.5 mM and this was added incrementally to the SH3 domain protein. The pH was checked following addition of the peptide and maintained at pH 6.0. Spectra from the SH3 domain with no site 1 peptide (shown in red), with 3 mM peptide (4 fold excess purple) and 6 mM (8 fold excess blue) peptide are represented. The numbering is consistent with mouse Tec IV protein.

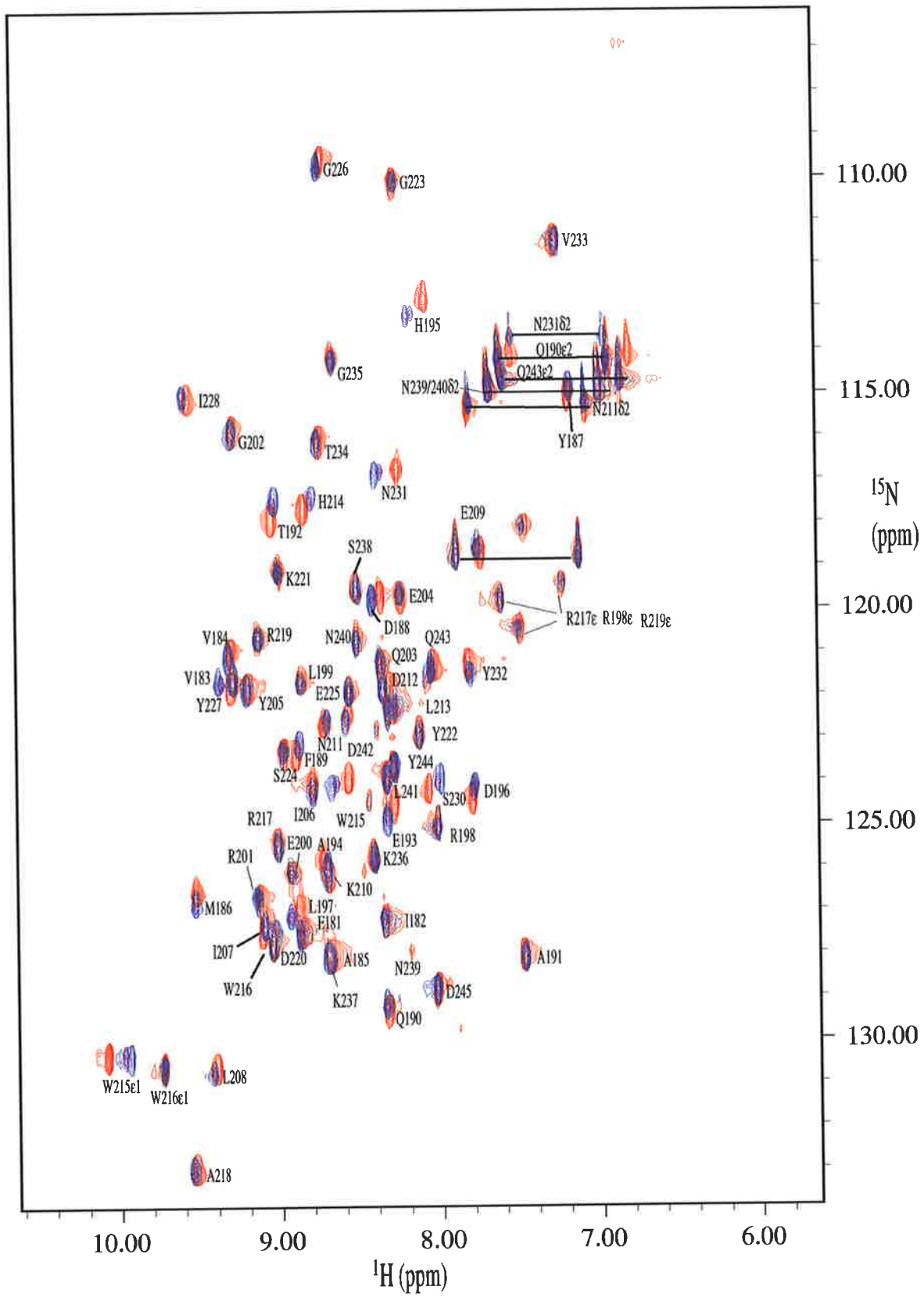
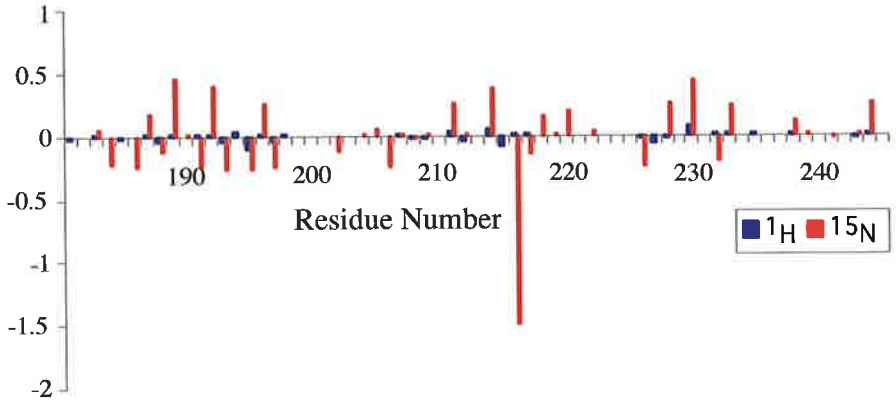

G226
G223
V233
H195
N231δ2
Q190ε2
Q243ε2
G235
N239/240δ2
I228
N211δ2
G202
T234
Y187
N231
H214
E209
T192
S238
K221
E204
R217ε R198ε R219ε
D188
V184
R219
N240
Q243
Q203
V183
L199
D212
Y232
Y227
Y205
E225
L213
Y222
N211
D242
F189
Y244
S224
L241
S230
D196
R217
I206
W215
E193
R198
E200
A194
K236
R201
K210
M186
L197
E181
I182
I207
D220
A185
N239
A191
W216
K237
D245
Q190
W215ε1
W216ε1
L208
A218
110.00
115.00
120.00
125.00
130.00
15N (ppm)
10.00
9.00
8.00
7.00
6.00
1H (ppm)

Figure 5.23

A. A bar graph showing the chemical shift deviation with ligand binding over the whole SH3 domain. The shifts represented in red are those in the ^{15}N dimension and those in blue are in the ^{1}H dimension.

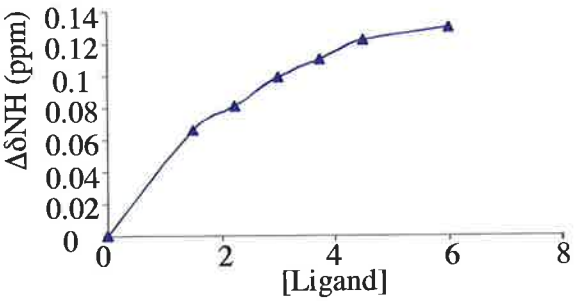
B. Chemical shift analysis of amino acids involved in binding to the PRR site 1 peptide. Amino acids Asp 231, His 195 and Trp 215ε1 are examples of amino acids shifted with increasing concentrations of site 1 peptide (mM). These and other amino acids were analysed to determine a kinetic dissociation constant by Scatchard analysis. Plots of the change in chemical shift (δ)/ligand concentration is plotted against the change in chemical shift. The Scatchard plots and the corresponding line equations are shown. The reciprocal of the slope is the K_D of interaction.

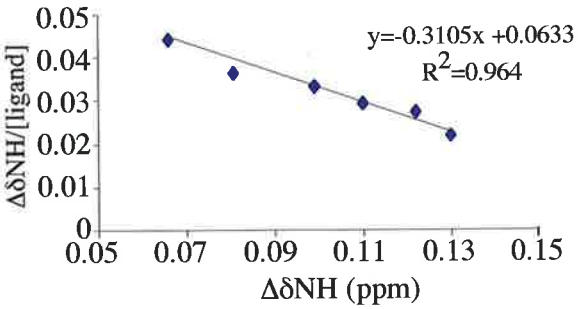
A



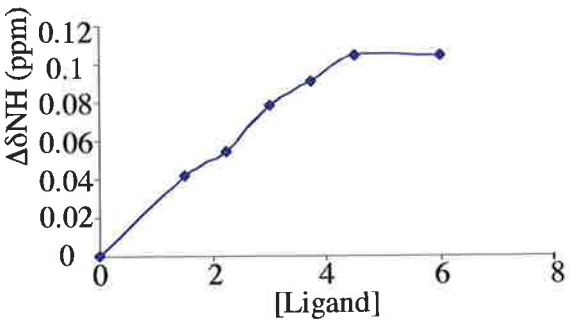
B

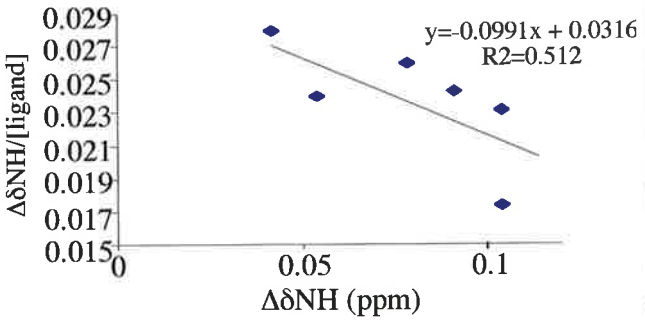
Asparagine 231



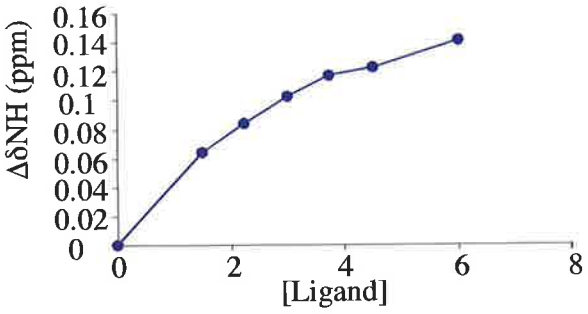


Histidine 195





Tryptophan 215ε



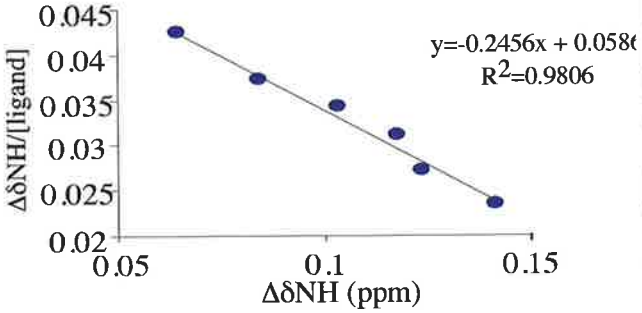
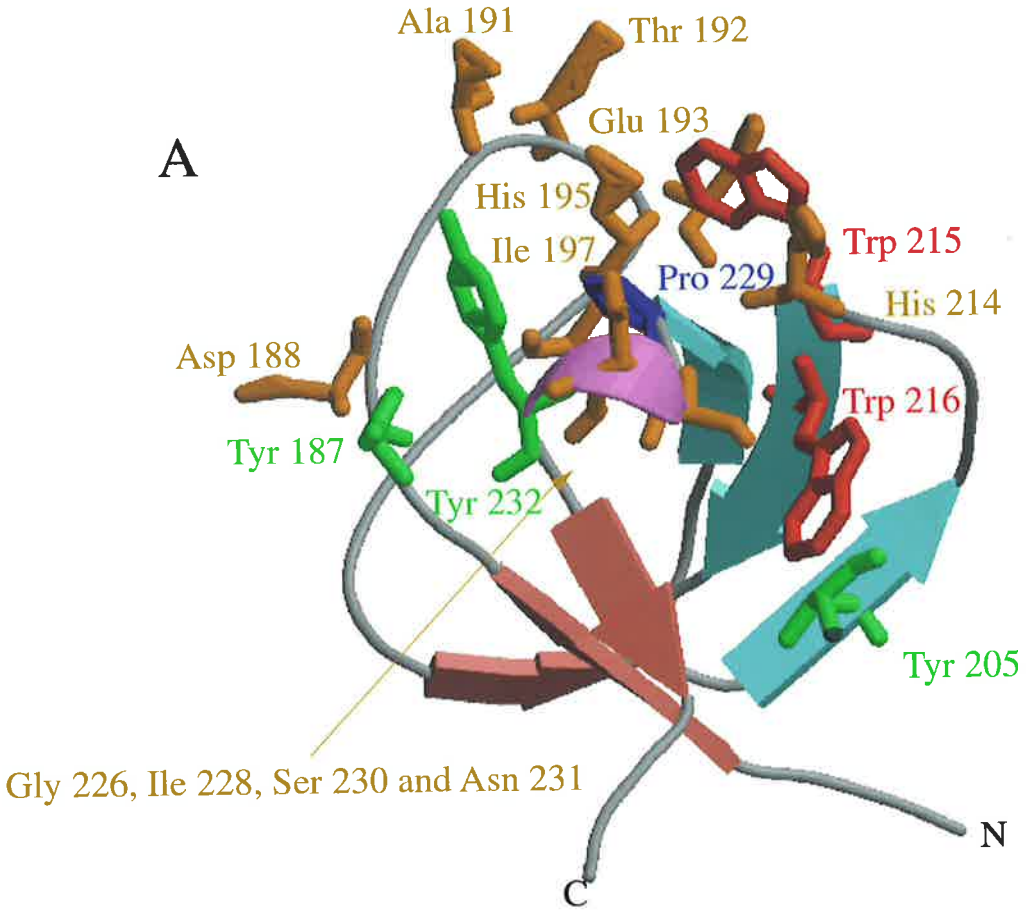


Figure 5.24

A. Ribbon diagram of Tec SH3 domain showing the amino acids detected to be involved in ligand binding by HSQC experiments. Trp 215 and Trp 216 are shown in red. Tyr 187, Tyr 205 and Tyr 232 are shown in green. Pro 229 is shown in blue while the remaining amino acids involved in ligand binding are shown in brown. The N- and C- termini are highlighted. The figure was generated in Molscript and Raster3D (Kraulis *et al.*, 1991; Merritt *et al.*, 1997).

B. Surface plot of Tec SH3 domain with the ligand-binding site generated in InsightII. . The plot has been rotated 90° along the Z axis relative to A. The binding site (magenta) was determined using NMR spectroscopy. Proline 51 is highlighted in yellow, as although it is not detected in an HSQC experiment it is likely to be involved in any interactions involving ligand. The autophosphorylated tyrosine is shown in white.


A
Ala 191
Thr 192
Glu 193
His 195
Ile 197
Pro 229
Trp 215
His 214
Asp 188
Trp 216
Tyr 187
Tyr 232
Tyr 205
Gly 226, Ile 228, Ser 230 and Asn 231
N
C

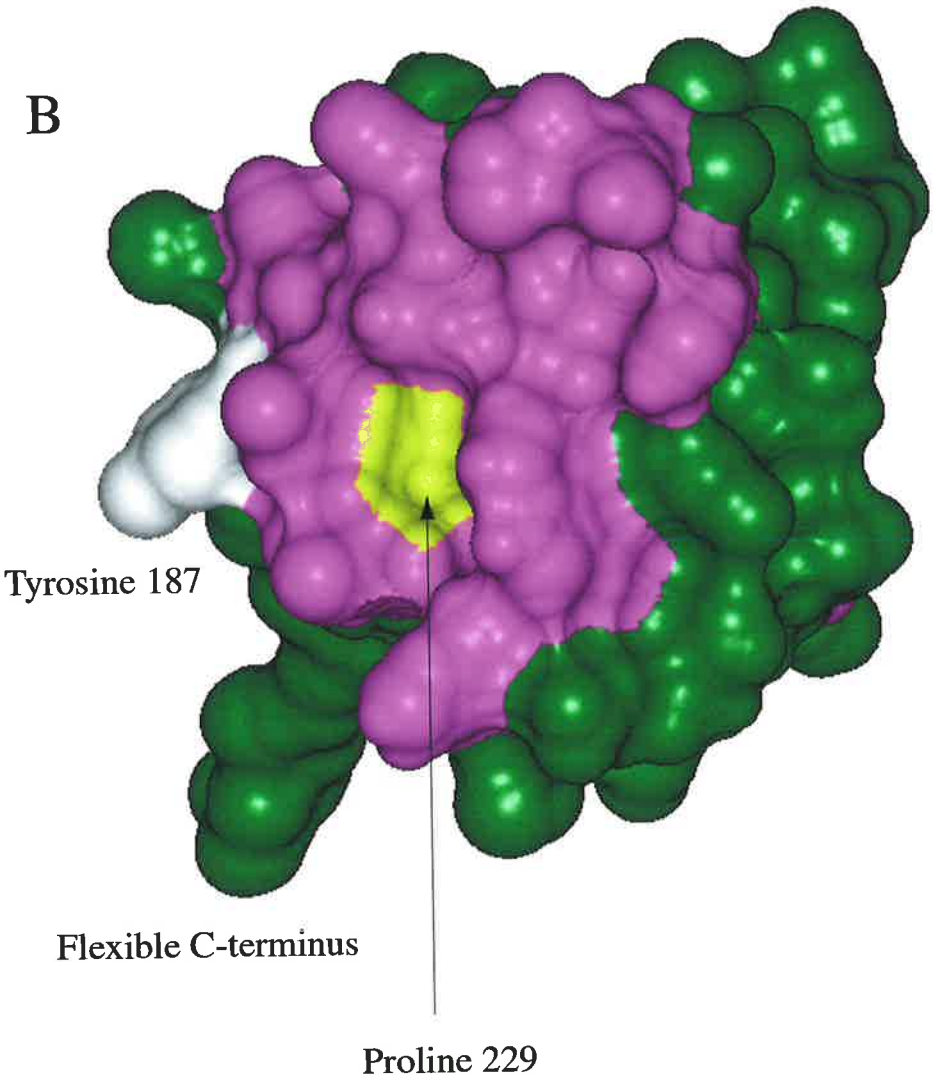

B
Tyrosine 187
Flexible C-terminus
Proline 229

Table 5.3 Chemical Shift values of Tec SH3 domain with the addition of PRR site 1 peptide

	No Addition of Site 1 peptide		Addition of 6 mM Site 1 peptide	
Residue No.	**^{1}H Chemical Shift**	**^{1}H Chemical Shift**	**^{15}N Chemical Shift**	**^{15}N Chemical Shift**
E181	127.324	8.854	127.328	8.879
I182	127.067	8.346	127.068	8.347
V183	121.678	9.289	121.625	9.278
V184	120.908	9.29	121.143	9.304
A185	128.095	8.672	128.093	8.689
M186	126.553	9.518	126.808	9.529
Y187	114.991	7.176	114.813	7.171
D188	119.615	8.359	119.742	8.413
F189	123.469	8.893	123.003	8.873
Q190	129.129	8.327	129.118	8.342
A191	127.838	7.487	128.088	7.474
T192	117.922	9.035	117.525	9.014
E193	124.497	8.281	124.756	8.327
A194	125.78	8.73	125.787	8.698
H195	112.93	8.073	113.19	8.177
D196	124.24	7.799	123.983	7.784
L197	126.81	8.88	127.058	8.936
R198	125.011	8.021	125.01	8.01
L199	121.67	8.854	121.673	8.864
E200	126.039	8.919	126.037	8.921
R201	126.553	9.128	126.55	9.141
G202	115.76	9.271	115.896	9.276
Q203	121.156	8.359	121.16	8.372
E204	119.615	8.229	119.609	8.241
Y205	121.84	9.183	121.786	9.193
I206	123.983	8.789	124.234	8.789
I207	127.319	9.102	127.313	9.098
L208	130.629	9.396	130.661	9.421
E209	118.587	7.734	118.582	7.758
K210	126.048	8.689	126.048	8.703
N211	122.698	8.724	122.449	8.699
D212	121.695	8.306	121.679	8.360
L213	122.18	8.255	122.191	8.269
H214	117.683	8.835	117.305	8.78
W215	123.983	8.568	123.982	8.665
W216	126.01	9.118	127.495	9.101
R217	125.268	9.01	125.414	9.006
A218	132.978	9.531	132.832	9.547
R219	120.645	9.12	120.64	9.122
D220	127.701	9.045	127.51	9.044
K221	119.101	8.984	119.098	8.998
Y222	122.955	8.125	122.918	8.124
G223	110.106	8.255	110.106	8.255
S224	123.212	8.958	123.213	8.973
E225	121.927	8.555	121.927	8.561
G226	109.601	8.704	109.849	8.724

...Cont'd

Table 5.3 cont'd *Chemical Shift values of Tec SH3 domain with the addition of PRR site 1 peptide*

	No Addition of Site 1 peptide		Addition of 6 mM Site 1 peptide	
Residue No.	**^{1}H Chemical Shift**	**^{1}H Chemical Shift**	**^{15}N Chemical Shift**	**^{15}N Chemical Shift**
Y227	121.67	9.3	121.671	9.373
I228	115.246	9.544	114.989	9.571
P229#	-	-	-	-
S230	124.24	8.073	123.797	7.995
N231	114.21	6.799	114.21	6.799
Y232	121.193	7.818	121.411	7.81
V233	111.622	7.257	111.394	7.24
T234	116.017	8.737	116.022	8.747
G235	114.219	8.652	114.218	8.646
K236	125.782	8.411	125.782	8.411
K237	127.84	8.699	127.84	8.699
S238	119.615	8.516	119.499	8.505
N239	127.849	8.19	127.839	8.191
N240	120.642	8.508	120.65	8.515
L241	123.726	8.32	123.756	8.331
D242	122.184	8.32	122.184	8.32
Q243	121.413	8.034	121.407	8.053
Y244	123.726	8.294	123.472	8.281
D245	128.608	8.024	128.606	8.028
190ε2	114.218	7.604	114.216	7.617
190ε2	114.218	6.927	114.224	6.941
203ε	118.824	7.885	118.824	7.885
203ε	118.844	7.122	118.844	7.122
211δ2	115.246	7.812	115.241	7.799
211δ2	115.246	7.07	115.244	7.083
215ε1	130.278	10.065	130.152	9.924
216ε1	130.659	9.726	130.407	9.727
231δ2	116.788	8.255	116.786	8.385
231δ2	114.218	7.539	114.218	7.539
239/240δ2	114.989	7.682	114.738	7.681
239/240δ2	114.941	6.985	114.767	6.989
243ε2	114.732	7.604	114.728	7.591
243ε2	114.732	6.849	114.726	6.855

Proline residues are not detected by HSQC experiment

suggesting a very weak interaction. The dissociation constant derived by NMR for the PRR site 1 interaction with Tec SH3 domain in solution is of a similar order to that derived for the Tec PRRSH3Δ2 protein interaction with the immobilised SH3 domain on the chip. HSQC experiments and chemical shift changes are a rapid method of determining the binding surface of a protein, however, it must be noted that this method can over estimate the ligand binding surface and only three dimensional structure determination of the complex will precisely show the molecular interactions involved in ligand binding (Ramos *et al.*, 2000).

5.6 PRRSH3Δ2 FORMS INTRAMOLECULAR INTERACTIONS

PRRSH3Δ2 protein dimerises at a low affinity K_D of 416.6 μM as determined by analytical ultracentrifugation and this interaction could be attributable to non-specific interactions. BIAcore analysis also indicated that an intramolecular interaction might form through PRR site 1 within PRRSH3Δ2 thus preventing an interaction of the PRR site 1 with the SH3 domain on the chip. To confirm an intramolecular interaction of PRR site 1 and the SH3 domain is occurring, a ^{15}N-labelled sample of PRRSH3Δ2 was produced. The yield of ^{15}N-labelled GST-PRRSHΔ2 was 20 mg/L culture medium. This sample was digested with thrombin and the GST removed by the addition of glutathione agarose. Following removal of the GST fusion partner, the sample was concentrated to 7 mg/ml and 10% D_2O added.

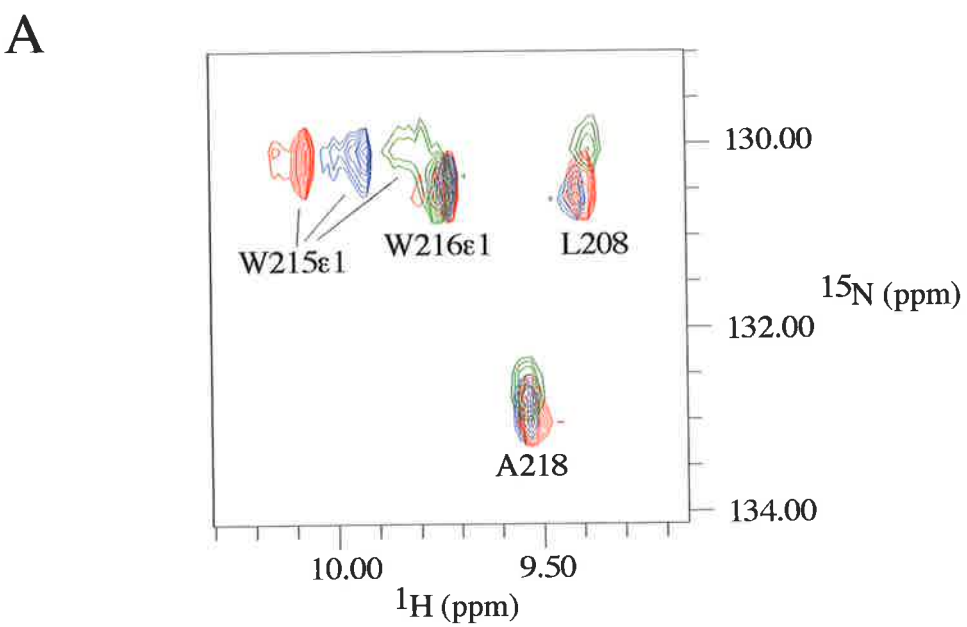
A heteronuclear HSQC spectrum was collected on a 0.63 mM sample of ^{15}N-PRRSH3Δ2 in PBS supplemented with 10% D_2O and the spectra compared with the HSQC spectra of the Tec SH3 domain alone (section 4.6) or with the PRR site 1 ligand (section 5.5.3). Figure 5.25 shows the HSQC peak for the Trp 215ε1 side chain relative to the Tec SH3 domain HSQC and the SH3 domain with 6 mM site 1 ligand added. Chemical shifts obtained in the PRRSH3Δ2 were located closer to the ligand bound form of the SH3 domain than the SH3 alone. In the case of Trp 215ε1 the magnitude of the chemical shift were greater than those observed in the SH3 domain titration suggesting the SH3 domain in PRRSH3Δ2 is more saturated than that observed in the PRR site 1 titration experiment (section 5.5.3). Figure 5.25C is the plot of chemical shift against fold increases of ligand where the PRRSH3Δ2 has been indicated with a star. Thus, PRRSH3Δ2 is likely to form intramolecular interactions that are mediated by the PRR site 1. Protein breakdown of PRRSH3Δ2 was observed following storage at 4°C and, thus, no further experiments were conducted on the PRRSH3Δ2 protein.

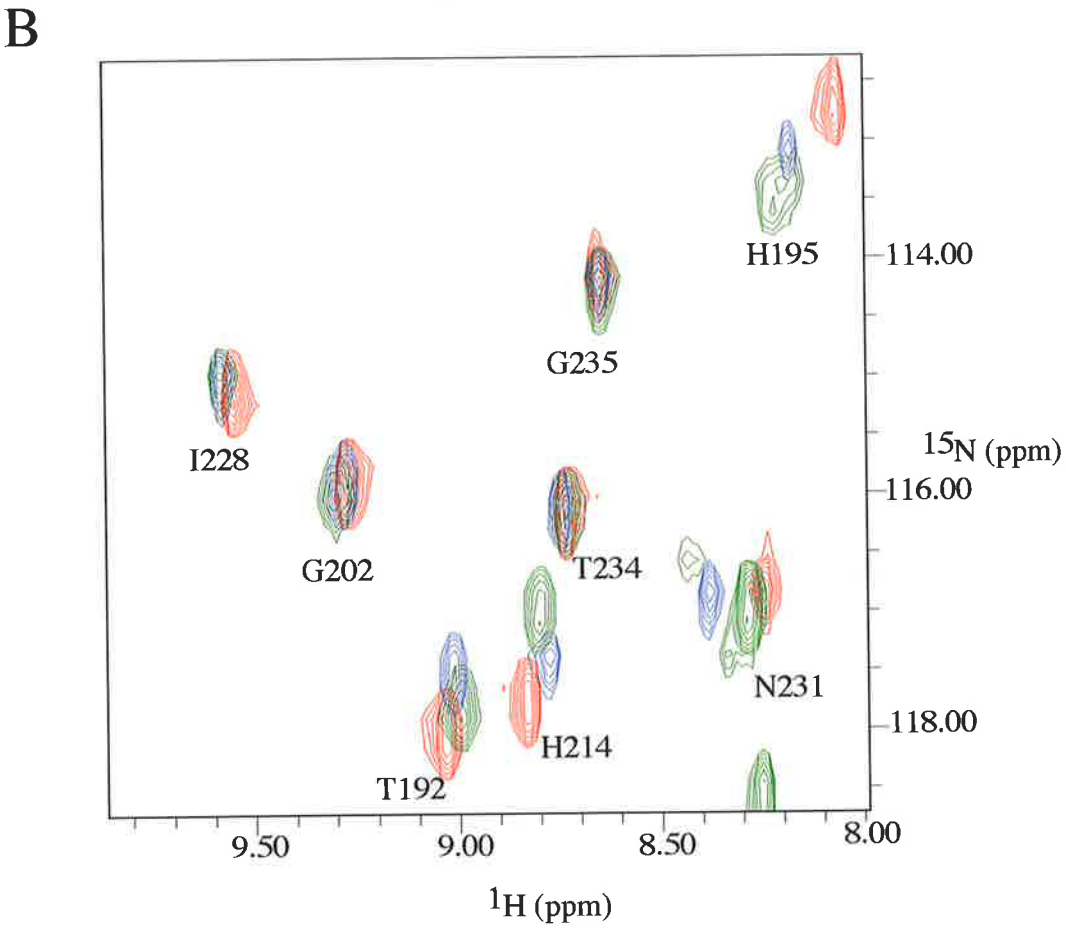
Figure 5.25

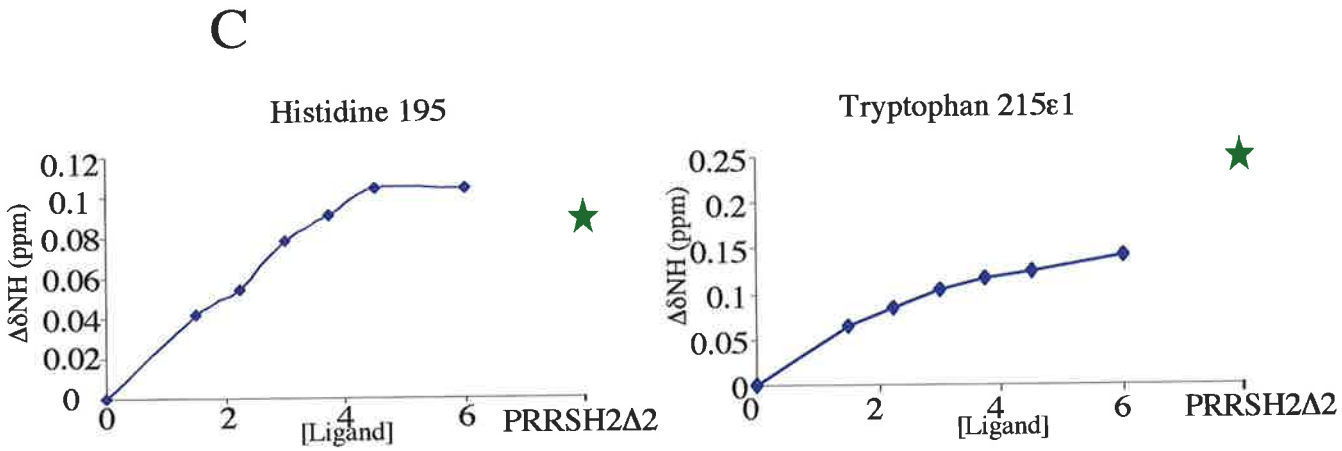
A. ^{1}H-^{15}N HSQC plots of ^{15}N labelled Tec SH3 domain compared to Tec PRRSH3Δ2 protein. The spectra were recorded on a Varian Inova spectrometer with spectral widths of 8000 Hz in the ^{1}H dimension and 2000 Hz in the ^{15}N dimension. The number of t_1 increments collected was 64 and the number of complex points was 2048 with an acquisition time of 0.128 seconds. 16 transients were collected. The spectra were collected at 25°C. Both dimensions in the spectra were processed with a sinebell function and linear prediction was used in the nitrogen dimension using the VNMR software. The final spectral size was 2048 and 256 and this was visualised and further analysed using XEASY (Bartels *et al.*, 1997). These spectra highlight the Trp 215 amino acid and the SH3 domain with no site 1 peptide (red), in the presence of 6 mM-8 fold excess peptide (blue) and the spectra obtained for ^{15}N PRRSH3Δ2 (green). The numbering is consistent with mouse Tec protein and has been assigned in chapter 4.

B. The spectral overlay of Tec SH3 domain (red), Tec SH3 domain with 6 mM site 1 peptide (blue) and PRRSH3Δ2 protein (green) between the chemical shifts of 8-10 ppm in the hydrogen dimension and 113-118 in the nitrogen dimension.

C. Plots of the change in chemical shift (δ)/ligand concentration versus change in chemical shift for the SH3 domain with increasing concentrations of site 1 peptide. The curves are for His 195 and Trp 215ε1 amino acids. The relative shift of the PRRSH3Δ2 protein has been indicated on the curve by a green star.


A
W215ε1
W216ε1
L208
A218
130.00
132.00
134.00
15N (ppm)
10.00
9.50
1H (ppm)


B
H195
G235
I228
G202
T234
N231
H214
T192
114.00
116.00
118.00
15N (ppm)
9.50
9.00
8.50
8.00
1H (ppm)


C
Histidine 195
ΔδNH (ppm)
0.12
0.1
0.08
0.06
0.04
0.02
0
0
2
4
6
[Ligand]
PRRSH2Δ2
Tryptophan 215ε1
ΔδNH (ppm)
0.25
0.2
0.15
0.1
0.05
0
0
2
4
6
[Ligand]
PRRSH2Δ2

5.7 DISCUSSION

The amino acid sequence of Tec kinase contains three complete class I SH3 ligand consensus sequences (+XXPXXP) and one incomplete consensus (PXXP) with three of the sites overlapping. The work reported in this chapter has shown that, using solution-based assays (ultracentrifugation and NMR) and a solid phase BIAcore assay, the Tec SH3 domain can form an interaction with the PRR.

Initial ultracentrifugation data highlighted that at solution equilibrium, Tec PRRSH3 wild-type protein will form dimers with an equilibrium dissociation constant of 125 μM. Tetramer production of Tec PRRSH3 protein was also observed by ultracentrifugation. The Btk PRRSH3 protein has also been observed to form dimers with a dissociation constant of 60 μM (Hansson *et al.*, 2001); two fold stronger than that observed for Tec kinase PRRSH3. In contrast, the Itk PRRSH3 domain forms an intramolecular interaction between the PRR and the SH3 domain (Andreotti *et al.*, 1997).

PRRSH3 proteins missing one or both of the SH3 domain ligand consensus sequences were generated by site-directed mutagenesis. Tec PRRSH3 wild-type and PRRSH3Δ1 behave similarly in both the ultracentrifugation and the BIAcore experiments. Both proteins readily form dimers and tetramers with dimer equilibrium dissociation constants in the order of 100 μM and interact with immobilised Tec SH3 domain on the surface of the BIAcore chip. Tec PRRSH3Δ2 and PRRSH3Δ12 exhibit distinct characteristics compared with PRRSH3 wild-type and PRRSH3Δ1 in the ultracentrifugation and BIAcore experiments. Dimerisation was much reduced compared with PRRSH3 wild-type and PRRSH3Δ1 and the interaction with the immobilised SH3 domain surface in the BIAcore experiments was also reduced. A peptide modelled on PRR site 1 was shown to bind the Tec SH3 domain through a weak fast exchanging interaction in the order of 3 mM by NMR. Thus, this work has shown that PRR site 2 promotes an intermolecular interaction whereas PRR site 1 does not facilitate dimerisation and, instead, could promote an intramolecular interaction.

Dimerisation of PRRSH3Δ2 observed by ultracentrifugation experiments indicated that the protein fragment PRRSH3Δ2 is essentially a monomer with a very weak dimer association. This suggests that PRR site 1 is either not capable of binding to the SH3 domain or that the equilibrium favours an intramolecular association. Tec PRR site 1 conforms to the class I consensus sequence for SH3 domain binding making an interaction

possible. It is therefore likely that PRR site 1 forms an intramolecular interaction with the SH3 domain that prevents any further intermolecular interactions. Three dimensional structure determination of the PRRSH3Δ2 protein will confirm the presence of the intramolecular association through PRR site 1. In contrast, PRR site 2 (PRRSH3 wild-type and PRRSH3Δ1 proteins) does not appear to associate intramolecularly with the SH3 domain, perhaps because it is adjacent to the SH3 domain. Thus, the PRR site 2 promotes a dimerisation event that is favoured over the intramolecular interaction through PRR site 1.

Other Tec family members contain one or more SH3 domain binding consensus sequences within the PRR of the protein that correspond to PRR site 1 and PRR site 2 in Tec (Figure 5.7). In the case of Btk, which contains both PRR site 1 and PRR site 2, the first motif (site 1) has been shown to bind the SH3 domain of Btk as well as the SH3 domains of Lyn, Fyn and Hck (Cheng *et al.*, 1994); (Alexandropoulos *et al.*, 1995). Pro-Ala mutations in the Btk PRR sites, individually and in concert, indicate that Btk PRR site 1 mediates these SH3 intermolecular associations *in vitro* and *in vivo.* A fragment of Btk protein incorporating the PRR and the SH3 domain (Btk PRRSH3) has been shown to form dimers (Hansson *et al.*, 2001). Dimerisation of Btk PRRSH3 is likely to depend on PRR site 2 as this site in Tec kinase is responsible for the dimerisation reported in this work.

Itk contains only one SH3 consensus binding site that corresponds to Tec PRR site 1. An Itk fragment consisting of the PRR and SH3 domain has been shown to form an intramolecular interaction between the SH3 domain and the PRR (Andreotti *et al.*, 1997). Tec PRRSH3Δ2 protein does not appear to dimerise and NMR spectroscopy has suggested the protein is bound in an intramolecular interaction. Thus, a role for PRR site 1 in intramolecular interaction of Tec is consistent with the work reported for Itk.

The two SH3 domain binding consensus sequences within the PRR appear to perform different functions and these functions are conserved between different Tec family members. The PRR site 1 appears to mediate intramolecular interactions and an interaction with other cellular ligands, whereas PRR site 2 mediates the dimerisation/tetramerisation reaction. Tec kinase has been shown to interact with c-Kit, and the SH3 domains of Lyn and Vav (Tang *et al.*, 1994) (Machide *et al.*, 1995; Mano *et al.*, 1994). From the evidence presented here, it is likely that these interactions are mediated by the PRR site 1. A further prediction from this work is that Txk, which lacks PRR site 1 but contains PRR site 2 will form dimers. This is intriguing for signalling downstream of the T cell receptor as both Itk and Txk are required for signalling although their roles are different. Itk contains a PRR site 1 and no PRR site 2 whereas Txk contains a PRR site 2 and no site 1. The possibility that

these two proteins share functions based on the presence or absence of SH3 domain binding site in the PRR is interesting and warrants further investigation.

Thus, work presented here implies that Tec family members might utilise an intra- or intermolecular association as a regulatory mechanism to control their kinase activity. The internal autophosphorylated tyrosine of Tec is located in the SH3 domain ligand binding site and must be phosphorylated for full enzymatic activity. This tyrosine is protected and not easily phosphorylated upon ligand binding. Thus, bound ligand may prevent full enzymatic activity of the protein. Tec kinases use a similar overall strategy to Src kinases to regulate their activity but utilise distinct interactions.

CHAPTER 6:

FINAL DISCUSSION AND FUTURE DIRECTIONS

6.1 FINAL DISCUSSION

Structural characterisation of proteins can provide important information regarding potential binding surfaces and mutational studies can ultimately link different proteins together that participate in the same signalling pathway.

Tec family kinases are essential for signal transduction following antigen engagement of T cell, B cell and Fcγ receptors (Yang *et al.*, 2000). The interactions involved have recently been elucidated and suggest equivalent roles for different Tec family members downstream of different receptors (section 1.6.2). A degree of redundancy of Tec family members has been observed. The atypical Tec family members Bmx and Txk can reconstitute the PLC-γ signalling in DT40 cells lacking Btk (Tomlinson *et al.*, 1999). *Btk*$^{-/-}$ mice have a severely compromised immune system resulting from a blockage in the development of B cells. Mice lacking Tec kinase show no obvious phenotype, however, mice lacking Tec and Btk have a slightly more severe phenotype than those lacking Btk alone (Ellmeier *et al.*, 2000). This suggests that Btk can partially substitute for Tec kinase, however, the reciprocal does not occur. Therefore, the predominantly expressed Tec family member (Btk in B cells) is essential for downstream signalling and its function cannot be compensated for by lesser expressed Tec family proteins (Tec in B cells). The function of the secondary Tec family member however, can be partially compensated for by the major Tec family member. In addition, overexpression of Tec family members can recover knockout phenotypes as family members substitute for each other.

Role of Tec PHTH domain

The PH domains of Tec family kinases are absolutely required for membrane localisation following receptor engagement in all haematopoietic lineages. The PH domain of Btk has been shown to bind to phosphatidylinositol phosphates and also the βγ subunit of G proteins, however, the structure and function of the TH domain was unknown (Baraldi *et al.*, 1999); (Touhara *et al.*, 1994). This thesis initiated an investigation into the structure of the PHTH region of Tec kinase.

The expression and purification of the Tec PHTH protein proved problematic. A variety of coding regions, purifications under both denaturing and non-denaturing conditions and different expression systems were tested. Generation of protein samples of sufficient quantity for structural investigations was not achieved. The yield of Trx-PHTH purified under denaturing conditions was improved compared with non-denaturing purification,

although subsequent refolding was not successful. Addition of zinc to the culture medium resulted in increased yields of the proteins PHTH, THSH3 and the TH alone but the final amount of each protein in these samples was still not sufficient for further analysis. The THSH3 protein contains the PRR and, based on the results of Chapter 5, most likely forms dimeric or tetrameric complexes, although this was not further investigated.

During the course of this work the crystal structure of Btk PHTH was determined with and without the Ins (1,3,4,5)P_4 ligand and revealed the TH domain was a zinc binding motif that formed contacts with one surface of the PH domain (Hyvonen and Saraste, 1997); (Baraldi *et al.*, 1999). The initial structure determination of Btk PHTH domain was conducted on the R28C mutant of the PH domain, as this was the only soluble protein generated that produced high quality crystals (Hyvonen and Saraste, 1997).

Further structural characterisation of this region of Tec kinase will require protein expression to be performed in either an insect expression system, mammalian expression system (eg Cho cells) or a cell free system to aid in the folding and solubility of the protein (Archer *et al.*, 1993); (Kigawa *et al.*, 1995). The difficulty of these systems is the expense of generating the ^{15}N, and ^{13}C labelled material required for NMR spectroscopy analysis.

6.1.1 The mechanism of Tec kinase regulation

Intramolecular interactions are important to maintain the Src tyrosine kinases in a closed and therefore inactive conformation. Deregulated tyrosine kinase activity of the Src family kinases results in constitutive activation of the cellular pathways controlling cell growth and differentiation. Thus, tight regulation of the enzymatic activity of tyrosine kinases is critical for maintaining normal cellular function and preventing oncogenic activation. Evidence from Src and Hck structural studies show that they are maintained in an inactive conformation by the presence of two weak interactions (Sicheri *et al.*, 1997); (Xu *et al.*, 1997). One interaction occurs between the SH2 domain and the regulatory tyrosine near the C-terminus and the other occurs between the SH3 domain and the polypeptide sequence between the SH2 and the kinase domain (SH2 kinase linker). Tec family members lack the C-terminal regulatory tyrosine found in Src kinas. However, evidence from studies on Itk and Btk suggest an inhibitory mechanism exists and that it involves the PRR and adjacent SH3 domain. The PRR of Itk can fold back on itself and intramolecularly bind the SH3 domain, thereby maintaining Itk in a closed inactive conformation (Andreotti *et al.*, 1997). However, Btk PRRSH3 domain has been shown to form dimers with an affinity of

60 μM (Hansson *et al.*, 2001). Tec kinase contains two SH3 consensus sequences within the PRR adjacent to the SH3 domain. These two sites, PRR site 1 and PRR site 2, conform to consensus SH3 binding ligands and an investigation was initiated to further analyse the potential for such an interaction in Tec.

To begin this investigation, the solution structure of Tec SH3 domain was determined. This structure was determined using two and three dimensional NMR spectroscopy techniques and the iterative approach of Nilges for NOE assignment (Nilges *et al.*, 1997). Although all SH3 domains share a common fold, the Tec SH3 domain more closely resembles a subclass of SH3 domains that includes Hck and Fyn tyrosine kinases. SH3 domains of this class are composed of a six stranded, two β-sheet β-barrel with the second strand shared between the two sheets. The amino acids at the ligand binding site were identified from chemical shift perturbations that occurred upon binding a peptide derived from the sequence of the SH2 kinase linker. The ligand binding site is a shallow cleft on the surface of the protein and is consistent with those of other SH3 domains. The affinity of the interaction between the Tec SH3 domain and the ligand was 4 mM.

The role of the PRR

The prolines in the +4 position (+XXPXXP) of each of the SH3 binding consensus sequences of the PRR were substituted with alanines, both separately and in tandem, using site directed matagenesis. The mutations were designed to prevent the formation of the polyproline type 2 helix required for binding to the Tec SH3 domain. An equivalent mutation in Itk prevents binding of the mutant PRR legend to the Itk SH3 domain (Andreotti *et al.*, 1997).

Mutagenesis and subsequent biophysical characterisation of the Tec PRRSH3 proteins showed that PRR site 1 (PRRSH3Δ2) formed an intramolecular interaction with the adjacent Tec SH3 domain while PRR site 2 (PRRSH3Δ1) formed intermolecular dimers and tetramers. In the wild-type PRRSH3 protein, a dimer, formed through an interaction between PRR site 2 and the SH3 domain, will predominate over monomers with intramolecular interaction, as the dimer association is greater than the corresponding intramolecular association. PRRSH3Δ2 protein contains only the site 1 SH3 binding site in the PRR. It's dimerisation dissociation constant, determined by analytical ultracentrifugation, was 417 μM indicating that below a concentration of 600 μM the monomeric form of the protein is in excess. Using NMR spectroscopy and a site 1 synthetic

peptide, the affinity for the intermolecular interaction of PRR site 1 and the Tec SH3 domain was determined to be low with a dissociation constant of 3 mM. This dissociation constant is consistent with the dissociation constant derived by BIAcore for PRRSH3Δ2 binding to the immobilised SH3 domain. Structural determination of a high resolution complex of Tec SH3-PRR site 1 was not possible because of the fast exchange regime of the ligand, producing insufficient build up of NOEs. NMR data confirmed that the intermolecular interaction between the PRR site 1 and the SH3 domain (PRRSH3Δ2) does not occur, and that this protein is possibly already bound through an intramolecular interaction thereby making it unavailable to bind in an intermolecular interaction.

BIAcore and ultracentrifugation experiments have shown a dimerisation interaction of high affinity between two reciprocal PRR site 2-SH3 (PRRSH3Δ1) molecules with a dimerisation dissociation constant K_D of 50 μM. Although not determined by NMR, the affinity of the interaction between the PRR site 2 and Tec SH3 domain is expected to be higher. Steric constraints sre predicted to prevent intramolecular interaction between Tec SH3 domain and PRR site 2 (T. D. Mulhern. unpublished results). Thus, it is unlikely that an intramolecular interaction between the Tec SH3 domain and adjacent PRR site 2 will occur in the context of the whole protein in a cellular environment. The affinity of the interactions observed was low. This is not surprising since PRR SH3 domain interactions are found in situations where the rapid recruitment or interchange of several proteins occurs (Kay *et al.*, 2000).

The observations described in this thesis indicate that the two SH3 domain binding consensus sequences within the PRR of Tec kinase perform distinct functions that are conserved between different Tec family members. PRR site 1 appears to mediate intramolecular interactions and an interaction with other cellular proteins whereas PRR site 2 is responsible for the dimerisation and tetramerisation of PRRSH3 region of Tec family proteins. Itk's PRR contains one SH3 consensus site (equivalent to PRR site 1 in Tec) that mediates an intramolecular interaction with the SH3 domain whereas Btk contains two SH3 consensus sites and the equivalent PRR site 1 has been shown to interact with cellular proteins. A PRRSH3 fragment of Btk has also been shown to dimerise (Hansson *et al.*, 2001).

The linker between the SH2 domain and the kinase domain of Tec contains the same amino acid characteristics as the equivalent region in Src kinase (Xu *et al.*, 1997); (Sicheri *et al.*, 1997). Thus, this region would be expected to form a polyproline type 2 helix required for interaction with the SH3 domain. This thesis has focussed on SH3 domain interactions

with the PRR, however, an interaction between the Tec SH3 domain and the SH2 kinase linker may also occur in the full-length protein. It is unclear if such an interaction would be sufficient to maintain Tec in an inactive conformation. The SH2 kinase linker could provide yet another interaction to an already complex system for the control of the kinase activity of Tec family members.

The intermolecular interaction between Tec SH3 domain and the PRR is predicted to sequester the SH3 domain and the PRR within the cell. It is possible that PRR site 1 binds intramolecularly to the SH3 domain while PRR site 2 is free to make other contacts. The cellular ligands for both the SH3 domain and the PRR site 1 would therefore be unable to bind to the Tec protein. It is also possible that an interaction occurs between Tec SH3 domain and the SH2 kinase linker, allowing the PRR to interact with other cellular proteins, perhaps via other SH3 domains, WW domains or Profilin.

Tec kinase has been shown to associate with c-kit and is tyrosine phosphorylated following binding of stem cell factor to c-kit. This association takes place via a region incorporating the PHTH domain of Tec kinase (Tang *et al.*, 1994). Tec also interacts with the SH3 domain of the Src family kinase Lyn (Mano *et al.*, 1994) and with Vav through the TH region (composed of the Btk motif and the PRR) (Machide *et al.*, 1995). The Tec PRR is likely to be responsible for these interactions. The PRR will be more accessible for binding these cellular ligands if there is no intermolecular interaction between the SH3 domain and the PRR. Although the SH3 domain could bind either the SH2 kinase linker or the PRR, it is still uncertain which will occur in the cell. The PRRSH3 dimerisation reaction was the strongest interaction observed and will clearly dominate over the intramolecular interaction between the PRR and the SH3 domain in the isolated domain. It could be postulated that following activation of PI3K there would be a high concentration of Tec kinase at the membrane that could foreseeably place two or more molecules of Tec together and an intermolecular interaction could take place. In fact, a dimerisation event, however transiently, must take place for the autophosphorylated tyrosine in the SH3 domain of Tec to become phosphorylated. This interaction may take place through reciprocal SH3 PRR interactions. Whether Tec kinase is present in the cytoplasm of the cell as a dimer or tetramer has not been determined.

Potential regulation mechanisms of Tec family kinases

Deletion of the SH3 domain of Tec kinase results in constitutive activation of Tec kinase activity indicating that this removes a negative regulatory mechanism (Yamashita *et*

al., 1996). Tec III protein is a naturally occurring Tec isoform that lacks the C-terminal 22 amino acids of the SH3 domain and is predicted to have a non-functional SH3 domain. Tec III is a constitutively active form of Tec kinase (Yamashita *et al.*, 1996). The loss of a functional SH3 domain results in deregulation of the kinase activity of Tec tyrosine kinase. Thus, the SH3 domain is crucial in the regulation of the Src and Tec family of tyrosine kinases by participating in both intramolecular interactions and intermolecular interactions. Several mechanisms of Tec kinase activation can be proposed based on the evidence presented in this thesis.

Following receptor activation, Tec kinase is translocated to the membrane through an interaction between the PH domain and phosphatidylinositol lipids (Figure 6.1). Translocation of the Tec protein to the membrane could sufficiently saturate the membrane with Tec protein allowing a dimerisation or tetramerisation event between reciprocal SH3/PRR regions to occur (Figure 6.1). PRRs in CD28 have been shown to bind the Tec SH3 domain and the SH3 domain-CD28 interaction could out-compete the PRR-SH3 domain interaction (Figure 6.2) (Yang *et al.*, 1999). Phosphorylation of the tyrosine in the active loop by Src kinases will partially activate Tec kinase. Autophosphorylation of the tyrosine in the SH3 domain would then break the SH3-CD28 interaction and result in the complete activation of the Tec protein. The full activation of Tec and phosphorylation of the regulatory tyrosine in the SH3 domain might then break the interaction of the SH3 domain and the PRR in CD28 (Figure 6.2).

Alternatively, the SH2 kinase linker and the PRR could act in tandem to regulate Tec kinase. In the inactive state, an interaction of the SH3 domain and the SH2 kinase linker and potentially other interactions in Tec kinase could maintain the inactive state of the protein. Membrane localisation of Tec kinase through the PH domain might result in a conformational change that facilitates the interaction between the SH3 domain and the PRR. This interaction may occur through either a dimerisation/tetramerisation event or an intramolecular interaction. Following activation via Src family kinases and phosphorylation of the SH3 domain tyrosine, the interaction of the PRR and the SH3 domain will again be broken activating the kinase. Following downstream activation of effectors, the PRR-SH3 interaction may reform facilitating the formation of the inactive kinase.

The PRR-SH3 domain interaction might also be a mechanism to simply facilitate the autophosphorylation of the tyrosine in the SH3 domain of Tec.

Finally, in the inactive state, Tec may exist as a dimer or tetramer linked by reciprocal SH3 domain PRR interactions. Data presented in this thesis has shown that the

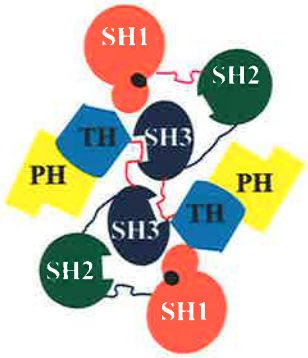
Figure 6.1

A. A schematic representation of interactions of Tec kinase that could maintain Tec in an inactive conformation. Both lobes of the SH1 domain (kinase domain) are shown in orange and the active site is shown in black. The SH2 domain is shown in green, the SH3 domain is shown in dark blue, the TH domain is shown in light blue and the PH domain is shown in yellow. The SH2 kinase linker is shown in magenta and the PRR is shown in red. There are two possible mechanisms to maintain Tec in an inactive conformation, those utilising the SH3-PRR interactions and those utilising the SH2-kinase linker. The dimer models have been depicted, however tetramers can also occur. The two mechanisms are labelled.

B. A schematic of possible interactions of Tec kinase following membrane localisation through the PH domain in T cells. Activation of PI3K produces PIP_3, a ligand for the PH domain of Tec family kinases. This facilitates the membrane localisation of these proteins. The TCR and CD4/CD8 receptors are shown in grey and Lyn tyrosine kinase is shown in pink. PIP_3 is shown by the green diamonds. The other domains are coloured as in A. The three possibilities shown here are dimerisation through a single SH3-PRR interaction, dimerisation through reciprocal SH3-PRR interactions and intramolecular interaction either between a SH3 domain and the PRR or the SH2-linker linker. The dimer models have been depicted, however tetramers can also occur.

A

Inactive Tec kinase

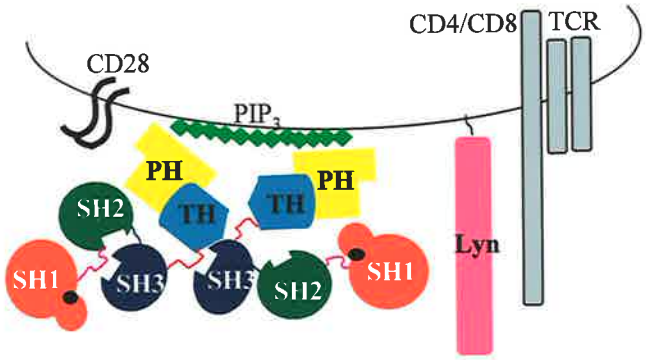


Dimerisation through reciprocal SH3 PRR interaction

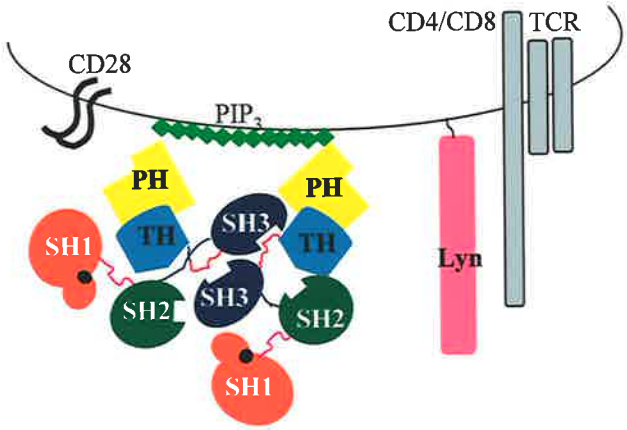
Intramolecular interactions through SH2 kinase linker and the SH3 domain

B

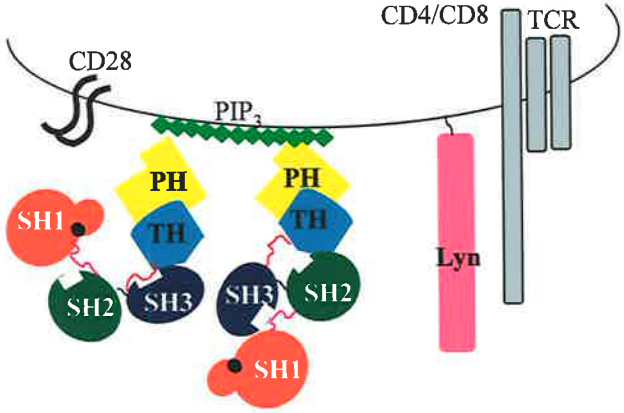
Membrane Localisation through a PH domain mediated interaction



Dimerisation through a single SH3-PRR interaction



Dimerisation through reciprocal SH3 PRR interactions



Intramolecular interactions through PRR and/or SH2 kinase linker and the SH3 domain

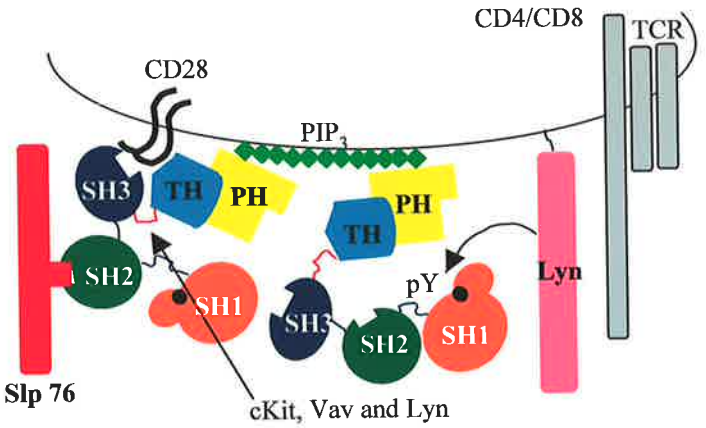
Figure 6.2

A. A schematic of possible interactions of Tec kinase following membrane localisation and activation by Lyn tyrosine kinase. Lyn kinase phosphorylates Tec family kinases in the catalytic (SH1) kinase domain. The SH2 domain of Tec kinases has been shown to interact with SLP-76 adapter proteins (red). This will potentially break an interaction between the SH2 kinase linker and the SH3 domain. The PRR of Tec kinases may then be available for binding to other cellular ligands including cKit, Vav and Lyn. The phosphorylation of Tec kinases and the interactions with the SLP family of adapters relaxes the structure such that autophosphorylation of the tyrosine in the SH3 domain is possible. The colours of the domain in Tec kinase are equivalent to those in the figure 6.1.

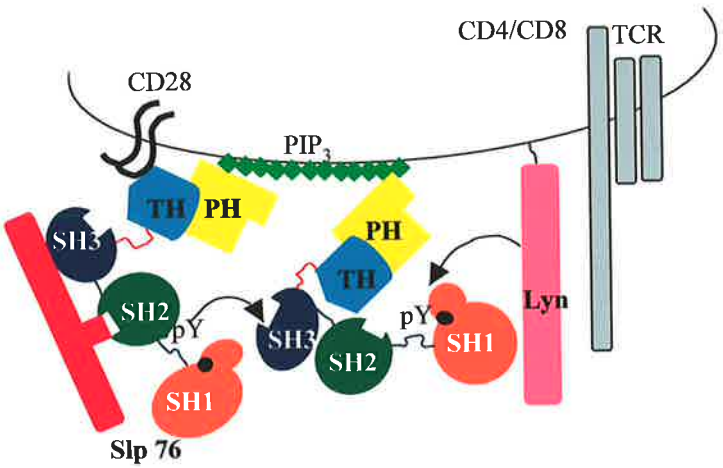
B. A schematic of autophosphorylation of Tec within the SH3 domain. The phosphorylation of the SH3 domain tyrosine (Tyr 187 in Tec) could break any remaining inhibitory interactions mediated by the SH3 domain rendering Tec kinase fully active. The colours of the domain in Tec kinase are equivalent to those in the figure 6.1.

C. A schematic of PLC-γ phosphorylation by Tec family members. Phosphorylation results in activation of PLCγ (blue) and the release of DAG and IP3 leading to protein kinase C activation and Ca^{2+} mobilisation, respectively. Dephosphorylation of PIP_3 by SHP-1 would release Tec kinases from the membrane and facilitate the return to an inactive conformation of these proteins. The colours of the domain in Tec kinase are equivalent to those in the figure 6.1.

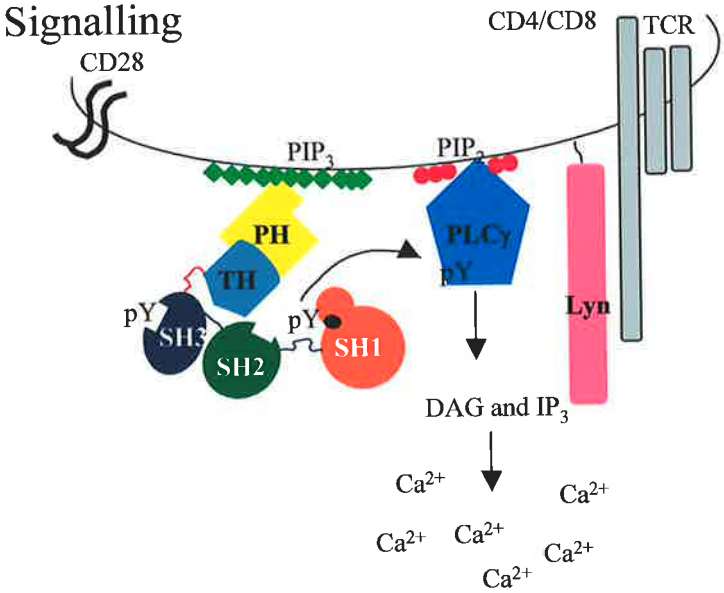
A Phosphorylation of Tec kinase domain by Lyn



B Autophosphorylation of Tec SH3 domain



C Phosphorylation of PLCγ and Ca^{2+} Signalling



PRRSH3 region of Tec kinase can form dimers *in vitro* at high affinity. Evidence from Btk also suggests that Btk PRRSH3 can also readily form dimers (Hansson *et al.*, 2001). It is also possible the SH3 domain PRR dimer interaction can occur *in vivo*. Thus, Tec kinase may be maintained in an inactive conformation by a dimerisation/tetramerisation event that requires activation of the protein to break. The autophosphorylation of the SH3 domain tyrosine, possible as these proteins are already dimers, might be the event that breaks the dimer interaction and produces the fully active conformation of Tec kinase.

Itk protein is constitutively bound to the cell membrane in T cells via the PH domain, removing the first stage necessary for the activation of Tec kinase (Yang *et al.*, 2000). Thus, a method of regulation utilising the SH3 domain could prove to be sufficient for kinase activation. Itk can be predicted not to form PRR-SH3 dimers (Andreotti *et al*, 1997) and, thus, an intramolecular interaction between the SH3 domain and the PRR could prevent incorrect activation. Removal of this inhibition would occur only when the tyrosine in the SH3 domain is phosphorylated and the kinase is active. Regulation of the atypical Tec family members Txk and Bmx is expected to vary. Txk lacks a PH domain and contains only one PRR (PRR site 2), which is suggested, from data presented here, to mediate dimer formation. Thus, this protein could utilise a dimerisation mechanism of regulation over the PRR-SH3 intramolecular mechanism. Bmx contains a non-standard SH3 domain and lacks a PRR, thus, if the SH3 domain was capable of binding a ligand then an inactive conformation could be maintained by a SH3 domain SH2 kinase linker interaction.

6.2 FURTHER EXPERIMENTS

A possible mechanism for regulation of Tec kinases has been proposed from the work conducted in this thesis. Initially, determination of the structure of the PRRSH3Δ2 protein will confirm the intramolecular interaction of PRR site 1 with the SH3 domain of Tec kinase. Experiments will then need to be conducted to confirm this proposed mechanism in cell-based assays or *in vivo*. This requires expression of full-length Tec kinase in mammalian tissue culture cells. Initial experiments will need to determine whether Tec kinase can be found present in the cell as a dimer. Transfection of a tagged version of the protein would be required, followed by standard immunoprecipitation experiments using the tag label to isolate the complex. Kinase assays can be conducted to differentiate between levels of Tec kinase activation. The activation of Tec kinase might be predicted to differ if the protein is present as a dimer or monomer. Equivalent PRRSH3Δ1, PRRSH3Δ2 and

PRRSH3Δ12 mutations in the full length protein will provide valuable information about the importance of these regions of Tec kinase. If the dimer through PRR site 1 were required to maintain Tec in an inactive conformation, expression of the PRRSH3Δ1 protein would be predicted to produce a constitutively active form of Tec kinase. Moreover, if the intramolecular interaction is important in regulation, then deletion of PRR site 2 in the full length protein might result in deregulation of the kinase activity of Tec. Should an interaction occur between the SH3 domain and the SH2 kinase linker, analysis of the PRR mutants might reveal more information about the binding partners for the PRRSH3 region of Tec kinase. If an interaction between the PRR and the SH3 domain occurs in the full length protein, then important information on the regulation of this family of tyrosine kinases will be further elucidated.

REFERENCES

7.0 REFERENCES

Modern Physical Methods in Biochemistry Part B, A. Neuberger and L. M. Van. Deenan., eds. Amsterdam.

Abram, C. L., and Courtneidge, S. A. **(2000)**. Src Family Tyrosine Kinases and Growth Factor Signalling. *Experimental Cell Research* **254**, 1-13.

Afar, D. E., Park, H., Howell, B. W., Rawlings, D. J., Cooper, J., and Witte, O. N. **(1996)**. Regulation of Btk by Src family tyrosine kinases. *Mol Cell Biol* **16**, 3465-71.

Alexandropoulos, K., and Baltimore, D. **(1996)**. Coordinate activation of c-Src by SH3- and SH2-binding sites on a novel p130Cas-related protein, Sin. *Genes Dev* **10**, 1341-55.

Alexandropoulos, K., Cheng, G., and Baltimore, D. **(1995)**. Proline-rich sequences that bind to Src homology 3 domains with individual specificities. *Proc Natl Acad Sci U S A* **92**, 3110-4.

Andreotti, A. H., Bunnell, S. C., Feng, S., Berg, L. J., and Schreiber, S. L. **(1997)**. Regulatory intramolecular association in a tyrosine kinase of the Tec family. *Nature* **385**, 93-7.

Aoki, Y., Isselbacher, K. J., and Pillai, S. **(1994)**. Bruton tyrosine kinase is tyrosine phosphorylated and activated in pre-B lymphocytes and receptor-ligated B cells. *Proc Natl Acad Sci U S A* **91**, 10606-9.

Archer, S. J., Bax, A., Roberts, A. B., Sporn, M. B., Ogawa, Y., Piez, K. A., Weatherbee, J. A., Tsang, M. L., Lucas, R., Zheng, B. L., and *et al.* **(1993)**. Transforming growth factor beta 1: NMR signal assignments of the recombinant protein expressed and isotopically enriched using Chinese hamster ovary cells. *Biochemistry* **32**, 1152-63.

August, A., Sadra, A., Dupont, B., and Hanafusa, H. **(1997)**. Src-induced activation of inducible T cell kinase (ITK) requires phosphatidylinositol 3-kinase activity and the Pleckstrin homology domain of inducible T cell kinase. *Proc Natl Acad Sci U S A* **94**, 11227-32.

Bar-Sagi, D., Rotin, D., Batzer, A., Mandiyan, V., and Schlessinger, J. **(1993)**. SH3 domains direct cellular localization of signaling molecules. *Cell* **74**, 83-91.

Baraldi, E., Carugo, K. D., Hyvonen, M., Surdo, P. L., Riley, A. M., Potter, B. V., O'Brien, R., Ladbury, J. E., and Saraste, M. **(1999)**. Structure of the PH domain from Bruton's tyrosine kinase in complex with inositol 1,3,4,5-tetrakisphosphate. *Structure* **7**, 449-60.

Bartels, C., Xia, T.H., Billeter, P., Guntert, P. and Wuthrich, K. **(1995)**. The Program XEASY for Computer-supported NMR spectral analysis of Biological Macromolecules. *Journal of Biomolecular NMR* **5**, 1-10.

Blomberg, N., Baraldi, E., Nilges, M., and Saraste, M. **(1999)**. The PH superfold: a structural scaffold for multiple functions. *Trends Biochem Sci* **24**, 441-5.

Blomberg, N., Gabdoulline, R. R., Nilges, M., and Wade, R. C. **(1999)**. Classification of protein sequences by homology modeling and quantitative analysis of electrostatic similarity. *Proteins* **37**, 379-87.

Bolen, J. B. **(1993)**. Nonreceptor tyrosine protein kinases. *Oncogene* **8**, 2025-31.

Bolland, S., Pearse, R. N., Kurosaki, T., and Ravetch, J. V. **(1998)**. SHIP modulates immune receptor responses by regulating membrane association of Btk. *Immunity* **8**, 509-16.

Booker, G. W., Breeze, A. L., Downing, A. K., Panayotou, G., Gout, I., Waterfield, M. D., and Campbell, I. D. **(1992)**. Structure of an SH2 domain of the p85 alpha subunit of phosphatidylinositol-3-OH kinase. *Nature* **358**, 684-7.

Booker, G. W., Gout, I., Downing, A. K., Driscoll, P. C., Boyd, J., Waterfield, M. D., and Campbell, I. D. **(1993)**. Solution structure and ligand-binding site of the SH3 domain of the p85 alpha subunit of phosphatidylinositol 3-kinase. *Cell* **73**, 813-22.

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

Brand, L., and Witholt, B. **(1967)**. Fluorescence Measurements. In Enzyme Structure-Methods in enzymology **11**, Hirs, C. H. W. ed.: Academic Press, pp. 776-856.

Briggs, S. D., Sharkey, M., Stevenson, M., and Smithgall, T. E. **(1997)**. SH3-mediated Hck tyrosine kinase activation and fibroblast transformation by the Nef protein of HIV-1. *J Biol Chem* **272**, 17899-902.

Brunger, A. T. **(1991)**. Solution of a Fab (26-10)/digoxin complex by generalized molecular replacement. *Acta Crystallogr A* **47**, 195-204.

Bubeck Wardenburg, J., Fu, C., Jackmani, J. K., Flotow, H., Wilkinson, S. E., Williams, D. H., Johnson, R., Kong, G., Chan, A. C., and Findelli, P. R. **(1996)**. Phosphorylation of SLP-76 by the ZAP-70 Protein-tyrosine Kinase Is Required for T-cell Receptor Function. *J Biol Chem* **271**, 19641-19644.

Bunnell, S. C., Diehn, M., Yaffe, M. B., Findell, P. R., Cantley, L. C., and Berg, L. J. **(2000)**. Biochemical interactions integrating Itk with the T cell receptor-initiated signaling cascade. *J Biol Chem* **275**, 2219-30.

Bunnell, S. C., Henry, P. A., Kolluri, R., Kirchhausen, T., Rickles, R. J., and Berg, L. J. **(1996)**. Identification of Itk/Tsk Src homology 3 domain ligands. *J Biol Chem* **271**, 25646-56.

Carman, C. V., Barak, L. S., Chen, C., Liu-Chen, L. Y., Onorato, J. J., Kennedy, S. P., Caron, M. G., and Benovic, J. L. **(2000)**. Mutational analysis of G-betagamma and phospholipid interaction with G protein-coupled receptor kinase 2. *J Biol Chem* **275**, 10443-52.

Cartwright, C. A., Eckhart, W., Simon, S., and Kaplan, P. L. **(1987)**. Cell transformation by pp60c-src mutated in the carboxy-terminal regulatory domain. *Cell* **49**, 83-91.

Chan, V. W., Meng, F., Soriano, P., DeFranco, A. L., and Lowell, C. A. **(1997)**. Characterization of the B lymphocyte populations in Lyn-deficient mice and the role of Lyn in signal initiation and down-regulation. *Immunity* **7**, 69-81.

Cheadle, C., Ivashchenko, Y., South, V., Searfoss, G. H., French, S., Howk, R., Ricca, G. A., and Jaye, M. **(1994)**. Identification of a Src SH3 domain binding motif by screening a random phage display library. *J Biol Chem* **269**, 24034-9.

Chen, X., Vinkemeier, U., Zhao, Y., Jeruzalmi, D., Darnell, J. E., Jr., and Kuriyan, J. **(1998)**. Crystal structure of a tyrosine phosphorylated STAT-1 dimer bound to DNA. *Cell* **93**, 827-39.

Cheng, G., Ye, Z. S., and Baltimore, D. **(1994)**. Binding of Bruton's tyrosine kinase to Fyn, Lyn, or Hck through a Src homology 3 domain-mediated interaction. *Proc Natl Acad Sci U S A* **91**, 8152-5.

Ching, K. A., Kawakami, Y., Kawakami, T., and Tsoukas, C. D. **(1999)**. Emt/Itk associates with activated TCR complexes: role of the pleckstrin homology domain. *J Immunol* **163**, 6006-13.

Ciccarelli, F. D., Acciarito, A., and Alberti, S. **(2000)**. Large and diverse numbers of human diseases with HIKE mutations. *Hum Mol Genet* **9**, 1001-7.

Cicchetti, P., Mayer, B. J., Thiel, G., and Baltimore, D. **(1992)**. Identification of a protein that binds to the SH3 region of Abl and is similar to Bcr and GAP-rho. *Science* **257**, 803-6.

Cleland, J. L., Builder, S. E., Swartz, J. R., Winkler, M., Chang, J. Y., and Wang, D. I. **(1992)**. Polyethylene glycol enhanced protein refolding. *Biotechnology (N Y)* **10**, 1013-9.

Clowes, R. T., Crawford, A., Raine, A. R., Smith, B. O., and Laue, E. D. **(1995)**. Improved methods for structural studies of proteins using nuclear magnetic resonance spectroscopy. *Curr Opin Biotechnol* **6**, 81-8.

Cooper, J. A., Gould, K. L., Cartwright, C. A., and Hunter, T. **(1986)**. Tyr527 is phosphorylated in pp60c-src: implications for regulation. *Science* **231**, 1431-4.

Craik, D. J., and Wilce, J. A. **(1997)**. Studies of Protein-Ligand Interactions by NMR. In Protein NMR Techniques, D. G. Reid, ed. (New Jersey: Humana Press), pp. 195-232.

Dai, Z., and Pendergast, A. M. **(1995)**. Abi-2, a novel SH3-containing protein interacts with the c-Abl tyrosine kinase and modulates c-Abl transforming activity. *Genes Dev* **9**, 2569-82.

Das, B., Shu, X., Day, G. J., Han, J., Krishna, U. M., Falck, J. R., and Broek, D. **(2000)**. Control of intramolecular interactions between the pleckstrin homology and dbl homology domains of vav and sos1 regulates rac binding. *J Biol Chem* **275**, 15074-81.

den Hertog, J., Pals, C. E., Peppelenbosch, M. P., Tertoolen, L. G., de Laat, S. W., and Kruijer, W. **(1993)**. Receptor protein tyrosine phosphatase alpha activates pp60c-src and is involved in neuronal differentiation. *Embo J* **12**, 3789-98.

Dhe-Paganon, S., Ottinger, E. A., Nolte, R. T., Eck, M. J., and Shoelson, S. E. **(1999)**. Crystal structure of the pleckstrin homology-phosphotyrosine binding (PH-PTB) targeting region of insulin receptor substrate 1. *Proc Natl Acad Sci U S A* **96**, 8378-83.

Dingjan, G. M., Maas, A., Nawijn, M. C., Smit, L., Voerman, J. S., Grosveld, F., and Hendriks, R. W. **(1998)**. Severe B cell deficiency and disrupted splenic architecture in

transgenic mice expressing the E41K mutated form of Bruton's tyrosine kinase. *Embo J* **17**, 5309-20.

Downing, A. K., Driscoll, P. C., Gout, I., Salim, K., Zvelebil, M. J., and Waterfield, M. D. **(1994)**. Three-dimensional solution structure of the pleckstrin homology domain from dynamin. *Curr Biol* **4**, 884-91.

Eck, M. J., Atwell, S. K., Shoelson, S. E., and Harrison, S. C. **(1994)**. Structure of the regulatory domains of the Src-family tyrosine kinase Lck. *Nature* **368**, 764-9.

Eck, M. J., Dhe-Paganon, S., Trub, T., Nolte, R. T., and Shoelson, S. E. **(1996)**. Structure of the IRS-1 PTB domain bound to the juxtamembrane region of the insulin receptor. *Cell* **85**, 695-705.

Egan, S. E., Giddings, B. W., Brooks, M. W., Buday, L., Sizeland, A. M., and Weinberg, R. A. **(1993)**. Association of Sos Ras exchange protein with Grb2 is implicated in tyrosine kinase signal transduction and transformation. *Nature* **363**, 45-51.

Ellmeier, W., Jung, S., Sunshine, M. J., Hatam, F., Xu, Y., Baltimore, D., Mano, H., and Littman, D. R. **(2000)**. Severe B cell deficiency in mice lacking the tec kinase family members Tec and Btk. *J Exp Med* **192**, 1611-24.

Erpel, T., and Courtneidge, S. A. **(1995)**. Src family protein tyrosine kinases and cellular signal transduction pathways. *Curr Opin Cell Biol* **7**, 176-82.

Fantl, W. J., Johnson, D. E., and Williams, L. T. **(1993)**. Signalling by receptor tyrosine kinases. *Annu Rev Biochem* **62**, 453-81.

Fedorov, A. A., Fedorov, E., Gertler, F., and Almo, S. C. **(1999)**. Structure of EVH1, a novel proline-rich ligand-binding module involved in cytoskeletal dynamics and neural function. *Nat Struct Biol* **6**, 661-5.

Feng, S., Chen, J. K., Yu, H., Simon, J. A., and Schreiber, S. L. **(1994)**. Two binding orientations for peptides to the Src SH3 domain: development of a general model for SH3-ligand interactions. *Science* **266**, 1241-7.

Feng, S., Kasahara, C., Rickles, R. J., and Schreiber, S. L. **(1995)**. Specific interactions outside the proline-rich core of two classes of Src homology 3 ligands. *Proc Natl Acad Sci U S A* **92**, 12408-15.

Ferguson, K. M., Lemmon, M. A., Schlessinger, J., and Sigler, P. B. **(1994)**. Crystal structure at 2.2 A resolution of the pleckstrin homology domain from human dynamin. *Cell* **79**, 199-209.

Finan, P. M., Soames, C. J., Wilson, L., Nelson, D. L., Stewart, D. M., Truong, O., Hsuan, J. J., and Kellie, S. **(1996)**. Identification of regions of the Wiskott-Aldrich syndrome protein responsible for association with selected Src homology 3 domains. *J Biol Chem* **271**, 26291-5.

Fluckiger, A. C., Li, Z., Kato, R. M., Wahl, M. I., Ochs, H. D., Longnecker, R., Kinet, J. P., Witte, O. N., Scharenberg, A. M., and Rawlings, D. J. **(1998)**. Btk/Tec kinases regulate sustained increases in intracellular Ca2+ following B-cell receptor activation. *Embo J* **17**, 1973-85.

Foster, M. P., Wuttke, D. S., Clemens, K. R., Jahnke, W., Radhakrishnan, I., Tennant, L., Reymond, M., Chung, J., and Wright, P. E. **(1998)**. Chemical shift as a probe of molecular interfaces: NMR studies of DNA binding by the three amino-terminal zinc finger domains from transcription factor IIIA*Journal of Biomolecular NMR* **12**, 51-71.

Franz, W. M., Berger, P., and Wang, J. Y. **(1989)**. Deletion of an N-terminal regulatory domain of the c-abl tyrosine kinase activates its oncogenic potential. *Embo J* **8**, 137-47.

Fushman, D., Cahill, S., Lemmon, M. A., Schlessinger, J., and Cowburn, D. **(1995)**. Solution structure of pleckstrin homology domain of dynamin by heteronuclear NMR spectroscopy. *Proc Natl Acad Sci U S A* **92**, 816-20.

Gale, N. W., Kaplan, S., Lowenstein, E. J., Schlessinger, J., and Bar-Sagi, D. **(1993)**. Grb2 mediates the EGF-dependent activation of guanine nucleotide exchange on Ras. *Nature* **363**, 88-92.

Gonfloni, S., Frischknecht, F., Way, M., and Superti-Furga, G. **(1999)**. Leucine 255 of Src couples intramolecular interactions to inhibition of catalysis. *Nat Struct Biol* **6**, 760-4.

Gosser, Y. Q., Zheng, J., Overduin, M., Mayer, B. J., and Cowburn, D. **(1995)**. The solution structure of Abl SH3, and its relationship to SH2 in the SH(32) construct. *Structure* **3**, 1075-86.

Goudreau, N., Cornille, F., Duchesne, M., Parker, F., Tocque, B., Garbay, C., and Roques, B. P. **(1994)**. NMR structure of the N-terminal SH3 domain of GRB2 and its complex with a proline-rich peptide from Sos. *Nat Struct Biol* **1**, 898-907.

Gout, I., Dhand, R., Hiles, I. D., Fry, M. J., Panayotou, G., Das, P., Truong, O., Totty, N. F., Hsuan, J., Booker, G. W., and et al. **(1993)**. The GTPase dynamin binds to and is activated by a subset of SH3 domains. *Cell* **75**, 25-36.

Grodberg, J., and Dunn, J. J. **(1988)**. ompT encodes the Escherichia coli outer membrane protease that cleaves T7 RNA polymerase during purification. *J Bacteriol* **170**, 1245-53.

Gronenborn, A. M., and Clore, G. M. **(1990)**. Protein structure determination in solution by two-dimensional and three-dimensional nuclear magnetic resonance spectroscopy. *Anal Chem* **62**, 2-15.

Grzesiek, S., Bax, A., Clore, G. M., Gronenborn, A. M., Hu, J. S., Kaufman, J., Palmer, I., Stahl, S. J., and Wingfield, P. T. **(1996)**. The solution structure of HIV-1 Nef reveals an unexpected fold and permits delineation of the binding surface for the SH3 domain of Hck tyrosine protein kinase. *Nat Struct Biol* **3**, 340-5.

Guinamard, R., Fougereau, M., and Seckinger, P. **(1997)**. The SH3 domain of Bruton's tyrosine kinase interacts with Vav, Sam68 and EWS. *Scand J Immunol* **45**, 587-95.

Guruprasad, L., Dhanaraj, V., Timm, D., Blundell, T. L., Gout, I., and Waterfield, M. D. **(1995)**. The crystal structure of the N-terminal SH3 domain of Grb2. *J Mol Biol* **248**, 856-66.

Hajduk, P. J., Meadows, R. P., and Fesik, S. W. **(1999)**. NMR-based screening in drug discovery. *Q Rev Biophys* **32**, 211-40.

Hansson, H., Mattsson, P. T., Allard, P., Haapaniemi, P., Vihinen, M., Smith, C. I., and Hard, T. **(1998)**. Solution structure of the SH3 domain from Bruton's tyrosine kinase. *Biochemistry* **37**, 2912-24.

Hansson, H., Okoh, M. P., Smith, C. I. E., Vihinen, M., and Hard, T. **(2001)**. Intermolecular interactions between the SH3 domain and the proline-rich TH region of Bruton's Tyrosine kinase. *FEBS letters* **489**, 67-70.

Harlan, J. E., Hajduk, P. J., Yoon, H. S., and Fesik, S. W. **(1994)**. Pleckstrin homology domains bind to phosphatidylinositol-4,5-bisphosphate. *Nature* **371**, 168-70.

Hartl, F. U. **(1996)**. Molecular chaperones in cellular protein folding. *Nature* **381**, 571-9.

Haslam, R. J., Koide, H. B., and Hemmings, B. A. **(1993)**. Pleckstrin domain homology. *Nature* **363**, 309-10.

Heinrich, P.C., Behrmann, I., Muller-Newen, G., Schaper, F. and Graeve, L. **(1998)**. Interleukin-6-type cytokine signalling through the gp130/Jak/STAT pathway. *Biochem J*, **334**, 297-314.

Hensmann, M., Booker, G. W., Panayotou, G., Boyd, J., Linacre, J., Waterfield, M., and Campbell, I. D. **(1994)**. Phosphopeptide binding to the N-terminal SH2 domain of the p85 alpha subunit of PI 3'-kinase: a heteronuclear NMR study. *Protein Sci* **3**, 1020-30.

Hinds, M. G., and Norton, R. S. **(1994)**. NMR spectroscopy of peptides and proteins. *Methods Mol Biol* **36**, 131-54.

Hinshelwood, S., Lovering, R. C., Genevier, H. C., Levinsky, R. J., and Kinnon, C. **(1995)**. The protein defective in X-linked agammaglobulinemia, Bruton's tyrosine kinase, shows increased autophosphorylation activity in vitro when isolated from cells in which the B cell receptor has been cross-linked. *Eur J Immunol* **25**, 1113-6.

Hirai, H., and Varmus, H. E. **(1990)**. Site-directed mutagenesis of the SH2- and SH3-coding domains of c-src produces varied phenotypes, including oncogenic activation of p60c-src. *Mol Cell Biol* **10**, 1307-18.

Hiroaki, H., Klaus, W., and Senn, H. **(1996)**. Determination of the solution structure of the SH3 domain of human p56 Lck tyrosine kinase. *J Biomolecular NMR* **8**, 105-22.

Holzinger, A., Phillips, K. S., and Weaver, T. E. **(1996)**. Single-step purification/solubilization of recombinant proteins: application to surfactant protein B. *Biotechniques* **20**, 804-6, 808.

Horita, D. A., Baldisseri, D. M., Zhang, W., Altieri, A. S., Smithgall, T. E., Gmeiner, W. H., and Byrd, R. A. **(1998)**. Solution structure of the human Hck SH3 domain and identification of its ligand binding site. *J Mol Biol* **278**, 253-65.

Howard, M. J. **(1998)**. Protein NMR spectroscopy. *Curr Biol* **8**, R331-3.

Hyvonen, M., and Saraste, M. **(1997)**. Structure of the PH domain and Btk motif from Bruton's tyrosine kinase: molecular explanations for X-linked agammaglobulinaemia. *Embo J* **16**, 3396-404.

Imai, K., Nonoyama, S., Miki, H., Morio, T., Fukami, K., Zhu, Q., Aruffo, A., Ochs, H. D., Yata, J., and Takenawa, T. **(1999)**. The pleckstrin homology domain of the Wiskott-Aldrich syndrome protein is involved in the organization of actin cytoskeleton. *Clin Immunol* **92**, 128-37.

Imamoto, A., and Soriano, P. **(1993)**. Disruption of the csk gene, encoding a negative regulator of Src family tyrosine kinases, leads to neural tube defects and embryonic lethality in mice. *Cell* **73**, 1117-24.

Johnson, M. L., Correia, J. J., Yphantis, D. A., and Halvorson, H. R. **(1981)**. Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques. *Biophys J* **36**, 575-88.

Karin, M. **(1992)**. Signal transduction from cell surface to nucleus in development and disease. *Faseb J* **6**, 2581-90.

Kawakami, Y., Yao, L., Tashiro, M., Gibson, S., Mills, G. B., and Kawakami, T. **(1995)**. Activation and interaction with protein kinase C of a cytoplasmic tyrosine kinase, Itk/Tsk/Emt, on Fc epsilon RI cross-linking on mast cells. *J Immunol* **155**, 3556-62.

Kay, B. K., Williamson, M. P., and Sudol, M. **(2000)**. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *Faseb J* **14**, 231-41.

Kigawa, T., Muto, Y., and Yokoyama, S. **(1995)**. Cell-free synthesis and amino acid-selective stable isotope labeling of proteins for NMR analysis. *J Biomolecular NMR* **6**, 129-34.

Kitanaka, A., Mano, H., Conley, M. E., and Campana, D. **(1998)**. Expression and activation of the nonreceptor tyrosine kinase Tec in human B cells. *Blood* **91**, 940-8.

Kmiecik, T. E., and Shalloway, D. **(1987)**. Activation and suppression of pp60c-src transforming ability by mutation of its primary sites of tyrosine phosphorylation. *Cell* **49**, 65-73.

Koch, C. A., Anderson, D., Moran, M. F., Ellis, C., and Pawson, T. **(1991)**. SH2 and SH3 domains: elements that control interactions of cytoplasmic signaling proteins. *Science* **252**, 668-74.

Kohda, D., Hatanaka, H., Odaka, M., Mandiyan, V., Ullrich, A., Schlessinger, J., and Inagaki, F. **(1993)**. Solution structure of the SH3 domain of phospholipase C-gamma. *Cell* **72**, 953-60.

Koradi, R., Billeter, M., and Wuthrich, K. **(1996)**. MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph* **14**, 51-5, 29-32.

Kraulis, P. **(1991)**. MOLSCRIPT: A Program to Produce Both Detailed and Schematic Plots of Protein Structures. *Journal of Applied Crystallography* **24**, 946-950.

Kurosaki, T., and Kurosaki, M. **(1997)**. Transphosphorylation of Bruton's tyrosine kinase on tyrosine 551 is critical for B cell antigen receptor function. *J Biol Chem* **272**, 15595-8.

Kurosaki, T., and Tsukada, S. **(2000)**. BLNK: connecting Syk and Btk to calcium signals. *Immunity* **12**, 1-5.

Kypta, R. M., Goldberg, Y., Ulug, E. T., and Courtneidge, S. A. **(1990)**. Association between the PDGF receptor and members of the src family of tyrosine kinases. *Cell* **62**, 481-92.

Laffargue, M., Monnereau, L., Tuech, J., Ragab, A., Ragab-Thomas, J., Payrastre, B., Raynal, P., and Chap, H. **(1997)**. Integrin-dependent tyrosine phoshorylation and cytoskeletal translocation of Tec in thrombin-activated platelets. *Biochem Biophys Res Commun* **238**, 247-51.

Langhans-Rajasekaran, S. A., Wan, Y., and Huang, X. Y. **(1995)**. Activation of Tsk and Btk tyrosine kinases by G protein beta gamma subunits. *Proc Natl Acad Sci U S A* **92**, 8601-5.

Lee, C. H., Leung, B., Lemmon, M. A., Zheng, J., Cowburn, D., Kuriyan, J., and Saksela, K. **(1995)**. A single amino acid in the SH3 domain of Hck determines its high affinity and specificity in binding to HIV-1 Nef protein. *Embo J* **14**, 5006-15.

Lee, C. H., Saksela, K., Mirza, U. A., Chait, B. T., and Kuriyan, J. **(1996)**. Crystal structure of the conserved core of HIV-1 Nef complexed with a Src family SH3 domain. *Cell* **85**, 931-42.

Li, T., Tsukada, S., Satterthwaite, A., Havlik, M. H., Park, H., Takatsu, K., and Witte, O. N. **(1995)**. Activation of Bruton's tyrosine kinase (BTK) by a point mutation in its pleckstrin homology (PH) domain. *Immunity* **2**, 451-60.

Li, Z., Wahl, M. I., Eguinoa, A., Stephens, L. R., Hawkins, P. T., and Witte, O. N. **(1997)**. Phosphatidylinositol 3-kinase-gamma activates Bruton's tyrosine kinase in concert with Src family kinases. *Proc Natl Acad Sci U S A* **94**, 13820-5.

Lian, L.-Y., and Roberts, G. C. K. **(1993)**. Effects of Chemical exchange on NMR Spectra. In NMR of Macromolecules, G. C. K. Roberts, ed.: IRL Press at Oxford University Press), pp. 153-182.

Liao, X. C., and Littman, D. R. **(1995)**. Altered T cell receptor signaling and disrupted T cell development in mice lacking Itk. *Immunity* **3**, 757-69.

Lim, W. A., Richards, F. M., and Fox, R. O. **(1994)**. Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains. *Nature* **372**, 375-9.

Linge, J. P., and Nilges, M. **(1999)**. Influence of non-bonded parameters on the quality of NMR structures: A new force field for NMR structure calculation. *Journal of Biomolecular NMR* **13**, 51-59.

Liu, X., Brodeur, S. R., Gish, G., Songyang, Z., Cantley, L. C., Laudano, A. P., and Pawson, T. **(1993)**. Regulation of c-Src tyrosine kinase activity by the Src SH2 domain. *Oncogene* **8**, 1119-26.

Lowell, C.A., Niwa, M., Soriano, P. and Varmus, H.E. **(1996)**. Deficiency of the Hck and Src tyrosine kinases results in extreme levels of extramedullary hematopoiesis. *Blood*, **87**, 1780-1792.

Lowell, C. A., and Soriano, P. **(1996)**. Knockouts of Src-family kinases: stiff bones, wimpy T cells, and bad memories. *Genes Dev* **10**, 1845-57.

MacArthur, M. W., Driscoll, P. C., and Thornton, J. M. **(1994)**. NMR and crystallography-complementary approaches to structure determination. *Trends Biotechnol* **12**, 149-53.

Machide, M., Mano, H., and Todokoro, K. **(1995)**. Interleukin 3 and erythropoietin induce association of Vav with Tec kinase through Tec homology domain. *Oncogene* **11**, 619-25.

Macias, M. J., Musacchio, A., Ponstingl, H., Nilges, M., Saraste, M., and Oschkinat, H. **(1994)**. Structure of the pleckstrin homology domain from beta-spectrin. *Nature* **369**, 675-7.

Mahajan, S., Fargnoli, J., Burkhardt, A. L., Kut, S. A., Saouaf, S. J., and Bolen, J. B. **(1995)**. Src family protein tyrosine kinases induce autoactivation of Bruton's tyrosine kinase. *Mol Cell Biol* **15**, 5304-11.

Mahoney, N. M., Rozwarski, D. A., Fedorov, E., Fedorov, A. A., and Almo, S. C. **(1999)**. Profilin binds proline-rich ligands in two distinct amide backbone orientations. *Nat Struct Biol* **6**, 666-71.

Maignan, S., Guilloteau, J. P., Fromage, N., Arnoux, B., Becquart, J., and Ducruix, A. **(1995)**. Crystal structure of the mammalian Grb2 adaptor. *Science* **268**, 291-3.

Manninen, A., Hiipakka, M., Vihinen, M., Lu, W., Mayer, B. J., and Saksela, K. **(1998)**. SH3-Domain binding function of HIV-1 Nef is required for association with a PAK-related kinase. *Virology* **250**, 273-82.

Mano, H., Ishikawa, F., Nishida, J., Hirai, H., and Takaku, F. **(1990)**. A novel protein-tyrosine kinase, *tec*, is preferentially expressed in liver. *Oncogene* **5**, 1781-1786.

Mano, H., Mano, K., Tang, B., Koehler, M., Yi, T., Gilbert, D. J., Jenkins, N. A., Copeland, N. G., and Ihle, J. N. **(1993)**. Expression of a novel form of Tec kinase in hematopoietic cells and mapping of the gene to chromosome 5 near Kit. *Oncogene* **8**, 417-24.

Mano, H., Ohya, K., Miyazato, A., Yamashita, Y., Ogawa, W., Inazawa, J., Ikeda, U., Shimada, K., Hatake, K., Kasuga, M., Ozawa, K., and Kajigaya, S. **(1998)**. Grb10/GrbIR as an in vivo substrate of Tec tyrosine kinase. *Genes Cells* **3**, 431-41.

Mano, H., Sato, K., Yazaki, Y., and Hirai, H. **(1994)**. Tec protein-tyrosine kinase directly associates with Lyn protein-tyrosine kinase through its N-terminal unique domain. *Oncogene* **9**, 3205-11.

Mano, H., Yamashita, Y., Sato, K., Yazaki, Y., and Hirai, H. **(1995)**. Tec protein-tyrosine kinase is involved in interleukin-3 signaling pathway. *Blood* **85**, 343-50.

Marengere, L. E., Okkenhaug, K., Clavreul, A., Couez, D., Gibson, S., Mills, G. B., Mak, T. W., and Rottapel, R. **(1997)**. The SH3 domain of Itk/Emt binds to proline-rich sequences in the cytoplasmic domain of the T cell costimulatory receptor CD28. *J Immunol* **159**, 3220-9.

Maroun, C. R., Moscatello, D. K., Naujokas, M. A., Holgado-Madruga, M., Wong, A. J., and Park, M. **(1999)**. A conserved inositol phospholipid binding site within the pleckstrin

homology domain of the Gab1 docking protein is required for epithelial morphogenesis. *J Biol Chem* **274**, 31719-26.

Marshall, A. J., Niiro, H., Lerner, C. G., Yun, T. J., Thomas, S., Disteche, C. M., and Clark, E. A. **(2000)**. A novel B lymphocyte-associated adaptor protein, Bam32, regulates antigen receptor signaling downstream of phosphatidylinositol 3-kinase. *J Exp Med* **191**, 1319-32.

Matsuda, T., Takahashi-Tezuka, M., Fukada, T., Okuyama, Y., Fujitani, Y., Tsukada, S., Mano, H., Hirai, H., Witte, O. N., and Hirano, T. **(1995)**. Association and activation of Btk and Tec tyrosine kinases by gp130, a signal transducer of the interleukin-6 family of cytokines. *Blood* **85**, 627-33.

Matsushita, M., Yamadori, T., Kato, S., Takemoto, Y., Inazawa, J., Baba, Y., Hashimoto, S., Sekine, S., Arai, S., Kunikata, T., Kurimoto, M., Kishimoto, T., and Tsukada, S. **(1998)**. Identification and characterization of a novel SH3-domain binding protein, Sab, which preferentially associates with Bruton's tyrosine kinase (BtK). *Biochem Biophys Res Commun* **245**, 337-43.

Mayer, B. J., and Baltimore, D. **(1994)**. Mutagenic analysis of the roles of SH2 and SH3 domains in regulation of the Abl tyrosine kinase. *Mol Cell Biol* **14**, 2883-94.

Mayer, B. J., Ren, R., Clark, K. L., and Baltimore, D. **(1993)**. A putative modular domain present in diverse signaling proteins. *Cell* **73**, 629-30.

McEvoy, M. M., de la Cruz, A. F., and Dahlquist, F. W. **(1997)**. Large modular proteins by NMR. *Nat Struct Biol* **4**, 9.

Merkel, A. L., Atmosukarto, II, Stevens, K., Rathjen, P. D., and Booker, G. W. **(1999)**. Splice variants of the mouse Tec gene are differentially expressed in vivo. *Cytogenet Cell Genet* **84**, 132-9.

Merritt, E. A., and J., B. D. **(1997)**. Raster3D: Photorealistic Molecular Graphics. *Methods in Enzymology* **277**, 505-524.

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

Miki, H., Miura, K., and Takenawa, T. **(1996)**. N-WASP, a novel actin-depolymerizing protein, regulates the cortical cytoskeletal rearrangement in a PIP2-dependent manner downstream of tyrosine kinases. *Embo J* **15**, 5326-35.

Miller, J. H. **(1972)**. *Experiments in Molecular Genetics*
Cold Spring Harbour Laboratory: Cold Spring Harbour New York).

Mitchell, D. A., Marshall, T. K., and Deschenes, R. J. **(1993)**. Vectors for the inducible overexpression of glutathione S-transferase fusion proteins in yeast. *Yeast* **9**, 715-22.

Moarefi, I., LaFevre-Bernt, M., Sicheri, F., Huse, M., Lee, C. H., Kuriyan, J., and Miller, W. T. **(1997)**. Activation of the Src-family tyrosine kinase Hck by SH3 domain displacement. *Nature* **385**, 650-3.

Moran, M. F., Koch, C. A., Anderson, D., Ellis, C., England, L., Martin, G. S., and Pawson, T. **(1990)**. Src homology region 2 domains direct protein-protein interactions in signal transduction. *Proc Natl Acad Sci U S A* **87**, 8622-6.

Morrogh, L. M., Hinshelwood, S., Costello, P., Cory, G. O., and Kinnon, C. **(1999)**. The SH3 domain of Bruton's tyrosine kinase displays altered ligand binding properties when auto-phosphorylated in vitro. *Eur J Immunol* **29**, 2269-79.

Morton, C. J., and Campbell, I. D. **(1994)**. SH3 domains. Molecular 'Velcro'. *Curr Biol* **4**, 615-7.

Morton, C. J., Pugh, D. J., Brown, E. L., Kahmann, J. D., Renzoni, D. A., and Campbell, I. D. **(1996)**. Solution structure and peptide binding of the SH3 domain from human Fyn. *Structure* **4**, 705-14.

Morton, T. A., Myszka, D. G., and Chaiken, I. M. **(1995)**. Interpreting complex binding kinetics from optical biosensors: a comparison of analysis by linearization, the integrated rate equation, and numerical integration. *Anal Biochem* **227**, 176-85.

Musacchio, A., Gibson, T., Lehto, V. P., and Saraste, M. **(1992)**. SH3--an abundant protein domain in search of a function. *FEBS Lett* **307**, 55-61.

Musacchio, A., Gibson, T., Rice, P., Thompson, J., and Saraste, M. **(1993)**. The PH domain: a common piece in the structural patchwork of signalling proteins. *Trends Biochem Sci* **18**, 343-8.

Musacchio, A., Noble, M., Pauptit, R., Wierenga, R., and Saraste, M. **(1992)**. Crystal structure of a Src-homology 3 (SH3) domain. *Nature* **359**, 851-5.

Musacchio, A., Saraste, M., and Wilmanns, M. **(1994)**. High-resolution crystal structures of tyrosine kinase SH3 domains complexed with proline-rich peptides. *Nat Struct Biol* **1**, 546-51.

Myszka, D. G. **(1997)**. Kinetic analysis of macromolecular interactions using surface plasmon resonance biosensors. *Curr Opin Biotechnol* **8**, 50-7.

Myung, P. S., Boerthe, N. J., and Koretzky, G. A. **(2000)**. Adapter proteins in lymphocyte antigen-receptor signaling. *Curr Opin Immunol* **12**, 256-66.

Nada, S., Yagi, T., Takeda, H., Tokunaga, T., Nakagawa, H., Ikawa, Y., Okada, M., and Aizawa, S. **(1993)**. Constitutive activation of Src family kinases in mouse embryos that lack Csk. *Cell* **73**, 1125-35.

Nam, H. J., Haser, W. G., Roberts, T. M., and Frederick, C. A. **(1996)**. Intramolecular interactions of the regulatory domains of the Bcr-Abl kinase reveal a novel control mechanism. *Structure* **4**, 1105-14.

Neidhardt, F. C., Bloch, P. L., and Smith, D. F. **(1974)**. Culture Medium for Enterobacteria. *Journal of Bacteriology* **119**, 736-747.

Nilges, M., Macias, M. J., O'Donoghue, S. I., and Oschkinat, H. **(1997)**. Automated NOESY interpretation with ambiguous distance restraints: the refined NMR solution structure of the pleckstrin homology domain from beta-spectrin. *J Mol Biol* **269**, 408-22.

Nilges, M., Habezettl, J., Brunger, A. T. and Holak, T. A. **(1991)** Relaxation matrix refinement of the solution structure of squash trypsin inhibitor. *J Mol Biol* **219**, 499-510.

Nore, B. F., Vargas, L., Mohamed, A. J., Branden, L. J., Backesjo, C. M., Islam, T. C., Mattsson, P. T., Hultenby, K., Christensson, B., and Smith, C. I. **(2000)**. Redistribution of Bruton's tyrosine kinase by activation of phosphatidylinositol 3-kinase and Rho-family GTPases. *Eur J Immunol* **30**, 145-54.

Ohya, K., Kajigaya, S., Kitanaka, A., Yoshida, K., Miyazato, A., Yamashita, Y., Yamanaka, T., Ikeda, U., Shimada, K., Ozawa, K., and Mano, H. **(1999)**. Molecular cloning of a docking protein, BRDG1, that acts downstream of the Tec tyrosine kinase. *Proc Natl Acad Sci U S A* **96**, 11976-81.

Ohya, K., Kajigaya, S., Yamashita, Y., Miyazato, A., Hatake, K., Miura, Y., Ikeda, U., Shimada, K., Ozawa, K., and Mano, H. **(1997)**. SOCS-1/JAB/SSI-1 can bind to and suppress Tec protein-tyrosine kinase. *J Biol Chem* **272**, 27178-82.

Okoh, M. P., and Vihinen, M. **(1999)**. Pleckstrin homology domains of tec family protein kinases. *Biochem Biophys Res Commun* **265**, 151-7.

Owen, D. J., Wigge, P., Vallis, Y., Moore, J. D., Evans, P. R., and McMahon, H. T. **(1998)**. Crystal structure of the amphiphysin-2 SH3 domain and its role in the prevention of dynamin ring formation. *Embo J* **17**, 5273-85.

Panayotou, G., Gish, G., End, P., Truong, O., Gout, I., Dhand, R., Fry, M. J., Hiles, I., Pawson, T., and Waterfield, M. D. **(1993)**. Interactions between SH2 domains and tyrosine-phosphorylated platelet-derived growth factor beta-receptor sequences: analysis of kinetic parameters by a novel biosensor-based approach. *Mol Cell Biol* **13**, 3567-76.

Park, H., Wahl, M. I., Afar, D. E., Turck, C. W., Rawlings, D. J., Tam, C., Scharenberg, A. M., Kinet, J. P., and Witte, O. N. **(1996)**. Regulation of Btk function by a major autophosphorylation site within the SH3 domain. *Immunity* **4**, 515-25.

Parrini, M. C., and Mayer, B. J. **(1999)**. Engineering temperature-sensitive SH3 domains. *Chem Biol* **6**, 679-87.

Pascal, S. M., Singer, A. U., Gish, G., Yamazaki, T., Shoelson, S. E., Pawson, T., Kay, L. E., and Forman-Kay, J. D. **(1994)**. Nuclear magnetic resonance structure of an SH2 domain of phospholipase C-gamma 1 complexed with a high affinity binding peptide. *Cell* **77**, 461-72.

Patel, H. V., Tzeng, S. R., Liao, C. Y., Chen, S. H., and Cheng, J. W. **(1997)**. SH3 domain of Bruton's tyrosine kinase can bind to proline-rich peptides of TH domain of the kinase and p120cbl. *Proteins* **29**, 545-52.

Pawson, T., Olivier, P., Rozakis-Adcock, M., McGlade, J., and Henkemeyer, M. **(1993)**. Proteins with SH2 and SH3 domains couple receptor tyrosine kinases to intracellular signalling pathways. *Philos Trans R Soc Lond B Biol Sci* **340**, 279-85.

Perkins, S. J. **(1986)**. Protein volumes and hydration effects. The calculations of partial specific volumes, neutron scattering matchpoints and 280-nm absorption coefficients for proteins and glycoproteins from amino acid sequences. *Eur J Biochem* **157**, 169-80.

Pisabarro, M. T., and Serrano, L. **(1996)**. Rational design of specific high-affinity peptide ligands for the Abl-SH3 domain. *Biochemistry* **35**, 10634-40.

Pitcher, J. A., Touhara, K., Payne, E. S., and Lefkowitz, R. J. **(1995)**. Pleckstrin homology domain-mediated membrane association and activation of the beta-adrenergic receptor kinase requires coordinate interaction with G beta gamma subunits and lipid. *J Biol Chem* **270**, 11707-10.

Piwnica-Worms, H., Saunders, K. B., Roberts, T. M., Smith, A. E., and Cheng, S. H. **(1987)**. Tyrosine phosphorylation regulates the biochemical and biological properties of pp60c-src. *Cell* **49**, 75-82.

Prasad, K. V., Janssen, O., Kapeller, R., Raab, M., Cantley, L. C., and Rudd, C. E. **(1993)**. Src-homology 3 domain of protein kinase p59fyn mediates binding to phosphatidylinositol 3-kinase in T cells. *Proc Natl Acad Sci U S A* **90**, 7366-70.

Prehoda, K. E., Lee, D. J., and Lim, W. A. **(1999)**. Structure of the enabled/VASP homology 1 domain-peptide complex: a key component in the spatial control of actin assembly. *Cell* **97**, 471-80.

Primrose, W. U. **(1993)**. Sample Preparation. In NMR of Macromolecules, R. G. C. K., ed.: Oxford University Press).

Punt, C. J. **(1992)**. Regulation of hematopoietic cell function by protein tyrosine kinase-encoding oncogenes, a review. *Leuk Res* **16**, 551-9.

Qiu, Y., Robinson, D., Pretlow, T. G., and Kung, H. J. **(1998)**. Etk/Bmx, a tyrosine kinase with a pleckstrin-homology domain, is an effector of phosphatidylinositol 3'-kinase and is involved in interleukin 6-induced neuroendocrine differentiation of prostate cancer cells. *Proc Natl Acad Sci U S A* **95**, 3644-9.

Rameh, L. E., Arvidsson, A., Carraway, K. L., 3rd, Couvillon, A. D., Rathbun, G., Crompton, A., VanRenterghem, B., Czech, M. P., Ravichandran, K. S., Burakoff, S. J., Wang, D. S., Chen, C. S., and Cantley, L. C. **(1997)**. A comparative analysis of the phosphoinositide binding specificity of pleckstrin homology domains. *J Biol Chem* **272**, 22059-66.

Ramos, A., Kelly, G., Hollingworth, D., Pastore, A., and Frenkiel, T. **(2000)**. Mapping the interface of protein-nucleic acid complexes using cross saturation. *Journal of the American Chemical Society* **122**, 11311-11314.

Rawlings, D. J. **(1999)**. Bruton's tyrosine kinase controls a sustained calcium signal essential for B lineage development and function. *Clin Immunol* **91**, 243-53.

Rawlings, D. J., Saffran, D. C., Tsukada, S., Largaespada, D. A., Grimaldi, J. C., Cohen, L., Mohr, R. N., Bazan, J. F., Howard, M., Copeland, N. G., and *et al.* **(1993)**. Mutation of unique region of Bruton's tyrosine kinase in immunodeficient XID mice. *Science* **261**, 358-61.

Rawlings, D. J., Scharenberg, A. M., Park, H., Wahl, M. I., Lin, S., Kato, R. M., Fluckiger, A. C., Witte, O. N., and Kinet, J. P. **(1996)**. Activation of BTK by a phosphorylation mechanism initiated by SRC family kinases. *Science* **271**, 822-5.

Rawlings, D. J., and Witte, O. N. **(1994)**. Bruton's tyrosine kinase is a key regulator in B-cell development. *Immunol Rev* **138**, 105-19.

Ren, R., Mayer, B. J., Cicchetti, P., and Baltimore, D. **(1993)**. Identification of a ten-amino acid proline-rich SH3 binding site. *Science* **259**, 1157-61.

Rickles, R. J., Botfield, M. C., Weng, Z., Taylor, J. A., Green, O. M., Brugge, J. S., and Zoller, M. J. **(1994)**. Identification of Src, Fyn, Lyn, PI3K and Abl SH3 domain ligands using phage display libraries. *Embo J* **13**, 5598-604.

Rickles, R. J., Botfield, M. C., Zhou, X. M., Henry, P. A., Brugge, J. S., and Zoller, M. J. **(1995)**. Phage display selection of ligand residues important for Src homology 3 domain binding specificity. *Proc Natl Acad Sci U S A* **92**, 10909-13.

Roberts, G. C. K. **(1993)**. Introduction. In NMR of Macromolecules, G. C. K. Roberts, ed. (New York: oxford University Press), pp. 1-5.

Robinson, D., He, F., Pretlow, T., and Kung, H. J. **(1996)**. A tyrosine kinase profile of prostate carcinoma. *Proc Natl Acad Sci U S A* **93**, 5958-62.

Rodrigues, G. A., and Park, M. **(1994)**. Oncogenic activation of tyrosine kinases. *Curr Opin Genet Dev* **4**, 15-24.

Rozakis-Adcock, M., Fernley, R., Wade, J., Pawson, T., and Bowtell, D. **(1993)**. The SH2 and SH3 domains of mammalian Grb2 couple the EGF receptor to the Ras activator mSos1. *Nature* **363**, 83-5.

Sabe, H., Hata, A., Okada, M., Nakagawa, H., and Hanafusa, H. **(1994)**. Analysis of the binding of the Src homology 2 domain of Csk to tyrosine-phosphorylated proteins in the suppression and mitotic activation of c-Src. *Proc Natl Acad Sci U S A* **91**, 3984-8.

Sabe, H., Knudsen, B., Okada, M., Nada, S., Nakagawa, H., and Hanafusa, H. **(1992)**. Molecular cloning and expression of chicken C-terminal Src kinase: lack of stable association with c-Src protein. *Proc Natl Acad Sci U S A* **89**, 2190-4.

Saharinen, P., Ekman, N., Sarvas, K., Parker, P., Alitalo, K., and Silvennoinen, O. **(1997)**. The Bmx tyrosine kinase induces activation of the Stat signaling pathway, which is specifically inhibited by protein kinase Cdelta. *Blood* **90**, 4341-53.

Saraste, M., and Musacchio, A. **(1994)**. Backwards and forwards binding. *Nat Struct Biol* **1**, 835-7.

Schaeffer, E. M., Debnath, J., Yap, G., McVicar, D., Liao, X. C., Littman, D. R., Sher, A., Varmus, H. E., Lenardo, M. J., and Schwartzberg, P. L. **(1999)**. Requirement for Tec kinases Rlk and Itk in T cell receptor signaling and immunity. *Science* **284**, 638-41.

Schaeffer, E. M., and Schwartzberg, P. L. **(2000)**. Tec family kinases in lymphocyte signaling and function. *Curr Opin Immunol* **12**, 282-8.

Scharenberg, A. M., El-Hillal, O., Fruman, D. A., Beitz, L. O., Li, Z., Lin, S., Gout, I., Cantley, L. C., Rawlings, D. J., and Kinet, J. P. **(1998)**. Phosphatidylinositol-3,4,5-trisphosphate (PtdIns-3,4,5-P3)/Tec kinase-dependent calcium signaling pathway: a target for SHIP-mediated inhibitory signals. *Embo J* **17**, 1961-72.

Schneider, H., Guerette, B., Guntermann, C., and Rudd, C. E. **(2000)**. Resting lymphocyte kinase (Rlk/Txk) targets lymphoid adaptor SLP-76 in the cooperative activation of interleukin-2 transcription in T-cells. *J Biol Chem* **275**, 3835-40.

Schuster, T. M., and Toedt, J. M. **(1996)**. New revolutions in the evolution of analytical ultracentrifugation. *Curr Opin Struct Biol* **6**, 650-8.

Seidel-Dugan, C., Meyer, B. E., Thomas, S. M., and Brugge, J. S. **(1992)**. Effects of SH2 and SH3 deletions on the functional activities of wild-type and transforming variants of c-Src. *Mol Cell Biol* **12**, 1835-45.

Shan, X., and Wange, R. L. **(1999)**. Itk/Emt/Tsk activation in response to CD3 cross-linking in Jurkat T cells requires ZAP-70 and Lat and is independent of membrane recruitment. *J Biol Chem* **274**, 29323-30.

Shaw, G. **(1993)**. Identification of novel pleckstrin homology (PH) domains provides a hypothesis for PH domain function. *Biochem Biophys Res Commun* **195**, 1145-51.

Shaw, G. **(1996)**. The pleckstrin homology domain: an intriguing multifunctional protein module. *Bioessays* **18**, 35-46.

Shi, Y., Alin, K., and Goff, S. P. **(1995)**. Abl-interactor-1, a novel SH3 protein binding to the carboxy-terminal portion of the Abl protein, suppresses v-abl transforming activity. *Genes Dev* **9**, 2583-97.

Sicheri, F., Moarefi, I., and Kuriyan, J. **(1997)**. Crystal structure of the Src family tyrosine kinase Hck. *Nature* **385**, 602-9.

Smart, J. E., Oppermann, H., Czernilofsky, A. P., Purchio, A. F., Erikson, R. L., and Bishop, J. M. **(1981)**. Characterization of sites for tyrosine phosphorylation in the transforming protein of Rous sarcoma virus (pp60v-src) and its normal cellular homologue (pp60c-src). *Proc Natl Acad Sci U S A* **78**, 6013-7.

Sommers, C. L., Huang, K., Shores, E. W., Grinberg, A., Charlick, D. A., Kozak, C. A., and Love, P. E. **(1995)**. Murine txk: a protein tyrosine kinase gene regulated by T cell activation. *Oncogene* **11**, 245-51.

Songyang, Z., Gish, G., Mbamalu, G., Pawson, T., and Cantley, L. C. **(1995)**. A single point mutation switches the specificity of group III Src homology (SH) 2 domains to that of group I SH2 domains. *J Biol Chem* **270**, 26029-32.

Sparks, A. B., Adey, N. B., Quilliam, L. A., Thorn, J. M., and Kay, B. K. **(1995)**. Screening phage-displayed random peptide libraries for SH3 ligands. *Methods Enzymol* **255**, 498-509.

Sparks, A. B., Quilliam, L. A., Thorn, J. M., Der, C. J., and Kay, B. K. **(1994)**. Identification and characterization of Src SH3 ligands from phage-displayed random peptide libraries. *J Biol Chem* **269**, 23853-6.

Stempfer, G., Holl-Neugebauer, B., and Rudolph, R. **(1996)**. Improved refolding of an immobilized fusion protein. *Nat Biotechnol* **14**, 329-34.

Su, Y. W., Zhang, Y., Schweikert, J., Koretzky, G. A., Reth, M., and Wienands, J. **(1999)**. Interaction of SLP adaptors with the SH2 domain of Tec family kinases. *Eur J Immunol* **29**, 3702-11.

Superti-Furga, G., and Courtneidge, S. A. **(1995)**. Structure-function relationships in Src family and related protein tyrosine kinases. *Bioessays* **17**, 321-30.

Superti-Furga, G., Fumagalli, S., Koegl, M., Courtneidge, S. A., and Draetta, G. **(1993)**. Csk inhibition of c-Src activity requires both the SH2 and SH3 domains of Src. *Embo J* **12**, 2625-34.

Takahashi-Tezuka, M., Hibi, M., Fujitani, Y., Fukada, T., Yamaguchi, T., and Hirano, T. **(1997)**. Tec tyrosine kinase links the cytokine receptors to PI-3 kinase probably through JAK. *Oncogene* **14**, 2273-82.

Takata, M., and Kurosaki, T. **(1996)**. A role for Bruton's tyrosine kinase in B cell antigen receptor-mediated activation of phospholipase C-gamma 2. *J Exp Med* **184**, 31-40.

Takeya, T., and Hanafusa, H. **(1983)**. Structure and sequence of the cellular gene homologous to the RSV src gene and the mechanism for generating the transforming virus. *Cell* **32**, 881-90.

Tamagnone, L., Lahtinen, I., Mustonen, T., Virtaneva, K., Francis, F., Muscatelli, F., Alitalo, R., Smith, C. I., Larsson, C., and Alitalo, K. **(1994)**. BMX, a novel nonreceptor tyrosine kinase gene of the BTK/ITK/TEC/TXK family located in chromosome Xp22.2. *Oncogene* **9**, 3683-8.

Tang, B., Mano, H., Yi, T., and Ihle, J. N. **(1994)**. Tec kinase associates with c-kit and is tyrosine phosphorylated and activated following stem cell factor binding. *Mol Cell Biol* **14**, 8432-7.

Terasawa, H., Kohda, D., Hatanaka, H., Tsuchiya, S., Ogura, K., Nagata, K., Ishii, S., Mandiyan, V., Ullrich, A., Schlessinger, J., and et al. **(1994)**. Structure of the N-terminal SH3 domain of GRB2 complexed with a peptide from the guanine nucleotide releasing factor Sos. *Nat Struct Biol* **1**, 891-7.

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. *Nucleic Acids Res* **22**, 4673-80.

Timm, D., Salim, K., Gout, I., Guruprasad, L., Waterfield, M., and Blundell, T. **(1994)**. Crystal structure of the pleckstrin homology domain from dynamin. *Nat Struct Biol* **1**, 782-8.

Tomlinson, M. G., Kurosaki, T., Berson, A. E., Fujii, G. H., Johnston, J. A., and Bolen, J. B. **(1999)**. Reconstitution of Btk signaling by the atypical tec family tyrosine kinases Bmx and Txk. *J Biol Chem* **274**, 13577-85.

Torigoe, T., R, O. C., Fagard, R., Fischer, S., Santoli, D., and Reed, J. C. **(1992)**. Regulation of SRC-family protein tyrosine kinases by interleukins, IL-2, and IL-3. *Leukemia* **6**, 94S-97S.

Touhara, K., Inglese, J., Pitcher, J. A., Shaw, G., and Lefkowitz, R. J. **(1994)**. Binding of G protein beta gamma-subunits to pleckstrin homology domains. *J Biol Chem* **269**, 10217-20.

Touhara, K., Koch, W. J., Hawes, B. E., and Lefkowitz, R. J. **(1995)**. Mutational analysis of the pleckstrin homology domain of the beta-adrenergic receptor kinase. Differential effects on G beta gamma and phosphatidylinositol 4,5-bisphosphate binding. *J Biol Chem* **270**, 17000-5.

Toyoshima, K., Yamanashi, Y., Inoue, K., Semba, K., Yamamoto, T., and Akiyama, T. **(1992)**. Protein tyrosine kinases belonging to the src family. *Ciba Found Symp* **164**, 240-8; discussion 248-53.

Tsukada, S., Rawlings, D. J., and Witte, O. N. **(1994)**. Role of Bruton's tyrosine kinase in immunodeficiency. *Curr Opin Immunol* **6**, 623-30.

Tsukada, S., and Witte, O. N. **(1994)**. X-linked agammaglobulinemia and Bruton's tyrosine kinase. *Adv Exp Med Biol* **365**, 233-8.

Tsukube, H., Furuta, H., Odani, A., Takeda, Y., Kudo, Y., Inoue, Y., Liu, Y., Sakamoto, H., and Kimura, K. **(1996)**. Determination of Stability Constants. In *Comprehensive Supramolecular Chemistry*, pp. 425-482.

Tzeng, S. R., Lou, Y. C., Pai, M. T., Jain, M. L., and Cheng, J. W. **(2000)**. Solution structure of the human BTK SH3 domain complexed with a proline-rich peptide from p120cbl. *J Biomol NMR* **16**, 303-12.

Varnai, P., Rother, K. I., and Balla, T. **(1999)**. Phosphatidylinositol 3-kinase-dependent membrane association of the Bruton's tyrosine kinase pleckstrin homology domain visualized in single living cells. *J Biol Chem* **274**, 10983-9.

Venkataraman, C., Muthusamy, N., Muthukkumar, S., and Bondada, S. **(1998)**. Activation of lyn, blk, and btk but not syk in CD72-stimulated B lymphocytes. *J Immunol* **160**, 3322-9.

Vetrie, D., Vorechovsky, I., Sideras, P., Holland, J., Davies, A., Flinter, F., Hammarstrom, L., Kinnon, C., Levinsky, R., Bobrow, M., and *et al.* **(1993)**. The gene involved in X-linked agammaglobulinaemia is a member of the src family of protein-tyrosine kinases. *Nature* **361**, 226-33.

Vihinen, M., Belohradsky, B. H., Haire, R. N., Holinski-Feder, E., Kwan, S. P., Lappalainen, I., Lehvaslaiho, H., Lester, T., Meindl, A., Ochs, H. D., Ollila, J., Vorechovsky, I., Weiss, M., and Smith, C. I. **(1997)**. BTKbase, mutation database for X-linked agammaglobulinemia (XLA) [published erratum appears in Nucleic Acids Res 1997 May 1;25(9):1874]. *Nucleic Acids Res* **25**, 166-71.

Vihinen, M., Kwan, S. P., Lester, T., Ochs, H. D., Resnick, I., Valiaho, J., Conley, M. E., and Smith, C. I. **(1999)**. Mutations of the human BTK gene coding for bruton tyrosine kinase in X-linked agammaglobulinemia. *Hum Mutat* **13**, 280-5.

Vihinen, M., Nilsson, L., and Smith, C. I. **(1994)**. Tec homology (TH) adjacent to the PH domain. *FEBS Lett* **350**, 263-5.

Wang, D., Feng, J., Wen, R., Marine, J. C., Sangster, M. Y., Parganas, E., Hoffmeyer, A., Jackson, C. W., Cleveland, J. L., Murray, P. J., and Ihle, J. N. **(2000)**. Phospholipase Cgamma2 is essential in the functions of B cell and several Fc receptors. *Immunity* **13**, 25-35.

Wang, J. Y. (1993) Abl tyrosine kinase in signal transduction and cell cycle regulation. *Curr Opin Genet Dev* **3**, *35-43.*

Weil, D., Power, M. A., Smith, S. I., and Li, C. L. **(1997)**. Predominant expression of murine Bmx tyrosine kinase in the granulo-monocytic lineage. *Blood* **90**, 4332-40.

Welch, P. J., and Wang, J. Y. **(1993)**. A C-terminal protein-binding domain in the retinoblastoma protein regulates nuclear c-Abl tyrosine kinase in the cell cycle. *Cell* **75**, 779-90.

Wen, S. T., and Van Etten, R. A. **(1997)**. The PAG gene product, a stress-induced protein with antioxidant properties, is an Abl SH3-binding protein and a physiological inhibitor of c-Abl tyrosine kinase activity. *Genes Dev* **11**, 2456-67.

Williamson, M. P., Havel, T. F., and Wuthrich, K. **(1985)**. Solution conformation of proteinase inhibitor IIA from bull seminal plasma by 1H nuclear magnetic resonance and distance geometry. *J Mol Biol* **182**, 295-315.

Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. **(1995)**. ^{1}H, ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids.
I. Investigations of nearest-neighbor effects. *Journal of Biomolecular NMR* **5**, 67-81.

Wishart, D. S., and Sykes, B. D. **(1994)**. The ^{13}C Chemical Shift Index: A simple method for the identification of protein secondary structure using 13C chemical shift data. *Journal of Biomolecular NMR* **4**, 171-180.

Wishart, D. S., Sykes, B. D., and M., R. F. **(1992)**. The Chemical Shift Index: A fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry* **31**, 1647-1651.

Wishart, D. S., and Wang, Y. **(1998)**. Facile measurement of polypeptide JHNH alpha coupling constants from HMQC-J spectra. *J Biomol NMR* **11**, 329-36.

Wittekind, M., Mapelli, C., Farmer, B. T., 2nd, Suen, K. L., Goldfarb, V., Tsao, J., Lavoie, T., Barbacid, M., Meyers, C. A., and Mueller, L. **(1994)**. Orientation of peptide fragments from Sos proteins bound to the N-terminal SH3 domain of Grb2 determined by NMR spectroscopy. *Biochemistry* **33**, 13531-9.

Wuthrich, K. **(1986)**. *NMR of proteins and nucleic acids* (New York: Wiley- Interscience).

Wuthrich, K. **(1995)**. NMR- this other method for protein and nucleic acid structure determination. *Acta crystal D* **51**, 549-270.

Wuthrich, K. **(1990)**. Protein structure determination in solution by NMR spectroscopy. *J Biol Chem* **265**, 22059-62.

Wuthrich, K. **(1989)**. Protein structure determination in solution by nuclear magnetic resonance spectroscopy. *Science* **243**, 45-50.

Wuthrich, K., von Freyberg, B., Weber, C., Wider, G., Traber, R., Widmer, H., and Braun, W. **(1991)**. Receptor-induced conformation change of the immunosuppressant cyclosporin A. *Science* **254**, 953-4.

Xu, W., Harrison, S. C., and Eck, M. J. **(1997)**. Three-dimensional structure of the tyrosine kinase c-Src. *Nature* **385**, 595-602.

Xu, Y., Xu, D., Crawford, O. H., and Einstein, J. R. **(2000)**. A computational method for NMR-constrained protein threading. *J Comput Biol* **7**, 449-67.

Yamabhai, M., Hoffman, N. G., Hardison, N. L., McPherson, P. S., Castagnoli, L., Cesareni, G., and Kay, B. K. **(1998)**. Intersectin, a novel adaptor protein with two Eps15 homology and five Src homology 3 domains. *J Biol Chem* **273**, 31401-7.

Yamadori, T., Baba, Y., Matsushita, M., Hashimoto, S., Kurosaki, M., Kurosaki, T., Kishimoto, T., and Tsukada, S. **(1999)**. Bruton's tyrosine kinase activity is negatively regulated by Sab, the Btk-SH3 domain-binding protein. *Proc Natl Acad Sci U S A* **96**, 6341-6.

Yamashita, Y., Miyazato, A., Ohya, K., Ikeda, U., Shimada, K., Miura, Y., Ozawa, K., and Mano, H. **(1996)**. Deletion of Src homology 3 domain results in constitutive activation of Tec protein-tyrosine kinase. *Jpn J Cancer Res* **87**, 1106-10.

Yamashita, Y., Miyazato, A., Shimizu, R., Komatsu, N., Miura, Y., Ozawa, K., and Mano, H. **(1997)**. Tec protein-tyrosine kinase is involved in the thrombopoietin/c-Mpl signaling pathway. *Exp Hematol* **25**, 211-6.

Yang, W., and Desiderio, S. **(1997)**. BAP-135, a target for Bruton's tyrosine kinase in response to B cell receptor engagement. *Proc Natl Acad Sci U S A* **94**, 604-9.

Yang, W., Malek, S. N., and Desiderio, S. **(1995)**. An SH3-binding site conserved in Bruton's tyrosine kinase and related tyrosine kinases mediates specific protein interactions in vitro and in vivo. *J Biol Chem* **270**, 20832-40.

Yang, W. C., Ching, K. A., Tsoukas, C. D., and Berg, L. J. **(2001)**. Tec kinase signaling in T cells is regulated by phosphatidylinositol 3-kinase and the Tec pleckstrin homology domain. *J Immunol* **166**, 387-95.

Yang, W. C., Collette, Y., Nunes, J. A., and Olive, D. **(2000)**. Tec kinases: a family with multiple roles in immunity. *Immunity* **12**, 373-82.

Yang, W. C., Ghiotto, M., Barbarat, B., and Olive, D. **(1999)**. The role of Tec protein-tyrosine kinase in T cell signaling. *J Biol Chem* **274**, 607-17.

Yang, Y. S., Garbay, C., Duchesne, M., Cornille, F., Jullian, N., Fromage, N., Tocque, B., and Roques, B. P. **(1994)**. Solution structure of GAP SH3 domain by 1H NMR and spatial arrangement of essential Ras signaling-involved sequence. *Embo J* **13**, 1270-9.

Yao, L., Janmey, P., Frigeri, L. G., Han, W., Fujita, J., Kawakami, Y., Apgar, J. R., and Kawakami, T. **(1999)**. Pleckstrin homology domains interact with filamentous actin. *J Biol Chem* **274**, 19752-61.

Yao, L., Kawakami, Y., and Kawakami, T. **(1994)**. The pleckstrin homology domain of Bruton tyrosine kinase interacts with protein kinase C. *Proc Natl Acad Sci U S A* **91**, 9175-9.

Yao, L., Suzuki, H., Ozawa, K., Deng, J., Lehel, C., Fukamachi, H., Anderson, W. B., Kawakami, Y., and Kawakami, T. **(1997)**. Interactions between protein kinase C and pleckstrin homology domains. Inhibition by phosphatidylinositol 4,5-bisphosphate and phorbol 12-myristate 13-acetate. *J Biol Chem* **272**, 13033-9.

Yoon, H. S., Hajduk, P. J., Petros, A. M., Olejniczak, E. T., Meadows, R. P., and Fesik, S. W. **(1994)**. Solution structure of a pleckstrin-homology domain. *Nature* **369**, 672-5.

Yu, H., Chen, J. K., Feng, S., Dalgarno, D. C., Brauer, A. W., and Schreiber, S. L. **(1994)**. Structural basis for the binding of proline-rich peptides to SH3 domains. *Cell* **76**, 933-45.

Yu, H., Rosen, M. K., Shin, T. B., Seidel-Dugan, C., Brugge, J. S., and Schreiber, S. L. **(1992)**. Solution structure of the SH3 domain of Src and identification of its ligand-binding site. *Science* **258**, 1665-8.

Yuzawa, S., Yokochi, M., Hatanaka, H., Ogura, K., Kataoka, M., Miura Ki, K., Mandiyan, V., Schlessinger, J., and Inagaki, F. **(2001)**. Solution Structure of Grb2 Reveals Extensive Flexibility Necessary for Target Recognition. *J Mol Biol* **306**, 527-537.

Zhang, P., Talluri, S., Deng, H., Branton, D., and Wagner, G. **(1995)**. Solution structure of the pleckstrin homology domain of Drosophila beta-spectrin. *Structure* **3**, 1185-95.

Zheng, J., Cahill, S. M., Lemmon, M. A., Fushman, D., Schlessinger, J., and Cowburn, D. **(1996)**. Identification of the binding site for acidic phospholipids on the pH domain of dynamin: implications for stimulation of GTPase activity. *J Mol Biol* **255**, 14-21.

Zheng, X. M., Resnick, R. J., and Shalloway, D. **(2000)**. A phosphotyrosine displacement mechanism for activation of Src by PTPalpha. *Embo J* **19**, 964-78.

Zheng, X. M., Wang, Y., and Pallen, C. J. **(1992)**. Cell transformation and activation of pp60c-src by overexpression of a protein tyrosine phosphatase. *Nature* **359**, 336-9.

Zhu, J., and Shore, S. K. **(1996)**. c-ABL tyrosine kinase activity is regulated by association with a novel SH3-domain-binding protein. *Mol Cell Biol* **16**, 7054-62.

Zhu, Q., Zhang, M., Rawlings, D. J., Vihinen, M., Hagemann, T., Saffran, D. C., Kwan, S. P., Nilsson, L., Smith, C. I., Witte, O. N., and *et al.* **(1994)**. Deletion within the Src homology domain 3 of Bruton's tyrosine kinase resulting in X-linked agammaglobulinemia (XLA). *J Exp Med* **180**, 461-70.

APPENDIX

APPENDIX 1: RESTRAINTS LIST FOR TEC SH3 DOMAIN

The amino acid numbers are full length Tec numbers.
The numbers following the brackets are the peak numbers in the XEASY peak list.

Intraresidue NOES: 468
Sequential NOES : 173
Medium Range NOES: 55
Long Range NOES : 260
Total NOEs : 956

```
  GLY -2

  SER -1
assign (resid    -1 and name HA  )(resid    -1 and name HB2 ) 344
assign (resid    -1 and name HA  )(resid    -1 and name HB1 ) 346
assign (resid    -1 and name HB1 )(resid    -1 and name HB2 ) 1295
assign (resid    -1 and name HB2 )(resid   181 and name HN  ) 369
assign (resid    -1 and name HA  )(resid   181 and name HN  ) 370
assign (resid    -1 and name HA  )(resid   206 and name HG2*) 345
assign (resid    -1 and name HB2 )(resid   206 and name HG2*) 1146
assign (resid    -1 and name HB1 )(resid   206 and name HG2*) 1147

  GLU 181
assign (resid   181 and name HN  )(resid   181 and name HG* ) 106
assign (resid   181 and name HN  )(resid   181 and name HA  ) 107
assign (resid   181 and name HN  )(resid   181 and name HB* ) 368
assign (resid   181 and name HA  )(resid   182 and name HG12) 938
assign (resid   181 and name HA  )(resid   182 and name HN  ) 356
assign (resid   181 and name HG* )(resid   182 and name HN  ) 668
assign (resid   181 and name HG* )(resid   183 and name HG1*) 1081
assign (resid   181 and name HA  )(resid   183 and name HG1*) 1087
assign (resid   181 and name HA  )(resid   207 and name HG2*) 1226
assign (resid   181 and name HG* )(resid   207 and name HG2*) 1232
assign (resid   181 and name HN  )(resid   206 and name HG2*) 367

  ILE 182
assign (resid   182 and name HA  )(resid   182 and name HB  ) 219
assign (resid   182 and name HA  )(resid   182 and name HG2*) 223
assign (resid   182 and name HA  )(resid   182 and name HG12) 228
assign (resid   182 and name HA  )(resid   182 and name HG11) 229
assign (resid   182 and name HB  )(resid   182 and name HG11) 893
assign (resid   182 and name HG12)(resid   182 and name HG11) 1050
assign (resid   182 and name HG2*)(resid   182 and name HG11) 1111
assign (resid   182 and name HG12)(resid   182 and name HD1*) 1504
assign (resid   182 and name HN  )(resid   182 and name HB  ) 99
assign (resid   182 and name HN  )(resid   182 and name HG11) 287
assign (resid   182 and name HN  )(resid   182 and name HG12) 357
assign (resid   182 and name HN  )(resid   182 and name HA  ) 779
assign (resid   182 and name HA  )(resid   206 and name HA  ) 300
assign (resid   182 and name HG2*)(resid   206 and name HA  ) 308
assign (resid   182 and name HD1*)(resid   236 and name HG* ) 1172

  VAL 183
assign (resid   183 and name HA  )(resid   183 and name HB  ) 193
assign (resid   183 and name HB  )(resid   183 and name HG1*) 772
assign (resid   183 and name HB  )(resid   183 and name HG2*) 773
assign (resid   183 and name HG1*)(resid   183 and name HA  ) 1075
assign (resid   183 and name HA  )(resid   184 and name HN  ) 190
assign (resid   183 and name HB  )(resid   184 and name HN  ) 492
assign (resid   183 and name HG2*)(resid   218 and name HA  ) 157
```

```
assign (resid   183 and name HG2*)(resid   233 and name HA  ) 167
assign (resid   183 and name HB  )(resid   204 and name HA  ) 173
assign (resid   183 and name HA  )(resid   235 and name HA1 ) 410
assign (resid   183 and name HA  )(resid   235 and name HA2 ) 465
assign (resid   183 and name HG1*)(resid   207 and name HA  ) 404
assign (resid   183 and name HG2*)(resid   216 and name HB2 ) 567
assign (resid   183 and name HG2*)(resid   216 and name HB1 ) 647
assign (resid   183 and name HG2*)(resid   205 and name HB2 ) 686
assign (resid   183 and name HG2*)(resid   205 and name HB1 ) 740
assign (resid   183 and name HB  )(resid   233 and name HA  ) 769
assign (resid   183 and name HB  )(resid   233 and name HG1*) 774
assign (resid   183 and name HB  )(resid   233 and name HG2*) 1210
assign (resid   183 and name HG2*)(resid   207 and name HB  ) 804
assign (resid   183 and name HG1*)(resid   218 and name HB* ) 1068
assign (resid   183 and name HG1*)(resid   216 and name HZ3 ) 1390
assign (resid   183 and name HG1*)(resid   216 and name HE3 ) 1456
assign (resid   183 and name HG1*)(resid   205 and name HB1 ) 1078
assign (resid   183 and name HG1*)(resid   207 and name HB  ) 1080
assign (resid   183 and name HG1*)(resid   218 and name HA  ) 1084
assign (resid   183 and name HG1*)(resid   206 and name HA  ) 1085
assign (resid   183 and name HG1*)(resid   216 and name HB1 ) 1088
assign (resid   183 and name HG2*)(resid   216 and name HE3 ) 1101
assign (resid   183 and name HG2*)(resid   233 and name HG2*) 1113
assign (resid   183 and name HG2*)(resid   207 and name HG2*) 1114
assign (resid   183 and name HA  )(resid   206 and name HA  ) 1289
assign (resid   183 and name HG2*)(resid   216 and name HZ3 ) 1452

  VAL 184
assign (resid   184 and name HA  )(resid   184 and name HB  ) 1359
assign (resid   184 and name HN  )(resid   184 and name HB  ) 192
assign (resid   184 and name HG2*)(resid   185 and name HN  ) 358
assign (resid   184 and name HA  )(resid   185 and name HN  ) 359
assign (resid   184 and name HA  )(resid   205 and name HD* ) 103
assign (resid   184 and name HG1*)(resid   204 and name HA  ) 176
assign (resid   184 and name HA  )(resid   204 and name HA  ) 187
assign (resid   184 and name HG1*)(resid   204 and name HG1 ) 746
assign (resid   184 and name HA  )(resid   204 and name HG2 ) 809
assign (resid   184 and name HB  )(resid   234 and name HG2*) 852
assign (resid   184 and name HB  )(resid   204 and name HA  ) 853
assign (resid   184 and name HG2*)(resid   234 and name HG2*) 1032
assign (resid   184 and name HG1*)(resid   204 and name HG2 ) 1096
assign (resid   184 and name HG1*)(resid   236 and name HA  ) 1116
assign (resid   184 and name HG2*)(resid   204 and name HA  ) 1121
assign (resid   184 and name HG1*)(resid   236 and name HN  ) 411
assign (resid   184 and name HG2*)(resid   203 and name HN  ) 473
assign (resid   184 and name HA  )(resid   205 and name HN  ) 481
assign (resid   184 and name HN  )(resid   235 and name HA1 ) 493
assign (resid   184 and name HN  )(resid   234 and name HN  ) 494

  ALA 185
assign (resid   185 and name HA  )(resid   185 and name HB* ) 267
assign (resid   185 and name HN  )(resid   185 and name HB* ) 101
assign (resid   185 and name HN  )(resid   185 and name HA  ) 105
assign (resid   185 and name HB* )(resid   186 and name HN  ) 349
assign (resid   185 and name HA  )(resid   186 and name HN  ) 350
assign (resid   185 and name HA  )(resid   233 and name HA  ) 243
assign (resid   185 and name HA  )(resid   233 and name HG2*) 271
assign (resid   185 and name HA  )(resid   230 and name HA  ) 697
assign (resid   185 and name HB* )(resid   205 and name HE* ) 864
assign (resid   185 and name HB* )(resid   201 and name HA  ) 885
assign (resid   185 and name HB* )(resid   200 and name HN  ) 131
assign (resid   185 and name HA  )(resid   200 and name HN  ) 133
assign (resid   185 and name HN  )(resid   204 and name HA  ) 360
assign (resid   185 and name HN  )(resid   205 and name HD* ) 361
```

```
assign (resid   185 and name HN  )(resid   205 and name HE* ) 362
assign (resid   185 and name HB* )(resid   203 and name HN  ) 474

  MET 186
assign (resid   186 and name HA  )(resid   186 and name HG* ) 434
assign (resid   186 and name HA  )(resid   186 and name HB* ) 435
assign (resid   186 and name HG* )(resid   186 and name HB* ) 996
assign (resid   186 and name HN  )(resid   186 and name HB* ) 92
assign (resid   186 and name HN  )(resid   186 and name HG* ) 93
assign (resid   186 and name HN  )(resid   186 and name HA  ) 94
assign (resid   186 and name HB* )(resid   187 and name HE* ) 121
assign (resid   186 and name HB* )(resid   187 and name HD* ) 1492
assign (resid   186 and name HN  )(resid   187 and name HN  ) 352
assign (resid   186 and name HB* )(resid   233 and name HN  ) 55
assign (resid   186 and name HG* )(resid   233 and name HA  ) 164
assign (resid   186 and name HB* )(resid   232 and name HA  ) 303
assign (resid   186 and name HE* )(resid   237 and name HE* ) 871
assign (resid   186 and name HG* )(resid   234 and name HG2*) 1030
assign (resid   186 and name HG* )(resid   234 and name HB  ) 1367
assign (resid   186 and name HN  )(resid   233 and name HG2*) 348
assign (resid   186 and name HN  )(resid   233 and name HA  ) 351
assign (resid   186 and name HN  )(resid   233 and name HG1*) 355
assign (resid   186 and name HN  )(resid   234 and name HN  ) 669
assign (resid   186 and name HN  )(resid   199 and name HD2*) 773

  TYR 187
assign (resid   187 and name HD* )(resid   187 and name HA  ) 118
assign (resid   187 and name HD* )(resid   187 and name HB2 ) 119
assign (resid   187 and name HD* )(resid   187 and name HB1 ) 747
assign (resid   187 and name HA  )(resid   187 and name HB1 ) 1466
assign (resid   187 and name HB1 )(resid   187 and name HB2 ) 751
assign (resid   187 and name HA  )(resid   187 and name HB2 ) 1349
assign (resid   187 and name HE* )(resid   187 and name HB1 ) 1490
assign (resid   187 and name HE* )(resid   187 and name HA  ) 1493
assign (resid   187 and name HB1 )(resid   188 and name HB2 ) 753
assign (resid   187 and name HB1 )(resid   188 and name HN  ) 465
assign (resid   187 and name HB2 )(resid   188 and name HN  ) 467
assign (resid   187 and name HB1 )(resid   232 and name HD* ) 756
assign (resid   187 and name HA  )(resid   201 and name HB1 ) 278
assign (resid   187 and name HA  )(resid   201 and name HG2 ) 279
assign (resid   187 and name HA  )(resid   201 and name HD1 ) 585

  ASP 188
assign (resid   188 and name HA  )(resid   188 and name HB1 ) 297
assign (resid   188 and name HA  )(resid   188 and name HB2 ) 298
assign (resid   188 and name HB2 )(resid   188 and name HB1 ) 708
assign (resid   188 and name HN  )(resid   188 and name HB2 ) 468
assign (resid   188 and name HN  )(resid   188 and name HB1 ) 175
assign (resid   188 and name HB1 )(resid   189 and name HN  ) 155
assign (resid   188 and name HA  )(resid   189 and name HN  ) 435
assign (resid   188 and name HB2 )(resid   189 and name HN  ) 671
assign (resid   188 and name HN  )(resid   189 and name HN  ) 672
assign (resid   188 and name HA  )(resid   200 and name HA  ) 1291
assign (resid   188 and name HN  )(resid   201 and name HG2 ) 464
assign (resid   188 and name HA  )(resid   198 and name HE  ) 733

  PHE 189
assign (resid   189 and name HD* )(resid   189 and name HA  ) 203
assign (resid   189 and name HD* )(resid   189 and name HB1 ) 26
assign (resid   189 and name HD* )(resid   189 and name HB2 )
assign (resid   189 and name HA  )(resid   189 and name HB1 ) 205
assign (resid   189 and name HA  )(resid   189 and name HB2 ) 206
assign (resid   189 and name HB1 )(resid   189 and name HB2 ) 587
assign (resid   189 and name HZ  )(resid   189 and name HE* ) 1335
```

```
assign (resid   189 and name HE* )(resid   189 and name HD* ) 1333
assign (resid   189 and name HN  )(resid   189 and name HB1 ) 156
assign (resid   189 and name HN  )(resid   189 and name HA  ) 157
assign (resid   189 and name HA  )(resid   190 and name HN  ) 353
assign (resid   189 and name HD* )(resid   190 and name HN  ) 354
assign (resid   189 and name HN  )(resid   190 and name HN  ) 440
assign (resid   189 and name HB1 )(resid   190 and name HN  ) 690
assign (resid   189 and name HE* )(resid   191 and name HA  ) 9
assign (resid   189 and name HE* )(resid   191 and name HB* ) 10
assign (resid   189 and name HE* )(resid   196 and name HB2 ) 12
assign (resid   189 and name HE* )(resid   196 and name HB1 ) 13
assign (resid   189 and name HD* )(resid   232 and name HE* ) 22
assign (resid   189 and name HD* )(resid   199 and name HB2 ) 24
assign (resid   189 and name HD* )(resid   199 and name HG  ) 1450
assign (resid   189 and name HD* )(resid   228 and name HG11) 30
assign (resid   189 and name HD* )(resid   199 and name HD1*) 31
assign (resid   189 and name HD* )(resid   232 and name HD* ) 78
assign (resid   189 and name HA  )(resid   232 and name HD* ) 79
assign (resid   189 and name HB2 )(resid   232 and name HD* ) 80
assign (resid   189 and name HA  )(resid   232 and name HE* ) 204
assign (resid   189 and name HB1 )(resid   199 and name HB1 ) 591
assign (resid   189 and name HB1 )(resid   199 and name HD1*) 592
assign (resid   189 and name HB2 )(resid   199 and name HB2 ) 0668
assign (resid   189 and name HB2 )(resid   199 and name HG  ) 669
assign (resid   189 and name HB2 )(resid   199 and name HD1*) 671
assign (resid   189 and name HB2 )(resid   199 and name HD2*) 672
assign (resid   189 and name HE* )(resid   215 and name HH2 ) 1329
assign (resid   189 and name HZ  )(resid   215 and name HH2 ) 1336
assign (resid   189 and name HB1 )(resid   199 and name HB2 ) 1428
assign (resid   189 and name HN  )(resid   199 and name HD1*) 436
assign (resid   189 and name HN  )(resid   199 and name HB1 ) 437
assign (resid   189 and name HN  )(resid   199 and name HG  ) 438
assign (resid   189 and name HN  )(resid   199 and name HB2 ) 439

  GLN 190
assign (resid   190 and name HA  )(resid   190 and name HG* ) 428
assign (resid   190 and name HG* )(resid   190 and name HB1 ) 755
assign (resid   190 and name HB2 )(resid   190 and name HG* ) 846
assign (resid   190 and name HB1 )(resid   190 and name HA  ) 1470
assign (resid   190 and name HN  )(resid   190 and name HB1 ) 95
assign (resid   190 and name HN  )(resid   190 and name HG* ) 96
assign (resid   190 and name HN  )(resid   190 and name HA  ) 97
assign (resid   190 and name HE21)(resid   190 and name HE22) 218
assign (resid   190 and name HE21)(resid   190 and name HG* ) 559
assign (resid   190 and name HA  )(resid   191 and name HB* ) 432
assign (resid   190 and name HB1 )(resid   191 and name HN  ) 568
assign (resid   190 and name HA  )(resid   191 and name HN  ) 569
assign (resid   190 and name HG* )(resid   191 and name HN  ) 570
assign (resid   190 and name HG* )(resid   192 and name HG2*) 763
assign (resid   190 and name HA  )(resid   198 and name HG2 ) 429
assign (resid   190 and name HA  )(resid   198 and name HG1 ) 431

  ALA 191
assign (resid   191 and name HA  )(resid   191 and name HB* ) 464
assign (resid   191 and name HN  )(resid   191 and name HB* ) 232
assign (resid   191 and name HN  )(resid   191 and name HA  ) 233
assign (resid   191 and name HA  )(resid   192 and name HN  ) 507
assign (resid   191 and name HN  )(resid   192 and name HN  ) 678
assign (resid   191 and name HB* )(resid   194 and name HN  ) 402
assign (resid   191 and name HB* )(resid   193 and name HN  ) 589
assign (resid   191 and name HA  )(resid   193 and name HN  ) 592
assign (resid   191 and name HB* )(resid   197 and name HA  ) 336
assign (resid   191 and name HN  )(resid   198 and name HA  ) 393
assign (resid   191 and name HB* )(resid   198 and name HA  ) 396
```

```
assign (resid   191 and name HA  )(resid   196 and name HB2 ) 462
assign (resid   191 and name HA  )(resid   196 and name HB1 ) 463
assign (resid   191 and name HB* )(resid   196 and name HB2 ) 657
assign (resid   191 and name HB* )(resid   196 and name HB1 ) 670
assign (resid   191 and name HB* )(resid   198 and name HB2 ) 817
assign (resid   191 and name HB* )(resid   198 and name HB1 ) 840
assign (resid   191 and name HB* )(resid   198 and name HN  ) 577
assign (resid   191 and name HN  )(resid   198 and name HG2 ) 691

  THR 192
assign (resid   192 and name HA  )(resid   192 and name HG2*) 0416
assign (resid   192 and name HB  )(resid   192 and name HA  ) 1292
assign (resid   192 and name HG2*)(resid   192 and name HB  ) 01042
assign (resid   192 and name HN  )(resid   192 and name HA  ) 198
assign (resid   192 and name HG2*)(resid   193 and name HG2 ) 1044
assign (resid   192 and name HG2*)(resid   193 and name HG1 ) 1045
assign (resid   192 and name HA  )(resid   193 and name HN  ) 240
assign (resid   192 and name HB  )(resid   193 and name HN  ) 241
assign (resid   192 and name HN  )(resid   193 and name HN  ) 594
assign (resid   192 and name HN  )(resid   196 and name HB2 ) 506

  GLU 193
assign (resid   193 and name HA  )(resid   193 and name HB2 ) 371
assign (resid   193 and name HG1 )(resid   193 and name HA  ) 859
assign (resid   193 and name HB1 )(resid   193 and name HG2 ) 915
assign (resid   193 and name HB1 )(resid   193 and name HA  ) 1364
assign (resid   193 and name HN  )(resid   193 and name HG2 ) 245
assign (resid   193 and name HB1 )(resid   194 and name HB* ) 1472
assign (resid   193 and name HG2 )(resid   194 and name HN  ) 404
assign (resid   193 and name HA  )(resid   194 and name HN  ) 405
assign (resid   193 and name HN  )(resid   196 and name HB2 ) 591

  ALA 194
assign (resid   194 and name HA  )(resid   194 and name HB* ) 448
assign (resid   194 and name HN  )(resid   194 and name HB* ) 134
assign (resid   194 and name HN  )(resid   194 and name HA  ) 135
assign (resid   194 and name HN  )(resid   195 and name HA  ) 406
assign (resid   194 and name HA  )(resid   197 and name HD* ) 443

  HIS 195
assign (resid   195 and name HD2 )(resid   195 and name HB1 ) 65
assign (resid   195 and name HA  )(resid   195 and name HB2 ) 230
assign (resid   195 and name HA  )(resid   195 and name HB1 ) 231
assign (resid   195 and name HB1 )(resid   195 and name HB2 ) 558
assign (resid   195 and name HD2 )(resid   196 and name HA  ) 62
assign (resid   195 and name HD2 )(resid   215 and name HZ3 ) 1326
assign (resid   195 and name HD2 )(resid   227 and name HB2 ) 1337
assign (resid   195 and name HD2 )(resid   227 and name HN  ) 488
assign (resid   195 and name HA  )(resid   217 and name HE  ) 687

  ASP 196
assign (resid   196 and name HA  )(resid   196 and name HB1 ) 179
assign (resid   196 and name HA  )(resid   196 and name HB2 ) 180
assign (resid   196 and name HA  )(resid   197 and name HN  ) 373
assign (resid   196 and name HA  )(resid   227 and name HB2 ) 178
assign (resid   196 and name HA  )(resid   227 and name HB1 ) 1481
assign (resid   196 and name HA  )(resid   227 and name HN  ) 487

  LEU 197
assign (resid   197 and name HA  )(resid   197 and name HB2 ) 334
assign (resid   197 and name HA  )(resid   197 and name HD* ) 340
assign (resid   197 and name HN  )(resid   197 and name HB1 ) 109
assign (resid   197 and name HN  )(resid   197 and name HA  ) 110
assign (resid   197 and name HA  )(resid   198 and name HN  ) 579
```

```
assign (resid   197 and name HD* )(resid   198 and name HN  ) 705
assign (resid   197 and name HA  )(resid   228 and name HG11) 337
assign (resid   197 and name HD* )(resid   226 and name HA1 ) 453
assign (resid   197 and name HB1 )(resid   228 and name HG11) 908
assign (resid   197 and name HB1 )(resid   228 and name HD1*) 1446
assign (resid   197 and name HN  )(resid   227 and name HN  ) 371
assign (resid   197 and name HN  )(resid   228 and name HG11) 375
assign (resid   197 and name HD* )(resid   227 and name HN  ) 484
assign (resid   197 and name HB2 )(resid   226 and name HN  ) 538

  ARG 198
assign (resid   198 and name HA  )(resid   198 and name HG1 ) 395
assign (resid   198 and name HD* )(resid   198 and name HB2 ) 573
assign (resid   198 and name HD* )(resid   198 and name HB1 ) 574
assign (resid   198 and name HD* )(resid   198 and name HG1 ) 576
assign (resid   198 and name HG2 )(resid   198 and name HG1 ) 947
assign (resid   198 and name HB1 )(resid   198 and name HG1 ) 948
assign (resid   198 and name HG2 )(resid   198 and name HD* ) 1419
assign (resid   198 and name HB2 )(resid   198 and name HB1 ) 1311
assign (resid   198 and name HN  )(resid   198 and name HB1 ) 237
assign (resid   198 and name HN  )(resid   198 and name HG1 ) 578
assign (resid   198 and name HN  )(resid   198 and name HA  ) 580
assign (resid   198 and name HN  )(resid   198 and name HG2 ) 704
assign (resid   198 and name HE  )(resid   198 and name HD* ) 734
assign (resid   198 and name HE  )(resid   198 and name HG1 ) 0736
assign (resid   198 and name HA  )(resid   199 and name HB2 ) 387

  LEU 199
assign (resid   199 and name HA  )(resid   199 and name HB1 ) 197
assign (resid   199 and name HA  )(resid   199 and name HD1*) 198
assign (resid   199 and name HA  )(resid   199 and name HD2*) 199
assign (resid   199 and name HA  )(resid   199 and name HB2 ) 200
assign (resid   199 and name HA  )(resid   199 and name HG  ) 201
assign (resid   199 and name HB2 )(resid   199 and name HD1*) 784
assign (resid   199 and name HB2 )(resid   199 and name HD2*) 1268
assign (resid   199 and name HB1 )(resid   199 and name HD1*) 1200
assign (resid   199 and name HB1 )(resid   199 and name HB2 ) 1020
assign (resid   199 and name HB1 )(resid   199 and name HD2*) 01271
assign (resid   199 and name HD1*)(resid   199 and name HD2*) 1274
assign (resid   199 and name HD2*)(resid   200 and name HN  ) 0396
assign (resid   199 and name HB1 )(resid   200 and name HN  ) 397
assign (resid   199 and name HA  )(resid   200 and name HN  ) 399
assign (resid   199 and name HA  )(resid   205 and name HE* ) 71
assign (resid   199 and name HD2*)(resid   205 and name HE* ) 77
assign (resid   199 and name HD2*)(resid   205 and name HD* ) 114
assign (resid   199 and name HD1*)(resid   228 and name HA  ) 225
assign (resid   199 and name HD2*)(resid   228 and name HA  ) 227
assign (resid   199 and name HG  )(resid   232 and name HB2 ) 579
assign (resid   199 and name HD1*)(resid   232 and name HB1 ) 633
assign (resid   199 and name HD1*)(resid   229 and name HD2 ) 716
assign (resid   199 and name HD1*)(resid   229 and name HD1 ) 726
assign (resid   199 and name HD1*)(resid   228 and name HG11) 1070
assign (resid   199 and name HD2*)(resid   228 and name HG11) 1072
assign (resid   199 and name HD2*)(resid   229 and name HB1 ) 1175
assign (resid   199 and name HD1*)(resid   232 and name HB2 ) 1418
assign (resid   199 and name HD1*)(resid   228 and name HG2*) 1199
assign (resid   199 and name HB1 )(resid   233 and name HG2*) 1213
assign (resid   199 and name HD2*)(resid   233 and name HG2*) 1215
assign (resid   199 and name HD2*)(resid   228 and name HD1*) 1275
assign (resid   199 and name HD1*)(resid   229 and name HG1 ) 1256
assign (resid   199 and name HD2*)(resid   228 and name HG2*) 1270
assign (resid   199 and name HD1*)(resid   205 and name HD* ) 1393
```

```
  GLU 200
assign (resid   200 and name HA  )(resid   200 and name HG* ) 232
assign (resid   200 and name HA  )(resid   200 and name HB2 ) 233
assign (resid   200 and name HA  )(resid   200 and name HB1 ) 234
assign (resid   200 and name HN  )(resid   200 and name HB2 ) 401
assign (resid   200 and name HN  )(resid   200 and name HA  ) 780
assign (resid   200 and name HB2 )(resid   201 and name HN  ) 390
assign (resid   200 and name HG* )(resid   201 and name HN  ) 391
assign (resid   200 and name HA  )(resid   201 and name HN  ) 394
assign (resid   200 and name HB1 )(resid   203 and name HE21) 68
assign (resid   200 and name HB1 )(resid   203 and name HB2 ) 826
assign (resid   200 and name HB1 )(resid   203 and name HG2 ) 828
assign (resid   200 and name HN  )(resid   203 and name HB2 ) 398
assign (resid   200 and name HG* )(resid   203 and name HE22) 628
assign (resid   200 and name HN  )(resid   205 and name HE* ) 400

  ARG 201
assign (resid   201 and name HA  )(resid   201 and name HG1 ) 544
assign (resid   201 and name HA  )(resid   201 and name HB2 )546
assign (resid   201 and name HA  )(resid   201 and name HB1 ) 547
assign (resid   201 and name HA  )(resid   201 and name HG2 ) 548
assign (resid   201 and name HD2 )(resid   201 and name HG2 ) 583
assign (resid   201 and name HD1 )(resid   201 and name HG2 ) 589
assign (resid   201 and name HD1 )(resid   201 and name HG1 ) 590
assign (resid   201 and name HD1 )(resid   201 and name HB1 ) 601
assign (resid   201 and name HG1 )(resid   201 and name HD2 ) 969
assign (resid   201 and name HN  )(resid   201 and name HB1 ) 388
assign (resid   201 and name HN  )(resid   201 and name HB2 ) 389
assign (resid   201 and name HN  )(resid   201 and name HD1 ) 392
assign (resid   201 and name HN  )(resid   201 and name HA  ) 393

  GLY 202
assign (resid   202 and name HA2 )(resid   202 and name HA1 ) 524
assign (resid   202 and name HN  )(resid   202 and name HA1 ) 197
assign (resid   202 and name HA2 )(resid   203 and name HN  ) 475
assign (resid   202 and name HA1 )(resid   203 and name HN  ) 476
assign (resid   202 and name HN  )(resid   203 and name HN  ) 503

  GLN 203
assign (resid   203 and name HA  )(resid   203 and name HG2 ) 358
assign (resid   203 and name HA  )(resid   203 and name HB2 ) 364
assign (resid   203 and name HA  )(resid   203 and name HB1 ) 365
assign (resid   203 and name HA  )(resid   203 and name HG1 ) 367
assign (resid   203 and name HB2 )(resid   203 and name HG2 ) 1415
assign (resid   203 and name HB2 )(resid   203 and name HG1 ) 677
assign (resid   203 and name HG1 )(resid   203 and name HG2 ) 754
assign (resid   203 and name HB1 )(resid   203 and name HB2 ) 794
assign (resid   203 and name HG2 )(resid   203 and name HB1 ) 1306
assign (resid   203 and name HN  )(resid   203 and name HB2 ) 182
assign (resid   203 and name HE22)(resid   203 and name HE21) 274
assign (resid   203 and name HE22)(resid   203 and name HA  ) 629
assign (resid   203 and name HE22)(resid   203 and name HG2 ) 630
assign (resid   203 and name HN  )(resid   203 and name HG2 ) 673
assign (resid   203 and name HE21)(resid   203 and name HG2 ) 681
assign (resid   203 and name HA  )(resid   204 and name HG1 ) 744
assign (resid   203 and name HB2 )(resid   204 and name HN  ) 470
assign (resid   203 and name HA  )(resid   204 and name HN  ) 471
assign (resid   203 and name HB1 )(resid   205 and name HE* ) 74
assign (resid   203 and name HG2 )(resid   205 and name HD* ) 1376
```

```
  GLU 204
assign (resid   204 and name HA  )(resid   204 and name HG1 ) 172
assign (resid   204 and name HA  )(resid   204 and name HG2 ) 174
assign (resid   204 and name HA  )(resid   204 and name HB* ) 175
assign (resid   204 and name HG1 )(resid   204 and name HG2 ) 807
assign (resid   204 and name HG1 )(resid   204 and name HB* ) 1386
assign (resid   204 and name HN  )(resid   204 and name HB* ) 177
assign (resid   204 and name HN  )(resid   204 and name HG2 ) 178
assign (resid   204 and name HN  )(resid   204 and name HA  ) 180
assign (resid   204 and name HA  )(resid   205 and name HD* ) 170
assign (resid   204 and name HG1 )(resid   205 and name HN  ) 184
assign (resid   204 and name HB* )(resid   205 and name HN  ) 479
assign (resid   204 and name HA  )(resid   205 and name HN  ) 480

  TYR 205
assign (resid   205 and name HD* )(resid   205 and name HB2 ) 108
assign (resid   205 and name HD* )(resid   205 and name HB1 ) 109
assign (resid   205 and name HD* )(resid   205 and name HA  ) 314
assign (resid   205 and name HA  )(resid   205 and name HB2 ) 320
assign (resid   205 and name HA  )(resid   205 and name HB1 ) 321
assign (resid   205 and name HB2 )(resid   205 and name HB1 ) 689
assign (resid   205 and name HN  )(resid   205 and name HB2 ) 185
assign (resid   205 and name HN  )(resid   205 and name HA  ) 186
assign (resid   205 and name HN  )(resid   205 and name HD* ) 482
assign (resid   205 and name HD* )(resid   206 and name HD1*) 113
assign (resid   205 and name HB2 )(resid   206 and name HN  ) 430
assign (resid   205 and name HB1 )(resid   206 and name HN  ) 431
assign (resid   205 and name HA  )(resid   206 and name HN  ) 432
assign (resid   205 and name HD* )(resid   219 and name HE  ) 4
assign (resid   205 and name HE* )(resid   220 and name HB1 ) 63
assign (resid   205 and name HE* )(resid   220 and name HA  ) 207
assign (resid   205 and name HD* )(resid   220 and name HA  ) 208
assign (resid   205 and name HD* )(resid   220 and name HB1 ) 106
assign (resid   205 and name HD* )(resid   218 and name HB* ) 111
assign (resid   205 and name HA  )(resid   220 and name HA  ) 210
assign (resid   205 and name HB2 )(resid   220 and name HA  ) 213
assign (resid   205 and name HB2 )(resid   218 and name HB* ) 685
assign (resid   205 and name HB1 )(resid   218 and name HB* ) 739
assign (resid   205 and name HD* )(resid   233 and name HG2*) 1207
assign (resid   205 and name HB2 )(resid   228 and name HD1*) 1471
assign (resid   205 and name HE* )(resid   221 and name HN  ) 674
assign (resid   205 and name HD* )(resid   221 and name HN  ) 675

  ILE 206
assign (resid   206 and name HA  )(resid   206 and name HG12) 302
assign (resid   206 and name HA  )(resid   206 and name HG11) 304
assign (resid   206 and name HA  )(resid   206 and name HG2*) 305
assign (resid   206 and name HB  )(resid   206 and name HG11) 891
assign (resid   206 and name HG11)(resid   206 and name HG12) 1052
assign (resid   206 and name HG2*)(resid   206 and name HG12) 1155
assign (resid   206 and name HG2*)(resid   206 and name HG11) 1156
assign (resid   206 and name HD1*)(resid   206 and name HA  ) 1168
assign (resid   206 and name HN  )(resid   206 and name HD1*) 151
assign (resid   206 and name HN  )(resid   206 and name HG11) 152
assign (resid   206 and name HN  )(resid   206 and name HG12) 434
assign (resid   206 and name HG2*)(resid   207 and name HN  ) 698
assign (resid   206 and name HA  )(resid   208 and name HD2*) 1505
assign (resid   206 and name HD1*)(resid   220 and name HA  ) 215
assign (resid   206 and name HB  )(resid   219 and name HA  ) 887
assign (resid   206 and name HD1*)(resid   220 and name HB1 ) 1482
assign (resid   206 and name HD1*)(resid   221 and name HA  ) 1495
assign (resid   206 and name HN  )(resid   219 and name HB1 ) 153
assign (resid   206 and name HN  )(resid   219 and name HN  ) 429
assign (resid   206 and name HN  )(resid   220 and name HA  ) 433
```

```
  ILE 207
assign (resid   207 and name HA  )(resid   207 and name HB  ) 414
assign (resid   207 and name HB  )(resid   207 and name HG11) 801
assign (resid   207 and name HG12)(resid   207 and name HG11) 878
assign (resid   207 and name HG11)(resid   207 and name HA  ) 1000
assign (resid   207 and name HG2*)(resid   207 and name HA  ) 1227
assign (resid   207 and name HG2*)(resid   207 and name HB  ) 1231
assign (resid   207 and name HG2*)(resid   207 and name HG11) 1236
assign (resid   207 and name HA  )(resid   207 and name HG12) 1442
assign (resid   207 and name HN  )(resid   207 and name HG12) 695
assign (resid   207 and name HN  )(resid   207 and name HG2*) 699
assign (resid   207 and name HG11)(resid   208 and name HN  ) 89
assign (resid   207 and name HD1*)(resid   208 and name HN  ) 342
assign (resid   207 and name HA  )(resid   208 and name HN  ) 344
assign (resid   207 and name HB  )(resid   208 and name HN  ) 667
assign (resid   207 and name HD1*)(resid   210 and name HA  ) 353
assign (resid   207 and name HD1*)(resid   210 and name HG1 ) 1241
assign (resid   207 and name HD1*)(resid   209 and name HN  ) 631
assign (resid   207 and name HD1*)(resid   216 and name HE3 ) 101
assign (resid   207 and name HA  )(resid   218 and name HA  ) 153
assign (resid   207 and name HG12)(resid   216 and name HE3 ) 865
assign (resid   207 and name HG2*)(resid   216 and name HZ3 ) 1223
assign (resid   207 and name HD1*)(resid   216 and name HB1 ) 1229
assign (resid   207 and name HA  )(resid   218 and name HB* ) 1353
assign (resid   207 and name HG11)(resid   217 and name HN  ) 147
assign (resid   207 and name HD1*)(resid   217 and name HN  ) 425
assign (resid   207 and name HN  )(resid   218 and name HB* ) 697

  LEU 208
assign (resid   208 and name HA  )(resid   208 and name HG  ) 373
assign (resid   208 and name HA  )(resid   208 and name HB* ) 375
assign (resid   208 and name HA  )(resid   208 and name HD2*) 379
assign (resid   208 and name HG  )(resid   208 and name HD1*) 926
assign (resid   208 and name HB* )(resid   208 and name HG  ) 998
assign (resid   208 and name HD2*)(resid   208 and name HB* ) 1184
assign (resid   208 and name HN  )(resid   208 and name HG  ) 90
assign (resid   208 and name HN  )(resid   208 and name HD1*) 343
assign (resid   208 and name HB* )(resid   209 and name HB1 ) 816
assign (resid   208 and name HN  )(resid   209 and name HN  ) 638
assign (resid   208 and name HD1*)(resid   209 and name HN  ) 632
assign (resid   208 and name HB* )(resid   209 and name HN  ) 633
assign (resid   208 and name HD1*)(resid   225 and name HA  ) 137
assign (resid   208 and name HD1*)(resid   219 and name HA  ) 1479
assign (resid   208 and name HG  )(resid   218 and name HA  ) 155
assign (resid   208 and name HD1*)(resid   218 and name HA  ) 158
assign (resid   208 and name HD2*)(resid   219 and name HD2 ) 563
assign (resid   208 and name HD1*)(resid   217 and name HD2 ) 623
assign (resid   208 and name HD1*)(resid   225 and name HG2 ) 741
assign (resid   208 and name HD1*)(resid   225 and name HG1 ) 798
assign (resid   208 and name HD1*)(resid   217 and name HD1 ) 1177
assign (resid   208 and name HD2*)(resid   218 and name HA  ) 1179
assign (resid   208 and name HD2*)(resid   225 and name HG2 ) 1180
assign (resid   208 and name HD2*)(resid   225 and name HG1 ) 1181
assign (resid   208 and name HD2*)(resid   219 and name HG2 ) 1182
assign (resid   208 and name HD1*)(resid   218 and name HN  ) 294
assign (resid   208 and name HN  )(resid   218 and name HA  ) 345
assign (resid   208 and name HN  )(resid   219 and name HN  ) 347
assign (resid   208 and name HD1*)(resid   225 and name HN  ) 451
assign (resid   208 and name HD1*)(resid   219 and name HN  ) 495
assign (resid   208 and name HD1*)(resid   226 and name HN  ) 536
assign (resid   208 and name HD2*)(resid   219 and name HE  ) 547
assign (resid   208 and name HD1*)(resid   217 and name HE  ) 764
```

```
  GLU 209
assign (resid   209 and name HA  )(resid   209 and name HB2 ) 388
assign (resid   209 and name HG* )(resid   209 and name HA  ) 1366
assign (resid   209 and name HB2 )(resid   209 and name HB1 ) 1312
assign (resid   209 and name HN  )(resid   209 and name HA  ) 268
assign (resid   209 and name HN  )(resid   209 and name HB2 ) 269
assign (resid   209 and name HN  )(resid   209 and name HB1 ) 635
assign (resid   209 and name HA  )(resid   210 and name HN  ) 139
assign (resid   209 and name HB2 )(resid   210 and name HN  ) 408
assign (resid   209 and name HB2 )(resid   211 and name HN  ) 449
assign (resid   209 and name HG* )(resid   217 and name HB1 ) 945
assign (resid   209 and name HB1 )(resid   217 and name HN  ) 426
assign (resid   209 and name HN  )(resid   217 and name HN  ) 637
assign (resid   209 and name HN  )(resid   217 and name HB1 ) 634
assign (resid   209 and name HN  )(resid   218 and name HA  ) 636
assign (resid   209 and name HN  )(resid   217 and name HB2 ) 679

  LYS 210
assign (resid   210 and name HA  )(resid   210 and name HB1 ) 352
assign (resid   210 and name HA  )(resid   210 and name HG2 ) 354
assign (resid   210 and name HA  )(resid   210 and name HG1 ) 357
assign (resid   210 and name HE* )(resid   210 and name HG1 ) 788
assign (resid   210 and name HE* )(resid   210 and name HD2 ) 958
assign (resid   210 and name HE* )(resid   210 and name HG2 ) 799
assign (resid   210 and name HD2 )(resid   210 and name HG1 ) 965
assign (resid   210 and name HB2 )(resid   210 and name HG2 ) 1244
assign (resid   210 and name HB2 )(resid   210 and name HG1 ) 1061
assign (resid   210 and name HB1 )(resid   210 and name HG1 ) 1064
assign (resid   210 and name HD1 )(resid   210 and name HE* ) 1160
assign (resid   210 and name HD1 )(resid   210 and name HD2 ) 1161
assign (resid   210 and name HD1 )(resid   210 and name HG1 ) 1282
assign (resid   210 and name HG2 )(resid   210 and name HD1 ) 1245
assign (resid   210 and name HG2 )(resid   210 and name HG1 ) 1286
assign (resid   210 and name HD2 )(resid   210 and name HG2 ) 1445
assign (resid   210 and name HB2 )(resid   210 and name HA  ) 1498
assign (resid   210 and name HN  )(resid   210 and name HG2 ) 137
assign (resid   210 and name HN  )(resid   210 and name HB1 ) 138
assign (resid   210 and name HN  )(resid   210 and name HG1 ) 410
assign (resid   210 and name HN  )(resid   211 and name HA  ) 407
assign (resid   210 and name HB1 )(resid   211 and name HN  ) 448
assign (resid   210 and name HA  )(resid   211 and name HN  ) 450
assign (resid   210 and name HG1 )(resid   213 and name HD1*) 1284
assign (resid   210 and name HG2 )(resid   213 and name HD1*) 1247
assign (resid   210 and name HE* )(resid   213 and name HD1*) 1314
assign (resid   210 and name HD2 )(resid   216 and name HZ2 ) 5
assign (resid   210 and name HA  )(resid   216 and name HD1 ) 15
assign (resid   210 and name HB2 )(resid   216 and name HD1 ) 18
assign (resid   210 and name HB1 )(resid   216 and name HD1 ) 19
assign (resid   210 and name HA  )(resid   216 and name HA  ) 351
assign (resid   210 and name HA  )(resid   216 and name HB2 ) 356
assign (resid   210 and name HA  )(resid   217 and name HN  ) 150

  ASN 211
assign (resid   211 and name HA  )(resid   211 and name HB1 ) 307
assign (resid   211 and name HN  )(resid   211 and name HB1 ) 160
assign (resid   211 and name HN  )(resid   211 and name HA  ) 161
assign (resid   211 and name HD21)(resid   211 and name HD22) 214
assign (resid   211 and name HD21)(resid   211 and name HB1 ) 556
assign (resid   211 and name HD22)(resid   211 and name HB1 ) 562
assign (resid   211 and name HN  )(resid   212 and name HN  ) 447
assign (resid   211 and name HA  )(resid   212 and name HN  ) 718
assign (resid   211 and name HB1 )(resid   212 and name HN  ) 720
```

```
  ASP 212
assign (resid   212 and name HA  )(resid   212 and name HB1 ) 241
assign (resid   212 and name HB2 )(resid   212 and name HB1 ) 543
assign (resid   212 and name HA  )(resid   212 and name HB2 ) 1440
assign (resid   212 and name HN  )(resid   212 and name HB2 ) 710
assign (resid   212 and name HA  )(resid   213 and name HD1*) 1093
assign (resid   212 and name HA  )(resid   213 and name HN  ) 620
assign (resid   212 and name HB2 )(resid   213 and name HN  ) 622
assign (resid   212 and name HB1 )(resid   213 and name HN  ) 623
assign (resid   212 and name HN  )(resid   213 and name HD1*) 721
assign (resid   212 and name HB2 )(resid   215 and name HB* ) 628
assign (resid   212 and name HA  )(resid   216 and name HD1 ) 1395
assign (resid   212 and name HB2 )(resid   215 and name HN  ) 442
assign (resid   212 and name HB2 )(resid   214 and name HN  ) 520

  LEU 213
assign (resid   213 and name HA  )(resid   213 and name HG  ) 425
assign (resid   213 and name HB2 )(resid   213 and name HA  ) 929
assign (resid   213 and name HB2 )(resid   213 and name HB1 ) 933
assign (resid   213 and name HB2 )(resid   213 and name HG  ) 934
assign (resid   213 and name HB1 )(resid   213 and name HA  ) 992
assign (resid   213 and name HG  )(resid   213 and name HB1 ) 1028
assign (resid   213 and name HD1*)(resid   213 and name HG  ) 1097
assign (resid   213 and name HD2*)(resid   213 and name HB2 ) 1117
assign (resid   213 and name HN  )(resid   213 and name HB1 ) 263
assign (resid   213 and name HN  )(resid   213 and name HA  ) 264
assign (resid   213 and name HN  )(resid   213 and name HB2 ) 625
assign (resid   213 and name HN  )(resid   213 and name HG  ) 761
assign (resid   213 and name HG  )(resid   214 and name HD2 ) 1477
assign (resid   213 and name HD2*)(resid   214 and name HN  ) 518
assign (resid   213 and name HG  )(resid   214 and name HN  ) 519
assign (resid   213 and name HA  )(resid   214 and name HN  ) 521
assign (resid   213 and name HN  )(resid   214 and name HN  ) 621
assign (resid   213 and name HB1 )(resid   214 and name HN  ) 775

  HIS 214
assign (resid   214 and name HD2 )(resid   214 and name HB2 ) 125
assign (resid   214 and name HA  )(resid   214 and name HB2 ) 398
assign (resid   214 and name HA  )(resid   214 and name HB1 ) 399
assign (resid   214 and name HB1 )(resid   214 and name HB2 ) 1299
assign (resid   214 and name HD2 )(resid   214 and name HB1 ) 1476
assign (resid   214 and name HN  )(resid   214 and name HB1 ) 201
assign (resid   214 and name HN  )(resid   214 and name HB2 ) 202
assign (resid   214 and name HN  )(resid   214 and name HA  ) 203
assign (resid   214 and name HB2 )(resid   215 and name HD1 ) 1374
assign (resid   214 and name HB1 )(resid   215 and name HE1 ) 337
assign (resid   214 and name HB1 )(resid   215 and name HN  ) 441
assign (resid   214 and name HA  )(resid   215 and name HN  ) 443
assign (resid   214 and name HN  )(resid   215 and name HN  ) 522
assign (resid   214 and name HA  )(resid   230 and name HB2 ) 857
assign (resid   214 and name HA  )(resid   230 and name HN  ) 756

  TRP 215
assign (resid   215 and name HZ2 )(resid   215 and name HH2 ) 21
assign (resid   215 and name HZ3 )(resid   215 and name HE3 ) 40
assign (resid   215 and name HZ3 )(resid   215 and name HA  ) 182
assign (resid   215 and name HH2 )(resid   215 and name HE3 ) 49
assign (resid   215 and name HA  )(resid   215 and name HB* ) 184
assign (resid   215 and name HB* )(resid   215 and name HD1 ) 625
assign (resid   215 and name HN  )(resid   215 and name HA  ) 159
assign (resid   215 and name HE1 )(resid   215 and name HD1 ) 335
assign (resid   215 and name HE1 )(resid   215 and name HZ2 ) 338
assign (resid   215 and name HN  )(resid   215 and name HD1 ) 444
assign (resid   215 and name HA  )(resid   216 and name HN  ) 692
```

```
assign (resid   215 and name HZ3 )(resid   227 and name HB1 ) 42
assign (resid   215 and name HZ3 )(resid   227 and name HD* ) 45
assign (resid   215 and name HE3 )(resid   228 and name HA  ) 128
assign (resid   215 and name HE3 )(resid   229 and name HD1 ) 129
assign (resid   215 and name HA  )(resid   229 and name HA  ) 183
assign (resid   215 and name HZ3 )(resid   229 and name HA  ) 510
assign (resid   215 and name HZ3 )(resid   227 and name HB2 ) 1357
assign (resid   215 and name HE3 )(resid   229 and name HD2 ) 1392
assign (resid   215 and name HA  )(resid   230 and name HN  ) 588

  TRP 216
assign (resid   216 and name HZ2 )(resid   216 and name HH2 ) 6
assign (resid   216 and name HD1 )(resid   216 and name HA  ) 14
assign (resid   216 and name HD1 )(resid   216 and name HB2 ) 16
assign (resid   216 and name HZ3 )(resid   216 and name HE3 ) 94
assign (resid   216 and name HA  )(resid   216 and name HB1 ) 150
assign (resid   216 and name HA  )(resid   216 and name HB2 ) 154
assign (resid   216 and name HB2 )(resid   216 and name HB1 ) 639
assign (resid   216 and name HB1 )(resid   216 and name HZ3 ) 636
assign (resid   216 and name HN  )(resid   216 and name HB2 ) 121
assign (resid   216 and name HN  )(resid   216 and name HA  ) 123
assign (resid   216 and name HB1 )(resid   217 and name HN  ) 422
assign (resid   216 and name HB2 )(resid   217 and name HN  ) 423
assign (resid   216 and name HA  )(resid   217 and name HN  ) 424
assign (resid   216 and name HH2 )(resid   233 and name HG1*) 47
assign (resid   216 and name HZ3 )(resid   230 and name HB1 ) 54
assign (resid   216 and name HE3 )(resid   233 and name HG1*) 96
assign (resid   216 and name HB1 )(resid   230 and name HB1 ) 641
assign (resid   216 and name HD1 )(resid   230 and name HB1 ) 1388
assign (resid   216 and name HN  )(resid   228 and name HN  ) 670

  ARG 217
assign (resid   217 and name HA  )(resid   217 and name HB1 ) 327
assign (resid   217 and name HA  )(resid   217 and name HG2 ) 328
assign (resid   217 and name HA  )(resid   217 and name HD2 ) 329
assign (resid   217 and name HD2 )(resid   217 and name HB1 ) 944
assign (resid   217 and name HD2 )(resid   217 and name HG2 ) 1009
assign (resid   217 and name HD1 )(resid   217 and name HD2 ) 1301
assign (resid   217 and name HD1 )(resid   217 and name HG2 ) 646
assign (resid   217 and name HD1 )(resid   217 and name HB2 ) 649
assign (resid   217 and name HB1 )(resid   217 and name HG2 ) 1007
assign (resid   217 and name HG2 )(resid   217 and name HG1 ) 1130
assign (resid   217 and name HG1 )(resid   217 and name HA  ) 1124
assign (resid   217 and name HG1 )(resid   217 and name HD1 ) 1128
assign (resid   217 and name HA  )(resid   217 and name HB2 ) 1362
assign (resid   217 and name HN  )(resid   217 and name HB1 ) 148
assign (resid   217 and name HN  )(resid   217 and name HB2 ) 149
assign (resid   217 and name HE  )(resid   217 and name HD1 ) 688
assign (resid   217 and name HE  )(resid   217 and name HG2 ) 766
assign (resid   217 and name HA  )(resid   218 and name HN  ) 293
assign (resid   217 and name HB1 )(resid   227 and name HE* ) 69
assign (resid   217 and name HG1 )(resid   227 and name HE* ) 70
assign (resid   217 and name HD1 )(resid   227 and name HD* ) 98
assign (resid   217 and name HB1 )(resid   227 and name HD* ) 99
assign (resid   217 and name HA  )(resid   227 and name HA  ) 326
assign (resid   217 and name HG1 )(resid   227 and name HA  ) 163
assign (resid   217 and name HA  )(resid   227 and name HE* ) 325
assign (resid   217 and name HG1 )(resid   227 and name HD* ) 1144
assign (resid   217 and name HD1 )(resid   227 and name HN  ) 187
assign (resid   217 and name HG1 )(resid   227 and name HN  ) 485
assign (resid   217 and name HG2 )(resid   226 and name HN  ) 537
assign (resid   217 and name HA  )(resid   226 and name HN  ) 769
assign (resid   217 and name HA  )(resid   228 and name HN  ) 776
```

```
  ALA 218
assign (resid   218 and name HA  )(resid   218 and name HB* ) 156
assign (resid   218 and name HN  )(resid   218 and name HA  ) 87
assign (resid   218 and name HN  )(resid   218 and name HB* ) 88
assign (resid   218 and name HB* )(resid   219 and name HN  ) 496
assign (resid   218 and name HA  )(resid   219 and name HN  ) 497
assign (resid   218 and name HB* )(resid   228 and name HD1*) 981
assign (resid   218 and name HN  )(resid   227 and name HA  ) 292
assign (resid   218 and name HN  )(resid   226 and name HN  ) 543
assign (resid   218 and name HN  )(resid   228 and name HB  ) 682

  ARG 219
assign (resid   219 and name HA  )(resid   219 and name HG1 ) 141
assign (resid   219 and name HA  )(resid   219 and name HB2 ) 144
assign (resid   219 and name HA  )(resid   219 and name HG2 ) 146
assign (resid   219 and name HA  )(resid   219 and name HB1 ) 147
assign (resid   219 and name HD2 )(resid   219 and name HB2 ) 765
assign (resid   219 and name HD2 )(resid   219 and name HB1 ) 561
assign (resid   219 and name HD2 )(resid   219 and name HG1 ) 562
assign (resid   219 and name HD1 )(resid   219 and name HD2 ) 599
assign (resid   219 and name HB2 )(resid   219 and name HB1 ) 767
assign (resid   219 and name HG1 )(resid   219 and name HB2 ) 920
assign (resid   219 and name HG1 )(resid   219 and name HD1 ) 1426
assign (resid   219 and name HG2 )(resid   219 and name HD1 ) 1412
assign (resid   219 and name HN  )(resid   219 and name HB1 ) 193
assign (resid   219 and name HE  )(resid   219 and name HD2 ) 545
assign (resid   219 and name HE  )(resid   219 and name HG1 ) 546
assign (resid   219 and name HE  )(resid   219 and name HB2 ) 548
assign (resid   219 and name HE  )(resid   219 and name HD1 ) 712
assign (resid   219 and name HE  )(resid   219 and name HB1 ) 762
assign (resid   219 and name HB1 )(resid   220 and name HN  ) 385
assign (resid   219 and name HB2 )(resid   220 and name HN  ) 386
assign (resid   219 and name HA  )(resid   220 and name HN  ) 387
assign (resid   219 and name HD2 )(resid   223 and name HA1 ) 363
assign (resid   219 and name HD1 )(resid   223 and name HA2 ) 586
assign (resid   219 and name HD2 )(resid   221 and name HB2 ) 898
assign (resid   219 and name HG1 )(resid   223 and name HN  ) 653
assign (resid   219 and name HB2 )(resid   223 and name HN  ) 654
assign (resid   219 and name HD1 )(resid   223 and name HN  ) 655
assign (resid   219 and name HA  )(resid   225 and name HA  ) 142
assign (resid   219 and name HA  )(resid   225 and name HG1 ) 145
assign (resid   219 and name HG1 )(resid   225 and name HG2 ) 923
assign (resid   219 and name HG1 )(resid   225 and name HG1 ) 796
assign (resid   219 and name HG1 )(resid   225 and name HB1 ) 848
assign (resid   219 and name HG2 )(resid   225 and name HG2 ) 873
assign (resid   219 and name HD1 )(resid   224 and name HN  ) 420
assign (resid   219 and name HA  )(resid   226 and name HN  ) 542

  ASP 220
assign (resid   220 and name HA  )(resid   220 and name HB1 ) 211
assign (resid   220 and name HA  )(resid   220 and name HB2 ) 212
assign (resid   220 and name HB1 )(resid   220 and name HB2 ) 530
assign (resid   220 and name HN  )(resid   220 and name HB2 ) 125
assign (resid   220 and name HN  )(resid   220 and name HB1 ) 126
assign (resid   220 and name HN  )(resid   220 and name HA  ) 127
assign (resid   220 and name HA  )(resid   221 and name HG2 ) 214
assign (resid   220 and name HB1 )(resid   221 and name HD2 ) 523
assign (resid   220 and name HB1 )(resid   221 and name HG2 ) 526
assign (resid   220 and name HB2 )(resid   221 and name HN  ) 513
assign (resid   220 and name HB1 )(resid   221 and name HN  ) 514
assign (resid   220 and name HA  )(resid   221 and name HN  ) 516
```

```
  LYS 221
assign (resid   221 and name HA  )(resid   221 and name HB2 ) 445
assign (resid   221 and name HA  )(resid   221 and name HG1 ) 452
assign (resid   221 and name HE* )(resid   221 and name HB1 ) 690
assign (resid   221 and name HE* )(resid   221 and name HD1 ) 967
assign (resid   221 and name HE* )(resid   221 and name HG1 ) 693
assign (resid   221 and name HB2 )(resid   221 and name HG1 ) 901
assign (resid   221 and name HB1 )(resid   221 and name HA  ) 919
assign (resid   221 and name HD2 )(resid   221 and name HE* ) 1464
assign (resid   221 and name HD2 )(resid   221 and name HG1 ) 952
assign (resid   221 and name HD1 )(resid   221 and name HG1 ) 1265
assign (resid   221 and name HG2 )(resid   221 and name HE* ) 1129
assign (resid   221 and name HG2 )(resid   221 and name HG1 ) 1141
assign (resid   221 and name HG1 )(resid   221 and name HB1 ) 1263
assign (resid   221 and name HN  )(resid   221 and name HA  ) 200
assign (resid   221 and name HN  )(resid   221 and name HG1 ) 510
assign (resid   221 and name HN  )(resid   221 and name HG2 ) 511
assign (resid   221 and name HG2 )(resid   222 and name HE* ) 39
assign (resid   221 and name HE* )(resid   222 and name HD* ) 92
assign (resid   221 and name HG1 )(resid   222 and name HD* ) 93
assign (resid   221 and name HB2 )(resid   222 and name HB1 ) 889
assign (resid   221 and name HN  )(resid   222 and name HN  ) 517
assign (resid   221 and name HA  )(resid   222 and name HN  ) 604
assign (resid   221 and name HB1 )(resid   222 and name HN  ) 605
assign (resid   221 and name HG1 )(resid   222 and name HN  ) 607
assign (resid   221 and name HG2 )(resid   222 and name HN  ) 778
assign (resid   221 and name HG2 )(resid   224 and name HB1 ) 436
assign (resid   221 and name HG2 )(resid   224 and name HB2 ) 1126
assign (resid   221 and name HG2 )(resid   224 and name HN  ) 418
assign (resid   221 and name HD2 )(resid   224 and name HN  ) 421
assign (resid   221 and name HA  )(resid   223 and name HN  ) 656

  TYR 222
assign (resid   222 and name HE* )(resid   222 and name HD* ) 89
assign (resid   222 and name HE* )(resid   222 and name HA  ) 34
assign (resid   222 and name HE* )(resid   222 and name HB2 ) 35
assign (resid   222 and name HE* )(resid   222 and name HB1 ) 36
assign (resid   222 and name HA  )(resid   222 and name HB2 ) 324
assign (resid   222 and name HB1 )(resid   222 and name HB2 ) 598
assign (resid   222 and name HN  )(resid   222 and name HB1 ) 253
assign (resid   222 and name HN  )(resid   222 and name HA  ) 255
assign (resid   222 and name HN  )(resid   222 and name HE* ) 603
assign (resid   222 and name HN  )(resid   223 and name HN  ) 706
assign (resid   222 and name HN  )(resid   223 and name HA2 ) 763

  GLY 223
assign (resid   223 and name HA1 )(resid   223 and name HA2 ) 366
assign (resid   223 and name HN  )(resid   223 and name HA2 ) 282
assign (resid   223 and name HN  )(resid   223 and name HA1 ) 283
assign (resid   223 and name HN  )(resid   224 and name HN  ) 657

  SER 224
assign (resid   224 and name HA  )(resid   224 and name HB1 ) 332
assign (resid   224 and name HA  )(resid   224 and name HB2 ) 333
assign (resid   224 and name HB1 )(resid   224 and name HB2 ) 466
assign (resid   224 and name HN  )(resid   224 and name HB2 ) 143
assign (resid   224 and name HN  )(resid   224 and name HB1 ) 144
assign (resid   224 and name HN  )(resid   224 and name HA  ) 145
assign (resid   224 and name HB2 )(resid   225 and name HN  ) 452
assign (resid   224 and name HB1 )(resid   225 and name HN  ) 453
assign (resid   224 and name HA  )(resid   225 and name HN  ) 454
```

```
  GLU 225
assign (resid   225 and name HA  )(resid   225 and name HG2 ) 133
assign (resid   225 and name HA  )(resid   225 and name HG1 ) 134
assign (resid   225 and name HA  )(resid   225 and name HB1 ) 135
assign (resid   225 and name HG2 )(resid   225 and name HG1 ) 795
assign (resid   225 and name HG2 )(resid   225 and name HB1 ) 849
assign (resid   225 and name HB2 )(resid   225 and name HA  ) 829
assign (resid   225 and name HB2 )(resid   225 and name HG2 ) 835
assign (resid   225 and name HN  )(resid   225 and name HG1 ) 163
assign (resid   225 and name HN  )(resid   225 and name HG2 ) 164
assign (resid   225 and name HN  )(resid   225 and name HA  ) 166
assign (resid   225 and name HN  )(resid   225 and name HB1 ) 455
assign (resid   225 and name HG1 )(resid   226 and name HN  ) 540
assign (resid   225 and name HA  )(resid   226 and name HN  ) 541

  GLY 226
assign (resid   226 and name HA2 )(resid   226 and name HA1 ) 1294
assign (resid   226 and name HN  )(resid   226 and name HA2 ) 209
assign (resid   226 and name HA2 )(resid   227 and name HN  ) 486

  TYR 227
assign (resid   227 and name HB2 )(resid   227 and name HA  ) 1354
assign (resid   227 and name HB1 )(resid   227 and name HA  ) 608
assign (resid   227 and name HA  )(resid   228 and name HG11) 1480
assign (resid   227 and name HA  )(resid   228 and name HN  ) 665

  ILE 228
assign (resid   228 and name HA  )(resid   228 and name HG12) 220
assign (resid   228 and name HA  )(resid   228 and name HG2*) 221
assign (resid   228 and name HA  )(resid   228 and name HG11) 222
assign (resid   228 and name HA  )(resid   228 and name HD1*) 226
assign (resid   228 and name HG12)(resid   228 and name HG11) 954
assign (resid   228 and name HG12)(resid   228 and name HD1*) 963
assign (resid   228 and name HN  )(resid   228 and name HD1*) 659
assign (resid   228 and name HN  )(resid   228 and name HG12) 661
assign (resid   228 and name HA  )(resid   229 and name HD2 ) 217
assign (resid   228 and name HA  )(resid   229 and name HD1 ) 218
assign (resid   228 and name HG2*)(resid   229 and name HD2 ) 712
assign (resid   228 and name HG11)(resid   229 and name HD2 ) 713
assign (resid   228 and name HG2*)(resid   229 and name HD1 ) 724
assign (resid   228 and name HG11)(resid   229 and name HD1 ) 725
assign (resid   228 and name HG2*)(resid   232 and name HB2 ) 580
assign (resid   228 and name HG2*)(resid   230 and name HN  ) 585
assign (resid   228 and name HB  )(resid   230 and name HN  ) 586

  PRO 229
assign (resid   229 and name HA  )(resid   229 and name HB2 ) 508
assign (resid   229 and name HA  )(resid   229 and name HG2 ) 509
assign (resid   229 and name HA  )(resid   229 and name HB1 ) 514
assign (resid   229 and name HD2 )(resid   229 and name HG2 ) 714
assign (resid   229 and name HD2 )(resid   229 and name HG1 ) 1255
assign (resid   229 and name HD1 )(resid   229 and name HD2 ) 723
assign (resid   229 and name HD1 )(resid   229 and name HG1 ) 727
assign (resid   229 and name HD1 )(resid   229 and name HG2 ) 730
assign (resid   229 and name HB2 )(resid   229 and name HG2 ) 1036
assign (resid   229 and name HB2 )(resid   229 and name HB1 ) 1162
assign (resid   229 and name HB2 )(resid   229 and name HG1 ) 1039
assign (resid   229 and name HB1 )(resid   229 and name HG1 ) 1253
assign (resid   229 and name HG2 )(resid   229 and name HG1 ) 1254
assign (resid   229 and name HB1 )(resid   230 and name HN  ) 583
assign (resid   229 and name HB2 )(resid   230 and name HN  ) 584
assign (resid   229 and name HA  )(resid   230 and name HN  ) 587
assign (resid   229 and name HB1 )(resid   231 and name HA  ) 339
assign (resid   229 and name HB1 )(resid   231 and name HN  ) 626
```

```
assign (resid   229 and name HB2 )(resid   232 and name HN  ) 648
assign (resid   229 and name HB1 )(resid   232 and name HN  ) 711
assign (resid   229 and name HB2 )(resid   231 and name HN  ) 713

  SER 230
assign (resid   230 and name HA  )(resid   230 and name HB2 ) 700
assign (resid   230 and name HA  )(resid   230 and name HB1 ) 701
assign (resid   230 and name HB2 )(resid   230 and name HB1 ) 960
assign (resid   230 and name HN  )(resid   230 and name HA  ) 238
assign (resid   230 and name HB1 )(resid   231 and name HN  ) 714
assign (resid   230 and name HA  )(resid   233 and name HN  ) 699
assign (resid   230 and name HA  )(resid   233 and name HG1*) 702

  ASN 231
assign (resid   231 and name HA  )(resid   231 and name HB1 ) 331
assign (resid   231 and name HB2 )(resid   231 and name HA  ) 638
assign (resid   231 and name HB2 )(resid   231 and name HB1 ) 645
assign (resid   231 and name HD22)(resid   231 and name HD21) 217
assign (resid   231 and name HN  )(resid   231 and name HB1 ) 265
assign (resid   231 and name HN  )(resid   231 and name HB2 ) 266
assign (resid   231 and name HN  )(resid   231 and name HA  ) 267
assign (resid   231 and name HD22)(resid   231 and name HB2 ) 554
assign (resid   231 and name HD22)(resid   231 and name HB1 ) 555
assign (resid   231 and name HD21)(resid   231 and name HB1 ) 560
assign (resid   231 and name HD21)(resid   231 and name HB2 ) 561
assign (resid   231 and name HN  )(resid   232 and name HN  ) 652
assign (resid   231 and name HB1 )(resid   232 and name HN  ) 649

  TYR 232
assign (resid   232 and name HD* )(resid   232 and name HB2 ) 82
assign (resid   232 and name HD* )(resid   232 and name HB1 ) 84
assign (resid   232 and name HA  )(resid   232 and name HB1 ) 301
assign (resid   232 and name HB2 )(resid   232 and name HB1 ) 578
assign (resid   232 and name HA  )(resid   232 and name HB2 ) 1348
assign (resid   232 and name HA  )(resid   232 and name HD* ) 1439
assign (resid   232 and name HN  )(resid   232 and name HB1 ) 278
assign (resid   232 and name HN  )(resid   232 and name HB2 ) 279
assign (resid   232 and name HN  )(resid   232 and name HA  ) 280
assign (resid   232 and name HN  )(resid   232 and name HD* ) 650
assign (resid   232 and name HB1 )(resid   233 and name HG2*) 1209
assign (resid   232 and name HB1 )(resid   233 and name HN  ) 565
assign (resid   232 and name HA  )(resid   233 and name HN  ) 566
assign (resid   232 and name HN  )(resid   233 and name HN  ) 651
assign (resid   232 and name HN  )(resid   233 and name HG2*) 647

  VAL 233
assign (resid   233 and name HA  )(resid   233 and name HB  ) 165
assign (resid   233 and name HA  )(resid   233 and name HG1*) 168
assign (resid   233 and name HA  )(resid   233 and name HG2*) 169
assign (resid   233 and name HB  )(resid   233 and name HG1*) 813
assign (resid   233 and name HB  )(resid   233 and name HG2*) 1211
assign (resid   233 and name HG2*)(resid   233 and name HG1*) 1214
assign (resid   233 and name HN  )(resid   233 and name HA  ) 224
assign (resid   233 and name HN  )(resid   233 and name HG2*) 563
assign (resid   233 and name HN  )(resid   233 and name HG1*) 564
assign (resid   233 and name HG1*)(resid   234 and name HA  ) 239
assign (resid   233 and name HA  )(resid   234 and name HN  ) 524
assign (resid   233 and name HB  )(resid   234 and name HN  ) 527
assign (resid   233 and name HG2*)(resid   234 and name HN  ) 529
assign (resid   233 and name HG1*)(resid   234 and name HN  ) 530
```

```
  THR 234
assign (resid   234 and name HA  )(resid   234 and name HB  ) 236
assign (resid   234 and name HA  )(resid   234 and name HG2*) 238
assign (resid   234 and name HG2*)(resid   234 and name HB  ) 1034
assign (resid   234 and name HN  )(resid   234 and name HB  ) 205
assign (resid   234 and name HN  )(resid   234 and name HG2*) 206
assign (resid   234 and name HG2*)(resid   235 and name HN  ) 533
assign (resid   234 and name HA  )(resid   235 and name HN  ) 534
assign (resid   234 and name HB  )(resid   237 and name HB2 ) 440
assign (resid   234 and name HA  )(resid   237 and name HN  ) 366

  GLY 235
assign (resid   235 and name HA1 )(resid   235 and name HA2 ) 412
assign (resid   235 and name HN  )(resid   235 and name HA2 ) 207
assign (resid   235 and name HN  )(resid   235 and name HA1 ) 208
assign (resid   235 and name HA2 )(resid   236 and name HN  ) 413

  LYS 236
assign (resid   236 and name HA  )(resid   236 and name HB* ) 420
assign (resid   236 and name HA  )(resid   236 and name HD* ) 421
assign (resid   236 and name HE* )(resid   236 and name HD* ) 652
assign (resid   236 and name HB* )(resid   236 and name HD* ) 877
assign (resid   236 and name HG* )(resid   236 and name HE* ) 979
assign (resid   236 and name HN  )(resid   236 and name HB* ) 141
assign (resid   236 and name HA  )(resid   237 and name HN  ) 365

  LYS 237
assign (resid   237 and name HA  )(resid   237 and name HG2 ) 374
assign (resid   237 and name HE* )(resid   237 and name HD* ) 621
assign (resid   237 and name HN  )(resid   237 and name HA  ) 104
assign (resid   237 and name HB1 )(resid   238 and name HN  ) 727

  SER 238
assign (resid   238 and name HA  )(resid   238 and name HB2 ) 361
assign (resid   238 and name HA  )(resid   238 and name HB1 ) 362
assign (resid   238 and name HN  )(resid   238 and name HB1 ) 731
assign (resid   238 and name HA  )(resid   239 and name HN  ) 739

  ASN 239
assign (resid   239 and name HB1 )(resid   240 and name HN  ) 726
assign (resid   239 and name HA  )(resid   241 and name HD1*) 1499

  ASN 240
assign (resid   240 and name HA  )(resid   240 and name HB2 ) 310
assign (resid   240 and name HN  )(resid   240 and name HA  ) 725
assign (resid   240 and name HB2 )(resid   241 and name HN  ) 249
assign (resid   240 and name HA  )(resid   241 and name HN  ) 599

  LEU 241
assign (resid   241 and name HG  )(resid   241 and name HA  ) 1352
assign (resid   241 and name HD1*)(resid   241 and name HG  ) 1089
assign (resid   241 and name HN  )(resid   241 and name HD2*) 595
assign (resid   241 and name HN  )(resid   241 and name HG  ) 597
assign (resid   241 and name HA  )(resid   242 and name HN  ) 615
assign (resid   241 and name HG  )(resid   242 and name HN  ) 616
assign (resid   241 and name HD2*)(resid   242 and name HN  ) 707
assign (resid   241 and name HD2*)(resid   244 and name HD* ) 1100

  ASP 242
assign (resid   242 and name HA  )(resid   242 and name HB* ) 705
assign (resid   242 and name HN  )(resid   242 and name HA  ) 261
assign (resid   242 and name HN  )(resid   242 and name HB* ) 262
assign (resid   242 and name HA  )(resid   243 and name HN  ) 644
assign (resid   242 and name HB* )(resid   243 and name HN  ) 645
```

```
  GLN 243
assign (resid   243 and name HA  )(resid   243 and name HG* ) 418
assign (resid   243 and name HA  )(resid   243 and name HB1 ) 419
assign (resid   243 and name HG* )(resid   243 and name HB1 ) 836
assign (resid   243 and name HB2 )(resid   243 and name HA  ) 821
assign (resid   243 and name HB1 )(resid   243 and name HB2 ) 831
assign (resid   243 and name HN  )(resid   243 and name HB1 ) 275
assign (resid   243 and name HN  )(resid   243 and name HA  ) 277
assign (resid   243 and name HB1 )(resid   244 and name HN  ) 596
assign (resid   243 and name HA  )(resid   244 and name HN  ) 601

  TYR 244
assign (resid   244 and name HD* )(resid   244 and name HE* ) 58
assign (resid   244 and name HD* )(resid   244 and name HB1 ) 59
assign (resid   244 and name HA  )(resid   244 and name HD* ) 311
assign (resid   244 and name HA  )(resid   244 and name HB1 ) 316
assign (resid   244 and name HB1 )(resid   244 and name HB2 ) 654
assign (resid   244 and name HA  )(resid   244 and name HB2 ) 1369
assign (resid   244 and name HN  )(resid   244 and name HA  ) 252
assign (resid   244 and name HA  )(resid   245 and name HN  ) 226

  ASP 245
assign (resid   245 and name HB2 )(resid   245 and name HA  ) 696
assign (resid   245 and name HB2 )(resid   245 and name HB1 ) 1303
assign (resid   245 and name HA  )(resid   245 and name HB1 ) 1351
assign (resid   245 and name HN  )(resid   245 and name HB1 ) 225
```