INVESTIGATING THE ROLE OF PIN1 AND PHOSPHORYLATED AICD IN ALZHEIMER'S DISEASE USING NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY, ISOTHERMAL TITRATION CALORIMETRY AND RATIONAL DESIGN OF SMALL MOLECULES

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirement for the Degree of

Doctor of Philosophy

by

Soumya De

January 2011

INVESTIGATING THE ROLE OF PIN1 AND PHOSPHORYLATED AICD IN ALZHEIMER'S DISEASE USING NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY, ISOTHERMAL TITRATION CALORIMETRY AND RATIONAL DESIGN OF SMALL MOLECULES

Soumya De, Ph. D.

Cornell University 2011

Alzheimer's disease (AD) is a progressive neurodegenerative disorder caused by the degeneration of synapse and neuronal cell death in brain. The hallmarks of AD are "amyloid plaques" and "neurofibrillary tangles". The amyloid plaques contain deposits of Aβ-peptide and the neurofibrillary tangles contain hyper phosphorylated tau proteins. Recently it has been found that the soluble oligomers (dimers, trimers and tetramers) of the Aβ-peptide are the most neurotoxic. The Aβ-peptide is generated by proteolytic processing of Amyloid Precursor Protein (APP). In fact, APP undergoes proteolytic processing by two competing pathways: amyloidogenic pathway, which produces the toxic Aβ-peptide and non-amyloidogenic pathway, which is the major pathway in a healthy brain. Identification of the signaling events that regulate APP processing pathways and Aβ-peptide levels in the cell is one of the key questions in AD research.

Phosphorylation of T668 in APP intracellular domain (AICD) is an important signaling event that regulates APP processing. T668 phosphorylation results in the partitioning of pT668-P669 peptide bond into *cis* and *trans* isomers. The slow isomerization of these isomers is catalyzed by the only known human enzyme Pin1. Pin1 plays a neuroprotective role in AD. This work attempts to elucidate how Pin1's enzymatic action influences the phosphorylation dependent signaling in AD.

Combined NMR and ITC methodologies were used to determine kinetics and thermodynamics of pAICD:Pin1-WW domain interaction. These studies led to the derivation of a mathematical framework that predicts the impact of isomerization rates on the effective binding affinities of isomer-specific interactions and on the time-dependence of concentrations of isomer-specific complexes that could serve as entry points into modular protein interaction networks in the cell. Pin1 activity is proposed to act as a switch that connects and modulates flux through *cis*/*trans* isomer-specific interaction networks which are otherwise insulated in its absence. This work offers for the first time a direct molecular mechanism to explain the known influence of Pin1 activity on diverse cellular functions under normal and disease conditions. Furthermore, peptide analogues were designed and synthesized for *cis* and *trans* isomers of pAICD. These analogues are currently being used to identify the isomer-specific binding partners in the cell.

BIOGRAPHICAL SKETCH

Soumya was born to a loving couple, Samir and Ruma, in the steel city of Bhilai, India. He had a happy childhood with his caring parents. In his early childhood days, Soumya somehow got fascinated with scientific discoveries and dreamed to be a scientist. In his high school days, Soumya got addicted to detective novels and his dream career would oscillate between becoming a detective or a scientist. The first two years of undergraduate studies in the Indian Institute of Technology Kharagpur, helped him a lot in making up his mind. He found science to be as fascinating and intellectually exciting as solving mysteries. Moreover, science did suit his quiet demeanor very well. After completing his bachelors and masters in Industrial Chemistry from IIT Kharagpur, he came to Cornell to pursue his graduate studies. He was very fortunate to join the Nicholson lab and be part of exciting research. Meanwhile the advent of internet in IIT hostels, during his undergraduate studies, brought in good fortune for Soumya in his personal life too. He met Sucheta over the internet and quickly they became good friends. They would chat over the internet, email, talk over phone and also write letters to each other for the next seven years. While in his penultimate year of graduate studies in Cornell, Soumya and Sucheta got married and are living happily thereafter.

To our Elders

ACKNOWLEDGMENTS

First and foremost I would like to thank my intellectual Guru and mentor, Linda Nicholson, for being one of the best advisors that a graduate student can ask for. I am very grateful to her for teaching me the basics of scientific research. From her I learnt how to carefully plan experiments with paying attention to the minute details. "The devil is in the details" she would say. I also learnt how to analyze and interpret data. While working with Linda, I understood a very important aspect of experimental science that "Data never lies, it is our understanding that lacks in interpreting it correctly". Apart from paying attention to the minute details, Linda also taught me how to step back and look at the "big picture". Working under her guidance has truly been a wonderful experience for me.

I would also like to thank my committee members, Robert Oswald and Brian Crane, for serving on my committee and helping me with their insightful thoughts. I also thank our collaborators, Kun Ping Lu and Evgueni Kovriguine. Special thanks to Bhargavi and Jennifer, former Nicholson lab members, for guiding me during the beginning of my graduate life in the lab. I had a great time working alongside Lea, Ross, Jolita, Alex, Sarah, Carolyn and Monique in the Nicholson lab. I also had the good fortune to mentor and work alongside a few bright and intelligent undergraduate and graduate student such as Zahrae, Mohammed, Aaron, Han and Leila.

I am also very fortunate to be blessed with wonderful family members and friends. I would like to convey my special thanks to Suddhasattwa, Tathagata, Subhadeep, Jaya, Lekha, Debamita, Debashree, little Debashree, Anupam and Parag. I would like to express my heartfelt gratitude to my parents, my brother Sourav and my wife Sucheta for their love, support and encouragement.

Finally, thanks to you, the reader, for going through this dissertation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Aβ	Amyloid beta peptide
ACh	Acetyl choline
AChE	Acetyl choline esterase
AD	Alzheimer's disease
ADAM	A Disintegrin And Metalloproteinase
ADI	Alzheimer's Disease International
AICD	APP intracellular domain
APP	Amyloid precursor protein
BACE	β-site APP cleaving enzyme
COSY	NMR Correlation spectroscopy
ER	Endoplasmic reticulum
FDA	US Food and Drug Administration
HMBC	Heteronuclear Multiple Bond Correlation NMR experiment
HMQC	Heteronuclear multiple quantum coherence
HSQC	Heteronuclear single quantum coherence
IPTG	Isopropyl β-thio galacto pyranoside
ITC	Isothermal titration calorimetry
NFT	Neurofibrillary tangles
NMDA	N-methyl-D-aspartate
NMR	Nuclear Magnetic resonance
NOE	Nuclear overhauser enhancement
NOESY	Nuclear overhauser enhancement spectroscopy
ODE	Ordinary differential equation
PKC	Protein kinase C

RIP	Regulated intramembrane proteolysis
TACE	TNFα converting enzyme
TGN	Trans golgi network
wt	Wild type

CHAPTER 1

INTRODUCTION TO ALZHEIMER'S DISEASE *

Alzheimer's disease (AD) is a progressive neurodegenerative disorder caused by the degeneration of synapses and neuronal cell death in the brain (Mattson, 2004). It is characterized by progressive cognitive impairment that affects memory, orientation, logic and ability to perform everyday activities. AD is the most common type of dementia, the other types being vascular dementia, fronto-temporal dementia and dementia with Lewy bodies. Definite diagnosis of AD is done by the postmortem analysis of the patient's brain (Dickson, 1997) and should contain "amyloid plaques" and "neurofibrillary tangles" (NFT). The amyloid plaques have been designated as the hallmark of this disease. These plaques primarily contain deposits of a peptide called the Aβ-peptide and the neurofibrillary tangles contain hyper phosphorylated tau proteins. The risk of AD increases dramatically in individuals above the age of 65. Since this illness affects the human mind and the quality of life, it has a devastating effect on both the patients and their dear ones. In 2010 the total worldwide cost of dementia is estimated to be US$604 billion, 70% of which was incurred in Western Europe and North America (www.alz.org). The organization Alzheimer's Disease International (ADI) estimates that currently 35.6 million people are living with dementia worldwide in 2010 and this number would increase to almost 115.4 million by 2050.

Despite these staggering figures, efforts in drug development for Alzheimer's disease have intensified only recently, even though the disease was first described in 1906 (Alzheimer, 1907). Currently only five FDA-approved drugs for treatment of

---

* Parts of this chapter are based on a manuscript submitted to *Nature Chemical Biology*.

AD are available (www.nia.nih.gov/Alzheimer's). These drugs temporarily slow the progression of the disease for an average of 6-12 months for approximately half of the individuals who take them. These five drugs are donepezil, rivastigmine, tacrine, which belong to the class of cholinesterase inhibitors, galantamine which is an acetylcholinesterase inhibitor and memantine which is a NMDA (N-methyl-D-aspartate) receptor uncompetitive antagonist. Currently 56 drugs (Mangialasche et al., 2010) are in various stages of clinical trials and 21 in preclinical trials. These drugs target various aspects of AD such as amyloid formation, β-secretase inhibition, γ-secretase inhibition and modulation, α-secretase activation, prevention of Aβ aggregation and oligomerization, promotion of Aβ clearance by immunotherapy, targeting tau protein, targeting mitochondrial dysfunction and targeted delivery of nerve growth factors. Application of such diverse strategies for drug development in AD underlines the complex nature of this disease. In 1970s, the "cholinergic hypothesis" (Bartus et al., 1982) emerged as the first attempt to explain the molecular basis of this disease. Four out of the five FDA approved drugs are based on this hypothesis. In mid-1980s the Aβ-peptide, which consists primarily of 40 or 42 amino acid residues, was identified from the meningeal blood vessels of patients with AD (Glenner and Wong, 1984b) and Downs syndrome (Glenner and Wong, 1984a). Later, this Aβ-peptide was identified as the major component of the senile plaque (Masters et al., 1985). The gene of the β-amyloid precursor protein (APP), whose proteolytic cleavage produces the Aβ-peptide, was localized to chromosome 21, (Goldgaber et al., 1987; Kang et al., 1987; Robakis et al., 1987; Tanzi et al., 1987) which is linked to Downs syndrome. Earlier it was observed that Downs syndrome invariably leads (Olson and Shaw, 1969) to the neuropathology of AD. All these observations led to the proposal of the "amyloid hypothesis" which suggested that the aggregation of the Aβ-peptide leads to the neuropathology of AD.

The Cholinergic Hypothesis

The cholinergic system (Carlton, 1963) is associated with the neurotransmitter acetyl choline (ACh) and its function in the central nervous system. It is involved in storing and retrieving information in recent learning (Deutsch, 1971). Age related studies in monkeys (Bartus et al., 1978; Bartus and Johnson, 1976) and humans (Bartus et al., 1982) have shown that dysfunction in this system is related to age dependent memory deficit. The depletion of the cholinergic neurons in the ventral forebrain, resulting in disruption of cholinergic function in neocortex and hippocampus correlates well with dementia (Davies and Maloney, 1976; Wilcock et al., 1982). The cholinergic hypothesis (Bartus et al., 1982) states that the deficiency in cholinergic neurotransmission is critical to the cognitive and behavioral symptoms in AD. This led to the development of cholinesterase inhibitors (Wilkinson et al., 2004) as the first drugs for treatment in AD. These inhibitors block the catabolic enzyme acetylcholine esterase (AChE), which degrades the neurotransmitter ACh, and thus, increase the lifetime of ACh in the synapse resulting in prolonged cholinergic neurotransmission. Unfortunately, these drugs showed modest results (Frolich, 2002) as symptomatic treatment of AD. The cholinesterase inhibitors are therapeutically valuable in modestly improving the conditions of patients with moderate dementia but fail in counteracting or significantly delaying their resurgence. The inefficacy of the cholinesterase inhibitors, however, does not belie the cholinergic hypothesis. The depletion of the cholinergic neurons leading to the impairment of the cholinergic system is an important event in the progression of AD and thus, is an important therapeutic target. The significance of this hypothesis has been excellently summed up by Bartus (Bartus et al., 1986): "The cholinergic hypothesis necessarily suffers from a certain degree of inherent oversimplification and possible myopia. Yet, advantages gained by the research activities it helps stimulate and direct must not be

underestimated. It seems clear that the opportunity it provides for increasing our understanding of age-related memory problems and for developing an effective means of reducing the severity of the cognitive symptoms of aged brain and AD, provide adequate testimony for the continued value of the hypothesis as an important heuristic tool".

The Amyloid Hypothesis

The amyloid hypothesis (Hardy and Selkoe, 2002) states that the accumulation of amyloid β-peptide (Aβ) results in the formation of amyloid plaques which causes neurodegeneration in the brain. Later it was found that the soluble oligomers (dimers, trimers and tetramers) of the Aβ-peptide impair synaptic plasticity and behavior (Selkoe, 2008). At picomolar concentrations these oligomers disrupt the basal synaptic transmission and long term potentiation in the hippocampus. They were also found to interfere with the memory of the complex learned behavior in rats (Cleary et al., 2005). Currently, the Aβ-peptide oligomers are considered to be the key to neurodegeneration in AD (Klein et al., 2001; Selkoe, 2002, 2008).

The Aβ-peptide is generated by the proteolytic processing of the Amyloid Precursor Protein (APP), which is a transmembrane protein with a large extracellular domain, a transmembrane domain and a small intracellular domain (Figure 1.1a). APP is expressed in various tissues including the neurons and glia of the brain. Its pre-mRNA is alternatively spliced resulting in three isoforms APP_{695}, APP_{751} and APP_{770} (Kitaguchi et al., 1988). APP_{695}, containing 695 amino acids, is the most abundant in neuronal cells (Tanaka et al., 1989). APP is processed through two different pathways (Pastorino and Lu, 2006): the non-amyloidogenic and the amyloidogenic pathway (Figure 1.1a). Non-amyloidogenic processing of APP involves cleavage by α-secretase between Lys613 and Leu614 in the extracellular domain which produces the

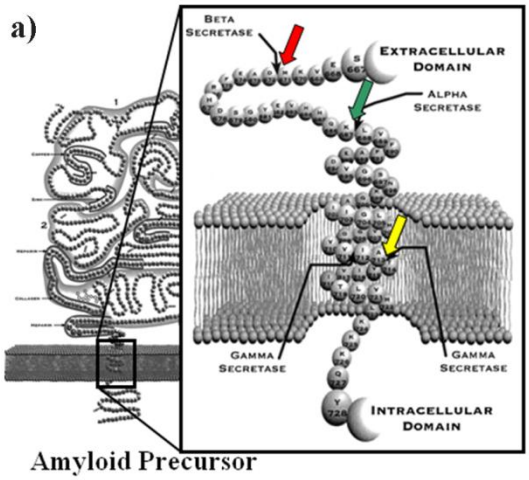
ectodomain (sαAPP) and an 83 residue transmembrane fragment C83. The potential candidates for the α-secretase are zinc metallanoproteinases such as ADAM9, ADAM10, TACE/ADAM17 and MDC9 (Allinson et al., 2003). The fragment C83 subsequently undergoes regulated intramembrane proteolysis (RIP) and is cleaved by γ-secretase. The γ-secretase complex (Shah et al., 2005) consists of four proteins: presenilin (PS1 or PS2), nicastrin, APH-1 (APH-1aL, APH-1aS or APH-1b) and PEN-2. Presenilin is the catalytic component of the γ-secretase complex (Wolfe et al., 1999). Thus two more fragments are produced: p3, a 23 amino acid non-toxic peptide and a 59 amino acid APP intracellular domain (AICD). The α-secretase cuts at the residue number 17 of the Aβ-peptide sequence resulting in p3 fragment which does not aggregate nor does it have any amyloidogenic activity. Amyloidogenic processing involves the cleavage of APP by β-secretase in the extracellular domain followed by γ-secretase cleavage in the transmembrane region resulting in the production of a soluble ectodomain (sβAPP), 40/42 amino acid Aβ-peptide and AICD. The major β-secretase identified in neurons is BACE1 (β-site APP cleaving enzyme 1) (Sinha et al., 1999; Vassar et al., 1999; Yan et al., 1999).

Subcellular localization of APP processing

APP is synthesized in the endoplasmic reticulum (ER) and then transported to the trans-Golgi-network (TGN) (Figure 1.1b), which is the steady-state residence site of APP in neurons (Greenfield et al., 1999; Hartmann et al., 1997). From the TGN the APP is transported to the cell surface by secretory vesicles. The APP can again get internalized into endosomes or lysosomes from the plasma membrane (Caporaso et al., 1994; Nordstedt et al., 1993). In neurons, APP is anterogradely transported along the axon which is mediated by its direct or indirect interaction to kinesin light chain subunit of kinesin-I (Kamal et al., 2000). The subcellular distribution of APP has

Figure 1.1: APP processing and its localization. a) The two proteolytic processing pathways of the transmembrane protein APP. The green, red and yellow arrows indicate the α, β and γ secretase cut sites on APP. The cleavage products generated in the two pathways are shown on the right. b) Location of APP and the three proteases in the cell. The major sites of activity of the three proteases are also shown. Figures a) and b) are adapted from Turner *et. al*. (Turner et al., 2003) with permission.

a)

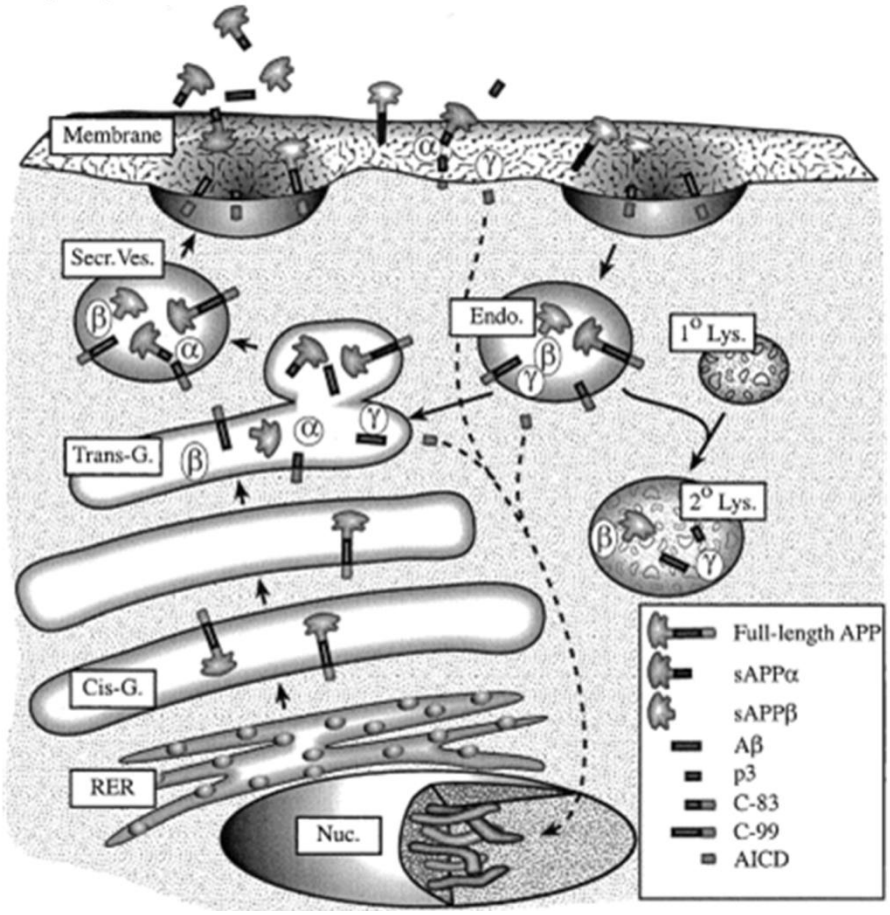


Amyloid Precursor Protein (APP)

Non-amyloidogenic Processing → Large extracellular domain + P3 fragment (23 a.a) + AICD

Amyloidogenic Processing → Large extracellular domain + Aβ-peptide (40 or 42 a.a) + AICD

b)



direct impact on its processing fate (Zhang and Xu, 2007) and hence, in the pathogenesis of AD (Figure 1.1b). The α-secretase cleavage of APP primarily occurs at the plasma membrane (Sisodia, 1992). The β-secretase BACE1 requires an acidic environment for optimal activity and is primarily localized in late Golgi/TGN and endosomes (Vassar et al., 1999). The γ-secretase complex resides mostly in the ER and TGN (Kovacs et al., 1996) and these are the major sites of APP processing by γ-secretase (Greenfield et al., 1999; Xu et al., 1997). The TGN, which is the major site of APP residence in neurons and also has acidic environment for β-secretase activity, is one of the major sites for the generation of secreted Aβ-peptide (Greenfield et al., 1999; Xu et al., 1997). The APP, which is not cleaved in the TGN, is transported to the plasma membrane where it can be cleaved by the α-secretase or endocytosed to the endosomal/lysosomal system. The endosomal/lysosomal system is another major site for Aβ production (Haass et al., 1993a; Haass et al., 1993b). In neurons APP is associated with axonal transport and the disruption of axonal transport has been shown to be an early pathogenic feature of AD (Lazarov et al., 2007; Stokin et al., 2005; Zhang et al., 2004). The mechanism of this disruption and its effect on AD is not clearly understood (De Vos et al., 2008). However, it is clear that the pathogenesis is mediated by increased Aβ production, since the disruption of the anterograde axonal transport of APP increases Aβ production (Stokin et al., 2005).

The APP Intracellular Domain (AICD)

The cytoplasmic tail of APP interacts with various cellular binding partners such as Fe65, Fe65L1, Fe65L2, Dab1, Mint1, Mint2, Mint3 and JNK-interacting protein family members (Ando et al., 2001; Borg et al., 1996; Duilio et al., 1998; Miller et al., 2006; Pietrzik et al., 2004; Sastre et al., 1998; Zambrano et al., 1998). It contains the widely recognized NPXY endocytosis signal which contributes to its

internalization (De Strooper and Annaert, 2000). Mutations within the YENPTY motif in APP inhibit its internalization and also decrease Aβ production (Perez et al., 1999). Binding to Fe65 facilitates the transport of AICD to the nucleus (Cao and Sudhof, 2001). The cytoplasmic tail also contains eight potential phosphorylation sites (Y653, T654, S655, T668, S675, Y682, T686 and Y687) (Lee et al., 2003b; Oishi et al., 1997; Zambrano et al., 2001). The residues T654 and S655 are phosphorylated by kinases such as protein kinase C (PKC) and calcium/calmodulin-dependent protein kinase II (CamKII). These phosphorylation events supposedly favor the non-amyloidogenic processing of APP (Buxbaum et al., 1993). On the other hand, phosphorylation of Thr668 has been found to promote the amyloidogenic processing of APP (Lee et al., 2003b), β-amyloid production (Phiel et al., 2003), neurodegeneration (Lee et al., 2003a; Leroy et al., 2002) and also neuronal differentiation (Ando et al., 1999). The T668 residue is followed by a proline residue and is phosphorylated by proline-directed kinases such as cyclin dependent kinases cdc2 and cdk5, the glycogen synthase kinase 3 (GSK3) and the c-Jun N-terminal kinase (JNK) (Lu et al., 2002a).

The phosphorylation of T668 in APP is an important regulatory event in the pathogenesis of AD. It was observed to be constitutively phosphorylated in the brain (Iijima et al., 2000). APP phosphorylated on T668 (pAPP) was found to colocalize with BACE1 in enlarged endosomes, the site of Aβ production, of AD hippocampal neurons (Lee et al., 2003b). The β-secretase cleavage product (C99) was found to be preferentially phosphorylated over the α-secretase cleavage product (C83) in AD brain and catecholaminergic cell lines (CAD cells). Use of kinase inhibitors and T668A mutation significantly reduced Aβ generation in rat cortical neurons (Lee et al., 2003b). The AICD is generated after the γ-secretase activity in both non-amyloidogenic and amyloidogenic processes but is mostly phosphorylated at T668

when processed through the amyloidogenic pathway (Lee et al., 2003b; Sodhi et al., 2008). The interaction complex of AICD and the adaptor protein Fe65 gets translocated to the nucleus and participates in gene transcription by forming a ternary complex with Tip60 (Cao and Sudhof, 2001). Several target gene candidates of this regulation have been reported such as tetraspanin *KAI1* (Baek et al., 2002), *GSK-3β* (Kim et al., 2003), *nephrilysin* (Pardossi-Piquard et al., 2005), *APP*, *BACE* and *Tip60* (von Rotz et al., 2004). It does not regulate the transcription of *ADAM10*, *PSEN1* and *Fe65* (von Rotz et al., 2004). It is not clear from these studies whether the AICD, in the ternary complex with Fe65 and Tip60, is phosphorylated at T668 or not. Deletion of the endocytic signal YENPTY in AICD or T668A mutation abolished nuclear transport of AICD (Chang et al., 2006). The use of inhibitors for the kinases cdc2, cdk5 and GSK-3β significantly reduced phosphorylation of T668 and nuclear translocation of AICD. Thus, the phosphorylation of T668 seems to be important for the nuclear transport and transcriptional activity of AICD which in turn increases APP phosphorylation and processing through the amyloidogenic pathway (Chang et al., 2006; Shin et al., 2007; Sodhi et al., 2008).

Pin1 in Alzheimer's disease

When T668 in APP is not phosphorylated, its local, TPEE, forms a transient helix capping box structure (Figure 1.2a) while most of the APP cytoplasmic tail transiently samples nascent helix and other conformations (Ramelot et al., 2000). The helix capping box structure stabilizes the *trans* conformation of the T668-P669 peptide bond. Normally, the *trans* conformation in the Xxx-Pro peptide bond is destabilized, as compared to other peptide bonds, due to steric hindrance of the imide bond and hence, a minor population corresponding to the *cis* conformation is observed along with the *trans* conformation (Fanghanel, 2003). Upon phosphorylation of T668, the

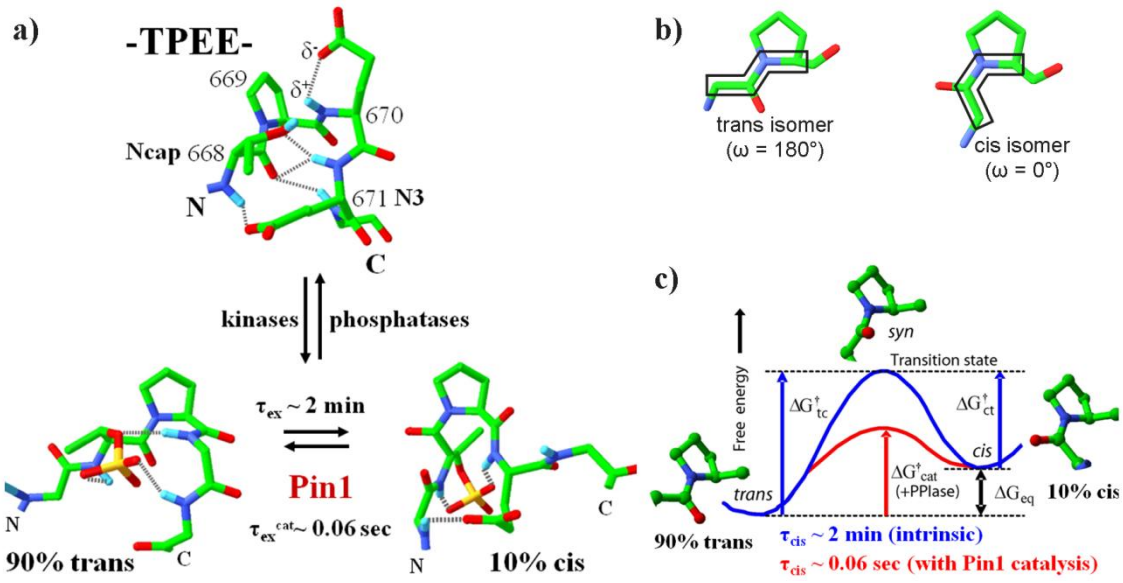


Figure 1.2 *Cis* and *trans* isomers. a) The TPEE region of the APP cytoplasmic tail forms a transient helix capping box structure which is predominantly *trans* with respect to the T668-P669 peptide bond. Upon phosphorylation, the helix capping box structure is disrupted and both the *trans* and *cis* isomers can be detected by NMR. Pin1 catalyzes their isomerization. b) The dihedral angles of the *trans* and the *cis* isomers for the Xxx-Pro peptide bond are shown. c) Double bond character of the peptide bond provides significant transition energy barrier between the inter-conversion of the two isomers. Enzymes like Pin1 lower this barrier and catalyze the isomerization.

helix capping box structure is disrupted and the minor *cis* population becomes detectable (Ramelot and Nicholson, 2001) (Figure 1.2a). The *cis* and the *trans* isomers in the Xxx-Pro peptide bond (Figure 1.2b) undergo slow isomerization (Fanghanel, 2003). Rotation about this peptide bond is restricted (Figure 1.2c) due to its partial double bond character and hence, is slow (Fischer, 2000). This isomerization is catalyzed by a class of enzymes known as peptidyl-prolyl isomerases (PPIases). Pin1 (Lu et al., 1996; Ranganathan et al., 1997) is a mammalian PPIase that specifically catalyzes the *cis-trans* isomerization of the pSer/Thr-Pro bond.

Pin1 has been shown to interact and catalyze the *cis-trans* isomerization of the pT668-P669 peptide bond in AICD (Pastorino et al., 2006) and is postulated to play a neuroprotective role in AD (Balastik et al., 2007). In healthy brains Pin1 is mostly expressed in the soluble fraction (Thorpe et al., 2004) while in diseased brains it co-localizes and co-purifies with NFTs leading to its depletion in soluble fraction (Lu et al., 1999a; Ramakrishnan et al., 2003). Also Pin1 knockout mice develop age-dependent neurodegeneration characterized by hyperphosphorylation of Tau protein leading to filament formation and Aβ-peptide formation and accumulation (Liou et al., 2003; Pastorino et al., 2006). Pin1 binds to the phosphorylated T231 (pT231) of the tau protein whose level of phosphorylation correlates with the progression of AD. Pin1's catalytic activity restores the ability of phosphorylated Tau to bind microtubules (Lu et al., 1999a) and also promotes Tau dephosphorylation by the phosphatase PP2A, which is *trans*-isomer specific (Zhou et al., 2000). In transgenic mouse models, Tau hyperphosphorylation leading to Tau aggregation and Tau filament formation has been observed in aged Pin1 knockout mice (Liou et al., 2003). Moreover, the subcellular localization of Pin1 also points towards its neuroprotective role in AD. Non-amyloidogenic processing of APP occurs at the plasma-membrane while the amyloidogenic processing occurs in the TGN and the endosomes/lysosomes.

Pin1 co-localizes with APP in the plasma-membrane and in clathrin coated vesicles but not in endosomes (Pastorino et al., 2006). The overexpression of Pin1 in cell culture experiments showed reduced Aβ secretion while a significant increase in Aβ42 was observed in the brains of Pin1 knockout mice (Pastorino et al., 2006). Taken together, these data point towards a novel regulatory role of Pin1 in AD pathogenesis through its interaction with the phosphorylated cytoplasmic tail of APP and phosphorylated tau protein.

Investigating the role of Pin1 in Alzheimer's disease

As presented above, there are compelling evidences that bring to light the role of Pin1 in AD, most probably a protective one. Nonetheless, the underlying molecular mechanism is still unclear. In fact, the direct interaction of Pin1 on APP still needs to be shown as influencing the APP processing pathways. So far, the evidences are circumstantial. Also the role of phosphorylation of T668 in APP cytoplasmic tail is not clear. Its impact on APP processing is well demonstrated but the underlying molecular mechanism is still not known. Phosphorylation of T668 in APP results in the emergence of a minor *cis* population with distinct structural features from the *trans* isomer. We hypothesized that the *cis* isomer may favor the amyloidogenic processing of APP and the *trans* isomer may favor the non-amyloidogenic processing (Pastorino et al., 2006). Pin1 activity ensures rapid re-equilibration between the two isomers whenever the equilibrium is disturbed due to dephosphorylation or phosphorylation, which are *trans*-isomer specific (Lu et al., 2002a), or due to isomer-specific binding events. Absence of Pin1 would result in prolonged nonequilibrium between the two isomers which might favor the amyloidogenic processing of APP. Thus, a key question would be the role of the *cis* isomer in the presence and absence of the *trans* isomer. Do these isomers, in the full-length APP, play some role in endocytocis?

Also, what are the isomer-specific binding partners and what are their roles in the cell? In the AICD fragment, when translocated to the nucleus, do these isomers transcribe different sets of genes favoring one or the other processing pathway of APP? These are all very intriguing questions which might go a long way in helping us understand the signaling pathways involved in AD and provide us with some therapeutic targets.

The enzyme Pin1 is involved in diverse cellular functions under normal and diseased conditions. It is involved in the regulation of cell-cycle progression, cellular stress response, immune response, neuronal differentiation and survival, and germ cell development (Lu and Zhou, 2007; Wulf et al., 2005). Although the central role of Pin1 in multiple biological and pathological processes is now clear, precisely how its enzymatic action influences these processes is still not known. The work presented here investigates the underlying molecular mechanism by which the enzymatic activity of Pin1 can influence such diverse processes. In chapters 2 and 3, we present a detailed study on the kinetics and thermodynamics of the interaction of pAPP659-682, a phophopeptide derived from the cytoplasmic tail of APP, and the isolated Pin1-WW domain. This study and subsequent derivation of a mathematical framework to predict the impact of *cis-trans* isomerization rates on effective binding affinities of isomer-specific interactions and on time-dependence of concentrations of isomer-specific complexes enabled us to propose a powerful and novel role of Pin1. We propose that Pin1 acts as a switch in coupling protein interaction network modules, modulates flux through them and also modulates the effective binding affinities of the isomer-specific binding partners. This provides for the first time a direct molecular mechanism to explain the known influence of Pin1 catalytic activity on diverse biological and pathological processes. In chapter 4 we focus our attention to the design of molecular tools and probes to study *cis* and *trans* isomer-specific binding interactions. We designed mimics for the *cis* and *trans* isomers of pAICD that are

being currently used in “fishing” experiments to identify the *cis* and *trans*-specific binding partners in the cell.

In summary, this work provides a novel molecular mechanism to explain the impact of the catalytic activity of Pin1 on diverse biological and pathological processes. We also provide NMR titration and lineshape analysis, ITC experiments and derived mathematical framework as a set of tools to investigate isomer-specific interactions and predict their impact on cellular functions. Additionally, *cis* and *trans* mimics are developed as chemical tools to identify the putative isomer-specific binding partners in the cell and investigate their functions. Thus, this work should pave the way for exciting and important research to understand the underlying signaling events in Alzheimer’s disease and develop therapeutics for its cure.

CHAPTER 2

KINETICS OF THE INTERACTION OF PIN1-WW DOMAIN AND PHOSPHORYLATED APP CYTOPLASMIC TAIL *

Introduction

The peptidyl prolyl isomerase Pin1 is composed of two domains (Figure 2.1), a catalytic domain (PPIase) and a substrate binding domain (WW) (Ranganathan et al., 1997). Pin1 belongs to a class of enzymes called PPIases which catalyze the *cis-trans* isomerization of the Xxx-Pro peptide bond (Figure 1.2b). PPIases are classified into three categories, namely, cyclophilins (Cyp) (Handschumacher et al., 1984), FK506 binding proteins (FKBPs) (Harding et al., 1989) and parvulins (Rahfeld et al., 1994). Pin1 belongs to the class of parvulins which is unique from the other two classes due to the presence of an additional binding domain apart from the catalytic PPIase domain. The binding domain, known as WW domain, is a peptide recognition module (PRM). PRMs are protein domains that have been extensively used by nature as functional units that recognize and bind specific motifs (Pawson and Nash, 2003). Multicellular organisms possess nearly a thousand such modules (Chothia et al., 2003) and employ them for protein-ligand interactions which form the basis of complex signaling pathways. PRMs are classified into families, each consisting of members that recognize specific motifs within the broad specificity of the family. The WW domain is typically between 38 to 40 amino acids in length and derives its name from two conserved tryptophan residues. The WW domain family is classified into five groups (I-V) based on specificity for different proline-rich recognition motifs (Verdecia et al., 2000). Among them group IV, which is present in Pin1 and other

---

* Parts of this chapter are based on a manuscript submitted to *Nature Chemical Biology*.

parvulins, is unique in that it is phosphorylation dependent, specifically recognizing phosphorylated Ser/Thr-Pro (pSer/Thr-Pro) motifs. As with other modular functional units, WW domains are typically found as a contiguous part of larger, multidomain proteins (Sudol, 1996). Such modular architecture of proteins where function (such as catalysis) and response to signal (such as binding activity) are physically separated, offer evolutionary advantages (Bhattacharyya et al., 2006; Jin et al., 2009) to the organism.

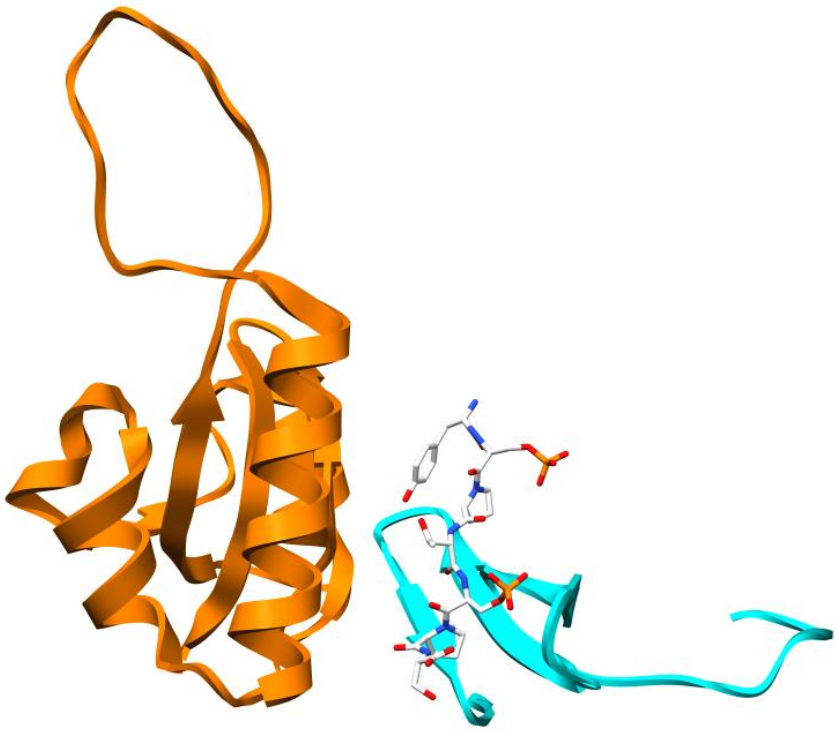


Figure 2.1: The ribbon structure of Pin1 (pdb code: 1F8A). The catalytic PPIase domain (golden) and the substrate binding WW domain (cyan) are shown along with a phosphorylated peptide bound to the WW domain.

The pSer/Thr-Pro motif is central to the proline-directed phosphorylation dependent signaling pathways (Lu et al., 2002a). Pin1 specifically catalyzes the isomerization of this motif and regulates diverse cellular processes such as progression of the cell-cycle, cellular stress response, growth-signal response, immune response and neuronal function (Lu and Zhou, 2007). Moreover, Pin1 has been implicated in

various diseases such as cancer (Lu et al., 2006) and Alzheimer's disease (Pastorino et al., 2006). Pin1 is found to be over expressed in human cancers (Lu et al., 2006) and is receiving significant attention as a drug target (Xu and Etzkorn, 2009). On the other hand, it plays a neuroprotective role and promotes the non-amyloidogenic processing of APP (Pastorino et al., 2006), thus preventing the production of neurotoxic Aβ-peptide, thought to be a major causative agent of Alzheimer's disease (Li et al., 2010). Even though Pin1 has been found to be important for various cellular processes, it still remains to be elucidated how its enzymatic action influences these signaling pathways. In order to understand the role of Pin1 and underlying molecular mechanism, it is important to elucidate the kinetics of its interaction with its substrates. The phosphorylated AICD (pAICD) is an important biological substrate of Pin1 and the elucidation of their interaction will provide us with a better understanding of Pin1's role in AD.

The interaction of Pin1 and its substrate involves six distinct states and can be represented by a reaction scheme (Figure 2.2a). In this reaction scheme, each step corresponds to substrate binding, on-enzyme catalysis or substrate transfer between two domains. Since the WW domain has no catalytic activity, the two WW domain bound states (C and F) are not connected (Lu et al., 1999b). We aim to determine the microscopic rates of each step and identify residues in Pin1 that regulate those steps. This would elucidate the detailed workings of this enzyme. Subsequent rational mutagenesis would allow us to understand and manipulate Pin1's influence on various signaling pathways. In order to characterize this complex reaction scheme, we utilized the logical approach of divide and conquer. Each of the two Pin1 domains are first studied independently and the resulting information is utilized as a restraint to model data obtained from the intact Pin1. While the WW domain is a key element of the Pin1 reaction scheme, it is not yet known whether its role is simply to increase the

effective concentration of substrate (Lu et al., 1999b), to facilitate direct transfer of substrate to the catalytic domain (Jacobs et al., 2003), to facilitate subcellular localization (Lu et al., 2002b) or some combination of these possibilities. However, it is clear that the WW domain is required for the in vivo functioning of Pin1, as shown by functional assay in yeast (Lu et al., 1999b). The characterization of the Pin1 reaction scheme (Figure 2.2a) should highlight the role of the WW domain in the functioning of the enzyme. Here we present the binding kinetics of the Pin1-WW domain and a peptide derived from the cytoplasmic tail of APP, phosphorylated at T668 corresponding to states A-C and F and their connecting steps (Figure 2.2).

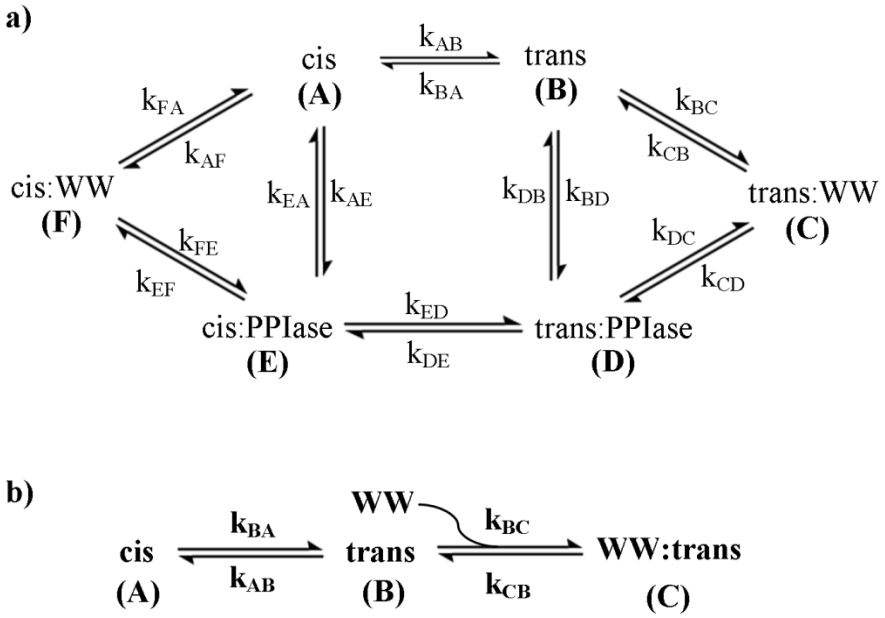


Figure 2.2: Pin1 reaction scheme. a) The reaction scheme for the interaction of Pin1 with its substrate. Each of the possible six states is denoted by a letter. The connections of the steps correspond to uncatalyzed (A-B) or catalyzed (D-E) isomerization, binding to WW domain (A-F and B-C) or PPIase domain (A-E and B-D) and substrate transfer (C-D and F-E). The microscopic rates for each step are shown. b) The reaction scheme for the interaction of isolated Pin1-WW domain with its substrate. The four microscopic rates of the two steps correspond to the same steps in the full-length Pin1 reaction scheme.

NMR Lineshape Analysis

We utilized NMR lineshape analysis as our primary tool to study the binding kinetics. The richness of information one might glean from the shape of an NMR peak has been known since the early days of NMR spectroscopy (Rao, 1989). However, this information has been largely untapped, having taken a backseat to the more high-profile applications of NMR to proteins, such as structure determination (Wuthrich, 1990) and relaxation-based NMR methods (Kay, 2005; Palmer, 2004; Palmer et al., 2005) for extracting motions on various timescales. NMR lineshape analysis has been sparsely used in the study of proteins, some examples being the determination of protein-ligand interaction kinetics (Gunther and Schaffhausen, 2002; Kovrigin and Loria, 2006) and enzyme catalysis (Kern et al., 1995). In comparison, the relaxation-based NMR methods such as CPMG (Korzhnev and Kay, 2008) and R1ρ (Massi et al., 2004) experiments have been widely applied to study μs-ms motions in proteins (Loria et al., 2008). While CPMG and R1ρ experiments can be used to characterize exchange between two or three states, lineshape analysis can handle complex reaction schemes involving multiple states. Additionally, lineshape analysis is well suited to study systems where titration experiments are of interest, such as protein-ligand interactions. Hence, the lineshape analysis method expands the scope of NMR studies of protein kinetics and thermodynamics, since it can be used to study any reaction scheme where the NMR lineshape can be systematically varied as a function of some known parameter such as ligand concentration, pH or temperature. Only one spectrum, such as ^{15}N-^{1}H HSQC, needs to be collected for each titration point and the collective dataset is simultaneously analyzed. The ^{1}H and ^{15}N lineshapes for each residue obtained from a ^{15}N-^{1}H HSQC experiment provide independent detectors of μs-ms motions for each residue giving additional restraints for the reaction scheme. It is important to note that since the number of fitted parameters

increases rapidly with the complexity of the reaction scheme and hence, for complex systems it is necessary to use additional independent experiments to restrain the parameter space.

The reaction scheme for the isolated WW domain is a simple model and served as an excellent test case for the lineshape analysis. We performed the analysis from both the protein and the peptide perspectives, which report on the same binding event, and compared the fitted parameters obtained. Excellent agreement in the fitted parameters from the two perspectives provided confidence in the applicability of this approach for characterization of more complex reaction schemes of the PPIase domain and, ultimately, intact Pin1. As long as enough experimental perspectives are accessible, accurate kinetic and thermodynamic parameters for a given multistate reaction can be determined.

Materials and Methods

Protein expression and purification

The WW domain vector was generously provided by Dr. Kun Ping Lu, Harvard Medical School. The gene sequence is encoded in pET28 a(+) expression vector, which has His_6 as the affinity tag, thrombin cleavage site and Kanamycin resistance. This construct (WWold) has residues 1-53 of full length Pin1 and has Gly-Ser-His as the three extra N-terminus residues after thrombin cleavage. A new construct was made by removing the extra N-terminus histidine residue and this construct was named as WW_Hdel. Furthermore, the three terminal residues Glu-Pro-Ala were also removed so that the construct has 50 WW domain residues and it was named WW50.

The vectors were transformed into BL21 DE3 *Escherechia coli* cells for protein synthesis. For the ^{15}N-labeled protein, the bacteria were grown in M9 minimal

media at 37°C with $^{15}NH_4Cl$ as the sole nitrogen source. For unlabeled protein, the bacteria were grown in 1L Lysogeny Broth (LB). The protein expression was induced with 1mM isopropyl β-D-1-thiogalactopyranoside when OD at 600 nm reached 0.8, and cells were harvested after 4 hours. The cells were lysed in 20ml lysis buffer (25mM HEPES, 150mM NaCl, 10mM Imidazole 100µl protease inhibitor cocktail, 5mg lysozyme and pH = 7.5) and sonicated on ice. The cell debris was removed by centrifugation and the supernatant passed through 2ml Ni-column pre-equilibrated with wash buffer (25mM HEPES, 150mM NaCl, 10mM Imidazole and pH = 7.5). The column was washed with 10 bed volumes of wash buffer and 5 bed volumes of cleavage buffer (20mM Tris-HCl, 150mM NaCl**,** 2.5mM $CaCl_2$ and pH=8.4) leaving 1 bed volume of the cleavage buffer in the column. Biotinylated thrombin (Novagen Thrombin Cleavage Capture Kit) was used to cleave the WW domain on the column. The protein concentration was measured by UV absorption at 280 nm and also by Bradford assay. The thrombin was removed by adding streptavidin-agarose beads which were removed by spin columns. Purity of the protein was verified by SDS polyacrilamide gel electrophoresis (PAGE). The purified ^{15}N labeled protein was dialyzed into final buffer (10mM Citric acid, 10mM NaCl, 1mM DTT, 5mM NaN_3, pH = 6.9) overnight at 4°C. The unlabeled protein was dialyzed into a final buffer containing 10mM HEPES, 10mM NaCl, 1mM DTT, 5mM NaN_3 and pH = 6.9. For the NMR experiments 5µl of PI cocktail was added to 300µl of protein sample.

Peptide Synthesis

The unlabeled peptide consists of the residues 659 to 682 and the labeled peptide consists of residues 659 to 679, respectively, from the cytoplasmic tail of APP. The unlabeled peptide is denoted as TAR2 and the labeled peptide as 15NTAR1. In earlier titration experiments, the unlabeled peptide used consisted of residues 659 to 679 and was denoted as TAR1. The purified 15NTAR1 and TAR2 phosphopeptides

were purchased from the Tufts University Core Facility, Boston. Along with the peptide, analytical reverse phase HPLC analysis (>99%) and MALDI-TOF mass spectrometry analysis was also obtained. The relevant peptide was weighed out, as required, dissolved in 200µl de-ionized H_2O, adjusted to pH ~7 using NaOH and lyophilized. The lyophilized peptide was dissolved in the final NMR buffer. The pH of the sample adjusts to the pH of the buffer. The concentration is measured using tyrosine extinction coefficient of 1400 $M^{-1}cm^{-1}$ (Creighton, 1993). Rough concentration measurements for the 15NTAR1 and TAR1 peptides were done based on mass as it lacks any tyrosine or tryptophan residues. A correction for the concentration of the peptide was used as a fitted parameter in the lineshape analysis.

NMR titration experiments

Two titration experiments were performed: ^{15}N peptide (15NTAR1) titrated with unlabeled WW-domain (WWold) and ^{15}N WW-domain (WW50) titrated with unlabeled peptide (TAR2). For the ^{15}N- peptide titration, the most saturated sample was prepared first containing 0.65 mM peptide and 1 mM WW domain. In the subsequent titration steps, part of the previous sample was diluted with 0.65 mM peptide to get the required molar ratio and a final volume of 300 µl. Eight titration points were collected; the ^{15}N peptide (15NTAR1) concentration kept constant at 0.65 mM while the WW (WWold) concentration was 0, 0.01, 0.05, 0.1, 0.2, 0.33, 0.5 or 1 mM. Similarly for the ^{15}N WW titration, the most saturated sample, 0.4 mM ^{15}N WW domain (WW50) and 6.6 mM peptide (TAR2), was prepared first. The subsequent steps were obtained by diluting with 0.4mM ^{15}N WW domain. Eleven titration points were collected with peptide concentrations of 0, 0.044, 0.088, 0.176, 0.264, 0.352, 0.44, 0.88, 2.2, 4.4 and 6.6 mM. For each titration point a ^{15}N-^{1}H HSQC spectrum was collected. All experiments were carried out at 25°C on a Varian Inova 600-MHz spectrometer. For the labeled protein titrations 2048×256 complex data points and for

the labeled peptide titrations 2048×128 complex data points were acquired. The spectra were processed using NMRpipe (Delaglio et al., 1995) and visualized with Sparky (Goddard, 2008). The exponential decay window function was applied to all data in order to preserve the lorentzian lineshape and then zerofilled.

The construct WWold was initially used for the titration experiments. Concentrated TAR1 was added to ^{15}N-WW for each titration point. The protein concentration was 1mM and the ratio of the peptide to protein used was 0, 0.156, 0.320, 0.493, 0.675, 0.867, 1.071, 1.824, 2.531 and 6.828.

Lineshape analysis

The lineshape analysis was performed according to the method described by Beach (Beach et al., 2005) and Gunther (Gunther and Schaffhausen, 2002) using the package BiophysicsLab (Kovrigin, E.L., manuscript in preparation, http://biophysicslab.net). BiophysicsLab treats 1D slices of each dimension (^{1}H, ^{15}N) of each peak from a titration series as an independent dataset. The area under each peak is normalized in order to eliminate the effect of line broadening in the orthogonal dimension. For the lineshape fitting, the equilibrium populations of the free *trans* isomer, free WW domain and bound WW domain were determined from the reaction scheme (Figure 2.2b). The populations (p_i where i is A, B, etc.) of the free and bound species are expressed in a column vector. The relaxation rate constant for each species (R_2) are calculated from the linewidths of the free and the bound peaks and expressed as diagonal elements of a square matrix (left term, equation 2.1a). Similarly, the chemical shifts (ν_i) of the free and the bound species are expressed in a matrix (right term, equation 2.1a). The rate constants (k_i) are expressed in the rate-matrix which is derived from the equilibrium rate equations of each species (middle term, equation 2.1a). Finally, the spectrum at the frequency ν is calculated using equation (2.1c). The independent variable chemical shift parameter (ν) which constitutes the

independent variable dimension in the simulated spectrum is represented by equation (2.1b).

$$M_1 = \begin{pmatrix} R_2^A & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & R_2^B \end{pmatrix} - \begin{pmatrix} -k_{AB} & \cdots & k_{BA} \\ \vdots & \ddots & \vdots \\ k_{AB} & \cdots & -k_{BA} \end{pmatrix} - 2\pi i \begin{pmatrix} \nu_A & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \nu_B \end{pmatrix} \quad (2.1a)$$

$$M_2 = 2\pi\nu \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix} \quad (2.1b)$$

$$\text{Spectrum} = \text{Real}\left(\sum \left(\text{Inv}(M_1 + M_2) \times \begin{pmatrix} p_A \\ \vdots \\ p_B \end{pmatrix} \right) \right) \quad (2.1c)$$

The quality of the fitting was determined from the sum of squares of the residuals (SS) which is determined as $SS = \sum_j^M \sum_i^{N_j} \left(f_{ij} - y_{ij}\right)^2 / \sigma_j^2$, where f_{ij} and y_{ij} are the calculated and actual NMR signal at a frequency ν_i respectively, σ_j is the noise rmsd of spectrum 'j', N_j is the number of data points in spectrum 'j' and M is the number of titration points. For the practical aspects of lineshape analysis using BiophysicsLab, see Appendix 1.

Results

Isomer-specific interaction of pAPPc with the Pin1 WW domain

In order to study the Pin1-WW domain (hereafter referred to as WW domain) in action we chose one of its natural substrates, a peptide (pAPP659-682) derived from the amyloid precursor protein (APP) phosphorylated at T668. We looked at the interaction from both the protein and the peptide perspective by titrating in the unlabeled peptide into ^{15}N-labeled WW domain and the unlabeled WW domain into the ^{15}N-labeled peptide respectively. For each titration point ^{15}N-^{1}H HSQC spectra

were collected. The pT668-P669 peptide bond of the substrate exists in two isomeric forms, *cis* and *trans.* The labeled peptide contains ^{15}N-labels for the amide nitrogen of the residues E667 and V670, which are sensitive to the isomeric state of the pT668-P669 peptide bond. The ^{15}N-^{1}H-HSQC spectra of this peptide shows distinct *cis* and *trans* peaks for both the residues (Figure 2.3). Upon titration with unlabeled WW domain the peaks corresponding to the *trans* isomer exhibit fast exchange relative to the NMR timescale, resulting in peak shifts from the free to the bound positions (Figure 2.3). The corresponding peaks of the *cis* isomer do not move appreciably and gradually diminish via conversion to the *trans* isomer as the concentrations of free *cis*, free *trans* and the bound *trans* equilibrate with increasing concentration of the WW domain. In the final titration spectrum only the bound *trans* peaks are observed. Thus, the interaction of this peptide to the WW domain can be represented by two steps, isomerization of the *cis* and *trans* isomers, and binding of the *trans* isomer to the WW domain (Figure 2.2b).

Previously, it was reported that addition of WW domain to this peptide caused significant peak shifts for both the *cis* and *trans* isomers, indicating similar affinity for these distinctly different structures (Pastorino et al., 2006). In this earlier work, the concentrations were lower (0.2 mM peptide, 0 and 0.2 mM WW), which would reduce the relative population of bound trans (c.a 0.7) and cause the trans peak corresponding to E670 to be broadened. Furthermore, the position of E670 *cis* peak is highly sensitive to pH in the mildly acidic to neutral pH range due to its interaction with the titrating phosphate group, while the V668 *cis* peak remains unperturbed (Figure 2.4). Accounting for pH and concentration effects, the previously reported and current results demonstrate that the *cis* interaction with the WW domain is weak, and the behavior of the peak corresponding to the E670 *cis* isomer is highly sensitive to sample conditions.

From the protein perspective the effect of the peptide binding is felt throughout the core of the protein. Out of 26 residues (G10 to E35) that constitute the core structure of the WW domain (Figure 2.5a), 22 backbone amides show appreciable chemical shift change in response to the peptide binding (Figure 2.5b). These chemical shift perturbations provide us with several reporters throughout the protein that report on the same binding event. In order to obtain the kinetics of this interaction, lineshape analysis was performed on both sets of titration data.

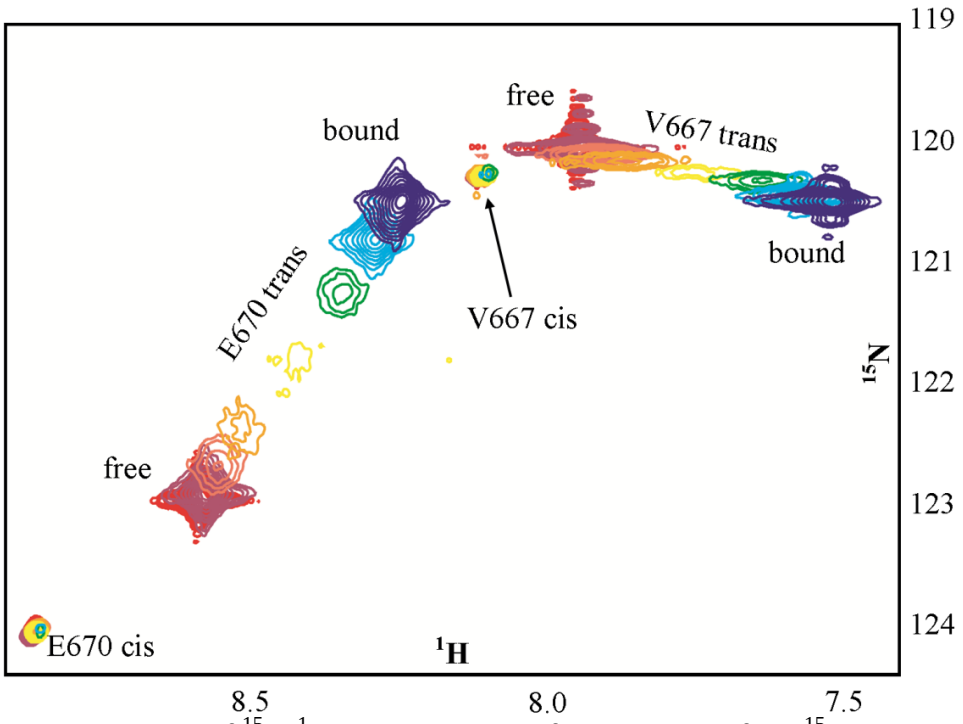


Figure 2.3: Overlay of ^{15}N-^{1}H HSQC spectra of the titration of the ^{15}N-labeled peptide with isolated WW domain. Very little peak shift is observed for the cis peaks whereas the trans peaks move considerably. The *cis* peaks are not visible in the final spectrum (dark blue) indicating complete conversion to bound *trans*.

Determination of kinetic parameters for interaction of *trans* isomer and WW domain

In addition to confirming the residues in the WW domain that are interacting with the pAPP659-682 peptide, the chemical shift changes along with the change in

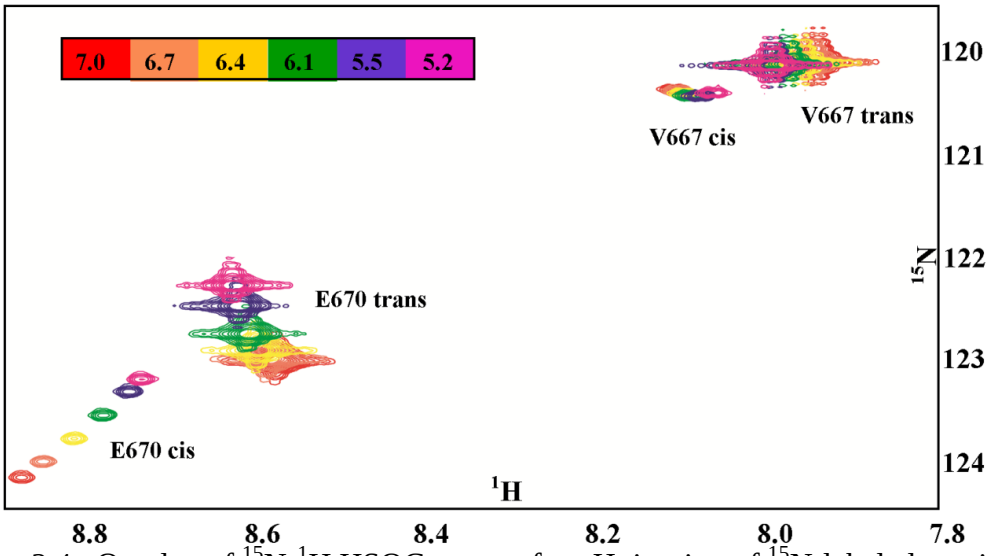


Figure 2.4: Overlay of ^{15}N-^{1}H HSQC spectra for pH titration of ^{15}N-labeled peptide (15NTAR1). E670 is highly sensitive to pH titration.

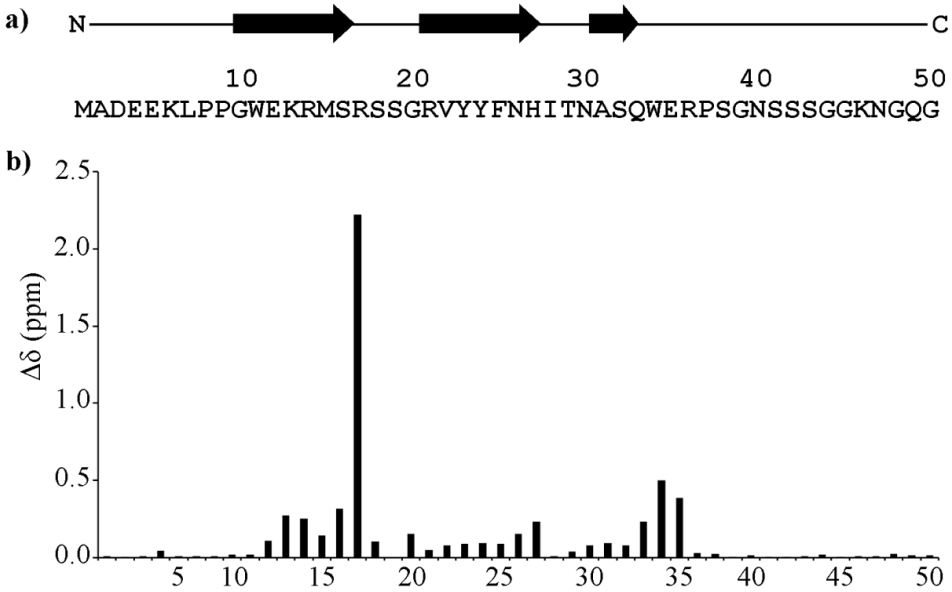


Figure 2.5: a) Primary sequence and secondary structure of the WW domain denoting the core of the protein. b) Composite chemical shift change (Ayed et al., 2001; Mulder et al., 1999) for each residue between the free and bound states, where $\Delta\delta = \sqrt{\left(\delta_H^{free} - \delta_H^{bound}\right)^2 + (0.154)^2 * \left(\delta_N^{free} - \delta_N^{bound}\right)^2}$. The chemical shifts δ^{free} and δ^{bound} are obtained from the free and the most saturated spectra of isolated ^{15}N WW domain respectively.

NMR lineshapes also provide information on the kinetics and binding affinity of this interaction. The chemical shift difference ($\Delta\omega$) between two exchanging states (e.g. free and bound) determines whether a chemical exchange process is in fast ($k_{ex} > \Delta\omega$) intermediate ($k_{ex} \sim \Delta\omega$) or slow ($k_{ex} < \Delta\omega$) timescale, where k_{ex} is the rate constant of exchange. In the fast timescale a single sharp peak is observed at the population weighted average chemical shift. Intermediate timescale results in a single broad peak corresponding to the *trans* isomer in the labeled peptide move from the chemical shift of the free state to that of the bound state (Figure 2.3), indicating that the interaction occurs on the fast timescale. During the titration the peaks progressively broaden with the addition of peptide and then narrow as saturation is reached. Similar behavior is observed for the lineshapes in the labeled WW domain upon interaction with the peptide. Information regarding the kinetics of the binding event is contained in these broad and narrow lineshapes and can be extracted by lineshape analysis(Kern et al., 1995).

Lineshape analysis on titration data from both perspectives was performed whereas in the slow timescale distinct peaks are observed at the respective chemical shifts of the exchanging species. Upon interaction with the WW domain, the peaks using the reaction scheme (Figure 2.2b) for the WW domain and pAICD interaction. The populations of the free *cis* isomer, the free *trans* isomer, the free WW domain and the bound WW domain were calculated according to this scheme. Since the intrinsic *cis* to *trans* inter-conversion rate is very slow (timescale on the order of minutes, vide infra), this rate does not affect the NMR lineshapes and only affects the populations of the two isomers. The titration experiments were performed under equilibrium conditions and hence, the free isomers maintain their equilibrium populations. The NMR lineshapes were fit to a two state model (free and bound states) while accounting for the thermodynamic coupling to the *cis-trans* equilibrium. The

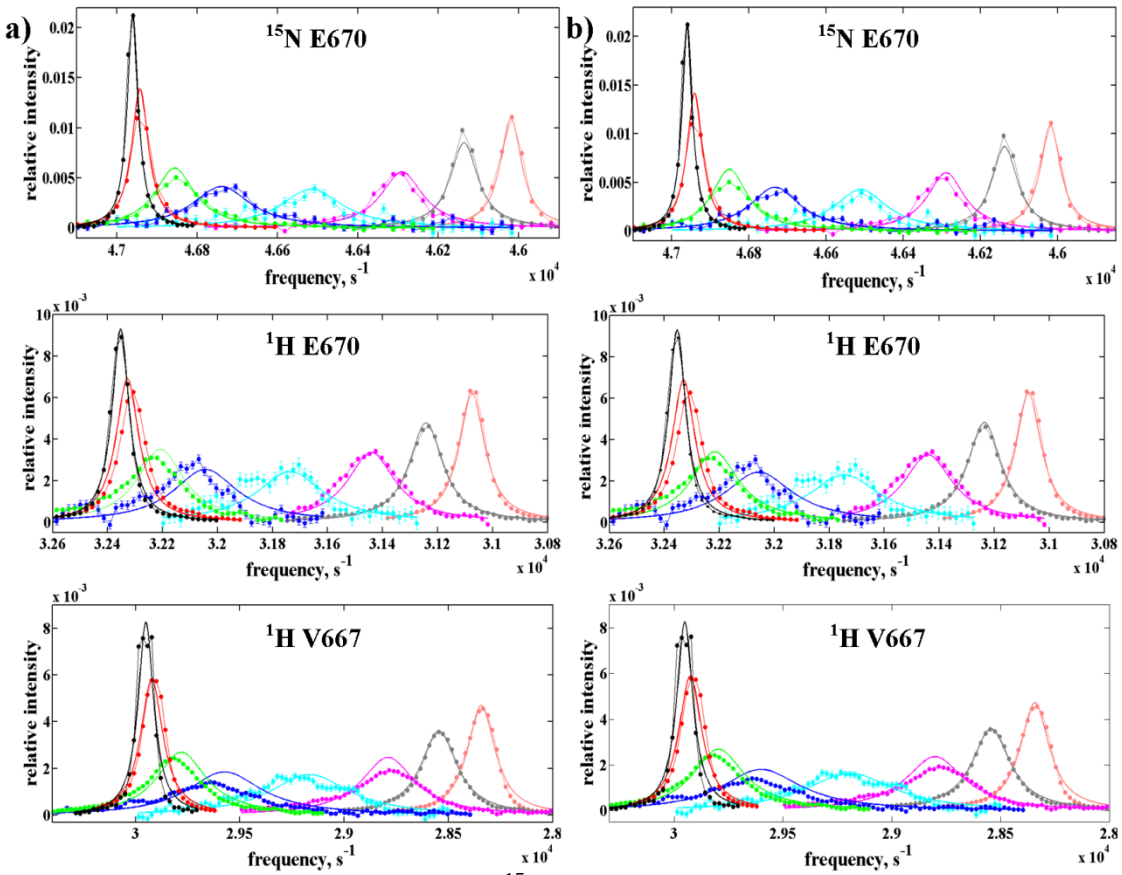


Figure 2.6: a) Individual fitting of the ^{15}N-peptide titration lineshapes. b) Simultaneous fitting of the three datasets. The quality of the fit are the same for individual and simultaneous fittings. The filled circles (•) represent the NMR data and the solid line is the simulated spectrum.

populations were determined from the volumes of the peaks corresponding to the *trans* and the *cis* isomers in the ^{15}N-^{1}H HSQC spectra of the peptide (Figure 2.3). The simulated lineshapes based on the two-state model fits the ^{15}N-labeled peptide titration data well (Figure 2.6a). The fitting parameters are the binding constant (K_D), the off rate (k_{CB}) and the chemical shift of the bound species (δ_{bound}), since saturation is not reached in these titration experiments. We designate the first two parameters (K_D and k_{CB}) as global parameters since they should be the same for all residues, and the chemical shift of the bound state, δ_{bound}, as a local parameter since it is unique to each

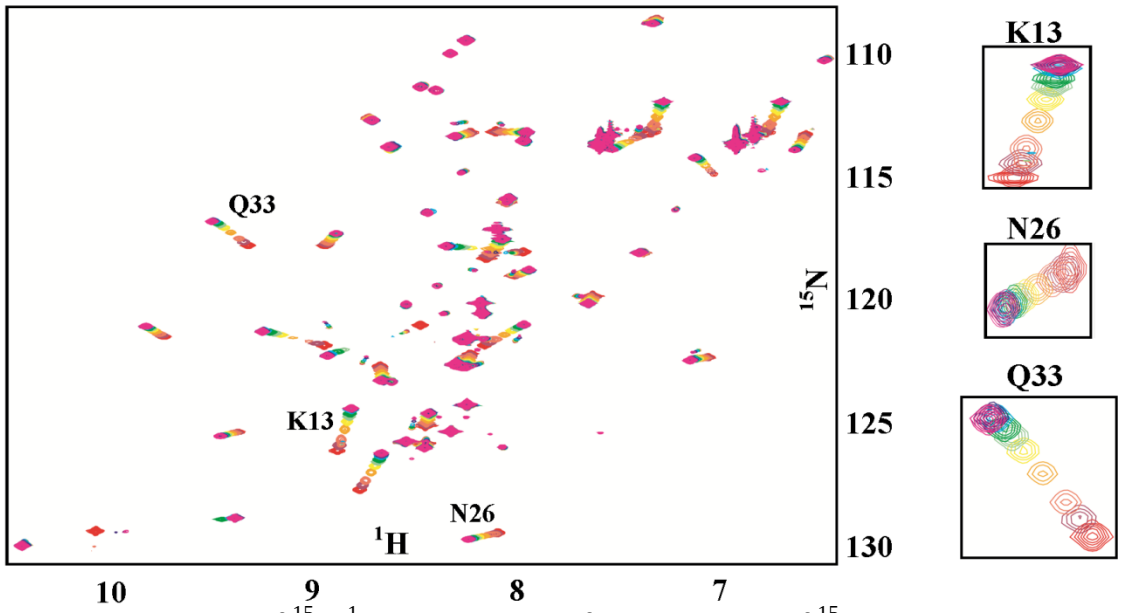


Figure 2.7 Overlay of ^{15}N-^{1}H HSQC spectra for the titration of ^{15}N-WW domain with unlabeled peptide. Peaks for three representative residues are shown on the right column. These peaks were used in simultaneous fitting.

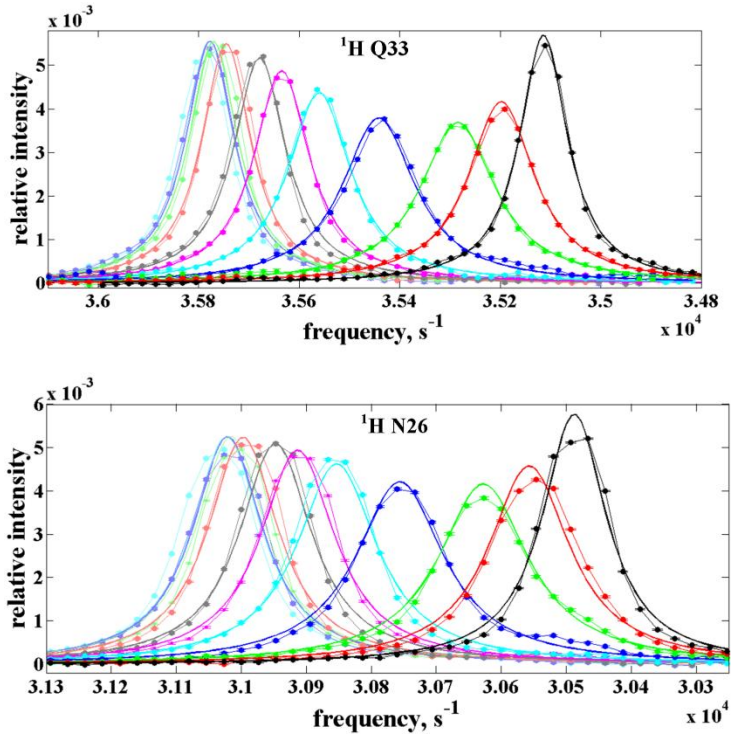


Figure 2.8: Representative examples of the individual fitting of the lineshapes from isolated ^{15}N-WW domain titration data.

Table 2.1 Fitted parameters from individual fitting of three datasets of the labeled peptide titrated with unlabeled WW domain

	k_{CB} (s^{-1})	k_{BC} ($M^{-1}s^{-1}$)	K_D (μM)
E22H	2511.80	4.9×10^7	50.7
E22N	1962.21	5.0×10^7	38.9
V19H	2461.65	4.8×10^7	50.3

Table 2.2: Comparison of the sum-of-squares (SS) for individual (I) and simultaneous (S) fitting.

Residue	**I (×10^5)**	**S (×10^5)**	**S/I**	**Residue**	**I (×10^5)**	**S (×10^5)**	**S/I**
E12H	3.95	4.81	1.219	N26H	4.46	4.66	1.045
K13N	11.8	20.9	1.776	H27N	33.1	61.5	1.857
R14N	43.4	78.9	1.819	H27H	4.59	7.47	1.627
M15H	5.92	9.19	1.551	N30N	33.8	38.5	1.138
S16H	1.46	2.45	1.676	N30H	13.0	22.4	1.719
G20H	3.17	6.92	2.183	A31H	8.77	15.2	1.731
V22N	38.2	69.3	1.815	S32H	9.86	9.97	1.012
Y23N	23.8	56.1	2.361	Q33H	2.50	3.50	1.400
Y24N	10.7	16.2	1.519	W34N	14.0	18.1	1.290
Y24H	2.41	3.28	1.358	W34H	1.19	3.70	3.100
F25H	3.83	5.57	1.454	W34eH	19.1	61.3	3.206
				E35H	9.85	12.8	1.302

residue. The two dimensions, ^{15}N and ^{1}H, for each peak are treated as separate datasets. The ^{15}N dimension of the V667 trans peak was not considered due to small chemical shift change upon binding. The three datasets for the two *trans* peaks were fit both individually (Table 2.1) and simultaneously (Figure 2.6b). The simultaneous fitting yields $K_D = 42.5 \pm 1.9$ μM, $k_{CB} = 2260 \pm 60$ s^{-1} and $k_{BC} = 5.3 \pm 0.2 \times 10^7$ $M^{-1}s^{-1}$ from the ligand's perspective, where best fit and the 95% confidence interval is reported for each parameter. The confidence intervals are computed from Monte-Carlo simulations.

To obtain the kinetics of the same binding event from the WW domain's perspective, lineshape analysis was performed on the ^{15}N WW domain titration data. Out of the 22 residues which display appreciable chemical shift change upon peptide binding, 17 residues gave rise to peaks that were resolved through the entire course of the titration (Figure 2.7). Individual fitting of these 17 residues yielded reasonable agreement in global parameters (Figure 2.8 and Table 2.2). Simultaneous fitting (Figure 2.9) was performed on datasets corresponding to seven residues (K13N, R14N, G20H, N26H, H27H, Q33H and W34H) that undergo the most chemical shift change and hence, contain the maximum information of the binding event. In excellent agreement with the global parameters obtained from the ligand perspective, the WW domain perspective simultaneous fitting yielded $K_D = 34 \pm 0.3$ μM (best fit ± 95% confidence interval) and $k_{CB} = 2170 \pm 40$ s^{-1}, and $k_{BC} = 6.4 \pm 0.1 \times 10^7$ $M^{-1}s^{-1}$, where the confidence intervals are computed from monte carlo simulation. In order to check whether these fitted parameters are truly global for all the 17 residues, these parameters were used to simulate their lineshapes. The quality of fit, measured by the sum of squares of the residuals (SS), were very similar to those obtained from the individual fitting of the 17 residues. Thus, the parameters obtained from the simultaneous fitting truly represent the global behavior of the binding interaction. As

a further step to global modeling of the binding mechanism, seven datasets from the ^{15}N WW domain and the three datasets from the ^{15}N peptide were simultaneously fit (Figure 2.10) which resulted in $K_D = 35.1 \pm 0.4$ μM, $k_{CB} = 2220 \pm 30$ s^{-1} and $k_{BC} = 6.3\pm 0.1 \times 10^7$ $M^{-1}s^{-1}$, where the confidence intervals are computed from Monte-Carlo simulations.

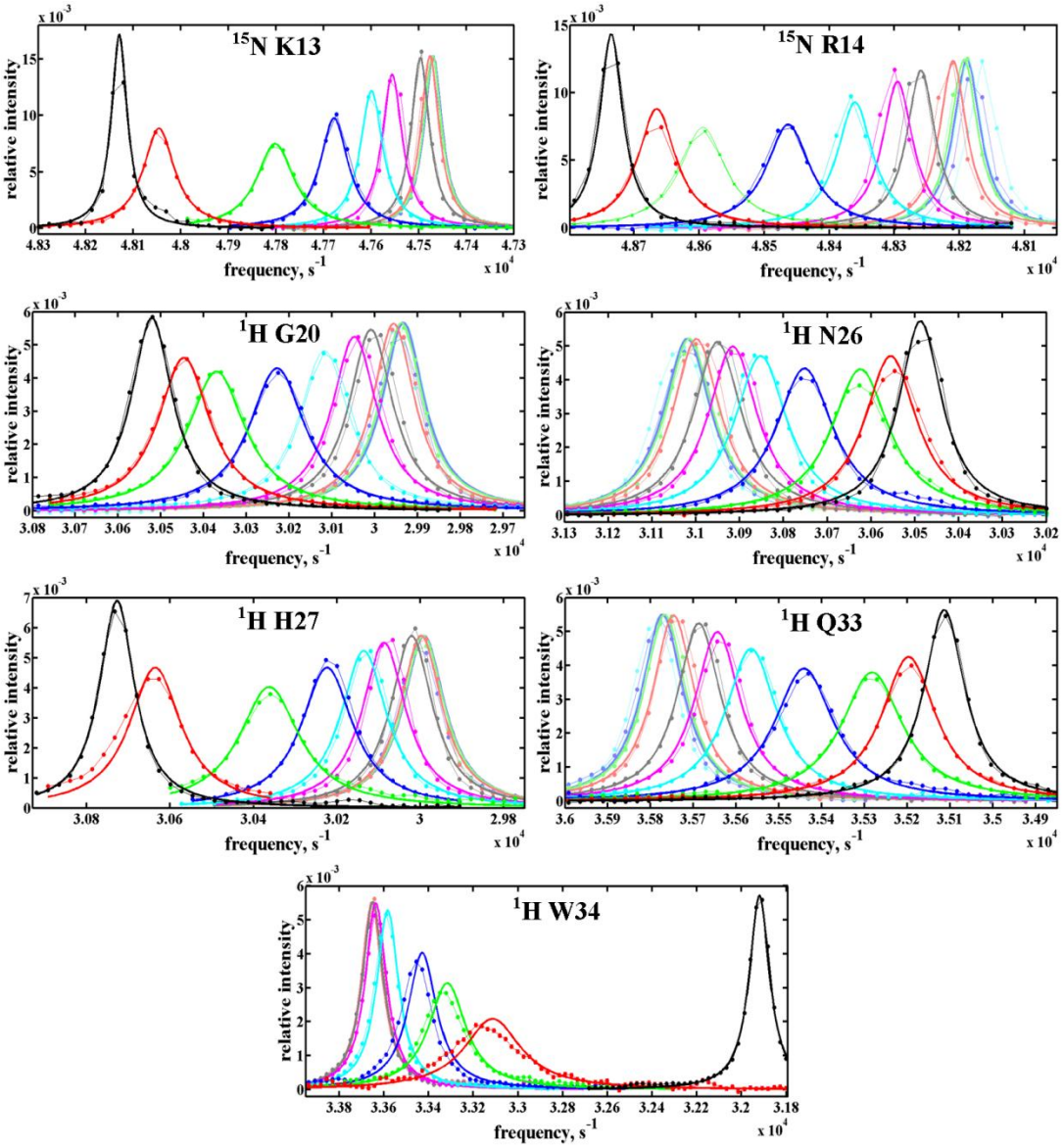


Figure 2.9: Simultaneous fitting of seven ^{15}N-WW domain datasets. Black peak represents free protein.

Interestingly, the initial titration experiments performed with ^{15}N-labeled WWold construct and unlabeled peptide, TAR1, displayed broader lineshapes for the free protein as compared to the saturated spectrum (Figure 2.11a). The WWold construct has 1-53 Pin1 residues, the last three being E51, P52 and A53. The last few residues in this construct are in a flexible tail (Figure 2.5b) and the E51-P52 peptide bond mimics the pS-P peptide bond. The peaks corresponding to the *cis* and the *trans* isomers are distinctly visible in the ^{15}N-^{1}H HSQC spectrum of this WW domain construct (Figure 2.11b). Hence, it was reasoned that the *trans* isomer of E51-P52 in the tail weakly interacts with the binding site of the WW domain. This results in the broadening of the core residues of the ^{15}N-WW domain due to fast exchange between the free and the bound states. When the phosphorylated peptide (TAR1) is titrated into this protein, it competes off the weak E51-P52 binding and under saturating condition there is only one bound state, i.e. ^{15}N-WW to TAR1, which has narrower lineshape than the free protein (Figure 2.11a). In order to confirm this hypothesis the last three residues were removed from the WW domain resulting in the construct WW50. The relative broadening of the free WW domain peaks as compared to the phosphopeptide bound peaks was eliminated in this construct (Figure 2.11c). This observation highlights the sensitivity of the NMR lineshapes even to weak interactions which can arise from intramolecular interactions, buffer components, etc. The above observations can be summarized in a model (Figure 2.11d). When the WWold is saturated with phosphopeptide the flexible tail is mostly free and the E51-P52 peptide bond only undergoes exchange between the two isomers. Hence, from the ratio of the peak heights, the equilibrium constant $K_{isom,EPA}$ was determined to be 7.9, which agrees well with reported value of Glu-Pro peptide bond (Reimer et al., 1998). In the free protein E51-P52 undergo exchange between the three states G, H and I. The state G gives rise to the *cis* peak in ^{15}N-^{1}H HSQC spectrum (Figure 2.11b). The states H

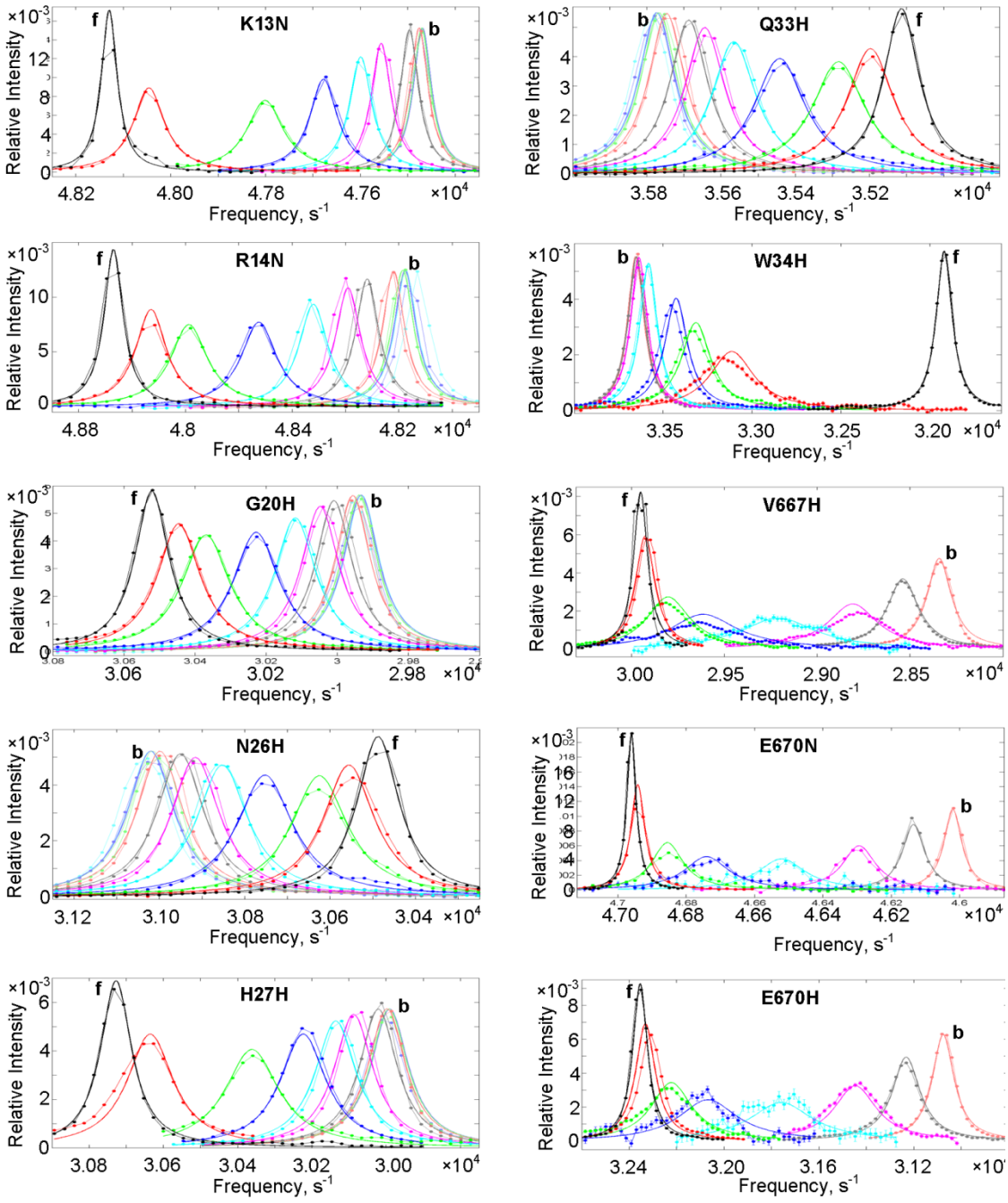


Figure 2.10: Simultaneous fitting of seven residues from the ^{15}N-WW domain and three residues from the ^{15}N-peptide.

and I are in fast exchange and give rise to the *trans* peak in the same spectrum. Hence, the following relation can be deduced between the ratio (R) of the *trans* and *cis* peaks, the equilibrium constants $K_{isom,EPA}$ and $K_{bind,EPA}$:

$$R = \frac{[H]+[I]}{[G]} = \frac{[H]}{[G]} + \frac{[I]}{[G]} = \frac{[H]}{[G]} + \frac{[I]}{[H]} \times \frac{[H]}{[G]}$$

$$= \frac{[H]}{[G]} \times \left(1 + \frac{[I]}{[H]}\right) = K_{isom,EPA} \times \left(1 + K_{bind,EPA}\right) \quad (2.2)$$

The ratio (R) was determined to be 8.95. Using equation (2.2), $K_{bind,EPA}$ was determined to be 0.136.

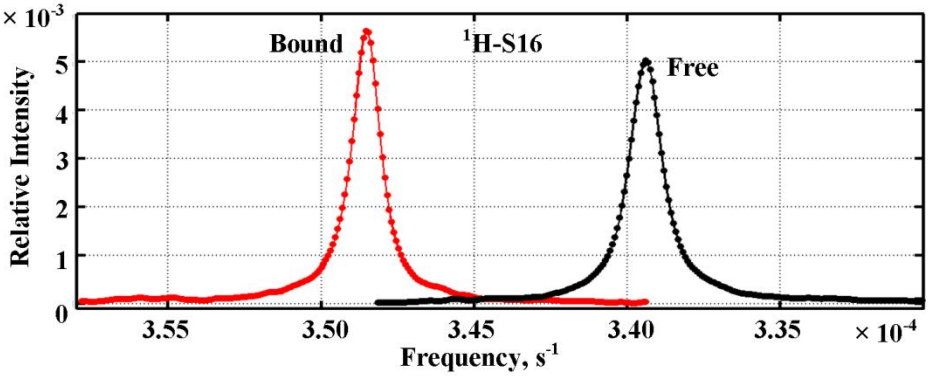
Discussion

Here we have used NMR lineshape analysis to characterize the kinetics of the interaction of the Pin1-WW domain with the cytoplasmic tail of APP, phosphorylated at T668. This binding interaction was investigated from both the protein and the peptide perspective by using ^{15}N-labeled protein and ^{15}N-labeled peptide, respectively. Use of the ^{15}N-labeled peptide revealed that the binding of the WW domain to the *cis* isomer is very weak and can be neglected. This suggests that the Pin1 reaction scheme can be simplified to a five-state system (Figure 2.12). The off-rates (k_{CB}) obtained from the two perspectives are in good agreement, validating our approach. To our knowledge, this is the first detailed kinetic study of the interaction of a WW domain with its substrate.

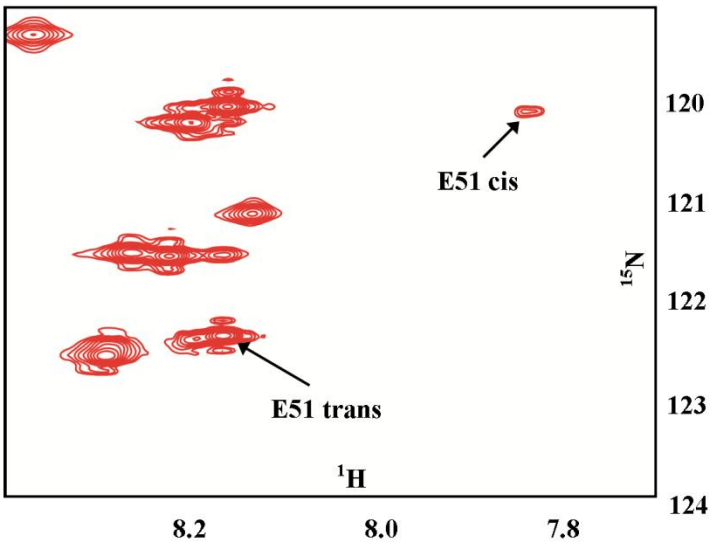
One of the most intriguing questions regarding the WW domain in Pin1 is its role in the functioning of the enzyme. It is well understood that enzymes utilize the enzyme-substrate binding energy for catalysis (Fersht, 1974). WW domain increases the binding affinity of Pin1 by several orders of magnitude and hence, some of its

Figure 2.11: Titration of WWold. a) Cross sections of the ^{1}H-dimension of S16 in free (black) WWold and saturated (red) with phosphopeptide. The peaks are normalized i.e. have unit area. Hence, the shorter peak has greater full-width at half-height. b) ^{15}N-^{1}H HSQC spectrum of the free WWold. The *cis* and the *trans* peaks of E51 are assigned. c) Cross sections of the ^{1}H-dimension of S16 in free (black) WW50 and saturated (red) with phosphopeptide. The black peak is slightly higher than the red peak. d) Model for WWold. The curvilinear line represents the C-terminal linker after the WW domain and the last last three residues are shown by single letter codes. The E51-P52 peptide bond in the flexible tail can exist as *cis* or *trans* isomer. The *trans* isomer binds the protein which is competed off by the phosphopeptide binding. Each state is denoted by a letter.

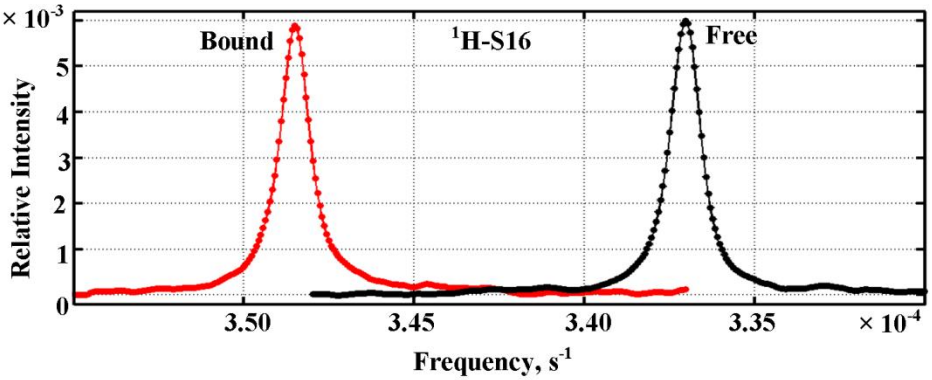
a)


× 10-3
Relative Intensity
Bound
1H-S16
Free
Frequency, s-1
× 10-4

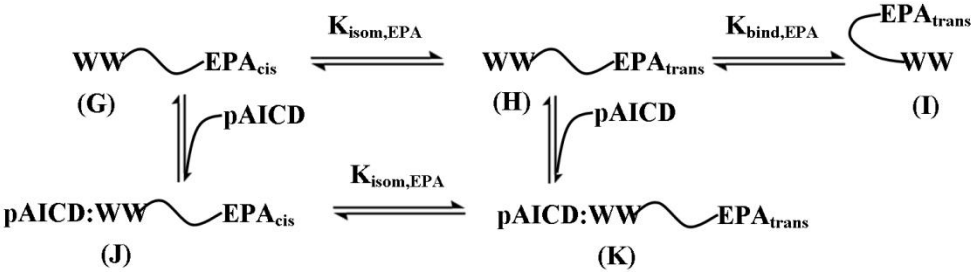
b)


E51 cis
E51 trans
15N
1H

c)


× 10-3
Relative Intensity
Bound
1H-S16
Free
Frequency, s-1
× 10-4

d)


Kisom,EPA
Kbind,EPA
EPAtrans
WW
EPAcis
(G)
pAICD
(H)
(I)
pAICD:WW
(J)
(K)

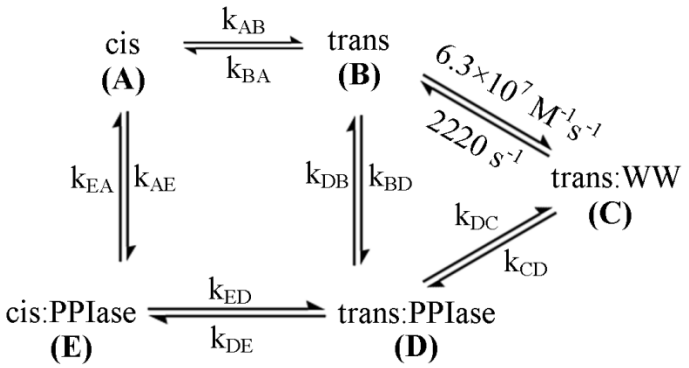


Figure 2.12: Revised reaction scheme of Pin1 and its substrate. The weak interaction of the cis peptide to WW domain has been eliminated from the reaction scheme and the rate constants determined for the binding of the trans peptide to the WW domain is shown.

binding energy should contribute to the overall catalytic efficiency of the enzyme. Another way to view this is by considering the internal equilibrium constant of an enzyme which should be close to 1 for a perfect enzyme (Albery and Knowles, 1976). The internal equilibrium constant is the equilibrium constant of the enzyme-bound species, i.e. the bound *cis* and bound *trans* in the case of Pin1. Since the WW domain differentially binds to these two isomers, its presence would change the internal equilibrium constant of Pin1 and it can be speculated that this change can bring it close to 1 as the presence of WW domain in Pin1 is evolutionarily favored. This would argue for the presence of a substrate transfer step (Figure 2.12) between the WW domain and the PPIase domain. Also it is interesting to note that the substrate off-rate (2200 s^{-1}) for the WW domain is in roughly the same time regime as the on-enzyme catalysis rate (1200 s^{-1}) of the isolated Pin1-PPIase domain (Labeikovsky et al., 2007) and these two steps are connected by the substrate transfer step (Figure 2.12). The kinetic synchronization of these two steps might be important in the overall catalytic activity of the enzyme and should be evident from the complete analysis of the reaction scheme of full length Pin1.

The NMR titration experiments and lineshape analysis of the titration data provide an excellent experimental tool to study kinetics of isomer-specific interactions. Also the excellent agreement between the ligand and the protein perspectives reporting on the same binding event provides us with confidence regarding the applicability of this method to the more complex reaction schemes of the PPIase domain and the full length Pin1. Additionally, our detection of a weak intramolecular interaction binding between the linker region (E51-P52) and the WW domain binding surface highlights the sensitivity of NMR lineshapes to highly transient interactions, and illustrates the importance of establishing experimental conditions in which the specific interaction(s) of interest dominate the observed changes in lineshape.

CHAPTER 3

THERMODYNAMICS OF THE INTERACTION OF PIN1-WW DOMAIN AND PHOSPHORYLATED APP CYTOPLASMIC TAIL *

Introduction

Isothermal titration calorimetry (ITC) was discovered almost 45 years ago (Christensen et al., 1966) and since has been the most favored technique to study thermodynamics of binding (Chaires, 2008). Almost all chemical reactions or physical interactions such as protein-ligand binding are accompanied with a change in heat. Since the calorimeter can monitor such change in heat, it can be viewed as a universal detector. Modern day calorimeters can detect heat effects as small as 0.1 μCal which allows the determination of binding affinities as large as 10^9 M^{-1} (Freyer and Lewis, 2008). Even though ITC was first developed as a technique to simultaneously determine the binding-energetics, i.e. free energy (ΔG), enthalpy (ΔH) and entropy (ΔS), from a single experiment, it has since been used to study kinetics, too (Todd and Gomez, 2001). Modern ITC instruments are capable of determining the Michaelis-Menten constant (K_M) and reaction rate (k_{cat}) in the ranges of 10^{-2}-10^3 μM and 0.05-500 s^{-1}, respectively (Freyer and Lewis, 2008).

We used ITC to independently determine the binding constant (K_D) of the pAICD derived peptide (pAPP659-682) to the isolated Pin1-WW domain and confirm the binding constant obtained from NMR lineshape analysis. The use of isothermal titration calorimetry (ITC) experiments provided a complete thermodynamic description of the interaction between pAICD and Pin1-WW, including changes in enthalpy and entropy of the system. Since the peptide exists in two isomeric states, *cis*

---

* Parts of this chapter are based on a manuscript submitted to *Nature Chemical Biology*.

and *trans*, and only the *trans* isomer binds significantly to the Pin1-WW domain (Figure 3.1), the existing binding models were not adequate to fit our ITC data. We derived a new set of equations and implemented the fitting in matlab. Monte-carlo simulations were used to estimate the errors in the fitting parameters.

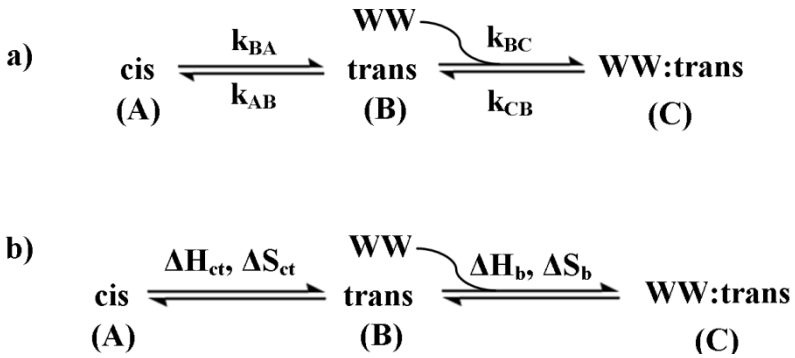


Figure 3.1: The reaction scheme for interaction of pAICD and isolated Pin1-WW domain. The kinetic (a) and thermodynamic (b) parameters for the isomerization and binding steps are shown.

We were also able to determine the uncatalyzed, slow *cis-trans* isomerization rates of pT668-P669 peptide bond in pAICD from analysis of the time-dependence of isolated ITC peaks. The determination of the slow isomerization rate, its speeding up in the presence of Pin1 (Pastorino et al., 2006) and the binding rate of an isomer-specific interaction between *trans* pAICD and Pin1-WW domain, led us to hypothesize on Pin1's influence on isomer-specific interactions. The two isomers have distinct structural features (Figure 3.2) and can potentially be recognized by isomer-specific binding partners in the cell. An example of such isomer-specific interaction is the *trans*-specific dephosphorylation by the phosphatase PP2A (Zhou et al., 2000). The interaction of the two isomers to their specific binding partners can signal different or even competing functions (Figure 3.2) in the cell such as the amyloidogenic and the non-amyloidogenic processing pathways of APP (Nicholson and Lu, 2007; Pastorino et al., 2006). Here we show that the catalytic activity of Pin1

can modulate the effective binding affinity of the putative isomer-specific binding partners and thus, can regulate the isomer-specific signaling events. In the absence of Pin1, the binding affinity of the isomer-specific interactions decreases in comparison to that in the presence of Pin1, when considered in a timescale of 1 to 10 seconds. Moreover, this effect is greater in the lesser populated isomer. In summary, this provided us with a detailed description of the thermodynamics of Pin1-WW domain and pAICD interaction and the role of Pin1 as a modulator of the effective binding affinities of the isomer-specific binding partners.

Materials and Methods

Protein expression and purification

Protein expression and purification for WW domain was carried out as described in the previous chapter. Pin1 gene is encoded in a PGex expression vector with GST-tag and thrombin cleavage site between the tag and the protein and was generously provided by Dr. Kun Ping Lu, Harvard Medical School. The protein was expressed in BL21 DE3 *E. coli* cells which were grown in 1L LB media at 37°C in the presence of 100µg/ml amphicillin. The protein expression was induced with 1mM IPTG at $O.D_{600}$ of 0.8 and harvested after 4 hours by centrifugation at 4920 g for 10 min at 4°C.

Pin1 purification was carried out at 4°C. The cells were lysed in 20ml lysis buffer (25mM HEPES, 150mM NaCl, pH = 7.5) containing 8mg of EDTA, 1mM DTT, 20µM of PMSF, 20µl protease inhibitor cocktail and 20µg Lysozyme. The lysed cells were sonicated on ice with 50% duty cycle for 5 min and the cell debris was removed by centrifugation for 40 min at 23,500 g. The supernatant was filtered using 0.80µm filter and passed over a packed column of 2ml glutathione Sepharose 4B(Amersham GE), pre-equilibrated with wash buffer (25mM HEPES, 150mM NaCl,

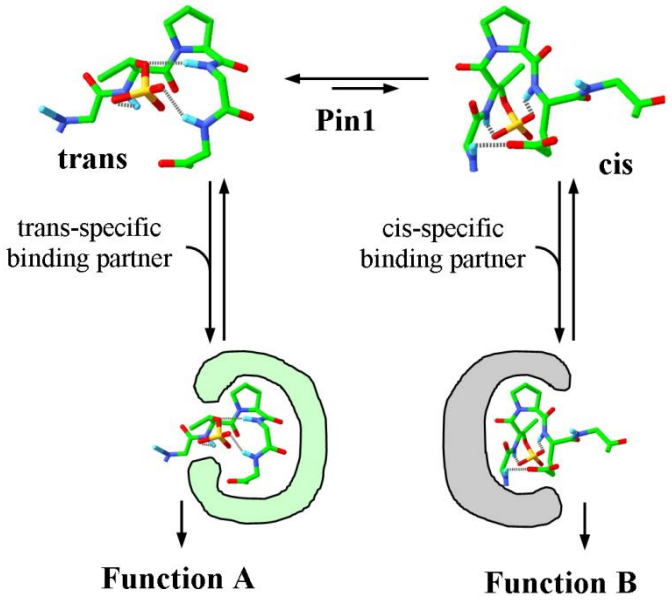


Figure 3.2: Phosphorylation of T668 results in the emergence of *trans* and *cis* isomers of the pT668-P669 peptide bond which have distinct structural features. Each isomer can be recognized by isomer-specific binding partners resulting in distinct cellular functions.

pH = 7.5). The column was washed with 10 bed volumes of wash buffer and 5 bed volumes of cleavage buffer (20mM Tris-HCl, 150mM NaCl**,** 2.5mM $CaCl_2$, pH=8.4) leaving 1 bed volume of the cleavage buffer in the column. The protease was prepared by adding 5µl thrombin to 95µl dilution buffer and all 100µl was added to the column. The column was rocked slowly at room temperature for 16 hrs. Cleaved Pin1 was collected by flowing out the buffer from the column. Two more fractions were collected by washing the column with 2 bed volumes of wash buffer each. The 3 fractions were checked for protein by UV absorption at 280 nm and also by Bradford assay. To remove thrombin 60µl streptavidin-agarose beads were added to the pooled fractions containing Pin1, rocked for 30 min and removed by spin columns. The purified protein was dialyzed into the final buffer (10mM HEPES, 10mM NaCl, 1mM DTT, 5mM NaN_3, pH = 6.9) overnight at 4°C.

Peptide Synthesis

The purified phosphopeptide (pAPP659-682) was purchased from the Tufts University Core Facility, Boston. It was weighed out, as required, dissolved in 200µl de-ionized H_2O, adjusted to pH ~7 using NaOH and lyophilized. The lyophilized peptide was dissolved in the final buffer. The pH of the sample adjusts to the pH of the buffer. The concentration is measured using tyrosine extinction coefficient of 1400 $M^{-1}cm^{-1}$ (Creighton, 1993).

Isothermal Titration Calorimetry (ITC)

The titration experiments were performed in a Calorimetry Sciences Corp. (Lindon, UT) Nano-ITC Series III instrument at 25°C. The protein and the peptide samples were degassed for 10 min prior to each experiment. The reaction cell was filled with 1.4ml of protein (WW-domain or Pin1) and 100µl or 250µl syringe was used to titrate in the peptide. For the WW domain:pAICD binding experiment, 170 µM of protein and 4.3 mM of peptide was used. The peptide was injected using the 250µl syringe. A total of 31 injections were made of 8µl each with an interval of 8 min between injections. The sample was stirred at 200rpm for rapid mixing. The binding experiment was also done in the presence of catalytic amount of Pin1. For this experiment 160µM of WW domain, 4.2µM of Pin1 and 4.3 mM of peptide was used. A total of 41 injections were made with 6µl each with an interval of 5 min between injections. To get the heat of dilution of the peptide, 4.3 mM peptide was injected into the sample buffer and the average heat of dilution was subtracted from the titration data. For the kinetics experiment 250 µl of 0.65 mM TAR2 was titrated into 70 µM of WW domain. Total 31 injections were made with 8 µl each. The injections were made at an interval of 8 min to ensure the complete return of the signal to the baseline.

Derivation of expressions for the binding isotherm

The following derivation is adapted from derivation of one-to-one binding by Lewis and Murphy (Lewis, 2005). The heat evolved from each injection (q_i) has two components: heat resulting from *trans* isomer binding to WW ($q_{b,i}$) and heat resulting from the conversion of *cis* isomer to *trans* isomer ($q_{ct,i}$). The heat of binding is given by the following expression

$$q_{b,i} = \langle \Delta H \rangle_i [M]_{tot,i} V_{cell} - \langle \Delta H \rangle_{i-1} [M]_{tot,i-1} \left(V_{cell} - V_{inj} \right)$$

where $\langle \Delta H \rangle_i = \Delta H_b \frac{K_A [T]_i}{1 + K_A [T]_i}$,

is the excess enthalpy per mole of macromolecule, K_A is the binding constant, ΔH_b is the binding enthalpy per mole and $[T]_i$ is the concentration of the free *trans* isomer. During the course of the experiment, the volume of solution that is ejected from the reaction cell equals the volume of ligand injected, which dilutes the concentration of the macromolecule as follows

$[M]_{tot,i} = n[M]_{tot,0} D_i$ where the dilution factor is $D_i = \left(1 - \frac{V_{inj}}{V_{cell}}\right)^i$

The concentration of the free *trans* isomer $[T]_i$ after each injection is calculated by solving the following set of equations

$$[L]_{tot,0} = [T]_i + [C]_i + [MT]_i,\ K_{isom} = \frac{[T]_i}{[C]_i},\ K_A = \frac{[MT]_i}{[M]_i [T]_i},\ [M]_{tot,0} = [M]_i + [MT]_i$$

The heat of isomerization is given by the following expression

$q_{ct,i} = \Delta H_{ct} \left(\frac{D[T]_{i-1}}{K_{isom}} + \frac{(1-D)L_{tot,0}}{1 + K_{isom}} - \frac{[T]_i}{K_{isom}} \right) V_{cell}$ where ΔH_{ct} is the enthalpy of *cis* to *trans* conversion. Thus, the total heat for injection 'i' is given by

$$q_i = q_{b,i} + q_{ct,i} \qquad (3.1)$$

In the above equations $[M]_{tot,i}$ is the total concentration of the macromolecule after the i^{th} injection, $L_{tot,0}$ is the ligand concentration in the syringe, V_{cell} is the reaction cell volume, V_{inj} is the injection volume of the peptide, 'n' is the number of binding sites in the protein, and K_{isom} is the free *trans* to *cis* equilibrium constant. $[T]_i$ and $[C]_i$ are the concentrations of the free *trans* and *cis* isomers respectively.

The fitting was performed in matlab. The non-linear least squares fitting routine 'lsqnonlin', available in matlab, was used to iteratively fit the ITC data to equation (3.1) and K_A, ΔH_b, ΔH_{ct} and n were the fitting parameters. The errors in the fitted parameters were estimated using Monte-Carlo simulation. Normally distributed random noise was added to the best-fit solution of the binding isotherm to generate 1000 synthetic datasets which were then fit to equation (3.1) to get 1000 sets of fitted parameters. The standard deviation of each parameter was reported as the error for that parameter. Hereafter, "code" refers to matlab programs written in-house. The iterative fitting routine is implemented in the code 'ITCfit'. The code 'myT' implements the analytical expression for the free *trans* isomer after each injection, 'myQ_calc' calculates the heat evolved from each injection using equation (3.1) and 'myQ' calculates the error between the experimental and calculated heats at each injection. The Monte-Carlo simulation is implemented in the code 'ShakeSimDataErrorEstimateMC'.

Derivation of approximate expressions for time-dependence

After each injection, the *trans* isomer binds rapidly to the free WW domain in the reaction cell which is followed by the slow conversion of free *cis* to free *trans* and its subsequent binding (Figure 3.1a). The time-dependence of each species i.e. free *cis* [C], free *trans* [T] and the bound WW domain [WW:T] can be expressed using

differential equations (3.2) which constitute a set of coupled non-linear ordinary differential equations (ODEs).

$$\frac{d[C]}{dt} = -k_{AB}[C] + k_{BA}[T] \quad (3.2a)$$

$$\frac{d[T]}{dt} = k_{AB}[C] - (k_{BA} + ([WW_0] - [WW:T])k_{BC})[T] + k_{CB}[WW:T] \quad (3.2b)$$

$$\frac{d[WW:T]}{dt} = ([WW_0] - [WW:T])k_{BC}[T] - k_{CB}[WW:T] \quad (3.2c)$$

where k_{AB} and k_{BA} are the rates for *cis* to *trans* and *trans* to *cis* conversion respectively, k_{BC} and k_{CB} are the on and off rates respectively and $[WW_0]$ is the concentration of the total WW domain. The set of coupled ordinary differential equations (3.2) were numerically solved and plotted (Figure 3.3a). For the numerical solution, the parameter values used are $k_{AB} = 0.01\ s^{-1}$, $k_{BA} = 0.001\ s^{-1}$, $k_{BC} = 2220\ s^{-1}$ and $k_{CB} = 6.3\times10^{7}\ s^{-1}$. From the plots it is clear that the binding of the *trans* isomer to WW domain is complete within the first few milliseconds. Then the system equilibrates according to the slow *cis* to *trans* conversion rate (Figure 3.3b). Since the time-dependence of the *cis* isomer, equation (3.2a), is only dependent on the rates k_{AB} and k_{BA} and not on the fast binding kinetics (k_{BC} and k_{CB}), an approximate expression for its time dependence can be obtained by neglecting the binding step and solving the resulting equations (3.3).

$$\frac{d[C]}{dt} = -k_{AB}[C] + k_{BA}[T] \quad (3.3a)$$

$$\frac{d[T]}{dt} = k_{AB}[C] - k_{BA}[T] \quad (3.3b)$$

$$[C(t)] = [C_{eq}]\frac{k_{AB}}{k_{BA}} - \left([C_{eq}] - \frac{L_0 k_{BA}}{k_{BA}+k_{AB}}\right)e^{-(k_{BA}+k_{AB})t} \quad (3.4)$$

The resulting time-dependence of the *cis* isomer is given by equation (3.4) which represents an exponential decay, where $[C_{eq}]$ is its final equilibrium concentration. In comparison to the time-course obtained from the numerically solved ODEs, the

analytical equation (3.4) represents the time-dependence of the *cis* isomer very well (Figure 3.3 c and d). After the initial fast binding, there is no appreciable *cis* to *trans* conversion and the free *trans* and the bound *trans* reach a pseudo-equilibrium state with concentrations [$T_{eq}{}^0$] and [$WW{:}T_{eq}{}^0$], respectively (Figure 3.3a). These pseudo-equilibrium concentrations were solved from the equations (3.5). Equation (3.5a) represents the steady-state condition, equation (3.5b) is the total *trans* concentration which excludes the *cis* isomer of the peptide that was initially present and equation (3.5c) is the total WW domain concentration.

$$\frac{d[T_{eq}^0]}{dt} = -k_{BC}[T_{eq}^0][WW_{eq}^0] + k_{CB}[WW{:}\,T_{eq}^0] = 0 \qquad (3.5a)$$

$$[T_{eq}^0] + [WW{:}\,T_{eq}^0] = \frac{[L_0]k_{AB}}{k_{BA}+k_{AB}} \qquad (3.5b)$$

$$[WW_{eq}^0] + [WW{:}\,T_{eq}^0] = [WW_0] \qquad (3.5c)$$

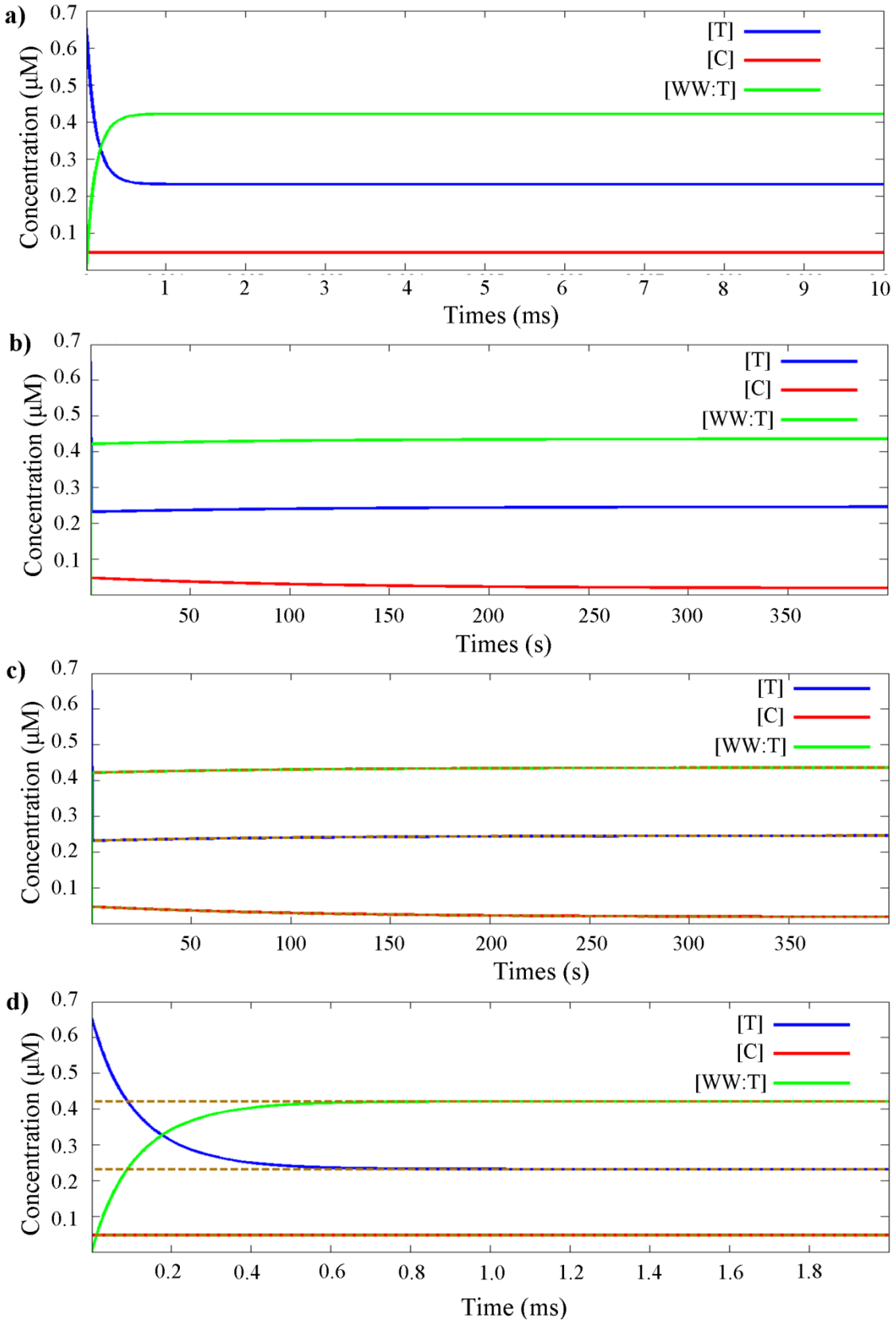
where [L_0] and [WW_0] are the total concentrations of the peptide and the WW domain, respectively. Similar to the free *cis* isomer, the slow relaxation of the free *trans* and the bound WW domain from [$T_{eq}{}^0$] and [$WW{:}T_{eq}{}^0$] to [T_{eq}] and [$WW{:}T_{eq}$], respectively, should also be exponential decaying functions. Hence, their time-dependence can be modeled by equations (3.6a) and (3.6b).

$$[T(t)] = [T_{eq}] - \left([T_{eq}] - [T_{eq}^0]\right)e^{-(k_{BA}+k_{AB})t} \qquad (3.6a)$$

$$[WW{:}\,T(t)] = [WW{:}\,T_{eq}] - \left([WW{:}\,T_{eq}] - [WW{:}\,T_{eq}^0]\right)e^{-(k_{BA}+k_{AB})t} \qquad (3.6b)$$

The equations (3.4) and (3.6) model the time-dependence of the free *cis*, free *trans* and bound WW domain concentrations very well in the post-fast binding period (Figure 3.3 c and d). The expression for the time-dependence of the *cis* isomer, equation (3.4), was obtained by neglecting the kinetics of the fast binding step, only treating it as a step that removes the *trans* isomer. The resulting expression is in excellent agreement

Figure 3.3: Approximate analytical expressions accurately predict the time-dependence for the free *cis*, free *trans* and bound *trans* species. a) Numerical solution of ODEs in equation 3.3 solving for the free *cis* (red), free *trans* (blue) and bound *trans* (green) is plotted. The plot of the first 10 milliseconds shows that the binding step is very fast and is complete in this time-scale. b) The system reaches equilibrium slowly due to the slow isomerization step. Beyond the first few milliseconds, the time-dependence of the three species behaves similarly. c) The approximate time-dependence for the three species, given by the equations (3.4) and (3.5), are plotted as broken brown lines. The approximate equations match the actual solutions very well. d) The approximate solution for the *cis* isomer matches its actual solution (solid red) even in the first few milliseconds. The approximate solutions for other two species match their actual solutions once the fast binding is complete (i.e. after ~0.6 ms).


a)
Concentration (μM)
Times (ms)
[T]
[C]
[WW:T]
b)
Concentration (μM)
Times (s)
[T]
[C]
[WW:T]
c)
Concentration (μM)
Times (s)
[T]
[C]
[WW:T]
d)
Concentration (μM)
Time (ms)
[T]
[C]
[WW:T]

to the actual time-dependence of the *cis* isomer of the entire time-period (Figure 3.3 c and d). After each injection, the heat (q) released during the slow equilibration period, which is considered to be after several seconds of the injection when the fast binding is complete, can be modeled by equation (3.7) where the first part is the heat of isomerization and the second part is the heat of binding.

$$\frac{dq}{dt} = \Delta H_{ct} V_{cell} \left(-\frac{d[C]}{dt}\right) + \Delta H_b V_{cell} \left(\frac{d[WW{:}T]}{dt}\right) \quad (3.7)$$

In the above equation, -d[C]/dt is the rate of depletion of the free *cis* isomer, d[WW:T]/dt is the rate of formation of the bound WW domain, V_{cell} is the volume of the reaction cell and ΔH_{ct} and ΔH_b are the enthalpies of *cis* to *trans* conversion and *trans* binding, respectively. The corresponding rates are obtained by differentiating equations (3.4) and (3.6b) and are given as

$$-\frac{d[C]}{dt} = \left(\frac{L_0 k_{BA}}{k_{BA}+k_{AB}} - [C_{eq}]\right)(k_{BA} + k_{AB})e^{-(k_{BA}+k_{AB})t} \quad (3.8a)$$

$$-\frac{d[WW{:}T]}{dt} = \left([WWT_{eq}] - [WWT^0_{eq}]\right)(k_{BA} + k_{AB})e^{-(k_{BA}+k_{AB})t} \quad (3.8b)$$

Since the time-dependence is only in the exponent, equation (3.8) can be simplified as follows,

$$\frac{dq}{dt} = q'e^{-tk_{ex}} \quad (3.9)$$

where $k_{ex} = k_{AB} + k_{BA}$. From the ratio of the volumes of peaks corresponding to the *cis* and *trans* isomers in the ^{15}N-^{1}H HSQC spectrum of the ^{15}N-labeled peptide, we obtained $K_{isom} = V_{trans}/V_{cis} = 14$.

$$\frac{k_{AB}}{k_{BA}} = K_{isom} \quad (3.10)$$

The rates of isomerization, k_{AB} and k_{BA}, were obtained by fitting the ITC data to equations (3.9) and (3.10).

The kinetic-analysis of the ITC data was performed on the signals from the 4^{th} injection to 10^{th} injection. To extract out the slow decaying portion of the isothermal signal, the starting 2.5 minutes of the data were ignored and the rest was fit to equation (3.9). Binding isotherms for interactions that do not involve a slow step are typically back to near-baseline after 2.5 min, indicating the thermal time-constant of the instrument is shorter than this. The standard deviations in k_{AB} and k_{BA} obtained from the seven peaks were reported as errors of those parameters.

The ordinary differential equations were numerically solved in matlab using the solver 'ode15s'. Chemical reactions often result in stiff ode, i.e. the step size taken during the numerical solution should be extremely small and the solver 'ode15s' in matlab is well suited to handle such problems. The *cis* isomer-specific binding is simulated using the code 'IsoBind_CisSpecific.m' and the differential equations are implemented in the code 'IsoBindfunc_CisSpecific.m'. Both codes are generated in-house.

Results

Thermodynamics of binding and isomerization

Isothermal titration calorimetry (ITC) was performed to determine the thermodynamic parameters such as the binding constant (K_D), enthalpy (ΔH) and entropy (ΔS) of binding of (pAPP659-682) to the WW domain. Since the binding reaction scheme (Figure 3.1b) involves isomerization and binding, the enthalpic contributions from both steps need to be considered. Upon injection of the peptide into the isolated WW domain exothermic peaks were observed (Figure 3.4a) which returned to the baseline very slowly, displaying a slow exponential decay (Figure 3.4b). The initial part of the peak corresponds to the fast binding of the free *trans* isomer to the WW domain and results in the depletion of free *trans*. The slow decay

results from the conversion of free *cis* to free *trans* isomer of the peptide and subsequent fast binding of the *trans* isomer to the WW domain. Based on this observation, equation (3.1) was derived to represent the binding isotherm. The binding isotherm data from the ITC experiment was fit to equation (3.1) (Figure 3.4c) to extract the binding constant (K_D), binding enthalpy (ΔH_b) and entropy (ΔS_b) and also isomerization enthalpy (ΔH_{ct}) and entropy (ΔS_{ct}).

The fitting yielded $K_D = 27 \pm 2$ μM, $\Delta H_b = -41 \pm 2$kJ/mol, $\Delta H_{ct} = -43 \pm 8$ kJ/mol and $n = 0.95 \pm 0.03$, where errors are standard deviations obtained from Monte-Carlo simulations. The corresponding entropies for binding and isomerization are $T\Delta S_b = -14.9$ kJ/mol and $T\Delta S_{ct} = -36.5$ kJ/mol respectively. The binding of the peptide to the WW domain is enthalpy driven, which is expected from the considerable electrostatic interactions in the interface. The *cis* to *trans* conversion is also enthalpy driven.

In order to confirm that the slow decay of the ITC signal (Figure 3.4b) is indeed due to the slow *cis* to *trans* isomerization, we did the same binding experiment in the presence of a catalytic amount of Pin1. Upon titration, similar exothermic peaks were observed (Figure 3.5a). The isolated peaks of the ITC binding experiments in the presence and absence of catalytic Pin1 were compared (Figure 3.5b and c). The ITC peak in the presence of catalytic Pin1 is almost four fold larger but still returns to the baseline much quicker. This convincingly demonstrates that the slow decay of the ITC signal is indeed due to the slow *cis* to *trans* isomerization.

Kinetics of isomerization

Apart from thermodynamic information, the ITC data also contains kinetic information (Fanghanel et al., 2006). The exponentially decaying portion of the peaks 4 to 10, in the raw ITC data, were extracted out and fit to equation (3.9) (Figure 3.6).

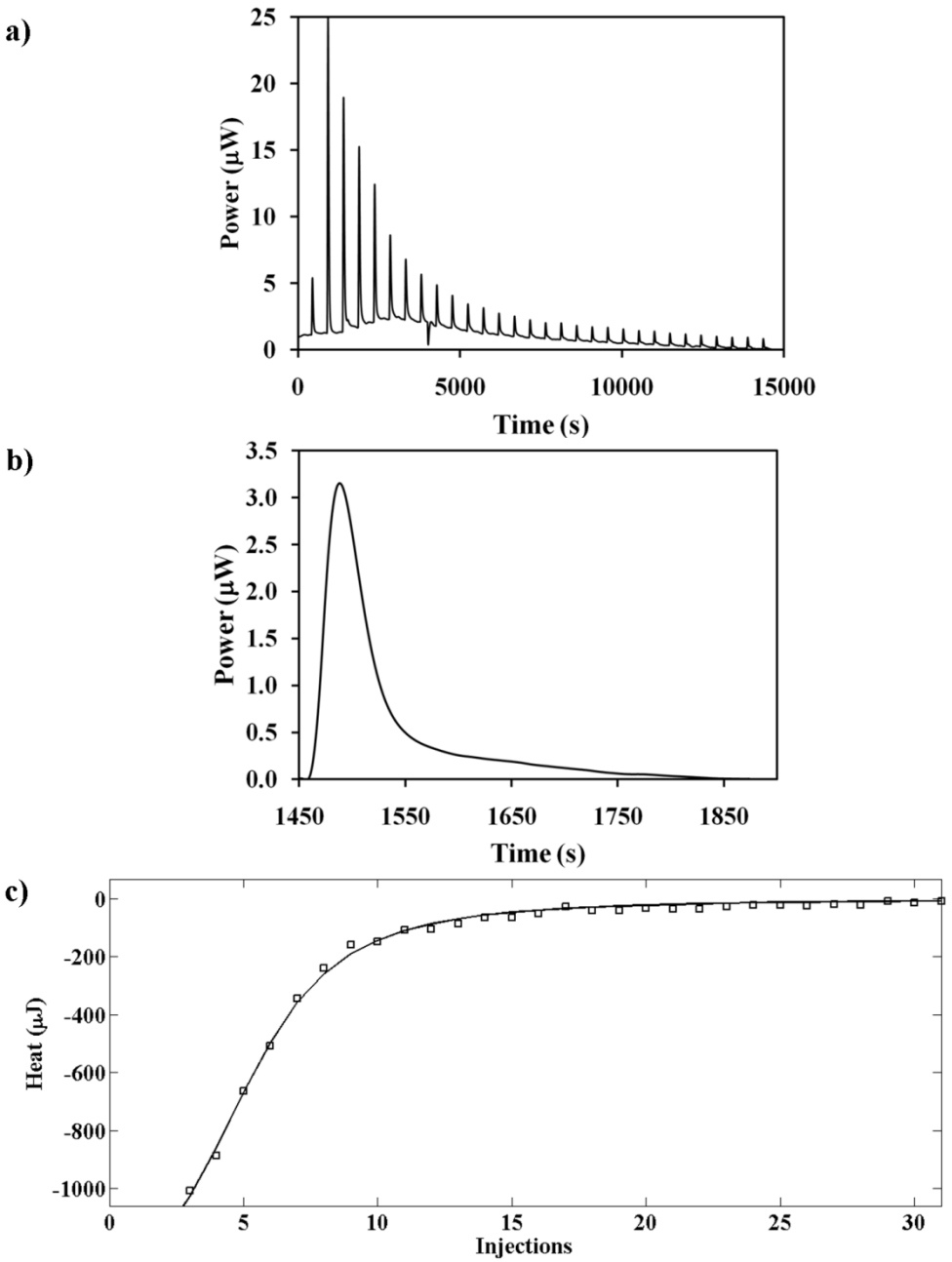


Figure 3.4: ITC experiment of pAICD and isolated WW binding a) Raw data of the binding isotherm of the interaction of pAICD and isolated WW domain. b) Peaks in this isotherm return slowly to the baseline due to the heat released from the slow isomerization of free *cis* to free *trans* isomer. c) Fitting of the heat evolved from each injection to equation (3.1) using in-house matlab codes. The enthalpy and entropy of both binding and isomerization were extracted from this fitting.

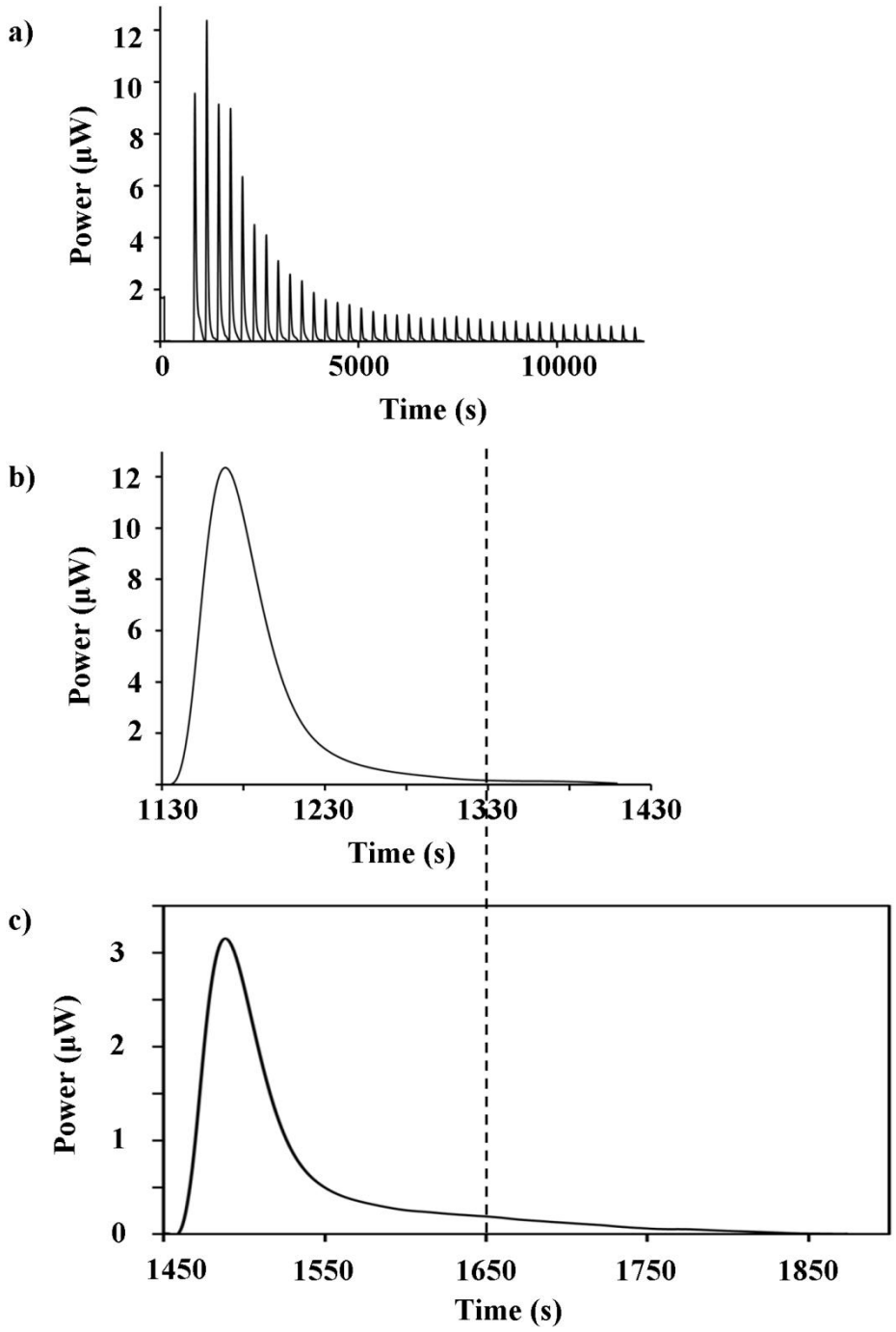


Figure 3.5: Interaction of pAICD and isolated WW domain in the presence of a catalytic amount of Pin1 a) Raw data for the binding isotherm from the ITC experiment performed to study the interaction of pAICD and WW domain in the presence of a catalytic amount of Pin1. b) Typical shape of an isolated peak in the above isotherm. The largest peak has been selected for display here. c) Isolated peak from an experiment without any Pin1. Comparison of b) and c) shows the faster return of the signal to baseline in the presence of catalytic amount of Pin1.

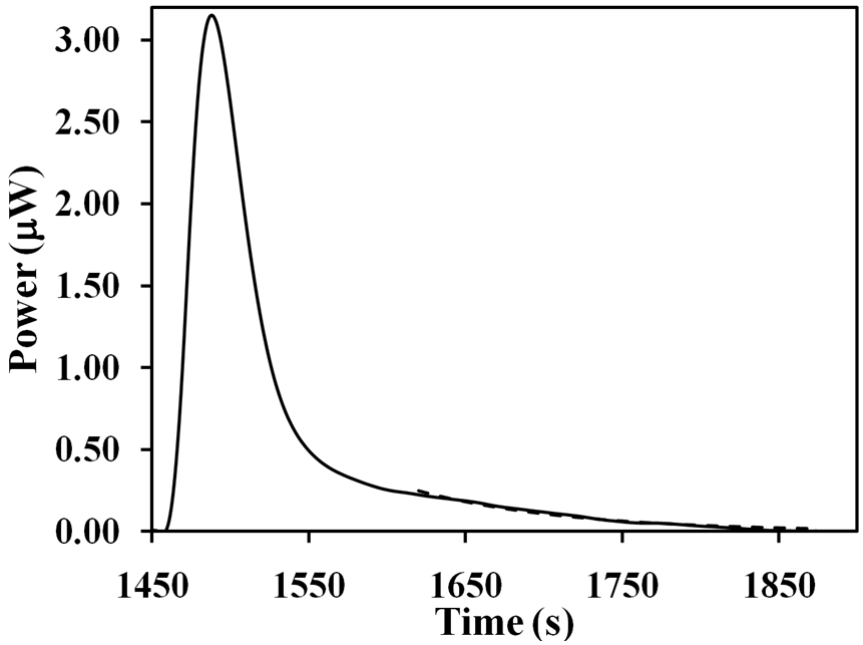


Figure 3.6: Fitting (broken line) of the slow conversion of *cis* to *trans* isomer to equation (3.9).

From the fitting we obtained $k_{AB} = 0.0089 \pm 0.0013 s^{-1}$ (mean ± standard deviation) and $k_{BA} = 0.00064 \pm 0.0001 s^{-1}$ which are in the same range as reported for other pSer/Thr-Pro peptides (Schutkowski et al., 1998). Thus, from the ITC and NMR experiments, the complete kinetics of the isomerization and binding reactions of the pAICD:WW interaction is now determined.

Pin1 as a connector and flux modulator

The kinetics from the ITC and NMR data clearly show the difference in the timescales of the uncatalyzed *cis-trans* isomerization ($k_{ex} = 0.00954\ s^{-1}$) and the isomer-specific binding event ($k_{off} = 2220\ s^{-1}$). A catalytic amount of Pin1 increases the rate of isomerization (Pastorino et al., 2006) to $k_{ex}^{cat} = 18\ s^{-1}$, a 2000 fold acceleration. In order to illustrate the effect of these different timescales on a protein interaction network, the smallest network with isomer-specific interactions is

considered (Figure 3.7a). In this network, *cis* and *trans* isomers of a phosphorylated Pin1 substrate have isomer-specific binding partners P and Q, respectively. In reality, there might be multiple isomer-specific binding partners with further downstream interactions but for illustrative purposes we will consider the simplest case of one binding partner for each isomer. We predict that the effect of Pin1 on coupling *cis* and *trans*-specific networks will be more pronounced if *cis* (less-populated) isomer has higher affinity or more abundant binding partners as shown here. For illustrative purposes, we assume the concentrations of the substrate (*cis* + *trans*) and the binding partners to be 1 μM each (reasonable cellular concentrations), the on-rates for both interactions to be 2×10^8 $M^{-1}s^{-1}$ (diffusion limited), the dissociation constants for the *cis* and *trans*-specific binding partners to be 10 nM and 100 nM (dissociation constants < cellular concentrations) and the corresponding off-rates to be 2 s^{-1} and 20 s^{-1} for *cis* and *trans* isomers, respectively. Initially, the free isomers are at equilibrium which is disturbed by the appearance of the binding partners, P and Q. The time-dependence of the concentrations of the four species i.e. free *cis* (pC), free *trans* (pT), bound *cis* (pC:P) and bound *trans* (pT:Q) are given by equations (3.11):

$$\frac{d[pC]}{dt} = -(k_{ct} + ([P_0] - [pC{:}P])k^c_{on})[pC] + k_{tc}[pT] + k^c_{off}[pC{:}P] \qquad (3.11a)$$

$$\frac{d[pT]}{dt} = k_{ct}[pC] - (k_{tc} + ([Q_0] - [pT{:}Q])k^t_{on})[pT] + k^t_{off}[pT{:}Q] \qquad (3.11b)$$

$$\frac{d[pC{:}P]}{dt} = ([P_0] - [pC{:}P])k^c_{on}[pC] - k^c_{off}[pC{:}P] \qquad (3.11c)$$

$$\frac{d[pT{:}Q]}{dt} = ([Q_0] - [pT{:}Q])k^t_{on}[pT] - k^t_{off}[pT{:}Q] \qquad (3.11d)$$

where k_{ct} and k_{tc} are the *cis-trans* inter-conversion rate constants, P_0 and Q_0 are the total concentrations of the binding partners and k^i_{on} and k^i_{off} are the on and off rate constants for the i^{th} isomer (i = *cis* or *trans*) (Figure 3.7a). The concentrations of the

four species are obtained from the numerical solution of equations (3.11). In the absence of Pin1, $k_{ct} = 0.0089\ s^{-1}$, $k_{tc} = 0.00064\ s^{-1}$ and the time-dependence of the concentrations are given by the solid lines while, in the presence of Pin1, $k_{ct} = 16.8\ s^{-1}$, $k_{tc} = 1.2\ s^{-1}$ and the time-dependence of the concentrations are given by the broken lines Figure 3.7b, c). The fast binding of the isomers to the respective binding partners is complete within one second irrespective of whether Pin1 is present Figure 3.7b). The system then re-equilibrates according to the isomerization rates (Figure 3.7c). The fast and slow phases in the simulation clearly demonstrate the significantly different timescales of these two processes i.e. binding and isomerization. The catalytic enhancement by Pin1 has significant impact on the equilibration of this smallest network. In the timescale of 10 seconds, the slow phase reaches equilibrium only in the presence of Pin1 as evident from the time-dependence of bound *cis* (pC:P, magenta) and bound *trans* (pT:Q, cyan). Thus, in the absence of Pin1, the isomer-specific interactions are insulated from each other (dotted boxes Figure 3.7a) in this timescale, whereas they are connected and form a single network module in the presence of Pin1 (solid box, Figure 3.7a). Pin1 activity also modulates the flux through the two isomer-specific network modules when the two bound species (pT:Q and pC:P) are the entry points into these distinct modules. After 10 seconds, the population of pT:Q is 10.3 fold that of pC:P in the absence of Pin1 whereas they are almost equal in the presence of Pin1 (Figure 3.7c). Hence, the effect of Pin1 activity on the protein interaction networks is more pronounced when the less populated isomer (*cis*) has higher affinity or more abundant interaction partners to compensate for the lower isomer population. It is important to note that the time-scale on the order of a few seconds is relevant, since cellular signaling activities, such as active or passive cellular diffusion, kinase or phosphatase reactions, or protein conformational changes, occur within this timescale (Papin et al., 2005).

From the above analysis, it is evident that upon appearance of binding partners P and Q there is no effective inter-conversion of the isomers in the absence of Pin1 in the short timescale, whereas re-equilibration to the new conditions is effectively complete in the presence of Pin1. Since the amount of bound species is different in the presence and absence of Pin1, it is informative to express the binding reactions in terms of effective dissociation constants ($K_{D,eff}$) for these two cases, where we assume that there is no inter-conversion of the isomers in the absence and complete inter-conversion in the presence of Pin1 in the timescale of few seconds. The effective dissociation constants for the *cis* and the *trans* isomer-specific binding partners in the absence of Pin1 is given by

$$K_{D,eff}^{cis} = K_D^{cis}(1 + K_{isom}) + K_{isom}[P] \quad (3.12a)$$

$$K_{D,eff}^{trans} = K_D^{trans}\left(1 + \frac{1}{K_{isom}}\right) + \frac{[Q]}{K_{isom}} \quad (3.12b)$$

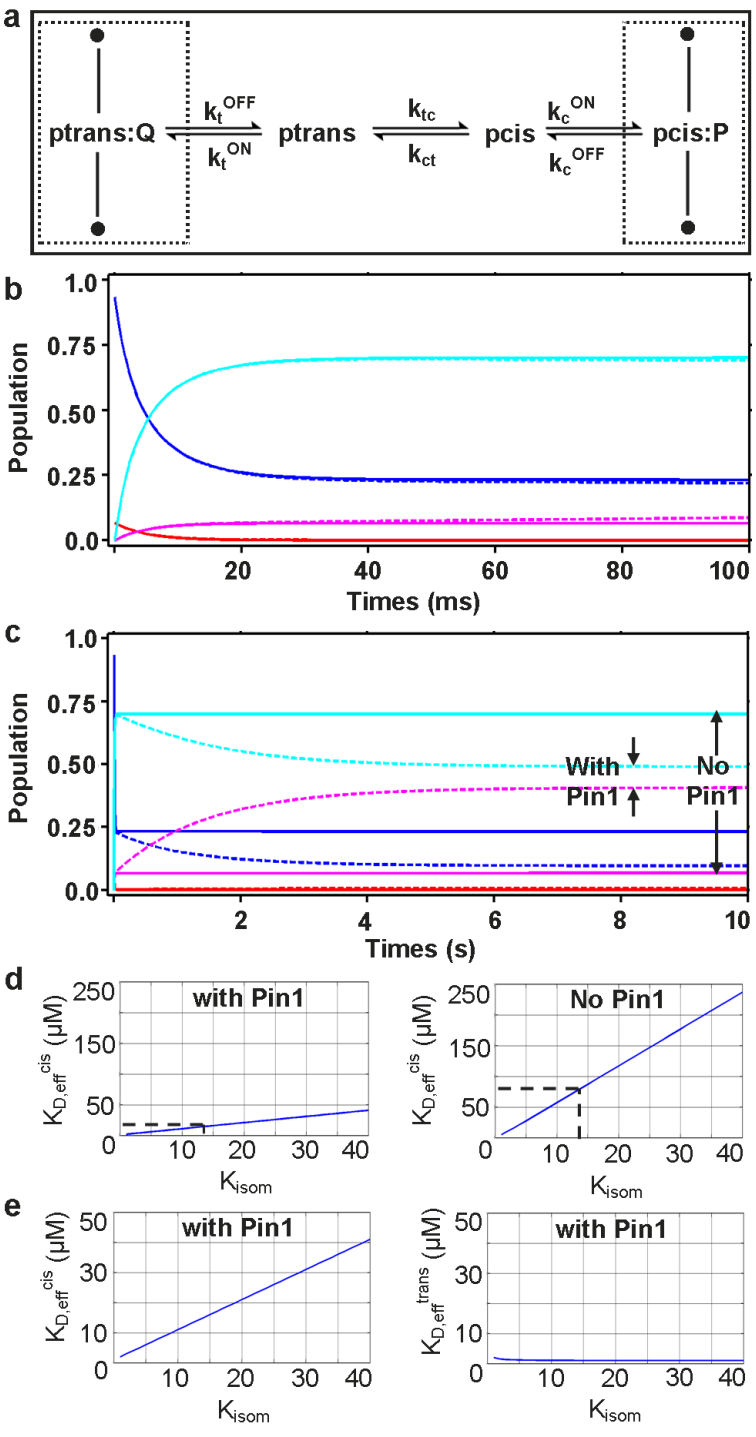
where K_D^{cis} and K_D^{trans} are the intrinsic dissociation constants for the *cis* and the *trans* isomers, respectively, K_{isom} (= $[T]_{free}/[C]_{free}$) is the equilibrium constant of the free *trans* and *cis* isomers, and [P] and [Q] are the concentrations of the free isomer-specific binding partners. The effective dissociation constants in the presence of Pin1 are

$$K_{D,eff}^{cis} = K_D^{cis}(1 + K_{isom}) \quad (3.13a)$$

$$K_{D,eff}^{trans} = K_D^{trans}\left(1 + \frac{1}{K_{isom}}\right) \quad (3.13b)$$

As discussed above, these expressions show that the effective dissociation constant for the isomer-specific binding partners increases in the absence of Pin1 (Figure 3.7d). Interestingly, this effect is more severe for the less-populated isomer (*cis*) and

Figure 3.7 Modulation of network connection and flux by Pin1. (a) The phosphorylated *cis* and *trans* isomers undergo isomerization and also bind to isomer-specific binding partners P and Q, respectively. The bound species pcis:P and ptrans:Q are the entry points to the two networks. The filled circles represent nodes and the lines represent connections in the networks. In the absence of Pin1, the two network modules (dotted box) are insulated from each other. In the presence of Pin1, the two modules are connected and act as a single module (solid box). (b) First 100 milliseconds of the time-dependence of free *cis* (red), free *trans* (blue), bound *cis* (magenta) and bound *trans* (cyan) in the presence (dotted lines) and absence (solid lines) of Pin1. (c) First 10 seconds of the time-dependence. The difference in the population of the bound species pcis:P and ptrans:Q in the presence and absence of Pin1 activity are shown by arrows. (d) Dependence of the effective dissociation constant for the *cis* isomer on the equilibrium constant (K_{isom}) in the presence and absence of Pin1. The dotted line is for pAICD, $K_{isom} = 14$. (e) The effect of the K_{isom} on the effective binding constant is more pronounced on the less populated isomer.


a
ptrans:Q
k_t^{OFF}
k_t^{ON}
ptrans
k_{tc}
k_{ct}
pcis
k_c^{ON}
k_c^{OFF}
pcis:P
b
Population
1.0
0.75
0.50
0.25
0.0
20
40
60
80
100
Times (ms)
c
Population
1.0
0.75
0.50
0.25
0.0
With Pin1
No Pin1
2
4
6
8
10
Times (s)
d
with Pin1
$K_{D,eff}^{cis}$ (μM)
250
150
50
0
10
20
30
40
K_{isom}
No Pin1
$K_{D,eff}^{cis}$ (μM)
250
150
50
0
10
20
30
40
K_{isom}
e
with Pin1
$K_{D,eff}^{cis}$ (μM)
50
30
10
0
10
20
30
40
K_{isom}
with Pin1
$K_{D,eff}^{trans}$ (μM)
50
30
10
0
10
20
30
40
K_{isom}

increases as a function of the equilibrium constant (K_{isom}) (Figure 3.7e). Thus, the catalytic activity of Pin1 can also dramatically modulate the effective dissociation constant of the isomer-specific binding partners in the relevant timescale of a signaling event, especially for the lower populated (*cis*) isomer.

Discussion

An emerging role of the prolyl *cis-trans* isomerization is that of a molecular timer where each isomer is associated with two distinct functions of the protein. For example, in the adaptor protein Crk, prolyl isomerization is involved in autoinhibition, where the *cis* conformation inhibits ligand binding (Sarkar et al., 2007). Another example is the prolyl isomerization driven activation of the filamentous phage fd infection of *Escherichia coli* cells (Eckert et al., 2005). PPIases can modulate these timers by tuning the lifetime of each isomer. Here we report the intrinsic, uncatalyzed exchange rate of the isomers in T668-phosphorylated APP as 0.0095 s^{-1} with the lifetime of the *cis* and the *trans* isomers being 2 min and 26 min, respectively. The peptidyl prolyl isomerase Pin1 speeds up this process 2000 fold (Pastorino et al., 2006) and thus, decreases the corresponding lifetimes to 60 ms and 0.5 s, respectively. The lifetimes of each isomer become relevant in the presence of isomer-specific interaction partners. For example, proline directed phosphatases are *trans* specific (Hamdane et al., 2006; Zhou et al., 2000) and the *cis* isomer is dephosphorylated only via its isomerization to *trans*. Thus, in the absence of Pin1 the phosphorylated *cis* isomer will be long-lived and have a higher probability of interacting with its binding partners, whereas the *trans* isomer will be rapidly depleted by phosphatases. Such tuning of the *cis-trans* molecular timer by Pin1 might be one of the key mechanisms by which it can influence various signaling pathways.

The results presented here show that the intrinsic, uncatalyzed isomerization rates (<0.01 s^{-1}) are very slow relative to the binding rates (k_{on}×[free protein] > 100 s^{-1}). These vastly different timescales have significant functional implications for protein-protein interaction networks that include isomer-specific binding partners, comprising a fundamental molecular switch that acts as a "circuit breaker", which is "open" in the absence and "closed" in the presence of Pin1 activity. In such networks the relative timescales of various events should be comparable to each other; otherwise the events with significantly different timescales would have no effective interaction (Papin et al., 2005). We have derived the mathematical framework for simulating the impact of *cis*-*trans* isomerization rates on two network modules that involve different isomer-specific interactions, one for *cis* and one for *trans* isomer. Without PPIase activity such network modules are insulated from each other on short timescales (<10 s), but are effectively coupled and modulated on this timescale when PPIase activity is present.

The addition of phosphorylation as a specificity determinant in the pS/T-P *cis*-*trans* molecular timer motif provides an additional level of regulation that is mediated by the temporally separated, competing activities of *trans*-specific (Brown et al., 1999; Weiwad et al., 2000) kinase and phosphatase enzymes. During the cell cycle, kinase and phosphatase activities must be separated into different periods, typically tens of minutes long (Novak et al., 2010). This temporal separation provides a mechanism by which the phosphorylated *cis* isomer can be produced during the kinase-active period, independent of Pin1 activity. However, in the absence of Pin1 activity, the connectivity between isomer-specific networks is lost (Figure 3.8a). The flux through the *trans*-specific network is higher but is rapidly shut-down due to *trans*-specific dephosphorylation during the phosphatase-active period, whereas the relatively lower flux in the *cis*-specific network persist much longer (Figure 3.8b). Thus, in the

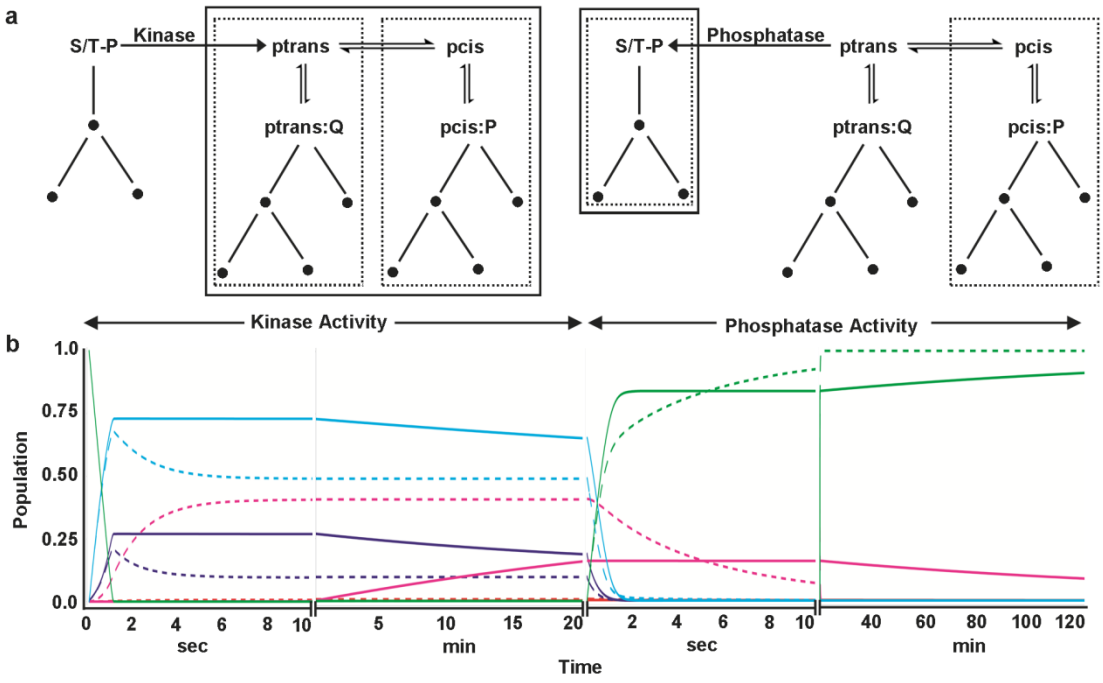


Figure 3.8 Effect of Pin1 catalysis on modular architecture of interaction networks. (a) The S/T-P motif is phosphorylated by proline directed kinases to generate phosphorylated *trans* isomer (ptrans) which isomerizes to *cis* isomer (pcis). The unphosphorylated and phosphorylated forms might have their own protein interaction networks. In the absence of Pin1 the isomer-specific networks are insulated modules (dotted box) while in its presence they form a single network module (solid box). In the absence of Pin1 *trans*-specific phosphatase activity results in two network modules (dotted box) while in its presence (solid box) the *cis*-specific interaction network is also rapidly shut-down. (b) Time-dependence of free S/T-P (green), free *cis* (red), free *trans* (blue), bound *cis* (magenta) and bound *trans* (cyan) in the presence (dotted lines) and absence (solid lines) of Pin1. The *trans*-specific kinases are assumed to be active for the first 20 minutes while the *trans*-specific phosphatases are active for the next 100 minutes. The *cis*-specific interaction network module persists for the entire phosphatase active period in the absence of Pin1, while the trans-specific network module shuts down rapidly.

absence of Pin1 activity the *cis*-specific network is "on" for the entire phosphatase-active period while the competing trans-specific network is turned "off". In the presence of Pin1 activity, the isomer-specific networks are connected and might have distinct function as compared to the disconnected networks. During the phosphatase active period, the connected networks are simultaneously turned "off". Thus, the *cis*-specific network only becomes dominant in the absence of Pin1 activity and during the phosphatase active period. This illustrates the impact of the combined regulatory effects of kinases, phosphatases, and Pin1 on the modular architecture of protein interaction networks that might result in distinct functional modules of a cell. Such modulation of isomer-specific interaction networks by Pin1 might be one of the key mechanisms by which it can influence various signaling pathways. These results offer for the first time a direct molecular mechanism to explain why the catalytic activity of Pin1 is necessary for its essential function in yeast (Zhou et al., 2000) or in all other Pin1-related functions identified so far (Nicholson and Lu, 2007).

In conclusion, these studies and analyses predict a powerful and novel role of Pin1 in coupling protein interaction network modules, modulating flux through them and also modulating the effective binding affinities of the isomer-specific binding partners. Thus, Pin1 activity can alter the modular architecture of the isomer-specific interaction networks and thereby explain its profound impact on biological and pathological processes. The challenge now is to identify such isomer-specific interaction networks for biological substrates of Pin1, such as APP phosphorylated at T668.

CHAPTER 4

DESIGN OF NOVEL PROBES TO STUDY ISOMER-SPECIFIC BINDING INTERACTIONS

The APP intracellular domain (AICD) plays important roles in APP trafficking (Sannerud and Annaert, 2009), APP processing (De Strooper and Annaert, 2000) and transcription of genes that are important in APP trafficking, phosphorylation and processing (Muller et al., 2008). Its function is also modulated by phosphorylation (Lee et al., 2003b). AICD has eight potential phosphorylation sites, of which, T668 phosphorylation has the most impact on its structure (Ramelot and Nicholson, 2001) and function (Lee et al., 2003b). Phosphorylation of T668 results in the emergence of *cis* and *trans* isomers, which have distinct structural features (Figure 1.2a). Even though the phosphorylation of T668 has been shown as an important event in AD, the distinct functional roles played by unphosphorylated AICD, *cis* pAICD and *trans* pAICD are not well understood (Pastorino et al., 2006). Moreover, the roles of these three species would also depend on two scenarios: AICD is in the full-length APP or AICD is a product of the γ-secretase processing. As part of full-length APP, AICD directs its trafficking and localization in the cell (Sannerud and Annaert, 2009), which in turn dictates the APP-processing pathway. On the other hand, as a proteolytic cleavage product, AICD binds to various cellular partners, is translocated into the nucleus and promotes transcription of certain genes (Muller et al., 2008). Deciphering the distinct roles played by the three AICD species would go a long way in illuminating the molecular mechanism of AD and provide us with specific therapeutic targets.

As discussed in Chapter 1, we hypothesized that the *cis* pAICD interacts with its specific binding partners and enhances the trafficking of APP into the endocytic

pathway. Moreover, isomer-specific interactions in the nucleus might also result in transcription of different sets of genes by the two isomers which might favor one or the other APP processing pathway. In order to investigate these possibilities, we aim to identify the pAICD isomer-specific binding partners in the cell. Such binding partners might be involved in secretory pathways, endocytic pathways, nuclear transport and gene transcription complexes. Their identification would enable us to investigate our hypothesis that the *cis* isomer of pAICD favors amyloidogenic processing of APP. The potential outcome of these studies is the identification of human proteins that bind specifically to the phospho-T668 form of the APP cytoplasmic tail, and the validation of one or more of these proteins as a therapeutic target in Alzheimer's disease.

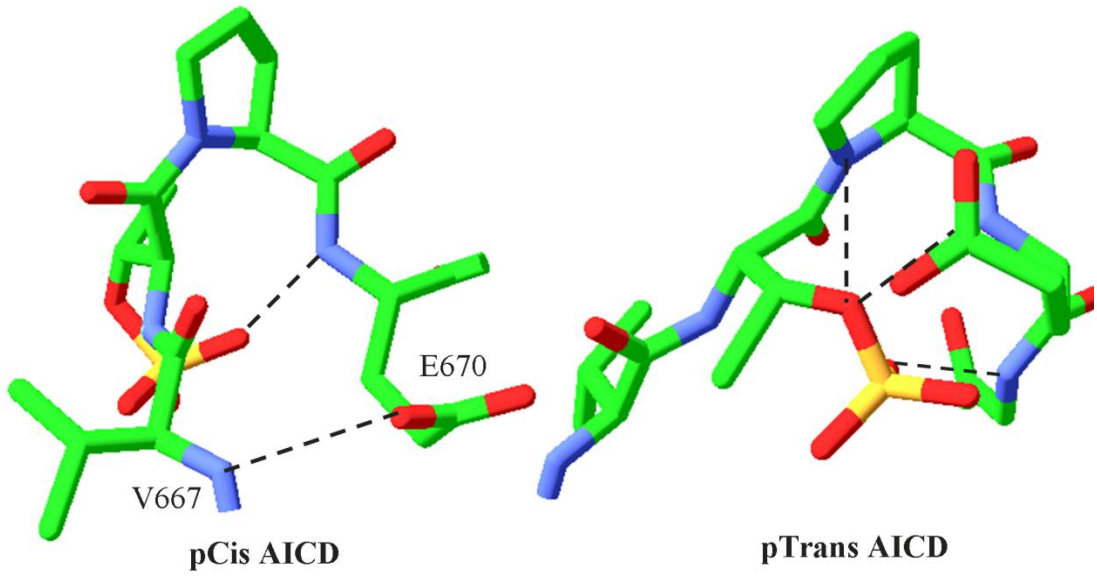


Figure 4.1: Structures of the *cis* and *trans* isomers of the VpTPE sequence in pAICD phosphorylated at T668. The H-bonds in each isomer are shown as dashed lines. The two isomers have distinct structural features.

Our hypothesis regarding the potential role of the *cis* and *trans* isomers of pAICD centers on putative isomer-specific recognition events. Since the two isomeric states are chemically identical, prior work such as proteomics studies would not have detected isomer specificity. Identification of such isomer-specific interactions would require molecular probes trapped in an isomeric state or inhibitors for isomer-specific binding partners. Since the *cis* and *trans* isomers isomerize quickly (within tens of minutes), their separation from each other is not possible. This makes it necessary to design mimics for each isomer. We utilized several strategies to design *cis* and *trans* isomer mimics. We also designed molecular probes and "fishing experiments" that would enable us to identify the isomer-specific binding partners in the cell.

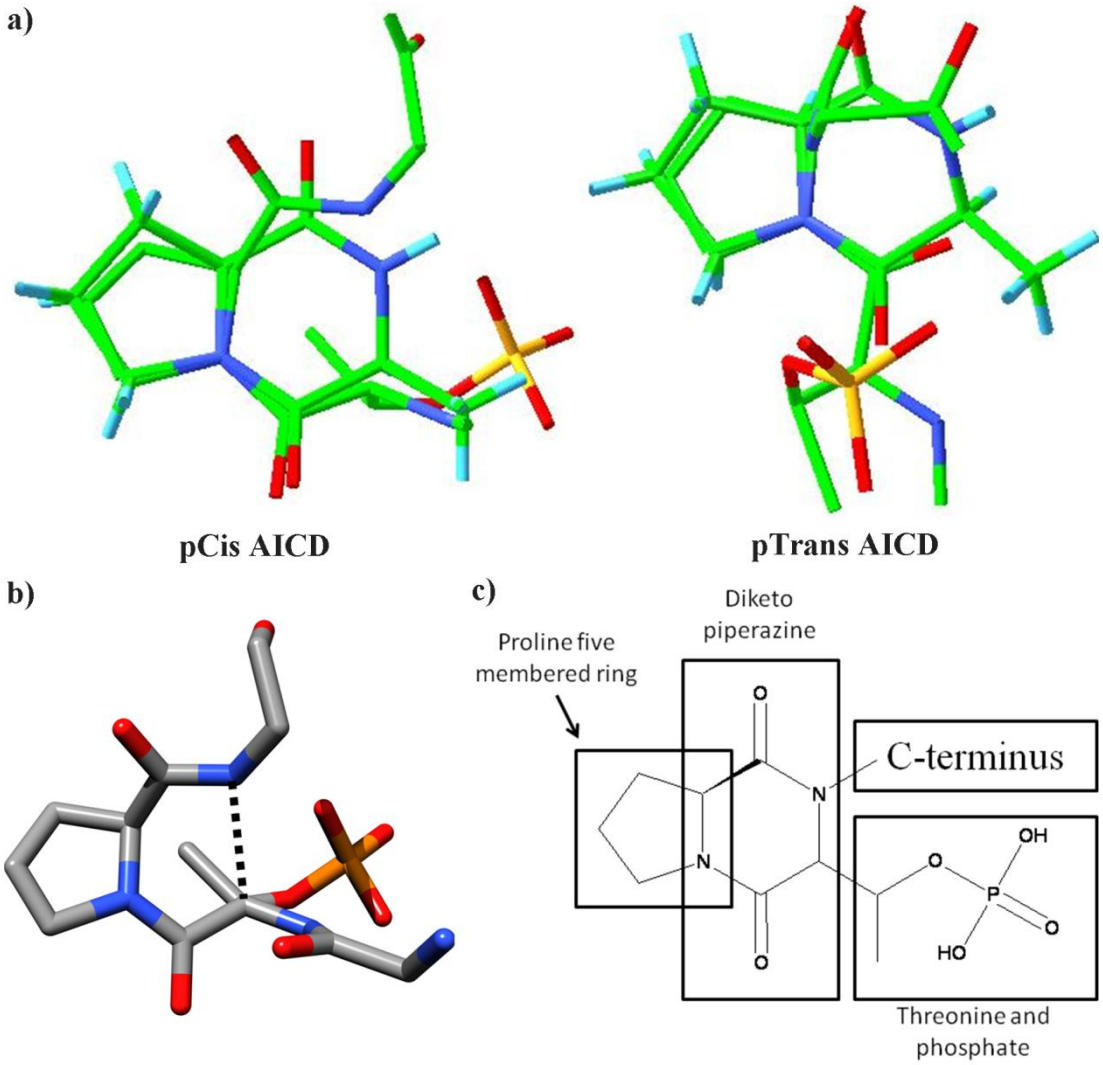
Cyclic peptide based on pAICD sequence as a *cis* isomer mimic

A *cis* enriched cyclic peptide was designed and synthesized based on the *cis* pAICD NMR structure (Figure 4.1). The distinctive features of the *cis* isomer as compared to the *trans* are:

1. Different hydrogen bonding patterns.
2. More extended backbone conformation in the *cis* isomeric state, aside from the kink of the peptide bond.
3. Presence of hydrogen bonding between V667 amide NH and E670 sidechain.

We reasoned that the *cis* isomer could be stabilized by forming a covalent bond between V667 amine NH_2 and E670 sidechain COOH. For this purpose, we used a six amino acid residue peptide with the sequence V667-pT-P-E-E-R672 derived from pAICD. A peptide bond between V667 amine NH_2 and E670 sidechain COOH was formed which increased the *cis* population.

Figure 4.2: a) Overlay of the phosphorylated cyclic dipeptide (pCDP) structure with structures of *cis* and *trans* isomers of pAICD. The pCDP backbone fits well to that of the *cis* isomer. b) Structure of the *cis* isomer of pAICD with the proposed covalent bond that would lock it in the *cis* conformation. c) The structure space of pCDP showing the four different dimensions that can be varied to generate a library of compounds based on pCDP. d) Comparison of the sidechains of alanine, proline and valine residues. The proline residue is replaced with alanine or valine in the pAICD sequence to design *trans*-locked peptides.


a)
pCis AICD
pTrans AICD
b)
c)
Diketo
piperazine
Proline five
membered ring
C-terminus
Threonine and
phosphate
d)
Alanine
Proline
Valine

Cyclic dipeptide as a *cis* isomer mimic

The *cis* isomer is well-mimicked in nature by proline containing cyclic dipeptides. The diketopiperazine core structure of these cyclic dipeptides overlay well with the *cis* conformer (Figure 4.2a). Cyclic dipeptides (CDP) are ubiquitous in nature and exhibit varied biological activities in mammals (Prasad, 1995). The CDP cycloHis-Pro has been shown to have biological activities involving endocrine, autonomic, analeptic and central nervous systems. Levels of cycloHis-Pro were found to be elevated in a group of patients diagnosed with various neurological/neuropsychiatric disorders (Wisniewski et al., 1994). The CDP cycloPhe-Pro has been shown to inhibit cancer cell growth and inhibit apoptosis (Brauns et al., 2004). A novel CDP 1-ARA-35b was shown to enhance motor and cognitive recovery after traumatic brain injury in rats and also showed neuroprotection (Faden et al., 2003). These are a few interesting examples of the diverse activities shown by the CDPs. Thus, the *cis* mimicking core structure of CDPs is well recognized by different biological pathways.

The phosphorylated Thr-Pro (pT-P) cyclic dipeptide should mimic the known structure of the *cis* isomer of the –pTPEERH– motif in the cytoplasmic tail of APP (Ramelot and Nicholson, 2001). The key feature of the *cis* isomer that is distinct from the structure of the *trans* isomer is the close proximity of the amide NH of T668 and CO of P669 which are within a bond-forming distance (Figure 4.2b). Closure of the pT-P backbone atoms to form a 6-membered ring is accomplished by addition of a second peptide bond which locks the pT-P into the *cis* conformation (Figure 4.2b). The resulting 5, 6-membered double-ring phosphorylated cyclic dipeptide (pCDP) (Figure 4.2c) is completely trapped in the *cis* conformation. It is important to note that this small molecule is a substrate analogue for Pin1, since it mimics the *cis* isomer of pAICD. Hence, it should not bind the enzyme tightly. Enzymes evolve to bind the

transition state of a reaction most tightly and bind the substrates and the products weakly. Accordingly, we expect that pCDP should have very weak affinity for Pin1. Since the WW domain of Pin1 does not interact with *cis* isomer, pCDP should interact only with the PPIase domain of Pin1. On the other hand, we expect much tighter interaction of pCDP with the cellular binding partners that specifically recognize the *cis* isomer of pAICD.

Modifications of pCDP for efficient cellular uptake

Phosphorylated compounds are not effective in penetrating the cell wall due to the negative charges on the phosphate group. Hence, for the efficient uptake of such compound by cells the phosphate group needs to be protected. The protecting group bis-pivaloyl-oxymethyl (bisPOM) has been found to be particularly effective in delivering phosphorylated drugs into cells (Rutschow et al., 2002). Once inside the cell the bisPOM group is removed by esterases resulting in the active phosphorylated compound. Thus, one strategy for delivering pCDP into the cell would be to protect the phosphate group with bisPOM group.

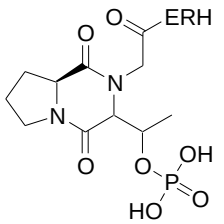
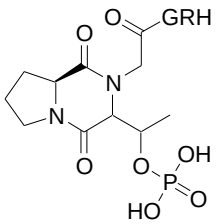
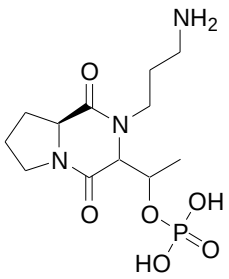
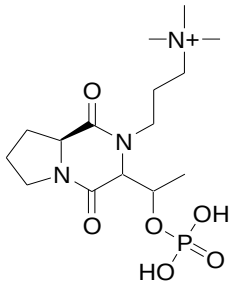
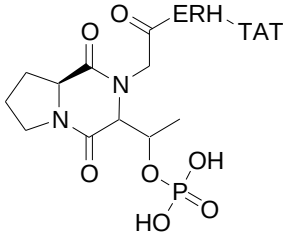
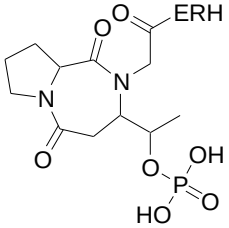
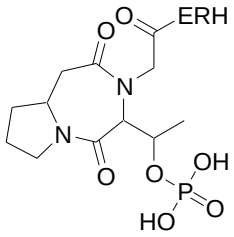
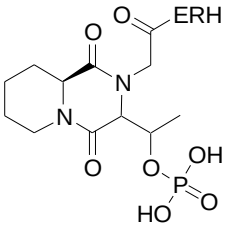
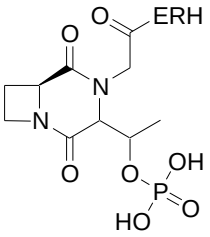
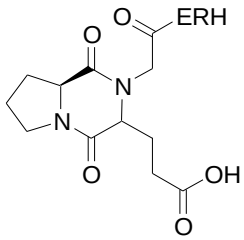
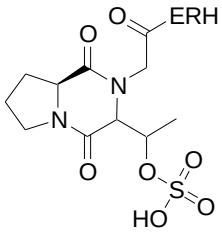
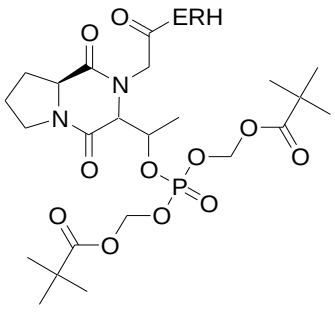
Another strategy to deliver pCDP into the cell is by coupling it to the TAT-tag (Brooks et al., 2005). The HIV-TAT derived peptide is a small basic peptide that has been efficiently used to deliver molecules of various sizes (small compounds to proteins and nucleic acids) into the cell. The TAT-tag contains a stretch of nine basic residues with the sequence RKKRRQRRR. In order to attach other residues, such as the TAT-tag, to pCDP it is necessary to attach a linker (Figure 4.2c) to this compound. The important feature of this linker is the presence of a carboxyl group, which would enable the use of this molecule in solid-phase peptide synthesis to add to any peptide sequence as the N-terminal residue. The attachment of the linker would also enable the use of pCDP as a bait to identify *cis* isomer-specific binding partners in the cell.

Cyclic dipeptide as an inhibitor

The pCDP compound mimicking the *cis* conformation of APP can also be used as a lead compound for designing inhibitors, which would specifically block the activity of the *cis*-isomer specific binding partners and thus, the amyloidogenic APP processing pathway. Exploration of the structure space of pCDP using these compounds should provide us with a library of compounds that can be screened for subsequent drug development for the prevention and treatment of Alzheimer's disease. The proposed compounds and variations thereof are described in a pending US Patent Application (US20080058276_A1) and in a recent Invention Disclosure (D-4784, Cornell Center for Technology, Enterprise & Commercialization). The core structure of the phosphorylated cyclic dipeptide (pCDP) features four major structural elements (Figure 4.2c). Each of these features can be varied to achieve optimal binding of the dipeptide to its target, corresponding to variation of pCDP along four dimensions in structure space. With regard to structure-activity relationships (SAR), the four dimensions are: C-terminus, diketopiperazine ring, proline ring and the phosphate moiety. A few representative examples of the variation of the structure space of pCDP are shown in Table 4.1. Various peptide sequences can be added to the C-terminus of pCDP such as the minimal AICD sequence ERH (compound **1**, Table 1), GRH (positive charges on R and H would neutralize the negatively charged phosphate, **2**), positively charged primary amine (**3**), tertiary amine (**4**) and TAT tag (Brooks et al., 2005) following the minimal AICD sequence which could also assist cellular uptake (**5**). The positive charge at the C-terminus on compounds **2**, **3** and **4** are designed to counter the negative charges of the phosphate group, which may facilitate uptake of these molecules by the cell.

The second dimension of structural variation concerns the diketopiperazine moiety (Figure 4.2c). The diketopiperazine ring is rich in heteroatoms and changing

Table 4.1: Variations of pCDP structure

1 	2 	3 
4 	5 	6 
7 	8 	9 
10 	11 	12 

achieve optimal binding of the dipeptide to its target, corresponding to variation of pCDP along four dimensions in structure space. With regard to structure-activity relationships (SAR), the four dimensions are: C-terminus, diketopiperazine ring, proline ring and the phosphate moiety. A few representative examples of the variation of the structure space of pCDP are shown in Table 4.1. Various peptide sequences can be added to the C-terminus of pCDP such as the minimal AICD sequence ERH (compound **1**, Table 1), GRH (positive charges on R and H would neutralize the negatively charged phosphate, **2**), positively charged primary amine (**3**), tertiary amine (**4**) and TAT tag (Brooks et al., 2005) following the minimal AICD sequence which could also assist cellular uptake (**5**). The positive charge at the C-terminus on compounds **2**, **3** and **4** are designed to counter the negative charges of the phosphate group, which may facilitate uptake of these molecules by the cell.

The second dimension of structural variation concerns the diketopiperazine moiety (Figure 4.2c). The diketopiperazine ring is rich in heteroatoms and changing this ring structure will place them in different orientation, thus, exploring the hydrogen bonding possibilities with the target. The six-member diketopiperazine ring can be expanded to a seven or eight-member ring by including β-amino acids (**5** and **6**). The third structural feature that can be optimized is represented by the backbone of the proline ring. In the design of Pin1 inhibitors, it has been observed that increasing the 5-membered ring to a 6-membered ring increases the affinity of the substrate to the enzyme (Zhang et al., 2007). Variations in this structure space can be made by using four (**9**), five (**7**), six (**8**) member rings, dimethyl-proline, fluoro-proline and unsaturated proline. Finally, the fourth dimension of structural variation is the threonine sidechain and the phosphate group. Variations include Glu (**10**) or Ser instead of Thr, sulfate (**11**) instead of phosphate and phosphate group protected by bis-pivaloyloxymethyl (POM) groups (**12**) (Hwang and Cole, 2004). An exhaustive

combination of these variations will give us a library of 5×4×6×4 i.e. 480 compounds based on pCDP.

Design of *trans*-locked peptides

As discussed in chapter 2, the TPEE region in unphosphorylated AICD forms a transient helix capping box structure which stabilizes the *trans* isomer (Ramelot and Nicholson, 2001). Upon phosphorylation the helix capping box structure is disrupted resulting in the emergence of *cis* isomer (~ 7%) along with the *trans* isomer. The *cis*-peptide bond has much higher free energy as compared to the *trans* peptide bond due to steric hindrance. In Xxx-Pro peptide bond, the *trans* isomer is relatively destabilized resulting in an appreciable amount of *cis* isomer. Hence, mutating P669 in pAICD should stabilize the *trans* isomer resulting in ~ 100% *trans* peptide. The proline residue was mutated to Ala or Val (Figure 4.2d) resulting in *trans* locked peptides *trans*-locked A or *trans*-locked B. The P669A mutation retains the C^{β} methylene feature but replaces the rest of the sidechain with H. The *trans* isomer in this sequence (-VpTAEE-) should be stabilized to ~ 100% and adopt the same structure as VpTPEE in pAICD. Although binding of pAICD-P669A to *trans*-specific cellular partners might be weakened due to the increased conformational flexibility of the free peptide (Ala has more conformational degrees of freedom than Pro) or by sub-optimal packing to hydrophobic pocket due to loss of Pro ring atoms. The P669V mutation retains both the C^{β} and C^{γ} positions of Pro but adds a CH_3 group close to the backbone. The interaction of pAICD-P669V with trans-specific binding partners might be weakened due to the bias of Val towards extended conformations and possible steric clashing in binding interfaces. Such weakened affinities can be overcome by the use of higher inhibitor concentration as long as these Pro-substituted peptides are specific for the *trans* isomers specific binding partners.

Identification of isomer-specific binding partners

There are potentially several different binding partners, most likely proteins, in the cell that bind to the VpTPEE region of AICD. The relative populations of *cis*-specific and *trans*-specific binding partners of the VpTPEE sequence in the cell is entirely unexplored and is therefore unknown. Due to the distinct structural features of the *cis* and *trans* isomers, it is unlikely that a single binding partner would tightly bind to both forms. Since we are anticipating a trafficking role for these binding partners, a tighter and highly specific interaction is expected. Anticipating proteins as the major binding partners, we utilized Benzophenylalanine (Bpa) (Figure 4.3a) as a photo-activable group (Dorman and Prestwich, 1994) to covalently capture the binding partners and facilitate their identification. The strategy of this "fishing experiment" is outlined in Figure 4.3b. In these experiments, the probe consists of three components:

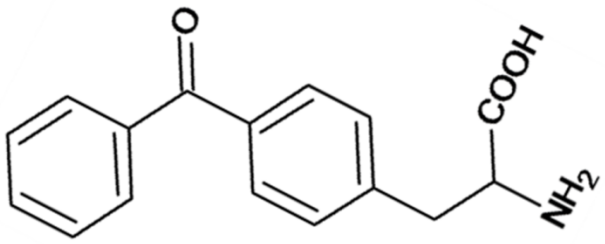
1. The "bait' is *cis* analogue, *trans* analogue or natural sequence (VpTPEER) from pAICD.
2. The "hook" is the Bpa which is photo-activated by UV irradiation at 350-360nm and covalently binds to nearby C^{α}-carbon of the binding protein.
3. Biotin as a tag to identify or purify the covalently bound binding protein.

The cell lysate is incubated with the probe allowing the equilibration of the isomer-specific interactions and then irradiated with UV light to covalently trap the binding partners. Cell lysate is then analyzed by Western blotting where horseradish peroxidase-linked streptavidin is used to identify the binding partners.

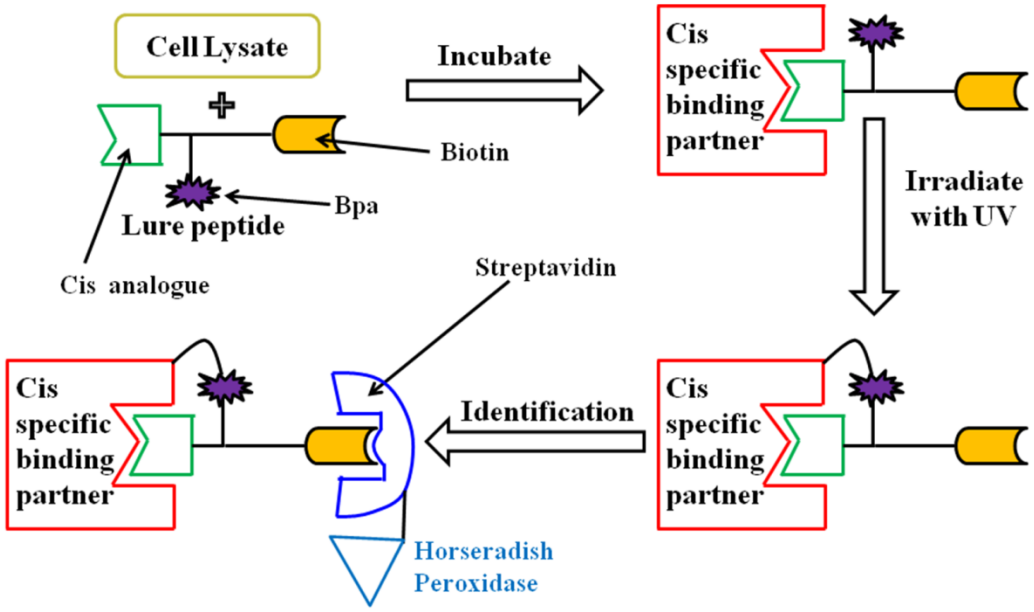
In order to identify the isomer-specific binding partners, two strategies are being used. First, the natural sequence (VpTPEER) is used as the bait and the *cis* (pCDP) or *trans* (*trans*-locked A or B) locked molecules are used as inhibitors. In this

Figure 4.3: a) The chemical structure of the photoactivable group, benzophenyl alanine (Bpa). This special amino acid can be easily put in any peptide sequence using standard peptide synthesis methods. b) Strategy of the fishing experiment. Cell lysate is incubated with the peptide probe which results in binding of the cellular binding partners with the 'bait' in the probe. The lysate is then irradiated by UV light (340-350 nm) which results in covalent linkage between the probe and the binding partner. The whole cell lysate is then run on a gel and probed with strepdavidin-horse radish paroxidase (HRP) to visualize and subsequently identify via mass spectrometry the binding partners. c) The probes and inhibitors (*trans*-locked and *cis*-locked) used in the fishing experiments. Probe1 has the natural pAICD sequence which has 93% *trans* and 7% *cis* population. This probe is used in conjunction with the *trans* or the *cis* locked inhibitors. Probe2 has the *cis*-locked pCDP as the bait. Probe3 has Val in place of Pro669 which would make it ~100%, i.e. it has the *trans*-locked peptide as the bait.

a)



b)



c)

Probes:

Probe1: Biotin-spacer-EVLFQGP-AA**VpTPEE**RH-Bpa-spacer-YPYDVPDYA-NH_2

Probe2: Biotin-spacer-EVLFQGP-**pCDP**-ERH-Bpa-spacer-YPYDVPDYA-NH_2

Probe3: Biotin-spacer-EVLFQGP-AA**VpTVEE**RH-Bpa-spacer-YPYDVPDYA-NH_2

Inhibitors:

***trans*-locked A:** GVVEVDAA**VpTAEE**RH-Bpa-SKMQQ

***trans*-locked B:** GVVEVDAA**VpTVEE**RH-Bpa-SKMQQ

pCDP

case the cell lysate will be probed with and without the inhibitors and both would be run and compared on a gel. It is important to note that the cell lysates can be prepared from various cell lines such as breast cancer cell line (HeLa cells), Chinese hamster ovary (CHO) cells or neuroglia cells (H4) each containing endogenous Pin1 or Pin1 knocked-down. The fact that VpTPEER has only 7% *cis* population should not hamper the identification process in the presence of *trans*-locked inhibitors. The inhibitors would block the *trans*-specific binding partners which would result in the interaction of the probe with the *cis*-specific binding partners. Binding of the *cis* isomer of the probe would deplete the free *cis* population, resulting in the isomerization of free *trans* to free *cis*. The *trans* to *cis* isomerization rate (k_{tc}) is 0.00064 s^{-1} in the absence of Pin1. To get ~99% conversion to *cis*, the probe, cell lysate and inhibitor should be incubated for at least 2hrs. The second strategy is to use the pCDP or *trans*-locked peptides themselves as probes.

The use of Bpa as the photo-activatable group in the probe of the fishing experiments provides the following advantages:

1. Bpa is chemically more stable than other groups such as diazo esters, aryl azides and diazirines. Moreover, Bpa is commercially available as a special aminoacid and can be easily attached to peptide probes using standard solid phase synthesis.
2. It is activated by non-damaging wavelengths for proteins.
3. It reacts preferentially to the unreactive C^{α}-H bond even in the presence of solvent water and other nucleophiles. This avoids the diffusion or release of the activated Bpa that can result in solvent trapping or nonspecific labeling.
4. The low lifetime (80 μs – 120 μs) of the Bpa excited state also reduces non-specific binding.

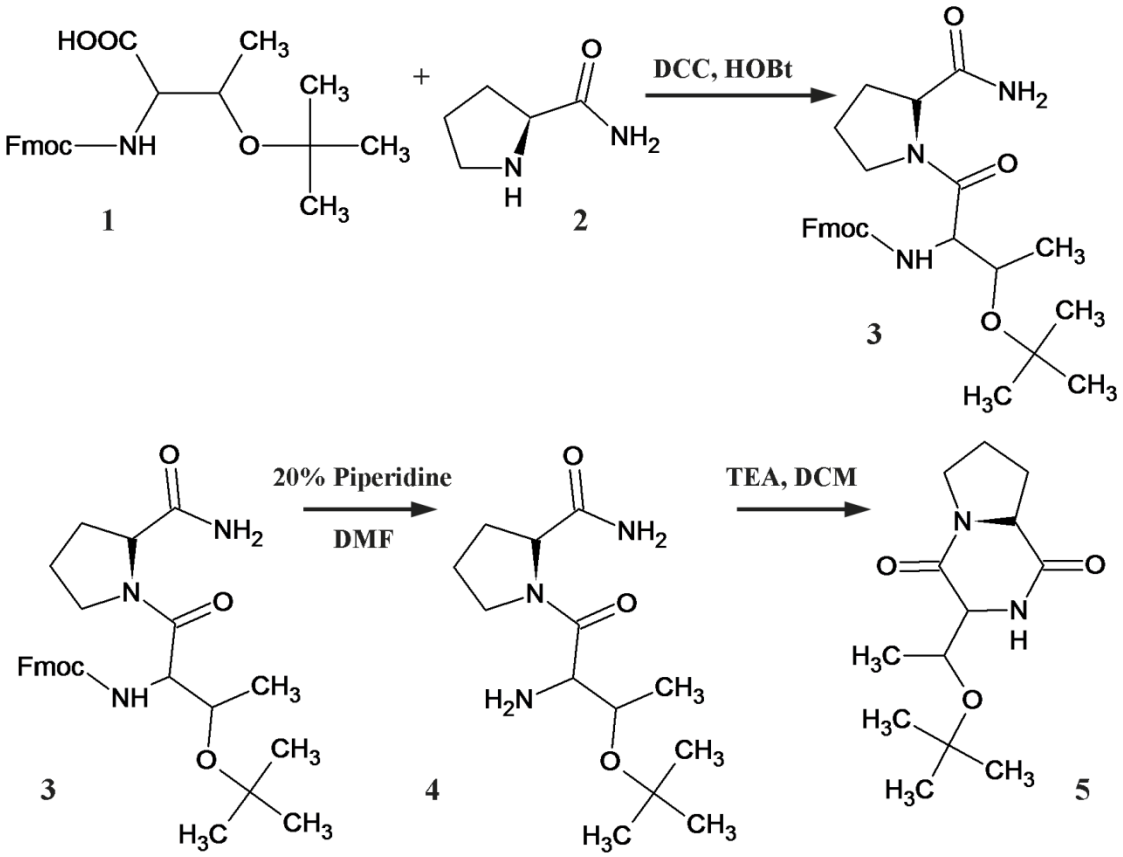


Figure 4.4: Synthesis of cyclic dipeptide. Fmoc and tBu protected threonine (**1**) was coupled to prolinamide (**2**) using DCC as a coupling reagent. The Fmoc protection group was removed from the dipeptide (**3**) using 20% piperidine and then cyclized using TEA as base to produce the protected cyclic dipeptide (**5**) as the final product.

Materials and Methods

Synthesis of six amino acid cyclic peptide

A six residue phosphopeptide corresponding to pAICD residues V667 to R672 was synthesized commercially (Tufts Core Facility) by solid phase synthesis and obtained on the resin. As part of the peptide design, the E670 side chain COOH was protected by orthogonal protection O-trimethylsilylethyl group (-OTMSEt) while all other functional groups in the peptide were protected by acid labile protection groups. The E670 sidechain was specifically deprotected by 1M TBAF prior to cyclization. The amide bond between E670 sidechain and V667 NH_2 was formed using 1e.q.

dicyclohexylcarbodiimide (DCC) (Sheehan and Hess, 1955) and 0.1 e.q. 1-hydroxy-benzotriazole (HOBT) in dimethylformamide (DMF) as solvent. The reaction was monitored using ninhydrin test which detects free N-terminus. Finally, the peptide was cleaved from the resin and all the protecting groups removed using trifluoroacetic acid (TFA). The peptide was purified by solvent extraction using water and diethyl ether and then analyzed by NMR and mass spectrometry.

Synthesis of phosphorylated cyclic dipeptide (pCDP)

The route for synthesis (Faden et al., 2003) of cyclic pThr-Pro dipeptide (pCDP) is shown in Figure 4.4. The solvent DMF was dehydrated overnight on silica beads. One eq. of Fmoc and tBu protected threonine (**1**) was dissolved in dimethylformamide (DMF) at 0°C (ice bath). Another 1 eq. of prolinamide (**2**) and 1eq. of DCC was added to it and stirred for 10 mins at 0°C. HOBt (1 eq) solution in DMF was added and the reaction mixture was stirred overnight at room temperature. The insoluble precipitate was filtered out. The filtrate has the product (**3**). The solvent DMF was removed using high vacuum and then the product (**3**) was dissolved in 20% piperidine in DMF to remove the Fmoc protection group from threonine amine. The reaction mixture was stirred at room temperature for 45 min. The solvent was again removed and the deprotected product (**5**) was dissolved in 5ml of dichloromethane (DCM) and 5ml of triethyl amine (TEA). This reaction mixture was stirred at room temperature overnight. Next day, the reaction mixture was diluted with DCM and washed with saturated $NaHCO_3$ solution (2 ×20 mL) followed by brine (20 mL). The organic layer was dried using anhydrous Na_2SO_4, concentrated under reduced pressure and purified on a silica gel column. Compound (**5**) is the protected cyclic dipeptide.

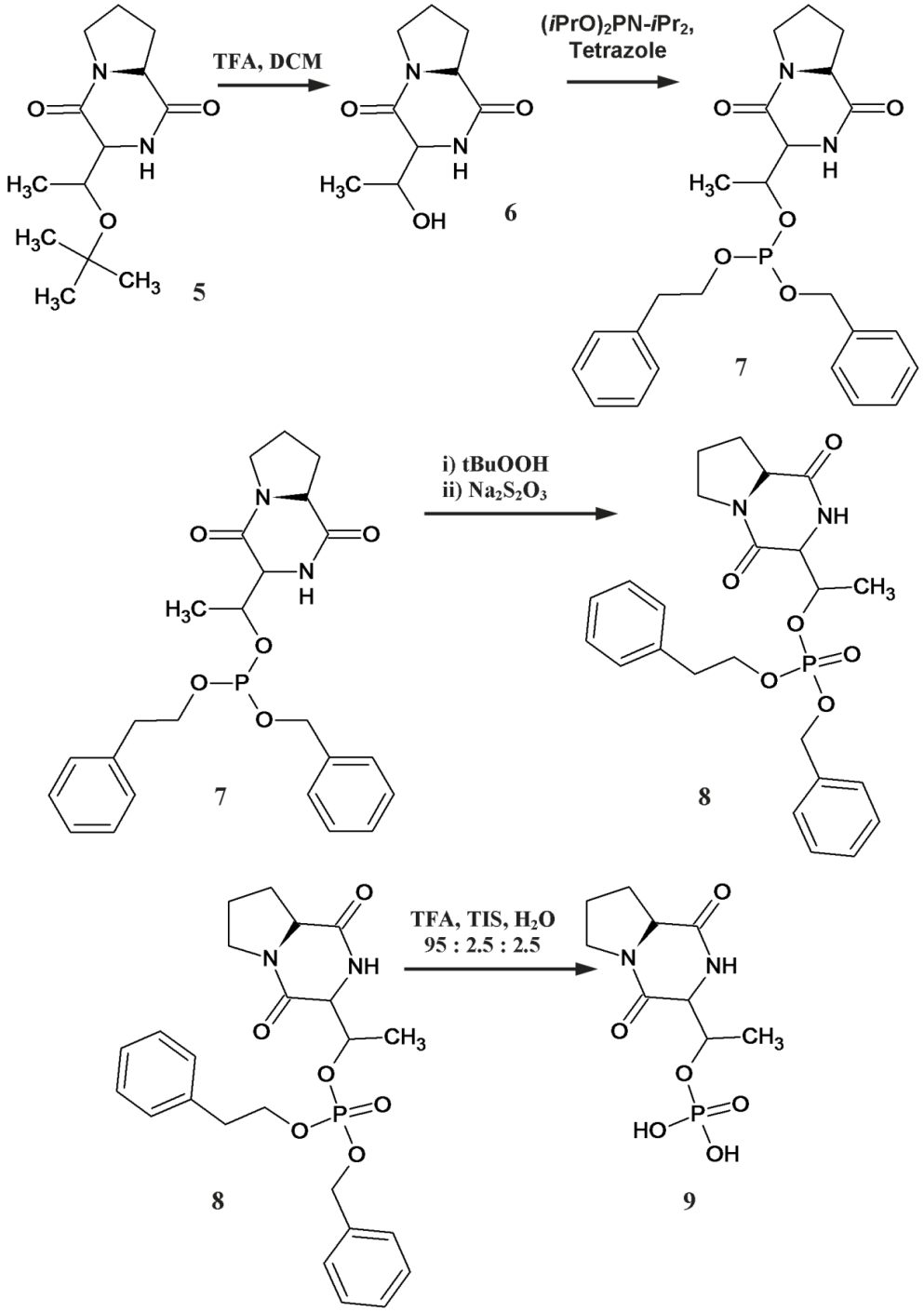
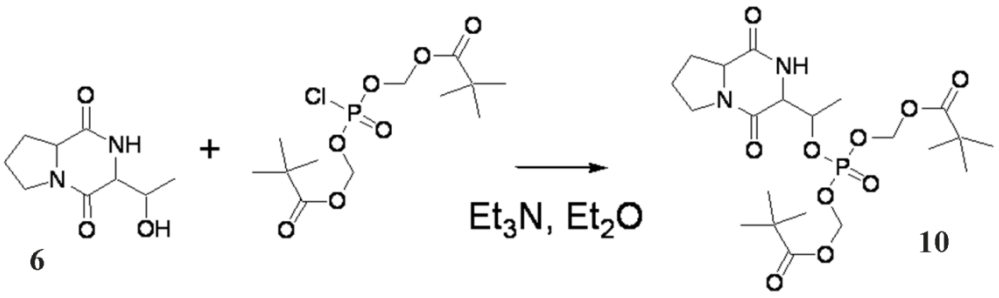


Figure 4.5: Phosphorylation of cyclic dipeptide. The tBu protection group was removed from the protected cyclic dipeptide (**5**) using 25% TFA. The reagent isopropyl phosphoramidite was used to put O-benzyl protected phosphite on the threonine sidechain. The phosphite (**7**) was oxidized to phosphate (**8**) using t-butyl hydroperoxide. The O-benzyl protection groups were removed to get the final product, phosphorylated cyclic dipeptide (**9**).

Phosphorylation of the cyclic dipeptide (Zhao and Etzkorn, 2007) was carried outas shown in Figure 4.5. The tBu protection group on theronine sidechain was removed by 25% TFA in DCM. The deprotected compound (**6**) was dissolved in 40ml acetonitrile and 4 eq. of tetrazole, and 2 eq. of $(i\text{PrO})_2\text{PN-}i\text{Pr}_2$ was added to it. The reaction mixture was stirred at room temperature for 20h. The reaction mixture was cooled at -40°C for 10 min and *tert*-Butylhydroperoxide (5.5 M in decane, 2 eq.) was added to it dropwise. The mixture was stirred at -40°C for 40 min. The cold bath was removed and the reaction mixture was stirred for an additional 45 min. It was again cooled back to 0°C and 10% aq. $Na_2S_2O_3$ (40 mL) was added to quench the reaction. It was stirred for 10 min and transfered to a separatory funnel using Et_2O (3×30 mL). The combined organic layer was washed with 10% aq. $Na_2S_2O_3$ (2×20 ml) and brine (20 mL), dried with Na_2SO_4 and concentrated to get (**8**). The protected phosphorylated cyclic dipeptide (**8**) was dissolved in 10mL of TFA:TIS:H_2O (95:2.5:2.5) and stirred for 2.5h. The TFA was evaporated by blowing dry nitrogen gas. It was then put under high vacuum to remove residual TFA and obtain the final product (**9**).



6

Figure 4.6: Proposed synthetic route for attaching bisPOM protected phosphate on cyclic dipeptide.

Attaching the bisPOM protection group

The protection of the phosphate group on pCDP by bis-pivaloyloxymethyl (bisPOM) can be accomplished according to the route shown in Figure 4.6. The

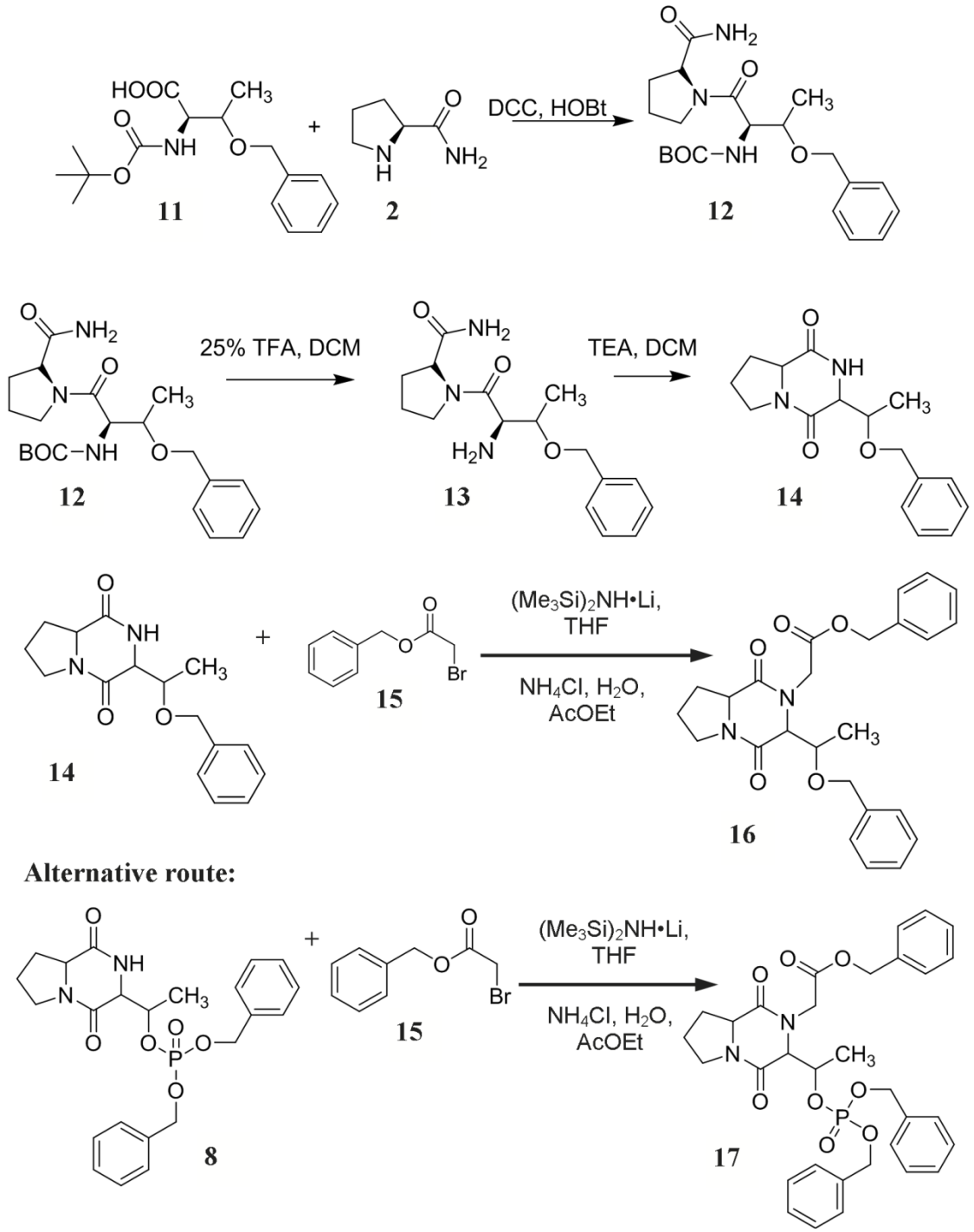


Figure 4.7: Two alternative routes for attaching a linker group to the phosphorylated cyclic dipeptide.

deprotected unphosphorylated cyclic dipeptide (**6**) is reacted with chloro-phosphoryl bisPOM reagent in the presence of triethylamine (Et_3N) in diethyl ether (Et_2O) (Hwang and Cole, 2004) to give the bisPOM protected cyclic dipeptide (**10**).

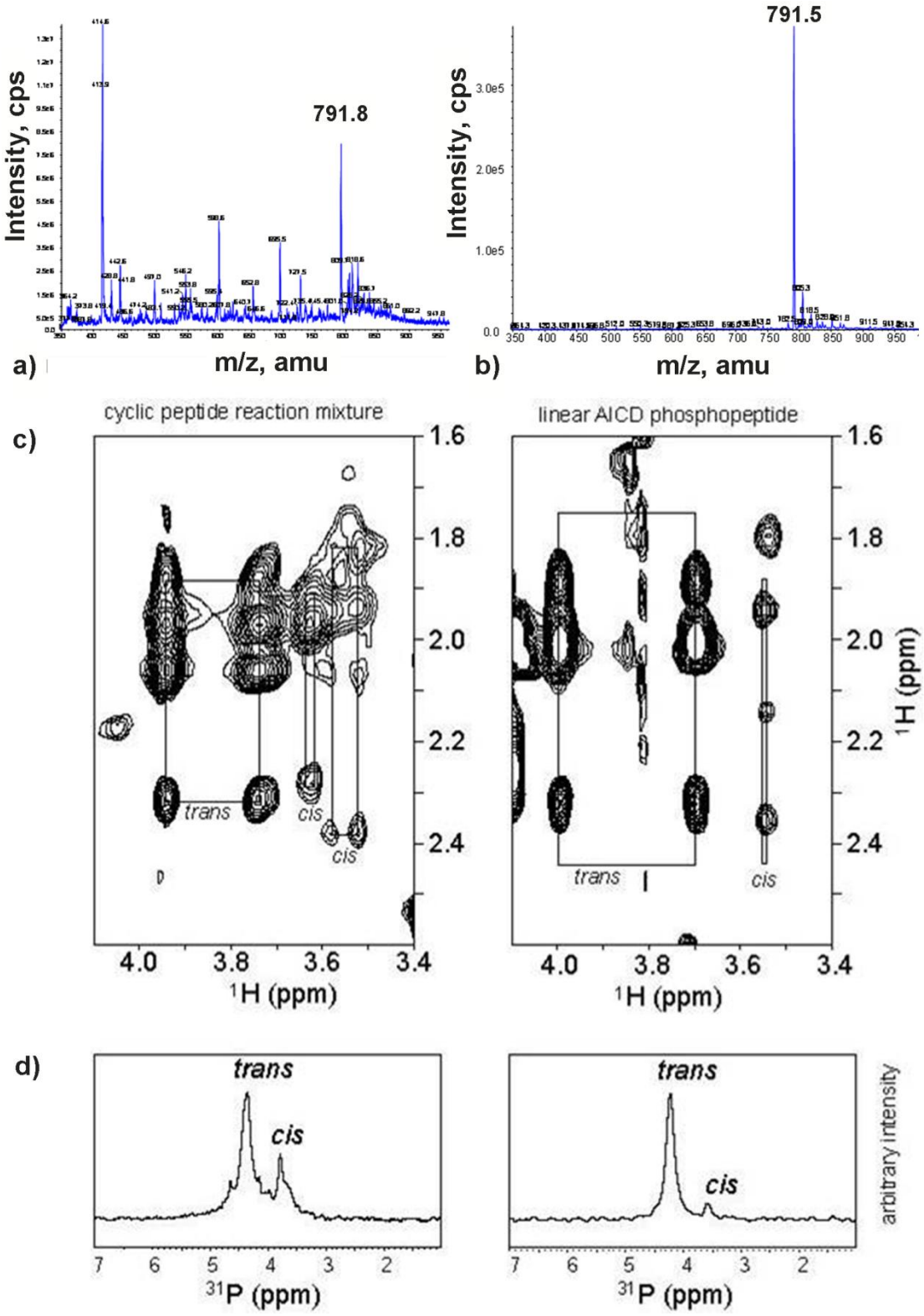
Proposed synthetic route for attaching linker on pCDP

Two alternative synthetic routes are proposed for the attachment of the linker (Estiarte et al., 2001) on pCDP (Figure 4.7). The already synthesized protected phosphorylated cyclic dipeptide (**8**) can be alkylated with benzyl bromoacetate (**15**) using lithium bis-trimethylsilyl amide (LHMDS) as base to get the product (**17**). Even though this is a one step synthesis, the chances of its success are limited. The strong base might degrade the compound (**8**) by reacting with the relatively labile O-benzyl protection group on phosphate. Hence, this is proposed as an alternative route. A better synthetic route would require a different starting material to synthesize the cyclic dipeptide as shown in Figure 4.7. The amine of threonine is protected by t-butyl carbamate (Boc) and the sidechain is protected by O-benzyl group. The resulting cyclic dipeptide (**14**) will have O-benzyl protection on the theronine sidechain which can be selectively removed by hydrogenolysis later. The linker group is added by alkylating with benzyl bromoacetate (**15**) using lithium bis-trimethylsilyl amide (LHMDS) as base to get the product (**16**). Selective deprotection of O-benzyl protection group from the threonine sidechain of (**16**) and subsequent phosphorylation using the method described earlier would give the final product (**17**).

NMR titration experiments

The six amino acid cyclic peptide was titrated into ^{15}N labeled WW domain (WWold construct). Solution of the cyclic peptide was prepared by dissolving 2.5mg of purified peptide in 310µl of NMR buffer which results in 10.2mM peptide,

Figure 4.8: a) LC-EMS of the reaction mixture of the synthesis of the cyclic VpTPEER peptide. Single-stage electrospray mass spectrometry spectrum shows peak at theoretical MW of cyclic peptide (791.8 amu) along with multiple impurities from reaction mixture. b) Selective detection of phosphate-containing species, showing cyclic peptide as by far the dominant component (linear form at 809 is less than 3%). c) ^{1}H-^{1}H TOCSY spectra of cyclic VpTPEER (left) and linear pAICD (right) phosphopeptides. d) The ^{31}P spectra of cyclic VpTPEER (left) and linear pAICD (right) phosphopeptides. The TOCSY and ^{31}P spectra show the enhancement of *cis* content in the cyclic peptide.


a)
Intensity, cps
791.8
m/z, amu
b)
Intensity, cps
791.5
m/z, amu
c)
cyclic peptide reaction mixture
linear AICD phosphopeptide
trans
cis
cis
trans
cis
1H (ppm)
1H (ppm)
d)
trans
cis
trans
cis
31P (ppm)
31P (ppm)
arbitrary intensity

assuming 100% purity. The ^{15}N WW domain was 0.43mM in NMR buffer consisting of 10mM HEPES, 10mM NaCl, 50mM Na_2SO_4, 1mM DTT and pH of 6.8. ^{15}N-^{1}H HSQC spectrum was collected for the apo protein. For each titration point, concentrated cyclic peptide was added to the same sample, thus, progressively increasing the concentration of the peptide. A total of 14 titration points were collected by adding 5, 5, 5, 5, 10, 10, 10, 10, 10, 80, 50, 15 and 30µl of cyclic peptide to the WW domain sample. ^{15}N ^{1}H HSQC spectrum was collected for each titration point. For each HSQC experiment 2048×256 complex data points and 32 scans were acquired. The spectra were processed using NMRpipe (Delaglio et al., 1995) and visualized with Sparky (Goddard, 2008). The shifted-sinebell window function was applied to all data while processing and then zero filled.

A similar titration experiment was also performed for cyclic the dipeptide (CDP) and the isolated PPIase domain of Pin1. The concentration of ^{15}N PPIase domain was 0.35mM. ^{15}N-^{1}H HSQC spectra were collected for four titration points with peptide (CDP) to protein ratio of 0, 0.5, 1, and 10. For each HSQC experiment 2048×256 complex data points and 48 scans were collected.

In order to titrate phosphorylated cyclic dipeptide (pCDP) into Pin1, a concentrated solution of pCDP was made by dissolving 11.4 mg of pCDP in 270µl H_2O and 30 µl D_2O with a final concentration of 137mM. This peptide was progressively titrated into ^{15}N labeled Pin1. The NMR buffer was 10mM HEPES, 10mM NaCl, 1mM DTT, 5mM NaN_3 and pH of 7.0. A total of six titration points were collected. The protein concentration was 0.35mM and the ratio of the peptide to protein used was 0, 1, 10, 30, 50 and 77. ^{15}N-^{1}H HSQC spectrum was collected for each titration point. For each HSQC experiment 2048×128 complex data points and 48 scans were collected.

Results and Discussion

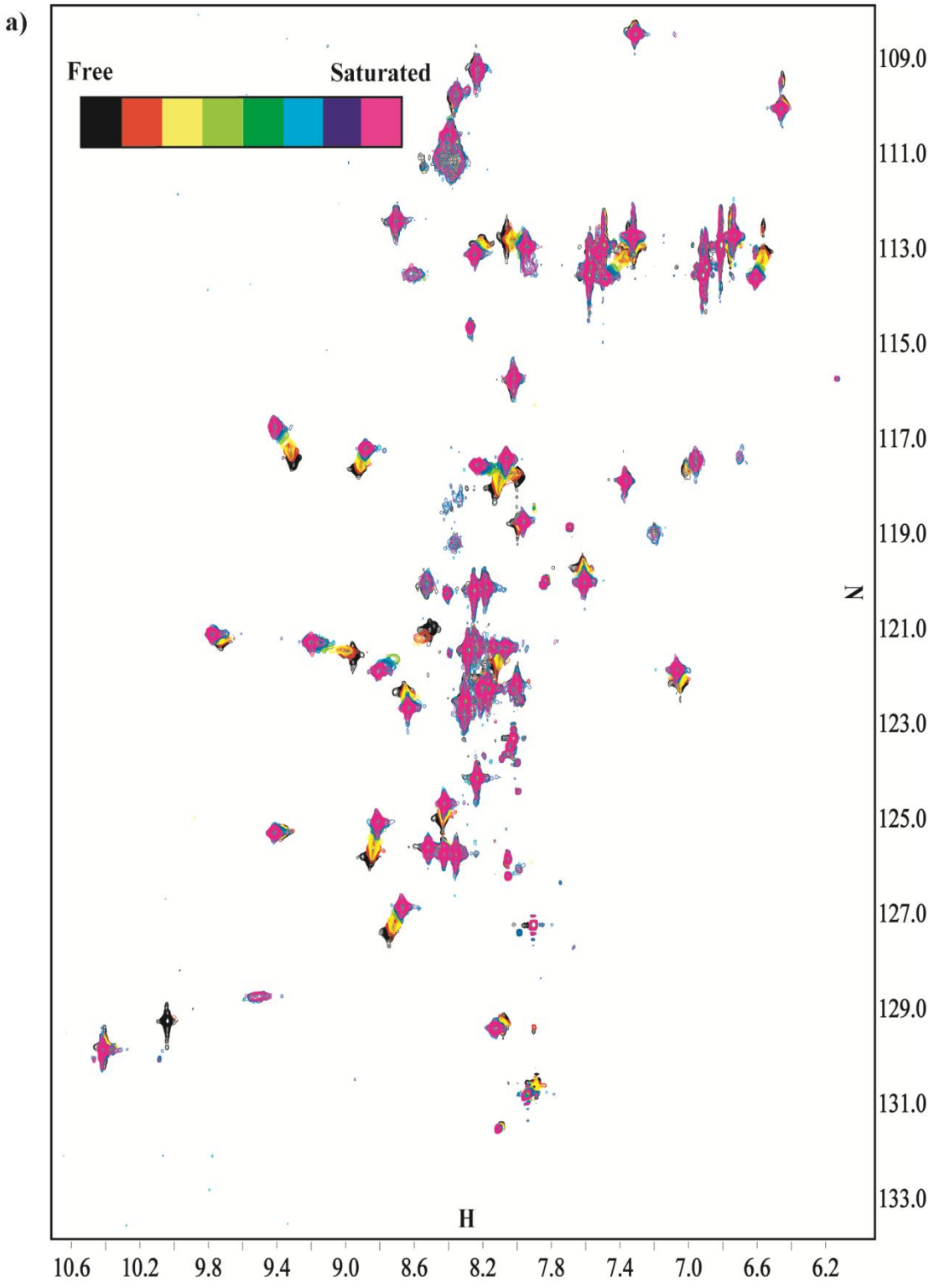
Cyclic peptide increases the *cis* content of VpTPEER sequence

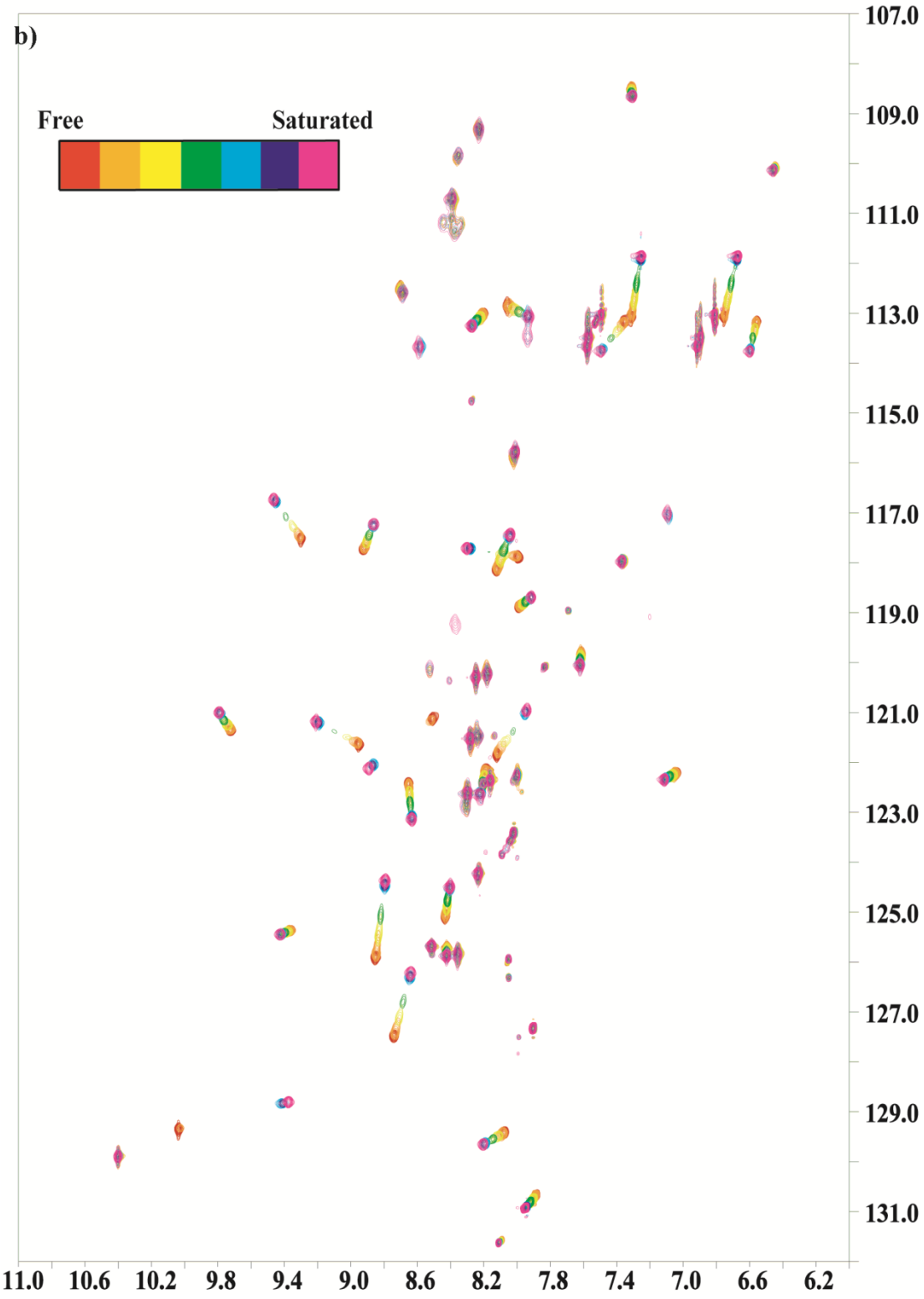
This six residue peptide (V-pT-P-E-E-R) was commercially synthesized as a linear peptide and was cyclized by forming an amide bond between the N-terminus NH_2 on V667 and sidechain COOH of E670. The extent of cyclization was quantified (Figure 4.8 a and b) by on-line LC/MS/MS using a triple quadrupole linear ion trap (4000Q Trap) performed in Cornell's BioResource Center. The cyclization was ca. ~100% complete. Since the reaction mixture contains almost 100% cyclic peptide, we analyzed the reaction mixture by NMR to determine *cis* and *trans* populations. 2D ^{1}H-^{1}H TOCSY (Figure 4.8c) and 1D ^{31}P (Figure 4.8d) spectra of the reaction mixture detect 30% *cis* content of the cyclic peptide which is considerable improvement on the 7% *cis* content of the linear peptide. The 2D ^{1}H-^{1}H TOCSY spectrum (Figure 4.8c) shows two *cis* populations that most likely reflect different hydrogen bond geometries adopted in the *cis* isomer.

Interaction of cyclic peptide with ^{15}N-labeled isolated WW domain

Titration of the cyclic peptide into isolated ^{15}N labeled WW domain showed chemical shift changes distinct from the titration of the linear peptide into isolated ^{15}N labeled WW domain (Figure 4.9 a and b). The chemical shift changes are due to the interaction of the *trans* conformation of the cyclic peptide which has a population of ~70%. Moreover, the differences in the trajectories of the peak shifts of the WW domain residues in response to linear and cyclic peptide titration indicate the different structures of the two *trans* peptides. No further experiments were performed with this cyclic peptide but it might be interesting to study the interaction of this cyclic peptide with *cis* specific binding partners, once they are identified.

Figure 4.9: a) Titration of cyclic peptide VpTPEER into ^{15}N labeled isolated WW domain. ^{15}N-^{1}H HSQC spectra acquired at each titration point are overlayed. b) Titration of the linear peptide TAR1 into ^{15}N labeled isolated WW domain. ^{15}N-^{1}H HSQC spectra acquired at each titration points are overlayed. Comparison of the two overlays (a and b) show less peak shifts in case of the cyclic peptide (a) and also different trajectories for some WW domain residues.


a)
Free
Saturated
109.0
111.0
113.0
115.0
117.0
119.0
121.0
123.0
125.0
127.0
129.0
131.0
133.0
N
H
10.6
10.2
9.8
9.4
9.0
8.6
8.2
7.8
7.4
7.0
6.6
6.2


b)
Free
Saturated
107.0
109.0
111.0
113.0
115.0
117.0
119.0
121.0
123.0
125.0
127.0
129.0
131.0
11.0
10.6
10.2
9.8
9.4
9.0
8.6
8.2
7.8
7.4
7.0
6.6
6.2

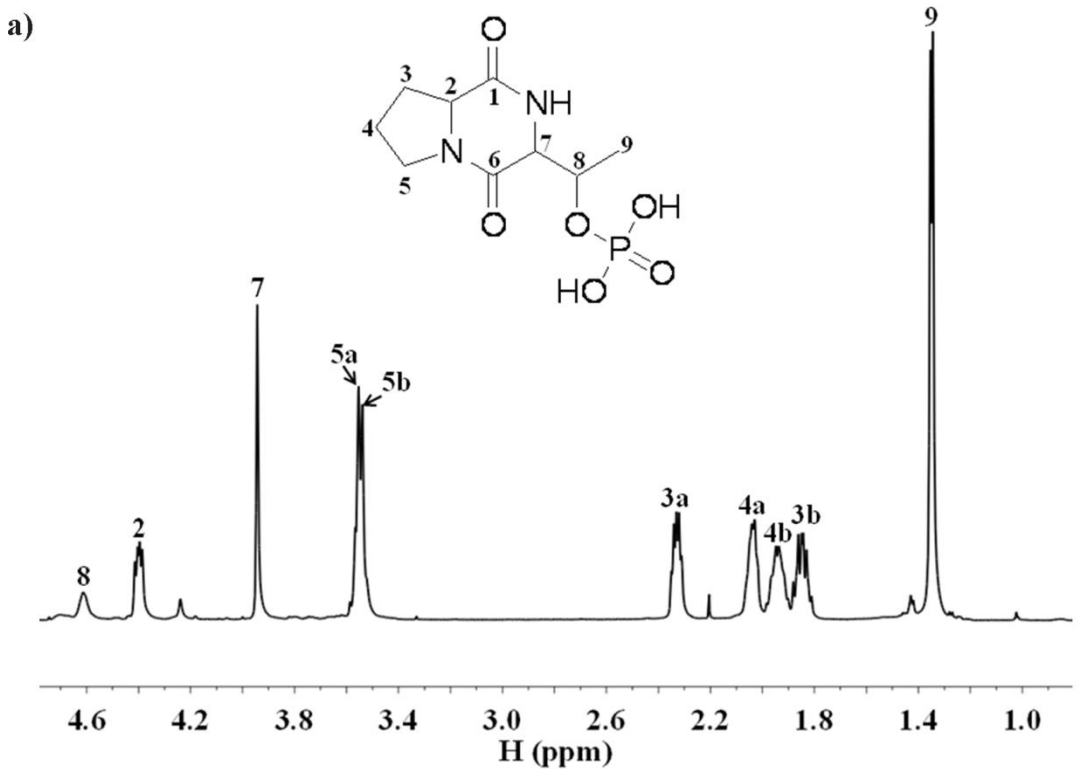
Successful synthesis of pCDP

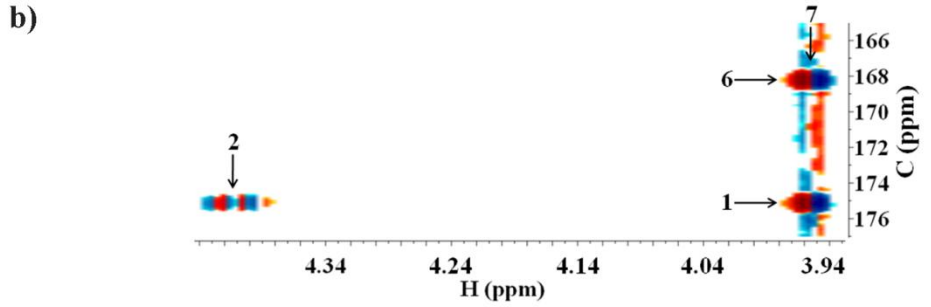
The phosphorylated cyclic dipeptide was successfully synthesized using the proposed scheme (Figures 4.5 and 4.6). 1D (^{1}H) and 2D (^{1}H-^{1}H COSY, ^{13}C-^{1}H HMQC and ^{13}C-^{1}H HMBC) experiments were used to verify the structure of the compound. The completely assigned 1D-spectrum of pCDP is shown in Figure 4.10a. The H^{α} (Figure 4.10b) of the threonine residue showed two-bond and three-bond coupling to the two carbonyl carbons, thus, validating the cyclic product. Further, ESI mass spectrum (Figure 4.10c) was obtained for the final product, and confirmed the correct mass for the pCDP compound (276.8 Da). Finally the doublet peak for threonine methyl was used to determine the purity of the product (Figure 4.10d). The final product was found to be greater than 95% pure.

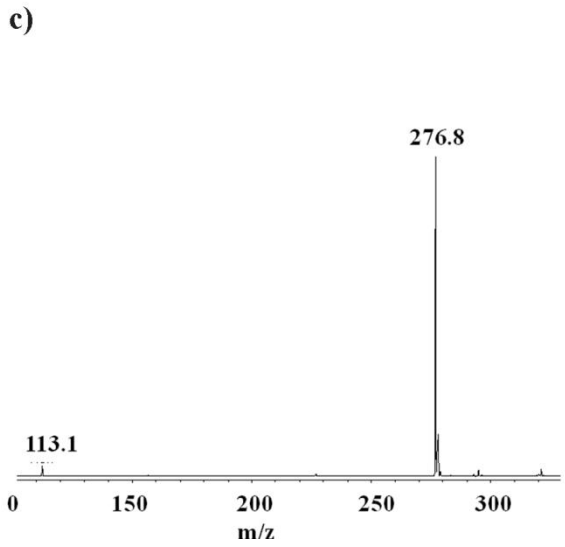
Interaction of unphosphorylated CDP with PPIase domain

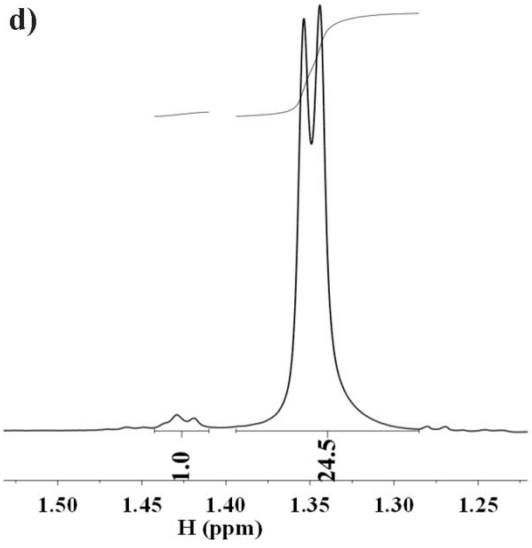
The WW domain of Pin1 binds specifically to the *trans* isomer of pSer/Thr-Pro motif whereas the PPIase domain seems to have higher affinity for *cis* than *trans* isomer. The first crystal structure (pdb code: 1PIN) solved for Pin1 had Ala-Pro dipeptide in the *cis* conformation bound to the PPIase domain (Ranganathan et al., 1997). Since the cyclic dipeptide is locked in the *cis* conformation, it should not be catalyzed by PPIase domain of Pin1 and should simply bind it. Moreover, we do not expect CDP to bind the PPIase domain tightly as CDP is a substrate analogue and not a transition state analogue. This provided us with an ideal opportunity to test the necessity of the phosphate group on substrates being recognized by this domain. We titrated in the unphosphorylated CDP into ^{15}N labeled PPIase domain. No peak shifts were observed even at 10:1 ratio of peptide to protein (Figure 4.11). Thus, the presence of phosphate group on the substrates is absolutely necessary for recognition by the PPIase domain of Pin1.

Figure 4.10: a) ^{1}H spectra of purified phosphorylated cyclic dipeptide (pCDP) in D_2O at 600MHz shows pure compound. Inset shows the structure of pCDP and the peaks in the spectrum are assigned as indicated. b) A section of the ^{13}C ^{1}H HMBC spectrum showing the two and three bond couplings of the α-protons to the carbonyl carbons. c) ESI-MS of pCDP confirms the purity of the sample and the success of the synthesis. The negative mode ESI mass spectrum was obtained for pCDP. d) The strong methyl doublet peak of the threonine sidechain was used to determine the purity of the sample. Peak integration shows the major product as more than 95%.


a)
O
3
2
1
NH
4
N
6
7
8
9
5
O
O
OH
P
HO
O
9
7
5a
5b
3a
4a
4b
3b
2
8
4.6
4.2
3.8
3.4
3.0
2.6
2.2
1.8
1.4
1.0
H (ppm)


b)
7
6
2
1
166
168
170
172
174
176
C (ppm)
4.34
4.24
4.14
4.04
3.94
H (ppm)


c)
276.8
113.1
0
150
200
250
300
m/z


d)
1.0
24.5
1.50
1.45
1.40
1.35
1.30
1.25
H (ppm)

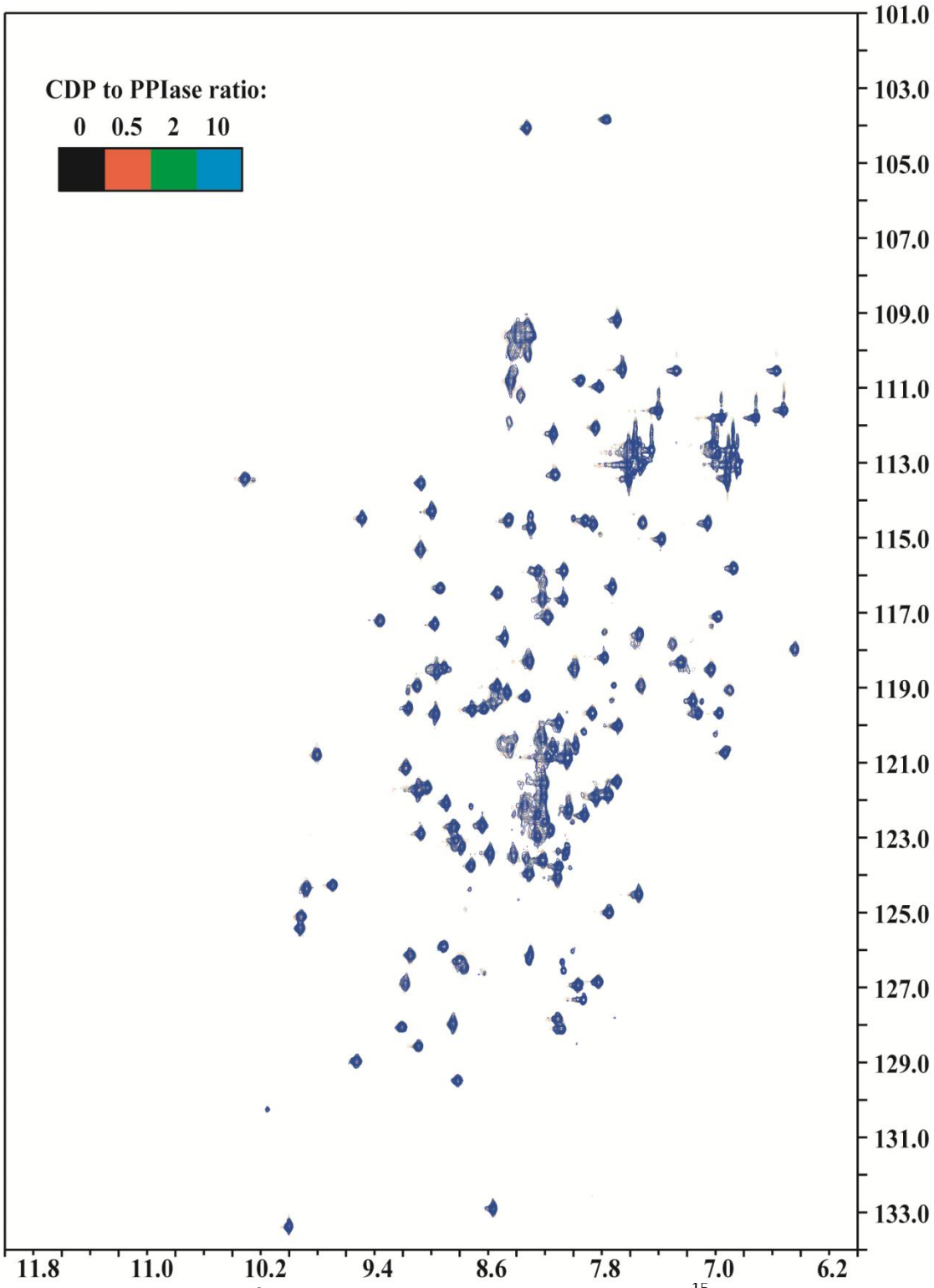


Figure 4.11: Titration of unphosphorylated cyclic dipeptide into ^{15}N labeled isolated PPIase domain of Pin1. No significant binding is observed.

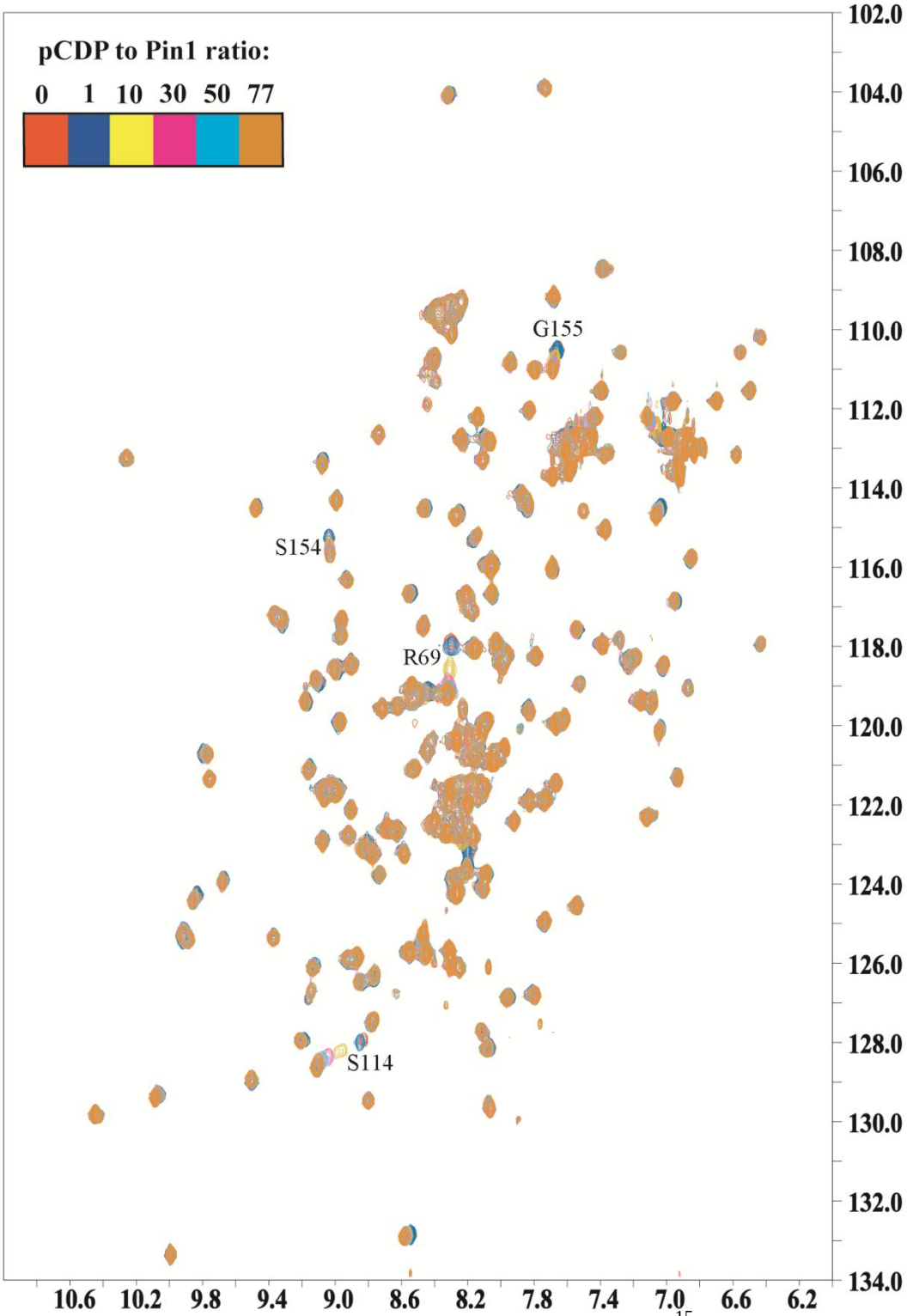


Figure 4.12: Titration of phosphorylated cyclic dipeptide into ^{15}N labeled isolated Pin1. Significant peak shifts are observed only in the PPIase domain.

Interaction of pCDP with Pin1

The phosphorylated cyclic dipeptide (pCDP) was titrated into ^{15}N labeled Pin1 in order to determine its specificity for the two domains of Pin1, PPIase and WW domain, and also the corresponding binding affinities. The titration experiment revealed that pCDP interacts specifically with the PPIase domain of Pin1. Chemical shift changes were not detected for the WW domain residues in this titration experiment, whereas the active-site residues of PPIase displayed considerable chemical shift changes (Figure 4.12). Residues in the PPIase domain that show considerable chemical shifts are R69, S114, S115, S154 and G155. The absence of interaction between pCDP and WW domain, and its binding to the active-site of PPIase domain (Figure 4.13) confirms our prediction that it mimics the *cis* isomer of pThr-Pro motif in pAICD. The chemical shift changes of the PPIase domain active-site residues from the titration experiment were used to determine the binding affinity of pCDP to the PPIase domain of Pin1. The residues R69, S114, S115, S154 and G155 displayed the largest chemical shift changes and were fit simultaneously (Figure 4.14) to determine the dissociation constant (K_D). The dissociation constant of pCDP to PPIase domain of Pin1 was determined to be 3mM. Since this small molecule mimics one of the substrates (*cis* isomer) of the enzyme Pin1, the weaker binding affinity is expected. Enzymes perform their catalytic activities by binding the transition-state tightly, thus, lowering the activation energy barrier and bind the substrates and products weakly for their efficient release. Thus, the small molecule pCDP is a true substrate analogue of Pin1 and mimics the *cis* isomer of pThr-Pro motif. This supports our prediction that pCDP and compound library based on this lead molecule should not inhibit Pin1. Importantly, these molecule are expected to bind tightly and competitively inhibit the *cis*-isomer specific binding partners in the cell.

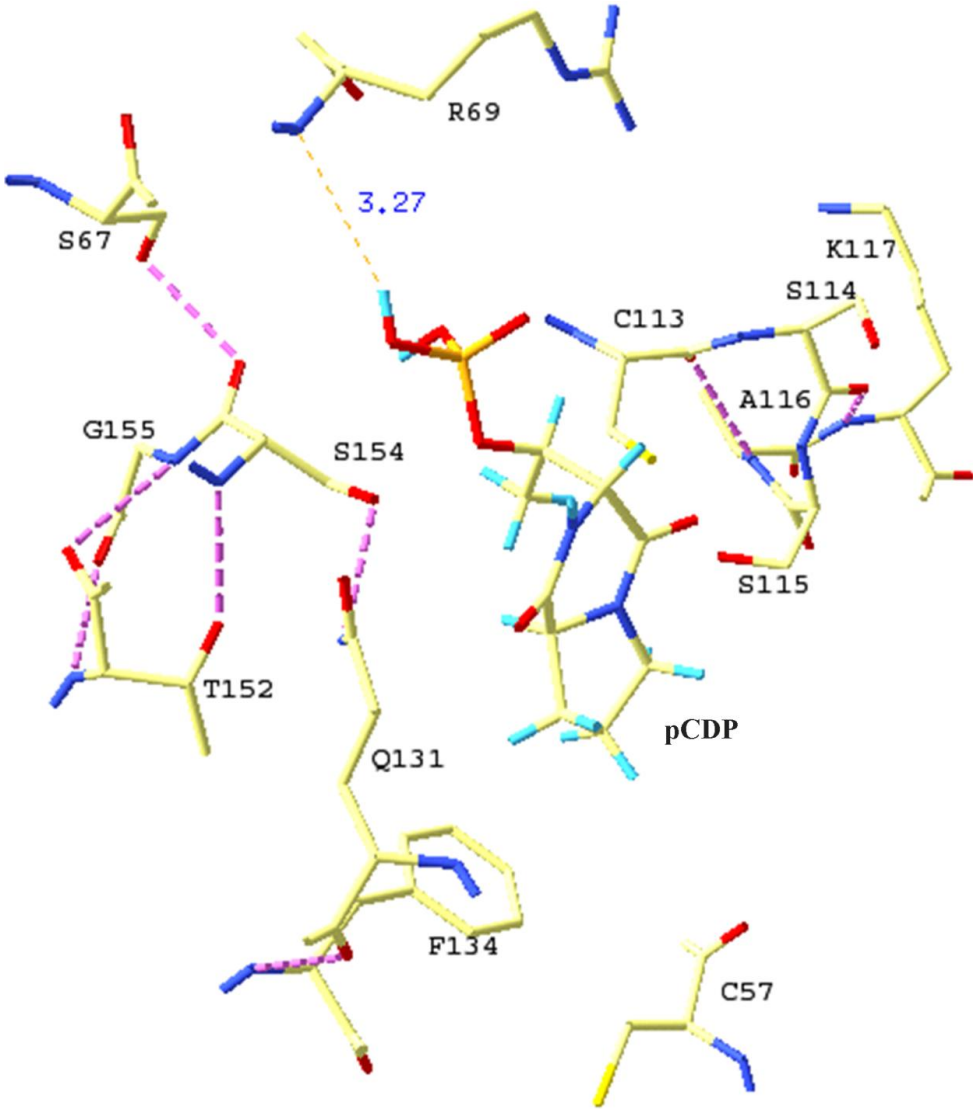


Figure 4.13: Interaction of pCDP with the active-site residues of the PPIase domain of Pin1. A molecular model of pCDP is docked into the active-site of the PPIase domain. The active-site residues are labeled and the H-bond interactions are shown as broken lines

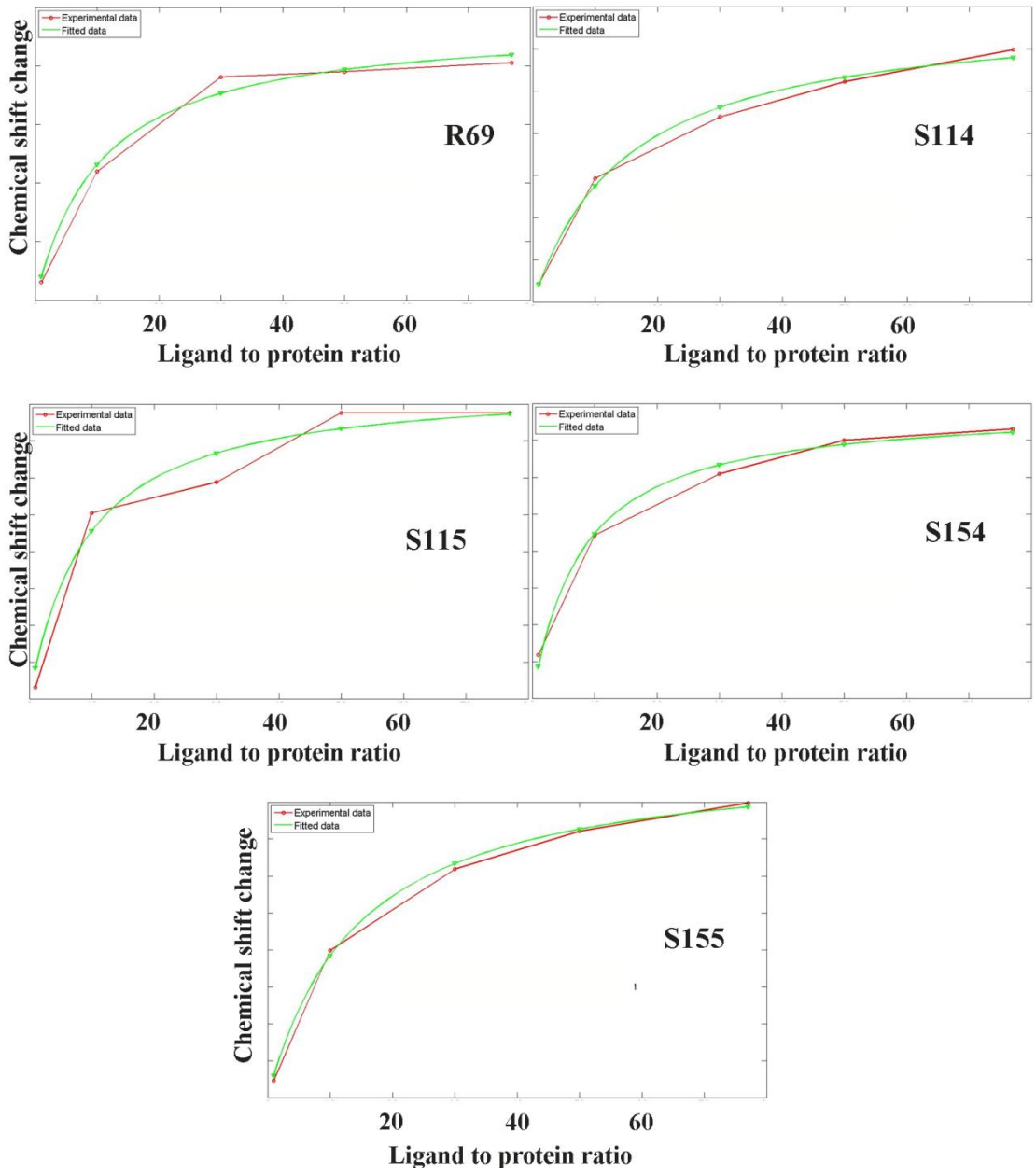


Figure 4.14: Fitting of the chemical shift changes due to titration with pCDP for the active-site residues of PPIase domain in Pin1 to determine the dissociation constant, K_D. The green lines are the fitted curve.

Effect of pCDP cyclic dipeptide on APP processing in neuroglioma cells.

The effect of pCDP on APP processing was studied in the lab of collaborator Dr. Kun Ping Lu (BIDMC, Harvard Medical School) by measuring the secreted products of APP processing, in a pCDP dose-dependent manner, in cultures of human neuroglioma H4 cells and stable Pin1 knockdown H4 cells. The assay is based on the quantification of αAPPs, the product of non-amyloidogenic processing of APP. Preliminary experiments performed by Suk Ling Ma in the Lu lab indicate that levels of αAPPs increase in human neuroglioma H4 cells after treatment with pCDP (Figure 4.15). This dose-dependent effect of pCDP was even more profound in stable Pin1 knockdown H4 cells, which is consistent with our previous results showing that more APP is processed by amyloidogenic pathway in Pin1 knockdown or knockout cells, as compared with Pin1 wild-type cells (Pastorino et al., 2006). These results indicate a shift toward the non-amyloidogenic processing pathway in a dose-dependent manner in response to treatment with the *cis* mimetic small molecule pCDP. Also, this result provides preliminary evidence justifying our hypothesis that the *cis* isomer of pAICD might favor the amyloidogenic processing of APP.

However, the high concentrations (>100 μM) required for efficacy signify that considerable improvement is required in order to acquire drug-level potency. The low potency observed for the pCDP could be either due to poor cellular uptake of the compound or to low binding affinity with its putative cellular target. The poor cellular uptake of pCDP is expected based on the anionic phosphate group. The fact that pCDP enhances the non-amyloidogenic processing despite its poor cellular uptake is very encouraging and provides impetus for the development of *cis* analogues based on this molecule. The proposed library of compounds should provide necessary improvement of the binding affinity to the cellular target(s). Moreover, identification of the actual targets would tremendously help in the rational design of inhibitors.

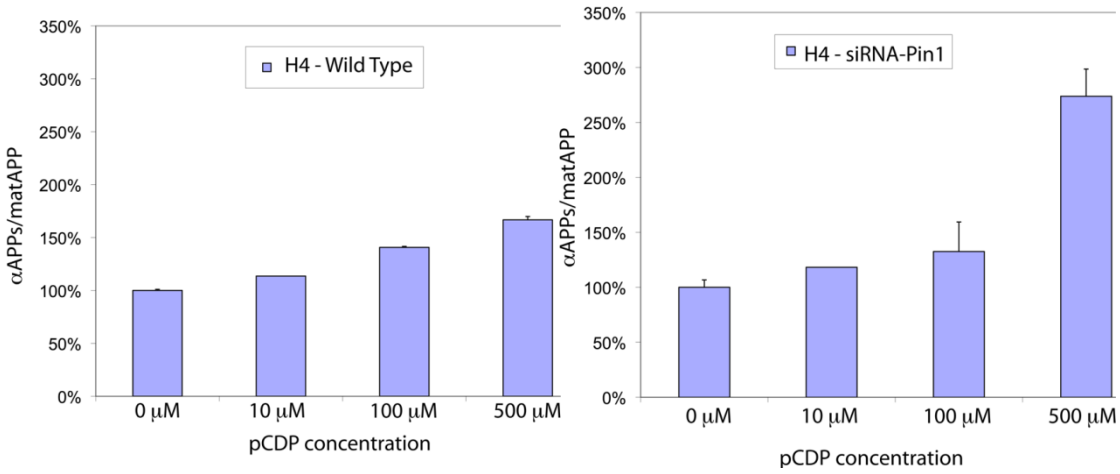


Figure 4.15: Phospho-Thr-Pro cyclic dipeptide promotes non-amyloidogenic APP processing in wild-type as well as stable Pin1 knockdown human neuroglioma H4 cells in a dose-dependent manner. Increase in αAPPs is observed for both H4-control cells expressing endogenous Pin1 and H4-Pin1-siRNA cells stably expressing the siRNA sequence for Pin1. Briefly, cells were grown in 6 well dishes, and each treatment was done in triplicate. Conditioned media were collected in presence of protease inhibitors, and centrifuged at 1400rpm for 10 min at 4C to pellet cell debris and avoid cross contamination with full length APP. Supernatants were transferred to a fresh tube and αAPPs was immunoprecipiated using 6E10 antibody (Covance). After elution from the beads, samples were run onto 6%SDS-PAGE, and blotted onto PVDF. Detection of both αAPPs and total APP was performed using the monoclonal antibody 22C11 by means of western blot (Chemicon). Total cell lysates were used to probe for total APP and Pin1 expression. These experiments were performed by Suk Ling Ma in Dr. Kun Ping Lu's laboratory in Harvard Medical School, Boston, MA. This figure was supplied by Suk Ling Ma.

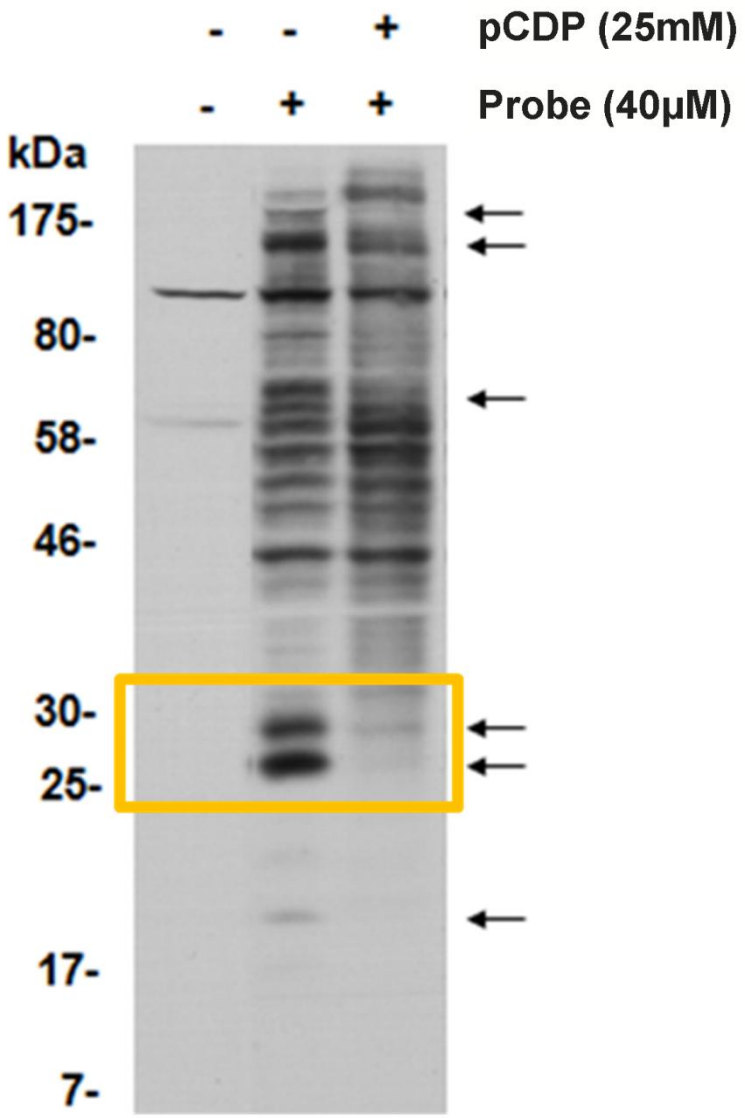


Figure 4.16: Western blot of the whole cell lysate visualized by enhanced chemiluminiscence. The protein bands that are different between lanes 2 and three are indicated by arrows. The two most prominent bands between 25 to 30 KDa are enclosed in a box. These 'fishing' experiments were performed by Ross Resnick, research support specialist, of the Nicholosn lab. This figure was supplied by Ross Resnick.

Identification of isomer-specific binding partners

In order to identify *cis*-specific binding partners, "fishing" experiments utilizing the above described probes and inhibitors are being pursued by Ross Resnick, research support specialist, in the Nicholson lab. As an example, pCDP was used as inhibitor, and a probe containing the natural sequence VpTPEER was used as the 'bait'. Whole cell lysates (WCL) were prepared from $Pin1^{-/-}$ mouse breast cancer cells and were preincubated with pCDP. The probe was added to the lysate, incubated and then irradiated by UV light. Proteins were resolved on 8-20% reducing/denaturing SDS-PAGE gel transferred to PVDF membrane and then probed with streptavidin-HRP (horseradish peroxidase) and the results were visualized by enhanced chemiluminescence (Figure 4.16). Lane 1 is a negative control where probe was not added to the whole cell lysate. The two bands visible in this lane are due to non-specific binding of streptavidin or HRP to proteins in these two bands. For all other bands in the next two lanes, a covalent linkage of the proteins to the peptide through Bpa is necessary for visualization. Lane 2 is without pCDP as the inhibitor. The different bands consist of proteins which recognize *cis* isomer, *trans* isomer or other region of the probe. Lane 3 is with pCDP as an inhibitor. Certain bands in these lanes can be seen missing when compared to lane 2 (noted by arrows). These bands most probably correspond to proteins that bind to pCDP. The next steps are the identification of the proteins in these bands and the validation of their binding and isomer-specificity to pAICD by NMR

CHAPTER 5

MODELLING THE PH-FOLDING SWITCH IN THE PSEUDOMONAS SYRINGAE EFFECTOR PROTEIN AVRPTO USING YANG-HONIG MODEL

Introduction

The Type Three Secretion System (T3SS) is a needle shaped nano-device used by gram-negative bacteria to infect host plant cells (Figure 5.1a). The pathogenic bacteria utilize the T3SS to inject effector proteins into the host cells (Galan and Collmer, 1999). These effector proteins suppress the basal and hypersensitive response (HR) defenses of plant cells, thus, enabling the pathogen to grow in the plant apoplast. The basal defense involves the formation of callose rich papilla below the cell wall, synthesis of reactive oxygen species (peroxide H_2O_2 and superoxide O_2^-) (Dangl and Jones, 2001) and upregulation of defense associated genes. The HR is much more specific and complicated. It is triggered by certain phytopathogen effectors resulting in localized cell death and increased defense of surrounding plant tissues. Hence, the efficient transport of the effector proteins through the T3SS into the host cell, in order to suppress these defense mechanisms, is one of the key steps for the virulent activity of the pathogens.

The T3SS is evolutionarily related to bacterial flagellum (Saier, 2004). It is composed of ring structures associated with the inner and outer membranes of the bacteria and a needle like structure called pilus reaching into the plant cell (Figure 5.1a**)** (Cornelis, 2006). A dodecameric ATPase is attached to the base of the ring structure and is postulated to take part in unfolding and translocation of the effector

---

This chapter is based on manuscript published courtesy of *Proceeding of the National Academy of Sciences of the United States of America* (Dawson et al., 2009).

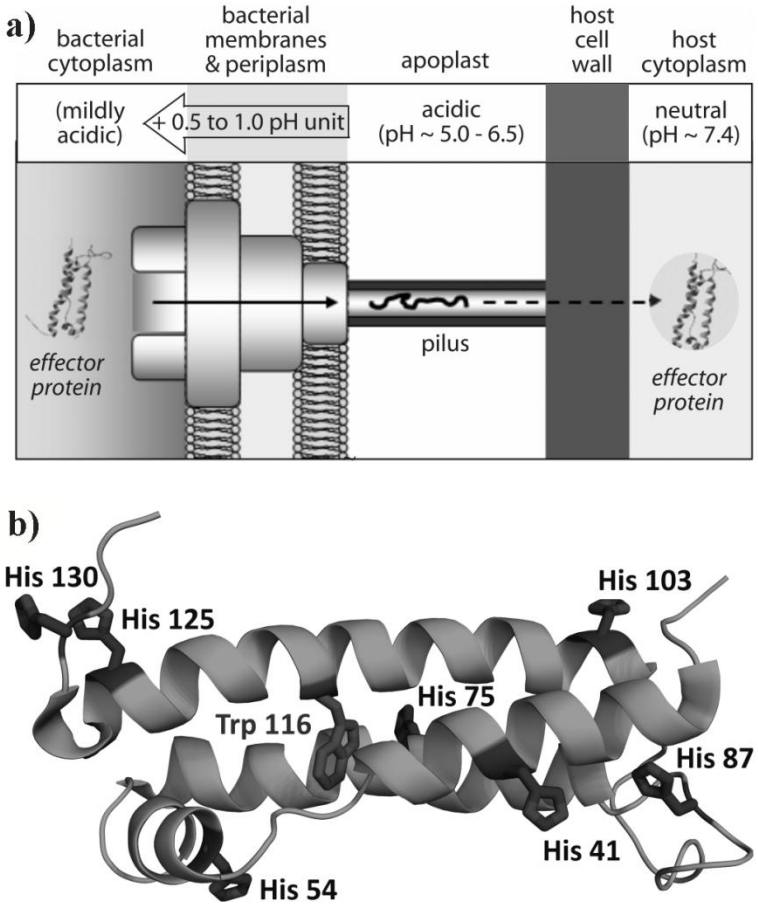


Figure 5.1: a) The T3SS of pathogenic bacteria spans the inner and outer membranes of the bacteria and the host cell wall. The effector proteins are in the mildly acidic bacterial cytoplasm which assists in their unfolding and translocation through the narrow pilus of the T3SS into the host cytoplasm. The host cytoplasm is at neutral pH which assists in the refolding of the effector proteins and subsequent signaling mechanisms leading to triggering or suppression of the hypersensitive response (HR). b) The TrAvrPto structure with the seven histidine residues and the lone tryptophan residue dispersed across it. This figure is reproduced with permission from Dawson et. al., 2009.

proteins (Galan and Wolf-Watz, 2006). The inner diameter of the pilus typically ranges from 2 to 3 nm (Cornelis, 2006). This presents a unique problem for the translocation of the effector proteins through the T3SS. Most of the effector proteins have a minimum structural dimension greater than 2 nm which makes it difficult to explain how the folded proteins get through the narrow pilus. In order to explain this

discrepancy it was hypothesized that the proteins have to partially unfold before getting into the T3SS (Feldman et al., 2002). The elegant work of Lee (Lee and Schneewind, 2002) and Feldman (Feldman et al., 2002) provided initial supporting evidence for this hypothesis. Feldman and co-workers showed that the N-terminal secretion signal (residues 1-16) of the effector protein YopE is unable to result in the secretion of the stable wild type dihydrofolate reductase (WT DHFR) while the same signal can affect the secretion of a destabilized mutant of DHFR. Lee and Schneewind showed similar results for ubiquitin. Furthermore, Feldman and co-workers proposed that the chaperones associated with T3SS are responsible for keeping their substrate in a partially folded state in order to facilitate secretion. Birtalan *et. al.* (Birtalan et al., 2002) found that the chaperones act locally by binding to the chaperone binding domain of the effector proteins and do not contribute to the global unfolding of the proteins. Also, not all effector proteins have an associated chaperone and therefore, are translocated in a chaperone independent manner (Buttner and Bonas, 2006). This necessitates the postulation of an alternative hypothesis that would explain the ability of effector proteins to unfold unassisted, translocate through the T3SS and refold once in the host cytoplasm.

Pseudomonas syringae pv. *tomato* is one of the most well-studied plant pathogens which causes bacterial speck disease in tomato (Mudgett, 2005). One of its effector proteins, AvrPto, has been studied as a model system for chaperone independent translocation through the T3SS (Alfano and Collmer, 2004). AvrPto acts as a virulence factor in non-resistant plants and thus, promotes the bacterial growth. In resistant plants it binds to the host Pto kinase and triggers its defense system (Scofield et al., 1996). The NMR structure of a truncated version of AvrPto (TrAvrPto) has been solved (PDB ID 1R5E) (Wulf et al., 2004), which reveals a three-helix bundle and a long Ω-loop (Figure 5.1b). TrAvrPto has been shown to retain

similar Pto-binding as AvrPto and TrAvrPto also elicits Pto-mediated immunity in resistant plants (Wulf et al., 2004). Detailed kinetic and thermodynamic studies on TrAvrPto revealed the inherent instability of the protein (Dawson and Nicholson, 2008). At pH = 6.1 the protein has 16% unfolded population and undergoes slow exchange between the folded and unfolded states. Studies at pH 6.1 and 7.0 showed that the TrAvrPto stability is highly dependent on pH. Thus, the folded core of TrAvrPto possesses a pH-regulated folding switch which can be utilized by the plant pathogens to translocate the effector protein.

It is essential for bacteria to adapt rapidly to their environment in order to grow. Plant pathogens such as *P. syringae* invade the host cell from the plant apoplast which is mildly acidic (pH 5.0 to 6.5 (Grignon and Sentenac, 1991)). There the bacteria assemble the T3SS which connects its interior to the host cell cytoplasm. The cytoplasmic pH of the bacteria has to be 0.5 – 1.0 pH unit higher than the apoplast as the T3SS requires a positive proton gradient (Minamino et al., 2003; Wilharm et al., 2004). The plant cell cytoplasm typically has a neutral pH (~7.4). Thus, a pH-gradient is present on the two extremes of the T3SS from mildly acidic (bacterium) to neutral (host cell), and nature has elegantly utilized this to translocate effector proteins by tuning their folding in this pH range (Figure 5.1a).

Materials and Methods

The circular dichroism (CD) and Trp fluorescence (Fl) thermal denaturation experiments were performed (Dawson et al., 2009) at different pH values to get the pH-dependence of the global folded population (p_F) of the WT TrAvrPto. The pH-dependence of p_F for the wild type protein was also obtained from the ^{15}N-^{1}H HSQC experiments at 26.6 °C. Since the pH-dependence of p_F obtained from the three different experiments report on the same physical phenomenon, the combined data

was fitted simultaneously to the two-state multi-titration model based on the Yang and Honig model (Yang and Honig, 1993). For the mutant H87Y, Fl data was collected at 25 °C to get the corresponding pH-dependence and was fitted to the same model (Dawson et al., 2009).

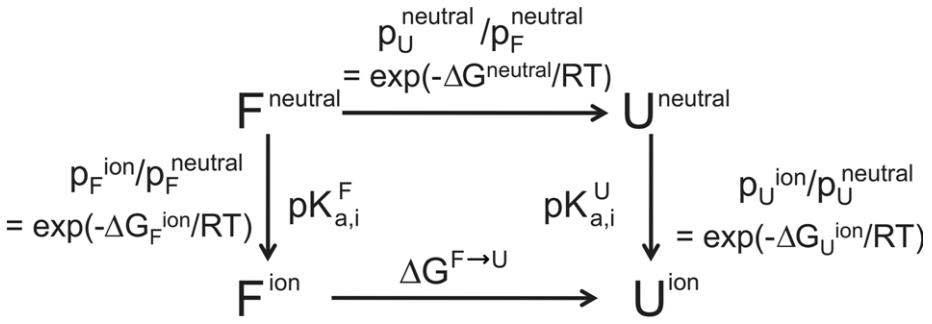


Figure 5.2: The multi-titration two-state Yang and Honig model. The denaturation free energy $\Delta G^{F \rightarrow U}$ is separated into pH-dependent ($\Delta\Delta G^{ion} = \Delta G_U^{ion} - \Delta G_F^{ion}$) and pH-independent ($\Delta G^{neutral}$) components. The populations of the F and U states are related to the pK_a of individual ionizable residues (i) in each state j via ΔG_j^{ion}. This figure is reproduced with permission from Dawson et. al., 2009.

The Yang and Honig model expresses p_F as a function of pH with the pK_a's of various ionizable groups in the unfolded (U) and folded (F) state of the protein and f, the ratio of neutral U and F, as model-parameters. This is achieved by separating the denaturation free energy ($\Delta G^{F \rightarrow U} = \Delta G^{neutral} + \Delta\Delta G^{ion}$) into pH-dependent ($\Delta\Delta G^{ion}$) and pH-independent ($\Delta G^{neutral} = -RT\ln(f)$) components, where $\Delta G^{neutral}$ is the denaturation free energy of the protein when all its ionizable groups are in the neutral state, and $\Delta\Delta G^{ion} = \Delta G_U^{ion} - \Delta G_F^{ion}$, where ΔG_j^{ion} (j = U or F) is the difference in free energy between the ionized and neutral forms of state j. Neglecting interactions between charged groups in each state, ΔG_j^{ion} is given by (Yang and Honig, 1993):

$$\Delta G_j^{ion} = -RT \sum_{i=1}^{N} \ln\left\{1 + \exp\left(-2.3\gamma_i\left(pH - pK_{a,i}^j\right)\right)\right\} \quad (5.1)$$

where N is the number of ionizable groups, γ_i is -1 or +1 for an acidic or basic group, respectively, and $pK_{a,i}^j$ is the pK_a of the i^{th} ionizable group in state j. Since ΔG_j^{ion}

depends only on the $pK_{a,i}^{j}$ values and pH, the above model expresses the pH-dependence of $\Delta G^{F\rightarrow U}$ (and hence, of p_F) for a given set of $pK_{a,i}^{j}$ values (Figure 5.2). It is worth mentioning that $\Delta\Delta G^{ion}$ depends on differences between pK_a values in the U and F states. Titrating groups with elevated or depressed pK_a^F relative to pK_a^U will dominate the pH-dependence of stability.

In order to determine the protein stability at each pH, populations were calculated from f ($f = p_U^{neutral} \Big/ p_F^{neutral}$), ΔG_U^{ion} and ΔG_F^{ion} (Figure 5.2). The total folded population as observed by the various physical techniques was calculated as $p_F^{obs} = p_F^{neutral} + p_F^{ion}$, the sum of neutral and ionized folded populations. The $pK_{a,His}^F$ for all the seven histidine sidechains were determined by NMR titration experiments. Hence, these values were not used as fitted parameters. For the unfolded protein $pK_{a,His}^U = 6.4$, typical for solvent exposed histidine sidechains (Nozaki and Tanford, 1967), was used for all histidines. Thus, the only fitted parameters were $pK_{a,COOH}^U$, $pK_{a,COOH}^F$ and f. The H87Y p_F vs. pH experimental data was fitted by omitting the contribution of H87 from the calculation of $\Delta\Delta G^{ion}$, but otherwise using the same pK_a values as for the wild type best-fit and varying f to optimize the fit.

Implementation of data-fitting by non-linear least-squares regression

Regression analysis is one of the most common methods of analyzing numerical data consisting of a set of dependent ($\mathbf{Y^{obs}}$) and independent variables ($\mathbf{X}$). In order to analyze the data, a model needs to be devised such that the dependent variable is defined as a function of the independent variables, equation (5.2). The function also consists of a set of parameters ($\mathbf{a}$) and an error term ($\boldsymbol{\varepsilon}$). The error term accounts for the deviation of the dependent variables from the model. The defined function, also called the regression equation, can be linear or non-linear in terms of the

model parameters. Hence, for a given dataset ($\mathbf{Y^{obs}}$, $\mathbf{X}$), a regression equation can be defined as follows:

$$\mathbf{Y^{calc}} = f(\mathbf{X}, \mathbf{a}) + \boldsymbol{\varepsilon} \quad (5.2)$$

where $\mathbf{Y^{calc}}$, $\mathbf{X}$, $\mathbf{a}$ and $\boldsymbol{\varepsilon}$ are the set of calculated dependent variables, independent variables, model parameters and errors respectively. Once the model is defined, the next step is to calculate the Weighted Sum of Squared Errors (WSSE) given by:

$$\text{WSSE} = \sum(\boldsymbol{\epsilon}/\boldsymbol{\sigma})^2 \quad (5.3)$$

where $\boldsymbol{\sigma}$ is the set of experimental errors associated to the dataset ($\mathbf{Y^{obs}}$, $\mathbf{X}$). The errors $\boldsymbol{\sigma}$ are normally determined from repeated measurements or by propagation of errors. From equation (5.3), it can be seen that datapoints (Y^{obs}, X) with larger error (σ) would contribute less to the WSSE that those with smaller errors. Hence, the datapoints with smaller errors would influence the fitting more than those with the larger errors. In the absence of experimental errors, a constant error (σ) is used, thus, giving equal weight to all the datapoints.

The best-fit parameters of the model (equation (5.2)) can be determined by minimizing the WSSE. This is easily done by using standard equations (Press, 1988) in the case of linear models. Non-linear models require an iterative approach involving the following steps:

1. An initial value needs to be guessed for the set of parameters and is given by $\mathbf{a_0}$.
2. Using $\mathbf{a_0}$ and $\mathbf{X}$, the set of dependent variables ($\mathbf{Y^{calc}}$) are calculated from the model such that $\mathbf{Y^{calc}} = f(\mathbf{X}, \mathbf{a_0})$.
3. The WSSE is determined, where $\text{WSSE} = \sum\left(\left(\mathbf{Y^{obs}} - \mathbf{Y^{calc}}\right)/\boldsymbol{\sigma}\right)^2$.
4. Again a new set of parameters $\mathbf{a_1}$ is guessed and WSSE calculated. The new parameter set $\mathbf{a_1}$ is accepted if $\text{WSSE}_{new} < \text{WSSE}_{old}$. There are several algorithms for adjusting the parameters, e.g. Gauss-Newton method, method of

steepest decent and Levenberg-Marquardt method. The Levenberg-Marquadt (Marquardt, 1963) method implements the best of the other two methods and is the one mostly used.

5. Step 4 is repeated until the change in WSSE is less than a tolerance value or a maximum number of iterations are reached. The parameter set ($\mathbf{a}_{\mathbf{bestfit}}$) for the best-fit result is reported along with the corresponding WSSE.
6. The precise value of $\mathbf{a}_{\mathbf{bestfit}}$ depends partially on the initial guess $\mathbf{a}_{\mathbf{0}}$, and hence, the iterations should be performed with at least three different initial guesses in order to prevent fitting artifacts.

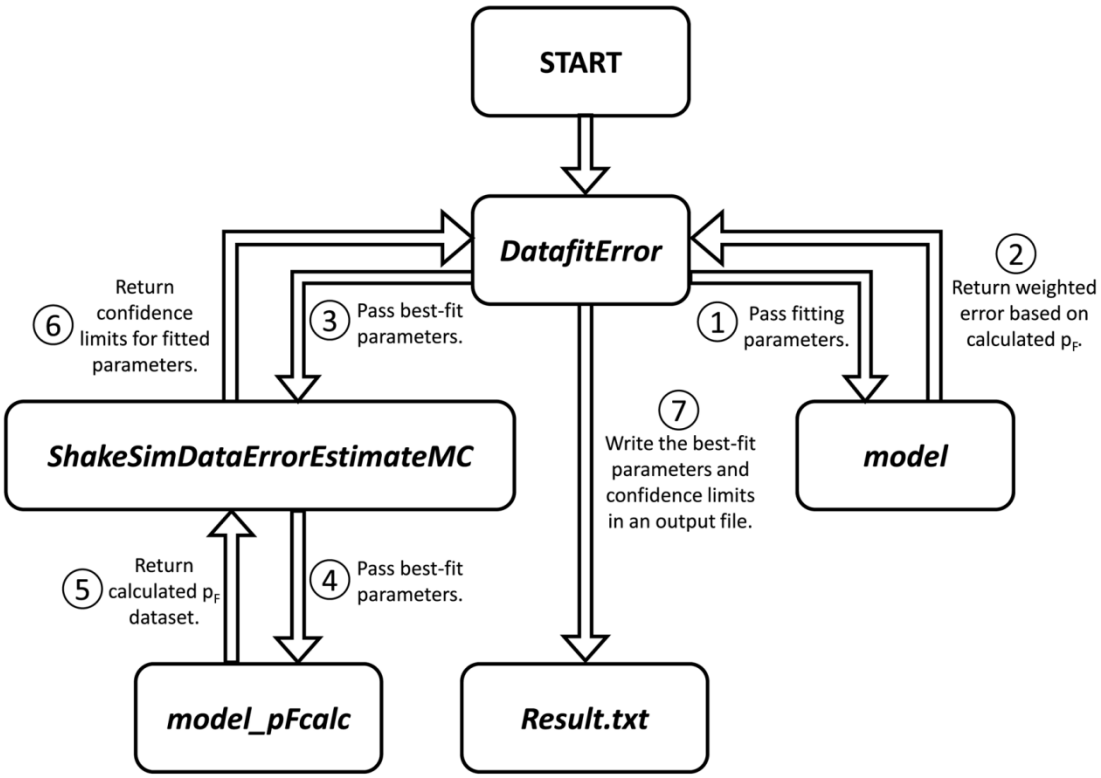


Figure 5.3: A flow chart depicting the implementation of the data fitting program in MATLAB. The start box indicates the program run by the user. All other boxes except Result.txt are MATLAB mechanism files and Result.txt is the final output file containing the results of the data fitting.

Determination of the best-fit parameters is only the first part of data-fitting to a model. It is also necessary to estimate the confidence limits of the parameters being fitted by regression. Again, for linear regression standard equations are available for the determination of the confidence limits (Press, 1988). For non-linear regression, the confidence limits can be estimated by Monte Carlo Simulation. This is a rigorous and computationally intensive method. Fitting the experimental dataset to the theoretical model gives us the best-fit parameters $\mathbf{a}_{\text{bestfit}}$ for that dataset. Varying the experimental dataset by adding random error to each point and fitting it to the same model would generate another set of fitted parameters $\mathbf{a}_i$. Thus, the variance of the fitted parameters **a** can be determined by generating a set of synthetic dataset and calculating the distribution of $\mathbf{a}_i$. Here it is assumed that the distribution of $\mathbf{a}_i$ about $\mathbf{a}_{\text{bestfit}}$ is the same as that about $\mathbf{a}_{\text{actual}}$, which is the set of actual values for the parameters **a** (Press, 1988).

The nonlinear least-squares regression routine to fit the Yang and Honig model (Yang and Honig, 1993) to the experimental data was implemented in MATLAB® 2007a. The routine *DatafitError.m* (Figure 5.3) implements the code for fitting of the experimental data by non-linear regression. The data fitting is started by running *DatafitError.m*, which passes the initial guess for the fitting parameters ($pK^{U}_{a,COOH}$, $pK^{F}_{a,COOH}$ and f) along with the fitting model to the MATLAB function *lsqnonlin*. The fitting model is defined in the routine *model.m*, which accepts the fitting parameters, calculates the folded populations ($\mathbf{p_F^{calc}}$) and returns an array containing the weighted error ($\left|\left(\mathbf{p_F^{expt}} - \mathbf{p_F^{calc}}\right)/\boldsymbol{\sigma}\right|$) at each pH. The function *lsqnonlin* minimizes the WSSE ($\sum\left(\left(\mathbf{p_F^{expt}} - \mathbf{p_F^{calc}}\right)/\boldsymbol{\sigma}\right)^2$) by iteratively changing the fitting parameters and returns the best-fit parameters. The Levenberg-Marquadt algorithm (Marquardt, 1963) is used to search for the best-fit parameters. The minimization routine can be optimized by

changing some of the options of *lsqnonlin*, e.g. maximum iterations and maximum function evaluations.

In order to estimate the confidence limits, the best-fit parameters are passed into the routine *ShakeSimDataErrorEstimateMC.m*, which in turn utilizes the routine *model_pFcalc.m* to calculate the populations ($\mathbf{p_F^{calc}}$) using the best-fit parameters and the fitting model. This routine calculates the standard deviation ($\sigma_{\text{fitting}} = \frac{1}{N-1}\sum\left(\mathbf{p_F^{expt}} - \mathbf{p_F^{calc}}\right)^2$) of the best-fit curve ($\mathbf{p_F^{calc}}$ vs. **pH**) from the experimental dataset ($\mathbf{p_F^{obs}}$ vs. **pH**) which is then used to generate simulated datasets with normally distributed random noise about the best-fit curve. Each simulated dataset i is fit to the same model as done previously for the experimental dataset yielding best-fit parameters $\mathbf{a_i}$. The standard deviation of the best-fit parameters $\mathbf{a_i}$ over all of the simulated datasets is determined and reported as the confidence limit for the fitted parameters $\boldsymbol{\sigma}_{\text{bestfit}}$. For a good estimate of the confidence limit, a minimum of 100 simulated datasets were used.

Table 5.1: The sidechain pK_a^F for histidine residues as determined by ^{15}N-^{1}H HMQC experiments (Dawson et al., 2009)

Residues	**H41**	**H54**	**H75**	**H87**	**H103**	**H125**	**H130**
$\mathbf{pK_a^F}$	6.32 ± 0.04	5.88 ± 0.01	6.89 ± 0.01	4.8 ± 0.1	6.28 ± 0.02	6.65 ± 0.01	6.28 ± 0.03

Results

In order to elucidate the molecular mechanism employed by nature to tune the stability of TrAvrPto in the correct pH range, its acid denaturation was extensively studied by NMR, circular dichroism (CD) and Trp fluorescence (Fl) spectroscopies (Dawson et al., 2009). Since AvrPto experiences a pH range of ~ 6 to 7.5 across the T3SS, it is necessary to identify residues that titrate in this pH range. The sidechain of

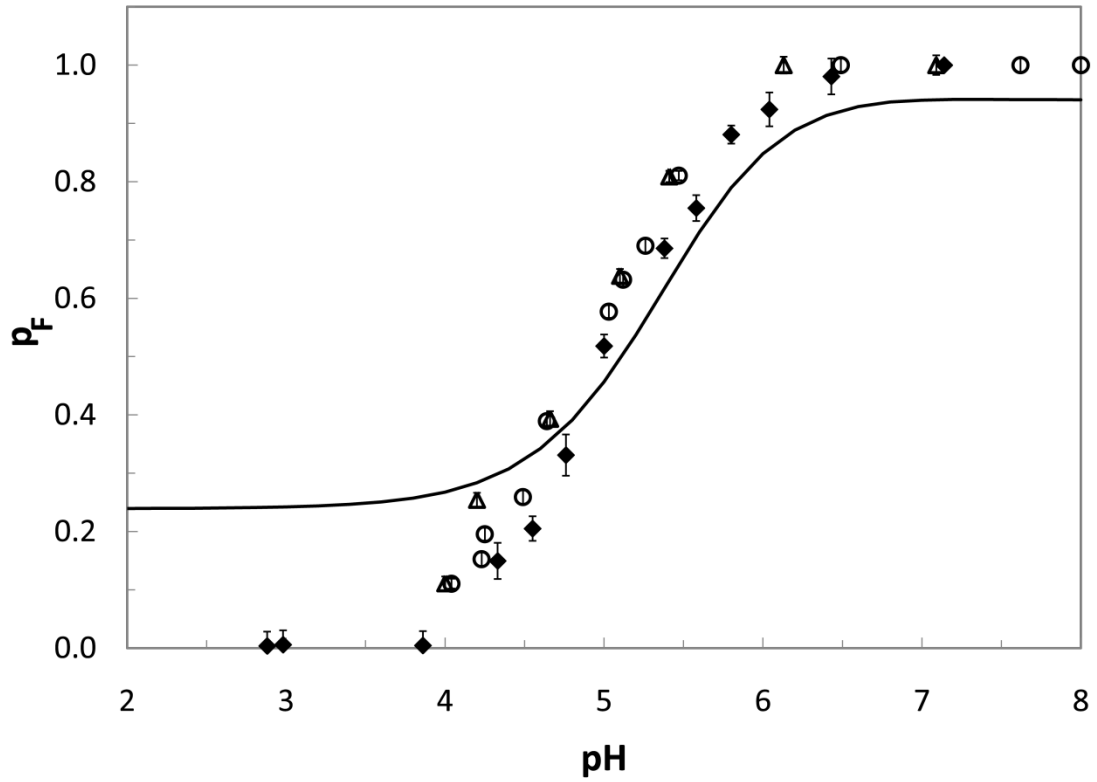


Figure 5.4: Folded populations (p_F) of WT TrAvrPto versus pH at 26.6 °C from NMR (♦), CD (○) and Fl (Δ) experiments. The solid line is the fitted curve to the Yang and Honig model considering the histidine sidechain titrations only. This model fails to predict the p_F at pH values less than 5.

free histidine residue has a pK_a of ~6.4 (Nozaki and Tanford, 1967) and hence, is the ideal candidate to provide the necessary pH-regulated folding. ^{15}N-^{1}H HMQC experiments were performed (Dawson et al., 2009) to determine the pK_a of all the seven histidine sidechains in folded TrAvrPto (Table 5.1). Histidines H41, H103 and H130 display pK_a values consistent with their solvent-exposed locations on the protein surface. Histidines H75 and H125 have elevated and H54 has depressed pK_a values. The remaining histidine H87 displays a dramatically low pK_a value. ^{15}N-^{1}H fHSQC experiments were performed to obtain the folded (p_F) and unfolded (p_U) populations at various pH and 26.6 °C (Dawson et al., 2009). The CD and Fl thermal denaturation experiments provided information on pH as well as temperature dependence of

TrAvrPto folding (Dawson et al., 2009). The data (Figure 5.4 and 5.5) from three different biophysical methods agreed well to each other and conclusively proved the role of pH as a folding switch for TrAvrPto. In order to elucidate the role of histidines as the tuners of this folding switch in the required pH-range, the model proposed by Yang and Honig (Yang and Honig, 1993) was employed.

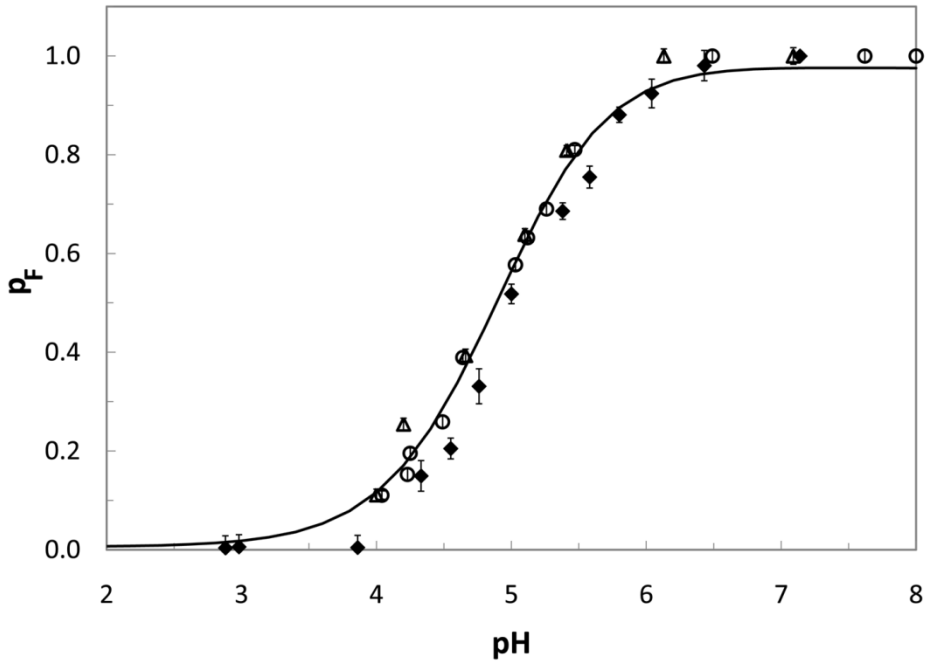


Figure 5.5: Folded populations (p_F) of WT TrAvrPto versus pH at 26.6 °C from NMR (♦), CD (○) and Fl (Δ) experiments. The solid line is the fitted curve to the Yang and Honig model considering the histidine sidechain and COOH titrations. This model nicely predicts the p_F at all pH values.

Yang and Honig (Yang and Honig, 1993) proposed a model for the pH-dependence of protein stability (Figure 5.2). The basis of this model is that the titratable groups in a protein should have different pK_a values in the folded (F) and unfolded (U) states of the protein. In the unfolded state, the titratable group in each type of amino acid should have the same pK_a value. Thus, all histidines should have a pK_a^U of ~6.4, glutamates a pK_a^U of ~ 4.8 and so on. On the other hand, in the folded

state, each residue would have a different pK_a^F value depending on whether it is surface exposed, buried, in a salt bridge or forming a hydrogen bond. For example, H87, which is a buried residue in TrAvrPto, has a pK_a^F of 4.8 in the folded state and pK_a^U of 6.4 in the unfolded state. Yang and Honig not only provided a sound physical model for the pH-dependent stability of proteins based on pK_a values but also were able to provide rigorous mathematical equations relating the folding free energy ($\Delta G^{F\rightarrow U}$) to the pK_a^F and pK_a^U of the titratable residues in the protein, pH of the system and $\Delta G^{neutral}$, the denaturation free energy of the protein when all its titratable residues are in the neutral state. The details of the model have been discussed in the Methods section.

The pH-dependence of the folded population (p_F) obtained from the CD, Fl and NMR experiments were fit to the Yang and Honig model (Figure 5.2). The titrations of all the seven histidines were considered simultaneously. The pK_a^F values obtained from the ^{15}N-^{1}H HMQC experiments (Table 5.1) and pK_a^U = 6.4 (Nozaki and Tanford, 1967) were used in the model. The only fitting parameter, $\Delta G^{neutral}$, was obtained to be 1.65 ± 0.03 kcal/mol. This model was able to fit the high pH titration points from pH 5 to 8 but was inadequate to predict the lower titration points (Figure 5.4). This implies that the histidines which have their pK_as in the range of 4.8 to 6.8 are not responsible for the titration of the protein in the lower pH range. This is consistent with the Yang and Honig model (Yang and Honig, 1993). While discussing the model, the authors have explained the unfolding transition of a protein in a pH-range as follows: "… the pK_a of the responsible group does not correspond to the midpoint of the titration as is sometimes assumed. Rather, there must be a group whose pK_a is shifted from one value that corresponds to the beginning of the titration to another value that corresponds to its endpoint. Thus, analyzing stability curves is

fundamentally different from analyzing standard titration curves since the endpoints rather than the midpoints serve as markers for the relevant pK_as."

Since the pK_as of the histidines in TrAvrPto range from 4.8 to 6.8, they tune the unfolding transition of the protein in that pH-range. For the titration points from pH 2.0 to 4.8, the responsible residues should have their relevant pK_as in the same range. The sidechains pK_as (Nozaki and Tanford, 1967) of aspartic acid (pK_a^U = 3.9) and especially glutamic acid (pK_a^U = 4.4) will be in the required range if their pK_a^F is close to 2.0. Hence, the model was modified to include two more fitting parameters $pK_{a,COOH}^F$ and $pK_{a,COOH}^U$ which are the pK_as for the required sidechain COOH group in the folded and unfolded protein respectively. For this model $\Delta G^{neutral}$ = -0.74 ± 0.5 kcal/mol, $pK_{a,COOH}^F$ = 2.7 ± 0.4 and $pK_{a,COOH}^U$ = 4.8 ± 0.05 was obtained and the model fits the experimental data well (Figure 5.5). A favourable $\Delta G^{neutral}$ indicates that the unfolded form is favored in the hypothetical neutral state (Figure 5.2) and hence, a contribution of favorable free energy would be required from the titratable residues in order to keep the protein folded in the neutral pH range. The contribution of each titratable residue to the stability of the protein is given by $\Delta\Delta G^{ion}$ (see methods and Figure 5.2). Calculating $\Delta\Delta G^{ion}$ for each residue as a function of pH (Figure 5.5) reveals that the COOH group and the His 87 are the major contributors to the protein's stability and destability, respectively. Furthermore, the $\Delta\Delta G^{ion}$ contribution of the COOH group (2.9 kcal/mol at pH 6.5 and above) is too high for a single COOH involved in solvent-exposed salt bridge (Horovitz et al., 1990) and suggests that multiple COOH groups are involved in stabilization of the folded protein. Thus, this modified model successfully explains the pH-dependent folding of TrAvrPto. It also identifies the H87 residue as the key regulator of the pH-dependent folding switch and several sidechain COOH groups, involved in salt-bridges, providing the basic structural stability to the protein.

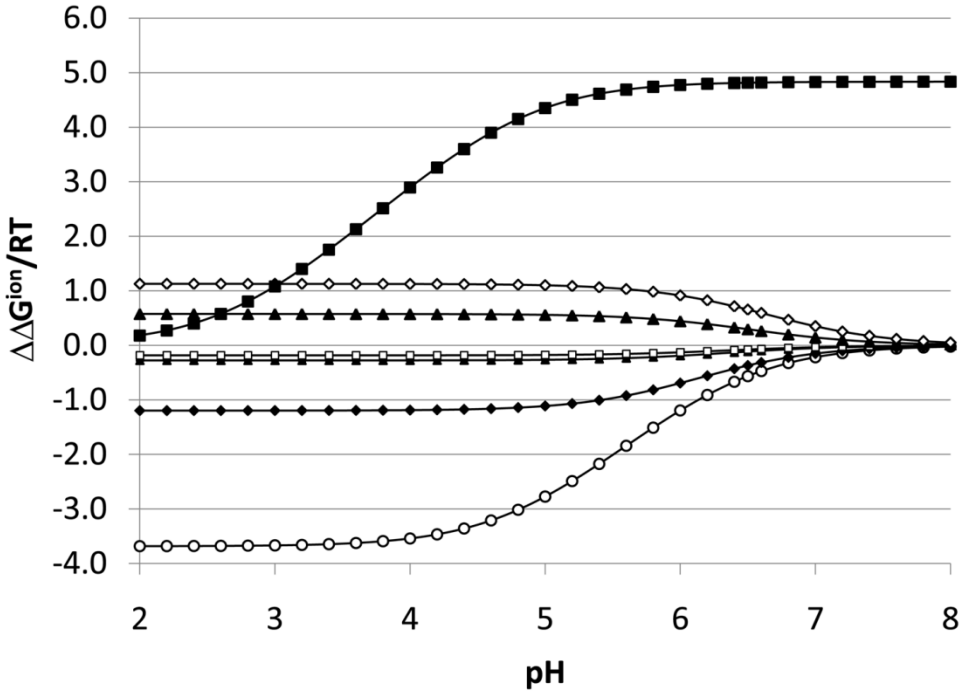


Figure 5.6: Corresponding contribution (based on equation 6.2) of each His residue (H41(□), H54(♦), H75(◊), H87(○), H103(Δ), H125(▲), H130(●)) and COOH(■) group to the $\Delta\Delta G^{ion}$ of the protein as a function of pH. This figure is adapted with permission from Dawson et. al., 2009.

The contribution of the H87 residue to the pH-dependent folding switch was tested by H87Y-TrAvrPto mutant (Dawson et al., 2009). Fl data of the pH-titration of the H87Y mutant shows a significantly lower acid denaturation (Figure 5.7). This data was modeled using the same pK_a values of all histidine residues (excluding H87), $pK^{F}_{a,COOH}$ and $pK^{U}_{a,COOH}$ of the WT TrAvrPto. The $\Delta G^{neutral}$ was allowed to vary and the model fits the H87Y fluorescence data well (Figure 5.7). The fitted $\Delta G^{neutral}$ (-0.68 ± 0.01 kcal/mol) is less than that of the wildtype protein (-0.74 kcal/mol) and indicates that the H87Y mutation has shifted the $F^{neutral} \rightarrow U^{neutral}$ equilibrium toward $F^{neutral}$. These results remarkably support the two-state multi-titration model (Figure 5.2) and confirm the central role of H87 in tuning the pH-dependent stability of TrAvrPto in the pH range of 5 to 7.

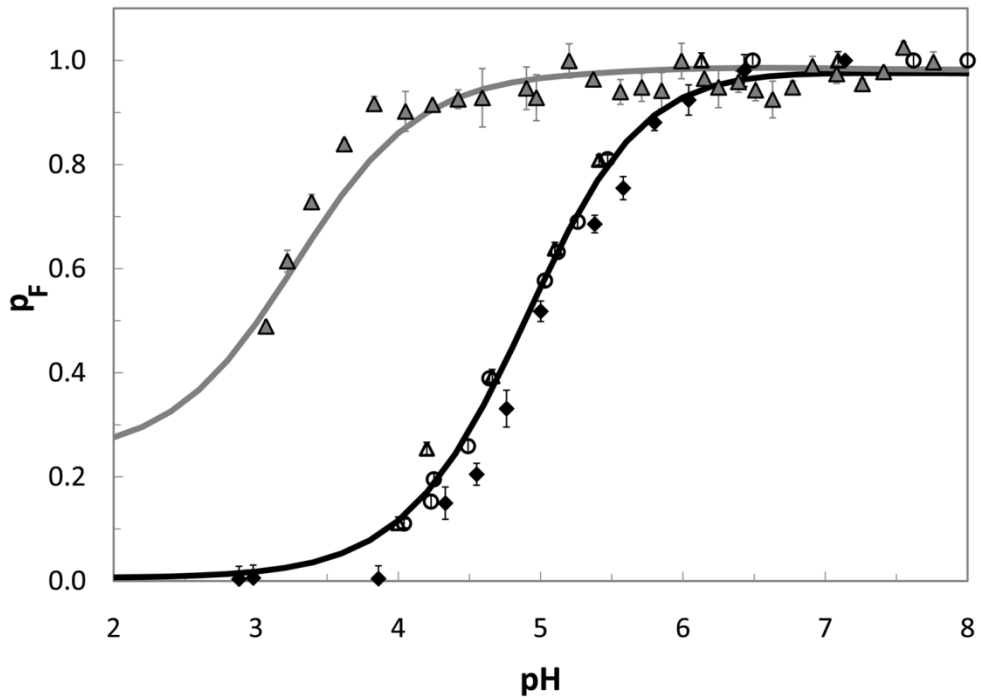


Figure 5.7: Folded populations (p_F) versus pH at 26.6 °C for the WT from NMR (♦), CD (○) Fl (△) experiments and for H87Y from Fl (▲) experiment. The solid grey line is the fitted curve to the Yang and Honig model considering the histidine sidechain (except H87) and COOH titrations. The solid black line is for the WT protein. The H87Y mutant shows markedly increased pH-dependent stability as compared to the WT. This figure is adapted with permission from Dawson et. al., 2009.

Discussion

The unfolding requirement of effector proteins in order to translocate through the T3SS is a widely debated topic in the literature. As discussed in several reviews (Galan and Wolf-Watz, 2006; Ghosh, 2004) of the T3SS, the results from various biochemical and in vivo experiments, done to prove or disprove this unfolding requirement, are not conclusive. Here, we have shown that a pH-regulated folding switch operates in AvrPto that would facilitate its unfolding in the bacterial cytoplasm and refolding in the host cytoplasm (Figure 5.1). A single histidine residue, H87, acts as a pH sensor and tunes the acid denaturation of the protein in the biologically

significant pH range of 6.0 to 7.5. The Yang and Honig model (Figure 5.2) elegantly explains

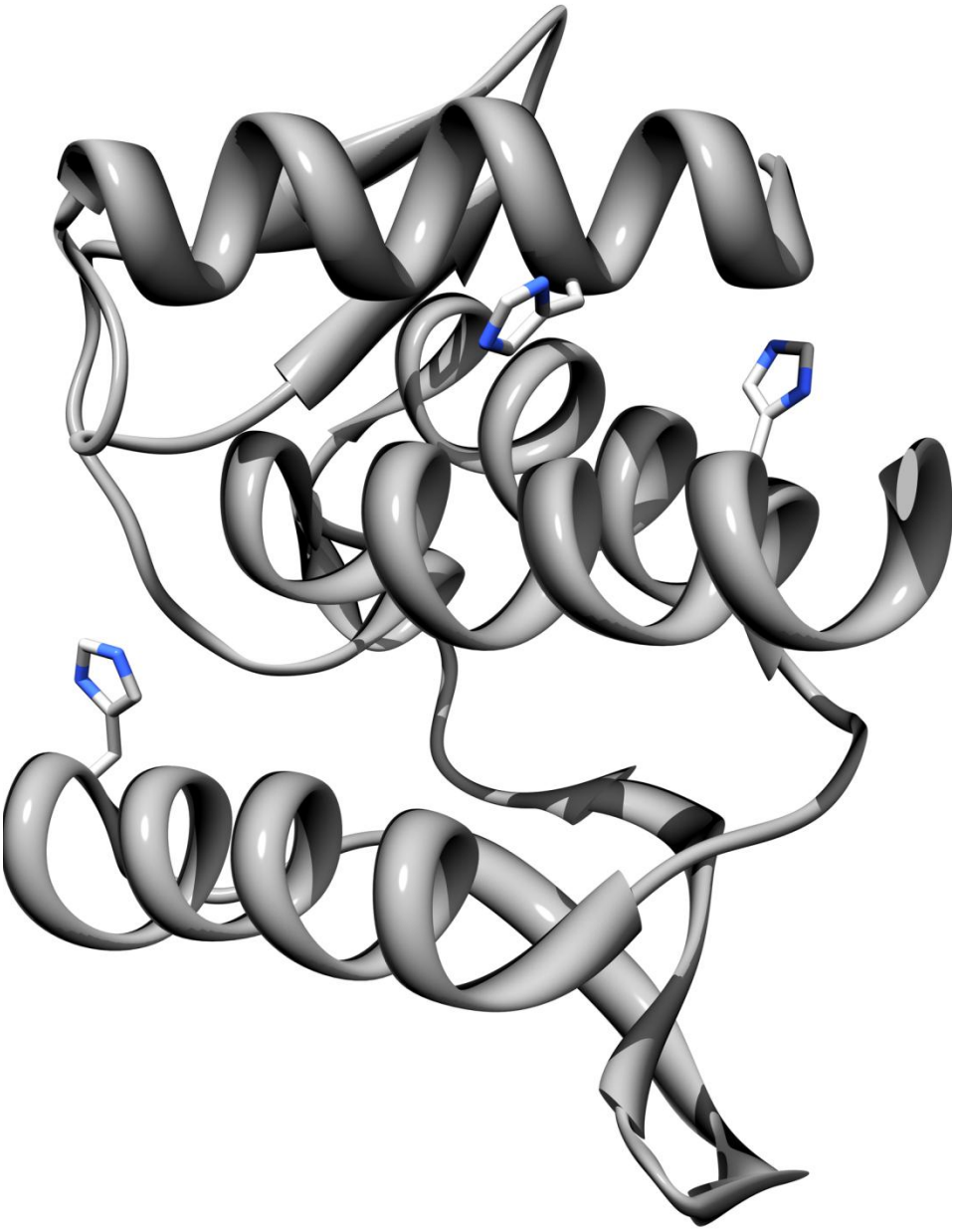


Figure 5.8: The N-terminus of the effector protein YopH (PDB ID 1HUF) showing the three surface exposed histidines.

these observations in terms of pK_a^F and pK_a^U of the titratable residues of the protein and highlights the role of buried histidine H87 as well as COOH salt-bridge in the protein's stability. Thus, this work beautifully illustrates the role of titratable groups in tuning the pH-dependent stability of effector proteins and it will be interesting to see whether the proposed pH-folding switch can explain some of the ambiguities discussed in the literature.

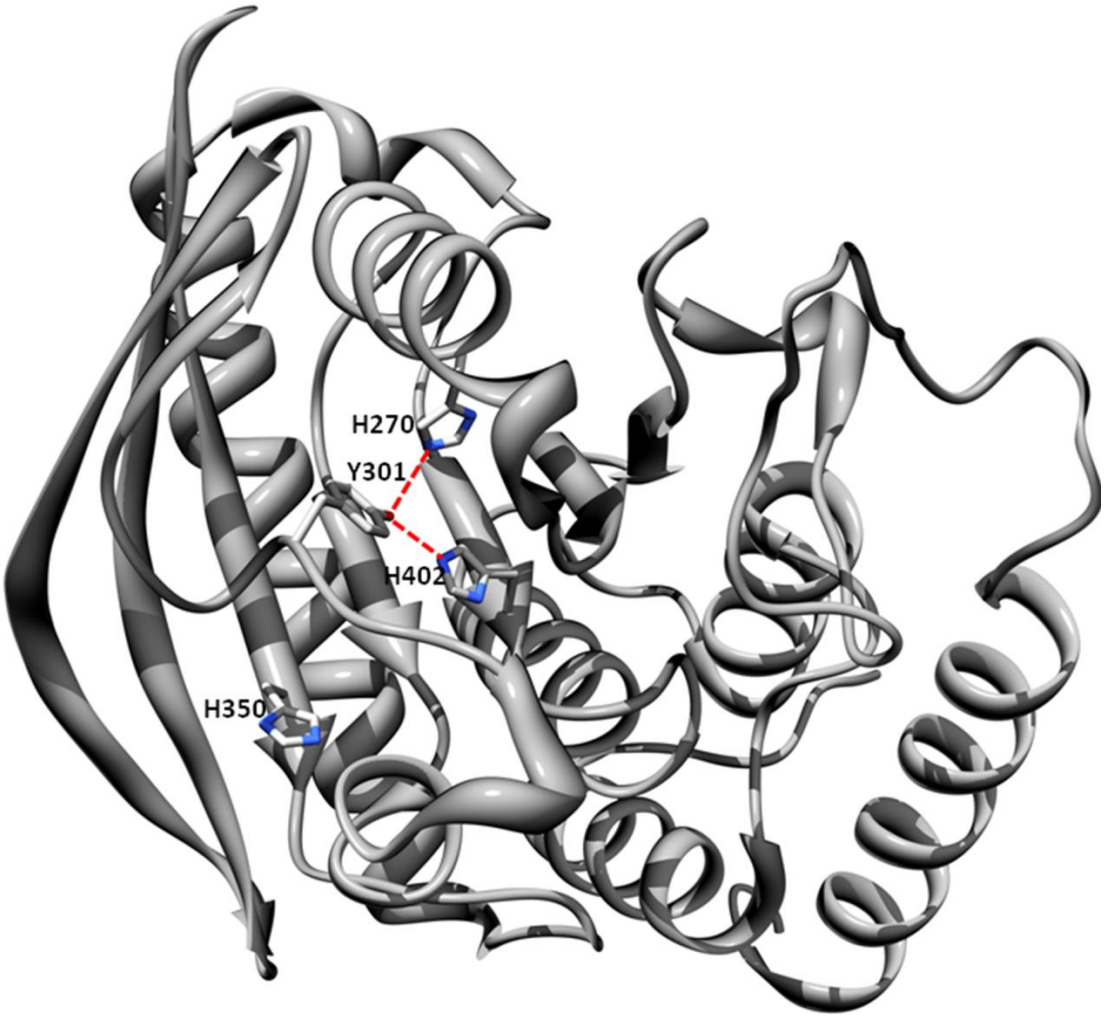


Figure 5.9: The phosphatase domain of the effector protein YopH (PDB ID 1QZ0) showing the exposed histidine H350. The histidines H270 and H402 are buried in the core pf the protein and also from hydrogen bonds with Y301 residue sidechain.

As mentioned earlier, fusion of the N-terminal secretion signal of YopE to the destabilized mutant of DHFR led to its secretion through the T3SS, while the stable WT DHFR was secretion incompetent. In a detailed review of T3SS (Ghosh, 2004) the author pointed out that the stability of WT DHFR (~5 to 6 kcal/mol) is similar to that of YopH (~6 kcal/mol) and the RhoGAP domain of YopE (~7 kcal/mol) effector proteins. This raises the question that with similar stability how do the effector proteins get secreted while the WT DHFR does not? As another example, the highly stable Green Fluorescent Protein (GFP) is readily secreted when fused with the N-terminus secretion signal. GFP has a beta-barrel structure with a diameter of ~3.5 nm, which makes it virtually impossible for GFP to go through the T3SS without unfolding. Keeping in mind the pH-folding switch, a detailed investigation on the pH-dependence of the stability of these proteins might answer these questions.

There are three histidines in each of the N-terminus (residues 3 to 125, PDB ID 1HUF, 1K46, 1MOV) and the phosphatase domain (residues 167 to 468, PDB ID 1Q70) of YopH. The histidines H9, H62 and H83 in the N-terminus are surface exposed and hence, should not contribute much to the pH-dependent stability of YopH (Figure 5.8). In the larger phosphatase domain only H350 is surface exposed. Histidines H270 and H402 are completely buried in the core of the domain and the Nε2 of each sidechain also form H-bonds with the sidechain of Y301 (Figure 5.9). It can be predicted that these two histidines will have very depressed pK_a^F and, based on the Yang and Honig model (Yang and Honig, 1993), should contribute to the pH-dependent unfolding of YopH. Also the stability of YopH (~6 kcal/mol) was measured at a pH = 8 (Birtalan et al., 2002), which being basic should favor the folded state.

For YopE, two crystal structures have been reported, one for N-terminus (PDB: 1HY5) and other for the chaperone binding region of YopE in complex with its

chaperone SycE (PDB: 1L2W). Not enough structural information is present in these fragments in order to make predictions regarding the effector's pH-dependent stability.

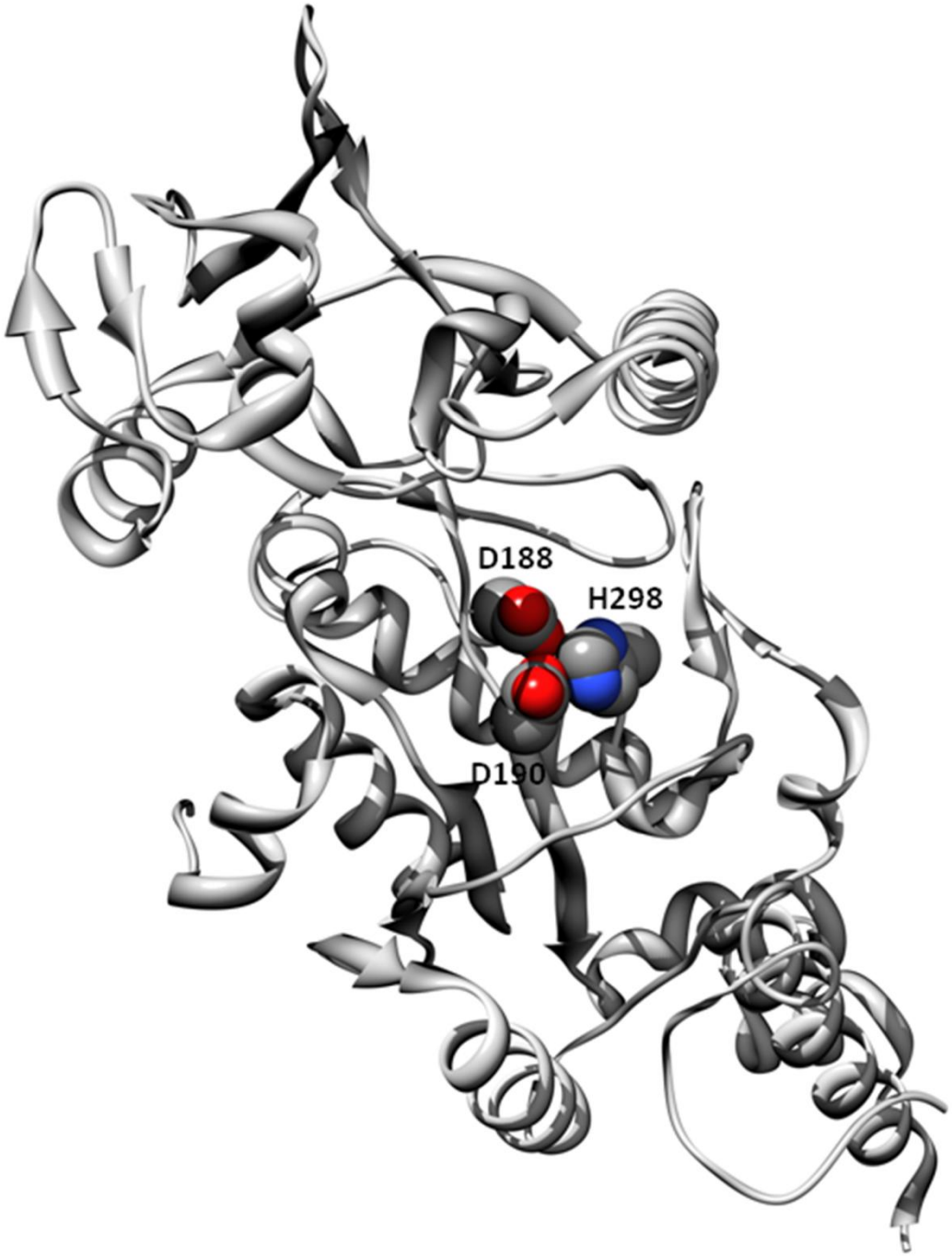


Figure 5.10: The core of Cya shows buried residues D188, D190 and H298.

Interestingly, GFP also displays pH-dependent stability (Alkaabi et al., 2005; Penna et al., 2004). Alkaabi and co-workers have shown that GFP is highly stable at pH 7.5 and 8.5 but has drastically reduced stability at pH of 6.5. The reduced stability of GFP at pH 6.5 (approximate internal pH of the pathogens) would facilitate its secretion through the T3SS when fused with the secretion signal. Thus, the pH-dependent folding switch can partially explain some of the ambiguous observations made in the literature regarding the translocation of effector proteins through the T3SS.

The biochemical assay for the *in vivo* detection of the effector protein translocation is done by the use of Effector-Cya fusion protein, where the Cya is an adenylate cyclase that produces cyclic AMP (cAMP) once inside the host cell and the detection of cAMP produced reports on the level of translocation. Cya is a calmodulin dependent enzyme (Botsford and Harman, 1992) and since prokaryotes lack calmodulin, the Cya fusion protein produces cAMP only after translocation to the host cell. This strategy was first used to study translocation of YopE from *Yersinia enterocolitica* into HeLa cells (Sory and Cornelis, 1994). A bacterial adenylate cyclase was used from *Bordetella pertussis*. The same strategy was used to develop a biochemical assay for the detection of translocation of AvrBs2 effector protein into plant cells (Casper-Lindley et al., 2002). The success of this assay depends on the efficient translocation of the Effector-Cya fusion protein. The crystal structure of Cya from *Bordetella pertussis* has been reported (PDB ID 1YRT). Cya has a minimum dimension of ~4.2nm and hence, has to unfold in order to pass through the T3SS. It has eight histidines out of which H108 is partially buried and H298 is completely buried. There are two aspartic acids D188 and D190 within 3.5Å of H298 and are also partially buried (Figure 5.10). D198 sidechain also forms a hydrogen bond with H298. Based on the Yang and Honig model (Yang and Honig, 1993) we can predict

that D188, D190 and H298 would impart pH-dependent stability to Cya and facilitate its facile translocation through the T3SS.

Thus, the pH-dependence of protein stability has added a new dimension to our understanding of the T3SS. Nature has elegantly utilized this pH-folding switch to unfold effector proteins inside the pathogens, deliver them through the T3SS and again refold them once inside the host cell. The same folding switch has, serendipitously, served us in developing biochemical assays to study the T3SS. It will be interesting to see how our detailed knowledge of the pH-folding switch helps us in understanding bacterial virulence and T3SS.

CHAPTER 6

FUTURE DIRECTIONS AND PERSPECTIVES *

This work studied, at the molecular level, the role of the peptidyl-prolyl isomerase Pin1 and its substrates the *cis*/*trans* isomers of pAICD (AICD phosphorylated at T668) in Alzheimer's disease (AD). One major outcome of this work is the proposed role of Pin1 as a connector and flux modulator of protein interaction network modules. This proposed role provides for the first time a molecular basis to explain the impact of Pin1's enzymatic action on various cellular processes and hinges upon the assumption that the substrate and product of Pin1 enzymatic action, i.e. *cis* and *trans* isomers, participate in isomer-specific interactions. Even though several examples of *trans*-specific interactions are known, identification of *cis*-specific interactions are yet to be made (here *trans* and *cis* refer to the phosphorylated substrates of Pin1). This work also contributed towards the design and synthesis of peptide analogues of the *cis* and *trans* isomers which are currently being used in the Nicholson lab for the identification of the pAICD *cis*/*trans* isomer-specific binding partners in the cell. Identification of these putative binding partners would open up exciting new frontiers of research. First, such isomer-specific binding partners, for both the isomers, can be used in rationally designed interaction network modules to experimentally show the impact of Pin1 activity as a connector and flux modulator of these modules. Second, identification of isomer-specific binding partners would provide important clues in understanding the signaling pathways that the *cis* and *trans* isomers of pAICD are involved in and should guide future experiments, both NMR and cell biology based. Finally, such binding partners can be

---

* Parts of this chapter are based on a manuscript submitted to *Nature Chemical Biology*.

targeted for therapeutic purposes in Alzheimer's disease.

Speculating the roles of *cis* and *trans* isomers of pAICD in AD

As described in detail in the first chapter, the phosphorylation of T688 in the cytoplasmic tail of APP is regarded as an important signaling event that plays a major role in deciding the processing fate of APP. The key question that arises is, how does the phosphorylated T668 influence APP processing? Two mechanisms can be postulated to answer this. First, T668 phosphorylation dictates the localization of APP in the cell and thus, influences its processing. For this case, the pT668 would have to be in the full length APP. The pT668-P669 *trans*-specific interaction network might be involved in APP trafficking to the plasma-membrane or out of the endosomes thus, favor the non-amyloidogenic processing whereas the corresponding *cis*-specific interaction network might be involved in endocytosis thus, favor the amyloidogenic processing. We have demonstrated using mathematical modeling that the *cis*-specific interaction network becomes dominant as compared to the *trans*-specific interaction network only in the absence of Pin1. Thus, it seems that this postulate can explain the observed increase in amyloidogenic processing of APP in the absence of Pin1. But there is a severe drawback in this postulate. Even though the *cis*-specific interaction network becomes dominant in the absence of Pin1, the flux through it is much smaller, owing to the small population of *cis* isomer, than the flux through the dephosphorylated APP interaction network which most probably favors the non-amyloidogenic processing of APP. This might be a key factor in the slow onset of the Alzheimer's disease. The hypothesized role of *cis*-isomer, that it enhances amyloidogenic processing of APP in the absence of Pin1, can also be explained if the longer lived signal (i.e. *cis*-isomer) gets amplified. This leads us to the second mechanism that the *cis*-isomer in the C-terminal fragment of APP, i.e. pAICD formed

after the γ-secretase activity, enhances the transcription of certain factors that take part in trafficking and processing of APP.

Under diseased condition, T668 phosphorylation increases (Lee et al., 2003b) and Pin1 is down regulated (Huang et al., 2010). Also it has been observed that APP processing increases in Pin1-knockdown cells (Pastorino et al., 2006). These observations can be explained if the *cis* or *trans* isomer of pAICD fragment promotes the transcription of α, β and γ-secretases. Taking account of all the possibilities presented above, we can postulate a hypothetical model for the signaling pathways involving *cis* and *trans* isomers of pAICD.

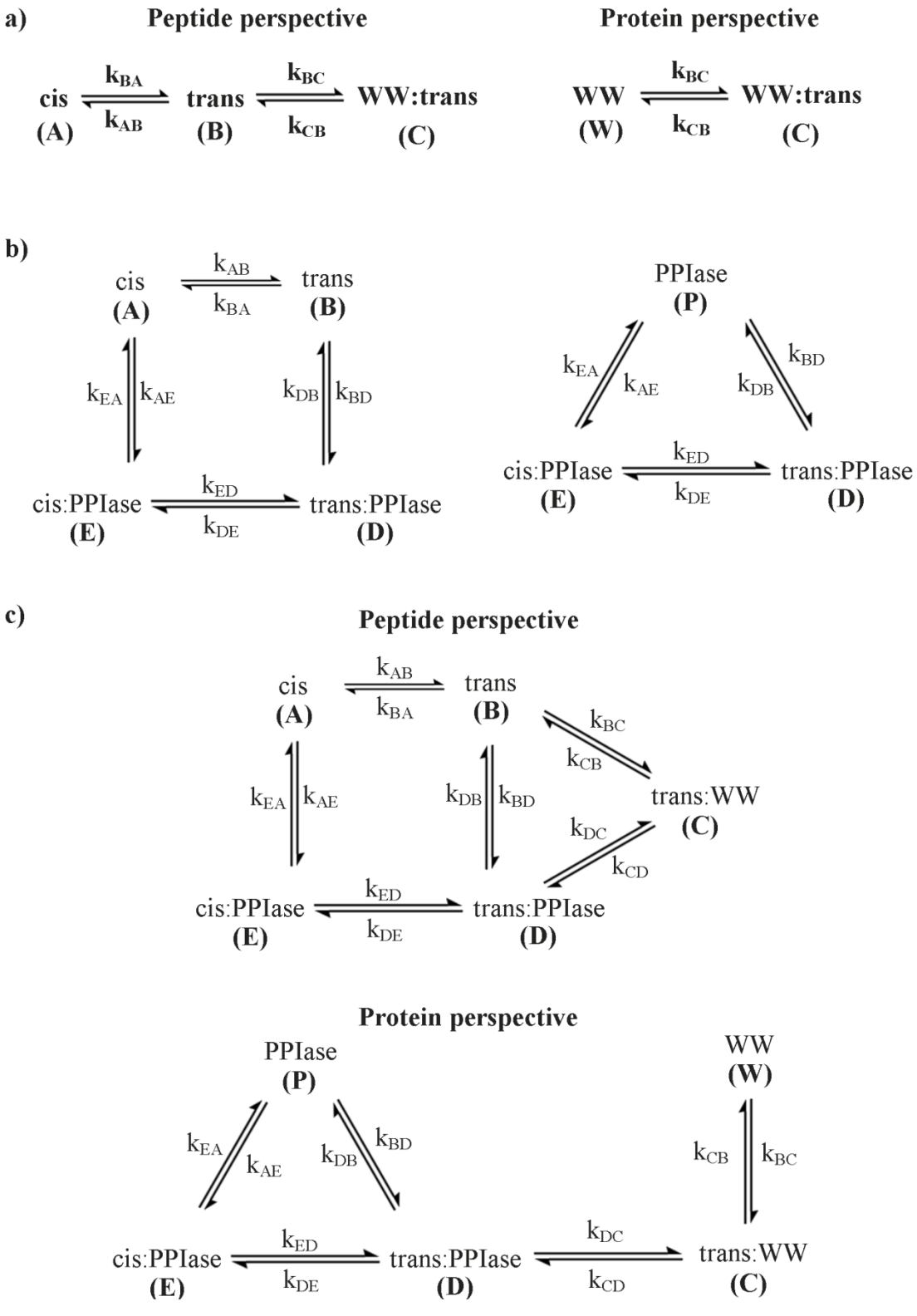
1. The *trans* isomer in full length APP is most probably involved in APP trafficking towards the plasma membrane and/or out of the endosomes. Thus, it favors the non-amyloidogenic processing of APP.
2. In the absence of Pin1, the insulated *trans*-specific interaction network module of pAICD fragment is involved in the transcription of α, β and γ-secretases. In the presence of Pin1, the *cis* and *trans* specific network modules are connected and the flux through the *trans* module is less, resulting in the lower levels of transcription of α, β and γ-secretases.
3. In the absence of Pin1, the insulated *cis*-specific interaction network module of pAICD fragment is involved in the transcription of cellular factors that recognize the endocytic signal in the cytoplasmic tail of APP and promote its endocytocis. Even though the flux through the *cis*-specific network module is lower in the absence of Pin1, its prolonged activation would eventually lead to greater transcription of the cellular factors involved in the endocytocis of APP. Thus, the *cis* isomer of pAICD, in the absence of Pin1, can enhance the amyloidogenic processing of APP.

4. There is no clear evidence in the literature that correlates down regulation of Pin1 and phosphorylation levels of T668 in APP and in tau proteins. It would be interesting to compare the phosphorylation levels of these proteins in wild type and Pin1 knockout (or knockdown) cells. Also, it would be interesting to compare the expression levels of the proline-directed kinases and phosphatases in these two cell types.

Lineshape analysis of PPIase domain and full length Pin1

The current work demonstrated the power, sensitivity and reliability of NMR lineshape analysis in determining the kinetics of protein-ligand interaction by studying the interaction of the isolated Pin1-WW domain and pAICD derived peptide pAPP659-682. The next step will be to study the interaction of isolated Pin1-PPIase domain and pAPP659-682. The interaction models for isolated Pin1-WW domain and isolated Pin1-PPIase domain are shown in Figure 6.1a and 6.1b, respectively. The *trans* isomer specificity of the WW domain simplified the lineshape analysis of this interaction. For the peptide perspective of this interaction, only the peaks corresponding to the *trans* isomer were analyzed while the slow isomerization of the free *cis* to free *trans* isomer was accounted for in the thermodynamic equations of the fitting model. Similar simplification is not possible for the PPIase domain, since it interacts with both the isomers and catalyzes their isomerization. Analysis of the protein perspective is relatively simpler. For each residue of the protein, there is only one peak that corresponds to the free state (Figure 6.1b, state P) and only one peak that corresponds to the bound state which is the population weighted average of the bound *cis* (Figure 6.1b, state E) and bound *trans* states (Figure 6.1b, state D) due to exchange in the fast timescale between them. Complications arise in the analysis of the peptide perspective of this interaction. For each residue of the peptide, there are two peaks for

Figure 6.1: Fitting models for NMR lineshape analysis. The peptide and protein perspective models for a) isolated WW domain, b) isolated PPIase domain, and c) full length Pin1. Only the states observable by NMR are accounted for in the models. The single letter symbols used to identify each state are consistent for all the models. The microscopic rates for each step are also consistent for all the models.


a)
Peptide perspective
cis
(A)
trans
(B)
WW:trans
(C)
Protein perspective
WW
(W)
WW:trans
(C)
b)
cis
(A)
trans
(B)
cis:PPIase
(E)
trans:PPIase
(D)
PPIase
(P)
cis:PPIase
(E)
trans:PPIase
(D)
c)
Peptide perspective
cis
(A)
trans
(B)
trans:WW
(C)
cis:PPIase
(E)
trans:PPIase
(D)
Protein perspective
PPIase
(P)
WW
(W)
cis:PPIase
(E)
trans:PPIase
(D)
trans:WW
(C)

the free state corresponding to free *cis* (Figure 6.1b, state A) and free *trans* (Figure 6.1b, state B), respectively. This poses some practical problems in terms of the implementation of lineshape analysis in BiophysicsLab. Since the *cis* and the *trans* isomers appear as two distinct peaks, their lineshapes are extracted as separate datasets (see Appendix 1) whereas for the actual fitting they need to be combined into a single dataset. This would require additional processing of the extracted datasets from the labeled peptide titration to be able to fit to the model (Figure 6.1b, peptide perspective) using lineshape analysis. This procedure would also be necessary for the fitting of the titration data of the full length Pin1 to the corresponding model (Figure 6.1c, peptide perspective).

For the lineshape analysis of the PPIase domain interaction with the peptide pAPP659-682, additional restraints for the microscopic rates (Figure 6.1b) can be obtained from catalytic rate constants measured from the ^{1}H-^{1}H ROESY experiments. The catalytic rate constants, k_{ct} and k_{tc}, represent the *cis* to *trans* and *trans* to *cis* conversions, respectively. These rate constants are related to the Michaelis-Menten constant (K_M) and rate constant (k_{cat}) as follows

$$k_{CT} = \frac{k_{CT}^{cat}}{K_M^{cis}}[E] \tag{6.1a}$$

$$k_{TC} = \frac{k_{TC}^{cat}}{K_M^{trans}}[E] \tag{6.1b}$$

In the above equations [E] is the free enzyme concentration. The Michaelis-Menten constant (K_M) and rate constant (k_{cat}) are related to the microscopic rate constants (Figure 6.1b) as follows

$$k_M^{trans} = \frac{k_{EA}k_{DB}+k_{EA}k_{DE}+k_{ED}k_{DB}}{k_{BD}(k_{EA}+k_{ED}+k_{DE})} \tag{6.2a}$$

$$k_M^{cis} = \frac{k_{EA}k_{DB}+k_{EA}k_{DE}+k_{ED}k_{DB}}{k_{AE}(k_{ED}+k_{DE}+k_{DB})} \tag{6.2b}$$

$$k_{TC}^{cat} = \frac{k_{EA}k_{DE}}{k_{EA}+k_{ED}+k_{DE}} \tag{6.2c}$$

$$k_{CT}^{cat} = \frac{k_{ED}k_{DB}}{k_{ED}+k_{DE}+k_{DB}} \quad (6.2d)$$

Substituting equations (6.2) into equations (6.1) we get

$$k_{CT} = \frac{k_{ED}k_{DB}k_{AE}}{k_{EA}k_{DB}+k_{EA}k_{DE}+k_{ED}k_{DB}}[E] \quad (6.3a)$$

$$k_{TC} = \frac{k_{EA}k_{DE}k_{BD}}{k_{EA}k_{DB}+k_{EA}k_{DE}+k_{ED}k_{DB}}[E] \quad (6.3b)$$

The free enzyme concentration [E] in the ROESY experiment will be approximately equal to the concentration of the total enzyme used in the experiment. Otherwise, it can also be calculated from the dissociation constants K_D^{cis} and K_D^{trans}, which should be fitting parameters in the lineshape analysis. Similar approach can be used for the lineshape analysis of the full length Pin1 titration data.

In conclusion, the studies completed thus far lay important groundworl for elucidating the microscopic rate and equilibrium constants for the Pin1 reaction scheme. Important insights have emerged from this work, including a novel proposed role of Pin1 as a connector and flux modulator of protein interaction network modules, and an appreciation of the power and sensitivity of NMR lineshape analysis. The kinetic and thermodynamic characterization of Pin1 will provide a foundation on which to elucidate detailed mechanisms involved in the binding and catalytic functions of this enzyme through similar studies using Pin1 mutants.

APPENDIX 1

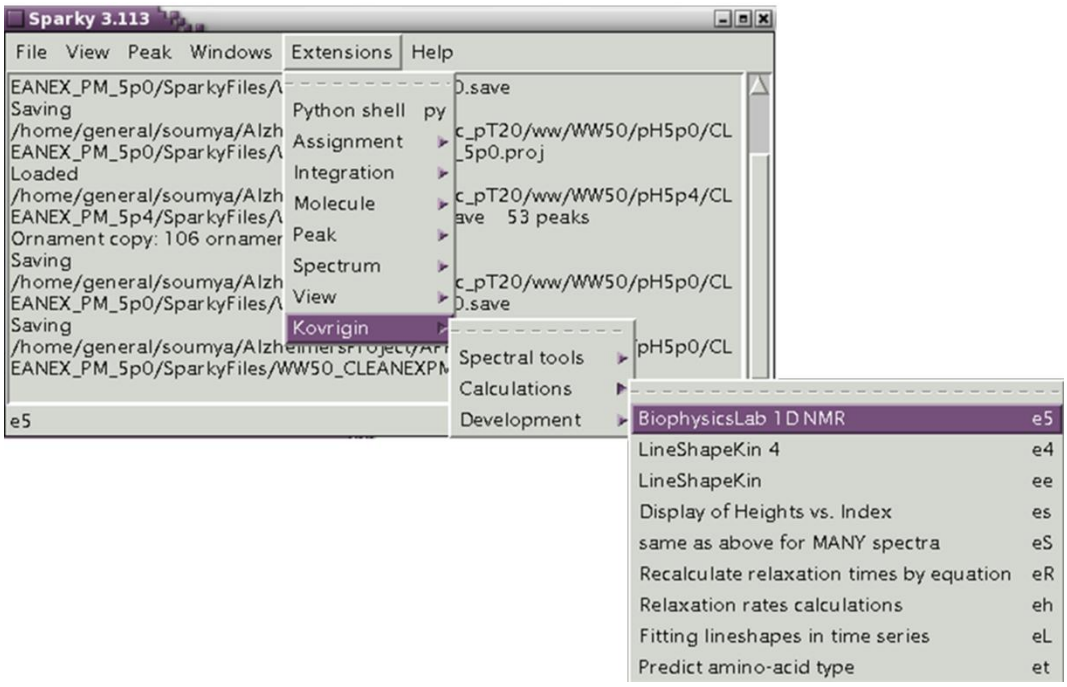
PRACTICAL ASPECTS OF LINESHAPE ANALYSIS

The first step is to perform the titration experiment and collect ^{15}N-^{1}H HSQC spectrum for each titration point. The following points should be considered while designing and performing the experiments:

1. If there is enough protein then always use the dilution method for titration. Start with the most saturated sample and keep on diluting it with the ^{15}N-labeled component, so that the concentration of the labeled species is constant for all titration points. The unlabeled component keeps getting diluted.
2. Run the extreme points first i.e. saturated and the apo spectra. This allows the determination of the minimum sweep widths for both dimension.
3. Adjust the sweep widths (sw, sw1 and dof2) to the minimum required. Use atleast np=2048 and ni=256. It is always good to have more points define a peak.
4. While processing data, use EM window function to maintain the lorentzian lineshape.
5. Set zerofilling to auto, more points would hamper global fitting. Zerofilling can be increased later to produce better figures for publication.

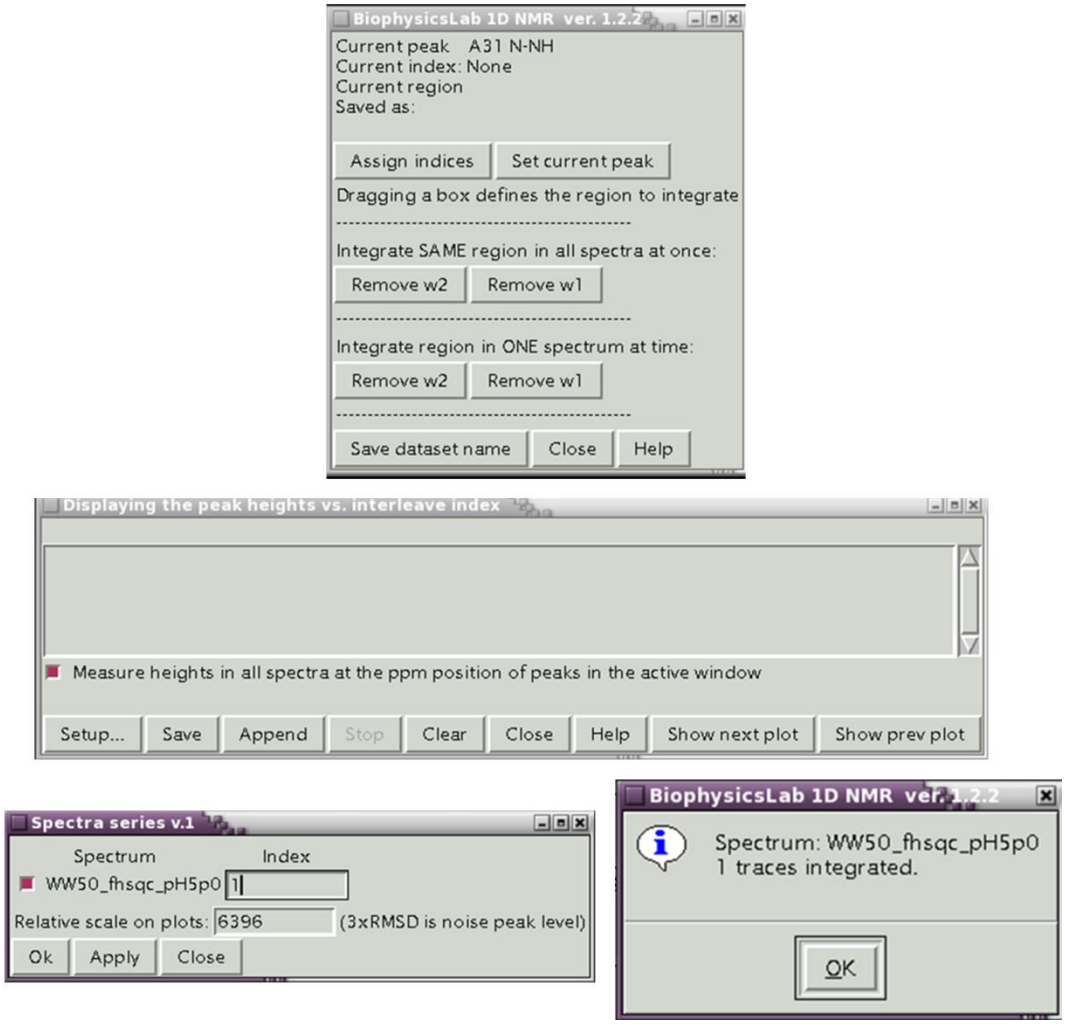
Once the data is collected, process it with nmrPipe and read it in sparky. We are using BiophysicsLab (http://biophysicslab.net/) to perform the lineshape analysis. The raw data for lineshape analysis is generated using sparky. Open all the spectra of

a titration experiment in sparky. Type ‘e5’ or alternatively click ‘BiophysicsLab 1D NMR’ as shown below.



Clicking the above options a new dialogue box as shown below. Once this dialogue box opens, click on ‘Assign indices’. This should open another dialogue box ‘Displaying the peak heights vs. interleave index’. In this dialogue box click the ‘Setup…’ button. A new dialogue opens which lists all the spectra opened in sparky. Check the button for all of them and number them from 1 to ‘n’, where ‘n’ is the number of titration points or spectra. The apo spectra should be 1 and the most saturated one ‘n’. Click ‘OK’ to close. Close the previous one too. Now select a peak to be analyzed. Then click ‘Set current peak’. Drag a box around the peak and make sure to cover enough region, so that the peak in all the spectra come within the box. Under ‘Integrate SAME region in all spectra at once:’ click ‘Remove w1’. This would get slices for the w2 (^{1}H) dimension from all the spectra. It would also indicate the number of slices integrated for each spectra (see figure below). This method should

only be used for well isolated peaks. Clicking on 'Remove w2' does the same for w1 (^{15}N) dimension. The above step is the first step for getting slices and should be used only once if slices from each spectrum are individually collected. The other option (see figure below) can be used to get slices from each spectrum. In this case, integrate one to three traces for each spectrum. The raw data is saved in a folder named 'Data_for_biophysicsLab'



Lineshape analysis using BiophysicsLab

BiophysicsLab is implemented in matlab. All the basic programs that constitute BiophysicsLab are in a folder named ‘code’ It is necessary to add this folder in the path of matlab. BiophysicsLab has two classes:

1) Dataset: It is a superclass for all types of data.

2) TotalFit: It contains methods that can load dataset objects and fit them to the user defined models.

NMRLineShapes1D is a derivative class of Dataset and contains all the predefined as well as user defined models. This is the only class that we need to modify when we come up with a new model. The important packages are: equilibrium_thermodynamic_equations, kinetic_matrices and line_shape_equations_1D. The class folders start with the ‘@’ sign and the package folders start with the ‘+’ sign and are located in ‘code’ folder. The working folder should have two folders:

1) Data_for_BiophysicsLab: This is created by sparky and has your NMR data.

2) BiophysicsLab: This is the data is fit. Create three subfolders for BiophysicsLab.

   I. 1_Preprocess
   II. 2_Individual_fits
   III. 3_Global_fit

Now setup the fitting model. Use MuPad to derive the thermodynamic equilibrium populations for all the species i.e. free ligand, free protein and bound protein. In ‘code’ go to the folder ‘+equilibrium_thermodynamic_equations’ and

create a file named ‘U_L_model.m’ and define all the equilibrium values. Determine the rate-matrix and write it in the file ‘U_L_model_K_matrix.m’ in the folder ‘+kinetic_matrices’. Create the file ‘U_L_model_1D.m’ in the folder ‘+line_shape_equations_1D’ to generate 1D lineshapes. Finally add the model to the file ‘NMRLineShapes1D.m’ in the folder ‘@NMRLineShapes1D’.

Before the actual lineshape fitting, it is necessary to preprocess the data. This is done in the folder ‘1_Preprocess’. An example processing file is ‘A31_w2.m’. Here load the data and visualize the slices, trim to remove overlap or extra baseline and finally normalize the slices to unit area. Any simple model can be used for the initial guess and the whole session is saved. Saving the sessions lets us to start from the point where we stop. The actual lineshape fitting for an individual dataset is done in the folder ‘2_Individual_fits’. An example of a processing file is ‘A31_w2_U_L_1.m’. The following are the important points to be kept in mind while setting up this file:

1. session.batch_link_parameters_in_series(series_name, [1 3 4 5 6 7 8 9 10 11]). Here list all the parameters in the model expect the second one which is the ligand-to-protein ratio and changes for each titration point.
2. ideal_parameters=[Rtotal 1e-4 6e4 0.08974359 2000 w0_1 w0_2 FWHH_1 FWHH_2 1 0.4]. The ideal parameters should be guessed well to get realistic solutions. If the fitting gives negetive value for some parameter, then change the initial value for that parameter and try again. Sometimes changing some other parameter also helps.
3. session.set_parameters_modes(1,[0 0 1 0 1 0 1 0 0 0 1]). This array has the same number of elements as that of the number of parameters in the

fitted model. The entry corresponding to the parameter that is being fit is ‘1’ while the parameter that is not being fit is ‘0’.

4. main_session.fit_only('Newton'). This is the fitting routine. Here the confidence limits of the fitted parameters are not determined.
5. In order to determine the confidence limits of the fitted parameters, set the range for the parameters. An example is the file ‘A31_w2_U_L_1_MC.m’.
6. min_values= [0 0 0 0 0 0 0 0 0 0 0]. max_values= [1 20 1e6 1 1e6 1e6 1e6 1e3 1e3 10 1]; This sets the range for all the parameters.
7. session.fit('simplex', 'single-core'). This is the fitting routine.
8. Remember that the positions of peaks govern the binding constant and their lineshapes govern the binding rates.

An example file for global fitting is ‘global_U_L_1.m’. The important features are as follows:

1. In global fitting it is important to define the global parameters (such as k_{OFF} and K_D) and make sure it is the same (i.e. linked) for all datasets.
2. main_session.link_parameters(); This routine links all the global parameters.

APPENDIX 2

DECIPHERING THE INFORMATION ON PIN1 CATALYSIS OBTAINED FROM NMR ^{1}H-^{1}H ROESY EXPERIMENT

The ^{1}H-^{1}H ROESY experiment is performed to obtain the rates of isomerization k_{ct} and k_{tc} between *cis* and *trans* isomers of pAICD derived peptides as shown in equation (1). The intrinsic conversion rates are too slow to be detected by ROESY experiments. Addition of catalytic amount of Pin1 accelerates this rate and then the catalytically enhanced rate constants k_{ct} and k_{tc} can be measured by ROESY experiments. In order to understand the physical meaning of these catalytically enhanced rate constants measured by ROESY, we have to consider the actual physical process taking place inside the NMR tube. In order to measure the rate constants, several ROESY experiments are performed with different mixing time (t_{mix}) which range from 5ms to 200ms. The NMR experiment is done under equilibrium conditions and it does not perturb the system to change the equilibrium populations. Hence, at any given time we have 7% *cis* and 93% *trans* populations of the two isomers. The two isomers have distinct chemical shifts. The molecules which do not isomerize during the mixing time contribute to the diagonal peaks, i.e. if a particular peptide is in *cis* conformation at the beginning and at the end of the mixing time of a ROESY experiment, then that molecule would contribute to the *cis* diagonal peak of the spectrum. On the other hand, those molecules that isomerize during the mixing time give rise to the cross-peaks.

As an illustration, let assume that there are 1000 peptide molecules, so that 70 are *cis* and 930 are *trans*. If we consider the *cis* molecules only then at the start of each acquisition we have 70 molecules as *cis*. After we wait for time t_{mix} some of those *cis* get converted to *trans*. As we increase t_{mix} we get more and more of the *cis* converted to *trans* (giving the signal I_{ct}) and also some of those *trans* being converted

back to *cis* (adding to the *cis* not yet converted to *trans* and giving the signal I_{cc}). If we increase t_{mix} to long enough time, then we will get equilibrium population i.e. 5 *cis* molecules (I_{cc}) and 65 *trans* molecules (I_{ct}). Now, we have to figure out the true meaning of k_{ct} and k_{tc} that we are measuring with ROESY experiment.

$$\textbf{Cis} \underset{k_{TC}}{\overset{k_{CT}}{\rightleftharpoons}} \textbf{Trans} \tag{1}$$

For the reaction where *cis* is converting to *trans*, the initial reaction velocity (v_{in}) is defined as:

$$v_{in}^{cis} = -\frac{d[C]}{dt} = k_{CT}[C] = \frac{[E_0][S]k_{cat}}{K_M + [S]} \tag{2}$$

This is the well-known Michaelis-Menten equation. For any substrate concentration, this equation can be rewritten as (Fersht p-111):

$$v_{in}^{cis} = \frac{k_{cat}}{K_M}[E][S] \tag{3}$$

Thus, k_{cat}/k_M is a second-order rate constant and $[E]k_{cat}/k_M$ is a pseudo-first order rate constant. This is a reasonable approximation since, under steady state condition the concentration of the enzyme-substrate complex is constant i.e. [ES] = constant. Therefore, the concentration of the free enzyme is also constant, $[E] = [E_0] - [ES]$ = constant. Comparing equations (2) and (3) we get

$$k_{CT}[C] = \frac{k_{cat}}{K_M}[E] \tag{4}$$

Now, we can derive an expression which we will use to fit the ROESY data and extract out k_{CT} and k_{TC}. For the *cis* to *trans* conversion, we have the following set of equations:

$$\frac{d[T]}{dt} = k_{CT}[C] - k_{TC}[T] \quad (5)$$

$$\frac{d[C]}{dt} = k_{TC}[T] - k_{CT}[C] \quad (6)$$

The boundary conditions for the above equations are:

[C] = 1 and [T] = 0 at time t = 0. (7)

Solving the set of equations (5) and (6) and applying the boundary conditions (7), we get

$$[C] = \frac{k_{CT}e^{-(k_{CT}+k_{TC})t}+k_{TC}}{k_{CT}+k_{TC}} \quad (8)$$

$$[T] = \frac{\left(1-e^{-(k_{CT}+k_{TC})t}k_{CT}\right)k_{CT}}{k_{CT}+k_{TC}} \quad (9)$$

Now I_{CC} α [C] and I_{CT} α [T], so I_{CT}/ I_{CC} = [T]/[C]. Therefore,

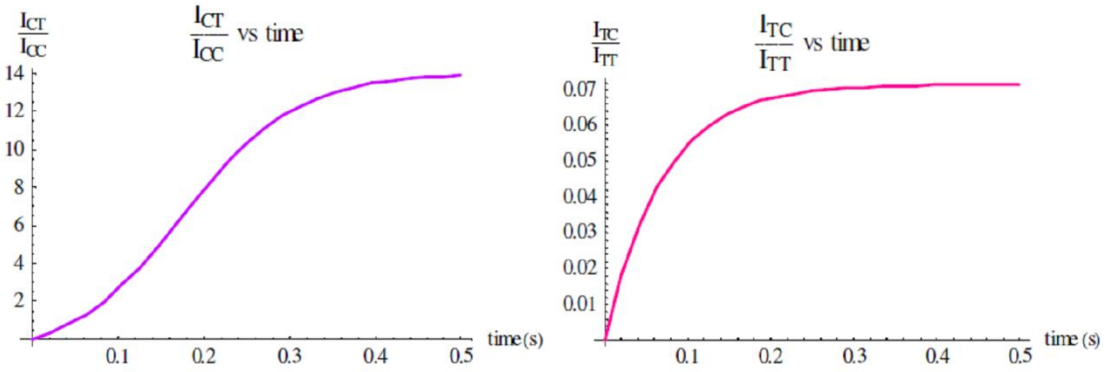
$$\frac{I_{CT}}{I_{CC}} = \frac{\left(-1+e^{(k_{CT}+k_{TC})t}\right)k_{CT}}{k_{TC}e^{(k_{CT}+k_{TC})t}+k_{CT}} \quad (10)$$

Similarly we get

$$\frac{I_{TC}}{I_{TT}} = \frac{\left(-1+e^{(k_{CT}+k_{TC})t}\right)k_{TC}}{k_{CT}e^{(k_{CT}+k_{TC})t}+k_{TC}} \quad (11)$$

Equations (10) and (11) are the expressions used to extract the exchange rate constants from ROESY data.

The plots for the above two expressions are shown below:



Therefore, in conclusion we have shown that the catalytic conversion of *cis* to *trans* and vice versa is a pseudo-first order reaction and valid for all enzyme concentrations. The rate constants obtained from the ROESY experiments are related to k_{cat} and k_M as follows:

$$k_{TC} = \frac{k_{TC}^{cat}}{K_M^{trans}}[E]$$

$$k_{CT} = \frac{k_{CT}^{cat}}{K_M^{cis}}[E]$$

REFERENCES

Albery, W.J., and Knowles, J.R. (1976). Evolution of enzyme function and the development of catalytic efficiency. Biochemistry *15*, 5631-5640.

Alfano, J.R., and Collmer, A. (2004). Type III secretion system effector proteins: double agents in bacterial disease and plant defense. Annu Rev Phytopathol *42*, 385-414.

Alkaabi, K.M., Yafea, A., and Ashraf, S.S. (2005). Effect of pH on thermal- and chemical-induced denaturation of GFP. Appl Biochem Biotechnol *126*, 149-156.

Allinson, T.M., Parkin, E.T., Turner, A.J., and Hooper, N.M. (2003). ADAMs family members as amyloid precursor protein alpha-secretases. J Neurosci Res *74*, 342-352.

Alzheimer, A. (1907). A new disease of the cortex (Ger). Allg Z Psychiatr *64*, 146-148.

Ando, K., Iijima, K.I., Elliott, J.I., Kirino, Y., and Suzuki, T. (2001). Phosphorylation-dependent regulation of the interaction of amyloid precursor protein with Fe65 affects the production of beta-amyloid. J Biol Chem *276*, 40353-40361.

Ando, K., Oishi, M., Takeda, S., Iijima, K., Isohara, T., Nairn, A.C., Kirino, Y., Greengard, P., and Suzuki, T. (1999). Role of phosphorylation of Alzheimer's amyloid precursor protein during neuronal differentiation. J Neurosci *19*, 4421-4427.

Ayed, A., Mulder, F.A., Yi, G.S., Lu, Y., Kay, L.E., and Arrowsmith, C.H. (2001). Latent and active p53 are identical in conformation. Nat Struct Biol *8*, 756-760.

Baek, S.H., Ohgi, K.A., Rose, D.W., Koo, E.H., Glass, C.K., and Rosenfeld, M.G. (2002). Exchange of N-CoR corepressor and Tip60 coactivator complexes links gene expression by NF-kappaB and beta-amyloid precursor protein. Cell *110*, 55-67.

Balastik, M., Lim, J., Pastorino, L., and Lu, K.P. (2007). Pin1 in Alzheimer's disease: multiple substrates, one regulatory mechanism? Biochim Biophys Acta *1772*, 422-429.

Bartus, R.T., Dean, R.L., 3rd, Beer, B., and Lippa, A.S. (1982). The cholinergic hypothesis of geriatric memory dysfunction. Science *217*, 408-414.

Bartus, R.T., Dean, R.L., and Fisher, S.K. (1986). Cholinergic treatment for age-related memory disturbances: dead or barely coming of age. Treatment development strategies for Alzheimer's disease, 421-450.

Bartus, R.T., Fleming, D., and Johnson, H.R. (1978). Aging in the rhesus monkey: debilitating effects on short-term memory. J Gerontol *33*, 858-871.

Bartus, R.T., and Johnson, H.R. (1976). Short-term memory in the rhesus monkey: disruption from the anti-cholinergic scopolamine. Pharmacol Biochem Behav *5*, 39-46.

Beach, H., Cole, R., Gill, M.L., and Loria, J.P. (2005). Conservation of mus-ms enzyme motions in the apo- and substrate-mimicked state. J Am Chem Soc *127*, 9167-9176.

Bhattacharyya, R.P., Remenyi, A., Yeh, B.J., and Lim, W.A. (2006). Domains, motifs, and scaffolds: the role of modular interactions in the evolution and wiring of cell signaling circuits. Annu Rev Biochem *75*, 655-680.

Birtalan, S.C., Phillips, R.M., and Ghosh, P. (2002). Three-dimensional secretion signals in chaperone-effector complexes of bacterial pathogens. Mol Cell *9*, 971-980.

Borg, J.P., Ooi, J., Levy, E., and Margolis, B. (1996). The phosphotyrosine interaction domains of X11 and FE65 bind to distinct sites on the YENPTY motif of amyloid precursor protein. Mol Cell Biol *16*, 6229-6241.

Botsford, J.L., and Harman, J.G. (1992). Cyclic AMP in prokaryotes. Microbiol Rev *56*, 100-122.

Brauns, S.C., Milne, P., Naude, R., and Van de Venter, M. (2004). Selected cyclic dipeptides inhibit cancer cell growth and induce apoptosis in HT-29 colon cancer cells. Anticancer Res *24*, 1713-1719.

Brooks, H., Lebleu, B., and Vives, E. (2005). Tat peptide-mediated cellular delivery: back to basics. Adv Drug Deliv Rev *57*, 559-577.

Brown, N.R., Noble, M.E., Endicott, J.A., and Johnson, L.N. (1999). The structural basis for specificity of substrate and recruitment peptides for cyclin-dependent kinases. Nat Cell Biol *1*, 438-443.

Buttner, D., and Bonas, U. (2006). Who comes first? How plant pathogenic bacteria orchestrate type III secretion. Curr Opin Microbiol *9*, 193-200.

Buxbaum, J.D., Koo, E.H., and Greengard, P. (1993). Protein phosphorylation inhibits production of Alzheimer amyloid beta/A4 peptide. Proc Natl Acad Sci U S A *90*, 9195-9198.

Cao, X., and Sudhof, T.C. (2001). A transcriptionally [correction of transcriptively] active complex of APP with Fe65 and histone acetyltransferase Tip60. Science *293*, 115-120.

Caporaso, G.L., Takei, K., Gandy, S.E., Matteoli, M., Mundigl, O., Greengard, P., and De Camilli, P. (1994). Morphologic and biochemical analysis of the intracellular trafficking of the Alzheimer beta/A4 amyloid precursor protein. J Neurosci *14*, 3122-3138.

Carlton, P.L. (1963). Cholinergic mechanisms in the control of behavior by the brain. Psychol Rev *70*, 19-39.

Casper-Lindley, C., Dahlbeck, D., Clark, E.T., and Staskawicz, B.J. (2002). Direct biochemical evidence for type III secretion-dependent translocation of the AvrBs2 effector protein into plant cells. Proc Natl Acad Sci U S A *99*, 8336-8341.

Chaires, J.B. (2008). Calorimetry and thermodynamics in drug design. Annu Rev Biophys *37*, 135-151.

Chang, K.A., Kim, H.S., Ha, T.Y., Ha, J.W., Shin, K.Y., Jeong, Y.H., Lee, J.P., Park, C.H., Kim, S., Baik, T.K., *et al.* (2006). Phosphorylation of amyloid precursor protein (APP) at Thr668 regulates the nuclear translocation of the APP intracellular domain and induces neurodegeneration. Mol Cell Biol *26*, 4327-4338.

Chothia, C., Gough, J., Vogel, C., and Teichmann, S.A. (2003). Evolution of the protein repertoire. Science *300*, 1701-1703.

Christensen, J.J., Izatt, R.M., Hansen, L.D., and Partridge, J.A. (1966). Entropy Titration. A Calorimetric Method for the Determination of G, H, and S from a Single Thermometric Titration1a, b. The Journal of Physical Chemistry *70*, 2003-2010.

Cleary, J.P., Walsh, D.M., Hofmeister, J.J., Shankar, G.M., Kuskowski, M.A., Selkoe, D.J., and Ashe, K.H. (2005). Natural oligomers of the amyloid-beta protein specifically disrupt cognitive function. Nat Neurosci *8*, 79-84.

Cornelis, G.R. (2006). The type III secretion injectisome. Nat Rev Microbiol *4*, 811-825.

Creighton, T.E. (1993). Proteins: structures and molecular properties (WH Freeman).

Dangl, J.L., and Jones, J.D. (2001). Plant pathogens and integrated defence responses to infection. Nature *411*, 826-833.

Davies, P., and Maloney, A.J. (1976). Selective loss of central cholinergic neurons in Alzheimer's disease. Lancet *2*, 1403.

Dawson, J.E., and Nicholson, L.K. (2008). Folding kinetics and thermodynamics of Pseudomonas syringae effector protein AvrPto provide insight into translocation via the type III secretion system. Protein Sci *17*, 1109-1119.

Dawson, J.E., Seckute, J., De, S., Schueler, S.A., Oswald, A.B., and Nicholson, L.K. (2009). Elucidation of a pH-folding switch in the Pseudomonas syringae effector protein AvrPto. Proc Natl Acad Sci U S A *106*, 8543-8548.

De Strooper, B., and Annaert, W. (2000). Proteolytic processing and cell biological functions of the amyloid precursor protein. J Cell Sci *113 (Pt 11)*, 1857-1870.

De Vos, K.J., Grierson, A.J., Ackerley, S., and Miller, C.C. (2008). Role of axonal transport in neurodegenerative diseases. Annu Rev Neurosci *31*, 151-173.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR *6*, 277-293.

Deutsch, J.A. (1971). The cholinergic synapse and the site of memory. Science *174*, 788-794.

Dickson, D.W. (1997). Neuropathological diagnosis of Alzheimer's disease: a perspective from longitudinal clinicopathological studies. Neurobiol Aging *18*, S21-26.

Dorman, G., and Prestwich, G.D. (1994). Benzophenone photophores in biochemistry. Biochemistry *33*, 5661-5673.

Duilio, A., Faraonio, R., Minopoli, G., Zambrano, N., and Russo, T. (1998). Fe65L2: a new member of the Fe65 protein family interacting with the intracellular domain of the Alzheimer's beta-amyloid precursor protein. Biochem J *330 (Pt 1)*, 513-519.

Eckert, B., Martin, A., Balbach, J., and Schmid, F.X. (2005). Prolyl isomerization as a molecular timer in phage infection. Nat Struct Mol Biol *12*, 619-623.

Estiarte, M., Diez, A., Rubiralta, M., and Jackson, R.F.W. (2001). Synthesis of a 3-aminopiperidin-2, 5-dione as a conformationally constrained surrogate of the Ala-Gly dipeptide. Tetrahedron *57*, 157-161.

Faden, A.I., Knoblach, S.M., Cernak, I., Fan, L., Vink, R., Araldi, G.L., Fricke, S.T., Roth, B.L., and Kozikowski, A.P. (2003). Novel diketopiperazine enhances motor and cognitive recovery after traumatic brain injury in rats and shows neuroprotection in vitro and in vivo. J Cereb Blood Flow Metab *23*, 342-354.

Fanghanel, J. (2003). Enzymatic catalysis of the peptidyl-prolyl bond rotation: are transition state formation and enzyme dynamics directly linked? Angew Chem Int Ed Engl *42*, 490-492.

Fanghanel, J., Wawra, S., Lucke, C., Wildemann, D., and Fischer, G. (2006). Isothermal calorimetry as a tool to investigate slow conformational changes in proteins and peptides. Anal Chem *78*, 4517-4523.

Feldman, M.F., Muller, S., Wuest, E., and Cornelis, G.R. (2002). SycE allows secretion of YopE-DHFR hybrids by the Yersinia enterocolitica type III Ysc system. Mol Microbiol *46*, 1183-1197.

Fersht, A.R. (1974). Catalysis, binding and enzyme-substrate complementarity. Proc R Soc Lond B Biol Sci *187*, 397-407.

Fischer, G. (2000). Chemical aspects of peptide bond isomerisation. Chemical Society Reviews *29*, 119-127.

Freyer, M.W., and Lewis, E.A. (2008). Isothermal titration calorimetry: experimental design, data analysis, and probing macromolecule/ligand binding and kinetic interactions. Methods in cell biology *84*, 79-113.

Frolich, L. (2002). The cholinergic pathology in Alzheimer's disease--discrepancies between clinical experience and pathophysiological findings. J Neural Transm *109*, 1003-1013.

Galan, J.E., and Collmer, A. (1999). Type III secretion machines: bacterial devices for protein delivery into host cells. Science *284*, 1322-1328.

Galan, J.E., and Wolf-Watz, H. (2006). Protein delivery into eukaryotic cells by type III secretion machines. Nature *444*, 567-573.

Ghosh, P. (2004). Process of protein transport by the type III secretion system. Microbiol Mol Biol Rev *68*, 771-795.

Glenner, G.G., and Wong, C.W. (1984a). Alzheimer's disease and Down's syndrome: sharing of a unique cerebrovascular amyloid fibril protein. Biochem Biophys Res Commun *122*, 1131-1135.

Glenner, G.G., and Wong, C.W. (1984b). Alzheimer's disease: initial report of the purification and characterization of a novel cerebrovascular amyloid protein. Biochem Biophys Res Commun *120*, 885-890.

Goddard, T.D., Kneller, D.G. (2008). Sparky 3, University of California, San Franscisco.

Goldgaber, D., Lerman, M.I., McBride, O.W., Saffiotti, U., and Gajdusek, D.C. (1987). Characterization and chromosomal localization of a cDNA encoding brain amyloid of Alzheimer's disease. Science *235*, 877-880.

Greenfield, J.P., Tsai, J., Gouras, G.K., Hai, B., Thinakaran, G., Checler, F., Sisodia, S.S., Greengard, P., and Xu, H. (1999). Endoplasmic reticulum and trans-Golgi network generate distinct populations of Alzheimer beta-amyloid peptides. Proc Natl Acad Sci U S A *96*, 742-747.

Grignon, C., and Sentenac, H. (1991). pH and ionic conditions in the apoplast. Annual Review of Plant Biology *42*, 103-128.

Gunther, U.L., and Schaffhausen, B. (2002). NMRKIN: simulating line shapes from two-dimensional spectra of proteins upon ligand binding. J Biomol NMR *22*, 201-209.

Haass, C., Hung, A.Y., Schlossmacher, M.G., Oltersdorf, T., Teplow, D.B., and Selkoe, D.J. (1993a). Normal cellular processing of the beta-amyloid precursor protein results in the secretion of the amyloid beta peptide and related molecules. Ann N Y Acad Sci *695*, 109-116.

Haass, C., Hung, A.Y., Schlossmacher, M.G., Teplow, D.B., and Selkoe, D.J. (1993b). beta-Amyloid peptide and a 3-kDa fragment are derived by distinct cellular mechanisms. J Biol Chem *268*, 3021-3024.

Hamdane, M., Dourlen, P., Bretteville, A., Sambo, A.V., Ferreira, S., Ando, K., Kerdraon, O., Begard, S., Geay, L., Lippens, G., *et al.* (2006). Pin1 allows for differential Tau dephosphorylation in neuronal cells. Mol Cell Neurosci *32*, 155-160.

Handschumacher, R.E., Harding, M.W., Rice, J., Drugge, R.J., and Speicher, D.W. (1984). Cyclophilin: a specific cytosolic binding protein for cyclosporin A. Science *226*, 544-547.

Harding, M.W., Galat, A., Uehling, D.E., and Schreiber, S.L. (1989). A receptor for the immunosuppressant FK506 is a cis-trans peptidyl-prolyl isomerase. Nature *341*, 758-760.

Hardy, J., and Selkoe, D.J. (2002). The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. Science *297*, 353-356.

Hartmann, T., Bieger, S.C., Bruhl, B., Tienari, P.J., Ida, N., Allsop, D., Roberts, G.W., Masters, C.L., Dotti, C.G., Unsicker, K., *et al.* (1997). Distinct sites of intracellular production for Alzheimer's disease A beta40/42 amyloid peptides. Nat Med *3*, 1016-1020.

Horovitz, A., Serrano, L., Avron, B., Bycroft, M., and Fersht, A.R. (1990). Strength and co-operativity of contributions of surface salt bridges to protein stability. J Mol Biol *216*, 1031-1044.

Huang, Q., Aluise, C.D., Joshi, G., Sultana, R., St Clair, D.K., Markesbery, W.R., and Butterfield, D.A. (2010). Potential in vivo amelioration by N-acetyl-L-cysteine of oxidative stress in brain in human double mutant APP/PS-1 knock-in mice: toward therapeutic modulation of mild cognitive impairment. J Neurosci Res *88*, 2618-2629.

Hwang, Y., and Cole, P.A. (2004). Efficient synthesis of phosphorylated prodrugs with bis(POM)-phosphoryl chloride. Org Lett *6*, 1555-1556.

Iijima, K., Ando, K., Takeda, S., Satoh, Y., Seki, T., Itohara, S., Greengard, P., Kirino, Y., Nairn, A.C., and Suzuki, T. (2000). Neuron-specific phosphorylation of Alzheimer's beta-amyloid precursor protein by cyclin-dependent kinase 5. J Neurochem *75*, 1085-1091.

Jacobs, D.M., Saxena, K., Vogtherr, M., Bernado, P., Pons, M., and Fiebig, K.M. (2003). Peptide binding induces large scale changes in inter-domain mobility in human Pin1. J Biol Chem *278*, 26174-26182.

Jin, J., Xie, X., Chen, C., Park, J.G., Stark, C., James, D.A., Olhovsky, M., Linding, R., Mao, Y., and Pawson, T. (2009). Eukaryotic protein domains as functional units of cellular evolution. Sci Signal *2*, ra76.

Kamal, A., Stokin, G.B., Yang, Z., Xia, C.H., and Goldstein, L.S. (2000). Axonal transport of amyloid precursor protein is mediated by direct binding to the kinesin light chain subunit of kinesin-I. Neuron *28*, 449-459.

Kang, J., Lemaire, H.G., Unterbeck, A., Salbaum, J.M., Masters, C.L., Grzeschik, K.H., Multhaup, G., Beyreuther, K., and Muller-Hill, B. (1987). The precursor of Alzheimer's disease amyloid A4 protein resembles a cell-surface receptor. Nature *325*, 733-736.

Kay, L.E. (2005). NMR studies of protein structure and dynamics. J Magn Reson *173*, 193-207.

Kern, D., Kern, G., Scherer, G., Fischer, G., and Drakenberg, T. (1995). Kinetic analysis of cyclophilin-catalyzed prolyl cis/trans isomerization by dynamic NMR spectroscopy. Biochemistry *34*, 13594-13602.

Kim, H.S., Kim, E.M., Lee, J.P., Park, C.H., Kim, S., Seo, J.H., Chang, K.A., Yu, E., Jeong, S.J., Chong, Y.H., *et al.* (2003). C-terminal fragments of amyloid precursor protein exert neurotoxicity by inducing glycogen synthase kinase-3beta expression. FASEB J *17*, 1951-1953.

Kitaguchi, N., Takahashi, Y., Tokushima, Y., Shiojiri, S., and Ito, H. (1988). Novel precursor of Alzheimer's disease amyloid protein shows protease inhibitory activity. Nature *331*, 530-532.

Klein, W.L., Krafft, G.A., and Finch, C.E. (2001). Targeting small Abeta oligomers: the solution to an Alzheimer's disease conundrum? Trends Neurosci *24*, 219-224.

Korzhnev, D.M., and Kay, L.E. (2008). Probing invisible, low-populated States of protein molecules by relaxation dispersion NMR spectroscopy: an application to protein folding. Acc Chem Res *41*, 442-451.

Kovacs, D.M., Fausett, H.J., Page, K.J., Kim, T.W., Moir, R.D., Merriam, D.E., Hollister, R.D., Hallmark, O.G., Mancini, R., Felsenstein, K.M., *et al.* (1996). Alzheimer-associated presenilins 1 and 2: neuronal expression in brain and localization to intracellular membranes in mammalian cells. Nat Med *2*, 224-229.

Kovrigin, E.L., and Loria, J.P. (2006). Enzyme dynamics along the reaction coordinate: critical role of a conserved residue. Biochemistry *45*, 2636-2647.

Labeikovsky, W., Eisenmesser, E.Z., Bosco, D.A., and Kern, D. (2007). Structure and dynamics of pin1 during catalysis by NMR. J Mol Biol *367*, 1370-1381.

Lazarov, O., Morfini, G.A., Pigino, G., Gadadhar, A., Chen, X., Robinson, J., Ho, H., Brady, S.T., and Sisodia, S.S. (2007). Impairments in fast axonal transport and motor neuron deficits in transgenic mice expressing familial Alzheimer's disease-linked mutant presenilin 1. J Neurosci *27*, 7011-7020.

Lee, C.W., Lau, K.F., Miller, C.C., and Shaw, P.C. (2003a). Glycogen synthase kinase-3 beta-mediated tau phosphorylation in cultured cell lines. Neuroreport *14*, 257-260.

Lee, M.S., Kao, S.C., Lemere, C.A., Xia, W., Tseng, H.C., Zhou, Y., Neve, R., Ahlijanian, M.K., and Tsai, L.H. (2003b). APP processing is regulated by cytoplasmic phosphorylation. J Cell Biol *163*, 83-95.

Lee, V.T., and Schneewind, O. (2002). Yop fusions to tightly folded protein domains and their effects on Yersinia enterocolitica type III secretion. J Bacteriol *184*, 3740-3745.

Leroy, K., Boutajangout, A., Authelet, M., Woodgett, J.R., Anderton, B.H., and Brion, J.P. (2002). The active form of glycogen synthase kinase-3beta is associated with granulovacuolar degeneration in neurons in Alzheimer's disease. Acta Neuropathol *103*, 91-99.

Lewis, E.A., Murphy, K.P. (2005). Protein-ligand interactions: methods and applications, Vol 305 (Humana Pr Inc).

Li, S., Shankar, G.M., and Selkoe, D.J. (2010). How do Soluble Oligomers of Amyloid beta-protein Impair Hippocampal Synaptic Plasticity? Front Cell Neurosci *4*, 5.

Liou, Y.C., Sun, A., Ryo, A., Zhou, X.Z., Yu, Z.X., Huang, H.K., Uchida, T., Bronson, R., Bing, G., Li, X., *et al.* (2003). Role of the prolyl isomerase Pin1 in protecting against age-dependent neurodegeneration. Nature *424*, 556-561.

Loria, J.P., Berlow, R.B., and Watt, E.D. (2008). Characterization of enzyme motions by solution NMR relaxation dispersion. Acc Chem Res *41*, 214-221.

Lu, K.P., Hanes, S.D., and Hunter, T. (1996). A human peptidyl-prolyl isomerase essential for regulation of mitosis. Nature *380*, 544-547.

Lu, K.P., Liou, Y.C., and Zhou, X.Z. (2002a). Pinning down proline-directed phosphorylation signaling. Trends Cell Biol *12*, 164-172.

Lu, K.P., Suizu, F., Zhou, X.Z., Finn, G., Lam, P., and Wulf, G. (2006). Targeting carcinogenesis: a role for the prolyl isomerase Pin1? Mol Carcinog *45*, 397-402.

Lu, K.P., and Zhou, X.Z. (2007). The prolyl isomerase PIN1: a pivotal new twist in phosphorylation signalling and disease. Nat Rev Mol Cell Biol *8*, 904-916.

Lu, P.J., Wulf, G., Zhou, X.Z., Davies, P., and Lu, K.P. (1999a). The prolyl isomerase Pin1 restores the function of Alzheimer-associated phosphorylated tau protein. Nature *399*, 784-788.

Lu, P.J., Zhou, X.Z., Liou, Y.C., Noel, J.P., and Lu, K.P. (2002b). Critical role of WW domain phosphorylation in regulating phosphoserine binding activity and Pin1 function. Journal of Biological Chemistry *277*, 2381.

Lu, P.J., Zhou, X.Z., Shen, M., and Lu, K.P. (1999b). Function of WW domains as phosphoserine-or phosphothreonine-binding modules. Science *283*, 1325.

Mangialasche, F., Solomon, A., Winblad, B., Mecocci, P., and Kivipelto, M. (2010). Alzheimer's disease: clinical trials and drug development. Lancet Neurol *9*, 702-716.

Marquardt, D.W. (1963). An algorithm for least-squares of nonlinear parameters. J Soc Appl Ind Math *11*, 431–441.

Massi, F., Johnson, E., Wang, C., Rance, M., and Palmer, A.G., 3rd (2004). NMR R1 rho rotating-frame relaxation with weak radio frequency fields. J Am Chem Soc *126*, 2247-2256.

Masters, C.L., Simms, G., Weinman, N.A., Multhaup, G., McDonald, B.L., and Beyreuther, K. (1985). Amyloid plaque core protein in Alzheimer disease and Down syndrome. Proc Natl Acad Sci U S A *82*, 4245-4249.

Mattson, M.P. (2004). Pathways towards and away from Alzheimer's disease. Nature *430*, 631-639.

Miller, C.C., McLoughlin, D.M., Lau, K.F., Tennant, M.E., and Rogelj, B. (2006). The X11 proteins, Abeta production and Alzheimer's disease. Trends Neurosci *29*, 280-285.

Minamino, T., Imae, Y., Oosawa, F., Kobayashi, Y., and Oosawa, K. (2003). Effect of intracellular pH on rotational speed of bacterial flagellar motors. J Bacteriol *185*, 1190-1194.

Mudgett, M.B. (2005). New insights to the function of phytopathogenic bacterial type III effectors in plants. Annu Rev Plant Biol *56*, 509-531.

Mulder, F.A., Schipper, D., Bott, R., and Boelens, R. (1999). Altered flexibility in the substrate-binding site of related native and engineered high-alkaline Bacillus subtilisins. J Mol Biol *292*, 111-123.

Muller, T., Meyer, H.E., Egensperger, R., and Marcus, K. (2008). The amyloid precursor protein intracellular domain (AICD) as modulator of gene expression, apoptosis, and cytoskeletal dynamics-relevance for Alzheimer's disease. Prog Neurobiol *85*, 393-406.

Nicholson, L.K., and Lu, K.P. (2007). Prolyl cis-trans Isomerization as a molecular timer in Crk signaling. Mol Cell *25*, 483-485.

Nordstedt, C., Caporaso, G.L., Thyberg, J., Gandy, S.E., and Greengard, P. (1993). Identification of the Alzheimer beta/A4 amyloid precursor protein in clathrin-coated vesicles purified from PC12 cells. J Biol Chem *268*, 608-612.

Novak, B., Kapuy, O., Domingo-Sananes, M.R., and Tyson, J.J. (2010). Regulated protein kinases and phosphatases in cell cycle decisions. Curr Opin Cell Biol.

Nozaki, Y., and Tanford, C. (1967). Examination of titration behavior. Methods Enzymol *11*, 715-734.

Oishi, M., Nairn, A.C., Czernik, A.J., Lim, G.S., Isohara, T., Gandy, S.E., Greengard, P., and Suzuki, T. (1997). The cytoplasmic domain of Alzheimer's amyloid precursor protein is phosphorylated at Thr654, Ser655, and Thr668 in adult rat brain and cultured cells. Mol Med *3*, 111-123.

Olson, M.I., and Shaw, C.M. (1969). Presenile dementia and Alzheimer's disease in mongolism. Brain *92*, 147-156.

Palmer, A.G., 3rd (2004). NMR characterization of the dynamics of biomacromolecules. Chem Rev *104*, 3623-3640.

Palmer, A.G., 3rd, Grey, M.J., and Wang, C. (2005). Solution NMR spin relaxation methods for characterizing chemical exchange in high-molecular-weight systems. Methods Enzymol *394*, 430-465.

Papin, J.A., Hunter, T., Palsson, B.O., and Subramaniam, S. (2005). Reconstruction of cellular signalling networks and analysis of their properties. Nat Rev Mol Cell Biol *6*, 99-111.

Pardossi-Piquard, R., Petit, A., Kawarai, T., Sunyach, C., Alves da Costa, C., Vincent, B., Ring, S., D'Adamio, L., Shen, J., Muller, U., *et al.* (2005). Presenilin-dependent transcriptional control of the Abeta-degrading enzyme neprilysin by intracellular domains of betaAPP and APLP. Neuron *46*, 541-554.

Pastorino, L., and Lu, K.P. (2006). Pathogenic mechanisms in Alzheimer's disease. Eur J Pharmacol *545*, 29-38.

Pastorino, L., Sun, A., Lu, P.J., Zhou, X.Z., Balastik, M., Finn, G., Wulf, G., Lim, J., Li, S.H., Li, X., *et al.* (2006). The prolyl isomerase Pin1 regulates amyloid precursor protein processing and amyloid-beta production. Nature *440*, 528-534.

Pawson, T., and Nash, P. (2003). Assembly of cell regulatory systems through protein interaction domains. Science *300*, 445-452.

Penna, T.C., Ishii, M., Junior, A.P., and Cholewa, O. (2004). Thermal stability of recombinant green fluorescent protein (GFPuv) at various pH values. Appl Biochem Biotechnol *113-116*, 469-483.

Perez, R.G., Soriano, S., Hayes, J.D., Ostaszewski, B., Xia, W., Selkoe, D.J., Chen, X., Stokin, G.B., and Koo, E.H. (1999). Mutagenesis identifies new signals for beta-amyloid precursor protein endocytosis, turnover, and the generation of secreted fragments, including Abeta42. J Biol Chem *274*, 18851-18856.

Phiel, C.J., Wilson, C.A., Lee, V.M., and Klein, P.S. (2003). GSK-3alpha regulates production of Alzheimer's disease amyloid-beta peptides. Nature *423*, 435-439.

Pietrzik, C.U., Yoon, I.S., Jaeger, S., Busse, T., Weggen, S., and Koo, E.H. (2004). FE65 constitutes the functional link between the low-density lipoprotein receptor-related protein and the amyloid precursor protein. J Neurosci *24*, 4259-4265.

Prasad, C. (1995). Bioactive cyclic dipeptides. Peptides *16*, 151-164.

Press, W.H. (1988). Numerical recipes in C : the art of scientific computing (Cambridge ; New York, Cambridge University Press).

Rahfeld, J.U., Schierhorn, A., Mann, K., and Fischer, G. (1994). A novel peptidyl-prolyl cis/trans isomerase from Escherichia coli. FEBS Lett *343*, 65-69.

Ramakrishnan, P., Dickson, D.W., and Davies, P. (2003). Pin1 colocalization with phosphorylated tau in Alzheimer's disease and other tauopathies. Neurobiol Dis *14*, 251-264.

Ramelot, T.A., Gentile, L.N., and Nicholson, L.K. (2000). Transient structure of the amyloid precursor protein cytoplasmic tail indicates preordering of structure for binding to cytosolic factors. Biochemistry *39*, 2714-2725.

Ramelot, T.A., and Nicholson, L.K. (2001). Phosphorylation-induced structural changes in the amyloid precursor protein cytoplasmic tail detected by NMR. J Mol Biol *307*, 871-884.

Ranganathan, R., Lu, K.P., Hunter, T., and Noel, J.P. (1997). Structural and functional analysis of the mitotic rotamase Pin1 suggests substrate recognition is phosphorylation dependent. Cell *89*, 875-886.

Rao, B.D. (1989). Nuclear magnetic resonance line-shape analysis and determination of exchange rates. Methods Enzymol *176*, 279-311.

Reimer, U., Scherer, G., Drewello, M., Kruber, S., Schutkowski, M., and Fischer, G. (1998). Side-chain effects on peptidyl-prolyl cis/trans isomerisation. J Mol Biol *279*, 449-460.

Robakis, N.K., Ramakrishna, N., Wolfe, G., and Wisniewski, H.M. (1987). Molecular cloning and characterization of a cDNA encoding the cerebrovascular and the neuritic plaque amyloid peptides. Proc Natl Acad Sci U S A *84*, 4190-4194.

Rutschow, S., Thiem, J., Kranz, C., and Marquardt, T. (2002). Membrane-permeant derivatives of mannose-1-phosphate. Bioorg Med Chem *10*, 4043-4049.

Saier, M.H., Jr. (2004). Evolution of bacterial type III protein secretion systems. Trends Microbiol *12*, 113-115.

Sannerud, R., and Annaert, W. (2009). Trafficking, a key player in regulated intramembrane proteolysis. Semin Cell Dev Biol *20*, 183-190.

Sarkar, P., Reichman, C., Saleh, T., Birge, R.B., and Kalodimos, C.G. (2007). Proline cis-trans isomerization controls autoinhibition of a signaling protein. Mol Cell *25*, 413-426.

Sastre, M., Turner, R.S., and Levy, E. (1998). X11 interaction with beta-amyloid precursor protein modulates its cellular stabilization and reduces amyloid beta-protein secretion. J Biol Chem *273*, 22351-22357.

Schutkowski, M., Bernhardt, A., Zhou, X.Z., Shen, M., Reimer, U., Rahfeld, J.U., Lu, K.P., and Fischer, G. (1998). Role of phosphorylation in determining the backbone dynamics of the serine/threonine-proline motif and Pin1 substrate recognition. Biochemistry *37*, 5566-5575.

Scofield, S.R., Tobias, C.M., Rathjen, J.P., Chang, J.H., Lavelle, D.T., Michelmore, R.W., and Staskawicz, B.J. (1996). Molecular Basis of Gene-for-Gene Specificity in Bacterial Speck Disease of Tomato. Science *274*, 2063-2065.

Selkoe, D.J. (2002). Alzheimer's disease is a synaptic failure. Science *298*, 789-791.

Selkoe, D.J. (2008). Soluble oligomers of the amyloid beta-protein impair synaptic plasticity and behavior. Behav Brain Res *192*, 106-113.

Shah, S., Lee, S.F., Tabuchi, K., Hao, Y.H., Yu, C., LaPlant, Q., Ball, H., Dann, C.E., 3rd, Sudhof, T., and Yu, G. (2005). Nicastrin functions as a gamma-secretase-substrate receptor. Cell *122*, 435-447.

Sheehan, J.C., and Hess, G.P. (1955). A new method of forming peptide bonds. Journal of the American Chemical Society *77*, 1067-1068.

Shin, R.W., Ogino, K., Shimabuku, A., Taki, T., Nakashima, H., Ishihara, T., and Kitamoto, T. (2007). Amyloid precursor protein cytoplasmic domain with phospho-Thr668 accumulates in Alzheimer's disease and its transgenic models: a role to mediate interaction of Abeta and tau. Acta Neuropathol *113*, 627-636.

Sinha, S., Anderson, J.P., Barbour, R., Basi, G.S., Caccavello, R., Davis, D., Doan, M., Dovey, H.F., Frigon, N., Hong, J., *et al.* (1999). Purification and cloning of amyloid precursor protein beta-secretase from human brain. Nature *402*, 537-540.

Sisodia, S.S. (1992). Beta-amyloid precursor protein cleavage by a membrane-bound protease. Proc Natl Acad Sci U S A *89*, 6075-6079.

Sodhi, C.P., Perez, R.G., and Gottardi-Littell, N.R. (2008). Phosphorylation of beta-amyloid precursor protein (APP) cytoplasmic tail facilitates amyloidogenic processing during apoptosis. Brain Res *1198*, 204-212.

Sory, M.P., and Cornelis, G.R. (1994). Translocation of a hybrid YopE-adenylate cyclase from Yersinia enterocolitica into HeLa cells. Mol Microbiol *14*, 583-594.

Stokin, G.B., Lillo, C., Falzone, T.L., Brusch, R.G., Rockenstein, E., Mount, S.L., Raman, R., Davies, P., Masliah, E., Williams, D.S., *et al.* (2005). Axonopathy and transport deficits early in the pathogenesis of Alzheimer's disease. Science *307*, 1282-1288.

Sudol, M. (1996). Structure and function of the WW domain. Prog Biophys Mol Biol *65*, 113-132.

Tanaka, S., Shiojiri, S., Takahashi, Y., Kitaguchi, N., Ito, H., Kameyama, M., Kimura, J., Nakamura, S., and Ueda, K. (1989). Tissue-specific expression of three types of beta-protein precursor mRNA: enhancement of protease inhibitor-harboring types in Alzheimer's disease brain. Biochem Biophys Res Commun *165*, 1406-1414.

Tanzi, R.E., Gusella, J.F., Watkins, P.C., Bruns, G.A., St George-Hyslop, P., Van Keuren, M.L., Patterson, D., Pagan, S., Kurnit, D.M., and Neve, R.L. (1987). Amyloid beta protein gene: cDNA, mRNA distribution, and genetic linkage near the Alzheimer locus. Science *235*, 880-884.

Thorpe, J.R., Mosaheb, S., Hashemzadeh-Bonehi, L., Cairns, N.J., Kay, J.E., Morley, S.J., and Rulten, S.L. (2004). Shortfalls in the peptidyl-prolyl cis-trans isomerase protein Pin1 in neurons are associated with frontotemporal dementias. Neurobiol Dis *17*, 237-249.

Todd, M.J., and Gomez, J. (2001). Enzyme kinetics determined using calorimetry: a general assay for enzyme activity? Anal Biochem *296*, 179-187.

Turner, P.R., O'Connor, K., Tate, W.P., and Abraham, W.C. (2003). Roles of amyloid precursor protein and its fragments in regulating neural activity, plasticity and memory. Prog Neurobiol *70*, 1-32.

Vassar, R., Bennett, B.D., Babu-Khan, S., Kahn, S., Mendiaz, E.A., Denis, P., Teplow, D.B., Ross, S., Amarante, P., Loeloff, R., *et al.* (1999). Beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE. Science *286*, 735-741.

Verdecia, M.A., Bowman, M.E., Lu, K.P., Hunter, T., and Noel, J.P. (2000). Structural basis for phosphoserine-proline recognition by group IV WW domains. Nat Struct Biol *7*, 639-643.

von Rotz, R.C., Kohli, B.M., Bosset, J., Meier, M., Suzuki, T., Nitsch, R.M., and Konietzko, U. (2004). The APP intracellular domain forms nuclear multiprotein complexes and regulates the transcription of its own precursor. J Cell Sci *117*, 4435-4448.

Weiwad, M., Kullertz, G., Schutkowski, M., and Fischer, G. (2000). Evidence that the substrate backbone conformation is critical to phosphorylation by p42 MAP kinase. FEBS Lett *478*, 39-42.

Wilcock, G.K., Esiri, M.M., Bowen, D.M., and Smith, C.C. (1982). Alzheimer's disease. Correlation of cortical choline acetyltransferase activity with the severity of dementia and histological abnormalities. J Neurol Sci *57*, 407-417.

Wilharm, G., Lehmann, V., Krauss, K., Lehnert, B., Richter, S., Ruckdeschel, K., Heesemann, J., and Trulzsch, K. (2004). Yersinia enterocolitica type III secretion depends on the proton motive force but not on the flagellar motor components MotA and MotB. Infect Immun *72*, 4004-4009.

Wilkinson, D.G., Francis, P.T., Schwam, E., and Payne-Parrish, J. (2004). Cholinesterase inhibitors used in the treatment of Alzheimer's disease: the relationship between pharmacological effects and clinical efficacy. Drugs Aging *21*, 453-478.

Wisniewski, T.L., Mendel, E., Morse, E.V., Hilton, C.W., and Svec, F. (1994). Relationship between serum cyclo(His-Pro) concentrations and the nutritional status of HIV-infected patients. South Med J *87*, 348-351.

Wolfe, M.S., Xia, W., Ostaszewski, B.L., Diehl, T.S., Kimberly, W.T., and Selkoe, D.J. (1999). Two transmembrane aspartates in presenilin-1 required for presenilin endoproteolysis and gamma-secretase activity. Nature *398*, 513-517.

Wulf, G., Finn, G., Suizu, F., and Lu, K.P. (2005). Phosphorylation-specific prolyl isomerization: is there an underlying theme? Nat Cell Biol *7*, 435-441.

Wulf, J., Pascuzzi, P.E., Fahmy, A., Martin, G.B., and Nicholson, L.K. (2004). The solution structure of type III effector protein AvrPto reveals conformational and dynamic features important for plant pathogenesis. Structure *12*, 1257-1268.

Wuthrich, K. (1990). Protein structure determination in solution by NMR spectroscopy. J Biol Chem *265*, 22059-22062.

Xu, G.G., and Etzkorn, F.A. (2009). Pin1 as an anticancer drug target. Drug News Perspect *22*, 399.

Xu, H., Sweeney, D., Wang, R., Thinakaran, G., Lo, A.C., Sisodia, S.S., Greengard, P., and Gandy, S. (1997). Generation of Alzheimer beta-amyloid protein in the trans-Golgi network in the apparent absence of vesicle formation. Proc Natl Acad Sci U S A *94*, 3748-3752.

Yan, R., Bienkowski, M.J., Shuck, M.E., Miao, H., Tory, M.C., Pauley, A.M., Brashier, J.R., Stratman, N.C., Mathews, W.R., Buhl, A.E., *et al.* (1999). Membrane-anchored aspartyl protease with Alzheimer's disease beta-secretase activity. Nature *402*, 533-537.

Yang, A.S., and Honig, B. (1993). On the pH dependence of protein stability. J Mol Biol *231*, 459-474.

Zambrano, N., Bruni, P., Minopoli, G., Mosca, R., Molino, D., Russo, C., Schettini, G., Sudol, M., and Russo, T. (2001). The beta-amyloid precursor protein APP is tyrosine-phosphorylated in cells expressing a constitutively active form of the Abl protoncogene. J Biol Chem *276*, 19787-19792.

Zambrano, N., Minopoli, G., de Candia, P., and Russo, T. (1998). The Fe65 adaptor protein interacts through its PID1 domain with the transcription factor CP2/LSF/LBP1. J Biol Chem *273*, 20128-20133.

Zhang, B., Higuchi, M., Yoshiyama, Y., Ishihara, T., Forman, M.S., Martinez, D., Joyce, S., Trojanowski, J.Q., and Lee, V.M. (2004). Retarded axonal transport of R406W mutant tau in transgenic mice with a neurodegenerative tauopathy. J Neurosci *24*, 4657-4667.

Zhang, Y., Daum, S., Wildemann, D., Zhou, X.Z., Verdecia, M.A., Bowman, M.E., Lucke, C., Hunter, T., Lu, K.P., Fischer, G., *et al.* (2007). Structural basis for high-affinity peptide inhibition of human Pin1. ACS Chem Biol *2*, 320-328.

Zhang, Y.W., and Xu, H. (2007). Molecular and cellular mechanisms for Alzheimer's disease: understanding APP metabolism. Curr Mol Med *7*, 687-696.

Zhao, S., and Etzkorn, F.A. (2007). A phosphorylated prodrug for the inhibition of Pin1. Bioorg Med Chem Lett *17*, 6615-6618.

Zhou, X.Z., Kops, O., Werner, A., Lu, P.J., Shen, M., Stoller, G., Kullertz, G., Stark, M., Fischer, G., and Lu, K.P. (2000). Pin1-dependent prolyl isomerization regulates dephosphorylation of Cdc25C and tau proteins. Mol Cell *6*, 873-883.