KINASE-INTERACTING FHA DOMAIN OF KINASE ASSOCIATED PROTEIN PHOSPHATASE: PHOSPHOPEPTIDE INTERACTIONS AND NMR-DETECTED DYNAMICS

A Dissertation

presented to

the Faculty of the Graduate School

University of Missouri-Columbia

In Partial Fulfillment

of the requirements for the Degree
Doctor of Philosophy

---

by

ZHAOFENG DING

Dr. Steven Van Doren, Dissertation Supervisor

MAY 2007

The undersigned, appointed by the Dean of the Graduate Scholl, have examined the dissertation entitled

KINASE-INTERACTING FHA DOMAIN OF KINASE ASSOCIATED PROTEIN PHOSPHATASE: PHOSPHOPEPTIDE INTERACTIONS AND NMR-DETECTED DYNAMICS

presented by Zhaofeng Ding

a candidate for the degree of Doctor of Philosophy

and hereby certify that in their opinion it is worthy of acceptance.

Dr. Steven R. Van Doren

Dr. John C. Walker

Dr. Michael T. Henzl

Dr. Xiaoqin Zou

Dr. John J. Tanner

ACKNOWLEDGMENTS

My graduate research has been conducted under the supervision of Dr. Steven R. Van Doren. He is a very patient, careful, and knowledgeable advisor. I sincerely thank him for his insightful mentoring which has carried me through the challenges of my graduate study. I would also like to thank the rest of my committee: Dr. John Walker, Dr. Michael Henzl, Dr. Xiaoqin Zou and Dr. John Tanner. I owe them much gratitude for their support during the past five years.

I have been in collaborations with many wonderful scientists, both in our laboratory or from other groups. I would like to thank my advisor, Dr. Steven Van Doren, and my collaborators, Dr. Xiangyang Liang and Dr. Gui-in Lee for sharing their relaxation data of free KI-FHA with me. I would like to give my special thanks to Dr. Michael Henzl and Dr. Sayeh Agah for their assistance in conducting isothermal titration calorimetry experiments; to Dr. Huachun Wang, Clayton Larue and Dr. Jiangqi Wen for their help in cloning of receptor-like protein kinases and performing western assays; to Dr. Xuelu Wang and Dr. Scott Peck for their generous gifts of RLK constructs; to Jason Doke for editing the draft of my dissertation; to Dr. Janine Crane for her guidance on MTS labeling and to Dr. Gitta Coaker for collaborating with me on the AvrRpt2 project. I am grateful to my best friend, Zhongji Liao, and my coworkers, Dr. Michael Riley, Dr. Gabio Gallazzi, Dr. Rajagopalan Bhaskaran, Dr. Xin Liu, Dr. Ravi Garimella, Dr. Anu Shende, Dr. Arunima, Mark Palmier, Nusayba Bagegni, and Jessica Blount for sharing their research experience, as well as happiness and fun, with me.

The degree of Doctor of Philosophy is priceless. I would like to share it with my family members oversea. They gave me life and supported my twenty-one years of education, from learning my ABC's all the way to the graduate program of Biochemistry at University of Missouri-Columbia, an excellent department located half way around the globe from my hometown. They send me greetings every week, reminding me to be strong, smart, persistent and confident from time to time. Without their support, I would never have been able to survive those difficult periods nor accomplish the highest academic degree.

This research has been supported by NSF grant MCB0111589, and a Monsanto-Interdisciplinary Plant Group Graduate fellowship. The Varian Inova 600 was purchased with funding from NSF grant DBI0070359, University of Missouri, and NIH grant GM52289 towards the cryogenic probe.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AIC	Akaike`s information criterion
BAK1	BRI1-associated kinase 1
BRI1	Brassinosteroid-insensitve 1
CHFr	checkpoint with FHA and ring finger protein
CLV1	Clavata1
ET	Evolutionary Trace analysis
FHA	forkhead associated
FID	free induction decay
FLS2	flagellin sensing 2
GST	glutathione-S-transferase
HPLC	high pressure liquid chromatography
HSQC	heteronuclear single quantum correlation
ITC	isothermal titration calorimetry
KAPP	kinase-associated protein phosphatase
KI-FHA	kinase-interacting FHA domain of KAPP

LRR	leucine rich repeats
MBP	maltose binding protein
NMR	nuclear magntic resonance
NOE	nuclear Overhauser effect
PBS	phosphate-buffered saline
PP2C	protein phosphatase, type 2C
R1	spin-lattice or longitudinal relaxation rate
R2	spin-spin or transverse relaxation rate
RAD53	radiation sensitive 53
RDC	residual dipolar coupling
RF	radio-frequency
RLK	receptor-like protein kinase
RPK	receptor protein kinase
SERK	somatic-embryogenesis receptor kinase
SVD	singular value decomposition
WAK	wall associated kinase1

KINASE-INTERACTING FHA DOMAIN OF KINASE ASSOCIATED PROTEIN PHOSPHATASE: PHOSPHOPEPTIDE INTERACTIONS AND NMR-DETECTED DYNAMICS

Zhaofeng Ding

Dr. Steven R. Van Doren, Dissertation Supervisor

ABSTRACT

FHA domains are phosphoThr recognition modules found in diverse signaling proteins. Kinase-associated protein phosphatase (KAPP) from the model plant Arabidopsis employs its FHA domain in its negative regulation of some receptor-like kinase (RLK) signaling pathways important in plant development and environmental response. The interactions between the kinase-interacting FHA (KI-FHA) domain of KAPP and RLK kinase domains were investigated by multiple biophysical and biochemical methods. To identify potential phospho-ligand binding site for the KI-FHA domain I predicted the functional surface of KI-FHA using Evolutionary Trace analysis. In a search for potential KAPP binding site in RLKs, three phosphoThr peptide fragments of KAPP-binding RLKs were found by isothermal titration calorimetry (ITC) and NMR to bind KI-FHA. NMR and ITC suggest affinities with K_d values of 8 to 30 μM. Thermodynamics study revealed that their affinities were driven by favorable enthalpy and the desolvation effect. Mutagenesis of these three threonine sites suggest Thr546 in the C-lobe of BAK1 kinase domain to be a principal but not sole site of KI-FHA binding *in vitro*. The phospho-

peptide binding site of KI-FHA determined by NMR titrations is very similar comparing to that predicted by Evolutionary Trace analysis. BRI1 kinase domain interacts with the same 3/4, 4/5, 6/7, 8/9, and 10/11 recognition loops of KI-FHA as do phosphoThr peptides, and with similar modest affinity.

The backbone mobility of KI-FHA, free and bound to a pThr peptide from its CLV1 receptor-like kinase binding partner, has also been investigated using ^{15}N NMR relaxation at 500 MHz and 600 MHz. This is the first backbone dynamic study of an FHA domain and an FHA in complex with its phospholigand. Binding of the CLV1 pThr peptide seems to reduce nsec-scale fluctuations of KI-FHA globally. In the psec to nsec timescale, KI-FHA residues that are critical for phosphopeptide recognitions are rigid. Peptide binding rigidifies KI-FHA at the binding site and remote sites across the β-sandwich. Peptide binding increases flexibility around the periphery of the binding site, perhaps relieving strain from the peptide association. The rigidity of this FHA domain appears to couple as a whole to pThr peptide binding.

AvrRpt2 is an effector protein from *Pseudomonas syringae* that plays an essential role in pathogen response of plant. The folding and activation processes of the cysteine protease AvrRpt2 in complex with its chaperone ROC1, have been characterized by fast protein liquid chromatography and NMR. AvrRpt2 was found to form a stable 140 kDa complex with ROC1 and to fold progressively in the presence of ROC1.

Chapter I: Introduction

I.1 Receptor-like protein kinase mediated signaling in Arabidopsis

I.1.1 Overview – Receptor protein kinases and receptor-like protein kinases

Receptor protein kinases (RPKs) are important mediators of signal transduction. RPKs are generally located in the plasma membrane, and usually contain an extracellular receptor domain, a single transmembrane domain and an intracellular protein kinase domain. Two major classes of RPKs include receptor tyrosine kinases (RTKs) and serine/theorine kinase receptors (STKRs). In animals, the vast majority of RPKs are receptor tyrosine kinases (RTKs). This family includes the receptors for insulin, ephrins, growth factors, and membrane-associated protein ligands (*1*). The transforming growth factor-β (TGF-β) receptor family, which functions in embryonic development and homeostasis, is the only class of receptors with serine/threonine kinase activities. The general mechanism of the activation of an RPK is that it exists in the membrane as an inactive monomer (*2*). Ligand binding induces dimerization, which allows the intracellular kinase domains to transphosphorylate and activate each another (*3*).

Genome research has identified 43 RPK genes in *C. elegans* (*4*), 25 in drosophila (http://flybase.bio.indiana.edu), 119 in yeast (*5*) and 70 in humans (*6*). In plants, numerous proteins with the predicted topologies of receptor protein kinases have been identified. These proteins are referred to as receptor-like protein kinases (RLKs). In the Arabidopsis genome, over 400 genes have been found as putative RLKs (*7*). Similar to RPKs, RLKs also contain three domains, an extracellular putative receptor domain, a transmembrane domain and a cytoplasmic catalytic domain (Figure I-1) (*8*).

RLKs in plants and RPKs in animals differ from each other in many aspects. First, almost all RLKs have serine/threonine kinase consensus sequences. Only a few exceptions have been identified. For example, the pollen receptor-like kinase1 (RPK1) (*9*) and the somatic embryo genesis receptor-like kinase (SERK) (*10*) may phosphorylate tyrosines, as well as serines and threonines. Second, in comparison to RPKs, RLKs function in a greater variety of signaling processes, including plant growth and development, hormone perception, self-incompatibility, and pathogen response. As more RLKs are studied, it is fully expected that RLKs function in many other processes. Furthermore, alternative splicing has been identified frequently in RLK transcripts (*3, 11*), which would add more complexity to the study of RLK-mediated signaling in plants.

I.1.2 Families of plant RLKs

The Arabidopsis genome contains over 400 RLK genes, and most of the RLKs from other plants have a closely related member in Arabidopsis (*7*). RLKs in plants can be classified based on their extra-cellular domains (Figure I-1) (*12*). Over 21 different classes of RLKs have been reported.

The most common class is the leucine-rich repeat receptor kinases (LRR-RLKs). There are over 200 LRR-RLKs in the Arabidopsis genome (*7, 13*). The LRR-RLK family includes 13 subfamilies, designated from LRR I to XIII (http://mips.gsf.de/proj/thal/db/index.html) and classified according to the organizations of the putative receptor domains (*12*). LRRs are thought to be involved in protein-protein interactions (*14*) and are known to function in plant development, hormone perception, and pathogen interactions.

Other major classes of RLK receptor domains include lectin motifs, epidermal growth factor (EGF) repeats, and cysteine-rich repeats (CRRs). The class with lectin motifs is the second largest, with over 40 members of the lectin-receptor kinases (LecRLKs) (*13*). A B-lectin motif is present in the receptor domains of RLKs with a cysteine-rich S-domain. The S-domain is found in RLKs involved in self-incompatibility (*13*).

I.1.3 RLK-mediated signaling cascades and regulation

The first plant RLK was identified from maize in 1990 (*15*). Since then, numerous RLKs have been found from *Arabidopsis thaliana* and other species. However, among the tremendously huge number of the putative RLKs in plants and the >200 LRR-RLKs in Arabidopsis, only several have been studied extensively (*16*). How RLK signaling is conveyed and regulated remains poorly understood.

While plant RLKs evolved independently from animal RPKs, the components of RLK signaling are similar to those of RPKs. Upon ligand binding to the receptor domain of an RLK, the intracellular kinase domain is activated by either autophosphorylation or transphosphorylation(*8, 16, 17*). The activation of an RLK typically involves recruiting proteins to the phosphorylated receptor complex and triggering the downstream signals to be conveyed into cell nuclei through cascades (*8*).

Phosphorylation and dephosphorylation are important for the regulation of the RLK signaling. It has been assumed that the basic mechanism of kinase domain activation is universal. Similar to RPKs, activation loops play essential roles in the activation and deactivation of RLKs. Juxtamembrane segments or RLK kinase domains are potentially important docking sites for substrates or cofactors. Except for phosphorylation, activities of RLKs are potentially regulated by other mechanisms, such as internalization and restriction of an RLK's localization (*8*).

I.1.4 Examples of RLK-mediated signaling in Arabidopsis

Most of the RLKs with known functions belong to the family of LRR-RLKs, and these RLKs are involved in diverse life processes in plants (Table I-1). Three of the best characterized RLKs in Arabidopsis, including BRI1, CLV1 and FLS2, illustrate the general model of RLK-mediated signaling.

I.1.4.1 Brassinosteroid signaling – the BRI1 pathway

Steroid hormones are very important for physiological and developmental regulation in plants. Brassinolide insensitive I (BRI1) is one of the few plant LRR-RLKs with a known hormone ligand and has a well-studied downstream pathway. In Arabidopsis, *BRI1* was identified by its brassinosteroid (BR)-deficient phenotype and the failure to be rescued by brassinolide (BL) treatment (*18*). BRI1 is globally expressed in all plant tissues (*19*). BRI1 binds BL, a type of BR, at the cell surface and regulates a variety of events in plant development. Another LRR-RLK, the BRI1-associated kinase 1 (BAK1), which belongs to the somatic-embryogenesis receptor kinase (SERK) family, has been identified to interact with BRI1. BAK1 is believed to play an important role in the activation of BRI1 kinase domain (*20*).

The currently hypothetical BR signaling pathway is shown in figure I-2. Without BL, BRI1 is kept in an inactive state by both its own carboxyl terminal domain and by an interaction with BKI1 (*21*). BL binding to the extracellular domain of BRI1 induces a

conformational change of the kinase domain, leading to the phosphorylation of the C-terminal domain of BRI1 and BKI1, the dissociation of BKI1 from the plasma membrane, and the full activation of BRI1. The activated BRI1/BAK1 signaling complex then inactivates the GSK3-like kinase BIN2 (*22*), which is an upstream regulator of BES1 and BRZ1. In the absence of BR, BIN2 is constitutively active and phosphorylates BES1 and BRZ1, leading to the degradation of BES1 and BRZ1. When not phosphorylated, BES1 and BRZ1 are localized to the nucleus where they activate transcription of brassinosteroid responsive genes (*16, 21*). Recent studies have revealed that kinase-associated protein phosphatase (KAPP) is another negative regulator of BRI1 signaling (*23*).

I.1.4.2 Meristem development – the CLV1 pathway

CLAVATA1 (CLV1) has a putative receptor domain containing 21 LRRs (*24*). CLV1's major function is to balance cell proliferation and differentiation, by promoting the shoot apical meristem to differentiate and by inhibiting meristem proliferation via down-regulating *WUS* expression (see below) (*25*). CLV2 is a receptor-like protein with a LRR motif and a short cytoplasmic tail (*26*). CLV3 is a 96-residue polypeptide that acts as the ligand for CLV1/CLV2 receptors (*27*). CLV3 binds to the heteromeric receptors of CLV1/CLV2 and forms a 450 kDa active CLV1 signaling complex at the cell membrane (*28*).

A hypothetical model for CLV1 signaling is illustrated at figure I-3. Upon binding of CLV3, CLV1 forms a heterodimer with CLV2 (*26*). Dimerization induces the autophosphorylation of serine or threonine residues on the kinase domain of CLV1. The activated CLV1 receptor complex recruits downstream components such as ROP, a small Rho-related GTPase (*28*), and transmits the signal through the mitogen-activated protein kinase (MAPK) cascade to the transcription factor WUSCHEL (WUS). Mutagenesis of the *WUS* gene suggested that WUS plays an antagonistic role to CLV1 (*29*). The CLV1 signaling pathway may lead to a suppression of *WUS* function (*30*). KAPP is an important negative regulators of the CLV1 pathway (*31-33*). Recent studies demonstrate that POLTERGEIST (POL) and PLL1 are two integral downstream components of the CLV1 signaling, which are important in regulating the balance between stem-cell maintenance and differentiation (*34*).

I.1.4.3 Pathogen response – the FLS2 pathway

Flagellin insensitive 2 (FLS2) is an LRR-RLK that functions in Arabidopsis pathogen defense (*35*). FLS2 is the receptor of the most conserved domain of flagellin, the flg22 peptide. The activated FLS2 initiates a signaling cascade with the MAPK signaling proteins, including MEKK1 and MAPK3/MAPK6, as downstream effectors (*36*). The final targets of the FLS2 pathway include the well-defined defense genes of plants, such as *PR5*, *PR1*, *PAL1*, *GST1*, and WRKY22/WRKY29 transcription factors (*16*). KAPP has been shown to interact with the C-terminal tail of FLS2 in a yeast two-hybrid assay and to negatively regulate the FLS2 signaling in Arabidopsis (*37*).

I.2 Kinase associated protein phosphatase: a common negative regulator of RLK signaling in Arabidopsis

The kinase associated protein phosphatase (KAPP) is the first characterized downstream regulator of RLKs in Arabidopsis. It was identified by screening an Arabidopsis cDNA expression library against the cytoplasmic kinase domain of an RLK, HAESA (*38*). *KAPP* is a single copy gene and no homologs have been identified in the Arabidopsis genome (*39*). *KAPP* has a wide expression pattern in different tissues and different developmental stages in Arabidopsis. KAPP has been found to interact with many RLKs, including HAESA (*38*), CLAVATA1 (*24*), RLK4 (*31*), TMK1 (*31*), WAK1 (*40*), FLS2 (*41*), BAK1 (*20*) (*23*), and SERK1 (*42*). KAPP orthologs identified in maize and rice also show binding to RLKs (*31*).

KAPP functions as a common negative regulator of RLK signaling. For example, KAPP has been shown to bind phosphorylated CLV1 *in vitro* and *in vivo*. The *clv1* mutant has an enlarged meristem. Reduction of KAPP transcript suppressed the *clv1* phenotype (*43*). Overexpression of KAPP mimicked the *clv1* mutant phenotype (*44*). Additional supporting evidence can be found from the studies of FLS2. Overexpression of KAPP mimicked the *fls2* mutant phenotype with respect to lack of growth inhibition and lack of production of reactive oxygen species in response to flg22 treatment (*41*). Recent studies reveal that KAPP also negatively regulates the BRI1 pathway (*23*).

KAPP is a multi-domain protein (Figure I-4), which contains a type I membrane anchor, a kinase-interacting FHA domain, and a type 2C protein phophatase (PP2C) domain (*38*). The N-terminal signal anchor targets KAPP to the cell plasma membrane. A 52-residue region in the 239-amino acid (residue 98-336) kinase-interacting domain shares sequence identity with FHA domains (*45*). The minimal functional phospho-protein binding module of KAPP comprises 119 residues (180-298). We designate the 119-residue region as the "KI-FHA" domain. 5 highly conserved residues of FHA domain include G211, R212, S226, H229, and N250 in KI-FHA numbering. Site-directed mutagenesis of G211, S226, H229, or N250 abolishes CLV1 binding of KI-FHA *in vitro*, which suggests that these residues are essential for KI-FHA's interactions with RLKs (*46*). *In vitro* protein phophatase assays show that KAPP containes a PP2C catalytic domain, with phosphoSer/Thr substrate specificities. Mg^{2+} and Mn^{2+} are required for the phophatase activity of KAPP, and KAPP is insensitive to a high concentration of okadaic acid, which is a specific inhibitor of PP1 and PP2A phosphatases (*38*).

How might KAPP interact with RLKs? The current hypothesis is that KAPP utilizes its KI-FHA domain to bind a phosphorylated RLK kinase domain, which places the PP2C domain close to dephosphorylate phosphoSer/Thr of the activated RLK kinase domain to attenuate the RLK signaling. A model of the CLV1 pathway is shown in figure I-3 to illustrate how the RLK signaling is negatively regulated by KAPP (*17*).

I.3 FHA domain: a class of phosphoprotein binding domain

I.3.1 Overview

FHA (Forkhead associated) domains were first identified in forkhead transcription factors as a class of phosphoprotein interacting domains (*45*). FHA domains have been found in over 1000 signaling proteins in both prokaryotes and eukaryotes, including protein kinases, protein phosphatases, adenylate cyclases, proteases, kinesins, zinc-finger proteins, and glycoproteins. FHA domains are associated with proteins which function in numerous processes, including intracellular signal transduction, transcription, protein transport, DNA repair, and protein degradation. Multiple sequence alignments show that FHA domains share low (< 30%) sequence indentities (http://www.sanger.ac.uk/Software/Pfam/data/jtml/full/PF00498.shtml.gz). The homologous regions of FHA domains comprise 55-75 amino acids, which are essential but not sufficient to form a functional phosphoprotein binding unit. Functional FHA domain modules usually contain 110-140 residues (*17*).

Similar to SH2, PTB and WW domains, FHA domains function as a scaffold in recognizing phosphorylated epitopes in signal transduction. For example, yeast protein Rad53 is a checkpoint protein that prevents cell division during DNA damage. Rad53 contains two FHA domains, FHA1 in the N-terminus and FHA2 in the C-terminus. In budding yeast, DNA damage is usually detected by a receptor or a sensor, which initiates the activation of a set of proximal kinases, such as TEL1 and MEC1 (*47*). These kinases

then activate a series of more distal kinases (i.e. Rad53) that phosphorylate and regulate other protein effectors involved in DNA damage responses. Rad53 recognizes its downstream target Rad9 via the two FHA domains and transmits the DNA damage signals (*48*). Similarly, FHA domains are also in a subset of forkhead-like transcription factors, such as FKH1/FKH2 and FHL1/FHL2 in yeast, and MNF and ILF1/2 in humans (*45*). These transcription factors use FHA domains to interact with other cofactors and form signaling complexes that regulate gene transcription.

I.3.2 Structures

The structures of the FHA domains from many signaling proteins, including plant phosphatase KAPP (*49*), yeast checkpoint protein Rad53 (*48, 50*), human checkpoint protein Chfr (*51*), human tumor suppressor Chk2 (*20*), and human nuclear protein Ki67 (*52*), have been determined by NMR or by X-ray crystallography. Almost all FHA domains are monomers under the experimental conditions *in vitro,* except the FHA domain of Chfr in crystals, which forms a homodimer through segment swapping at β-hairpin 7/8 (*51*). Despite the low sequence homology, structure-based sequence alignment (Figure I-5) shows that FHA domains share high structural similarity. FHA domains also share a similar fold with the MH2 domain of tumor suppressor SMAD4, a phosphoSer recognition module.

The NMR structure of KI-FHA of KAPP has been determined and represents FHA domains. As shown in figure I-6, KI-FHA is a sandwich with a 5-stranded mixed β-sheet

(β-strand 4, 3, 5, 6, 9) and a 6-stranded anti-parallel β-sheet (β-strands 1, 2, 11, 10, 7, 8) (*49*). The 5 highly conserved residues are located in the loop regions (Figure I-6).

I.3.3 Functions: recognition of phosphopartners

The structures of phosphopeptide bound FHA domains (*51, 53-56*) suggest that the phosphopeptide recognizing site is located at the 3/4, 4/5, 6/7 and 10/11 loops in an extended conformation. FHA domains bind their optimal phosphopeptides weakly with dissociation constants of 1-50 μM. A striking feature of FHA domains is the apparent binding specificity for pThr peptides *in vitro*.

Low sequence homology suggests that FHA domains have very diverse phosphopeptide binding specificities. An oriented phosphoThr peptide library screening against 6 different FHA domains suggests that the pT+3 residue in the phosphopeptide is of most importance to specificity. The pT+1 position is of the second importance (*57*). Different FHA domains select for different amino acids at the pT+3 position. For example, Rad53 FHA1 selects strongly for Asp at the pT+3 position of a phosphoThr peptide, whereas Chk2 FHA prefers Ile and KAPP KI-FHA selects for Ser or Ala (Table I-2).

The interaction of Rad53 FHA1 and the peptide SLEVpTEADATFAKK from Rad9 features an extensive network of hydrogen bonds, which is similar to the binding mechanisms of pSer or pTyr peptides by a 14-3-3 domain or an SH2 domain (*58*). Among the 5 higly conserved residues in the recognition loops of Rad53, Arg70, Ser85

and His88 interact directly with the peptide. The phosphate of pThr accepts 5 hydrogen bonds from the conserved Arg70 in the 3/4 loop, the conserved Ser85 and the non-conserved Asn86 in the 4/5 loop, and T106 in the 6/7 loop of the Rad53 FHA1 domain (*53*). The side chains of the conserved Arg70, Asn107 and the non-conserved Asn86 donate hydrogen bonds to main chain carbonyl oxygens of Glu at pT+1, Glu at pT-2 and Ser at pT-4 positions of the peptide (*53*). The above interactions of the phosphoThr peptide with the conserved residues of Rad53 FHA1 are likely to be widespread in FHA domains' interactions of phosphorylated binding partners (Liang, X. et al, unpublished).

For years, the failure to identify a pSer peptide that associates with an FHA domain has been a puzzle. KAPP KI-FHA interacts with phosphorylated but not unphosphorylated RLK kinase domains *in vitro* and *in vivo.* Phospho amino acid assays show that the autophosphorylation of BRI1 kinase domain occurs at threonines or serines (*59*), and that the phosphorylation of CLV1 kinase domain might occur exclusively at serine residues (*43*). This evidence suggests that KI-FHA may bind to pSer peptides, as well as pThr partners. A recent study shows the binding of a 44-residue phosphopeptide to the FHA domain of Ki67. The long peptide, $hNIFK_{226-269}$, was triply phosphorylated in a sequential manner on Thr238, Thr234, and Ser230 residues. The solution structure of Ki67 FHA in complex with $hNIFK_{226-269}$ shows that the peptide becomes structured upon binding to the FHA domain. The pSer230 contacts Ki67 FHA domain and contributes to affinity marginally (*60*). This suggests that the additional phosphorylation on Ser230 is not important for the peptide recognition of this FHA domain.

I.3.4 Folding and stability

The folding of the KAPP KI-FHA domain has recently been characterized by NMR and other biophysics techniques. The folding mechanism of KI-FHA appears to be in a non-two-state (at least three-state: native, intermediate and unfolded) manner. The stable folding core of KI-FHA (Figure I-7) comprises β-strands 1, 5, 6, 9, 10 and 11, plus the 1/2 and 9/10 loops (*61*). This FHA domain is hypothesized to fold progressively, starting with the stable 6-stranded hydrophobic core, followed by the additions of the less stable flanking β-strands and loops. This exploration provides the first evidence of the existence of partially unfolded forms of a β-sandwich protein. The global folding stability of the KI-FHA domain has been shown to be dependent on pH. KI-FHA is significantly more stable at pH 7.3 than at pH 6.3 (*61*).

Table I-1: Functions of LRR-RLKs in plants (*16*)

(This table is taken from Diévart, A. and Clark, S., 2004, *Development* 131, 251-261)

	Genes	Functions	Organism	References
Development	CLAVATA1 (CLV1)	Meristem differentiation	Arabidopsis	(Clark et al., 1997)
	Phytosulfokine receptor (PSK receptor)	Peptide hormone binding	Tomato	(Matsubayashi et al., 2002)
	ERECTA (ER)	Overall plant shape	Arabidopsis	(Torii et al., 1996)
	Ipomocea nil receptor protein kinase 1 (INRPK1)	Short-day photoperiodic floral induction	Ipomocea nil	(Bassett et al., 2000)
	HAESA/RLK5	Floral organ abscission	Arabidopsis	(Jinn et al., 2000)
	Excess microsporocytes 1 (EMS1)/Extra sporogenous cells (EXS)	Endosperm and pollen development	Arabidopsis	(Zhao et al., 2002a; Canales et al., 2002)
	Somatic embryogenesis receptor 1 (AtSERK1)	Ovule development and early embryogenesis	Arabidopsis	(Hecht et al., 2001)
	Brassinosteroid insensitive 1 (BRI1) and BRI1 associated receptor kinase 1 (BAK1)	Perception of BR	Arabidopsis	(Li and Chory, 1997; Li et al., 2002; Nam and Li, 2002)
	VASCULAR HIGHWAY1 (VH1)	Leaf patterning	Arabidopsis	(Clay and Nelson, 2002)
	Systemin receptor (SR160)/CURL3/tBRI1	Perception of BR and systemin	Tomato	(Montoya et al., 2002; Scheer and Ryan, 2002)
Symbiosis	Nodule autoregulation receptor kinase (GmNARK)/ Hyper autoregulation of nodulation receptor 1 (HAR1)	Autoregulation of nodulation	Soybean *Lotus japonicus*	(Searle et al., 2003; Krusell et al., 2002; Nishimura et al., 2002)
	Symbiosis receptor-like kinase (SYMRK)/ Nodulation receptor kinase (NORK)	Root nodule formation	*Lotus japonicus* *Medicago sativa*	(Stracke et al., 2002; Endre et al., 2002)
Host defense	FLAGELLIN SENSITIVE 2 (FLS2)	Plant defense/pathogen recognition	Arabidopsis	(Gomez-Gomez and Boller, 2000)
	Xa21	Fungal perception00	Rice	(Song et al., 1995)

Table I-2: Selections of phosphothreonine peptide motifs for FHA domains (*62*). GST fusion proteins of the indicated FHA domains were screened with a peptide library containing the sequence MAXXXXpTXXXXAKKK where X indicates all amino acids except Cys and Trp. Enrichment values, shown in parentheses, were obtained by normalizing the affinities of the optimal amino acids to the average affinities of the remaining amino acids as described by (*63*). Residues in the +3 position, and those showing particularly strong enrichment, are underlined. [a] The FHA1 (R70A) mutant bound less than 25% of peptides compared with the wild type.

(This table is taken from Durocher, D. *et al.*, *Mol Cell.* 2000 Nov; 6(5):1169-82.)

	−3	−2	−1		+1	+2	+3
Rad53-FHA1	Y (1.5) L (1.3)	V (1.3) A (1.3)	K (1.5) P (1.4) A (1.4)	pT	*V* (1.8) I (1.6) E (1.6) L (1.4) Q (1.4)	V (1.6) T (1.5) I (1.4) K (1.3)	*D* (2.3) V (1.4) I (1.3) L (1.2)
S. pombe-Cds1	Y (1.8) F (1.6) L (1.3)	V (1.7) I (1.4) Q (1.2)	K (1.4) A (1.4)	pT	*V* (1.9) L (1.8) E (1.7) I (1.6) Q (1.6)	L (1.6) V (1.5) R (1.5) K (1.4)	*D* (1.4) Y (1.3)
Rad53-FHA1[a] R70A Mutant	K (1.7) I (1.6) F (1.3)	A (1.2)	K (1.7) A (1.7) L (1.3) I (1.3)	pT	K (1.4) I (1.4) L (1.3)	K (1.6) E (1.3)	*D* (1.7) V (1.4) I (1.3) L (1.3)
RAD53-FHA2	L (1.7) I (1.7) K (1.7) T (1.3)	Q (1.3) A (1.3)	K (1.9) L (1.8) A (1.7) Y (1.3)	pT	I (1.4) V (1.3) Q (1.2)	*K* (2.1) R (1.5) I (1.5)	*I* (3.0) L (1.8)
KIAA0170	Y (1.4) P (1.2) M (1.2)	A (1.5)	P (1.5) K (1.2)	pT	Q (1.5) V (1.4) E (1.4) L (1.4) M (1.3) I (1.3)	L (1.4)	*L* (2.2) *I* (1.8) M (1.5) V (1.4)
Y127_MYCTU	L (1.7) I (1.3) M (1.3)	I (1.4) V (1.3) Q (1.3)	P (1.2) K (1.2)	pT	*V* (2.2) *Q* (2.0) E (1.8) I (1.6)	L (1.5) Y (1.2) K (1.2)	*Y* (2.1) *M* (1.8) L (1.7) I (1.6) F (1.4)
KAPP	A (1.4) Y (1.4) F (1.3)	A (1.4)	Y (1.3) F (1.2)	pT	V (1.5) Q (1.5) E (1.3) L (1.3) I (1.2)	X	*S* (1.9) *A* (1.7)

Figure I-1: Leucine rich repeats receptor – like protein kinases (LRR-RLKs)

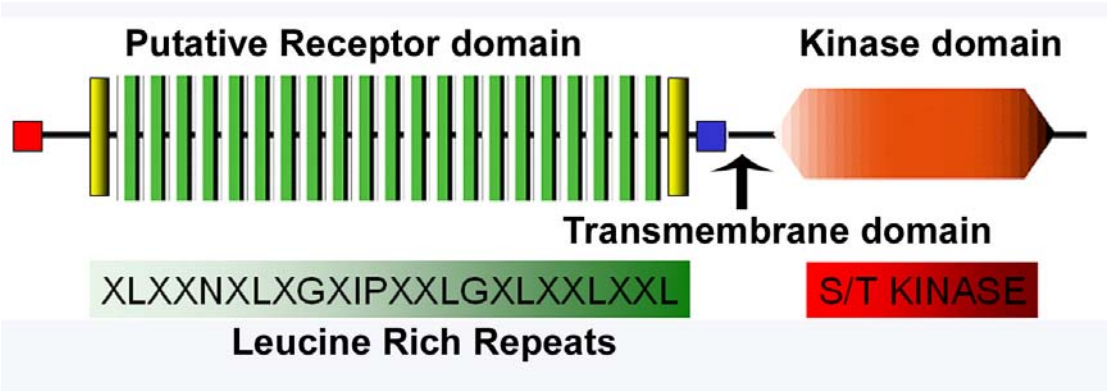


Figure I-2: Model for the BRI1 signaling pathway. Without Brassinosteriod (BR), BRI1 is kept in an inactive state by both its own carboxyl terminal domain and by an interaction with BKI1. BR binding to the extracellular domain of BRI1 induces a conformational change of the kinase domain, leading to the phosphorylation of the C-terminal domain of BRI1 and BKI1, and the dissociation of BKI1 from the plasma membrane (*64*). These events lead to the full activation of BRI1 and its association with BAK1 or other substrates. The activated BRI1 signaling complex then inactivates the GSK3-like kinase BIN2, which is an upstream regulator of BES1 and BRZ1. In the absence of BR, BIN2 is constitutively active, and phosphorylates BES1 and BRZ1, leading to their degradation. When BES1 and BRZ1 are not phosphorylated, they are localized to the nucleus where they activate transcription of brassinosteroid responsive genes (*16, 64*).

(This figure is taken from Wang, X. et al., 2006 *Science* 313: 1118)

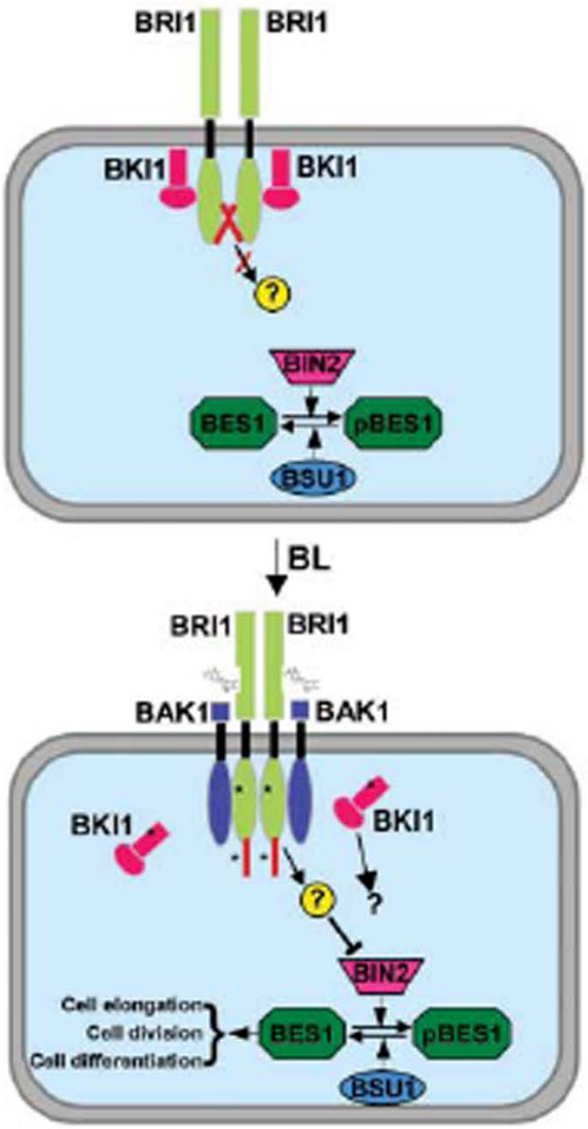

BRI1
BRI1
BKI1
BKI1
?
BIN2
BES1
pBES1
BSU1
BL
BRI1
BRI1
BAK1
BAK1
BKI1
BKI1
?
?
BIN2
Cell elongation
Cell division
Cell differentiation
BES1
pBES1
BSU1

Figure I-3: Model for the role of the KI-FHA domain in CLV1 signaling pathway. A putative ligand, CLV3 interacts specifically with the extracellular domains of CLV1 and CLV2 (*28*), which triggers conformational changes in the cytoplasmic kinase domain of CLV1 and leads to autophosphorylation of the CLV1 catalytic domain on multiple serine residues. The KAPP FHA domain selectively recognizes and interacts with a pThr containing region. The PP2C domain of KAPP then dephosphorylates CLV1 and attenuates CLV1 signaling. A Rho-GTPase-related protein, ROP, was also found in the assembled CLV1 signaling complex, which suggests that CLV1 transduces its signal through a MAPK pathway. The question marks represent unknown components (*17*).

(This figure is taken from Li, J. *et al., J Cell Sci*. 2000, 113: 4143-4149)

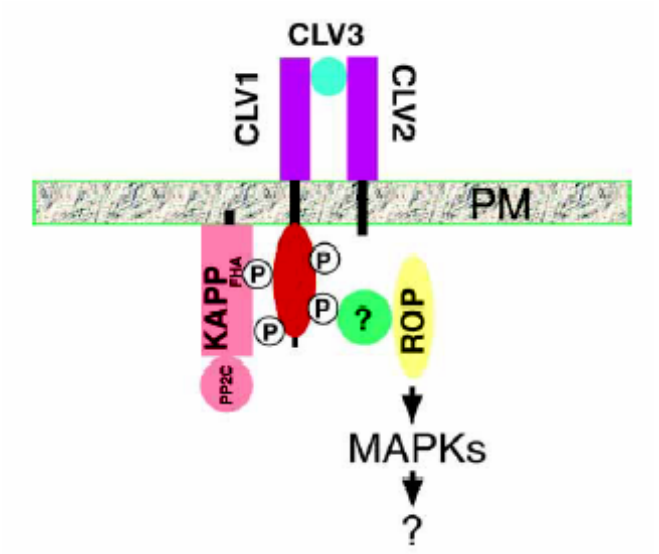


Figure I-4: Domains of the kinase associated protein phosphatase

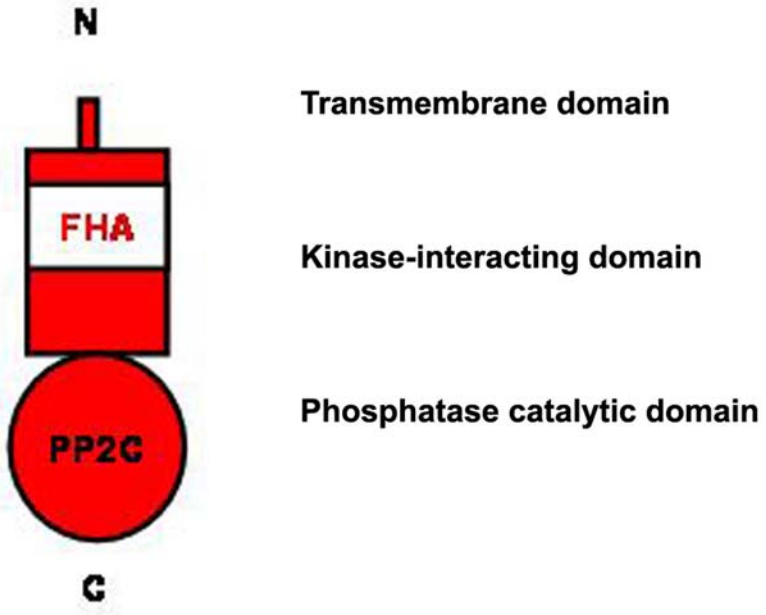


Figure I-5: Structure based sequence alignment of FHA domains. Locations of the secondary elements are marked above the aligned sequences. The Protein Data Bank (PDB) accession codes are as follows: lg6g-A for FHA1-RAD53; lqu5-A for FHA2-RAD53; 1lgp-A for CHFr; and lygs for MH2-SMAD4. Side chains with the backbone amide protons protected from hydrogen exchange are highlighted in yellow. Positions at which hydrophobic residues are conserved in all FHA folds are shown in red while hydrophilic residues are in blue. The numbers following Δ are the numbers of residues not shown in this figure. Highly conserved residues are highlighted in green and absolutely conserved glycines are in red (*49*).

(This figure is taken from Lee, G. *et al., PNAS* 2003, 100: 11261-11266)

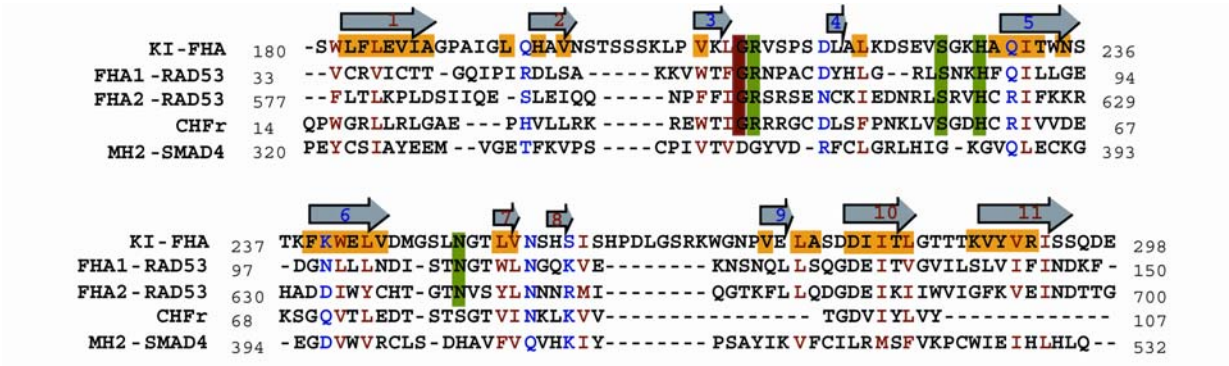


Figure I-6. NMR structure of KI-FHA domain (a) Superposition of the 30 accepted KI-FHA models. The two β-sheets are shown in red and blue respectively. β-strands 1, 2, 7, 8, 10, and 11 are in red. The loops between red strands, the C terminus, and Ser 180 at the N terminus are in orange. β-strands 3, 4, 5, 6, and 9 are in blue. The loops between blue strands are in light blue. The loops between the two β-sheets, namely the 2/3, 6/7 loop, and 8/9 loops are in violet. Five foreign residues at the N terminus are in gray. (b) Ribbon plot of the backbone structure of KI-FHA. Representative model number 1 is drawn using the color code for the strands used in (a). The side chains of highly conserved Arg 212, Ser 226, His 229, and Asn 250 in green are found in close space on the edge of a β-sheet (*49*).

(This figure is taken from Lee, G. *et al., PNAS* 2003, 100: 11261-11266)

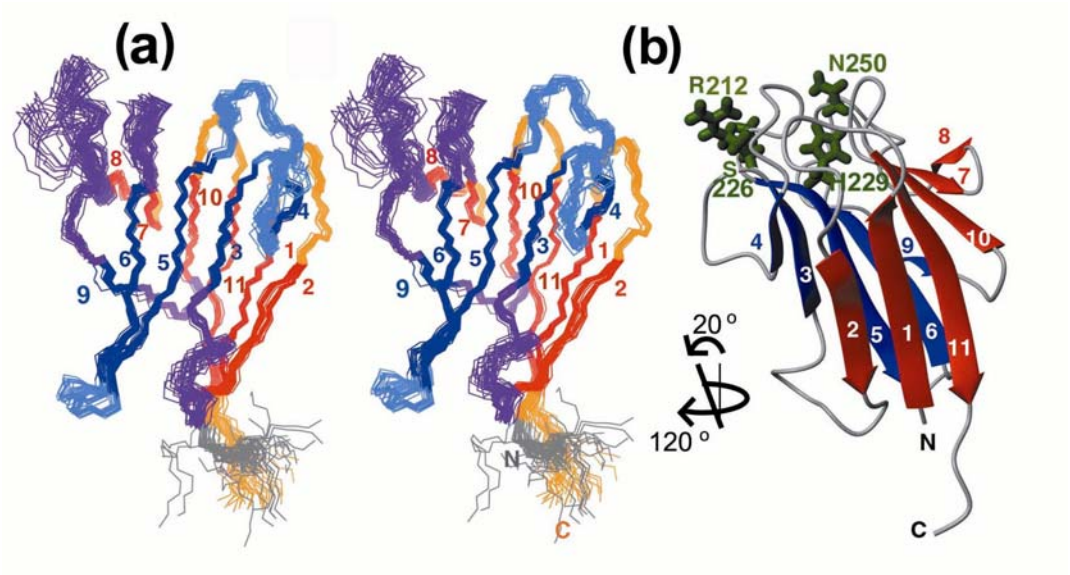
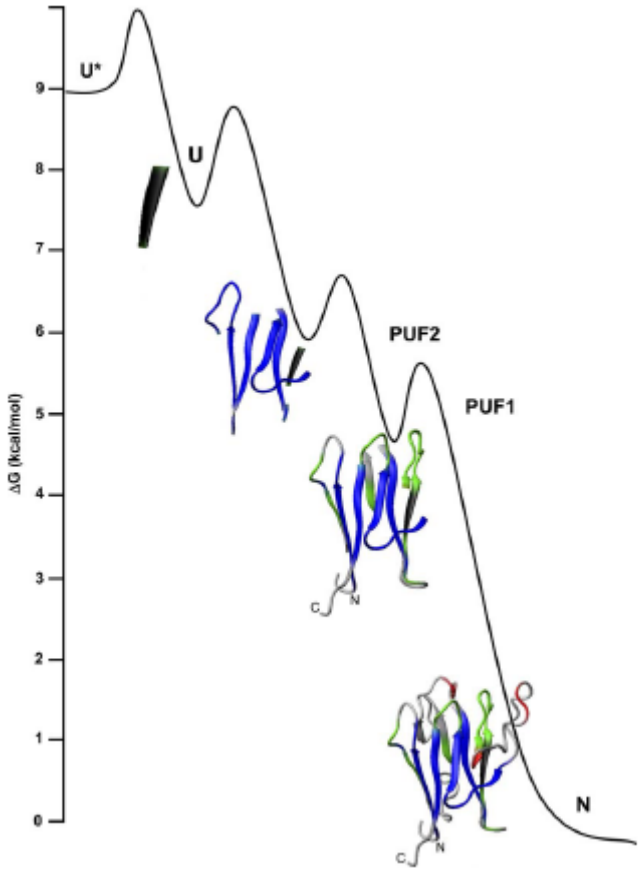


Figure I-7. Proposed model for the energy landscape of KI-FHA folding and structure of its PUFs minutely populated at equilibrium. NHX suggests two near-native, partially unfolded states (PUF1 and PUF2), an unfolded state U open to HX, and a higher energy unfolded state U* with residual structure melted. The structural states use red and green for progressively higher subglobal stabilities, blue for global stability and black for superprotection (*61*).

(This figure is taken from Liang, X. *et al., J Mol Biol.* 2006, 364(2):225-40)



Chapter II: Technical background

II.1 Nuclear Magnetic Resonance (NMR)

Nuclear magnetic resonance (NMR) is a physical phenomenon characterizing the magnetic properties of nuclei. When nuclei of certain atoms are placed in a static magnetic field and are exposed to a second oscillating magnetic field at radio frequency (RF), the responses to the RF pulse are recorded in nuclear magnetic resonance spectroscopy. NMR spectroscopy is one of the major techniques that can be used to obtain physical, chemical, and structural information about a biomolecule. It is the only technique that can provide atom-specific information on the three-dimensional structure of biological molecules in solution.

II.1.1 Origin of NMR signals

Spin is a fundamental property of a nucleus. Spin is in multiples of 1/2 and it can be + or -. For example, each individual unpaired electron, proton, or neutron has a spin of 1/2.

Nuclei possessing non-zero nuclear spins in the ground state have an intrinsic magnetic moment, which can be active in NMR. The most commonly used nuclei in NMR are ^{1}H, ^{15}N, and ^{13}C. Those are nuclei containing 1/2 total spin, which have high sensitivity and can produce high resolution magnetic signals in NMR.

In NMR, the frequency of the photon (υ) is in the radio frequency (RF) range.

$$\upsilon = \gamma B \qquad (1)$$

where B is the magnetic field strength, and γ is the gyromagnetic ratio. When a RF pulse is applied to a molecule, energy is absorbed by the spins which make a transition from the lower energy state to the higher energy state. After the magnetic pulse is removed, the excited spins simultaneously make a transition from the higher energy state to the lower energy state. The emission is proportional to the population difference between two states at the RF frequency. This generates an oscillating electric field, named free induction decay (FID), which can be recorded by a sensitive radio frequency detector. It is the resonance or exchange of energy at a specific frequency between the spins and the spectrometer that gives NMR signals. The time domain signal, FID, can be converted into a frequency domain spectrum through Fourier Transforms (Figure II-1), which generates an NMR spectrum.

The overall magnetization vector (M) of a molecule is the sum of all the individual spin vectors of the nuclei. In an external magnetic field B_0, M is aligned according to the

orientation of B_0. After an RF frequency with magnetic field component B_1 is applied, M responds to this pulse by rotating around the direction of the applied B_1 field. The rotation angle θ can be expressed as

$$\theta = 2\pi\gamma B_1 t \qquad (2)$$

t is the length of time when B_1 is on. B_0 is along the Z axis. A 90° pulse with RF field strength B_1 is one which rotates the magnetization vector M clockwise by 90° around the X axis and down to the Y axis. Thus it tips down M from the Z axis to the XY plane. An 180° pulse will rotate M by 180° to the -Z axis.

II.1.2 Spin relaxation

Motions of molecules in solution which result in time-dependent magnetic field changes cause spin relaxation, which can be classified into two categories, spin-lattice relaxation and spin-spin relaxation.

At equilibrium, the net magnetization vector M_0 aligns along the direction of the static magnetic field B_o. The longitudinal magnetization M_Z equals M_o. When an 180° pulse is issued, it is possible to make $M_z(t) = - M_z$. After the external pulse is stopped, $M_Z(t)$ will return to its equilibrium state M_0. The process is called spin-lattice relaxation or longitudinal relaxation. The time constant describing the process is called the spin-lattice relaxation time (T_1). The equation quantifying this behavior is

$$M_z = M_o (1 - e^{-t/T_1}) \quad (3)$$

T_1 is dependent on magnetic field strength B_0. T_1 is inversely proportional to the density of molecular motions at the Larmor frequency. The rate of the spin-lattice relaxation is expressed as

$$R_1 = 1/T_1 \quad (4)$$

If a 90° pulse is applied and the net magnetization M is placed in the XY plane. It will rotate around the Z axis at a frequency equal to the frequency of a photon which can cause a transition between the two energy levels of the spin. This frequency is called the Larmor frequency. At the same time, M starts to dephase along the XY plane and returns to its equilibrium state M_{xyo}. The process is called spin-spin relaxation or transverse relaxation. The spin-spin relaxation time, T_2, can be expressed as

$$M_{XY} = M_{XYo}\, e^{-t/T2} \quad (5)$$

T_2 is always less than or equal to T_1. T_2 is inversely proportional to the number of molecular motions less than or equal to the Larmor frequency. The rate of the spin-spin relaxation is expressed as

$$R_2 = 1/T_2 \quad (6)$$

II.1.3 Chemical shift

When a nucleus is placed in an external magnetic field, its electron circulates around the direction of the applied magnetic field, which generates a small magnetic field offsetting

the external field (shielding). The local effective magnetic field B is therefore generally less than the applied field B_0 by a fraction of σ.

$$B = B_o (1-\sigma) \qquad (7)$$

The electron density around different nucleus in a molecule varies according to the types of nuclei and the local chemical environment in the molecule. The opposing magnetic field and B of each nucleus will be different. This is called the chemical shift phenomenon. The chemical shift of a nucleus can be quantified as the difference between the resonance frequency of the nucleus and a standard. This quantity δ is reported in parts per million (ppm).

$$\delta = (\nu - \nu_{Ref}) \times 10^6 / \nu_{Ref} \qquad (8)$$

Tetramethylsilane (TMS), $Si(CH_3)_4$, is generally used as the reference in ^1H NMR spectroscopy.

The chemical shift δ of an atom is independent from the external field strength and is very sensitive to the local environment (i.e. Electronegativity, magnetic susceptibility of neighboring groups). It is predominantly dependent on intramolecular interactions, but it can also arise from intermolecular interactions as well. Important structural information around a nucleus can be obtained from the value or changes of its chemical shift.

II.1.4 Multi-Dimensional NMR

II.1.6.1 Fourier transformation and one-dimensional NMR spectra

A Fourier transform is a mathematical technique which converts functions (i.e. FID) from time domains to frequency domains (Figure II-1). The effect of Fourier transformation is to make the individual component of FID visible and plot the frequencies. The definition of the Fourier transform is shown as below:

$$f(\omega) = \sum_{-\infty}^{+\infty} f(t)[\cos(\omega t) - i \cdot \sin(\omega t)] \qquad (9)$$

A one-dimensional NMR spectrum, which is generated from an FID through Fourier transform and phase corrections, typically contains limited information of peak locations, peak splitting, and peak height in NMR spectra. Direct information on the connectivity of nuclei is generally not available.

II.1.6.2 Two-dimensional NMR

A 2D NMR spectrum is generated from a 2-dimensional Fourier transform. A 2D array of time domains have a t' and a t" dimension (Figure II-2a). A Fourier transform is first performed on the data in one dimension and then in the other. The first set of Fourier transforms (Figure II-2b) are performed in the time domain t' to yield a frequency domain

f' by t'' set of data. The second set of Fourier transforms (Figure II-2c) are performed in the t'' dimension to yield an f' by f'' set of data (*65*).

In a 2D NMR spectrum, two axes contain the chemical shifts information of each nucleus. If two nuclei interact during the mixing time, a peak will arise in the spectrum. Peaks with the same frequencies in the two dimensions are called “Diagonal peaks”. Peaks with different frequencies in the two axes are named “Cross peaks”, which generally contain structural information. For examples, the COSY or TOCSY NMR experiment correlates bonded homonuclear resonances that are coupled together through J-Coupling. In a 2D COSY or a 2D TOCSY spectrum, each cross-peak represents a pair of coupled spins. The nuclear Overhauser effect (NOE) is an incoherent process in which two nuclear spins cross-relax through space. In a 2D NOESY spectrum, each cross peak indicates that the nuclei resonating at the two frequencies are within 5 Å in space.

II.1.6.3 Three, four - dimensional NMR

Because of the increased number of nuclei in a biomolecule, the 1D or even the 2D spectra become crowded with overlapping signals to an extent where analysis can become impossible. 3D or 4D NMR experiments have been created to deal with this problem. Typically, a 3D spectrum, which can be visualized as a cube, is acquired by incrementing the values of two pulse sequence delays t_1 and t_2 independently. The evolutions of a 4D NMR experiment are repeated three times. It can be described as a linear array of cubes. Adding a new dimension decreases the experimental sensitivity and

increases resolution. In order to perform a multi-dimensional NMR experiment, a sample should be isotopically (i.e. ^{13}C and ^{15}N) labeled. An *n*-dimensional experiments can involve up to *n* types of different nuclei.

II.2 Investigations of protein dynamics by NMR

In recent decades, protein structures have been determined by multidimensional NMR and X-ray crystallography. These efforts shed light on structural information to atomic resolution, and provide insights into how proteins function. Structural information on proteins has been widely used in various areas, such as structure-based drug design, mutagenesis, folding and unfolding. However, the determined 3D structures usually do not contain comprehensive information on protein motions, such as amplitudes and timescales. In some cases, structure alone is insufficient to explain precisely the working mechanism of a protein. Many lines of evidence show that protein motions are also closely related to their biological functions. A variety of techniques have been developed to study protein dynamics, including low-angle X-ray diffraction, H/D exchange mass spectrometry, and NMR. Among these techniques, NMR is capable of probing molecular motions ranging from psec to days, providing the most comprehensive residue-specific information of protein motions (*66, 67*).

II.2.1 Protein motions

Protein fluctuations cover a wide range of amplitudes (10^{-2}-10^{2} Angstrom), timescales (10^{-15}-10^{4} second), and energies (10^{-1}-10^{2} kcal·mol^{-1}) (*68*). Proteins undergo random rotations in solution, which is typically in the nsec timescale. The timescales of the random rotations depend upon the molecular size, the solvent viscosity, and temperature. Proteins also undergo translational motions, with timescale ranging from msec to sec. The internal motions usually play more important roles for protein functions. Types of internal fluctuations of a certain bond of a protein residue (Table II-3) include molecular vibrations (psec or shorter timescale), rotations (μsec to msec), and chemical exchange (nsec to sec or longer) (*69*). Dynamic processes on μsec to msec and psec to nsec timescales are particularly important for protein functions.

Modern NMR has been widely used to probe the residue-specific or bond-specific fluctuations of proteins (Table II-1). Using multi-dimensional NMR, linkages between protein dynamics and enzymatic substrate turnover rate have been elucidated for adenylate kinases (Adk) (*70*). Correlated networks of protein motions have been proposed to be important for functions (*71-75*). The effects of residue-specific fluctuations on ligand-recognition, thermodynamics, folding and stabilities of proteins have also been investigated (*61, 76*). Typical NMR methods that have been routinely used to characterize protein motions will be introduced as follows.

II.2.2 The model-free approach

The model-free approach, first introduced by Lipari and Szabo (*77*), allows a complete characterization of the fast (psec to nsec timescales, faster than the global tumbling) internal motions of proteins. Model-free analysis is usually performed through the Modelfree program that fits an extended model-free spectral density function with NMR relaxation data of certain bond vectors of protein residues. This includes the spin-lattice relaxation rate (R_1), the spin-spin relaxation rate (R_2), and the heteronuclear steady-state ^{15}N $\{^1H\}$ nuclear Overhauser effect (NOE). Model-free calculations yield three parameters, the generalized order parameters (S^2), the internal correlation time (τ_e), and the conformational exchange term R_{ex}, which probe internal motions of bond vectors (N-H, C-C, C-H) of protein residues at psec to nsec and μsec to msec timescales respectively. The analysis is designated "model-free" because the parameters are derived without the need to invoke a specific model of internal motions.

The relaxation of backbone ^{15}N nuclei of proteins is mediated by dipolar couplings to the directly bonded amide ^1H and by chemical shift anisotropy (CSA). The spectral density functions of R_1, R_2, and ^{15}N $\{^1H\}$-NOE are given by

$$1/T_1 = R_1 = d^2[J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N) + c^2J(\omega_N)] \qquad (10)$$

$$1/T_2 = R_2 = 0.5d^2[4J(0) + J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H) +$$

$$6J(\omega_H + \omega_N)] + (1/6)c^2[3J(\omega_N) + 4J(0)] + R_{ex} \quad (11)$$

$$NOE = 1 + (\gamma_H / \gamma_N)d^2[6J(\omega_H + \omega_N) - J(\omega_H - \omega_N)]/R_1 \quad (12)$$

where $d = \mu_0 h\gamma_H\gamma_N\langle r_{NH}^{-3}\rangle/(8\pi^2)$ for dipole-dipole interaction; c (CSA) = $\omega_N (\sigma_{\|} - \sigma_{\perp})/\sqrt{3}$; μ_0 is the permeability of free space; *h* is Planck's constant; γ_H , γ_N are the gyromagnetic ratios; ω_H, ω_N are the Larmor frequencies of ^1H and ^{15}N nuclei respectively; r_{NH} is the internuclear ^1H - ^{15}N distance (1.02 Å), $\sigma_{\|}$, $\sigma_{\perp}$ are the parallel and perpendicular components of the chemical shift tensor [$\sigma_{\|} - \sigma_{\perp}$ = - 170 ppm] on average (*78, 79*).

A typical ^{15}N model-free analysis on backbone N-H vectors of a protein usually includes the following steps: (i) Collect basic relaxation data (T_1, T_2, ^{15}N {^1H} steady-state NOE) by NMR; (ii) Collect T_1, T_2, or NOE at a second magnetic field (i.e. 500/600MHz). This step is desirable for yielding accurate S^2 at more mobile sites; (iii) Determine global hydrodynamics for the protein, using software such as TENSOR2 (*80*); (iv) Select spectral density function models for each residue (*81*). The five expressions of spectral density functions of model-free are listed in Table II-2 (*77, 82*). Comparing relaxation-compensated T_2 at a fast and slow CPMG π-pulsing rate helps identify sites with μsec to

msec timescale of motions (*76*); (v) Optimization and error analysis using Monte Carlo simulations; (vi) Analyze results - amplitudes and timescales of backbone motions for each residue.

II.2.3 Other NMR-based methods

II.2.3.1 Relaxation dispersion analysis

Carr-Purcell-Meiboom-Gill (CPMG) based relaxation dispersion NMR measurements probe protein motions on μsec to msec timescale quantitatively and with higher sensitivity than the traditional transverse relaxation experiments (*83*). The effective decay rate R_2^{eff} of a bond vector of a protein residue increases as the external field strength increases. Relaxation dispersion profiles, R_2^{eff} versus υ_{cpmg}, can be measured. By fitting the relaxation dispersion profiles, one can derive the conformational exchange term (R_{ex}), the exchange rate (k_{ex}), the relative populations of the exchanging states (p_A and p_B), and the chemical shift changes of residues ($\Delta\omega$) (*74, 84*).

$$R_{ex} = p_A p_B \frac{\Delta\omega^2}{k_{ex}} \left(1 - \frac{4\nu_{CPMG}}{k_{ex}} \tanh \frac{k_{ex}}{4\nu_{CPMG}}\right) \qquad (13)$$

II.2.3.2 Reduced spectral density function analysis

Reduced spectral density function analysis is commonly used to probe the fluctuations of each N-H bond of protein residues at frequency ω_0, ω_N, and $0.87\omega_H$. Using the following equations, one can calculate the values of the spectral densities at the above three frequencies (*85, 86*).

$$\sigma_{NH} = R_1 (\text{NOE} - 1)\, \gamma_N / \gamma_N \qquad (14)$$

$$J(0) = (6R_2 - 3R_1 - 2.72\sigma_{NH})/(3d^2 + 4c^2) \qquad (15)$$

$$J(\omega_N) = (4R_1 - 5\sigma_{NH})/(3d^2 + 4c^2) \qquad (16)$$

$$J(0.87\omega_H) = 4\omega_{NH}/(5d^2) \qquad (17)$$

where $d = \mu h \gamma_N \gamma_H r^{-3}$ and $c = \omega_N \Delta\sigma \times 3^{-1/2}$. Heteronucleus NOE is defined as the ratio of the peak height of the saturated state and that of the nonsaturated state (I_{sat} / I_{nonsat}). A $J(0)$ value of lower than $^2/_5\tau_m$ indicates sub-nsec internal flexibility of an N-H bond. An increased $J(0)$ value indicates μsec to msec timescale of fluctuation. The μsec to msec motions do not affect the values of $J(\omega_N)$ or $J(\omega_H)$ spectral densities. Therefore, $J(\omega_N)$ or $J(\omega_H)$ may be used as a control to confirm the μsec to msec motions observed in the $J(0)$ spectral densities. Large $J(\omega_H)$ spectral densities represent psec to nsec timescale of motions (*86, 87*).

II.3 Investigations of ligand interactions of proteins by NMR

NMR can be employed to investigate associations between proteins or between a protein and its ligand. In the context of NMR, "chemical exchange" means any process in which a certain nucleus changes between two or more different local chemical environments and during the process, the NMR parameters (i.e. chemical shift) of the nucleus differ. Chemical exchange of protein residues can be the consequence of ligand binding perturbations, and it can be divided into three categories based on the exchange rate: fast, intermediate, or slow. Different timescales of chemical exchange lead to different characteristic effects (i.e. changes in line shape or chemical shift) on an NMR spectrum. For example, for an ionization process accompanied by a change in chemical shift of 250 Hz, an exchange rate of >1000 sec^{-1} is regarded as "fast"; an exchange rate around 250 sec^{-1} is considered as "intermediate"; and an exchange rate of < 100 sec^{-1} is considered as "slow" (*88*). Figure II-3 illustrates changes between two equally populated environments, in chemical shifts and linewidths in the presence of a chemical exchange from slow to fast.

NMR titrations can be used to determine the binding affinity of the interaction of two molecules and to map the ligand binding site of a protein. In an NMR titration experiment, the ligand is added into the protein solution and a series of ^{1}H-^{15}N HSQC (heteronuclear single quantum correlation) spectra are collected. In the HSQC spectrum of a protein, each amino acid residue (except proline) provides one peak that corresponds to its backbone N-H group. Ligand binding pertubations will change the chemical environment of protein residues that interact with the ligand. The chemical shifts of these ligand contacting residues will be changed in proportion to the amount of protein bound

by the ligand. By measuring the chemical shift differences (see equation 2 of Chapter III) between the ligand saturated state and free state of a protein, residues with larger chemical shift changes can be indentified, and the ligand binding site of the protein can be determined (See III.3.5 and III.4.7 for examples). Binding isotherms can also be derived by plotting chemical shift changes versus ligand concentrations. The affinity of the interaction can be determined by fitting the curve, using non-linear regression analysis shown below (See III.3.4 and III.4.5 for details).

$$\delta_{obs} - \delta_P = (\delta_{PL} - \delta_P)\frac{(K_d + L_t + P_t) - \sqrt{(K_d + L_t + P_t)^2 - 4P_t L_t}}{2P_t} \qquad (18)$$

II.4 Evolutionary trace analysis

Evolutionary Trace (ET) analysis is a bioinformatics method to highlight functionally important residues for a protein. ET analysis relies on two principles. First, the functional sites of a protein are more conserved than other surfaces of the protein. Second, a protein family can be divided into functional classes based on sequence, since differences in protein functions correlate with systematic variations in sequence (*89, 90*). Using a multiple sequence alignment, the phylogenetic tree (Figure II-4) is generated, and protein sequences are divided into classes, with the branching determined by the "trace" level of sequence identity. Class consensus residues are those conserved within a class. Conserved residues are class consensus sequences conserved throughout all the classes

(Figure II-4). Class-specific residues are class consensus sequences that vary among some or all of the classes. Nearly class-specific residues are, in most cases, class consensus sequences that vary among some or all of the classes. Class-specific and nearly class-specific residues may be important for the signature functions of a protein. When mapped upon a representative 3D protein structure, class-specific and conserved residues may cluster at the functional sites (*89, 91*). See III.4.6 for examples.

Table II-1: NMR can probe backbone fluctuations on various timescales

Timescales	Experiments	Significance
psec – nsec nsec psec->1/10 msec	S^2 order parameter from model-free τ_e, $J(\omega_N)$ RDC	Ligand-recognitions
μsec – msec	Line broadening, R_{ex} from model-free Relaxation dispersion	Recognition, allostery, catalysis.
minutes - days	$^1H/^2H$ exchange, protection factor	Stability, global & local

Table II-2: Expressions used to fit NMR relaxation data to the extended Lipari-Szabo forms of the spectral density functions (*77, 82*).

Expression	Fit parameters	Spectral density function
1	S^2	$J(\omega)=2/5\{(S^2\tau_m)/[1+(\omega\tau_m)^2]\}$
2	S^2, τ_e	$J(\omega)=2/5\{(S^2\tau_m)/[1+(\omega\tau_m)^2]+[(1-S^2)\tau]/[1+(\omega\tau)^2]\}$
3	S^2, R_{ex}	$J(\omega)=2/5\{(S^2\tau_m)/[1+(\omega\tau_m)^2]\}$
4	S^2, τ_e, R_{ex}	$J(\omega)=2/5\{(S^2\tau_m)/[1+(\omega\tau_m)^2]+[(1-S^2)\tau]/[1+(\omega\tau)^2]\}$
5	S_s^2, S_f^2, τ_e	$J(\omega)=(2S_f^2/5)\{(S_s^2\tau_m)/[1+(\omega\tau_m)^2]+[(1-S_s^2)\tau]/[1+(\omega\tau)^2]\}$

Table II-3: Typical motional timescales for biomolecules. (*69*)

fsec to psec	psec to nsec	nsec to μsec	μsec to sec
Molecular vibrations	Molecular and bond rotations	Chemical exchange	Chemical exchange Macroscopic diffusion Flow

Figure II-1: Fourier transform converts FID from time domain to frequency domains.

(This figure is taken from Hornak, J. P. *The basics of NMR*, http://www.cis.rit.edu/htbooks/nmr/bnmr.htm)

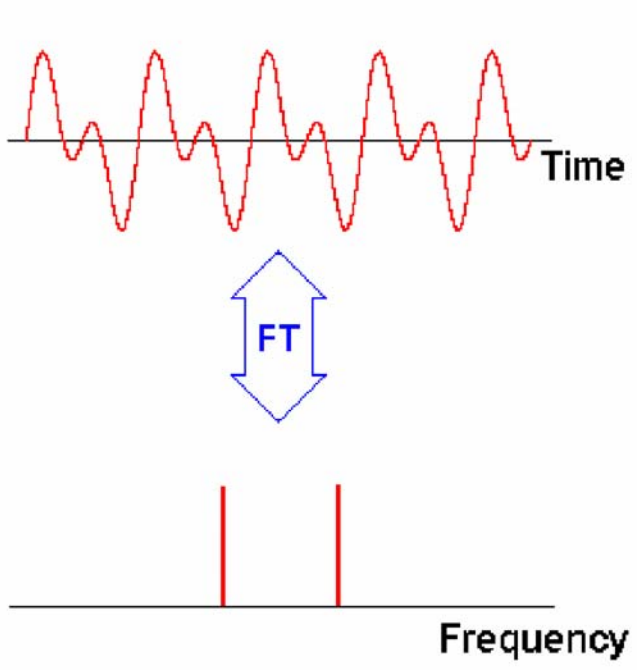


Figure II-2: Two-dimensional Fourier transform (*65*). (This figure is taken from Hornak, J. P. *The basics of NMR*, http://www.cis.rit.edu/htbooks/nmr/bnmr.htm)

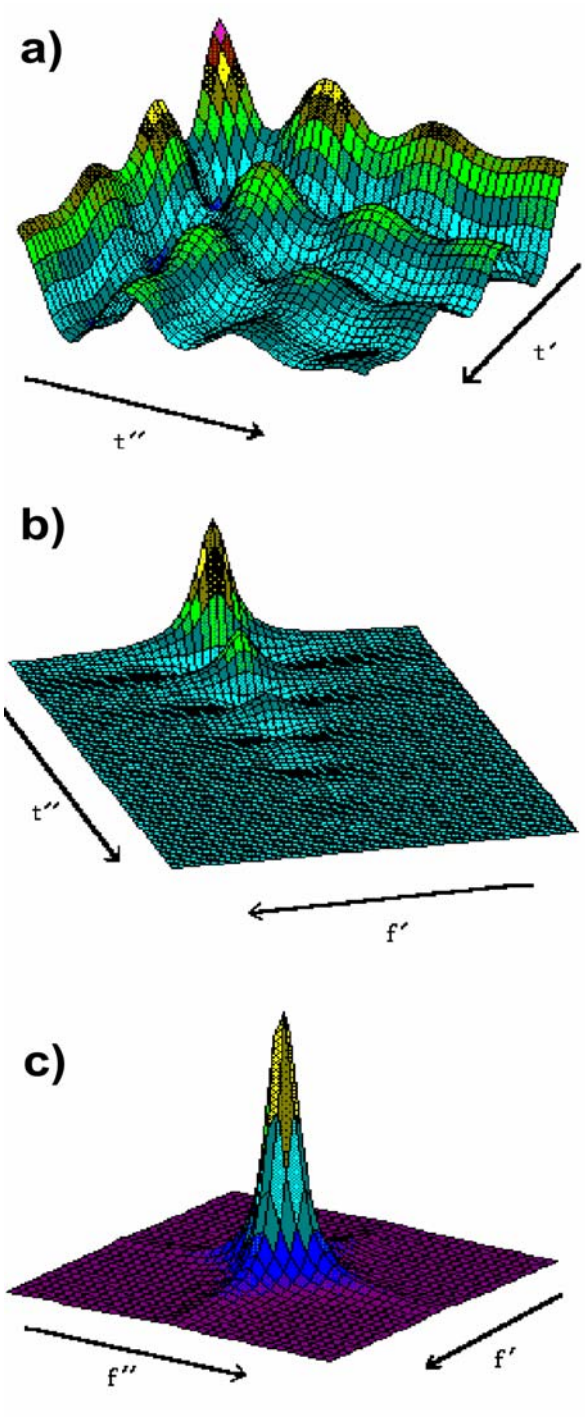


Figure II-3: Change in chemical shifts and linewidths in the presence of chemical exchange between two equally populated environments. Bottom to top: increasing rates of chemical exchanges.

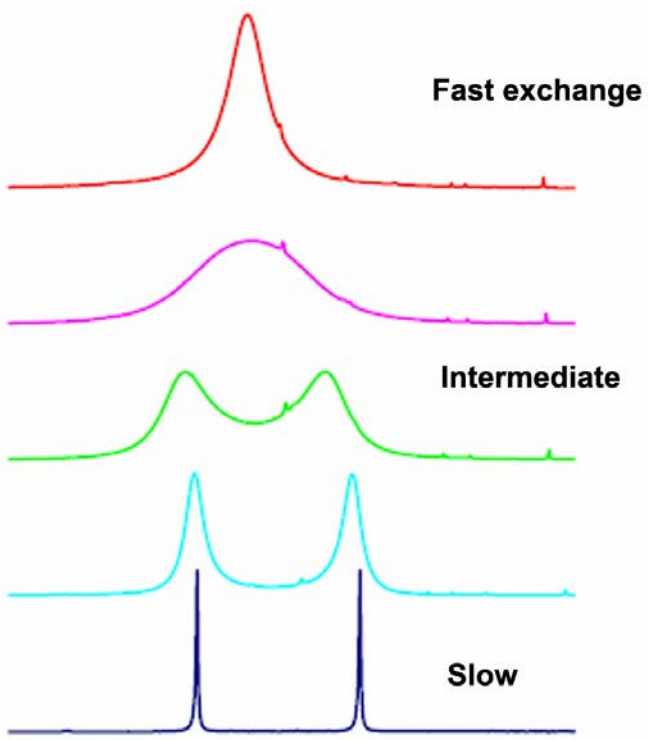
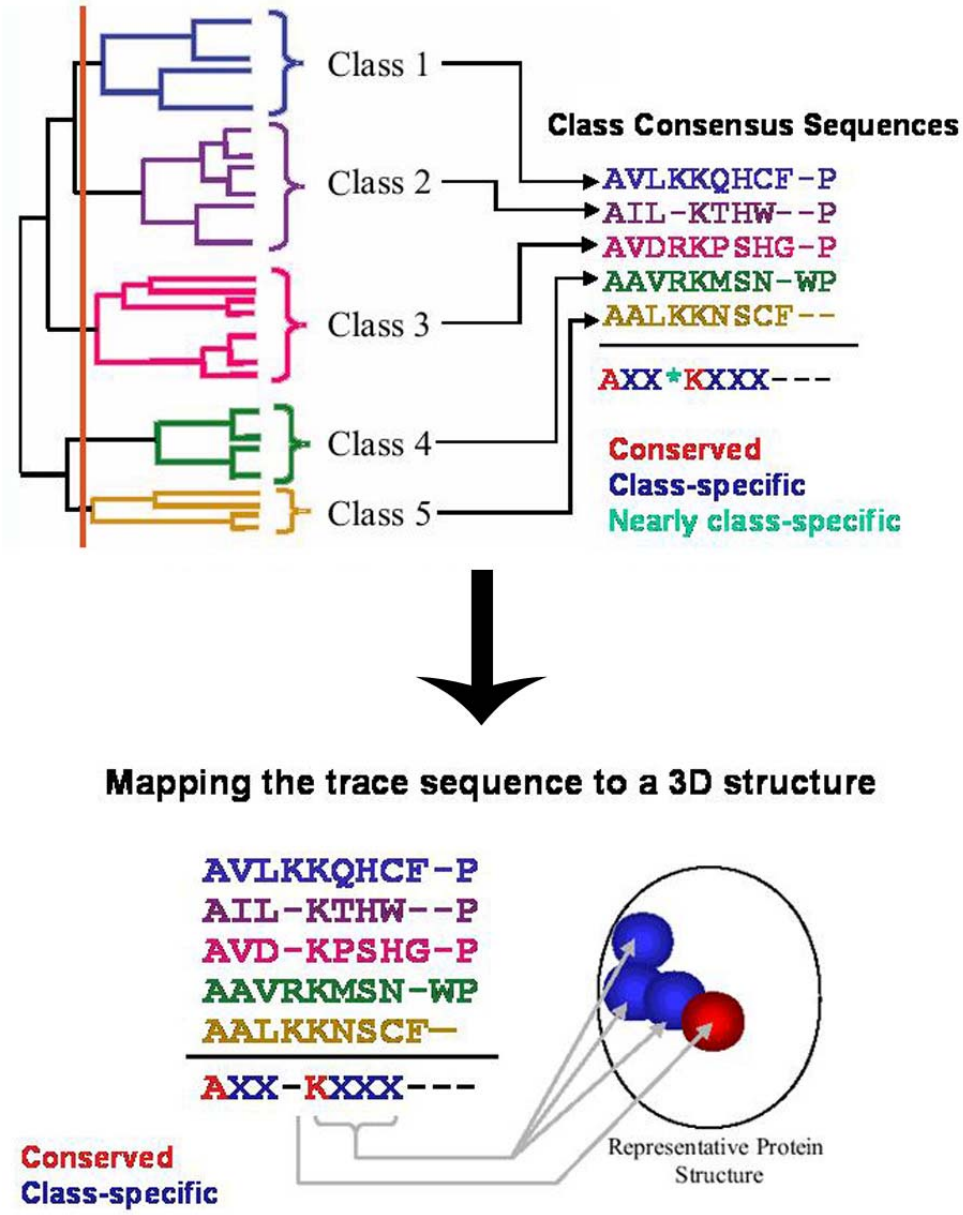


Figure II-4: Evolutionary trace analysis. At a certain trace level, protein sequences are partitioned into classes or subfamilies. Sequences conserved within each subfamily are called class consensus sequences. Class consensus sequences conserved through all the classes are the conserved residues (red). Class consensus sequences that vary among the classes are class-specific (blue). Class consensus sequences that vary among nearly all classes are nearly class- specific (green). Conserved, class-specific residues and nearly class-specific residues can be mapped onto the 3D structure of a protein to identify the functional surface shared by the same protein subfamily (*91*).

(This figure is taken the website of Lichtarge, O., Baylor College of Medicine, http://ncmi.bcm.tmc.edu/homes/msowa/ET.html)



Chapter III: Phosphoprotein and Phosphopeptide Interactions with KI-FHA

NMR titration experiments of GST-BRI1's binding with ^{15}N KI-FHA were performed by Dr. Xiangyang Liang. I have used his data for a comparison with the results of my phosphopeptide NMR titrations. After a major revision by Prof. Steven Van Doren and me, this research has been submitted to *Biochemistry* (*23*).

III.1 Abstract

FHA domains are phosphoThr recognition modules found in diverse signaling proteins, including kinase-associated protein phosphatase (KAPP) from Arabidopsis. Kinase-Interacting FHA domain (KI-FHA) of KAPP targets it to function as a negative regulator of some receptor-like kinase (RLK) signaling pathways important in plant development and environmental response. To aid in the identification of potential binding sites for the KI-FHA domain we predicted the functional surfaces of RLK kinase domains using

Evolutionary Trace analysis. We selected phosphopeptides from KAPP-binding Arabidopsis RLKs for *in vitro* studies of association with KI-FHA. The results cast doubts on KI-FHA binding either to the activation loop or juxtamembrane region. Three phosphoThr peptide fragments from the kinase domain of CLV1 or BAK1 were found to bind KI-FHA with K_D values of 8 to 40 μM, by NMR or titration calorimetry. Their affinity is driven by favorable enthalpy and the solvation entropy. Mutagenesis of these three threonine sites suggests Thr546 in the C-lobe of BAK1 kinase domain to be a principal but not sole site of KI-FHA binding *in vitro.* ^{15}N-labeled KI-FHA was titrated with GST-BRI1 kinase domain and monitored by NMR. BRI1 interacts with the same 3/4, 4/5, 6/7, 8/9, and 10/11 recognition loops of KI-FHA, with similar affinity as the phosphoThr peptides.

III.2 Introduction

Plants use RLKs to sense developmental signals and changing environmental conditions and to trigger signaling cascades to bring about developmental and adaptive responses (*92*). RLKs such as CLV1, BRI1, and BAK1 are critical to plant growth and development (*20, 93-96*). BRI1 and BAK1 are essential components of the Brassinosteroid receptor (BR) in brassinosteroid-dependent signal transduction in Arabidopsis (*20, 94-96*). CLV1, BRI1 and BAK1 are membrane-localized and belong to the leucine-rich repeat (LRR) family of RLKs of Arabidopsis *thaliana*. Each of these RLKs contains a receptor domain,

a single transmembrane helix, and an intracellular serine/threonine protein kinase domain (*97, 98*).

KAPP is a common negative regulator of RLKs in Arabidopsis (*38*). KAPP was found to interact with several Arabidopsis RLKs, including HAESA (*38*), RLK4 (*31*), CLV1 (*44*), FLS2 (*41*), and SERK1 (*42*). KAPP interacts with RLKs autophosphorylated on serine and/or threonine residues (*42, 43, 99, 100*). The KI-FHA domain of KAPP binds phosphorylated RLKs *in vitro* and fails to bind dephosphorylated RLKs (*31, 38, 44, 46*). It has been hypothesized that KAPP utilizes its KI-FHA domain to bind the phosphorylated RLK kinase domain to place the PP2C domain in proximity to dephosphorylate phosphoSer/Thr residues of RLK kinase domains on the intracellular face of the plasma membrane. This results in decreased RLK kinase activity and attenuated RLK signaling (*17*). Much of how KAPP binds some RLKs and how KAPP-RLK signaling complexes assemble remains unknown. Even with peptide library screening data available for the FHA domain of KAPP, it remains unclear what site(s) in an RLK are recognized by the KI-FHA domain of KAPP.

We searched for potential KI-FHA binding sites on KAPP-binding RLKs by screening peptide fragments of the kinase domains containing phosphoThr/Ser. We describe the association of three pThr peptides with KI-FHA and their location in the predicted structure of the CLV1 kinase domain. *In vitro*, the FHA domain of KAPP interacts with the kinase domain of either BRI1 or BAK1. NMR-detected titrations indicate that the

BRI1 kinase domain and phosphoThr peptides bind the same surface of KI-FHA with qualitatively similar affinity.

III.3 Materials and methods

III.3.1 Evolutionary trace analysis

A total of 301 sequences of FHA domains, spanning β-strands 3–10, and aligned at www.sanger.ac.uk/cgi-bin/Pfam/getacc?PF00498, were combined with 12 Arabidopsis FHA sequences from the Munich Information Center for Protein Sequences (MIPS) Arabidopsis thaliana database (http://mips.gsf.de/proj/thal). To improve the multiple sequence alignment (MSA) using CLUSTALW (www.ebi.ac.uk/clustalw), we removed the most divergent FHA domains, such as those of kinesins. Truncated sequences that interfere with MSA were deleted. The MSA was submitted to the ET server (www-cryst.bioc.cam.ac.uk/~jiye/evoltrace/evoltrace.html). The dendrogram was divided into 10 trace levels of sequence identity. Residues that are conserved within each subfamily, but vary among subfamilies, are considered class-specific residues at the chosen trace level.

III.3.2 Preparation of KI-FHA and GST fusion with BRI1 kinase domain

GST-tagged KI-FHA, containing residues 180-313 of Arabidopsis KAPP, was expressed in *E. coli* and purified as described (*101*). The N-terminal GST tag was removed with 3C protease (PreScission^{TM}, GE Healthcare). The sequence encoding the kinase domain from residue 816 to 1196 of Arabidopsis BRI1 (GI: 2392895) was amplified by PCR and subcloned into the vector pET42a as described (*102*). The fusion protein was expressed in BL21 in LB medium for 3 hours at 28 ˚C after induction with 250 μM IPTG. The following protease inhibitors were added to the cell lysate: 1 μM Pepstatin A, 10 μM Leupeptin, 0.1 mM TPCK, 0.5 mM AEBSF and 5 μM E-64 (Sigma). After disrupting the bacteria culture with a French Press, the soluble supernatant was dialyzed against PBS buffer (pH 7.4) to eliminate endogenous glutathione. The fraction was incubated with glutathione agarose (Sigma) at 4 ˚C. BRI1 was eluted with 30 mM reduced glutathione (Sigma) in 50 mM Hepes (pH 7.4). Purified GST-BRI1 was dialyzed and concentrated to 0.36 mM (24.6 mg/ml) in 20 mM sodium phosphate (pH 6.3) with 120 mM NaCl and 7% D_2O for NMR-detected titration.

III.3.3 Isothermal titration calorimetry of KI-FHA's interactions with phosphoThr/Ser peptides

43 phosphopeptides (Table III-1), chosen based on the criteria described in Results, were synthesized as peptides with phosphorylation of the central threonine or serine as described (*76*). A positive control for KI-FHA binding was also synthesized using idealized sequence of: AAYAYpTQASAAKKK (*62*). Forty-four phosphoThr/Ser

peptides (40-70% pure) (Table III-1) were initially assayed for binding to KI-FHA by ELISA and subsequently by isothermal titration calorimetry (ITC). Phosphopeptides that interact with KI-FHA were purified to > 85% by HPLC and used for further ITC with KI-FHA from KAPP.

The experimental design and data analysis of ITC have been described (*103-106*). We used a VP-ITC MicroCalorimeter from MicroCal, Inc. To ensure the same pH, KI-FHA and the phosphopeptides were dialyzed into PBS buffer (pH 7.5) prior to ITC, using Float-A-Lyzer® devices with 500 Da cutoff (Spectrum Laboratories). Titrations were conducted at 25 ˚C. GST-KI-FHA or free KI-FHA samples, with concentrations ranging from 0.03-0.1mM, were placed in the sample cell. The phosphopeptide solutions, with concentrations in 5 to 10 molar excess over KI-FHA concentration, were placed in the syringe and repeatedly injected at constant intervals. In order to estimate heat of mixing, a control titration was performed by titrating the phosphopeptide into buffer only. Baseline correction was done by subtracting the heat changes from the peptide-to-buffer control titration from the peptide-to-KI-FHA titration. After baseline correction, the sigmoidal curves were fit to an equation for an association with one binding site using MicroCal Origin version 5.0. This revealed ΔH, ΔG, $T\Delta S$ and stoichiometry (N-value) for the association (Table III-2).

III.3.4 ^{15}N KI-FHA NMR titrations with phosphopeptides

^{15}N-labeled KI-FHA was prepared with a concentration of 0.53 mM to 0.6 mM in 20 mM sodium phosphate buffer (pH 6.3), with 120 mM NaCl and 9% D_2O. Phosphopeptides were dissolved in the same buffer. The pH of each phosphopeptide stock solution was measured using a micro-pH electrode and adjusted to 6.3. Undissolved materials were removed by centrifugation. The concentrations of the peptide stock solutions were determined to be of 7 to 8 mM by quantitative amino acid analysis. The purified phosphoThr peptides were then titrated into ^{15}N-labeled KI-FHA at 22 ˚C. A Varian Inova 600 MHz spectrometer was used with a high-sensitivity 5 mm ^{1}H cryogenic probe with actively shielded Z-gradient coil. A series of 2D ^{15}N TROSY spectra were collected with molar ratios of ^{15}N KI-FHA to peptide of 1:0, 1:0.25, 1:0.5, 1:0.75, 1:1, 1:1.33, 1:1.67, 1:2, 1:2.5, 1:3, and 1:4. NMR spectra were processed using NMRPipe 2.3 (*107*). The titration curves were analyzed using nonlinear regression analysis (*88, 108*). Chemical shift changes of a KI-FHA residue were plotted against total peptide concentration for each titration. Binding affinity was determined by fits to the following equation:

$$\delta_{obs} - \delta_P = (\delta_{PL} - \delta_P)\frac{(K_d + L_t + P_t) - \sqrt{(K_d + L_t + P_t)^2 - 4P_t L_t}}{2P_t} \qquad (1)$$

P_t, and L_t are the total concentrations of the protein ^{15}N KI-FHA and peptide ligand, respectively. δ_{obs} is the observed NMR peak position or chemical shift; δ_P and δ_{PL} are the chemical shifts of the free KI-FHA and the complex. K_D is the dissociation constant.

Microcal Origin version 6.0 was used with its global fitting option to fit the binding isotherms of all affected residues' ^{1}H peak positions simultaneously.

III.3.5 Mapping phosphopeptide and GST-BRI1 binding sites on KI-FHA by NMR

For mapping the sites on KI-FHA where the phosphopartners bind, the ^{15}N TROSY spectra in absence and 4-fold excess were compared. The chemical shift changes of the amide peaks were calculated as the radius:

$$\Delta\omega_{NH} = (\Delta\omega_H^2 + (\Delta\omega_N/6)^2)^{1/2} \qquad (2)$$

where $\Delta\omega_H$ and $\Delta\omega_N$ are the changes in ppm in the ^{1}H and ^{15}N dimensions. The factor of six normalizes the shifts in the ^{15}N dimension down to the scale of the changes in the ^{1}H dimension. For mapping the site of binding of BRI1, ^{15}N-TROSY spectra were collected at 22°C and 600MHz ^{15}N KI-FHA, free and in the presence of additions of GST-BRI1. The conditions used in order of nominal KI-FHA to BRI1 ratio, KI-FHA concentration, and NMR signal averaging time were: 1:0, 0.4 mM, 0.5 hr; 1:0.25, 0.31 mM, 1 hr; 1:0.5, 0.26 mM, 8 hr; and 1:0.75, 0.22, 32 hr. With the greater additions of GST-BRI1 kinase domain to KI-FHA, GST-BRI1 kinase domain precipitated heavily.

III.3.6 Conformational entropy change upon binding: crude estimates from NMR order parameters

NMR order parameters, S_{LZ}, are available for KI-FHA, free and bound to pT868 CLV1 (*76*). Changes of configurational entropy ΔS_{conf} between bound (B) and free (F) states were proposed to relate to the Lipari-Szabo order parameter S_{LZ} describing psec to nsec bond reorientations (*109*):

$$\Delta S_{conf} = S_{conf,\,\mathbf{B}} - S_{conf,\,\mathbf{F}} = N\,k\,\ln((3-(1+8S_{LZ,\,\mathbf{B}})^{1/2})/(3-(1+8S_{LZ,\,\mathbf{F}})^{1/2})) \qquad (3)$$

where k is the Boltzmann constant and N is Avogadro's number. However, such estimates can only be considered qualitative due to many limitations listed by Arumugam et al. (*75*).

III.3.7 Far-western assays

Site-directed mutations were introduced to the kinase domains of CLV1 and BAK1 using the QuikChange mutagenesis kit (Stratagene, Cat. 200518) as described by the manufacturer. Sequencing confirmed the mutations. MBP-CLV1, MBP-mCLV1 (K720E), MBP-CLV1 (T868A), MBP-mBAK1 (K317E), GST-BAK1, GST-BAK1 (T312A), GST-BAK1 (T546A), MBP-BRI1 and MBP-mBRI1 (K911E), were expressed and purified as described (*110*). mCLV1 (K720E), mBAK1 (K317E) and mBRI1 (K911E) are inactive mutant forms of receptor-like kinases unable to autophosphorylate. Purified MBP-BRI1 and MBP-mBRI1 (K911E) were autophosphorylated *in vitro* with non-radioactive ATP (*20*) and loaded onto identical SDS-PAGE gels. One gel was stained. The bands of the other gel were transferred onto nitrocellulose (NC) membrane. The NC membrane was then blocked with 5% non-fat milk (w/v) and rinsed with PBS-T

buffer. 1 μg of GST-KI-FHA was added and incubated with the NC membrane overnight in 4˚C. Anti-glutathione S-transferase (Sigma) and anti-rabbit (Amersham) antibodies were used. Far-western assays were also performed on the purified MBP-CLV1, MBP-mCLV1 (K720E), MBP-mBAK1 (K317E), GST-BAK1, GST-BAK1 (T312A) and GST-BAK1 (T546A) as described above, except that a polyclonal antibody directed against KAPP (*111*) was used.

III.4 Results

III.4.1 Functional surfaces predicted for LRR-RLK protein kinase domain

No experimental structure of an RLK protein kinase domain has been reported so far. Such a structural model would aid exploration of potential surfaces that KAPP might recognize. Dr. Jeff Skolnick predicted the tertiary structure of the CLV1 kinase domain that has been well characterized in development, biochemistry, and recognition of KAPP (*24, 26, 44, 111-114*). According to the predicted structural model (Figure III-1), the CLV1 kinase domain has two subdomains, the N-lobe and the C-lobe. The N-lobe contains a 5-stranded β-sheet and two α-helices. The C-lobe begins with two helices followed by two central β-hairpins. Following the second β-hairpin in the C-lobe is the activation loop that precedes seven more α-helices (*23*).

For a set of 61 Arabidopsis RLKs, comprising the LRR II and LRR X-XIII families (*13*) , I conducted Evolutionary Trace analysis to forecast regions of their kinase domains important to function. This set includes five RLKs that interact with KAPP: CLV1, HAESA, SERK1, BAK1 and BRI1. The ET approach identifies "class-specific" sites that differ among subfamilies at functionally important sites (*89*), in this case among LRR-RLKs. Such positions that systematically differ may tune the specificities that distinguish subfamilies (*89, 115*). I reasoned that perhaps at least one of these loci might contribute to a KAPP binding site. At least ten class-specific segments are exposed on the surface of the structural model of the kinase domain of the representative family member CLV1 (Figure III-1). The surface-exposed, class-specific sites (Figure III-2) are found in kinase subdomains I through V of the N-lobe, as well as subdomains VIa – IX and XI of the C-lobe (Figure III-3). The Roman numbering scheme for kinase subdomains has been described (*116*).

III.4.2 Phosphopeptides from RLKs screened for interaction with KI-FHA domain of KAPP

I sought experimental evidence for sites on or near RLK kinase domains where KI-FHA might bind. Mapping using mutagenesis of RLK kinase domains from CLV1, HAESA, BAK1 or BRI1 expressed in bacteria was hampered by limited stability, solubility, expression, and for some mutants their kinase activity. We found that a phosphopeptide-directed approach made mapping more tractable. We screened candidate binding sites

among KAPP-binding RLKs that include BRI1, FLS2, RLK4, SERK1, CLV1 and BAK1. As models of candidate binding sites for KI-FHA, we selected 43 threonine or serine-containing peptides from their kinase domains (Table III-1) and phosphorylated them during synthesis.

I selected a first set of threonine or serine-containing sequences (Table III-1) from BAK1 and CLV1 as candidates for phosphorylation and screening. I considered four criteria to suggest potential for KAPP binding. First, we favored sequences with residues in the pT-3 to pT+3 positions preferred by KAPP (*62*) (see legend of Table III-1). Second, I favored threonines or serines found by Evolutionary Trace analysis to be class-specific (Figure III-2 and Table III-1), i.e. characteristic of RLK subfamilies (*89*) (Figure III-1). Third, I only considered sequences where the threonine or serine is likely to be on the surface of the RLK and available to bind KAPP, judged by inspecting the structural model of CLV1 (Figure III-1). Fourth, I preferred that a candidate's threonine or serine be conserved among most of the RLKs observed to bind KAPP, i.e. CLV1, HAESA, BAK1, BRI1, SERK1, WAK1, FLS2, and KIK1.

Reports of phosphorylation sites (*59, 99, 117*) suggested a second set of peptides for screening for affinity for KI-FHA. *In vivo* phosphorylation sites were identified in Arabidopsis membrane proteins that include ~50 RLKs (*117*). About 75% of these phosphorylation sites in RLKs are in the juxtamembrane region (between transmembrane helix and kinase domain) and at the C-terminus. *In vitro* and *in vivo* phosphoThr/Ser

sites were identified in recombinant BRI1 kinase domain (*59, 99*). We considered threonine and serines of the C-terminal fragment of the FLS2 kinase domain reported to interact with KAPP (*41*) and to be phosphorylated *in vivo* (*117*).

None of the phosphopeptides selected from activation loops of RLKs interact with KI-FHA, i.e. pS848 CLV1, pT449 BAK1, pT450 BAK1, pT1039 BRI1, pS1042 BRI1 (Table III-1). This suggests that the activation loop may not be recognized by the KI-FHA domain portion of KAPP. Also, KI-FHA binds none of the peptides screened simply because they are known to be sites of phosphorylation (Table III-1).

III.4.3 Locations in RLKs of phosphoThr peptides that bind KI-FHA

3 of the 43 peptides (Table III-1) were found to have unambiguous affinity for KI-FHA. These are phosphoThr peptides spanning residues 863-875 of CLV1 (named pT868 CLV1), residues 307-317 of BAK1 (named pT312 BAK1), residues 542-553 of BAK1 (named pT546 BAK1) (Table III-2). The locations of these three KI-FHA binding pThr peptides are marked on the CLV1 structural model (Figure III-4a). pT312 BAK1 is located in the N-lobe at a loop and β-strand at the beginning of kinase subdomain II. The kinase subdomains are indicated on the alignment of RLKs in Figure III-8. pT868 CLV1 lies in the C-lobe at the end of kinase subdomain VIII; this is in the first loop after the

activation loop. Peptide pT546 BAK1 lies in kinase subdomain XI, closer to the C-terminus (Figure III-4). All three of the peptides have residues (in pT-3 through pT+3 positions) observed to be preferred from a combinatorial library for binding KAPP (*62*). Notably, pT868 CLV1 and pT546 BAK1 coincide with sites suggested by Evolutionary Trace analysis to be important in function (Figures III-1 and III-4).

III.4.4 Energetics of binding of phosphopeptides to KI-FHA

The interactions of KI-FHA with the purified pThr peptides, including pT312 BAK1, pT546 BAK1, pT868 CLV1, and a positive control pThr peptide, were characterized by isothermal titration calorimetry (ITC). ITC is a biophysical technique used to determine the thermodynamic parameters of interactions in solution. In an ITC experiment, KI-FHA sample were input into sample cells, and peptide solutions were placed in the syringe. Peptide was titrated into KI-FHA solutions, and the heat change of each titration was monitored. The integrated heat was plotted against the ligand-to-protein molar ratio, to generate a binding isotherm (Figure III-5). Thermodynamics parameters can then be obtained by nonlinear regression analysis (Table III-2). The positive control (Table III-2) uses residues preferred in the pT-3 through pT+3 positions from the affinity assays of a combinatorial library (*62*); see Table III-1. The titrations' raw binding isotherms and baseline-corrected, integrated points with best fits are displayed in Figure III-5. The *N*-value of each of the titrations is near one (Table III-2), suggesting 1:1 stoichiometry. At pH 7.5 and 25 ºC, the dissociation constants between KI-FHA and the three phosphoThr

peptides derived from RLKs range from 8 to 40 μM (Table III-2). The K_D for KI-FHA and the control phosphoThr peptides with optimized sequence is 4 μM. The enthalpic change provides from about half to about three-quarters of the favorable free energy of association in each case (Table III-2). The four pThr peptides bind KI-FHA with favorable entropic terms $-T\Delta S$ ranging from -1.8 to -3.5 kcal·mol^{-1} (Table III-2).

The net entropy gain linked to binding is interesting considering the loss of entropy expected from the loss of freedom of diffusion on binding, loss of flexibility of the peptide once it binds, and the net increase of rigidity observed in the backbone of KI-FHA once pT868 CLV1 binds (*76*). We investigated whether release of ordered water might contribute to the entropy gain favoring the association. We performed calorimetric titrations of pT868 CLV1 with KI-FHA at 10, 14, 18 and 25 ºC. ΔH_{obs} drops linearly with increasing temperature. ΔC_p from that slope is -230 ± 8 cal K^{-1} mol^{-1} (Figure III-6). This ΔC_p can be used to estimate crudely the entropy of solvation from the expression $\Delta S_{solv}=\Delta C_p \ln(T/T_s^*)$, where T_s^* is the reference of 385 K (*118, 119*). This suggests a rough estimate of the solvation entropy of binding at 25 ºC of -17.6 kcal mol^{-1}. This is consistent with a large net release of ordered water molecules upon binding. This desolvation effect may be large enough to enhance affinity significantly, by offsetting the entropic costs from the increased structural rigidity in the complex.

III.4.5 Binding of phosphopeptides to KI-FHA characterized by NMR titrations

We also monitored the affinity and binding sites of KI-FHA for the four phosphoThr peptides. A series of ^{15}N TROSY spectra were collected with ^{15}N KI-FHA-to-peptide molar ratio of 1:0, 1:0.25, 1:0.5, 1:0.75, 1:1, 1:1.33, 1:1.67, 1:2, 1:2.5, 1:3, and 1:4. In each NMR titration, the additions of the pThr peptides shifted progressively the amide peak positions of KI-FHA affected until ^{15}N KI-FHA was saturated (Figure III-7). The binding isotherms were constructed by plotting ^{1}H chemical shift changes against total peptide concentrations (Figure III-7); these two parameter are designated δ_{obs} and L_t , respectively in equation 1. Each K_D was obtained by globally fitting each titration with equation 1. ^{15}N KI-FHA becomes saturated when each pThr peptide reaches about a two-fold molar excess. The RLK-derived peptides bind ^{15}N KI-FHA with K_D values of 8 to 40 μM (Figure III-7). The K_D with the control peptide is 3.4 μM. Despite the lower pH of 6.3, the NMR-derived Gibbs free energies of association agree within 0.2 kcal mol^{-1} with those from ITC, except for pT546 BAK1 (Table III-2).

III.4.6 Evolutionary trace analysis of phosphoprotein-binding surface of KI-FHA

FHA domain surfaces important in recognition and specificity were first predicted by Evolutionary Trace analysis (ET) (*89*) (see II.4 for an introduction of ET). Class-specific residues may confer distinctive specificities to FHA family members. ET used an alignment of 209 sequences of FHA domains (Figure III-8) that excludes divergent kinesins. Because of the great sequence diversity among FHA domains, clusters of class-

specific residues did not emerge until the 9th highest of ten trace levels, i.e., at sequence identities of > 74 % within each class (*49*). Many class-specific FHA residues at trace level 9 (Figure III-9) lie at or near the phoshoprotein-binding surface of KI-FHA, i.e., Arg212, Ser214 - Ala219, Lys221 – Ser223, Val225 – His229, and Met246 – Leu253 (blue in Figure III-10*a* and Figure III-9). Highly conserved residues of the FHA domains and residues flanking these were shown in other FHA domains to contact phosphopeptide ligands and are seen to be class-specific.

III.4.7 Recognition loops of KI-FHA identified by NMR titrations

KI-FHA residues with large chemical shift changes cluster at the apparent phosphopeptide binding site. The maximum amide peak shifts resulting from saturating additions of the pThr peptides were measured radially using equation 2. Binding of pT546 BAK1 peptide causes significant chemical shift changes (Figure III-12a, d) at Gly211-Val213 of the 3/4 loop; Lys221-Asp222, Val225 and Gly227-His229 of the 4/5 loop; Met246 and Ser248-Thr252 of the 6/7 loop; Ser260, Asp263, Gly265, Arg267 and Trp269 of the 8/9 loop; Gly284-Thr287 of the 10/11 loop. Binding of the control pThr peptide causes peak shifts of similar size and location (Figure III-12b, e) at Gly211-Val213 of the 3/4 loop; Lys221-Ser223, Val225, Gly227 and His229 of the 4/5 loop; Ser248-Gly251 of the 6/7 loop; Ser260, Asp263, Gly265, Arg267 and Trp269 of the 8/9 loop; Gly284 and Thr286 of the 10/11 loop. The pT312 BAK1 and pT868 CLV1 peptides cause similar effects on amide spectra of ^{15}N KI-FHA (*49, 76*). We verified the previous hypothesis that residues that are class-specific among FHA domains (blue in Figure III-

10*a*) via ET analysis or conserved in KAPP of plants (yellow in Figure III-10*a*) interact with RLK partners. There is excellent correspondence between the class-specific residues (blue in Figure III-10*a*) and KAPP-conserved residues (yellow in Figure III-10b) and the pT peptide-perturbed surface (red and pink in Figure III-10b). The 3/4, 4/5, 6/7, 8/9 and 10/11 loops are consistently implicated in the phosphopeptide recognition surface of KI-FHA, regardless of pThr peptide ligand.

III.4.8 Evaluation of the three prospective binding sites by mutagenesis

We investigated whether any of the three pThr peptides identified correspond to a site that KI-FHA can bind in an intact RLK kinase domain. We introduced single alanine substitutions for Thr868 of CLV1, Thr312 of BAK1 or Thr546 of BAK1. We screened the MBP or GST fused protein with each mutation for capture of KI-FHA of KAPP using Far-Western assays. Bindings were detected by a primary antibody against KAPP (*43*). The T868A mutant of MBP-CLV1 does not differ from wild-type MBP-CLV1 in capture of KI-FHA within the uncertainty. Results are compared for fusion proteins of BAK1 kinase domain with the wild-type, the inactivated, the T312A and the T546A variants (Table III-3 and Figure III-11). GST-BAK1 can clearly be seen to bind KI-FHA *in vitro*. The T312A mutant consistently fails to differ from wild type BAK1 kinase domain in binding of KI-FHA, within the uncertainties. The T546A variant of BAK1, however, consistently retains only about 40% binding to KI-FHA. For *E. coli*-expressed GST-BAK1 *in vitro*, the greatest share of KI-FHA from KAPP appears to bind at Thr546 in BAK1. KI-FHA evidently also binds elsewhere in the recombinant BAK1 kinase domain.

Or other Thr/Ser residues located close to Thr546 could also contribute to the binding of BAK1 to KI-FHA.

III.4.9 Surface of KI-FHA that interacts with the kinase domain of BRI1

BRI1 kinase domain was demonstrated to interact with KI-FHA *in vivo* and *in vitro* in a phosphorylation dependent manner (*23*). We used NMR to examine the interaction of an FHA domain with a large globular domain of a physiological phosphoprotein binding partner. For the globular partner, we selected the 43 kDa kinase domain of BRI1 shown in Figure III-12c and 12f to interact with KAPP. We found BRI1 to be stable and soluble to 0.36 mM (25 mg/ml) when fused to GST. However, GST is well-known to dimerize, doubling the anticipated MW of the GST-BRI1 construct to about 136 kDa. Association of such a large construct with ^{15}N-labeled KI-FHA of 15 kDa was expected to introduce considerable line broadening to amide spectra of KI-FHA. To compensate for such challenges, we exploited the ^{15}N line narrowing of ^{15}N TROSY detection, the high sensitivity of a cryogenic probe and long signal averaging (see Materials and Methods for details). ^{15}N-^{1}H TROSY spectra of ^{15}N KI-FHA were collected without and with the addition of 0.25, 0.5 and 0.75 equivalents of GST-BRI1. Since GST-BRI1 precipitated partially and increasingly with greater additions to KI-FHA, the actual BRI1 to KI-FHA molar ratios were $\leq$ 0.25:1, < 0.5:1 and < 0.75:1. The size of the peak shifts in the KI-FHA spectrum increased with the additions of GST-BRI1, implying the fast chemical exchange regime. This suggests moderate affinity of GST-BRI1 for KI-FHA similar to that observed in the pThr peptide titrations also in fast exchange. The radial chemical

shift differences between ^{15}N TROSY spectra of free ^{15}N KI-FHA and KI-FHA in the presence of nominally 0.75 equivalents of GST-BRI1 are plotted in Figure III-12c. The smaller size of these chemical shift differences, compared to pThr peptide titrations, is consistent with the *inability* to saturate the KI-FHA with the GST-BRI1 at the high concentrations that exceeded 0.2 mM for KI-FHA. Significant chemical shift changes of KI-FHA residues occurred at Gly211-Arg212 of the 3/4 loop; Leu220-Asp222, Val225 and Lys228-His229 of the 4/5 loop; Asp245-Ser248 and Asn250 of the 6/7 loop; His261, Asp263-Gly265 and Arg267 of the 8/9 loop; Leu283, Thr285-Thr286 and Lys288 of the 10/11 loop (Figure III-12c,f). Thus, the surface of KI-FHA that binds the BRI1 kinase domain is composed of the same 3/4, 4/5, 6/7, 8/9 and 10/11 loops that bind the pThr peptides. In addition, the BRI1 construct introduces small chemical shift perturbations distant from this surface at the far ends of β–strands 3, 8, 9 and 11 and neighboring 7/8 and 9/10 loops. These peak shifts far from the main binding surface suggest either conformational adjustment at long-range or perhaps some non-specific interaction.

III.5 Discussion

III.5.1 Prospective binding site for KI-FHA in C-lobe of BAK1

Mutagenesis of the Thr546 site of BAK1 appears to disrupt reproducibly more than half of KI-FHA binding to the BAK1 kinase domain expressed in *E. coli*. This binding occurs despite the appearance that Thr546 may be partly buried and only partly exposed on the

surface of the predicted structural model in Figure III-4. The phosphorylation of Thr546 is unlikely to allow this side chain to remain buried, due to its charge. Presumably phosphorylated Thr546 is genuinely on the surface, either due to its charge or due to uncertainty inherent to the predicted structural model. The Thr546 site in the C-lobe lies on the same side of BAK1 kinase domain as the activation loop. This prospective binding site for the FHA domain would place its neighboring PP2C domain from KAPP very well for dephosphorylating the activation loop of the kinase. Considering the length and potential flexibility of the linker joining the FHA and PP2C domains of KAPP (*49*), the *in vitro* binding site for KI-FHA seems neither too close nor too far from the activation loop of BAK1 for the PP2C domain to reach the activation loop readily. The Thr546 position is conserved (Figure III-3) with these KAPP-binding RLKs: HAESA, SERK1, WAK1 and FLS2, but not with BRI1 or CLV1. The prospective KAPP binding site in the C-lobe of one or more RLKs is consistent with the precedent of diverse protein kinases being inhibited by protein partners docked to their respective C-lobes (*120*).

Almost 40% of KI-FHA binding signal remaining using T546A-substituted BAK1 (Table III-3) suggests that KI-FHA may bind to one or more other phosphorylated sites as well. The three phosphopeptide fragments of RLKs found to bind KI-FHA are consistent with its ability to bind some breadth of sequences. Isolated KI-FHA is monomeric (*76*). Yet it is possible that full-length KAPP might form a dimer with its dimeric RLK partner. If so, the FHA domain of a second chain of KAPP in a complex might be able to bind an additional and different region of an RLK.

III.5.2 Phosphopeptides that fail to bind KI-FHA may narrow down alternatives for binding sites

About 75% of the flagellin-elicited phosphorylation of RLKs *in vivo* was found in the juxtamembrane and C-terminal regions (*117*). Those observations coupled with the importance of the phosphorylated juxtamembrane region of type I TGF-β receptor in a regulatory interaction with an FHA-like domain of SMAD2 (*121*), suggested the question: Might KI-FHA of KAPP interact analogously with phosphorylated juxtamembrane domains of RLK partners? PhosphoThr/Ser peptides from BRI1 corresponding to its sites of phosphorylation in the juxtamembrane region (*59, 102*) show at best only very weak and ambiguous binding to KI-FHA (Table III-1). This casts considerable doubt on the juxtamembrane region binding KAPP. C-terminal phosphoSer peptides from FLS2 are similarly marginal in affinity for KI-FHA. The phosphorylated peptides from activation loop sequences we screened also failed to bind KI-FHA; see S852, S857 of CLV1 and T449, T450, T455 of BAK1 in Table III-1. To sum up, the known sites of RLK phosphorylation tested lack affinity for KI-FHA.

III.5.3 Energetics of pThr peptide binding to KI-FHA

The Gibbs free energy of pThr peptide affinity for KI-FHA, with K_D values of the order of 10 μM, comes at least half from favorable enthalpy of binding (Table III-2). The favorable enthalpy can be attributed to optimal van der Waals close contacts, hydrogen-bonding and electrostatics at the interface. The KI-FHA residues (conserved or near the pT+3 position) that are most important for affinity for phosphopartners are remarkably rigid (*76*). Such rigidity was hypothesized to promote enthalpically favorable van der Waals contacts with the phosphopeptide (*76, 122, 123*).

It is noteworthy that the affinities of the pThr peptides for KI-FHA are also entropically favored. This comes despite the unfavorable entropic costs of flexible peptide becoming more rigid upon binding, the loss in diffusional degrees of freedom, and the net increase in rigidity of the backbone of KI-FHA (*76*). Just the entropic cost, ΔS_{conf}, of increased rigidity of KI-FHA upon binding is considerable. This entropic contribution $-T\Delta S_{conf}$ is estimated qualitatively to be 14 $kcal \cdot mol^{-1}$ unfavorable, using equation 3 and the reported changes in rigidity of KI-FHA upon binding pT868 CLV1 (*75*). Even though such an estimate is very crude due to several limitations (*75*), it is sufficient to point out a significant energetic impediment to binding of the pThr peptide. How can this cost be paid such that the entropic term actually reverses and promotes binding? This cost would appear to be paid by a still larger favorable entropy gain, which worths roughly -18 $kcal \cdot mol^{-1}$ (see above), from ordered waters released upon binding. The large size of this solvation entropy and the heat capacity may raise a question as to whether water is released from sites in addition to the interface. It appears that the favorable desolvation

effect may suffice to overcome the increased rigidity of KI-FHA to provide net entropy gain to promote binding.

III.5.4 Attributes predictive of KI-FHA binding site in an RLK

The success in finding a prospective KI-FHA binding site in an RLK kinase domain via three phosphopeptides found to bind KI-FHA provides a retrospective on the best sequence predictors. Sequence similarity to KAPP's residue preferences among a library of pThr peptides (*62*) clearly correlates with the phosphopeptides found to bind KI-FHA (Table III-1). Evolutionary Trace analysis also anticipated the functional importance of the binding site found at Thr546 of BAK1. ET also suggests functional importance of the Thr868 segment of CLV1 not confirmed by assay of the lone T868A point mutant. ET has been very successful in predicting functional sites (*115, 124*) and can be applied more readily than screening of peptide libraries (*62*).

In summary, the surface of KI-FHA from KAPP that recognizes BRI1's kinase domain coincides with the surface that recognizes phosphoThr peptides (Figure III-12). The affinity of phosphoThr peptides is driven by the enthalpy of favorable contacts and probably also by the desolvation effect. KI-FHA failed to bind phosphopeptides tested from either the activation loop of three RLKs or the phosphorylated sites of the juxtamembrane region of BRI1. Instead, KI-FHA *in vitro* appears to bind Thr546 from BAK1, in the C-lobe on the same side of the kinase domain as the activation loop. This

prospective binding site for the FHA domain of KAPP would appear to support positioning of the PP2C domain of KAPP strategically for dephosphorylation of BAK1's activation loop.

III.6 Acknowledgments

We thank Prof. Michael Henzl for providing access to the VP-ITC MicroCalorimeter and for helpful directions. We thank Prof. George P. Smith for instructions on ELISA assays. We are grateful to Dr. Xuelu Wang for the GST-BRI1 fusion construct.

Table III-1: Summary of properties of peptides from KAPP-binding RLKs

Peptides[1]	Position[2]	Class[3]	Preferred aa[4]	Surface[5]	Conserved[6]	Kd
		Specific		access	Thr or Ser	(μM)
CLV1 and BAK1 peptides selected based on sequence properties						
MAAYAYpTQASAAKKK	control	N/A	Best	N/A	N/A	4.0
CLV1						
KRLVGRGpTGRSDHGF	T727	n	1.9(+3)	y	4	No[7]
VGRGTGRpSDHGFTAE	S730	y	N/A	y	1	No[8]
GFTAEIQpTLGRIRHR	T740	y	1.3(+1)	y	4	No[7]
YEYMPNGpSLGELLHG	S773	y	1.3(+1)	y	6	No[7]
AYLHHDApSPLILHRD	S810	y	N/A	y	3	No[8]
ILHRDVKpSNNILLDS	S820	y	N/A	n	3	No[8]
SNNILLDpSDFEAHVA	S827	n	N/A	n	3	No[8]
FLVDGAApSEAMSSIA	S848	n	1.4(-2)1.3(+1)	y	3	No[7]
GAASEAMpSSIAGSYG	S852	y	1.9(+3)	y	6	No[7]
AMSSIAGpSYGYIAPE	S857	n	N/A	N/A	7	No[7]
IAPEYAYpTLKVDEKS	T868	y	1.4(-3)1.3(-1)1.3(+1)	y	5	16
TLKVDEKpSDVYSFGV	S875	y	N/A	n	5	No[7]
DEKSDVYpSFGVVLLE	S879	y	1.3(-1)	n	5	No[7]
EEAAARPpTMREVVHM	T959	n	1.4(-3)	y	3	No[7]
BAK1						
KGRLADGpTLVAVKRL	T312	n	1.4(-3)1.3(+1)1.7(+3)	y	4	9.2
KRLKEERpTQGGELQF	T324	n	1.5(+1)	y	4	No[7]
QTEVEMIpSMAVHRNL	S339	y	N/A	y	3	No[8]
YPYMANGpSVASALRE	S370	y	1.4(-3),1.9(+3)	y	6	No[7]
RQRIALGpSARGLAYL	S397	y	1.4(-3)	y	2	No[8]
DYKDTHVpTTAVRGTI	T449	y	N/A	y	2	No[8]
YKDTHVTpTAVRGTIG	T450	y	N/A	y	7	No[7]
VTTAVRGpTIGHIAPE	T455	y	1.2(+1)	N/A	7	No[7]
HIAPEYLpSTGKSSEK	S465	y	N/A	y	1	No[8]
IAPEYLSpTGKSSEKT	T466	y	1.4(-3),1.9(+3)	y	5	No[7]
YLSTGKSpSEKTDVFG	S470	y	1.3(+1)	y	4	No[8]

TGKSSEKpTDVFGYGV	T473	y	N/A	n	6	No[7]
VMLLELIpTGQRAFDL	T487	y	N/A	y	6	No[7]
IQVALLApTQSSPMER	T546	y	1.5(+1),1.9(+3)	y	4	9.9
QIENEYPpSGPR	S612	y	N/A	N/A	1	No[8]
BRI1, SERK1 & RLK4 peptides selected based on phosphorylation of threonine or serine						
YAEGHGNpSGDRTANN	S838(JM)[10]	y	N/A	N/A	1	220[9]
HGNSGDRpTANNTNWK	T842(JM)[10]	n	N/A	N/A	1	1150
GDRTANNpTNWKLTGV	T846(JM)[10]	n	1.4(-3)	N/A	1	370[9]
TGVKEALpSINLAAFE	S858(JM)[10]	y	N/A	y	2	200[9]
EKPLRKLpTFADLLQA	T872(JM)[10]	n	N/A	y	2	80[9]
RLMSAMDpTHLSVSTL	T1039	n	1.4(-3),1.9(+3)	n	3	No[7]
SAMDTHLpSVSTLAGT	S1042	y	N/A	n	3	1300
SERK1						
KLMDYKDpTHVTTAVR	T459	n	1.4(-3)	n	3	770
DYKDTHVpTTAVRGTI	T462	y	N/A	n	3	1160
RLK4						
NVIVNpSDpTLGEK	S465,T467	N/A	1.3(+1)	y	1	No[7]
FLS2 peptides selected from C-terminal 60-residue peptide that binds KAPP						
LRQLVEKpSIGNGRKG	S1096	N/A	N/A	y	2	No[7]
MELGDpSIVpSLKQEE	S1115,1118	N/A	N/A	y	1	144
MELGDpSIVSLKQEE	S1115	N/A	N/A	y	1	60[9]
MELGDSIVpSLKQEE	S1118	N/A	N/A	y	1	1000

1: phosphoThr/Ser peptides of KAPP-binding RLKs

2: The position of the phosphorylated Thr/Ser residue in the sequence of the full length RLK

3: Whether the phosphorylated Ser/Thr residue is class-specific in the Evolutionary Trace analysis of the 61 LRR-RLK kinase domains

4: The preferred amino acids of KI-FHA in each pThr peptide position (i.e. -3,-2,-1, +1,+2,+3) deduced using the phosphoThr peptide library (*62*).

Position	-3	-2	-1	pT	+1	+2	+3
Selectivity	A(1.4),Y(1.4), F(1.3)	A(1.4)	Y(1.3),F(1.2)	---	V(1.5),Q(1.5),E(1.3), L(1.3),I(1.2)	---	S(1.9),A(1.7)

GST tagged KI-FHA domain was used to screen a peptide library containing the sequence MAXXXXpTXXXXAKKK where X indicates all amino acids except Cys and Trp. Enrichment values, represented in numerical scores (i.e. 1.3, 1.7), were obtained by normalizing the affinities of the optimal amino acids by the average affinities of the remaining amino acids (*63*).

[5]: Surface accessibility according to the predicted structural model of CLV1 kinase domain (Figures III-1, and 2)

[6]: Number of Thr or Ser residues conserved among the seven aligned sequences of KAPP-binding RLKs, including CLV1, HAESA, BAK1, BRI1, WAK1 and FLS2 from Arabidopsis and KIK1 from maize

[7]: No binding to KI-FHA in ITC assays, but appeared to interact with GST-KIFHA in ELISA assays.

[8]: No binding to GST-KIFHA in ELISA experiments

[9]: Binding is no longer detectable after subtracting heat change of the peptide-to-buffer titration from that of the peptide-to-protein titration.

[10]: Juxtamembrane region of RLK kinase domain

Table III-2: Thermodynamics of phosphopeptide associations with KI-FHA monitored by ITC and NMR

pThr peptide	RLADGT(p)LVAVK aa307-317	PEYAYT(p)LKVDEKS aa863-875	ALLAT(p)QSSPMER aa542-553	AAYAYT(p)QASAAKKK
Name, RLK	**pT312 BAK1**	**pT868 CLV1**	**pT546 BAK1**	**Control**
K_D, μM NMR [1]	8.0 ± 1.6	20.0 ± 2.6	40.0 ± 4.5	3.4 ± 2.0
ΔG, kcal/mol NMR [1]	-6.9 ± 0.6	-6.3 ± 0.3	-5.9 ± 0.2	-7.4 ± 2.8
ΔG, kcal/mol ITC [2]	-7.1 ± 0.4	-6.5 ± 0.3	-6.8 ± 0.4	-7.4 ± 0.7
ΔH kcal/mol	-4.3 ± 0.1	-4.5 ± 0.3	-3.3 ± 0.1	-5.6 ± 0.2
-TΔS kcal/mol	-2.8 ± 0.2	-2.0 ± 0.3	-3.5 ± 0.1	-1.8 ± 0.2
***N*-value ITC [2]**	0.81 ± 0.02	1.0 ± 0.05	0.79 ± 0.02	0.83 ± 0.01
K_D, μM ITC [2]	9.2 ± 1.1	16.0 ± 2.0	9.9 ± 1.3	4.0 ± 0.6

[1] NMR titrations were done at pH 6.3.

[2] ITC experiments were done at pH 7.5.

Table III-3: Comparison of KI-FHA binding by threonine-directed mutations of BAK1 kinase domain

	GST-BAK1	MBP-BAK1, inactivated by K317E [2]	GST-BAK1 T312A	GST-BAK1 T546A
Relative capture of KI-FHA[1]	100%	6.5 ± 9%	106 ± 57%	38 ± 13%

[1] See Figure III-6 for an example of one of the three replicates of Far Western assays quantified and averaged here. The density of the KI-FHA and antibody-probed RLK band was normalized by the density of the Coomassie-stained RLK band of an equally loaded companion gel. The normalized density is expressed as a percentage of the wild-type, positive control band.

[2] Mutation of this conserved and essential lysine in protein kinases including RLKs abrogates activity (*20*). This lesion makes this construct a negative control, since auto-phosphorylation of the RLK's kinase domain is necessary for KI-FHA to bind.

Figure III-1: Stereo view of the predicted structure of the kinase domain of CLV1, This model was predicted by Shashi Pandit and Dr. Jeff Skolnick (*23*). CLV1, representing LRR-RLKs, is colored to mark regions predicted to be important for function. The backbone coordinates of this threaded structural model in PDB format have been listed in Appendix 2. From the Evolutionary trace analysis of LRR-RLK kinase domains (Figure III-2), the class-specific or subfamily-characteristic residues are colored yellow and the conserved residues red. Class-specific Thr/Ser residues are represented with yellow spheres and conserved ones with red spheres. The activation loop is magenta. The N-lobe and C-lobe are colored cyan and blue, respectively. The N-terminal juxtamembrane region and C-terminal tail, defined previously (*102, 116*), are gray.

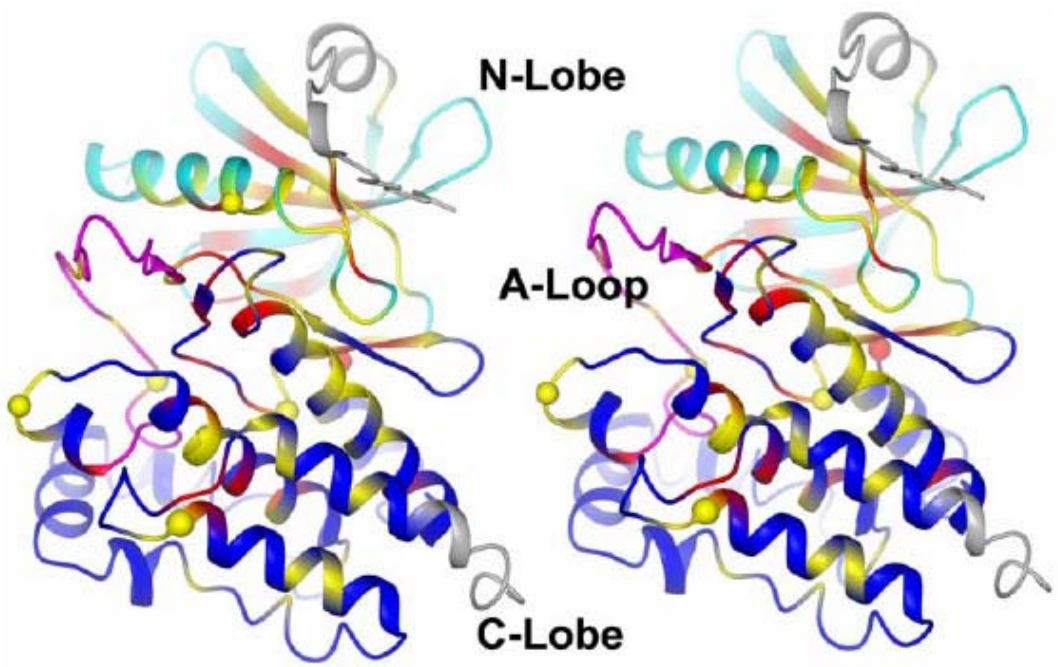


Figure III-2: Class-specific residues of RLK kinase domain sequences at trace level 7. Sixty-one sequences of Arabidopsis RLKs, comprising the LRR II and LRR X-XIII families (*13*), were partitioned into the 10 classes present at trace level 7. Class consensus residues conserved through all classes are conserved (red). "Class-specific" residues (yellow) are cases where class consensus residues differ in at least one class.

```
Group 1: At1g55610
Group 1: At3g13380

Group 2: At1g31420
Group 2: At2g35620

Group 3: At5g16000
Group 3: At3g25560

Group 4: At4g30520
Group 4: At2g23950

Group 5: At5g10290
Group 5: At5g65240

Group 6: At1g71830
Group 6: At1g34210
Group 6: BAK1
Group 6: At2g13800
Group 6: At2g13790

Group 7: AT5G07180
Group 7: At5g62230

Group 8: AT4G26540
Group 8: AT5G56040

Group 9: At4g20140
Group 9: At5g44700

Group 10: At3g49670
Group 10: At5g65700
```

```
GROUP_1          PSA-.-M-GSGGFG-VYKA-L-D..GSVVAIKKLI--TGQGDR....EFMAEMETI.GKIKHRN
GROUP_2          LNEE.HIIGCGGFGTVYKL-MDD..G-VFALKRI-KLNEGFDR....FFERELEILG.SIKHRY
GROUP_3          PSSK.NL-GKGG-GNVYKG-L-D..----AVKRLKD----GGE-...QFQTE-EMIS.LAVHRN
GROUP_4          PSSK.-ILGAGGFGNVYRGK-GD..GT-VAVKRLKD-NGTSG-S...QFR-ELEMIS.LAVH-N
GROUP_5          PSEK.NVLGQGGFGKVYKG-L-D..-TKVAVKRLTDFE-PGGD-...AFQREVEMIS.VAVHRN
GROUP_6(BAK1)    PS--.N-LG-G-FG--YKGRLAD..--LVAVKRL-EERT-GGEL...QFQTEVEMIS.MAVHRN
GROUP_7          L-EK.-IIGYGASSTVYKC--K-..SRPIAIKR-YNQYP-N-R....EFETELETIG.SIRHRN
GROUP_8          LTSA.NVIGTGSSGVVYR-TIPS..GE-LAVKKMWSKEE-.......AFNSEI-TLG.SIRHRN
GROUP_9          L-EE.FMIGSGGSGKVYKAEL-N..GET-AVKKILWKDDLMSN...KSP-REVKTLG.-IRHRH
GROUP_10         LKED.NIIGKGGAGIVYKG-MP-..GDLVAVKRLA-MS-GSSH..DHGFNAEIQTLG.RIRHRH
SUMMARY          X-------G-G--X--YX----------AXKX----------XXXP--E--XXXX-XXH-X

GROUP_1          LVPLLGYCK-GEER..LLVYEYMK-GSLETVLHEK-.KKGGI-L-W-ARKKIAIGAARGLAFLH
GROUP_2          LVNLRGYCNSPTSK..LLLYDYLPGGSLDEALH-RG..EQ...LDWDSRVNIIIGAAKGL-YLH
GROUP_3          LLRLYGFC-T--E-..LLVYFYMSNGSVASR-KA......KPVLDW--RKRIA-GA-RGL-YLH
GROUP_4          LLRLIGYCA-S-ER..LLVYFYM-NGSVAS-LK-......KPALDWN-RK-IAIGAARGL-YLH
GROUP_5          LLRLIGFCTTQTER..LLVYFPMQNLS-A--LREIK..-GDFVLDW--RK-IALGAARG-EYLH
GROUP_6(BAK1)    LLRLRGFCMTPTER..LLVYFYMANGSVASCLRERP..-----L-W--R--IALGSARGL-YLH
GROUP_7          IVSLHGYALSP-GN..LLFYDYMENGSLWDLLHG--..KK.VKLDWETRLKIAVGAAQGLAYLH
GROUP_8          I-RLLGWCSNRNLK..LLFYDYLPNGSLSS-LHG.A.GKG-G--DWEARYDVVLGVAHALAYLH
GROUP_9          LVKLMGYCSSK--GLNLLIYEYM-NGS-WDWLH-----KKK--L-WE-RL-IA-GLAQGVEYLH
GROUP_10         IVRLLGFCSNHETN..LLVYEYMPNGSLGEVLH....GKKGGHLHW-TRYKIALEAAKGLCYLH
SUMMARY          X-XLXGXX------XXLLXYXXX--XS-X---X------------W--R--XX-XX-XX--XLH

GROUP_1          HSCIPHIIHRDMKSSNVLLD-DF-ARVSDFGMARLVSALD.....THLSVSTLAGTPGYVPP.E
GROUP_2          HDCSPRIIHRDIKSSNILLDGNLEARVSDFGLAKLLEDEE.....SHITT.IVAGTPGYLAP.E
GROUP_3          EQCDPKIIHRDVKAANILLDDY-EAVVGDFGLAKLLDH-......SHVTT.AVRGTVGHIAP.E
GROUP_4          EQCDPKIIHRDVKAANILLDE-FEAVVGDFGLAKLLNH-D.....SHVTT.AVRGTVGHIAP.E
GROUP_5          EHCNPKIIHRDVKAANVLLDEDFEAVVGDFGLAKLVDVRR.....TNVTT.QVRGTMGHIAP.E
GROUP_6(BAK1)    DHCD-KIIH-DVKAANILLDEEFEAVVGDFGLA-LM-Y-D.....-HVTT.AVRGTIGHIAP.E
GROUP_7          HDCTPRIIHRDIKSSNILLD-NFEA-LSDFGIAKSIPA-K.....T-AST.YVLGTIGYIDP.E
GROUP_8          HDCLP-I-HGDVKAMNVLLG--FE-YLADFGLA---SG---TD--K--NRPP-AGS-------E
GROUP_9          -DCVPPIVHRDIKSSNVLLDSN-EAHLGDFGLAK-LT-N-...DTNT-SNT-FA-SYGYIAP.E
GROUP_10         HDCSPLIVHRDVKSNNILLDSNFEAHVADFGLAKFLQ.....DSGTSECMSAIAGSYGYIAP.E
SUMMARY          -XCX--I-H-DXKXXNXLLX------XXDFGXA--------XX-----XXX--X-X-------E

GROUP_1          YYQS.FRCTAKGDVYSYGVILLELLSGKKP.IDP-EFGE..DNNLVGWAKQ......LYREKRG
GROUP_2          YMQS.GRATEKTDVYSFGVLVLEVLSGK-P.TD..ASFIEKG-N-VGWL-F......LISE-R-
GROUP_3          YLST.GQSSEKTDVFGFGILLLEL-TG-RA.-EFGKAANQ-.G--LDWVKK.......--QEKKL
GROUP_4          YLST.GQSSEKTDVFGFGILLLELITG-RA.LEFGK-VSQK.GAMLEWVRK......LH-EMKV
GROUP_5          --ST.GKSSE-TDVFGYGIMLLELVTGQRA.IDPSRLEEEDDVLLLDHVKK......LEREKRL
GROUP_6(BAK1)    YLST.GKSSEKTDVFGYG-MLLELITGQ-A.FDLARLANDDD-MLLDWVK-......-LKEKKL
GROUP_7          YART.SR-NEKSDIYSFGIVLLELLTGKKA.VD..NEANLHQ-ILS...........KADDNTV
GROUP_8          HASM.Q-ITEKSDVYSYGVVLLEVLTGKHP..LDPDLP..GGAHLV-WVRD.....HLA-KKDP
GROUP_9          YAYS.LKATEKSDVYSMGIVLMEIVTGKMP..T---F-..-E-DMVRWVET-L-----..--AR
GROUP_10         YAYT.LKVDEKSDVYSFGVVLLEL-TG-KP..VGEFG...DGVDIVQWVR....MTD.SNKD-V
SUMMARY          --XX-X--XX-XDXXXXG-XXXEX-XG--X--X-----XX-X---X-XX---X-----------

GROUP_1          AEILDPELVT.DKSGDV...EL-HYLKIASQCLDDRPFKRPTMIQ-M-MFKE------E-...
GROUP_2          --IVD--CEG...---ES...LDALLSIAT-CVS-SP-ERPTMHRVVQLLESEVMTPCP.S-FY
GROUP_3          E--VDK-L-----YD-IE...--EMV-VALLCTQYLP-HRPKMSEVVRMLEGDG.L-EKWEAS-
GROUP_4          EEL-DRELGT-..YD-IE...VGEMLQVALLCTQ-LPAHRPKMSEVV-MLEGDG.LAERWAASH
GROUP_5          --IVDK-LD--..YIKEE...VEMMIQVALLCTQ--PE-RP-MSEVVRMLEGEG.LAERWEEWQ
GROUP_6(BAK1)    E-LVD--L----..Y---E...-EQ-IQ-ALLCTQ-S-MERPKMSEVVRMLEGDG.LAE-W-EWQ
GROUP_7          MEAVD-EV-VT.CMD-GH...I-KTPQLALLCTKRNPLERPTM-EVSRVLLSLVPS----KKLP
GROUP_8          ---LDPRL-GR-D-.I.M.HEMLQTLAV-FLCVSNKA--RP-MKD-VAML-EIR--D--RSE--
GROUP_9          -KLID--LK-LLP..-EE.-AA-QVLEIALQCTK--PQERPSSRQA---LL-V-N.........
GROUP_10         LKV-D-RLSS-P.....-.HEVTHVFYVA-LCVEEQAVERPTMREVVQILTEIPK-P-SK-Q--
SUMMARY          ----D--X------------X--X-X-----CXX-----RP-X-X----X--------------
GROUP_1          ..SLDEP-LKETPLVEESRDKEP.......................................
GROUP_2          DSS.SD.........................................................
GROUP_3          ---------.................---NE--SSS-R...YSDLTDDSS-LVQAMELSGPR
GROUP_4          -HSHFYHAN-S--TI-S------N--T---PGSS-....F-D-DD-Q-LDSFAMELSGPR
GROUP_5          N-E......................VTR--.EF-RLQRR.FDWGED.S--NQDAIELSGGR
GROUP_6(BAK1)    K-E.....................-----------------W--------------SGPR
GROUP_7          S----....................--------E---.....D---QWFVQFRE-ISKSS
GROUP_8          -IK-..........G-C---....-PQ-----------GS..SNCSFA-SD-SV.....
GROUP_9          ......................................NR-A-Y----------......
GROUP_10         ----.......................TE-AP---E-SP-SG--SPPDLL--.........
SUMMARY          --------X-X--XX-X-------------------------------------XXXX
```

Figure III-3: Alignment of the kinase domains of the nine KAPP-binding RLKs, including CLV1 (At1g75820), BAK1 (At4g33430), BRI1 (At4g39400), SERK1 (At1g71830), FLS2 (AT5g46330), HAESA (AT4g28490), WAK1 (At1g21250) from Arabidopsis, KIK1 (U82481) from maize and CrRLK1 *(Z*73295) from *Madagascar periwinkle*. The class-specific residues from the evolutionary trace analysis of the sixty-one RLKs are labeled with "x", and the conserved residues labeled with "X". The Thr or Ser sites of CLV1, BAK1, FLS2 and BRI1 that were tested for binding to KI-FHA are colored in light blue. The threonines of the three pThr peptides that bind KI-FHA (pT868 CLV1, pT312 BAK1 and pT546 BAK1) are colored in darker blue. The Thr or Ser residues of BRI1, BAK1, SERK1 and FLS2, which have been identified to be phosphorylated *in vivo* or *in vitro*, are underlined and colored red. The Thr or Ser residues, which may have been phosphorylated *in vivo* or *in vitro*, are italicized and colored yellow (*59, 99, 117*). The conserved motifs I to XI (*125*) and activation loop of protein kinase domains are indicated below the sequences**.**

```
                x      X X  X  Xx         XxXx           x         x
BAK1    -------FSNKNILGRGGFGKVYKGRLADGTLVAVKRLKEERTQGGELQFQ---------
SERK1   -------FSNKNILGRGGFGKVYKGRLADGTLVAVKRLKEERTPGGELQFQ---------
KIK1    IRTATSNFSDSNKLGEGGFGPVYMGTLPGGEEVAVKRLCRNSGQG-LEEFK---------
CLV1    E-DVLECLKEENIIGKGGAGIVYRGSMPNNVDVAIKRLVGRGTGRSDHG---------FT
HAESA   H-EIADCLDEKNVIGFGSSGKVYKVELRGGEVVAVKKLNKSVKGGDDEYSSDSLNRDVFA
BRI1    LLQATNGFHNDSLIGSGGFGDVYKAILKDGSAVAIKKLIH-VSGQGDRE---------FM
WAK1    MKKATNGYAESRILGQGGQGTVYKGILPDNSIVAIKKARLGDSSQ--VEQFI--------
CrRLK1  -------FSENRVIGIGGFGKVYKGVFKDGTKVAVKRGISCSSSKQGLSEFR--------
FLS2    LEQATDSFNSANIIGSSSLSTVYKGQLEDGTVIAVKVLNLKEFSAESDKWFY--------
                          I              II

         X  xxx xxX xx xXxXxx       XXxxxxx  xX x   x             X
BAK1    TEVEMISMAVHRNLLRLRGFCMT-PTERLLVYPYMANGSVASCLRERPESQPPLDWPKRQ
SERK1   TEVEMISMAVHRNLLRLRGFCMT-PTERLLVYPYMANGSVASCLRERPPSQPPLDWPTRK
KIK1    NEVILIAKLQHRNLVRLLGCCIP-REEKILVYEYMPNKSLDAFLFN-PEKQRLLDWKKRF
CLV1    AEIQTLGRIRHRHIVRLLGYVAN-KDTNLLLYEYMPNGSLGELLHGSKGG--HLQWETRH
HAESA   AEVETLGTIRHKSIVRLWCCCSS-GDCKLLVYEYMPNGSLADVLHGDRKGGVVLGWPERL
BRI1    AEMETIGKIKHRNLVPLLGYCKV-GDERLLVYEFMKYGSLEDVLHDPKKAGVKLNWSTRR
WAK1    NEVLVLSQINHRNVVKLLGCCLE-TEVPLLVYEFITNGTLFDHLHG-SMIDSSLTWEHRL
CrRLK1  TEVELLSQFRHRHLVSLIGYCDE-KNEMIIIYEFMENGTLRDHLYG-SDKPK-LNWRKRV
FLS2    TEAKTLSQLKHRNLVKILGFAWESGKTKALVLPFMENGNLEDTIHG--SAAPIGSLLEKI
        III         IV                  V

         xx xx xx  xXX xXx  X X X XxxXxXXx      xxXXXxX
BAK1    RIALGSARGLAYLHDHCDPKIIHRDVKAANILLDEEFEAVVGDFGLAKL---MDYKDTHV
SERK1   RIALGSARGLSYLHDHCDPKIIHRDVKAANILLDEEFEAVVGDFGLAKL---MDYKDTHV
KIK1    DIIEGIARGLLYLHRDSRLRVVHRDLKASNILLDADMKPKISDFGMARM---FGGDQNQF
CLV1    RVAVEAAKGLCYLHHDCSPLILHRDVKSNNILLDSDFEAHVADFGLAKF---LVDGAASE
HAESA   RIALDAAEGLSYLHHDCVPPIVHRDVKSSNILLDSDYGAKVADFGIAKVG-QMSGSKTPE
BRI1    KIAIGSARGLAFLHHNCSPHIIHRDMKSSNVLLDENLEARVSDFGMARL---MSAMDTHL
WAK1    KIAIEVAGTLAYLHSSASIPIIHRDIKTANILLDVNLTAKVADFGASRLI-P--MDKEEL
CrRLK1  EICIGSAKGLHYLHTGTMKRIIHRDVKSANILLDENLMAKVADFGVSKTG-PDHFDQTHV
FLS2    DLCVHIASGIDYLHSGYGFPIVHCDLKPANILLDSDRVAHVSDFGTARILGFREDGSTTA
            VIa                 VIb                 VII    A-loop

        xx   x x      X  xxx      Xxxx X   xX      x
BAK1    TT-AVRGTIGHIAPEYLSTGKSSEKTDVFGYGVMLLELITGQRAFDLARLANDDDVMLLD
SERK1   TT-AVRGTIGHIAPEYLSTGKSSEKTDVFGYGIMLLELITGQRAFDLARLANDDDVMLLD
KIK1    NTNRVVGTFGYMSPEYAMEGIFSVKSDVYGFGVLILEIITGKRAVSFH--CHEDSLNIAG
CLV1    CMSSIAGSYGYIAPEYAYTLKVDEKSDVYSFGVVLLELIAGKKPVG----EFGEGVDIVR
HAESA   AMSGIAGSCGYIAPEYVYTLRVNEKSDIYSFGVVLLELVTGKQPTDS---ELGD-KDMAK
BRI1    SVSTLAGTPGYVPPEYYQSFRCSTKGDVYSYGVVLLELLTGKRPTDSP--DFGD-NNLVG
WAK1    ET-MVQGTLGYLDPEYYNTGLLNEKSDVYSFGVVLMELLSGQKALCFK--RPQSSKHLVS
CrRLK1  ST-AVKGSFGYLDPEYLTMQKLTEKSDVYSFGVVMLEILTGRPVIDPS--KPREMVNLVE
FLS2    STSAFEGTIGYLAPEFAYMRKVTTKADVFSFGIIMMELMTKQRPTSLND-EDSQDMTLRQ
            VIII                IX                                X

                                x  x              x x     Xxx     XX x
BAK1    WVKGLLKEKK-------LEALVDVDLQGN----YKDEEVEQLIQVALLCTQSSPMERPKM
SERK1   WVKGLLKEKK-------LEMLVDPDLQTN----YEERELEQVIQVALLCTQGSPMERPKM
KIK1    YAWRQWNEDN-------AAELIDPVIRAS----CSVRQVLRCIHIALLCVQDHADERPDI
CLV1    WVRNTEEEITQPSDAAIVVAIVDPRLTGY-----PLTSVIHVFKIAMMCVEEEAAARPTM
HAESA   WVCTALDKCG-------LEPVIDPKLDLK-----FKEEISKVIHIGLLCTSPLPLNRPSM
BRI1    WVK-QHAKLR-------ISDVFDPELMKEDP--ALEIELLQHLKVAVACLDDRAWRRPTM
WAK1    YFATATKENR-------LDEIIGGEVMNE----DNLKEIQEAARIAAECTRLMGEERPRM
CrRLK1  WAMKCSRKG---------EEIVDSDIVNE----VRPESLIKFQETAEKCLAERGVDRPTM
FLS2    LVEKSIGNGRKG-----MVRVLDMELGDSIVSLKQEEAIEDFLKLCLFCTSSRPEDRPDM
                                                                  XI

         x    x
BAK1    SEVVRMLEGDGLAERWEEWQKEEMFRQDFNYP-THHPAVSGWIIGDSTSQIENEYPS---
SERK1   SEVVRMLEGDGLAEKWDEWQKVEILREEIDLS-PNP--NSDWIL-DSTYNLHAVELS---
KIK1    PTVILMLSNDSSSLPNPRPPTLMLRGREIESSKSSEKDRSHSIGTVTMTQLHGRSITRKI
CLV1    REVVHMLTNPP--------------KSVAN-LIAF-------------------------
HAESA   RKVVIMLQEVSGAVPCSSPNTSKRSKTGGK-LSPYYTEDLNSV-----------------
BRI1    VQVMAMFKEIQAGSGIDSQSTIRSIEDGGFSTIEMVDMSIKEVPEGKL------------
WAK1    KEVAAKLEALRVEKTKHKWSDQYPEENEHLIGGHI----LSAQGETSSSIGYDSIKNVAI
CrRLK1  GDVLWNLECALQLQGKQKENEQ-PEEMRDVSATEISLGSMADLAAVSMSKVFSELVKAQG
FLS2    NEILTHLMKLRGKANSFREDRNEDREV--------------------------------
```

Figure III-4: Locations within the CLV1 kinase domain of pThr peptides that bind KI-FHA are shown. Panel a) shows the stereo view of the structure predicted for the representative CLV1 kinase domain. Panel b) shows a cartoon of RLK kinase domains. Spheres correspond to threonines that were phosphorylated during peptide synthesis. The color code is the same as Figure III-1. Sites corresponding to BAK1 pT312 and CLV1 pT868 are yellow. The T546 site of intact BAK1 kinase domain suggested by mutagenesis to interact with KI-FHA *in vitro* is colored orange.

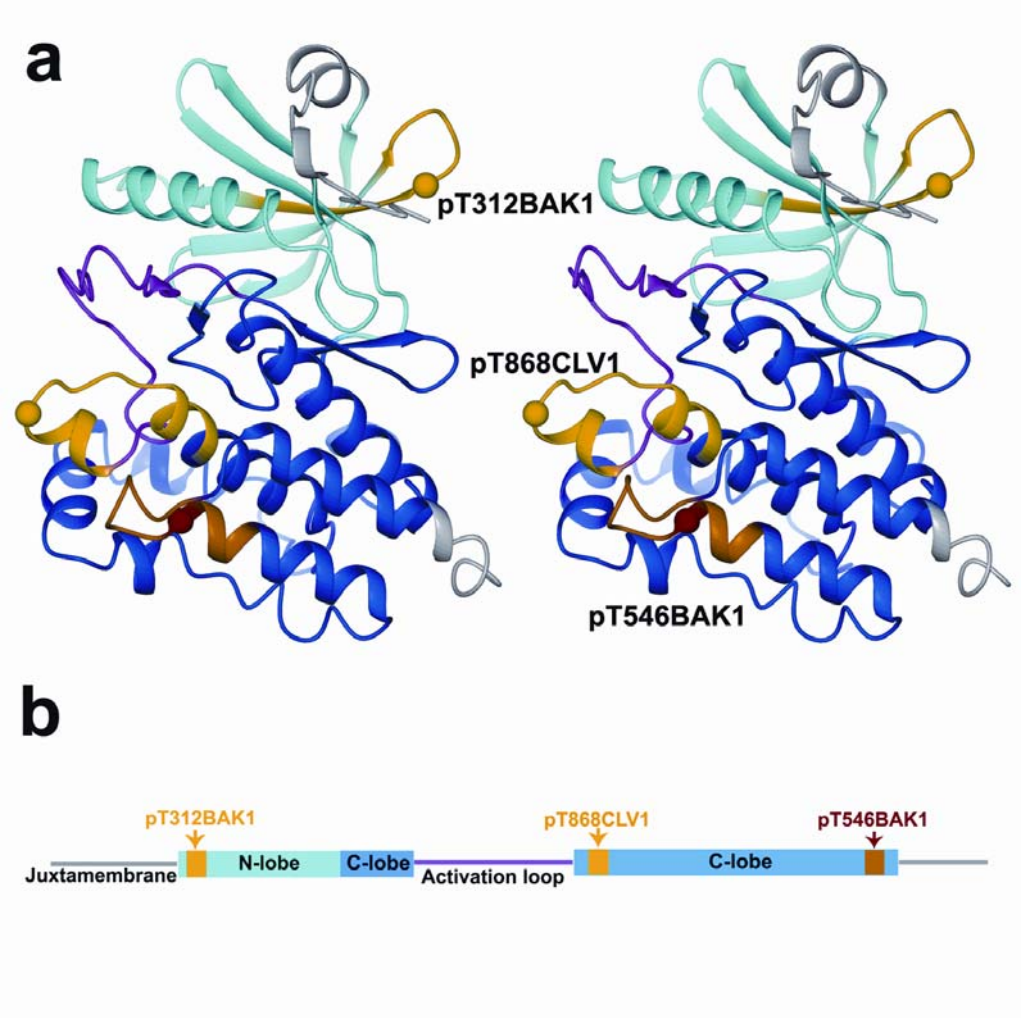


Figure III-5: Calorimetric titrations with KI-FHA of the three pThr peptides derived from RLKs. The upper trace and left axis of each panel describe the raw binding isotherms, while the lower squares and right axis describe the integrated heats. The pThr peptides include: a) pT868 CLV1, b) pT312 BAK1, c) pT546 BAK1, and d) a positive control peptide with sequence of AAYAYpTQASAAKKK optimized from preferences reported from a library (*62*).

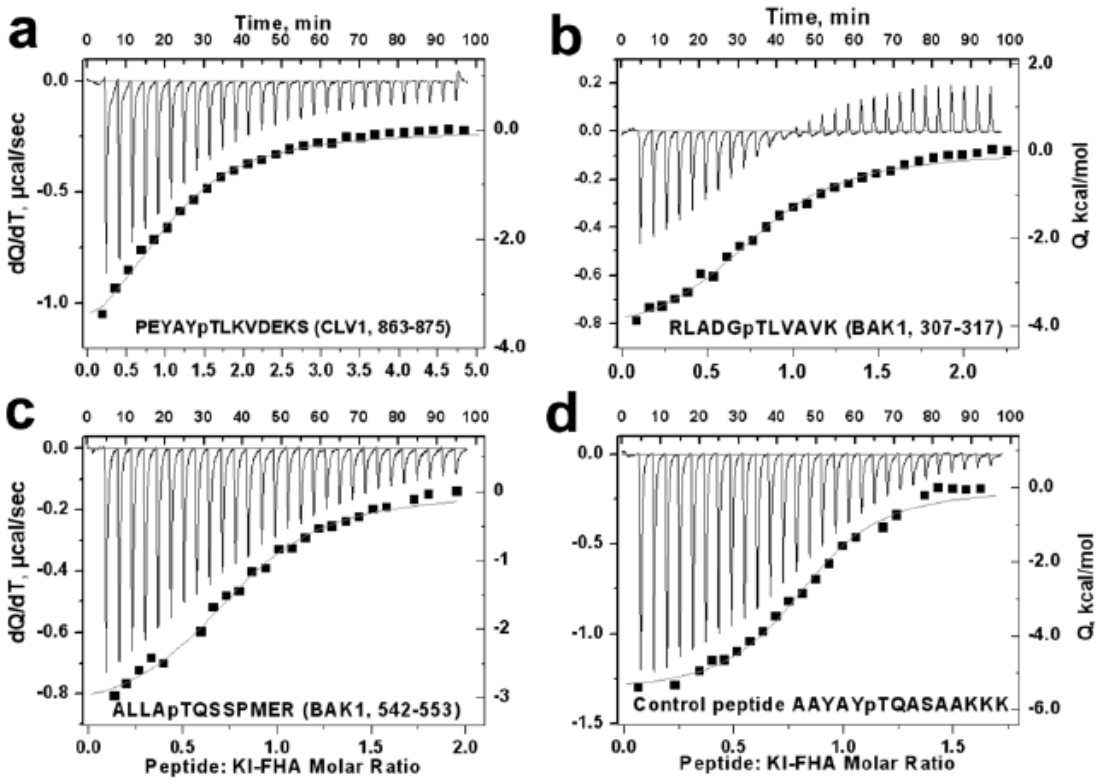


Figure III-6: Determination of the ΔCp of the interaction of the biotinylated pT868CLV1 peptide to KI-FHA in PBS (pH 7.5). KI-FHA, with the concentration of 0.06 mM, was put in the sample cell. The solution of the biotinylated pT868CLV1 peptide (Bio-X-IAPEYAYp**T**LKVDEKS) was put into the spinning syringe. ITC experiments were done at 10, 14, 18 and 25 ^{o}C. ΔH_{obs} was plot against temperature. The plot was fitted by a least-squares analysis. ΔC_p from the slope is -144.5 ± 4.5 cal K^{-1} mol^{-1}.

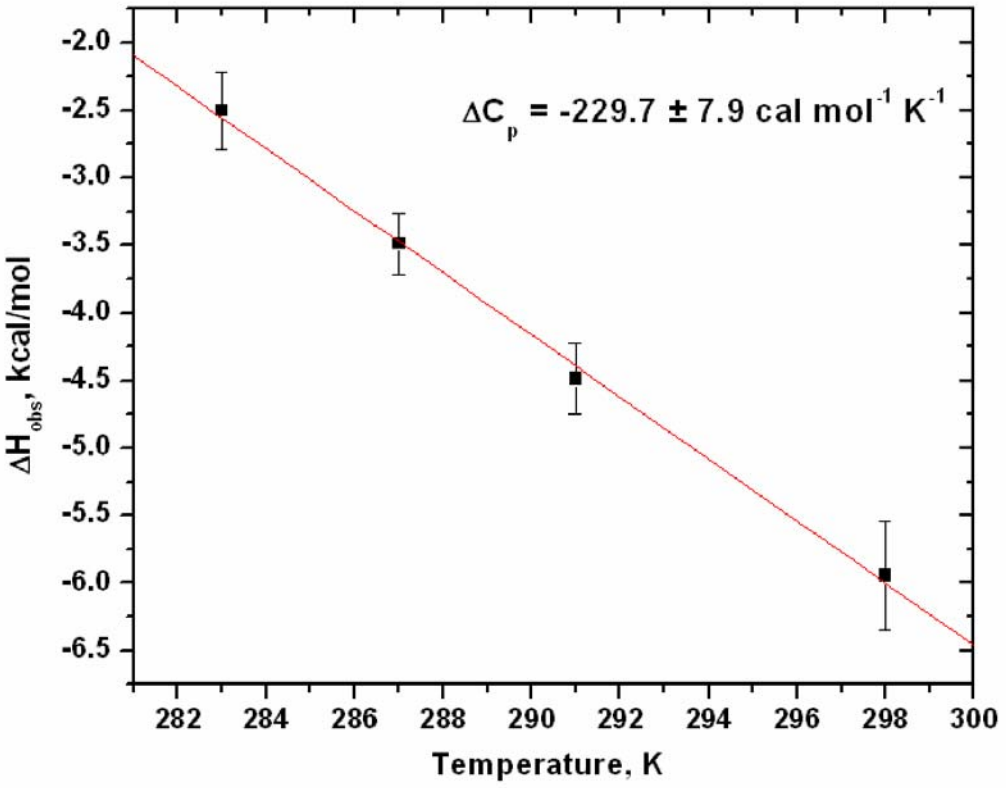


Figure III-7: NMR-detected binding isotherms for the associations of ^{15}N-labeled KI-FHA with pThr peptides of RLKs: a) pT868 CLV1, b) pT312 BAK1, c) pT546 BAK1 and d) the control peptide with sequence of AAYAYpTQASAAKKK. $^{1}H_{N}$ chemical shift changes from ^{15}N TROSY spectra at pH 6.3 are plotted vs. peptide concentrations for representative, affected residues. The fitted lines represent the global, simultaneous fit to the changes of all residues affected by that peptide. The color code for each residue is shared among the panels.

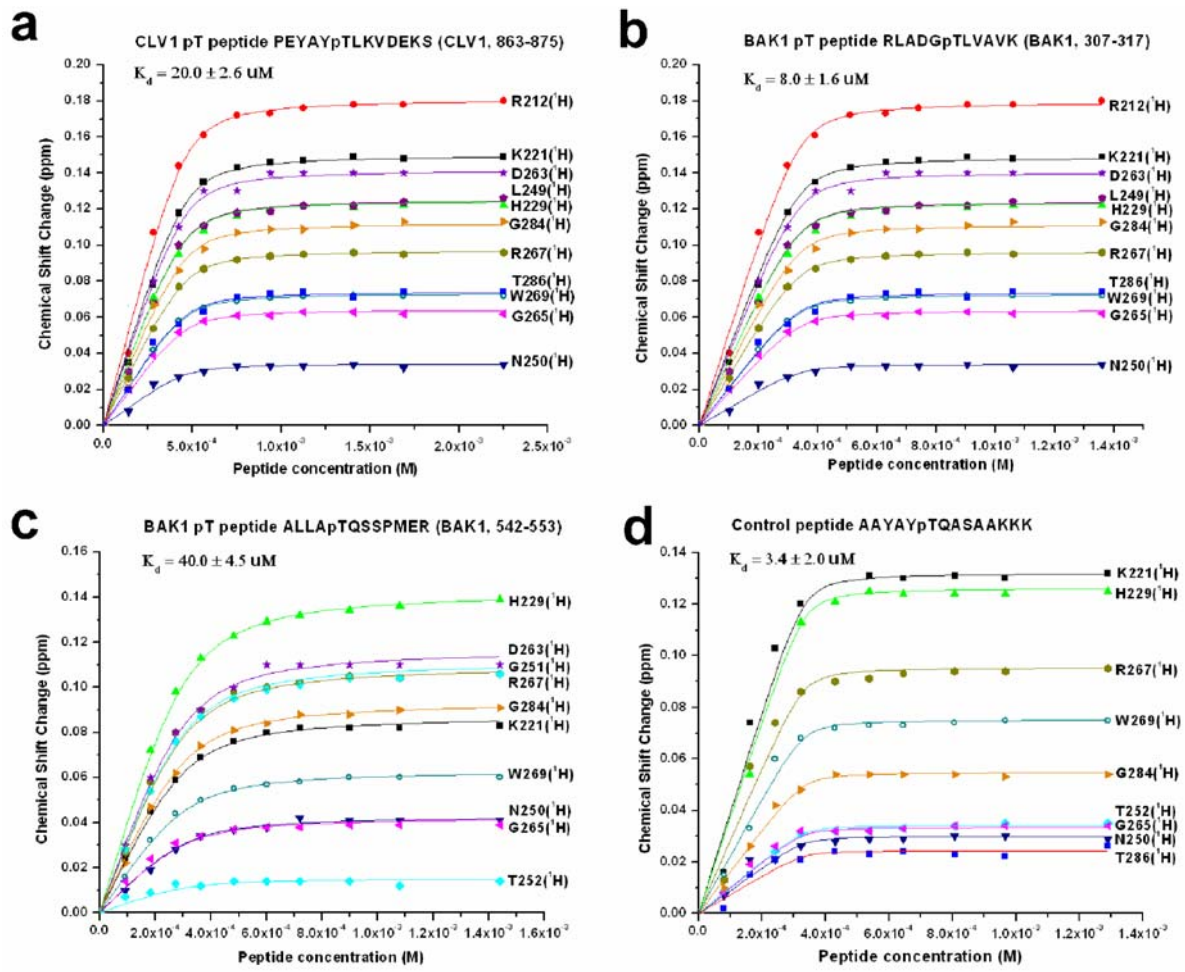


Figure III-8: The phylogenetic tree of 209 FHA domains. Sequences were obtained from the Pfam alignment at www.sanger.ac.uk/cgi-bin/Pfam/getacc?PF00498 and from 12 *Arabidopsis thaliana* sequences from the Munich Information Center for Protein Sequences (MIPS) *Arabidopsis thaliana* database (http://mips.gsf.de/proj/thal), each corresponding to β -strands 3–10 of known FHA structures. Divergent sequences such as the kinesins and truncated sequences that interfere with alignment were deleted prior to alignment by using CLUSTALW. Vertical lines 1–10 represent the partition identity cutoffs that define the evolutionary trace level. Sequence identities from CLUSTALW are listed at the bottom. Characterized FHA domains are labeled in bold. Accession numbers such as Q8W3L2 are from Swiss-Prot. Protein entry codes such as At1g75530 are from the MIPS database (*49*).

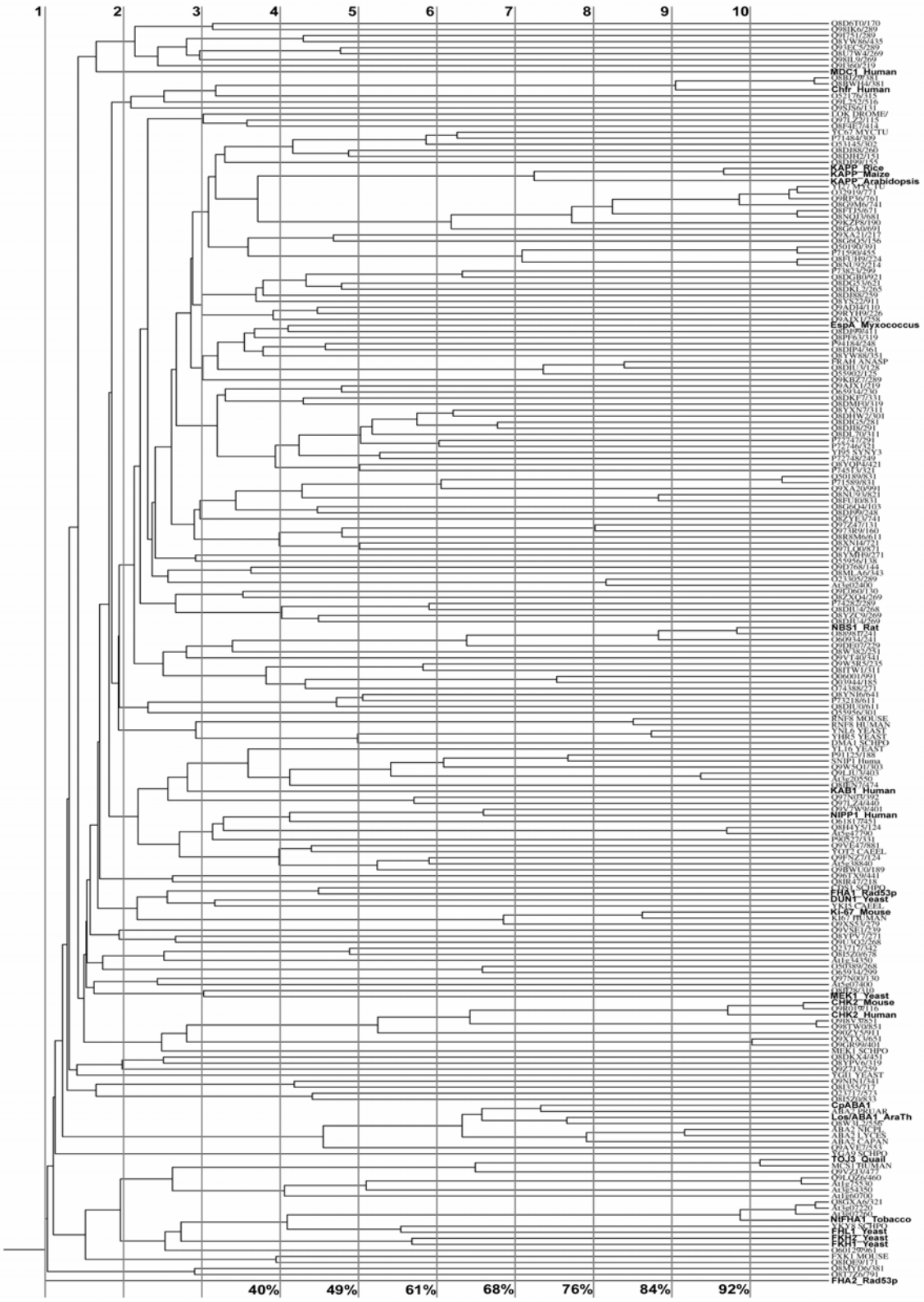

1
2
3
4
5
6
7
8
9
10
MDC1_Human
Chf_Human
KAPP_Rice
KAPP_Maize
KAPP_Arabidopsis
EspA_Myxococcus
NBS1_Rat
KAB1_Human
NIPP1_Human
FHA1_Rad53p
DUN1_Yeast
KI-67_Mouse
MEK1_Yeast
CHK2_Mouse
CHK2_Human
CpABA1
Los/ABA1_AraTh
TQ3_Quail
NtFHA1_Tobacco
FKH1_Yeast
FKH2_Yeast
FHA2_Rad53p
40%
49%
61%
68%
76%
84%
92%

Figure III-9: Class-specific residues of FHA domain sequences at trace level 9. FHA sequences were partitioned into the 17 classes present at trace level 9 of the phylogenetic tree of figure III-8. Note that as the trace level (sequence identity) increases, the sequences are divided into more classes, each containing fewer FHA domain sequences. At each trace level, a "class consensus" residue is conserved throughout its own class (1). Class consensus residues conserved through all classes are conserved (red). "Class-specific" residues (light blue) are cases where class consensus residues differ in at least one class (1). Because of the low homology among FHA domains, we consider a residue to be class-specific, even if there is a single exception of one class lacking conservation at that position, which is shown with a hyphen. Dots represent absence of a residue. The serine marked red appears conserved here, but is only 90% conserved among the 209 sequences of the alignment (*49*).

```
Trace 9                                  Class Consensus Sequence

CLASS1          WTIGRRR...GCDLSFPS.NKL.........VSGDHC---VDEΔ4.VTLEDTS.T-GTVIN..KL-VVΔ3.T-PL--GD-IYLV
CLASS2(KAPP)    ITLGRV...PPSD..LVLKDSE.........VSGKHA-INWN.Δ5.WE-VDMGSLNGTF-NS-...AVΔ12PAELA-GDIITLG
CLASS3          TSAGRH...PDSD..IFLDD.........VTVSRRHAEFR--.Δ3.F-VVDVGSLNGTYVNREPVDSA....VLA.NGDEVQIG
CLASS4          TTAGRH...PESD..IFLDD.........VTVSRRHAEFRIN.Δ3.FEVVDVGSLNGTYVNREPRN-Q....V-Q.TGDEIQIG
CLASS5          NI-GRG...QDAQ..FRLPDTG.........VSRRHL.EIRW.Δ4.ALL-DLNSTNGTTVNNA...PVΔ4....LADGDVIRLG
CLASS6          NIIGRS...NDAD..LRLPDTG.........VSRQH-.EITW.Δ4.AILVDLKSTNGTTVNDT...PVΔ4....LADGDVITVG
CLASS7          VLIGRA...DDST..LVLTDDY.........AS-RHARLS...Δ5.WYVEDLGSTNGTYLD.....RAΔ3TAVRVP-GTPVRIG
CLASS8          Y-VGR....KNC-ILIENDQS..........ISRNHAVL-VNFΔ13L-IKDN.SKYGTF-N......EΔ6LS-TLKTGDRVTFG
CLASS9          YLFGRE...RRIAD.IPTDHPS.........CSKQHAVIQYR.Δ16PYIMDLGSTNKTYIN--P----Δ1.--E----------
CLASS10         HIFGRQ...---CD.FVLDHQS.........VSRQHAAVVPH.Δ4.I-VIDLGS-HGTFVANERLTKDΔ1.PVELEVGQSLRFA
CLASS11         YWFGR...DKSCEYCFD-PLL-RTDKYRT..YSKKHFRIFRE-Δ7.-Y-EDH.SGNGTFVNTEL-GKGΔ1.R-PL-NNSEIALS
CLASS12         YVFGR...DKKCDYTFDIPVLNQTDRYKT..YSKRHFRIFQELΔ7.A-IEDL.SGNGTFVNKEIIGKGΔ1.TLPLTNNAEIALS
CLASS13         FVCGRGS-DAPTNFNFS.-VA-DVGLY-F..ISKIQFSIDRDTΔ4.IYLHDH.SRNGTLVN-EMIGKGΔ1.SRELMNGDLISI-
CLASS14         C--GS-SH-NIPGKS-V-PLPQ.........VSEMHARIS-K.Δ3.FFVTDLRSEHGTW-TDN.EGRRΔ6FPTRFHPSD-IEFG
CLASS15         ITLGR....ATKDNQIDVDL-LEGPA...WKISRKQGVIKL..Δ5.FFIANEG.RRPIYIDGRPVL-G..-KW-L-NNSVVEIA
CLASS16         VLLGR....ATGEYPVDIDLGRSGSE...TRFSRRQALIKL..Δ5.FEIKNLG.KFSIWMNDEEINHG..EVVILKNNCLIQ--
CLASS17         IILGR....NSKK-TVDVDLSSLGGG...M-ISR-HARIFYD.Δ5.F-LEVLG.KNGC-VEGVLHLPG.NP--KLDSQDLLQIG
Summary(9)      X--GX----XXXX--XX-XXX---------XSXXXX-X--X----X--XXXXXXXXXX-X------X------X--XX-XX--
```

Class1: Q8BJZ9
Class1: Q8BWH4
Class1: Chfr_Human

Class2: KAPP_Rice
Class2: KAPP_Maize

Class3: YI27_M.tuberculosis
Class3: O32919
Class3: Q9RP36
Class3: Q8G9M6

Class4: Q8FTJ5
Class4: Q8NQJ3

Class5: Q50190
Class5: P71590

Class6: Q8NU92
Class6: Q8FUH9

Class7: Q50189
Class7: P71589

Class8: Nbs1_Rat
Class8: O88981

Class9: Q9LJU3
Class9: At3g20550

Class10: Q8H4Y5
Class10: At5g47790

Class11: CHK2_Mouse
Class11: Q9R019
Class11: CHK2_Human

Class12: Q9I8V3
Class12: Q98TW0

Class13: Q9XTX3
Class13: Q9GR99

Class14: ABA2_Tobacco
Class14: ABA2_Tomato

Class15: TOJ3_Quail
Class15: MCS1_Human

Class16: Q9LQZ6
Class16: At1g75530

Class17: Q8GXA6
Class17: At3g07220
Class17: At3g07260
Class17: NtFHA1_Tobacco

Figure III-10: Phosphoprotein binding surface of KI-FHA. a) The class-specific residues at trace level 9 of the ET analysis of FHA domains are blue. The residues identical among KAPP of maize, rice, and Arabidopsis are yellow. (b) KI-FHA residues with amide NMR peaks most shifted by saturating amounts of pThr312BAK1 peptide (2 mM peptide: 0.5 mM KI-FHA), i.e. with $\Delta\omega_{HN} > 0.16$ ppm are red. Those residues with 0.16 ppm > $\Delta\omega_{HN} > 0.06$ ppm are pink (*49*).

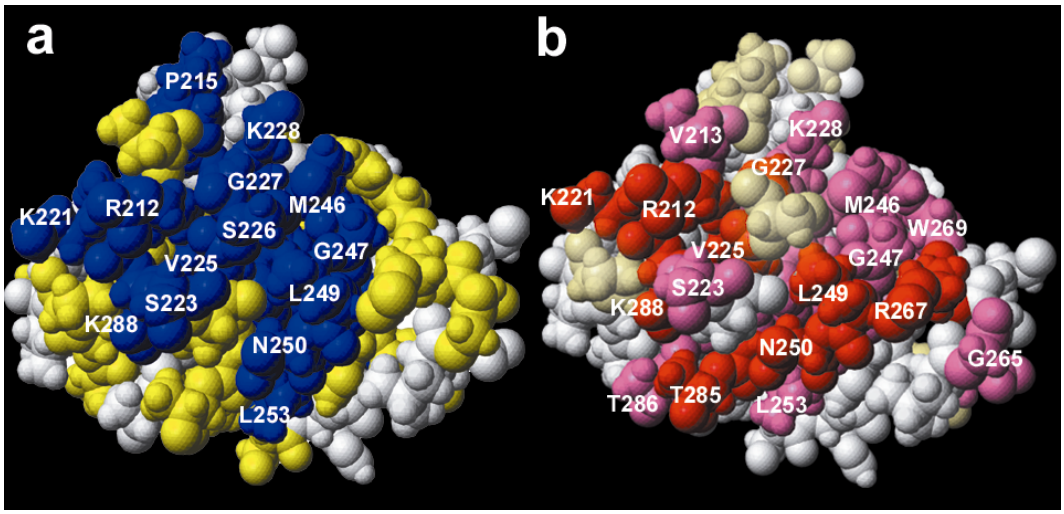


Figure III-11: T546A substitution of BAK1 diminishes binding of KI-FHA. The purified fusion proteins of variants of BAK1 kinase domain were run on duplicate SDS-PAGE gels (upper panel). The lower panel shows a Far-Western blot in which one gel was electroblotted to NC membrane, incubated with KI-FHA, and subsequently probed with antibody directed against KAPP.

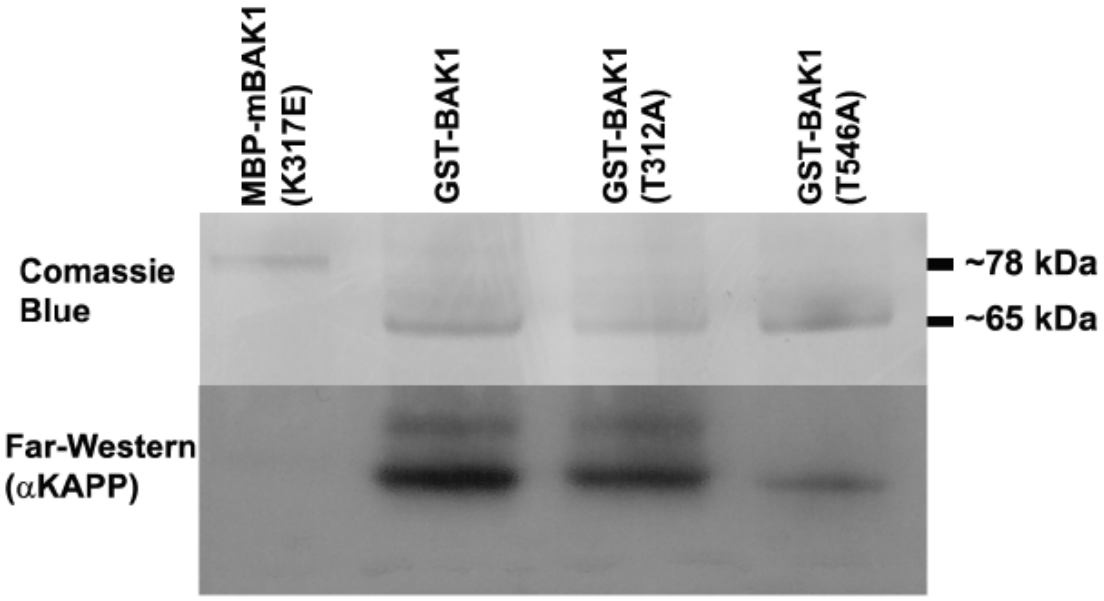
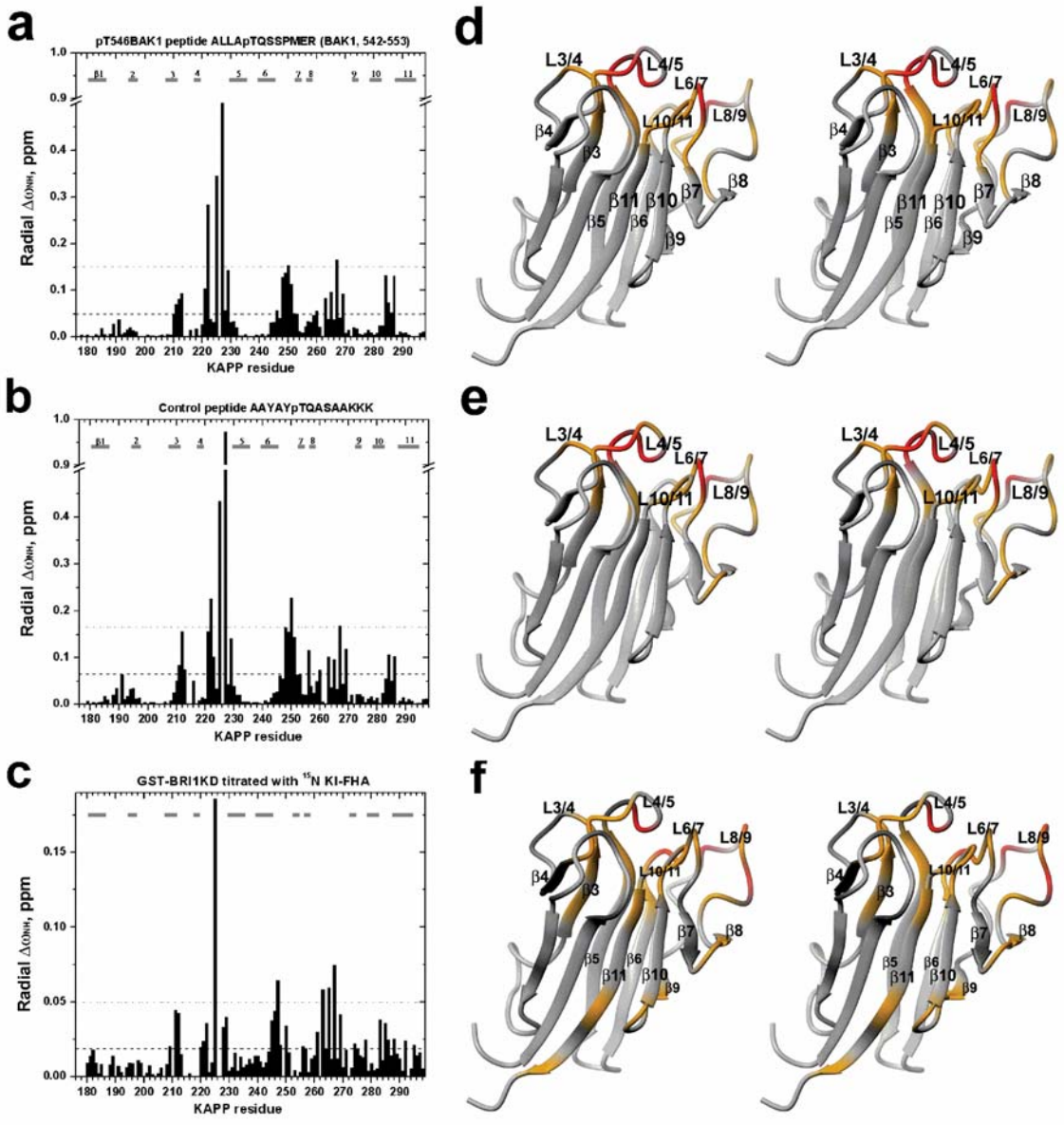


Figure III-12: pThr peptides and BRI1 kinase domain are recognized by the same surface of KI-FHA. Maximal binding-dependent changes observed in amide NMR peak shifts are plotted vs. KAPP residue number for pT546 BAK1 peptide at saturation (a), the control peptide at saturation (b) and 0.75 (nominal) equivalents of GST-BRI1 per equivalent of KI-FHA (Performed by Dr. Xiangyang Liang). (c) The binding-induced changes are plotted as radial shifts $\Delta\omega_{NH}$ according to equation 2. Sites of binding-induced peaks shifts $\Delta\omega_{NH}$ are mapped onto the stereo view of the backbone ribbon of the free KI-FHA structure. KI-FHA residues are colored red where $\Delta\omega_{NH} > 0.15$ and yellow where $0.15 > \Delta\omega_{NH} > 0.05$. (d) Residues are colored red where $\Delta\omega_{NH} > 0.165$ and yellow where $0.165 > \Delta\omega_{NH} > 0.065$. (e) Residues are colored red where $\Delta\omega_{NH} > 0.05$ and yellow where $0.05 > \Delta\omega_{NH} > 0.0188$. (f) Unobserved residues are colored a darker shade of gray.


a
pT546BAK1 peptide ALLApTQSSPMER (BAK1, 542-553)
Radial Δω_NH, ppm
KAPP residue
b
Control peptide AAYAYpTQASAAKKK
Radial Δω_NH, ppm
KAPP residue
c
GST-BRI1KD titrated with 15N KI-FHA
Radial Δω_NH, ppm
KAPP residue
d
e
f
L3/4
L4/5
L6/7
L8/9
L10/11
β3
β4
β5
β6
β7
β8
β9
β10
β11

Chapter IV: Backbone dynamics of peptide bound KI-FHA and comparisons with the free state

NMR relaxation experiments, ^{15}N-^{1}H steady state NOE, and relaxation-compensated CPMG experiments of the free KI-FHA, were performed by Dr. Gui-in Lee, Prof. Steven Van Doren, Dr. Xiangyang Liang, and Dr. Arunima. For purposes of comparison in this chapter, I have used Prof. Van Doren's Modelfree simulations, reduced spectra density function analysis, and residual dipolar coupling data of free KI-FHA. This research has been published in *Biochemistry* in 2005 (*76*).

IV.1 Abstract

A net increase in the backbone rigidity of the kinase-interacting FHA domain (KI-FHA) from the Arabidopsis receptor kinase-associated protein phosphatase (KAPP), free and bound to a pThr868 peptide from its CLV1 receptor-like kinase partner, is suggested by ^{15}N NMR relaxation at 11.7 and 14.1 T. All of the loops of free KI-FHA display evidence of nsec-scale motions. Many of these residues have residual dipolar couplings that deviate from structural predictions. Binding of the CLV1 pT868 peptide seems to reduce

nsec-scale fluctuations of all loops, including half of the residues of recognition loops. Residues important for affinity are found to be rigid, i.e. conserved residues and residues of the subsite for the key pT+3 peptide position. This behavior parallels SH2 and PTB domain recognition of pTyr peptides. PhosphoThr peptide binding increases KI-FHA backbone rigidity (S^2) of three recognition loops, a loop nearby, 7 strands from the β-sandwich, and a distal loop. Compensating the trend of rigidification, binding enhances fast mobility at a few sites in four loops on the periphery of the recognition surface and in two loops on the far side of the β-sandwich. Line broadening evidence of μsec to msec-scale fluctuations occurs across the 6-stranded β-sheet and nearby edges of the β-sandwich. This forms a network connected by packing of interior side chains and H-bonding. A patch of the slowly fluctuating residues coincides with the site of segment-swapped dimerization in crystals of the FHA domain of human Chfr. Phosphopeptide binding introduces μsec to msec-scale fluctuations to more residues of the long 8/9 recognition loop of KI-FHA. The rigidity of this FHA domain appears to couple as a whole to pThr peptide binding.

IV.2 Introduction

FHA domains bind phosphoThr/Ser-containing partners in diverse eukaryotic signaling pathways (*58*). Flexibility often appears to correlate with binding events, including affinity of protein modules for peptides (*66, 126, 127*). Residue-specific effects on the

affinity of SH2 domains for phosphotyrosine peptides are not fully explained by the structures but do correlate with rigidity of interfacial side chains (*122*). Peptide-binding dependent changes in backbone motions of the c-Src SH3 domain help explain its unfavorable entropy of peptide binding and show that the binding stabilizes the SH3 domain both near and far from the interface (*128*). Crystallographic B factors of FHA domains are low, especially in the β-strands (*55*), suggesting rigidity. Phosphopeptide binding does not appear to alter the structure of FHA domains significantly (*48, 51, 53, 129, 130*). This suggests that FHA domains may use a rigid, pre-formed surface for recognizing phosphoprotein partners. Investigation of the dynamic character of an FHA domain should add insights into behaviors affecting binding and stability.

This first investigation of binding-linked mobility changes of an FHA domain compares the backbone dynamics of representative KI-FHA from Arabidopsis KAPP, free and bound to a pThr peptide we identified from its CLV1 receptor-like kinase partner. ^{15}N relaxation data from two magnetic fields are interpreted using both the model-free formalism (*77, 79, 82*) and a reduced spectral density function. Structural fitting of residual dipolar couplings provides an independent assay sensitive to additional timescales. All the loops manifest fluctuations in nsec in the free state, but fewer such cases in the bound state. In the bound state, many residues have increased rigidity or S^2 whether near or far from the phosphopeptide binding site, but also around seven residues have clearly decreased rigidity or S^2. Fluctuations on the μsec to msec scale map to a larger swath and a smaller patch of the 6-stranded β-sheet.

IV.3 Materials and Methods

IV.3.1 Preparation of KI-FHA

GST-tagged KI-FHA, containing Arabidopsis KAPP residues 180-313, was expressed in *E. coli*, enriched with ^{15}N, and purified as described (*101*). After proteolytic removal of the N-terminal GST tag, a linker of Gly-Pro-Leu-Gly-Ser remained at the N-terminus of KI-FHA; the NMR-observable amides of the Leu-G1y-Ser are assigned KAPP residue numbers of 177 - 179. Purified KI-FHA was exchanged into 20 mM sodium phosphate (pH 6.3), 120 mM NaCl and 7% D_2O for NMR. NMR samples of KI-FHA were 0.33 to 0.6 mM, the solubility limit.

IV.3.2 Synthesis, purification, and NMR sample uses of CLV1 pT868 peptide

The pT868 peptide, comprising residues 863-875 of the receptor kinase CLAVATA1 of Arabidopsis, was synthesized with an Advanced ChemTech 396 multiple peptide synthesizer (Louisville, KY) using standard Fmoc chemistry and solid phase synthesis. Each Fmoc amino acid was coupled at least twice, except for the building blocks for the phosphothreonine and the subsequent amino acid that were coupled three times. Cleavage and side chain deprotection were achieved by treating the resin with 87.5% trifluoroacetic acid / 2.5% thioanisole / 2.5% phenol / 2.5% water / 2.5% ethanedithiol / 2.5% triisopropylsilane. All reagents were HPLC or peptide synthesis grade and obtained from Fluka (Milwaukee, WI), Novabiochem (San Diego, CA), or VWR Scientific (St. Louis,

MO). The crude product obtained was characterized by HPLC (Beckmann Coulter, Fullerton, CA) and LC-ESI-MS (Thermo Finnigan, San Jose, CA). The product was purified to greater than 80% by semi-preparative HPLC using an in-house-optimized multi-step gradient. ^{15}N-^{1}H TROSY at 22°C monitored the effects of the CLV1 pT868 peptide on the KI-FHA amide spectrum when titrated up to a 4:1 ratio of peptide:protein. A sample of 0.795 mM pT868 peptide and 0.53 mM KI-FHA, providing ~94% saturation at the K_D of 20 μM, was used to measure a full set of standard relaxation data at each of two magnetic field strengths. A sample of 1.16 mM pT868 peptide and 0.33 mM KI-FHA, providing ~98% saturation at the K_D of 20 μM, was used to measure standard as well as relaxation-compensated R_2 at 600 MHz.

IV.3.3 ^{15}N NMR relaxation measurements

Spectra monitoring ^{15}N R_1 and R_2 relaxation rates and ^{15}N{^{1}H} steady-state NOE were acquired at 25°C at both 600 and 500 MHz. Relaxation dispersion experiments were collected at 22°C at 600 MHz. A Varian Inova 600 MHz spectrometer (Univ. of Missouri) was used with either a high-sensitivity 5 mm ^{1}H {^{13}C/^{15}N} cryogenic probe with shielded z gradient coil or a 5 mm HCN triple resonance probe fitted with an xyz gradient coil. The Bruker DRX-500 spectrometer at the Univ. of Missouri was used with an 8 mm triple resonance probe with an actively shielded z gradient coil (Nalorac, Martinez, CA). The Bruker DMX-500 at NMRFAM was used for ^{15}N{^{1}H} steady-state NOE and R_1 at 500 MHz as it is fitted with a very sensitive 5 mm Cryoprobe™ with shielded z gradient coil. The temperature settings of Missouri's 500 and 600 MHz

instruments were calibrated using ethylene glycol from room temperature upward and using methanol below room temperature. The actual temperatures are within 1°C of the nominal temperatures and within 1°C between the two instruments. The 600 MHz R_1 and R_2 pulse sequences used sensitivity-enhanced gradient-based coherence selection (*131*), implemented in Biopack. Cross-correlation was suppressed from R_1 and R_2 spectra (*132*). For good water suppression with cryogenic probes and economy of pulses for the 8 mm probe, home-implemented 3-9-19 WATERGATE (*133*) versions of sequences of ref (*134*) were used for $^{15}N\{^{1}H\}$ steady-state NOE, 500 MHz R_1 and R_2. The relaxation-compensated CPMG pulse sequence of ref (*135*) was used on the Inova 600 with τ_{cp} values of 2 msec and 7.5 msec to identify sites of exchange broadening on the msec scale.

Using the relaxation-compensated CPMG sequence with τ_{cp} = 1 ms (π pulse spacing of 2 ms), total relaxation delays were 16, 32, 48, 64, 80, 96 ms, whereas at τ_{cp} = 3.75 ms (π pulse spacing of 7.5 ms) total delays were 30, 60, 90, and 120 ms. Relaxation delay periods at 500 MHz were 20*, 60, 120*, 240, 400*, 600 and 860* ms for R_1 and 16*, 32, 48*, 64, 80*, 112, 144* ms for R_2. Asterisks (*) indicate points duplicated for estimating uncertainties in peak heights. 500 MHz R_1 data of free KI-FHA were found to reproduce well between Missouri's DRX-500 with conventional probe and NMRFAM's DMX-500 with Cryoprobe™. Delay times at 600 MHz were 10, 20*, 50, 100*, 200, 350*, 500, and 700* ms for R_1 and 10*, 30, 50*, 70, 90*, 110, and 130* ms for R_2. For KI-FHA in the free state, the 600 MHz R_1 and R_2 results from the conventional probe were repeated with the cryogenic probe and averaged. NMR spectra were processed using Sybyl 6.8 or NMRPipe 2.3 (*107*). Uncertainties in R_1 and R_2 relaxation rates were generated by

relaxation fitting simulations with SPARKY 3 (*136*) using 40 iterations with addition of Gaussian-distributed random noise. To obtain ^{15}N $\{^1H\}$ NOEs, spectra were measured with and without 3.5 seconds of proton saturation, in an interleaved manner to maintain identical sample conditions. The non-saturated spectra reference spectra had relaxation delays of 9.2 seconds at 500 MHz and 12 seconds at 600 MHz to allow water magnetization to recover completely and avoid saturation transfer. Identical pairs of NOE spectra were collected in triplicate, except for the free state at 500 MHz collected in duplicate, to evaluate standard deviations for each residue. Standard rules of statistical error propagation were applied to arithmetic combinations of relaxation rates.

IV.3.4 Reduced spectral density analysis of ^{15}N relaxation

The spectral density function, $J(\omega)$, for an N-H bond vector describes the frequency spectrum of its reorientation. $J(\omega)$ decreases monotonically with frequency ω and has nearly zero slope at high frequency, i.e. $J(\omega_H+\omega_N)=J(\omega_H)= J(\omega_H-\omega_N)$. (Recall that ω_N is negative.) Consequently, three spectral density functions for each amide are sufficient and well-approximated from measurements of just ^{15}N R_1 and R_2 and $\{^1H\}$-^{15}N NOE enhancement. The three spectral density expressions resulting are (*85, 137, 138*):

$$J_{eff}(0) = [6/(3d^2 + 4c^2)](-R_1/2 + R_2 - 3\sigma/5) \quad (1)$$

$$J(\omega_N) = [4/(3d^2 + 4c^2)](R_1 - 7\sigma/5) \quad (2)$$

$$J(0.87\omega_H) = 4\sigma/(5d^2) \quad (3)$$

$$\sigma = (NOE-1)(\gamma_N/\gamma_H)R_1 \quad (4)$$

where $d = \mu_0 h\gamma_N\gamma_H\langle r_{NH}^{-3}\rangle /(8\pi^2)$; $c = \omega_N(\sigma_\parallel - \sigma_\perp)/\sqrt{3}$; $\sigma_\parallel - \sigma_\perp$ is the chemical shift anisotropy estimated to be –170 ppm; γ_H and γ_N are the gyromagnetic ratios of ^{1}H and ^{15}N; and ω_H and ω_N are the Larmor frequencies of ^{1}H and ^{15}N. J(ω) at ω=0, ω_N, and $0.87\omega_H$ were estimated from both 500 and 600 MHz relaxation data. The reduced spectral mapping approach needs no models of internal motion or of anisotropy in Brownian tumbling of the protein.

IV.3.5 Model-free analysis of ^{15}N relaxation

Model-free calculations were performed using the ModelFree 4.1 suite of programs provided by Prof. Arthur G. Palmer (*139*). The program Pdbinertia was used to estimate the ratio of the principal moments of inertia of representative model 1 from the high accuracy solution structure of KI-FHA (PDB accession code 1MZK)(*49*). Initial estimates of τ_m and anisotropy of the rotational diffusion tensor were obtained from R_2/R_1 ratios using the programs R2R1_diffusion or TENSOR2 (*80*), after excluding those for residues undergoing motion on the ps to ns timescale or chemical exchange in μs to ms (*78, 139*). The ^{15}N CSA was set to the approximation of -170 ppm (*140*). To anticipate systematic effects upon S^2 of the unknown variation in ^{15}N CSA (*141*) of KI-FHA, parallel model-free calculations were fixed to a range of other CSA values. At –150 ppm, S2 averages 0.015 ± 0.019 above S^2 at –170 ppm. At –190 ppm, S^2 averages 0.029 ± 0.026 below S^2 at –170 ppm. The uncertainties for R_1 and R_2 were set to 5% of their values as suggested (*142*). After selection of an expression for all residues, a final optimization was performed in which the overall rotational diffusion model and the

internal motional parameters for each NH vector were optimized simultaneously. Uncertainties in the dynamics parameters were obtained using 300 steps Monte Carlo simulations carried out using Modelfree 4.1. Under the assumption of a wobbling-in-a-cone motional model, the amplitude of the maximal cone semi-angle of N-H bond excursions can be estimated from order parameter S using the relationship (*77*):

$$S = \frac{\cos(\theta) + \cos^2(\theta)}{2} \tag{5}$$

IV.3.6 Selection of spectral density expressions

Relaxation rates were fitted to one of five forms of spectral density functions (Table IV-1), where τ_m and $D_{\parallel}/D_{\perp}$ were fixed to initial estimates from the R_2/R_1 ratios. To accommodate data from two magnetic fields leading to higher model-free χ^2 errors and to avoid under-fitting and bias (*81*), long-established Akaike`s Information Criterion (AIC) (*143*) was used as the principal method of statistical model selection as described (*81*). Akaike`s information criterion is computed as χ^2 + 2k, where k is the number of parameters evaluated from among S_s^2, τ_e, R_{ex} and S_f^2 (*81, 143*). AIC was used to test need to let τ_e range up to 2 nsec in expressions 2 and 4 of Table IV-1. By monitoring R_1^{500MHz} / R_1^{600MHz} after a correction for psec internal motion, the PINATA algorithm suggested nsec-scale motions in additional residues with ratio < 0.96 (*144*). The τ_e term of these residues was allowed to range up to 1 nsec. The suitability of use of 1 to 2 nsec τ_e was confirmed by elevated $J(\omega_N$=61 MHz) (eq. 2) since the baseline at each spectral density frequency marks the rigid structural core (*127*). To justify use of the exchange broadening term R_{ex} (expressions 3 and 4 in Table IV-1), F-testing of model-free χ^2,

using data from the two fields, was required to show a P value < 0.25 in addition to satisfying AIC. Alternatively, relaxation dispersion data with increased line broadening at slower pulsing, i.e. R_2 (7.5 ms) – R_2 (2 ms) beyond uncertainty above the baseline, was considered clear evidence of exchange on the msec scale and sufficient justification for use of R_{ex}.

IV.4 Results

IV.4.1 ^{15}N relaxation in free and phosphoThr peptide bound states

^{15}N NMR relaxation parameters of KI-FHA, free and bound to an excess of a 13-mer pThr peptide from CLV1, were determined at 298 K, at both 500 MHz and 600 MHz to enhance insight into the amplitude and timescale of backbone dynamics. The peptide, dubbed CLV1 pT868, comprises residues 863-875 of the Arabidopsis CLV1: PEYAY(pT)LKVDEKS. Multiple sequence alignment of kinase domains of RLKs indicates that the threonine we phosphorylated in this CLV1 pT868 peptide corresponds to Ser1060 of the BRI1 RLK, one of two neighboring residues phosphorylated at the C-terminal end of the activation loop of kinases (*59*). This suggests pT868 of CLV1 might also be phosphorylated *in vivo*. At 298K, CLV1 pT868 binds KI-FHA with apparent K_D of 21 μM by NMR at pH 6.3 and 16 μM by titration calorimetry at pH 7.5; details have been described in Chapter III. Of 114 observable backbone amide peaks, 99 in the free state and 102 in the bound state were resolved well enough for quantitative analysis of their relaxation and dynamics. Single exponentials fit well the R_1 and R_2 relaxation rates.

The exception is in 500 MHz R_1 spectra where four to five residues have anomalously fast relaxation (marked with gray squares in Figures IV-1b, 2b) that approaches or exceeds the theoretical limit of 3.14 s^{-1} when fitted with a single exponential; their biphasic relaxation is possibly an artifact of the exchange broadening that these particular backbone peaks undergo. In the free state of KI-FHA, the peaks of the following were too overlapped for reliable fitting: I187 / L210 / S223, D217 / I293, N235 / L283, V198 / S200 / E224 / S248, and M246 / L264. In the CLV1 pT868-bound state of KI-FHA, the spectral overlap remains in the first two peak clusters listed. With the 0.33 to 0.6 mM KI-FHA concentrations employed, particularly high S/N was assured by detection with a cryogenic probe at 600 MHz and a cryogenic probe for some of the 500 MHz data.

High backbone ^{15}N $\{^1H\}$ NOEs (Figures IV-1a, 2a and Table IV-2) indicate that KI-FHA is generally rigid in free and CLV1 pT868-bound states. Negative NOE values and low R_2 values (Figure IV-1a, c) at the termini indicate their very high mobility and that the folded domain extends from KAPP residue Ser180 through Ser295. ^{15}N $\{^1H\}$ NOEs (especially at 600 MHz) lower than 0.65 suggest such residues to be enriched in fluctuations on the psec to nsec scale. In free KI-FHA, such residues with higher amplitude fast motions are found in the 2/3 loop, 3/4 loop, 4/5 loop, and the long 8/9 loop (Figure IV-1a). Elevated R_1 values were shown to correlate with nsec internal fluctuations (*145*). In the free state, there are residues in all the loops and termini that have higher-than-average R_1 values, suggesting they may undergo nsec fluctuations. In the pT868-bound state of KI-FHA, all the loops show less of this trend of elevation of R_1 (Figure IV-2b).

Sites of high R_2 can suggest fluctuations on the μsec to msec timescale. In the free state, residues with at 600 MHz R_2 exceeding 13.8 s^{-1}, i.e. > 1.5 σ above the coarse-filtered $\overline{R_2} = 12.6 s^{-1}$ (Table IV-2), came under consideration for conformational exchange broadening processes on the μsec to msec scale. These residues lie in the 1/2 β–hairpin, the 3/4 loop, β4, the 4/5 loop, the 7/8 loop, β8, and 10/11 β–hairpin (Figure IV-1c). (The term *β-hairpin* refers to both anti-parallel strands and the intervening loop). Relaxation-compensated CPMG R_2 measurements comparing averaged R_2 between faster and slower pulsing of CPMG π pulses can exaggerate the line broadening of msec-scale chemical exchange processes (*135*). Elevated $\Delta R_2 = R_2$ (7.5 ms) – R_2 (2 ms) in Figure IV-1d clearly suggest msec exchange broadening to occur in the 1/2 β̃ hairpin, the 4/5 loop, β 8, the 9/10 loop, 10/11 β̃ hairpin, and C-terminus. Most of these residues evidently with msec-scale motions are found in a swath across the 6-stranded β-sheet, i.e. its β1, β2, β10, and β11 strands. Similar inspection of elevated 600 MHz R2 values with CLV1 pT868-bound present at 3.5-fold excess suggest the same segments of bound KI-FHA to undergo exchange broadening, plus the 6/7 recognition loop (Figure IV-2c). For the bound state, elevated $\Delta R_2 = R_2$ (7.5 ms) – R_2 (2 ms) in Figure IV-2d suggests similarity to the free state in patterns of msec exchange broadening in the 1/2 β̃ hairpin, 10/11 β̃ hairpin, and C-terminus. In the bound state, however, msec-scale exchange broadening in the central 4/5 recognition loop (Figure IV-2d, 1d) appears to be largely quenched.

IV.4.2 $J(\omega_N)$ reduced spectral density evidence of nsec motions in KI-FHA

Reduced spectral density fitting of the relaxation data (*85, 137*) give temporal insight into the backbone mobility of free and bound KI-FHA. Elevated $J(\omega_N)$ values, where ω_N=60.8 MHz, in free KI-FHA (Figure IV-3a) suggest that all of its loops and its termini have sites enriched in nsec-scale fluctuations. $J(\omega_N)$ values of pT868-bound KI-FHA (Figure IV-3a) are lower overall with fewer elevated excursions. Much of the general trend to lower $J(\omega_N)$ values in the bound state may be a consequence of its slower tumbling and higher basal $J_{eff}(0)$. In the pThr peptide bound state, however, locally more attenuated $J(\omega_N)$ values (Figure IV-3a) suggest that all of the loops have residues with nsec fluctuations apparently quenched. The following β-strands (Figure IV-3a) have residues with apparently fewer nsec fluctuations: especially β1, but also one to two residues in each of β-strands 6 through 11.

IV.4.3 Fluctuations in the loops of free KI-FHA are corroborated by deviations of ^{15}N-^{1}H residual dipolar couplings

Residual dipolar couplings (RDCs) can respond to motional averaging over a wide, continuous range of timescales from psec to tens of msec (*146, 147*), an advantage over spin relaxation studies. Deviations of measured RDCs from RDCs predicted from the alignment of the protein structure can result from static or dynamic discrepancy from the protein structural coordinates (*146, 147*). All $^{1}D_{NH}$ RDCs were fitted to the whole of each of the ten lowest energy structures from the KI-FHA ensemble (PDB code 1mzk) using PALES (*148*). Deviations of individual $^{1}D_{NH}$ RDCs from those predicted by fits using

singular value decomposition (SVD) are represented in Figure IV-3b as a χ^2-like quantity that is normalized by the Da description of the range of observed RDCs. Several loops have $^1D_{NH}$ RDCs that deviate from RDCs predicted by the structure, where L is for loop: L1/2, L2/3, L4/5, L5/6, L6/7, L7/8, L8/9, and L10/11 loops. β8 and the C-terminus also have residues with deviating $^1D_{NH}$ RDCs (Figure IV-3b). A majority of the sites of $^1D_{NH}$ deviation have higher-than-average RMSD in the ensemble of structures. Most cases of $(^1D_{NH,\ meas} - {}^1D_{NH,\ calc})/D_a$ deviations in the free state (Figure IV-3b) have elevated $J(\omega_N)$ evidence of nsec-scale motions (Figure IV-3ba). The $(^1D_{NH,\ meas} - {}^1D_{NH,\ calc})/D_a$ parameter appears to respond independently to many but not all possible sites of apparent nsec scale motion.

IV.4.4 Hydrodynamics of KI-FHA, free and bound to CLV1 pT868 peptide

Rotational correlation time τ_m (overall tumbling) and rotational diffusion anisotropy can be estimated from R_2/R_1 ratios (and atomic coordinates) using rigid residues where the R_2/R_1 ratio depends only on τ_m in the limit of small amplitude fast internal motions ($\tau_e \ll \tau_m$) (*78, 149*). An attractive two-stage procedure for identifying such residues without discarding R_2/R_1 ratios made high or low by anisotropy of diffusion is that of ref (*150*). The results of the coarse filtering stage using those authors' NormaDyn software with the 500 and 600 MHz data collected for free and bound states are presented in Table IV-2; typically about 80 candidate residues having acceptably high $^{15}N\ \{^1H\}$ NOE and acceptably low R_2 remained after coarse filtering. The purpose of the subsequent fine filtering stage is to remove additional residues poorly fit by simple model-free

expressions 1 and 2 (Table IV-1) due to motions on the nsec or μsec to msec scales (*150*). We enhanced the fine filtering stage by removing residues (1) with msec scale exchange broadening according to relaxation dispersion (Figures IV-1d, 2d), (2) likely to have nsec motion due to $J(\varpi_N) \geq \overline{J(\varpi_N)} + 1\sigma$ where ω_N= 60.8 MHz (Figure IV-3a), or (3) requiring the exchange broadening term R_{ex} in preliminary model-free simulations according to statistical model selection. Around 40 residues (Table IV-2) survived this enhanced fine filtering.

The fine-filtered data from 600 MHz with cryogenic probe were fitted to the NMR structure of free KI-FHA (PDB accession code 1MZK) to assess correlation time and anisotropy of rotational diffusion using the TENSOR2 program of ref (*80*). The ratio of the principal moments of inertia of free KI-FHA is 1.00:0.91:0.55. The free state is estimated to have τ_m of 7.57 ± 0.08 nsec with $D_{\parallel}/D_{\perp}$ =0.81 ± 0.07 (oblate) and the pT868-bound state to have τ_m of 10.70 ± 0.07 with $D_{\parallel}/D_{\perp}$ =1.21 ± 0.06 (prolate). (Note that a prolate to oblate shift was reported for peptide binding to an SH3 domain (*128*).) For the free and bound states, both oblate and prolate axially symmetric fits are acceptable with high statistical confidence and have S^2 similar to within 10%. Such two-minimum behavior suggests the possibility of fully anisotropic diffusion (*151*). Indeed, a fully anisotropic diffusion model is statistically significantly better for the bound state. The limitation to axial symmetry of Modelfree 4.1 used here and widely could add a very modest systematic bias, but little difference in S^2 values results from switching between oblate and prolate models for KI-FHA (not shown).

These τ_m estimates for KI-FHA at 25ºC are comparable to the τ_m values of monomers of similar size and temperature (*149, 152, 153*). The unlabeled free state of KI-FHA is 124 residues and 13.3 kDa by ESI-MS while the complex is 12% larger at 137 residues and 14.9 kDa. Saturation with the 1.6 kDa, 13-mer pT868 peptide increases R_2/R_1 more than expected (Table IV-2). The resulting increase in apparent Stokes radius with the pT868 peptide bound is around 9%. Why does the increase in τ_m upon saturation with the 1.6 kDa pT868 peptide appear to be 1 to 1.5 nsec more than expected? To investigate whether peptide binding promotes self-association, a complex of 0.53 mM KI-FHA with 1.5-fold excess of pT868 peptide was diluted 2.3-fold, resulting in perhaps a 4% average drop in R_2 values (Table IV-2). This decrease parallels the decrease in pT868 peptide saturation from 94% at 0.53 mM KI-FHA to 88% at 0.23 mM KI-FHA, i.e. the enhanced weighting of the free state contribution to apparent τ_m when diluted. If self-association were substantial, a larger drop in R_2 and estimated τ_m ought to have accompanied 2.3-fold dilution. A more likely alternative explanation could be that the greater extent of nsec fluctuations of the loops in the free state may decrease its effective Stokes radius relative to the bound state. (Note that a case of greater slowing of rotational diffusion upon peptide-binding to a monomer has been reported, where coupling between altered internal motion and diffusional shape seems larger (*154*).)

IV.4.5 Spectral density function selection for model-free simulations

Model-free simulations of the relaxation data proceeded with hydrodynamics parameters τ_m and $D_{\parallel}/D_{\perp}$ fixed at the values listed in Table IV-2. Use of relaxation data from two magnetic fields in the model-free approach more reliably defines the S^2, R_{ex} (excess line broadening), and τ_e (internal correlation time) parameters. Data at two magnetic fields also provides enough measurements to use three-parameter fit where needed. *Five* measurements were used for both the free and bound states of KI-FHA, namely 600 MHz R_1, R_2, and NOE and 500 MHz R_1 and NOE. Omission of 500 MHz R_1 gave very similar results. The spectral density functions (Table IV-1) most appropriate for use in model-free simulations were chosen primarily using AIC-based statistics (*81*) augmented by other experimental identification of msec motion from relaxation dispersion parameter $\Delta R_2(\tau_{cp})$ (ref *135*) (Figures IV-1d, 2d) and of nsec motion from the reduced spectral density function $J(\omega_N$=60.8 MHz) (ref *87*) (Figure IV-3a). The spectral functions selected, among the choices of Table IV-1, are listed in Table IV-3, by residue number for free and bound states of KI-FHA.

IV.4.6 General model-free outcomes for free and bound states

Excluding the termini, S^2 values appear to average 0.902 and 0.946 for the free and pT868-bound states of KI-FHA, respectively. For the non-terminal residues compared in both states, $\Delta S^2_{bound\text{-}free}$ averages 0.042 with a median $\Delta S^2_{bound\text{-}free}$ of 0.036 and average uncertainty of 0.037. These suggest that phosphoThr binding results in enhanced overall

rigidity on psec to nsec scale. Model-free τ_e values (Figure IV-4c) are joined by elevation of both $J(\omega_N)$ (ref *87*) (Figure IV-3a) and R_1 (ref *145*) (Figure IV-1b) in suggesting nsec motions in all of the loops and termini of free KI-FHA except for the 1/2 loop. Fewer of these residues in the CLV1 pT868-bound state of KI-FHA show such evidence of nsec motions (Figures IV-2b, 3a, 5c, 7), suggesting the peptide binding quenches nsec motions globally. Where the loops are enriched with nsec scale fluctuations, especially in the free state, the S^2 values can be regarded as overestimates, as established by methods that account for coupling of internal nsec motion with rotational diffusion (*155, 156*). The relaxation data and model-free results for KI-FHA, free and bound to CLV1 pT868, were deposited under BMRB accession codes 5841 and 6474, respectively, at www.bmrb.wisc.edu.

IV.4.7 pThr peptide binding surface and its slower dynamics

Titration of ^{15}N-enriched KI-FHA with the CLV pT868 peptide shifted the amide peak positions of loops on one edge of the β-sandwich of KI-FHA. Significant amide chemical shift changes occur at Leu210 - Val213 of the 3/4 loop; Lys221 - Ser223, Val225, and Gly227 - His229 of the 4/5 loop; Met246 – Asn250 and Thr252 of the 6/7 loop; Asp263, Gly265, Arg267, and Trp269 of the 8/9 loop; and Gly284 - Thr287 of the 10/11 loop (Figure IV-6c, d). nsec scale motions characterize the free state of several residues quite perturbed by the pT868 peptide: Arg212, Val213, Ser223, Val225, Lys228, Leu249, Asn250, Arg267, Thr285, and Thr286 (Figures IV-3a, 4c). Much slower conformational fluctuations over μsec to msec are shared by the free and bound states of these pT868-

perturbed residues: Val213, Ser216, Lys228, Arg267, and Thr286 (Figures IV-4b, 5b). Exchange broadening > 1 s^{-1} suggests that fluctuations in the μsec to msec regime also occur at Ser223 and Glu224 in the free state (Figure IV-4b) and at His229 and Gly251 in the bound state (Figure IV-5b)

IV.4.8 Binding effects on flexibility on fast timescale in recognition interface

Largely conserved recognition loop residues shown to be essential for binding phosphorylated receptor kinase domain (*46*) are clearly rigid. Their apparent S^2 values in free / bound states are: 0.99 / 0.999 for Gly211, 0.92 / 0.92 for His229, and 0.96 / 0.99 for Asn250 (Figures IV-4a, 5a). Binding appears to quench nsec fluctuations of Asn250 (Figures IV-3a, 4c, and 5c). Near the five most conserved residues in the recognition loops are non-conserved residues with greater mobility and significant changes of S^2 upon binding of the pT868 peptide from CLV1. Fluctuations in the free state of greater-than-average amplitude are suggested by apparent $S^2 \leq 0.8$ of residues of the 3/4, 4/5, and 8/9 recognition loops (Figure IV-4a): Ser216, Asp222, Glu224, Lys228, Ser260, and Ser266. Lys228 and Glu224 of the 4/5 loop, Ser266 of the 8/9 loop, as well as Ser216 and Val213 of the 3/4 loop (Figure IV-6) undergo striking increases in rigidity on the psec-nsec scale upon pT868 peptide binding. Leu220, Asp263, and Leu283 of these recognition loops are rigidified to a smaller degree. Also at the interface, Gly251 of the 6/7 loop, Arg267 of the 8/9 loop, and T286 of the 10/11 loop (Figure IV-6) in contrast become more flexible on the fast timescale in the complex.

IV.4.9 Long-range effects of binding on fast timescale motions

For a number of residues outside this binding surface for phosphorylated RLKs (Figure IV-6d), mobility of KI-FHA on the psec to nsec scale changes upon binding the pT868 peptide from the CLV1 RLK. In the loops of the 1/2 and 10/11 β hairpins pointing toward the interface, residues where binding increases the amplitude of psec to nsec motions (drop in S^2) are Ala191 and Thr286 (Figure IV-6a,b). In the loop of the 7/8 β hairpin distant from the interface, binding may have increased the fast motion of Ser256 (Figure IV-6). By contrast, pT868 peptide binding appears to have increased rigidity in β-strands of all four of the β-hairpins. This is especially evident in both strands of the 1/2 β hairpin, both strands of the 5/6 β hairpin, β8 of the 7/8 β hairpin, and more subtly in both strands of 10/11 β hairpin (Figure IV-6a, b). The apparent changes in S^2 in the highly disordered termini (Figure IV-6a) are unlikely to be significant since the termini lack the globular diffusion behavior assumed by the model-free approach. The 2/3 loop and 9/10 loop each cross between the two β-sheet layers of the sandwich and are most distant from the binding surface. The peptide binding effect on the 2/3 loop (Figure IV-6a) includes mobilization of Ser202 and increased rigidity of Ser203 and Leu206. The phosphoThr peptide binding rigidifies Leu275 - Asp278 of the remote 9/10 loop.

IV.4.10 Binding effects on flexibility on slow timescale

Exchange broadening suggests motions on the μsec to msec timescale for around 49 residues in the free state and 51 residues in the bound state of KI-FHA, by relaxation dispersion (Figures IV-1d, 2d) or statistical need for R_{ex} term in model-free calculations

(Figure IV-7). The slow fluctuations affect a contiguous swath of residues across β2, β1, β11, and β10 of the 6-stranded sheet (appearing left to right in Figures IV-7, IV-8). Adjacent to this central swath are other clusters of slowly exchanging residues. At the edge of the 5-stranded sheet, β4 and the 3/4 and 4/5 recognition loops are affected (Figure IV-7a). At the extreme opposite edge of the 6-stranded sheet, residues of β8, the 7/8 loop, and the 8/9 loop manifest motions on the μsec to msec scale (Figure IV-7a). Also exchanging are residues in the 1/2 and 10/11 loops near the binding site, the faraway 2/3 and 9/10 loops, and C-terminus. With the CLV1 pT868 peptide present at 3.5-fold excess, the peptide is expected to saturate 98% of KI-FHA based on the K_D of 20 μM. At 98% saturation, residues Lys221 - Ser223 of the 4/5 central recognition loop no longer show exchange broadening while residues Gly265 and Lys268 -Gly270 of the 8/9 peripheral recognition loop and Gly251 of the 6/7 recognition loop then have exchange broadening (Figures IV-4b, 5b, and 7). Gly227 and Thr285, with peaks most shifted and broadened during titration with the pT868 peptide (Figure IV-6a), are too broad in the peptide saturated state for relaxation fitting.

IV.4.11 Dissociation of CLV1 pT868 peptide from KI-FHA

Intermediate exchange broadening during titration with the pT868 peptide can be exploited to estimate the exchange rate k_{ex}. The off-rate k_{off} is equivalent to k_{ex}, since k_{on} is not at all limiting, being of the order of $\sim 5 \cdot 10^7$ $M^{-1}s^{-1}$. The range of k_{ex} and k_{off} possible is defined by the binding-induced chemical shift changes of Gly227 of $\Delta\nu_H$ =

569 Hz and Thr285 of $\Delta\nu_H$ = 356 Hz at 600 MHz; spectral and thermodynamic details will be published elsewhere. Gly227 in the slow-intermediate exchange regime limits k_{ex} to the range of $2.2*\Delta\nu_{H,G227} > k_{ex} > 0.01*\Delta\nu_{H,G227}$, i.e. 1250 $s^{-1} > k_{ex} > 5.7$ s^{-1}. Thr285 in the fast-intermediate exchange regime limits k_{ex} to this range: $100*\Delta\nu_{H,T285} > k_{ex} > \Delta\nu_{H,T285}$, i.e. 35,600 $s^{-1} > k_{ex} > 356$ s^{-1}. Combining these two ranges, leads to the conservative estimate that k_{off} and k_{ex} lie between 356 s^{-1} and 1250 s^{-1} for the pT868 interaction with KI-FHA. When pT868 is present at 1.5-fold excess resulting in only 94% saturation of KI-FHA, additional exchange broadening is seen in recognition loops in a comparison of 600 MHz R_2 values at 94% and 98% saturation. This is most evident in the 6/7 loop, followed by the 8/9 loop, the 4/5 loop, and conserved Gly211 (Figure IV-9). A set of model-free results were obtained from relaxation measurements at the 1.5-fold excess of pT868 that results in ~94% saturation of KI-FHA. At 94% saturation, more residues in recognition loops manifest chemical exchange broadening than at 98% saturation. At 94% saturation, such chemical exchange broadening is evident at 6/7 loop residues Met246, Leu249, Asn250, and Thr252, as well as conserved Gly211 and Asp263 and Ser266 of the 8/9 loop (Figures IV-9 and 10). Since each of these particular residues undergo chemical shift changes upon titration with the pT868 peptide (Figure IV-6a), it appears likely that their broadening at 94% saturation results from the peptide's apparent off-rate of between 350 and 1250 s^{-1}.

IV.5 Discussion

IV.5.1 Rigidity of residues conserved in the recognition site of FHA domains

Along the edge of the 5-stranded β-sheet, conserved Gly211, Arg212 (3/4 loop), Ser226, His229 (4/5 loop), and Asn250 (6/7 loop) are highly rigid (Figures IV-4a, 5a) and critical for forming the phosphoprotein binding site of FHA domains; the arginine, serine, and asparagine contact the phosphoThr of the ligand (*46, 49-51, 55, 62, 130*). The high rigidity of Gly211 and Arg212 of the 3/4 loop and His229 of the 4/5 loop (Figure IV-5) is consistent with the structure-stabilizing role proposed for conserved Gly211 and His229 (*57*). The unique conformation of absolutely conserved Gly211 could be important for positioning Arg212 to interact with the phosphate of the partner. Gly211 and Arg212 being recessed and largely buried under the interaction surface may confer their rigidity. When the S^2 values of Gly211, Arg212, His229, and Asn250 are interpreted with a diffusion-in-a-cone model of motion (*79*), the estimated amplitudes of N-H bond reorientations of their psec – nsec have cone semi-angles of 5° / 2°, 11° / NA, 14° / 13°, and 10° / 4°, for the free / bound states respectively. The backbone rigidity of Gly211, Arg212, His229, and Asn250 in the free state decreases energetic costs of loss of configurational entropy upon binding.

IV.5.2 Non-conserved residues of the phosphoprotein binding surface

Comparison of FHA domain structures of divergent sequence (*50, 57*) suggested that KI-FHA residues Glu224 in the 4/5 loop, Leu249 in the 6/7 loop, and Thr285 in the 10/11

loop may contribute to the key pT+3 site and KI-FHA's distinctive phosphopeptide specificity (*49*). In yeast Rad53 FHA domains, two more residues appear to contact the pT+3 peptide position (*157*); the first is equivalent to Gly284 and the second possibly to Thr286 of the 10/11 loop of KAPP. The amide peaks most shifted upon addition of the CLV1 pT868 peptide (with radial $\Delta\omega_{HN}$ > 0.15 ppm; Figure IV-6c) suggest a broad phosphoThr peptide recognition surface (Figure IV-6d). This same broad phosphoprotein (RLK)-binding surface is suggested by Evolutionary Trace analysis that has been introduced in Chapter III. Most residues of this surface are not, however, widely conserved among FHA domains (*49*). These non-conserved residues of the recognition loops can be placed in three groups of psec-nsec mobility (S^2). The first group is very rigid in both free and pT868-bound states: Lys221, Gly227, Leu249, Trp269, and Gly284. The second group is more flexible in the free state, undergoes restriction in the bound state, and is listed here with estimates of wobbling-in-cone semi-angles (see eq. 5) for free / bound states in parentheses: Val213 (18° / 8º), Ser216 (35º / 23º), Glu224 (28º / 17º), Lys228 (34º/ 10º), and Ser266 (27º / 13º). These are from solvent exposed portions of the 3/4, 4/5, and 8/9 loops (Figure IV-6), with the central 4/5 loop seeming to be in especially intimate contact with the phosphopeptide. The third group at the edge of the recognition surface has increased mobility in the pT868-bound state and is listed with estimated cone semi-angles: Gly251 (5° / 16°) of the 6/7 loop, Arg267 (21° / 26°) of the 8/9 loop, and Thr286 (18º / 25º) of the 10/11 loop.

IV.5.3 Rigidity and role in affinity of residues at the pT+3 position and conserved neighbors

Rigid portions of the phosphoprotein binding surface could confer affinity to this FHA domain and others. This hypothesis derives from observations on modules binding pTyr peptides: First, the residues of a PTB domain that contact that a relevant pTyr peptide are rigid, presumably promoting enthalpically favorable van der Waals contacts (*123*). Second, in the association of SH2 domains with pTyr peptides, high rigidity is characteristic of sites conferring affinity (*122*). Clearly, mostly conserved Gly211, Arg212, Ser226, His229, and Asn250 (KAPP numbering) are rigid in the free state (Figure IV-4) and apparently essential for affinity for receptor-like kinase domains (*46*). This behavior is consistent with the studies of pTyr peptide recognition. The pT+3 position of peptide ligands is the most important determinant of their affinity for FHA domains (*56, 62*). The non-conserved residues that divergent FHA domains place around the pT+3 position appear to be important for affinity (*157*). These residues of KI-FHA of KAPP, namely Gly227, Leu249, Gly284, and Thr285 are rigid (Figures IV-4, 5), consistent with the theme of high rigidity of residues conferring affinity (*122, 123*).

IV.5.4 Significance of pThr peptide binding-dependent flexibility changes remote from the RLK binding site

Energetic costs of pThr peptide binding include the loss of conformational entropy of the peptide, at sites in the 3/4, 4/5, and 8/9 recognition loops, at sites in 7 β–strands, and in three other loops (1/2, 2/3, and 9/10 loops) (Figure IV-6a, b). In the free state, there

appears to be more nsec-scale motion than in the bound state in all five recognition loops, other loops, and β8 (Figures IV-3a, 4c, 5c). Since S^2 is more likely to be underestimated at these locations (*155, 156*) in the free state, the binding-linked increases in rigidity and the entropic costs at these recognition loop residues could be underestimated. The modest increases in rigidity of seven of the 11 β-strands may not represent just costs in configurational entropy. The binding-dependent increases in S^2 appear to occur across networks of hydrogen-bonded residues across both β-sheets; one such network includes β3-β5-β6-β9 while another network includes β2-β1-β10-β11 (Figure IV-6b). The increased rigidity seen in both β-sheets suggests the possibility of improved hydrogen-bonding and side chain packing or van der Waals contacts, for a favorable change in the enthalpy of the bound state. This could provide part of the compensating favorable enthalpy driving the association of pThr peptides to KI-FHA and to another FHA domain (*57*). Peptide binding was also observed to rigidify and stabilize an SH3 domain both near and far from the interface (*128*). Precedent for enthalpic stabilization at long range was suggested by antibody binding to lysozyme increasing its hydrogen exchange protection to the distal side (*158*). Sites of binding-enhanced fast motions of KI-FHA, seen in the 1/2, 2/3, 6/7, 7/8, 8/9, and 10/11 loops, provide favorable increases in conformation entropy to compensate the entropic costs listed above.

IV.5.5 Flexibility of recognition surface in view of breadth of specificity

Several non-conserved residues of the recognition loops are more flexible than average: Ser216 (3/4 loop); Asp222, Glu224, Lys228 (4/5 loop); Ser260, Gly265 - Arg267 (8/9

loop); and Thr286 (10/11 loop). The mobility of these in the free state appears to be on the nsec scale (Figures IV-3a, 4c). The studies of SH2 domain affinity for pTyr peptides suggest that flexibility correlates well with contact sites of little importance to affinity. Mobility on the psec to nsec scale in a recognition surface may decrease the contribution of the mobile groups to affinity by diminishing favorable, strongly distance-dependent van der Waals interactions (*122*). Flexibility in the phosphoprotein binding surface of KI-FHA and other FHA domains might confer their breadth of specificity. An example of such flexibility could be the 8/9 loop of KAPP KI-FHA that distinguishes it from the phosphoprotein binding surface of other FHA domains. This 8/9 loop of KI-FHA is clearly longer, more solvent-exposed, and mobile on the nsec scale. The prominence and mobility of the 8/9 loop may be important in KAPP's recognition of several RLK targets in plants. Arabidopsis KAPP interacts in a phosphorylation-dependent manner with the following receptor-like kinases critically important to plant development and defense against infection: CLV1 (*28*), HAESA (*159*), WAK1 (*40*), BAK1 (*20*), SERK1 (*42*), and FLS2 (*41*). KAPP being promiscuous enough to attenuate multiple RLK-dependent signaling pathways in plants is supported moreover by KAPP being a single gene product without any known paralog of overlapping function (*31*). Yet, KAPP failed to interact with a number of kinases tested (*31*). The psec to nsec mobility of the RLK-binding surface of KI-FHA could foster the limited diversity of RLK partners recognized by KAPP. Perhaps an analogy could be drawn with the cytokine IL-2 where the flexible portion of its ligand-binding surface binds a far greater diversity than does the rigid portion (*160*). Both the psec-nsec flexibility and binding-inducible μsec to msec mobility

of the recognition loops could enable KI-FHA of KAPP to adjust to a minority of the 417 RLKs encoded by the Arabidopsis genome (*13*).

IV.5.6 Significance of μsec to msec fluctuations of 6-stranded β-sheet remote from phosphoprotein binding surface

Slow fluctuations of β2, β4, and β8 on the edges of the β-sandwich of KI-FHA may simply result from their solvent exposure. The largest line broadening evidence of fluctuations on the μsec to msec scale occurs at Ser256 of the 7/8 loop (Figures IV-4b, 5b, 7). Neighboring residues of the 7/8 β–hairpin, β10, and the 9/10 loop (Figures IV-4b, 5b, and 7) also display line broadening evidence of conformational exchange on the μsec to msec scale. This patch of slowly fluctuating residues coincides with a potential interaction surface predicted by Evolutionary Trace analysis to include many of the same residues of the β7/8 hairpin, β10, and the 9/10 loop (*49*). It is unlikely that a phosphoprotein partner can wrap 80° around from the phosphorecognition surface to this face of FHA domains. The β7/8 - L9/10 - β10 patch coincides with the site of segment-swapped dimerization in crystals of the FHA domain of human Chfr (*51*). (In concentrated solutions of the Chfr FHA domain, the dimer is a minor form relative to the predominant monomer (*51*).) Perhaps the slow conformational fluctuations of the β7/8 - L9/10 - β10 patch facilitate rearrangements such as the straightening of the 7/8 β–hairpin of the segment-swapped Chfr FHA dimer. The question has been raised as to whether this independent patch could be a site of dimerization of FHA domains more generally

(*49*). An alternative postulate regarding the patch is that it might play a role in self-recognition of another part of the same protein chain or signaling assembly (*49*). Slow fluctuations over μsec to msec have been observed at a number of protein-protein interaction sites (*153, 161-164*).

A swath of contiguous residues across the 6-stranded sheet undergo exchange broadening of their amide lines (sheet in foreground of Figure IV-7). Most of these have interior side chains arrayed in two packed rows across β2, β1, β11, and β10, plus β4 of the other sheet (Figure IV-8). Among these nine side chains, four—Leu184, Ile281, Val289, and Val291-- are among a dozen hydrophobic positions characteristic of the folding core of all FHA and SMAD MH2 domains (*49*). The upper row comprises Leu218 of β4, Leu194 of β2, Val186 of β1, Val289 of β11, and Ile281 of β10, progressing left-to-right across Figure IV-8. The lower row comprises His196 of β2, Leu184 of β1, Val291 of β11, and Asp279 of β10. The tight packing of this cluster of interior side chains plus the inter-strand hydrogen bonds might coordinate the backbone motions on the μsec to msec scale.

IV.5.7 The correlated slow internal motion and pH-dependent stability

The folding stability of KI-FHA has been estimated to be more than 3 kcal·mol^{-1} greater at pH 7.3 than at pH 6.3. Raising the pH to 7.3 where KI-FHA is more stable (*61*) causes the correlated slow motions at the 6-stranded anti-parallel β-sheet of KI-FHA to vanish (Figure IV-11). The quenching of these slow motions and greater stability at pH 7.3 may be intimately related. Perhaps KI-FHA is more stable at pH 7.3 because the absence of

slow correlated motions allows for better packing of the affected side chains, conferring more favorable enthalpy of folding stability through the improved Van der Waals contacts.

To recap, the conformational fluctuations of the loops on the nsec scale in the free state of KI-FHA are diminished (Figure IV-3) when the pT868 peptide fragment of the CLV1 receptor kinase partner is bound. A network of residues across the 6-stranded β-sheet undergoes conformational fluctuations on the μsec to msec scale independent of binding state. μsec to msec scale fluctuations are found at in a patch of residues at β7/8 - L9/10 - β10 corresponding to the site of crystallization-promoted, segment-swapped dimer formation in the FHA domain of human Chfr. With the phosphoThr peptide bound, some adjustments in sites of μsec to msec scale motion occur in the recognition loops, including net loss of slowly fluctuating sites in the 4/5 loop and a gain of sites in the 6/7 and 8/9 loops. Most of the broadening in the 6/7 loop seems to result from the peptide's off-rate of several hundred per sec. Peptide binding-enhanced rigidity of ten or more residues of recognition loops and adjoining 1/2 loop (Figure IV-7) is compensated by flexibility increases on the fast timescale of six residues on the periphery of the active site. The net overall increase in rigidity of KI-FHA phosphoThr peptide bound, with average $\Delta S^2 = 0.04$, includes increased rigidity at sites in seven β-strands (Figure IV-6). This suggests the possibility of long-range enhancement of favorable enthalpy to compensate entropic costs of binding. Favorable binding enthalpy is still more likely

from the backbone rigidity at the conserved residues clearly important to affinity and at neighboring residues most likely to contact the key pT+3 position that influences affinity.

IV.6 Acknowledgement

We thank NMRFAM (Madison, WI) for access to the cryoprobe™-equipped Bruker DMX-500.

Table IV-1: Spectral density functions used to fit relaxation data in the extended Lipari-Szabo model-free approach (*77, 82*).

Model	Fit parameters	Spectral density function
1	S^2	$J(\omega)=2/5\{(S^2\tau_m)/[1+(\omega\tau_m)^2]\}$
2	S^2, τ_e	$J(\omega)=2/5\{(S^2\tau_m)/[1+(\omega\tau_m)^2]+[(1-S^2)\tau]/[1+(\omega\tau)^2]\}$
3	S^2, R_{ex}	$J(\omega)=2/5\{(S^2\tau_m)/[1+(\omega\tau_m)^2]\}$
4	S^2, τ_e, R_{ex}	$J(\omega)=2/5\{(S^2\tau_m)/[1+(\omega\tau_m)^2]+[(1-S^2)\tau]/[1+(\omega\tau)^2]\}$
5	S_s^2, S_f^2, τ_e	$J(\omega)=(2S_f^2/5)\{(S_s^2\tau_m)/[1+(\omega\tau_m)^2]+ [(1-S_s^2)\tau]/[1+(\omega\tau)^2]\}$

$\tau = \tau_e\tau_m/(\tau_e+\tau_m)$, where τ_m is the isotropic rotational correlation time and τ_e is the effective correlation time for internal motions. $S^2 = S_f^2S_s^2$ is the square of the generalized order parameters. S_f^2 and S_s^2 are the squares of the order parameters for the internal motions on the fast and slow timescales.

Table IV-2: Average filtered relaxation parameters of KI-FHA, free and bound to pT868

			coarse-	filtered[A]				fine-	filtered[B]		
state of KI-FHA	*Field, Tesla*	*n* [A]	$^{15}N\{^1H\}$ $\overline{nOe}$	$\overline{R}_1$, s^{-1}	$\overline{R}_2$, s^{-1}	*n* [B]	$^{15}N\{^1H\}$ $\overline{nOe}$	$\overline{R}_1$, s^{-1}	$\overline{R}_2$, s^{-1}	τ_m, [C] *nsec*	$D_\parallel/D_\perp$ [C]
Free	11.74 14.1	83 80	0.76 0.78	2.07 1.89	11.6 12.6	46 43	0.76 0.79	2.05 1.85	11.5 12.3	7.57±0.08	0.81±0.07
Bound to pT868[D]	11.74 14.1	57 84	0.75 0.80	1.91 1.41	14.4 16.9 [E] (16.2) [F]	39 42	0.72 0.80	1.87 1.35	14.2 16.4 [E] (16.1) [F]	10.70±0.07	1.21±0.06

NOE, R_1, R_2 and R_2/R_1 were trimmed prior to averaging.

[A] Coarse filtering employed Normadyn software to remove residues with NOE < 0.65 or with $T_2 \le \overline{T_2} + \sigma_{T2}$ unless $T_1 \ge \overline{T_1} + \sigma_{T1}$ (*150*). Number of residues after filtration is *n*.

[B] Fine filtering removed residues poorly fit by simple model-free expressions 1 and 2 (Table IV-1). Residues with exchange broadening suggested by relaxation dispersion at 600 MHz or statistical model section were removed. Residues suggested to have nsec-scale motions by $J(\varpi_N = 61MHz) \ge \overline{J(\varpi_N)} + \sigma_{J(\varpi N)}$ were also removed. Number of residues after filtration is *n*.

[C] Rotational correlation time τ_m and the $D_{\parallel}/D_{\perp}$ ratio were estimated using the fine-filtered set of residues, their 600 MHz R_1 and R_2 relaxation data, Tensor2 (*80*), and PDB coordinates 1MZK.

[D] Except where noted, KI-FHA was present at 0.53 mM and the peptide at 1.5-fold excess, suggesting 94% peptide saturation of KI-FHA since K_D is ~20 μM.

[E] These values result from increasing the peptide excess to 3.5-fold, boosting peptide saturation of the 0.33 mM KI-FHA to 98%.

[F] Parenthesized values derive from R_2 measured after dilution of KI-FHA to 0.23 mM with 0.35 mM pT868 peptide. These conditions decreased peptide saturation of KI-FHA to ~88%. Uncertainties in R_2 at 0.23 mM exceed differences from R_2 at higher concentration and saturation.

Table IV-3: Spectral density functions (see Table IV-1) selected for model-free simulations of KI-FHA from Arabidopsis KAPP

Number	Residue	Function, Free state	Function, bound	Number	Residue	Function, Free state	Function, bound
177	L	2	2	239	F	3	2
178	G	2	4	240	K	4	1
179	S	2	2	241	W	2	2
180	S	2	4	242	E	2	1
181	W	1	4	243	L	1	2
182	L	2	4	244	V	2	1
183	F	4	3	245	D	2	2
184	L	4	3	246	K	N/A	2
185	E	4	3	247	G	N/A	2
186	V	3	4	249	L	2	2
188	A	4	3	250	N	2	1
189	G	1	1	251	G	2	4
191	P	3	4	252	T	2	2
192	A	4	3	253	L	1	1
193	I	4	4	254	V	2	1
194	G	4	3	255	N	2	1
195	L	4	2	256	S	4	4
196	H	4	3	257	H	2	3
197	A	4	3	258	S	4	2
198	V	N/A	4	259	I	4	4
199	N	4	4	260	S	4	4
200	S	2	3	261	H	2	4
201	T	4	4	263	D	2	2
202	S	2	4	264	L	2	N/A

203	S	4	4	265	G	2	4
205	K	2	3	266	S	4	2
206	L	4	1	267	R	4	4
208	V	2	1	268	K	2	4
209	K	2	2	269	W	2	3
211	G	2	2	270	G	2	3
212	R	2	N/A	271	N	2	1
213	V	4	3	273	V	2	1
216	S	4	4	274	E	2	2
218	L	4	3	275	L	4	1
219	A	3	3	276	A	3	3
220	L	2	1	277	S	4	1
221	K	3	1	278	D	4	1
222	D	4	N/A	279	D	4	3
223	S	4	N/A	280	I	2	1
224	E	4	4	281	I	3	3
225	V	2	3	282	T	1	2
227	G	3	N/A	283	L	1	1
228	K	4	3	284	G	1	1
229	H	2	4	286	T	2	4
230	A	1	1	288	K	2	1
231	Q	1	1	289	V	4	3
232	I	2	1	290	Y	4	3
233	T	2	1	291	V	4	3
234	W	2	1	292	R	4	1
235	N	1	2	294	S	4	1
236	S	2	4	295	S	4	4
237	T	2	2	296	Q	4	4
238	K	2	1	297	N	4	4

Figure IV-1: ^{15}N relaxation data of KI-FHA measured at 500 and 600 MHz at 25 °C and plotted vs. residues from the KAPP sequence: a) $^{15}N\{^{1}H\}$ NOE plotted as I_{sat} / I_{nonsat},, b) spin-lattice relaxation rate constant, R_1, c) spin-spin relaxation rate constant, R_2 and d) at 22 °C, relaxation-compensated CPMG measurement of $\Delta R_2 = R_2$ (7.5 ms) - R_2 (2 ms), where 7.5 and 2 ms are the spacing between π pulses of the CPMG trains. Uncertainties were estimated as described in Methods. In b), gray squares mark four cases of poor single exponential fits due to the clear presence of a second, fast, T_2-like exponential decay; for the four affected residues, the fit of the slow decaying component is shown with black squares. The locations of β-strands are marked with gray bars in each panel. Filled squares represent 500 MHz data. Open circles represent 600 MHz data, except that in panel (d) 600 MHz data are represented with filled squares.

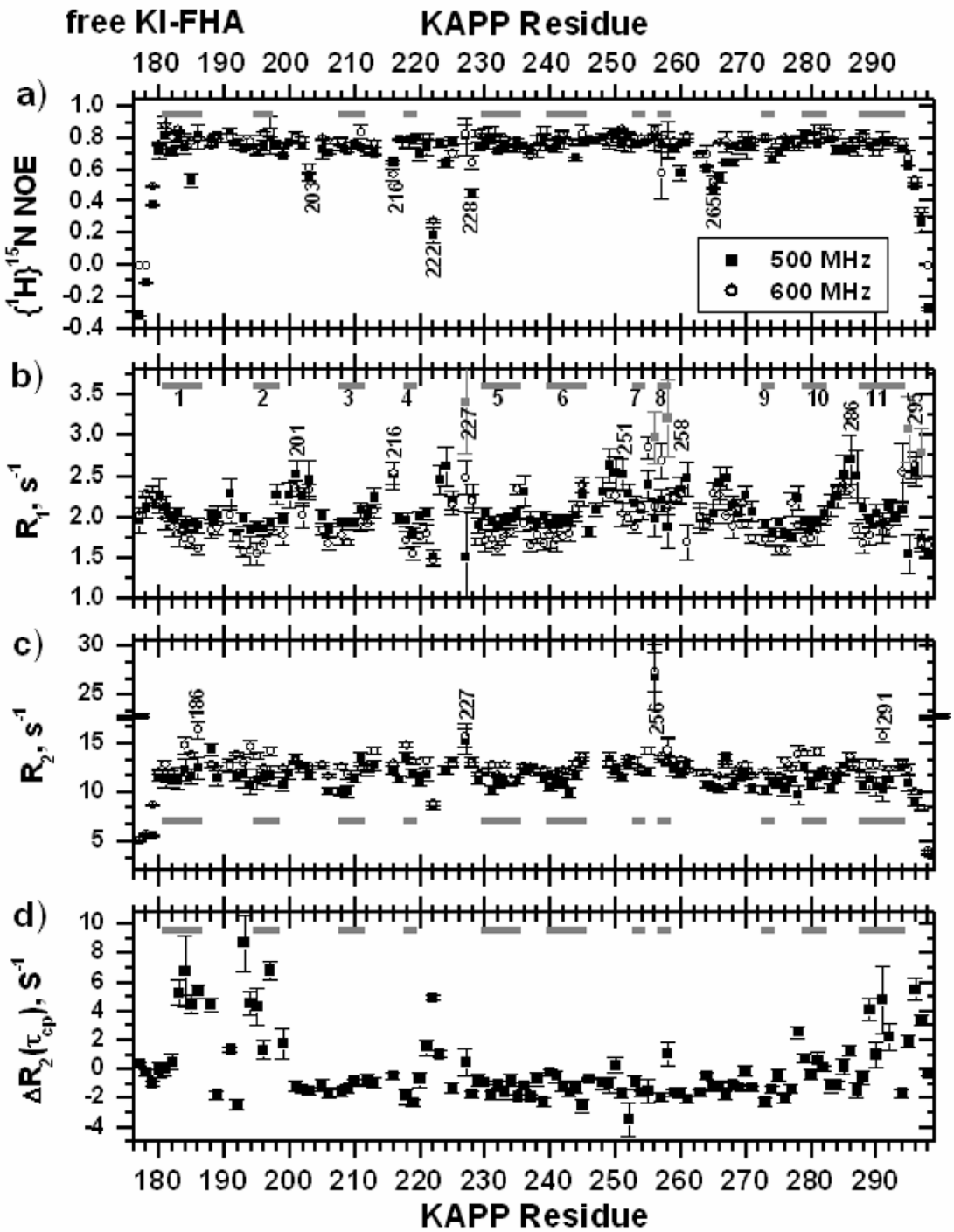

free KI-FHA
KAPP Residue
a)
{1H}15N NOE
500 MHz
600 MHz
b)
R1, s-1
c)
R2, s-1
d)
ΔR2(τcp), s-1
KAPP Residue

Figure IV-2: ^{15}N relaxation data of KI-FHA bound to pThr868CLV1 at 25 °C, and plotted vs. residue number from the KAPP sequence: a) $^{15}N\{^{1}H\}$ NOE, b) R_1, c) R_2 and d) $\Delta R_2 = R_2$ (7.5 ms) - R_2 (2 ms). The symbol code of Figure 1 is used.

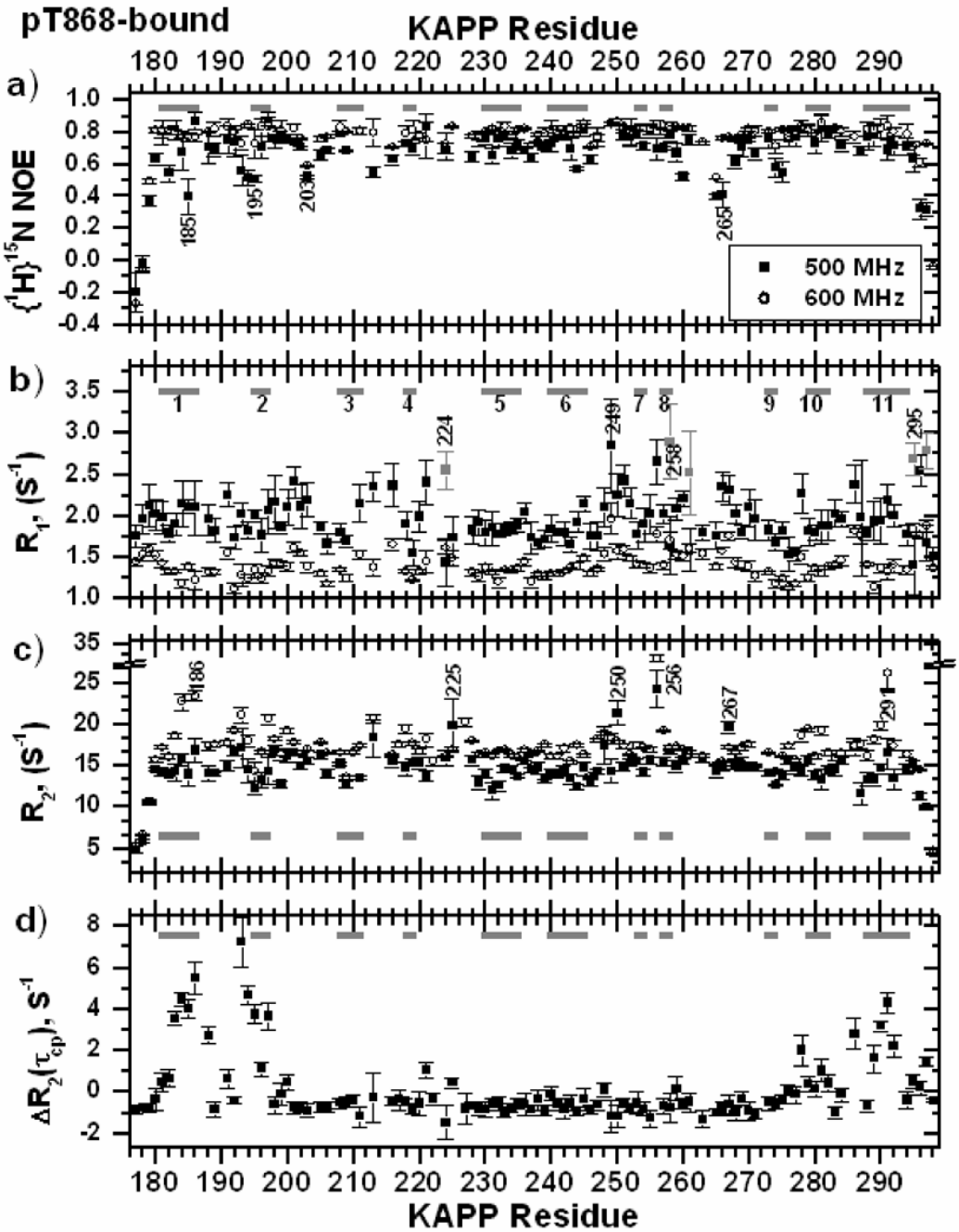

pT868-bound
KAPP Residue
a)
{1H}15N NOE
500 MHz
600 MHz
b)
R1, (S-1)
c)
R2, (S-1)
d)
ΔR2(τcp), S-1
KAPP Residue

Figure IV-3: Evidence for nsec backbone motions. a) Reduced spectral density function $J(\omega_N)$ plotted vs. residue, where $\omega_N=3.8*10^8$ rad/sec since the relaxation data were collected at 600 MHz. Open squares represent the free state and filled triangles represent the state bound to pT868 from CLV1. b) Sites in free KI-FHA where residual dipolar couplings (RDCs) deviate from those predicted by SVD fits to the NMR structural coordinates of KI-FHA (PDB code 1mzk). Squared deviations of individual $^1D_{NH}$ measured RDCs from those predicted by SVD fits of all RDCs to the entire backbone are normalized by D_a, the axial component of the powder pattern distribution of RDCs observed. This quantity is averaged for the ten lowest energy NMR structures of KI-FHA.

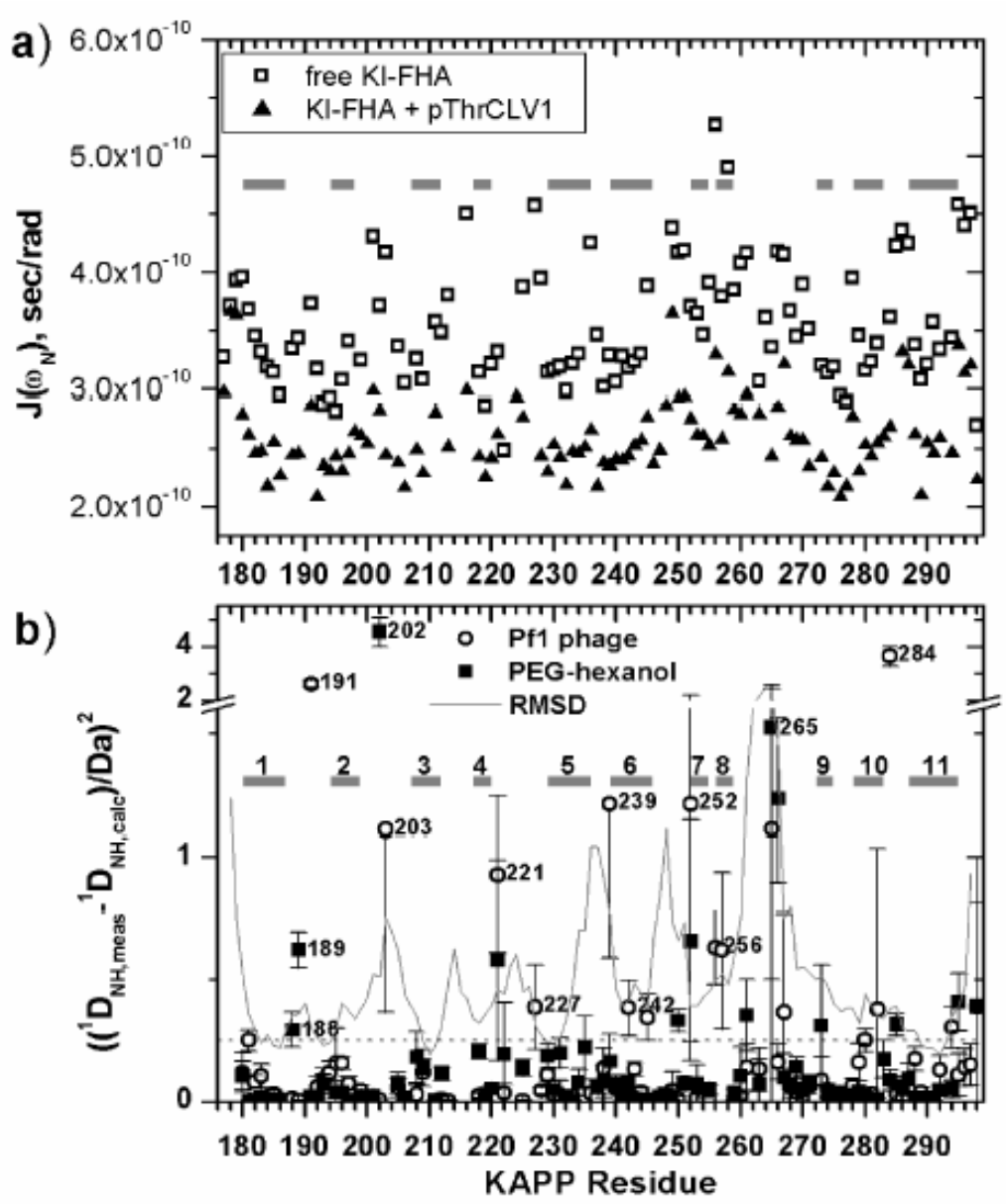

a)
free KI-FHA
KI-FHA + pThrCLV1
J(ω_N), sec/rad
6.0x10^-10
5.0x10^-10
4.0x10^-10
3.0x10^-10
2.0x10^-10
180 190 200 210 220 230 240 250 260 270 280 290
b)
Pf1 phage
PEG-hexanol
RMSD
((1D_NH,meas - 1D_NH,calc)/Da)^2
1 2 3 4 5 6 7 8 9 10 11
202
191
284
265
239
252
203
221
189
256
227
242
188
KAPP Residue

Figure IV-4: Model-free dynamics results for the free state of KI-FHA from fits of ^{15}N relaxation at two fields, 500 MHz and 600 MHz. a) The generalized order parameter S^2, b) the line broadening term R_{ex} and c) the internal correlation time τ_e are plotted vs. sequence position in KI-FHA. Locations of β-strands are marked with bars.

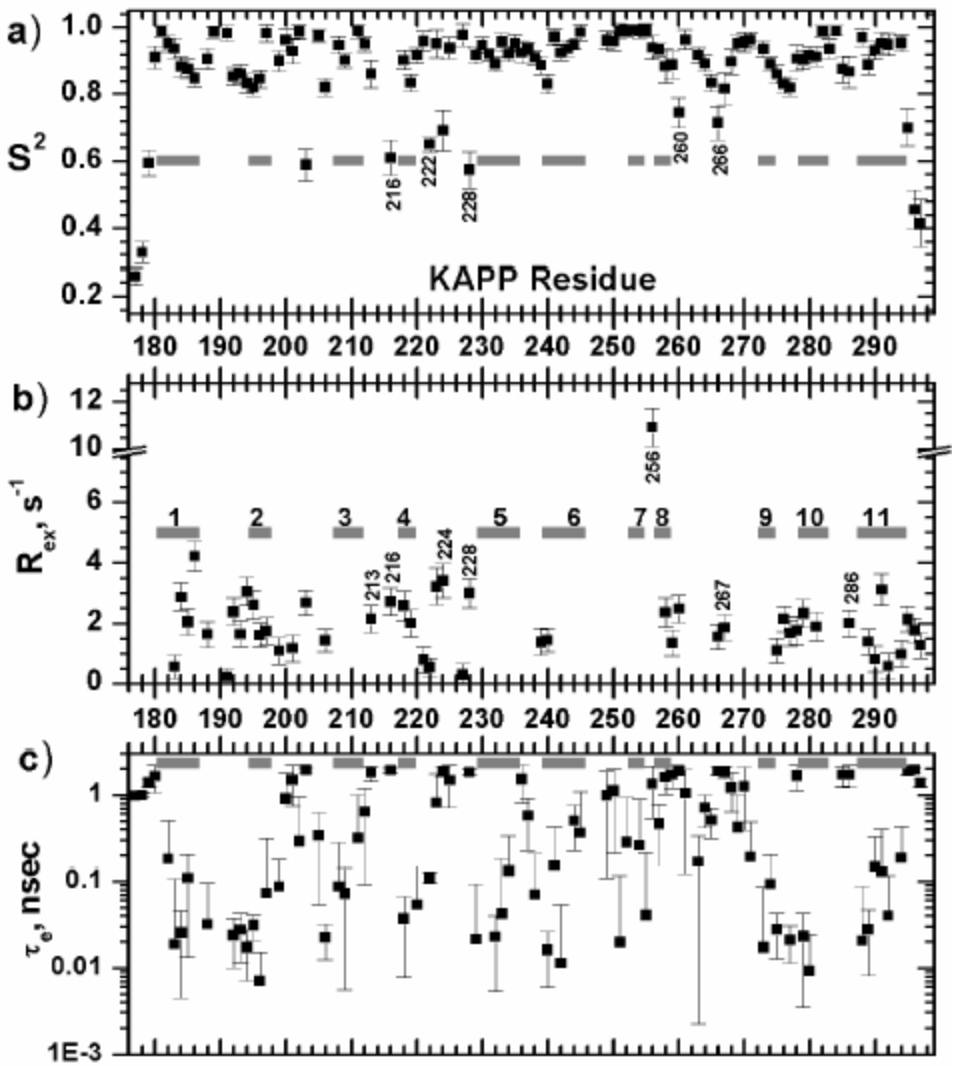


Figure IV-5: Model-free dynamics results for KI-FHA bound to CLV1 pT868 using ^{15}N relaxation collected at both 500 MHz and 600 MHz. a) The generalized order parameter S^2, b) the line broadening term R_{ex} and c) the internal correlation time τ_e are plotted vs. sequence position in KI-FHA. Locations of β-strands are marked with bars.

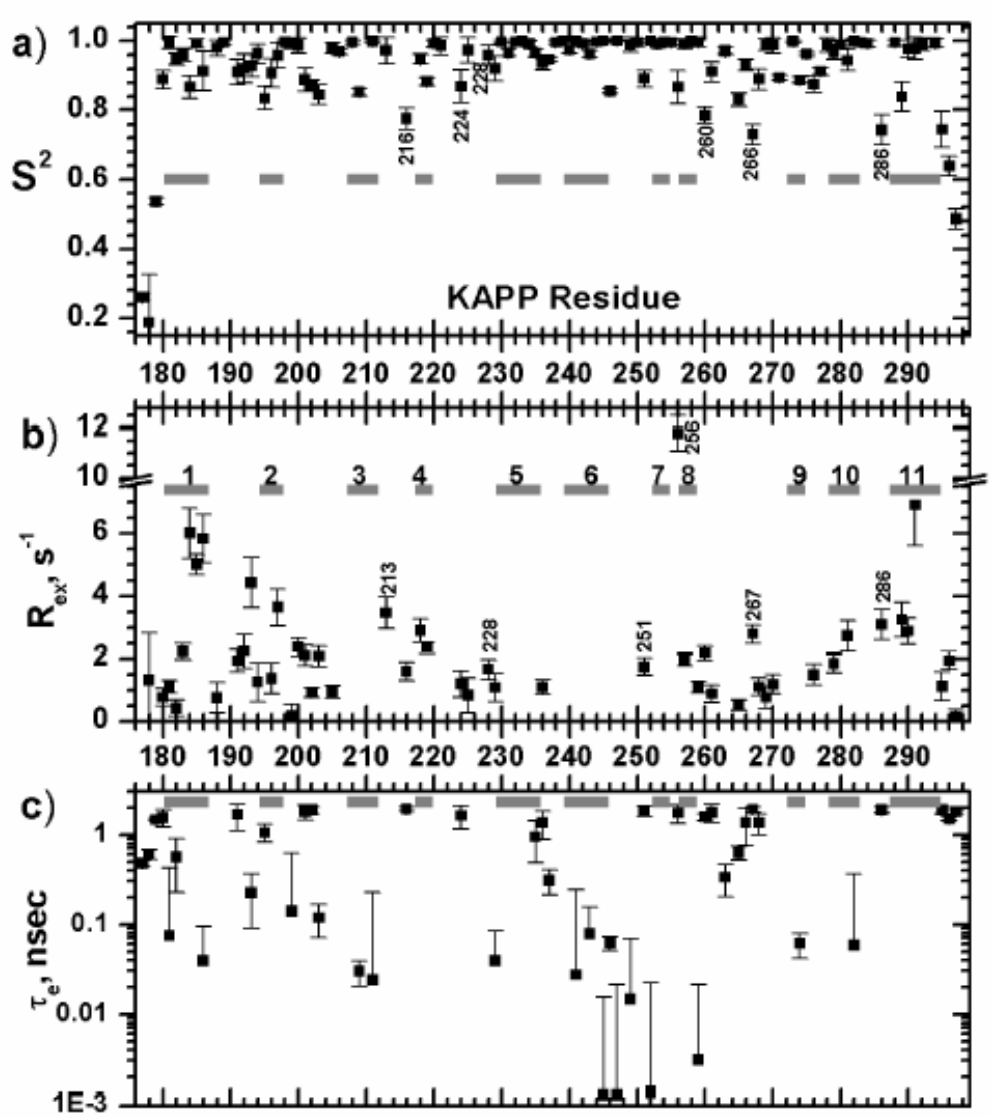


Figure IV-6: Comparison of pT868 peptide-induced changes in KI-FHA flexibility over psec-nsec and in amide NMR peak positions at the interface. a) Changes in the generalized order parameter S^2 upon the binding of the pT868 peptide are plotted vs. residue of KAPP. The locations of β-strands are marked with gray bars. Residues with $S^2_{bound} - S^2_{free} \geq 0.065$, marked by a dashed line, represent the upper 30-percentile of increases. b) Sites of ΔS^2 are colored on the stereo view of the backbone of free KI-FHA with tube width proportional to $1 - S^2_{free}$, i.e. for the free state. Residues rigidified by pT868 peptide binding are colored dark blue where $S^2_{bound} - S^2_{free} > 0.14$ in (a) and lighter blue where $0.14 > \Delta S^2 > 0.065$ in (a). Residues mobilized by pT868 peptide binding are colored red where $S^2_{bound} - S^2_{free} \leq -0.065$. b) Shifts of amide NMR peaks of KI-FHA upon addition of a 4-fold excess of the CLV1 pT868 peptide are plotted vs. location in KAPP. The radial shift of the peak $\Delta\omega_{NH} = (\Delta\omega_H^2 + (\Delta\omega_H/6)^2)^{1/2}$, where $\Delta\omega_H$ and $\Delta\omega_N$ are the changes in ^1H and ^{15}N dimensions in ppm. In d), pT868-induced chemical shift changes, $\Delta\omega_{NH}$, are mapped onto the stereo view of the backbone of the free KI-FHA structure with tube width proportional to $1 - S^2_{bound}$, representing the bound state. Residues are colored in red where $\Delta\omega_{NH} > 0.35$, orange where $0.35 > \Delta\omega_{NH} > 0.15$, and gold where $0.15 > \Delta\omega_{NH} > 0.067$. A few of the most shifted residues are labeled. In c) and d), unobservable residues are colored a darker shade of gray.

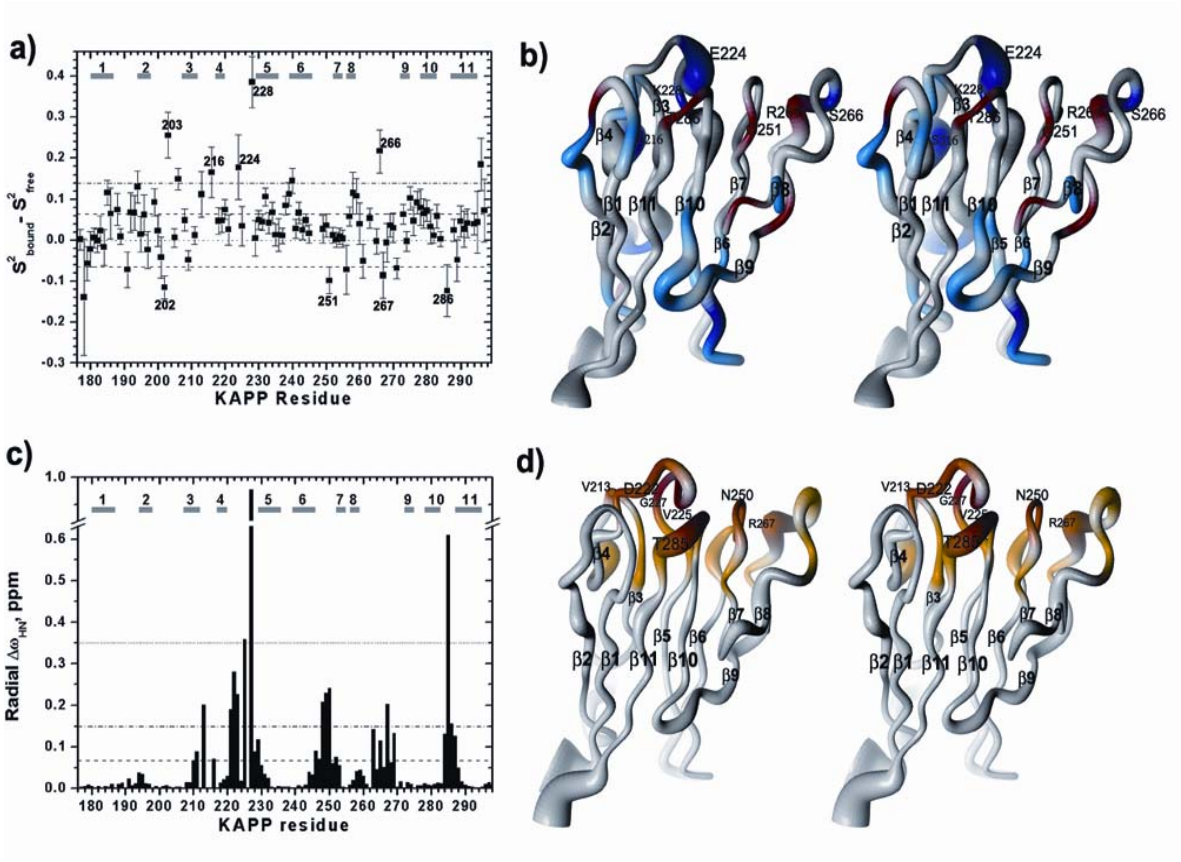

a)
b)
c)
d)
KAPP Residue
KAPP residue
Radial Δω_HN, ppm
E224
S266
N250
V213
V225
R267

Figure IV-7: Mapping of sites of μsec to msec scale fluctuations and of nsec scale motions, of backbone amide peaks upon the KI-FHA NMR structure for the free state (a) and pT868-bound state (b). In (a) and (b), residues with $R_{ex} > 1\ s^{-1}$ exchange broadening evidence of the slower motions are colored red and those with $1\ s^{-1} > R_{ex} > 0\ s^{-1}$ are colored orange. Residues with apparent $\tau_e > 0.7$ nsec are marked with blue spheres. Unobservable residues are colored a darker shade of gray. Structural elements with R_{ex} in free KI-FHA are labeled in (a). In (b), residues are labeled where binding of the pT868 peptide introduces R_{ex}. The binding surface for phosphoThr peptides or phosphoproteins (receptor-like kinases such as CLV1 in plants) is indicated with red arrows. Residues in the bound state undergoing too much chemical exchange broadening for relaxation fitting are labeled in pink; these residues exhibit intermediate exchange broadening during titration with the pT868 peptide.

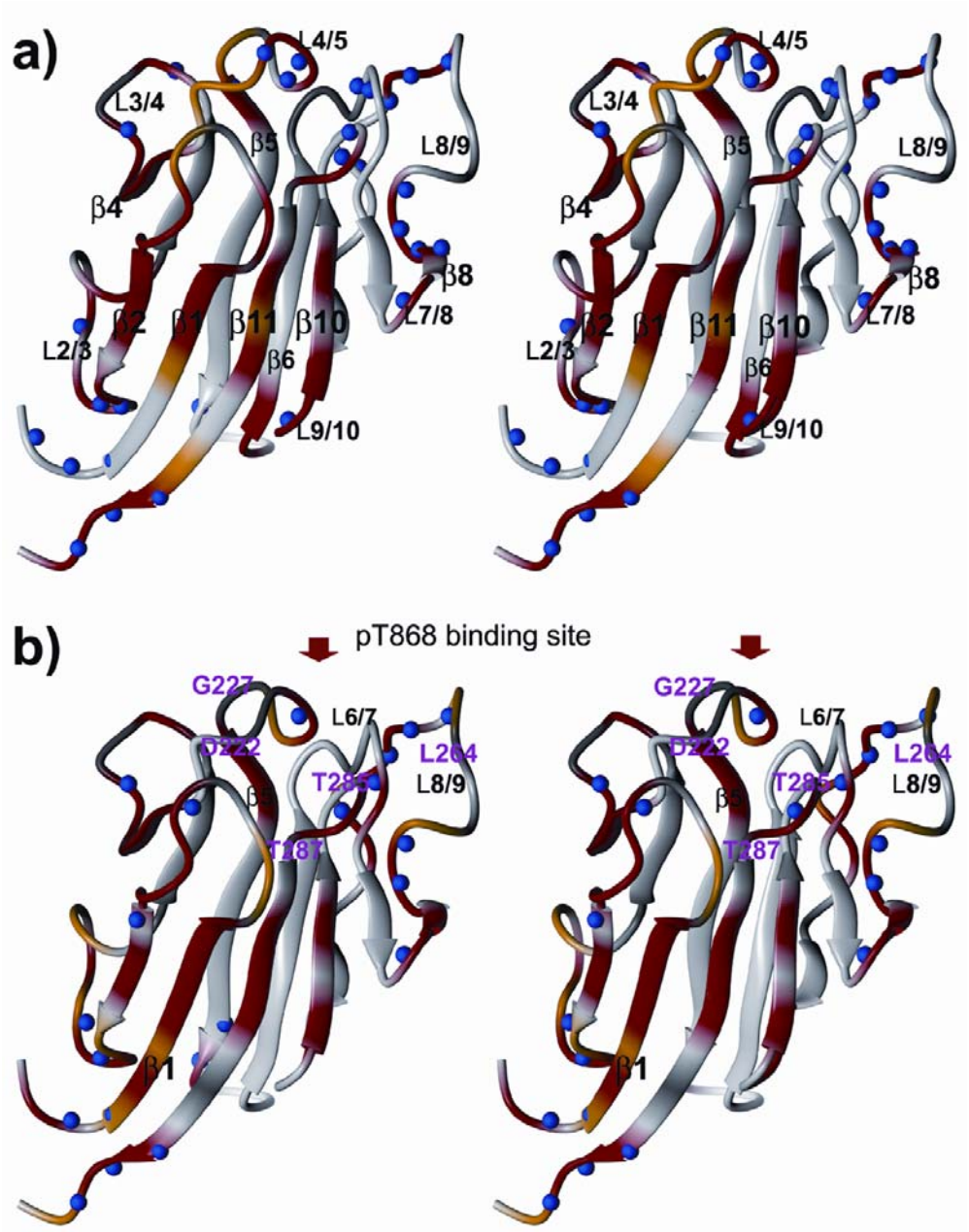

a)
L4/5
L3/4
β5
L8/9
β4
β8
L7/8
β2
β1
β11
β10
L2/3
β6
L9/10
b)
pT868 binding site
G227
L6/7
D222
L264
T285
L8/9
β5
T287
β1

Figure IV-8: A network of two tiers of packed interior side chains, plotted in blue, where the backbone is affected by conformational exchange detected as line broadening. The color code of the backbone is that of Figure 7.

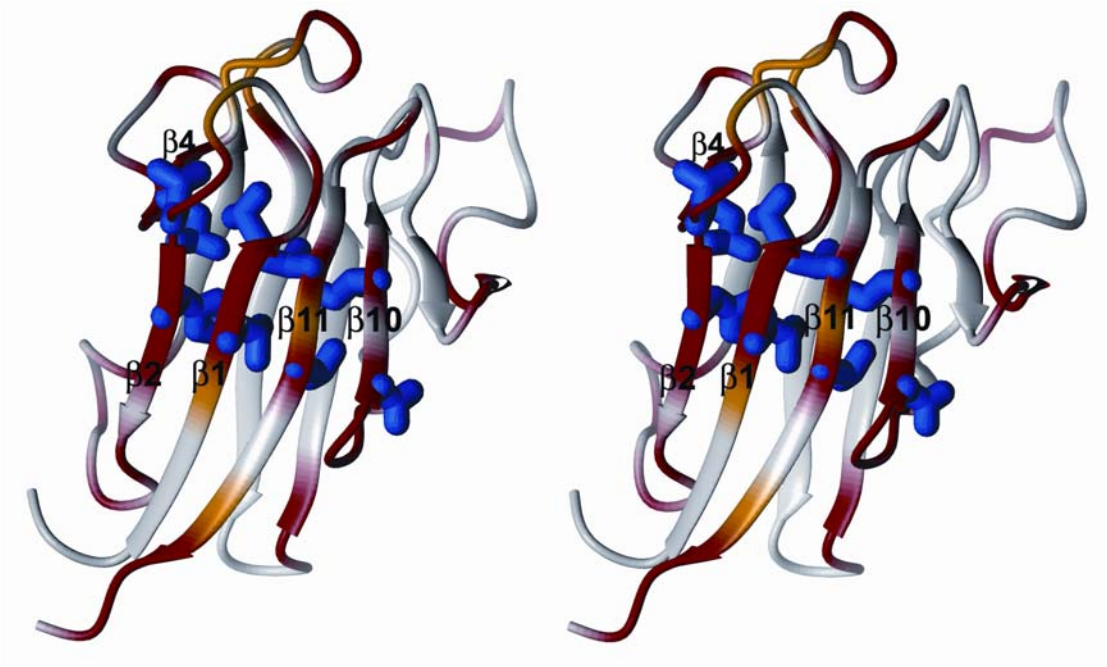


Figure IV-9: Greater extent of exchange broadening when KI-FHA is 94% saturated with pT868 peptide present at 1.5-fold excess. Panel a) overlays 600 MHz R2 collected on a 0.53 mM sample of KI-FHA with 1.5-fold excess of pT868 peptide (open blue circles; 94% saturation with peptide) with 600 MHz R2 collected on a 0.33 mM sample of KI-FHA with 3.5-fold excess of the peptide (filled squares; 98% saturation with peptide). Model-free dynamics results on slower motions of KI-FHA, 94% saturated by the 1.5-fold excess of CLV1 pT868, are given in b) the exchange broadening term Rex and in c) the internal correlation time τ_e. Locations of β-strands are marked with bars. At 94% saturation with pT868 peptide, the number of residues requiring a R_{ex} term is increased to 63. The added sites with chemical exchange broadening are found not only in the 4/5, 6/7, 8/9, and 10/11 recognition loops, but also in the nearby 1/2 loop, β-strands 3, 6, and 7, and in the remote 2/3 and 5/6 loops. The additional chemical exchange broadening in the 6/7, 4/5, and 10/11 recognition loops is attributed to the peptide's off-rate of several hundred per sec.

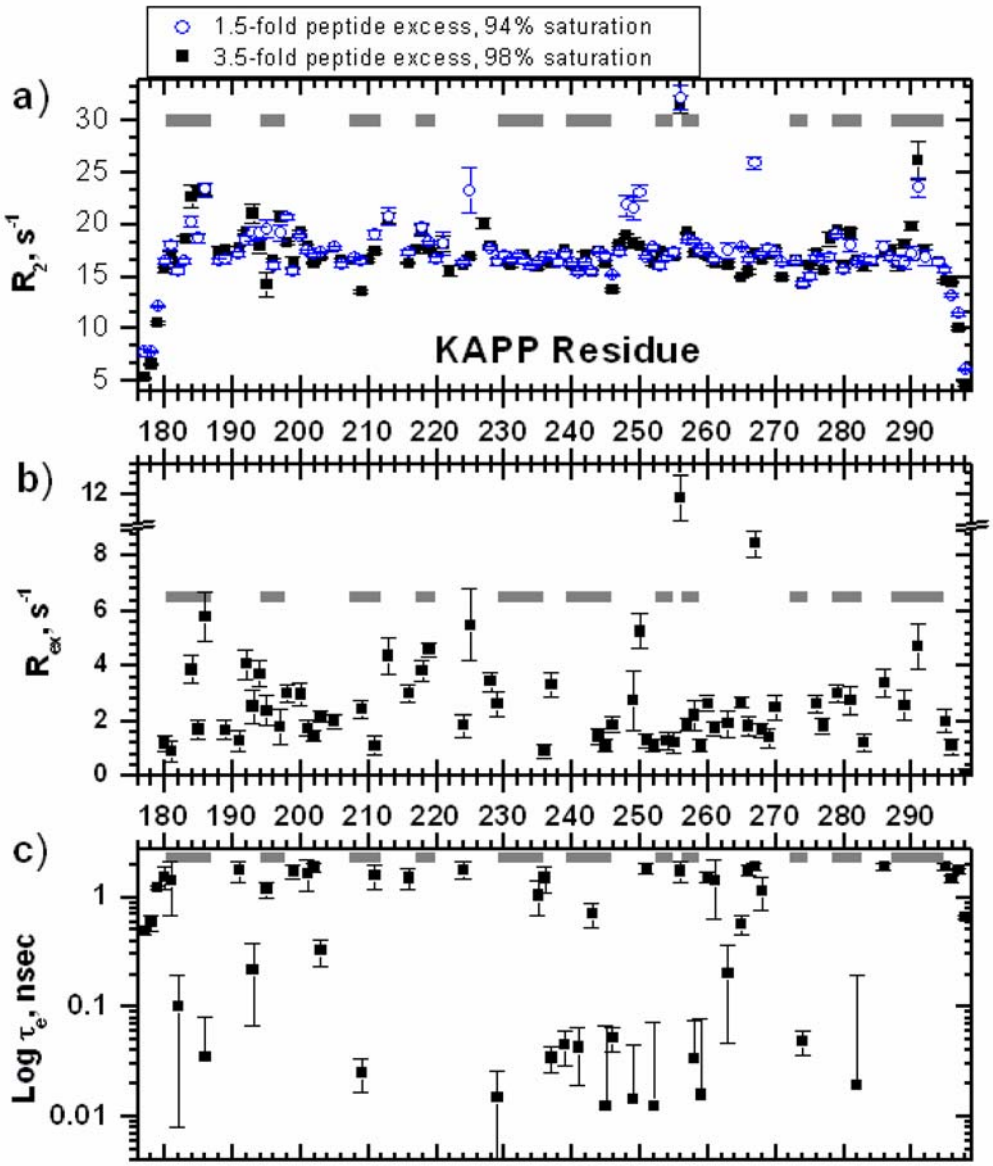

1.5-fold peptide excess, 94% saturation
3.5-fold peptide excess, 98% saturation
a)
R_2, s^{-1}
KAPP Residue
b)
R_{ex}, s^{-1}
c)
Log τ_e, nsec

Figure IV-10: Depiction of the extent of exchange broadening when KI-FHA is 94% saturated with pT868 peptide present at 1.5-fold excess. Residues with $R_{ex} > 2.6\ s^{-1}$ exchange broadening evidence of the slower motions are colored red. Residues with $2.6\ s^{-1} > R_{ex} > 1.0\ s^{-1}$ are colored orange. Residues with apparent $\tau_e > 0.7$ nsec are marked with blue spheres. Unobservable residues are colored a darker shade of gray. Structural elements are labeled where binding of the pT868 peptide introduces R_{ex}. The binding surface for phosphoThr peptides or phosphoproteins (receptor-like kinases such as CLV1 in plants) is indicated with red arrows. Residues in the bound state undergoing too much chemical exchange broadening for relaxation fitting are labeled in pink; these residues exhibit intermediate exchange broadening during titration with the pT868 peptide.

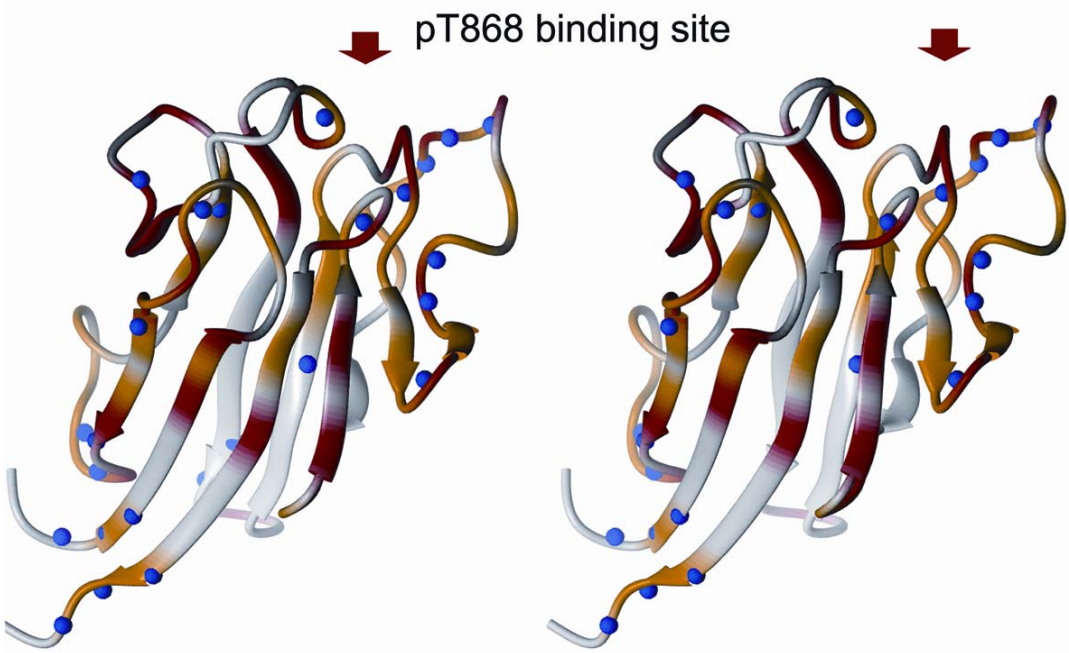
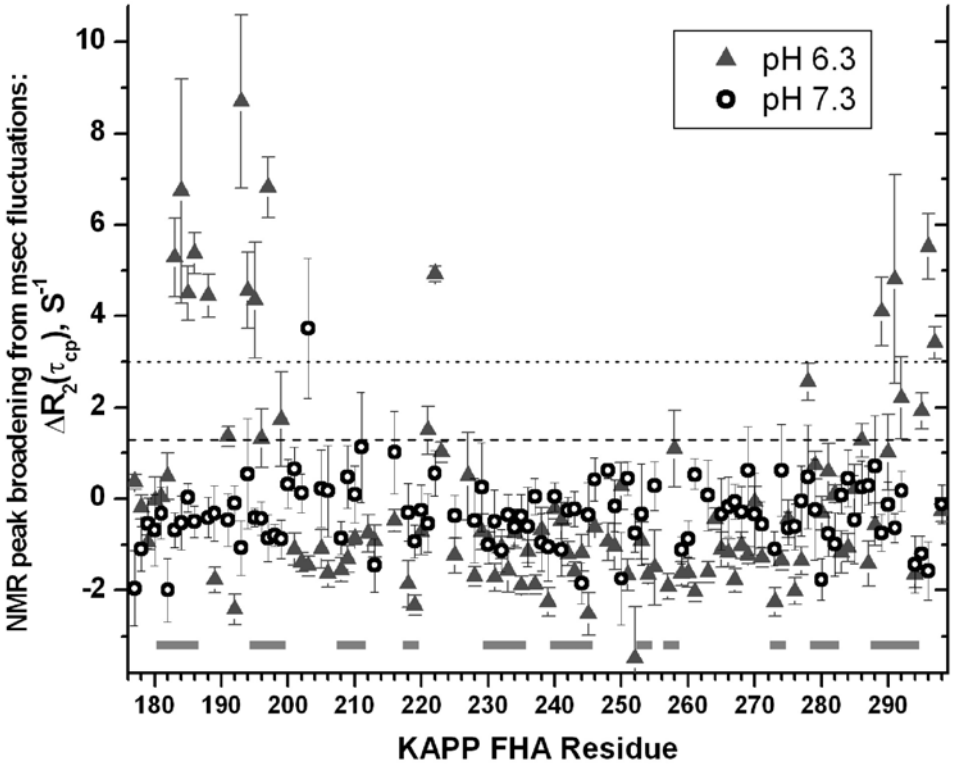


Figure IV-11: Slow fluctuations of KI-FHA at pH 6.3 and pH 7.3. Relaxation-compensated CPMG measurement of $\Delta R_2 = R_2$ (7.5 ms) - R_2 (2 ms), where 7.5 and 2 ms are the spacing between π pulses of the CPMG trains, was plotted against KI-FHA residues. CPMG data collected at pH 6.3 and pH 7.3 are represented with filled triangles and open spheres, respectively. β-strands are represented with gray bars. Uncertainties were estimated as described in Methods.



Chapter V: Preliminary results and future directions

Besides the two primary projects, phosphopeptide interactions and backbone dynamics of the KAPP KI-FHA domain, I have also participated in several preliminary studies for the purpose of grant initiations or collaborations. Although not complete scientific stories, these preliminary investigations are of great importance and interest for people who will follow the data and develop these projects further. Therefore, in this chapter I will describe those preliminary but interesting results and discuss future research directions.

V.1 Preparations of RLK kinase domains and preliminary structural studies

V.1.1 Introduction

BRI1-associated kinase 1 (BAK1) and Flagellin insensitive 2 (FLS2) are Arabidopsis LRR-RLKs that function in BRI1-mediated BR signaling and in pathogen defense (*20,*

35). The roles of BAK1 and FLS2 in signal transduction in plants have been introduced in Chapter I. Both RLKs interact with the KI-FHA domain of KAPP. A long-term goal in our laboratory has been to elucidate the mechanism of the interactions between KAPP and RLK kinase domains or to determine how the RLK-KAPP signaling complex is assembled. No experimental structure of any plant RLK kinase domain has been reported to date, perhaps because RLK kinase domains are typically insoluble and unstable *in vitro*. Futhermore, it is difficult to prepare a homogenously phosphorylated RLK kinase domain. Therefore, efforts have been made to prepare samples of the kinase domains of BAK1 and FLS2 suitable for structural studies.

V.1.2 Material and methods

V.1.2.1 Molecular cloning of the kinase domain of FLS2

Recombinant GST-FLS2 (F870 – V1173) in pBSK vector was generously provided by Dr. Scott Peck (University of Missouri-Columbia) and was used as template for PCR. The forward primer was 5'-GGTTTCATATGTTCAACAGTGCCAACATCATTG-3' containing an XhoI endonuclease site and the reverse primer was 5'-ATCCGCTCGAGCTAAACTTCTCGATCCTCGTTACG-3' harboring an NdeI site. PCR products were extracted from DNA agarose, cleaved by endonuclease XhoI/NdeI and ligated with pET-15b plasmid predigested with XhoI/NdeI. The ligated recombinant

vector was transformed into Top10 competent cells. The purified recombinant vector was then sequenced and transformed into BL21 (DE3 RIL) cells for expression.

V.1.2.2 Preparations of RLK kinase domains from inclusion bodies

HIS-FLS2 (in pET15b) and GST-BAK1 (in pGex-6p-1) were inoculated into 200 ml of LB media, induced with 0.2 mM IPTG at $O.D._{600}$ ~0.7, and expressed at 30 ^{o}C for 4 hours. Cells were pelleted by centrifugation, and resuspended in 10 ml of PBS buffer (pH 7.5). Cell lysis was performed via French Press. Pellets were washed in 40 ml of ice-cold buffer containing 50 mM Tris and 1 mM EDTA (pH8.0). Cells were pelleted again and washed in 40 ml of ice-cold buffer with 50 mM Tris, 1 mM EDTA (pH8.0) and 1% Triton X-100. This step was repeated once. Cells were pelleted again and resuspended in 40 ml of ice-cold buffer with 50 mM Tris, 1 mM EDTA (pH 8.0), and 0.5 M Guanidinium HCl. This step was performed two times. The following two methods were then used to recover the kinase domains of RLK from the pellets.

(1) Denaturing and refolding

The pellets were dissolved in 10-20 ml of denaturing buffer (50 mM Tris, 1 mM EDTA, 8 M Urea, 1 mM DTT, pH 8.0). The mixture was incubated with gentle shaking for one hour in room temp and diluted into 5-10 X volumes of rapidly stirring refolding buffer (50mM Hepes pH 7.5, 0.2 M NaCl, 1 mM DTT, 0.5M NDSB201 (*165*)) in 4 ^{o}C. The mixture was incubated overnight with constant stirring and dialyzed into PBS buffer (pH 7.5) in 4 ^{o}C. Precipitation was removed by centrifugation and cocktail protease inhibitor was added.

(2) Detergent (N-Lauroylsarcosine) solubilization

The pellets were resuspended in 10 ml of ice-cold buffer (1% N-Lauroylsarcosine [Sigma], 25 mM Triethanolamine [Sigma], 1 mM EDTA pH 8.0) and incubated for 15 minutes in 4 ^{o}C. Precipitation was removed by centrifugation, and the supernatant was dialyzed into PBS buffer (pH 7.5) with 1 mM DTT. Cocktail protease inhibitor was added.

V.1.2.3 Protein autophosphorylation assay

Autophosphorylation was assayed on the kinase domains of BAK1 and FLS2 as described previously (*20*) and with minor modifications. Each reaction was set up in 20 μl volume with 2 μg sample, 1x kinase buffer (50 mM HEPES [pH 7.4], 10 mM $MgCl_2$, 10 mM $MnCl_2$, 1 mM DTT, and 10 μM ATP), and 1 μl [γ-^{32}P] ATP. The kinase buffer used for the autophosphorylation of FLS2 kinase domain did not contain $MnCl_2$, to avoid the precipitation of FLS2. The reaction was incubated at room temperature for one hour and then terminated by adding SDS-PAGE sample buffer. The gel was then stained, dried, and autoradiographed.

V.1.3 Results and discussion

V.1.3.1 Kinase domains of Arabidopsis RLKs are mostly insoluble or unstable

To prepare the FLS2 kinase domain sample suitable for structural studies, optimal conditions for expression were explored. A single colony of BL21 harboring recombinant FLS2 in the pET-15b vector was inoculated into a seed culture in a small amount of LB medium and was grown overnight at 37 ^{o}C. The overnight culture was then inoculated into 50 ml of LB medium, which was incubated at 37 ^{o}C with shaking, until $O.D._{600}$ reached 0.6-0.7. IPTG was added to a final concentration 200 μM. Induction was performed for 3, 6, 9, and 22 hours at 18 ^{o}C. Supernatants and pellets of samples were run on SDS-PAGE gels. The expression of HIS-FLS2 (Figure V-1a) was excellent. However, over 95% of the HIS-FLS2 expressed was insoluble. The same behavior was also observed for other RLK kinase domains. For example, 1 liter of bacteria culture only yields 500 μg of GST-BAK1 protein from the soluble fractions. Several different expression conditions, such as varying concentrations of IPTG during induction and expression at low temperatures, have also been explored. Most of the expressed BAK1 still existed in the inclusion bodies.

Alternative methods for purifying RLK kinase domains directly from soluble fractions have also been used. First, kinase activities may be lethal to bacteria. High yield bacterial expression of Abl and Src tyrosine kinases has been achieved by coexpression with the YopH phosphatase (*166*). The inactive GST-mFLS2 (1064 G->R) (*35*) with abolished autophosphorylation activity was transformed into BL21 DE3 strain. Similar expression trials were performed at 20 ^{o}C for 4, 6, 11, and 22 hours (Figure V-1b). Most (>85%) of the expressed GST-mFLS2 was insoluble. Second, GST tagged or MBP tagged BRI1

kinase domain was found to be soluble, and 1 liter of bacterial culture can yield 5-10 mg of protein, following the standard procedures. However, after removing the affinity tags, BRI1 disappeared in solution, probably because of its poor solubility (data not shown). Third, the absence of RLK kinase domains in the soluble fractions may be because plant RLKs are foreign to bacteria that lack necessary chaperones of RLKs. To address this potential problem, a vector containing GroEL/ES cDNA was co-transformed with HIS-FLS2 into BL21 DE3 strain. Expression trials showed that most of the FLS2 kinase domain was not present in the soluble fractions (data not shown).

V.1.3.2 High yields were achieved by preparing RLK kinase domains from insoluble fractions

Because purifying RLK kinase domains directly from the soluble fractions of bacteria cultures is very difficult, we tried to prepare RLK kinase domains (i.e. GST-BAK1 and HIS-FLS2) from insoluble pellets by denaturing and refolding, or by detergent (N-Lauroylsarcosine) treatment. These methods resulted in high yield of RLK kinase domains with sufficient purity. Protocols have been introduced in Material and Methods.

For GST-BAK1, 100 ml original bacteria culture yielded 6 mg of protein (Figure V-2, lane 2) by denaturing and refolding. However, the refolded GST tag failed to bind glutathione affinity resin. The GST tag could be removed in solution by incubating with 3C PreScission protease (Bio-rad) and separated BAK1 kinase domain can be purified by

other types of chromatographies (i.e. Ion exchange, Gel-filtration). The kinase activity of BAK1 has not been tested.

For HIS-FLS2, 15 mg of sample was obtained from 500 ml of bacteria culture by detergent (N-Lauroylsarcosine) treatment (Figure V-1a, lane 11). The detergent treated FLS2 showed ambiguous autophosphorylation activity (Data not shown).

V.1.3.3 FLS2 kinase domains from both insoluble and soluble fractions were unstructured

To investigate the structural quality of the detergent-treated ^{15}N HIS-FLS2, 2D ^{15}N-^{1}H TROSY spectra were collected at 22°C and 600 MHz Varian with a cryo-probe. Fewer than 100 peaks were observed for the 303-residue FLS2 kinase domain (Figure V-3a), and most of the peaks clustered in the region of random coil with chemical shifts ranging from 7.9 ppm to 8.5 ppm. This indicated that ^{15}N HIS-FLS2 purified from the insoluble fractions was unfolded. Interestingly, by expressing recombinant HIS-FLS2 in pET-15b at 30 ^{o}C for 4 hours, ^{15}N HIS-FLS2 with purity higher than 80% was obtained from the soluble fractions. The ^{15}N-^{1}H TROSY spectrum (Figure V-3b) showed that ^{15}N HIS-FLS2 purified from supernatant was also unstructured. Adding the kinase buffer to the NMR samples of HIS-FLS2 did not promote their folding according to TROSY spectra.

Our efforts to prepare a soluble and stable RLK kinase domain, including GST-BAK1, HIS-FLS2, HIS-mFLS2, and BRI1, from bacteria for structural determination were also

not successful. However, a few alternative methods can be used in the future. First, our data does suggest that RLK kinase domains are usually more stable when expressed and purified with fusion tags (i.e. GST, MBP). A small solubility-enhancement tag (SET) (*167*) and NusA fusion tag have been shown to increase protein stabilities and could be used. Crystallization trials could be conducted on RLK kinase domains fused with MBP or GST tag via a rigid linker. Second, yeast looks more capable of expressing properly folded eukaryotic proteins than *E. coli*. Most crystallographic structures of kinases have been determined by using Sf9 insect cell expression. Therefore, preparations of RLK kinase domains could be carried out in eukaryotic cells, although the expense will be higher. Third, preparations of RLK kinase domains with co-expressed GroES/EL chaperone or an appropriate phosphatase can be developed further.

V.2 An investigation of the KAPP binding site of BRI1

V.2.1 Introduction

Brassinolide insensitive I (BRI1) plays an essential role in Arabidopsis brassinosteroid signaling. The pathway of the BRI1 signal cascade has been well studied and introduced in Chapter I. Recent studies revealed that KAPP was a negative regulator of BRI1, and KI-FHA of KAPP interacted with BRI1 kinase domain *in vitro* and *in vivo*. The GST-BRI1 binding site of KI-FHA is very similar to the phosphopeptide interacting site (*23*). Preliminary experiments have been conducted to identify the surface of BRI1 kinase

domain that binds KI-FHA, and to measure distances between single spin labels linked to both BRI1 and KI-FHA for building a preliminary model of the BRI1-KAPP signaling complex.

V.2.2 Materials and Methods

V.2.2.1 Preparation of GST-BRI1

Recombinant GST-BRI1 (E816->L1196) in pGex-6p-1 vector was transformed into BL21 (Codonplus RIL) competent cells with kanamycin and chloraphenicol double resistances. A single colony was isolated and seeded into a small LB culture, followed by an overnight incubation at 37 ^{o}C. The culture was then inoculated into a larger volume of LB medium and incubated at 37 ^{o}C with shaking until $O.D._{600}$ reached 0.8. IPTG was added to a final concentration of 200 uM. After an overnight induction at 20°C, cells were pelleted and resuspended in PBS buffer (pH 7.5) supplemented with cocktail protease inhibitors. Bacteria cells were lysed with a chilled French Press Cell and centrifuged at 10,000 rpm for 20 minutes. The supernatant was dialyzed against PBS buffer (pH 7.5). The glutathione agarose was added to the supernatant, incubating by gentle shaking in 4 ^{o}C for one hour. The GST-BRI1 bound agarose was packed in columns and washed with >20X volumes of PBS buffer (pH 7.5). GST-BRI1 was eluted with the elution buffer (50 mM Hepes, 15 mM reduced glutathione, pH 7.4). All steps should be performed at 4 ^{o}C.

V.2.2.2 Cleavage of GST-BRI1 by KI-FHA mutants conjugated to FeBABE.

Single substitutions of KI-FHA, including H261C, V213C, M246C, K288C and K228C, were engineered by Huachun Wang (John C. Walker group, University of Missouri). The mutants of KI-FHA were conjugated with FeBABE, an EDTA-chelated iron atom linked to a sulfhydryl-reactive moiety or Fe (III) (s)-1-(p-Bromoacetamidobenzyl) ethylenediaminetetraacetic acid (Pierce). The modified KI-FHA mutants were then incubated with GST-BRI1, and the results were analyzed by SDS-PAGE. See ProFound Protein Interaction Mapping Kit (Pierce, product No. 32223) for details.

V.2.2.3 MTS spin labeling

DTT was added into the cysteine mutants of KI-FHA to a final concentration of 2 mM. KI-FHA mutants were concentrated to 1 mg/ml, with the final volume ~0.5 ml. The solvent was exchanged into phosphate buffer (pH 6.7) with 120 mM NaCl via dialysis or using Nap-10 gel-filtration columns (GE Healthcare). The MTS spin label (1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl)-methanethiosulfonate (Toronto Research Chemicals, Inc) dissolved in acetonitrile (100 mM) was added to 5-10x molar excess, with a MTS/protein volume-ratio 1:100. The mixture of the reaction was chilled on ice and incubated at 4°C overnight. Free MTS was removed by running the samples through the Nap-10 columns. The labeled KI-FHA mutants were concentrated to 100-200 uM. The sample homogeneity was tested by electrospray mass spectrometry.

V.2.3 Results and discussion

V.2.3.1 Far-western assay of the single cysteine mutants of KI-FHA

To define the KI-FHA interacting site of an RLK kinase domain, six single cysteine substitutions of KI-FHA (Figure V-4), including G189C, V213C, K228C, M246C, H261C, and K288C, were introduced to the periphery of its phospho-ligand binding site. The single cysteines enable introduction of either a chelated Fe (III) (FeBABE) for proteolysis of a nearby RLK site, or a nitroxide spin label for distance measurement to the future spin labels in an RLK by NMR or EPR. In order to avoid interference in KI-FHA's interaction with RLKs, the sites of mutagenesis were not designed within the ligand binding interface. GST tagged KI-FHA mutants were prepared by glutathione affinity chromatography (*101*). Five of the single mutations except G189C, had yields as good as the wild type.

To test the binding of the single cysteine mutants of KI-FHA to GST-BRI1, far-Western assays were performed as described in Materials and Methods of Chapter III. Purified KI-FHA mutants were run on SDS-PAGE. The gel was electroblotted to nitrocellulose membrane and incubated with GST-BRI1. Binding was detected by anti-GST antibody. Figure V-5a shows that V213C and M246C interacted well with GST-BRI1 *in vitro* and that H261C and K228C bound to an intermediate degree.

V.2.3.2 GST-BRI1 was digested by FeBABE-conjugated M246C into two fragments

To identify the KI-FHA interacting site of GST-BRI1, GST-BRI1 was incubated with FeBABE-conjugated M246C and V213C for 30 seconds, and for 10 minutes. Binding competent M246C was conjugated via disulfide with the FeBABE (Pierce) protein-cutting reagent. The results were analyzed by SDS-PAGE. Bands of apparent molecular weight of 50 kDa and 38 kDa appeared (Figure V-5b, lane 4 and 5). This suggests one or more sites of cleavage in the BRI1 kinase domain. Hypomobility of the protein bands may result from autophosphorylation of BRI1.

The next step is to define the KI-FHA binding region of the BRI1 kinase domain. GST-containing cleavage fragments can be identified by Western-blot using anti-GST antibody. N-terminal sequencing or mass spectrometry can be employed to determine the cleavage site.

V.2.3.3 Docking an RLK kinase domain to KI-FHA

The second purpose of constructing the signal cysteine mutants of KI-FHA is to enable distance measurements between spin labels by EPR or NMR and subsequently build a model of a KI-FHA/RLK signaling complex. An MTS nitroxide spin label can be attached to KI-FHA cysteine mutants as described in Materials and Methods. The labeling procedure can also be applied to KI-FHA binding RLK kinase domains. Recent studies showed that the T546A mutant of BAK1 only had 38% of the extent of KI-FHA

binding of the wild type. T546 is located near the C-terminus of the BAK1 kinase domain (*23*). Therefore, perhaps cysteine substitutions should be introduced to the corresponding region at the C-lobe of an RLK kinase domain.

The nitroxide spin label has been heavily used by the Wayne Hubbell group for EPR-based structural studies (*168*). By using one nitroxide per protein, EPR can be used to qualitatively estimate distances between two spin labels from 8 to 20 or 25 Å. Moreover, NMR can further be used to measure distances of numerous spin labels to a single nitroxide spin label (*169*). After obtaining the distance information, a crude rigid body docking model of a KI-FHA/RLK complex can be built, using software such as HADDOCK (*170, 171*). However, before the spin labeling strategy can proceed, existing cysteines in KI-FHA, RLK kinase domains and affinity tags should be mutated out.

V.3 Folding and activation of AvrRpt2 in pathogen defense of Arabidopsis

V.3.1 Introduction

Innate immunity in higher plants has culminated in a sophisticated surveillance system capable of recognizing bacterial effector proteins directly secreted into host cells. An indirect mode of pathogen recognition of plants is illustrated in that a plant resistance

protein recognizes the alteration of a second plant protein activated by a pathogen effector protein. AvrRpt2, a cysteine protease, is such an effector protein from *Pseudomonas syringae* (*172*). AvrRpt2 is delivered to plant cells during infection as an inactive cysteine protease and is activated by eukaryotic cyclophilin ROC1. ROC1 is a proline isomerase that promotes AvrRpt2's self-removing its N-terminal 7 kDa fragment. The activated 22 kDa AvrRpt2 is able to cleave the plant protein RIN4 (*173*). The RIN4 protein is then removed, and the suppressed RPS2-mediated pathogen-resistance pathway is activated (*173, 174*). The early goal of our preliminary study was to use NMR and other biochemical techniques to determine the solution structure of the activated 22 kDa AvrRpt2, to characterize the folding process of AvrRpt2, and to investigate the mechanism of AvrRpt2's activation by ROC1.

V.3.2 Materials and Methods

V.3.2.1 Preparation of AvrRpt2 and GST-ROC1

Samples of AvrRpt2 and GST-ROC1 were expressed and purified as previously described by Gitta Coaker (Brian Staskawicz group, Berkeley) (*173*). GST-ROC1 was expressed in BL21 (DE3) pLysS cells and was isolated under native conditions. Clones were grown in LB medium containing 50 μg/ml carbenicillin and 35 μg/ml chloramphenicol until $O.D._{600}$ reaches 0.4. Protein expression was induced for 4 hours at 37 ^{o}C with 0.5 mM IPTG. Cells were lysed in a buffer containing 140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1% Triton X-100, pH 8.0. Recombinant GST-ROC1 was affinity

purified in batch format by glutathione sepharose affinity chromatography (GE Healthcare) and eluted from the glutathione sepharose with 10 mM reduced glutathione in 50 mM Tris-HCl, pH 8.0. For preparation of ^{15}N labeled protein samples, M9 media supplemented with 15% (v/v) ^{15}N Celtone were used.

V.3.2.2 NMR TROSY experiments

2D ^{15}N-^{1}H TROSY spectra were collected at 25°C and 600 MHz on free ^{15}N HIS-AvrRpt2-FLAG (0.1mM) and the ^{15}N HIS-AvrRpt2-FLAG incubated with GST-ROC1 at a AvrRpt2/ROC1 molar ratio 45:1 for over 9 hours. The acquisition times were 0.5 hour for the free AvrRpt2 and 8 hours for the AvrRpt2 mixed with ROC1. A Varian Inova 600 MHz spectrometer (Univ. of Missouri) was used with either a high sensitivity 5 mm ^{1}H {^{13}C/^{15}N} cryogenic probe with shielded z gradient coil. NMR spectra were processed using NMRPipe 2.3 (*107*).

V.3.2.3 Gel-filtration assay

Protein samples included (i) 30 ul purified ^{15}N labeled AvrRpt2 (22 kDa) and GST-ROC1 (47 kDa) complex in Tris buffer (20 mM Tris, 150 mM NaCl, 2 mM DTT, 2 mM Cysteine, pH 7.5), 10 mg/ml. (ii) 80 ul processed AvrRpt2 (22 kDa) in phosphate buffer (20 mM sodium phosphoate, 150 mM NaCl, 2 mM DTT, pH 6.8), 1.1 mg/ml and (iii) 60 ul GST-ROC1 in phosphate buffer (20 mM sodium phosphoate, 175 mM NaCl, 1 mM 2-

mercaptoethanol, pH 7.0), 2.5 mg/ml. Samples were loaded onto a SEC-300 HR gel filtration column equilibrated and eluted with PBS buffer (pH 7.4). The SEC-300 column has a void volume around 6 ml, and a bed volume around 12 ml. The absorbance of the eluted fractions at 280 nm (A_{280}) was recorded and plotted against the fraction number. Samples were fractionated at 0.5 ml/minute at room temperature. Fractions were then collected at 0.5 ml per fraction, concentrated and resolved by SDS-PAGE gels.

V.3.3 Results and discussion

V.3.3.1 AvrRpt2 is folded and activated in the presence of GST-ROC1

To test whether ROC1 is required for the folding and activation of AvrRpt2, a series of 2D ^{15}N-^{1}H TROSY spectra were collected at 25°C and 600 MHz on ^{15}N labeled HIS-AvrRpt2-FLAG (~277 residues) incubated with GST-ROC1 for 0, 1, 2, 3, 4, 5, 7, 9 and 20 hours, at a molar ratio of 45:1. TROSY of free ^{15}N HIS-AvrRpt2-FLAG has only around 50 backbone amide peaks (Figure V-6, upper panel). Most peaks are located in the region of random coil chemical shifts from 7.9 ppm to 8.5 ppm. This suggests that free ^{15}N HIS-AvrRpt2-FLAG is unfolded, since an additional 150 backbone amide peaks are not observed. They may be broadened away by chemical exchange between unfolded substrates. In the presence of GST-ROC1, more and more folded peaks progressively appeared over time in the TROSY spectra. After incubating for 9 hours, around 150 peaks are observed (Figure V-6, lower panel). This suggests that AvrRpt2 was folded and activated in the presence of GST-ROC1. After 9 hours, no further increase in number of

peaks was seen. Biochemical studies verified that the ability of AvrRpt2 to cleave its substrate RIN4 was dependent on the presence of ROC1 (*173*). AvrRpt2 is activated by cyclophilin ROC1 via its PPIase activity (*175*).

In the activated mixture of AvrRpt2 and ROC1, peaks at the central random coil region are strongest (Figure V-6), while the folded peaks on the downfield and upfield flanks of the spectrum are weaker. This weakness could be explained by the folded portion either being dilute or high in molecular weight. A high molecular weight would occur if the AvrRpt2 and GST-ROC1 form a stable complex. This possibility has been tested by gel filtration chromatography.

V.3.3.2 AvrRpt2 and ROC1 form a stable protein complex

NMR spectroscopy clearly demonstrates ROC1-induced structure formation of AvrRpt2. We employed analytical gel filtration (with an SEC-300 HR column) to help determine if AvrRpt2 and ROC1 can form a stable protein complex, or if ROC1 plays a transient catalytic role in folding AvrRpt2. Processed and purified AvrRpt2 runs in gel-filtration column with high mobility and low mobility peaks (middle panel of Figure V-7). The NMR spectrum of Figure V-6a suggests free AvrRpt2 peaks to be unfolded. The fast-migrating AvrRpt2 could be extended or aggregated while the slow-migrating “25 kDa” AvrRpt2 could be a compact, unfolded form. Free GST-ROC1 runs with an apparent molecular weight of 85 kDa (lower panel of Figure V-7), suggesting it to be a dimer as GST is known to dimerize. Analytical gel filtration of a mixture of the previously

processed AvrRpt2 with GST-ROC1 reveals a clear alteration in mobility. The chromatogram of the mixture is dominated by a new peak, containing both proteins of 22 and 47 kDa by SDS-PAGE, migrating together with an apparent molecular weight of roughly 140 kDa. This apparent molecular weight suggests formation of a stable complex of the GST-ROC1 dimer with AvrRpt2. The complexes could generally contain two 22 kDa chains of AvrRpt2, but the peak shape suggests other stoichiometries could also be present. While AvrRpt2 and GST-ROC1 clearly form a stable complex *in vitro* (Figure V-7), small percentages of GST-ROC1 are sufficient to activate AvrRpt2, suggesting the proteins must also dissociate. Together the NMR spectra and gel filtration chromatography suggest that AvrRpt2 may only be structured and active when ROC1 or another cyclophilin is bound. This may account for the weakness of the NMR peaks of folded AvrRpt2 in the presence of a few percent of GST-ROC1 (Figure V-6b).

We investigated cyclophilin-induced folding of AvrRpt2 by NMR. NMR spectroscopy readily monitors the degree of protein folding in solution. The detection of peaks of folded AvrRpt2 in the NMR TROSY spectra demonstrates ROC1-induced structure formation of AvrRpt2. The NMR spectra and gel filtration analysis together suggest that AvrRpt2 may only be structured and active when ROC1 or another cyclophilin is bound in order to maintain one or more proline residues in the appropriate isomerization state.

It will be of interest to know the 3D structure of AvrRpt2. Despite the vast number of phytopathogenic bacterial effectors that have been cloned, only a few have been

crystallized to determine their structures. By determining the structure by NMR in solution, mechanistic insights into their mode of action may be elucidated. The resulting structure could then be compared with crystal structures and uncover conservation to other cloned bacterial effectors. TROSY experiments at different pH showed that the ^{15}N labeled 22 kDa AvrRpt2 largely unfolded without GST-ROC1 and that it could not be restructured by titrating ROC1 into the NMR samples (data not shown). This suggests that ROC1 may be required to maintain the folding of the activated 22 kDa AvrRpt2 *in vitro*. Therefore, the first choice for the sample preparation of AvrRpt2 for structure determination is to incubate the ^{15}N labeled full-length HIS-AvrRpt2 with GST-ROC1 *in vitro* and remove the HIS tagged 7 kDa N-terminal fragment by incubating with nickel agarose column. The GST tag on ROC1 should also be removed by incubating with 3C PreScission protease (GE Healthcare). The second strategy is to coexpress HIS-AvrRpt2 and ROC1 in bacteria and purify both proteins via affinity chromatography. The third way is to study the full-length HIS-AvrRpt2 (C122A), an AvrRpt2 mutant that can be activated by ROC1 but will retain the 7 kDa N-terminal fragment, instead of the wild type. Furthermore, crystal structures of cysteine proteases (i.e. Staphopain, PDB entry number: 1CV8) are available. Structure prediction via homology modeling can be performed on AvrRpt2 to aid NMR structure determination.

V.4 Acknowledgement

I would like to thank Dr. Huachun Wang (John Walker group, University of Missouri) and Dr. Xuelu Wang for providing me with the constructs of GST-BAK1 and GST-BRI1.

I thank Dr. Scott Peck for giving me the constructs of FLS2 and mFLS2. I am grateful to Dr. Jennine Crane (Linda Randall group) for sharing protocols and information about MTS spin labeling with me. I would also like to thank Dr. Xiangyang Liang for discussing the preparations of GST-BRI1 with me. Dr. Gitta Coaker (Brian Staskawicz group, Berkeley) and Prof. Steven Van Doren contributed many valuable suggestions in the study of AvrRpt2 and ROC1. I learned a lot from them. Finally, I offer my special thanks to Dr. Michael Riley (University of Missouri) for teaching me run SEC-300 HR gel filtration column in his FPLC system.

Figure V-1: Expression and solubilization of the FLS2 kinase domain in *E. coli*. a) HIS-FLS2 in pET-15b was transformed into BL21 (DE3) strain, and cells were induced with 0.2 mM IPTG at 18 ^{o}C for 3 (lane 1, 2), 6 (lane 3, 4), 9 (lane 5, 6) and 22 hours (lane 7, 8). The gel bands of the detergent (N-Lauroylsarcosine) solubilized and refolded FLS2 were shown in lane 11 and lane 12 respectively. b) GST tagged kinase domain of the inactive FLS2 (GST-mFLS2) in pGex-6p-1 was transformed into BL21 (DE3) strain, and cells were induced with 0.2 mM IPTG at 20 ^{o}C for 4 (lane 1, 2), 6 (lane 3, 4), 11 (lane 7, 8) and 22 hours (lane 9, 10). Cell fractions were clarified by centrifugation. Supernatants (labeled with "s") and pellets (labeled with "p") were run on SDS-PAGE gels. Gels were then stained and destained.

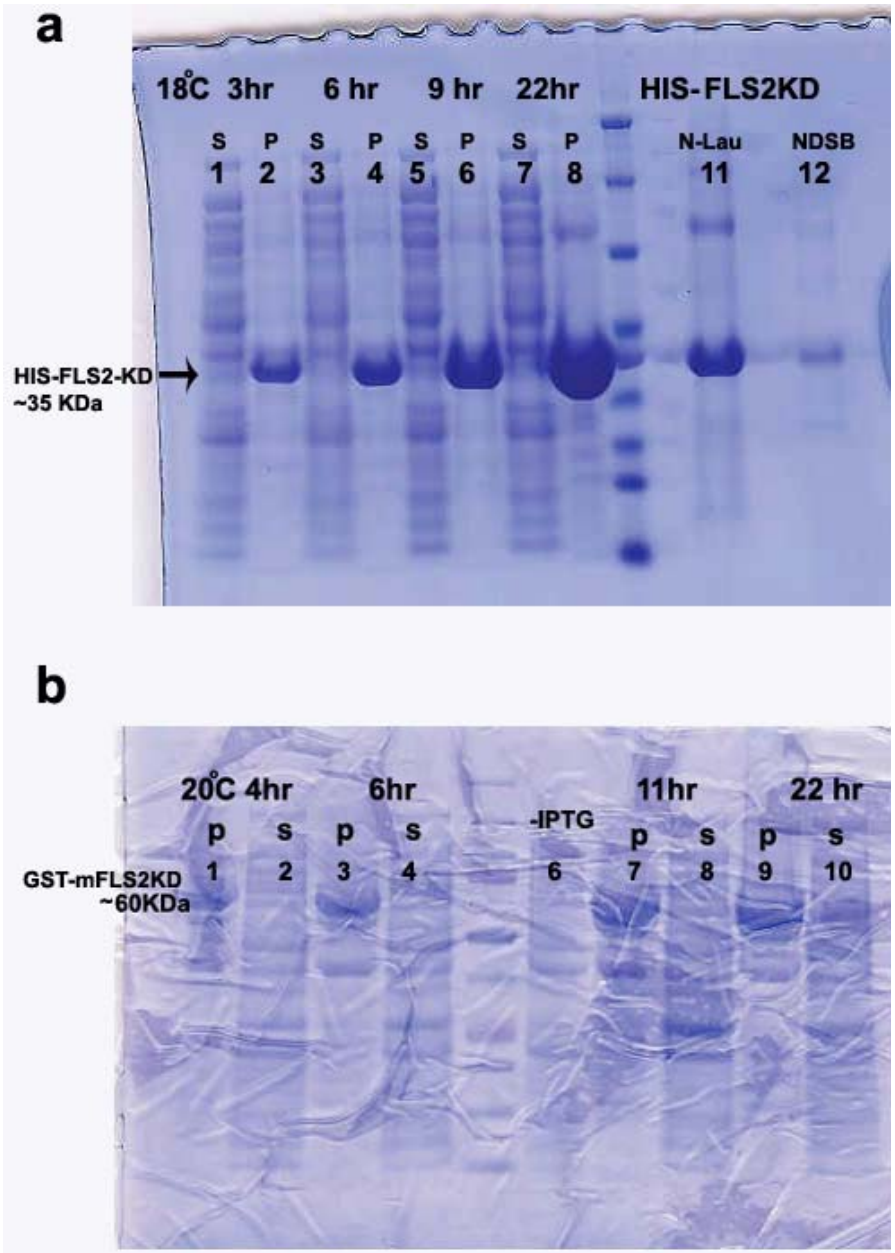

a
18°C 3hr
6 hr
9 hr
22hr
HIS-FLS2KD
S P S P S P S P
N-Lau
NDSB
1 2 3 4 5 6 7 8
11
12
HIS-FLS2-KD
~35 KDa
b
20°C 4hr
6hr
-IPTG
11hr
22 hr
p s p s
p s p s
1 2 3 4
6 7 8 9 10
GST-mFLS2KD
~60KDa

Figure V-2: The refolded BAK1 kinase domain. Lane 1, precision plus protein standards (Bio-rad); Lane 2, refolded GST-BAK1; Lane 3, refolded GST-BAK1 after in solution cleavage by 3C PreScission protease (Bio-rad).

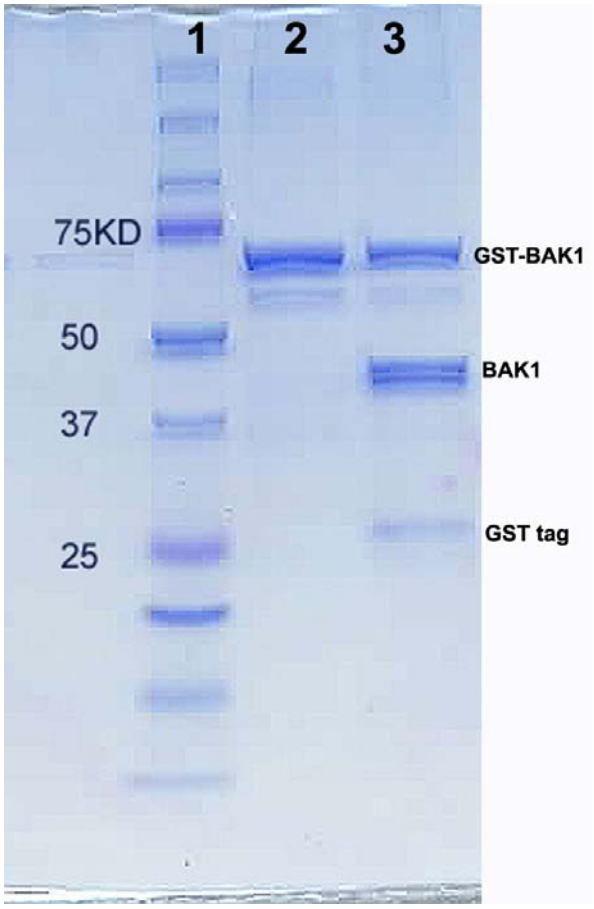


Figure V-3: TROSY spectra of ^{15}N FLS2 kinase domain. a) HIS tagged ^{15}N FLS2 kinase domain (FLS2-KD) was solubilized with N-Lauroylsarcosine from the cell pellet, dialyzed into PBS buffer (pH 7.5), and concentrated to roughly 7-10 mg/ml (~0.2-0.4 mM). TROSY experiment was performed on Varian 600 MHz with cryo probe. Data acquisition time was around 25 minutes. b) TROSY experiment was performed on HIS tagged ^{15}N FLS2 kinase domain (FLS2-KD) from the soluble fractions (~0.1 mM).

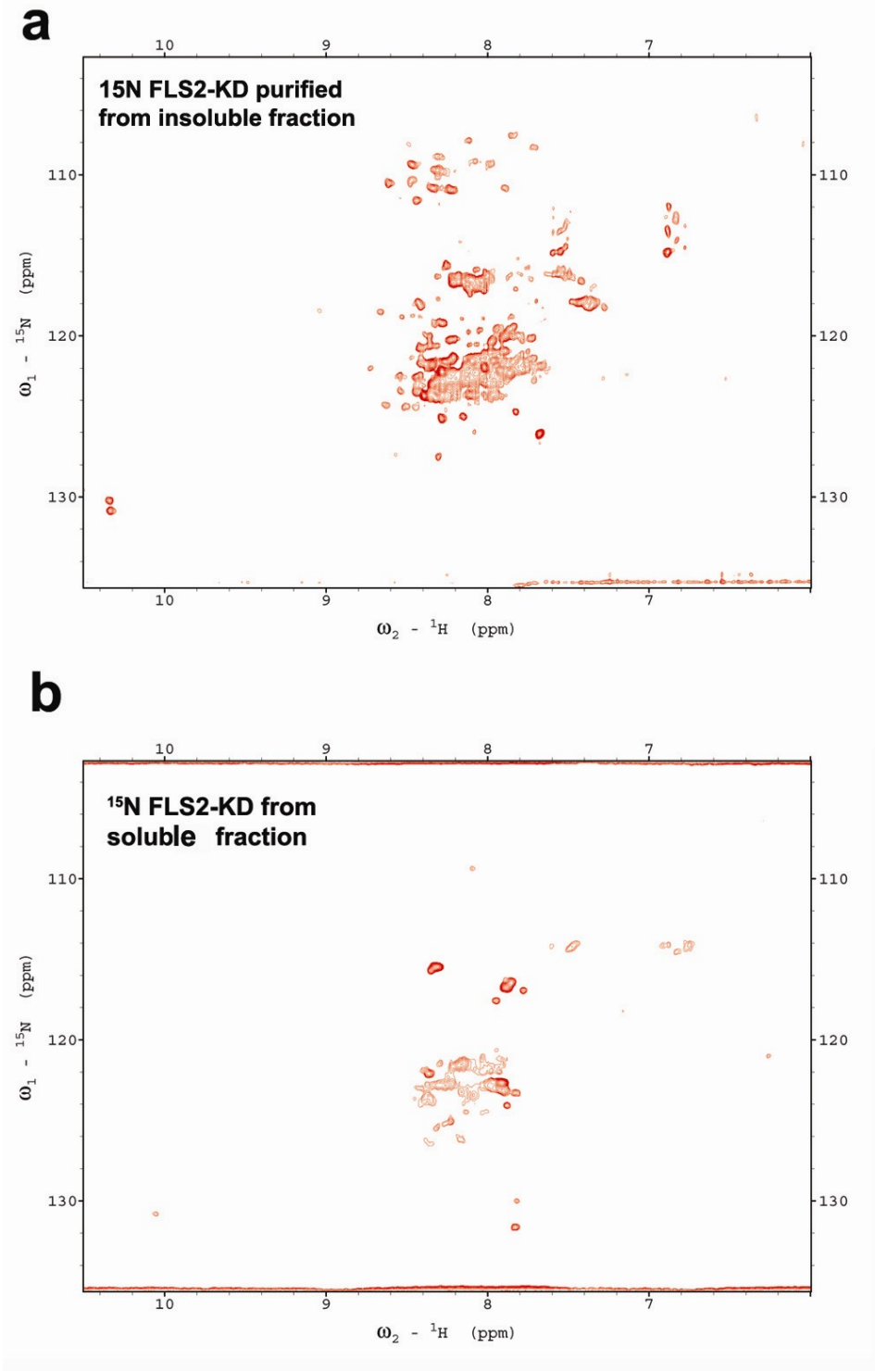

a
15N FLS2-KD purified
from insoluble fraction
ω1 - 15N (ppm)
ω2 - 1H (ppm)
b
15N FLS2-KD from
soluble fraction
ω1 - 15N (ppm)
ω2 - 1H (ppm)

Figure V-4: Locations of the single cysteine substitutions on KI-FHA. KI-FHA residues mutated to Cys (one per construct) are red. They surround the residues in the ligand binding site, colored green.

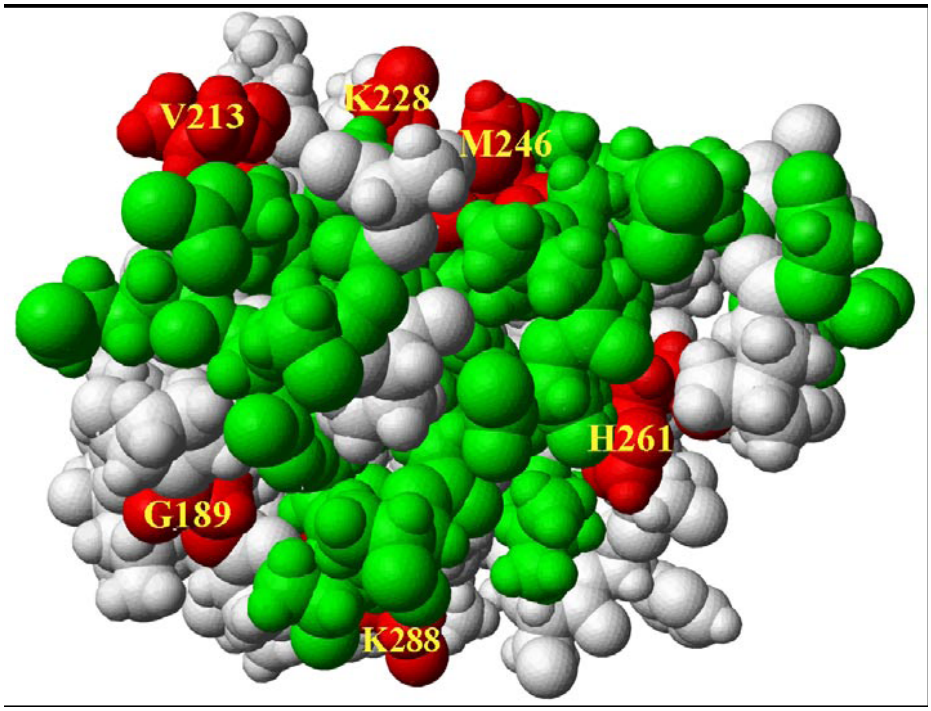


Figure V-5: The single cysteine mutants of KI-FHA interact with GST-BRI1 *in vitro* and GST-BRI1 was cleaved by FeBABE-conjugated KI-FHA. a) Binding of KI-FHA with single cysteine mutations to GST-BRI1 by Far-western assay. Purified KI-FHA mutant H261C (lane 1), V213C (lane 2), M246C (lane 3), K288C (lane 4), and K228C (lane 5) were run on an SDS-PAGE gel. The gel was transferred to nitrocellulose membrane and incubated with GST-BRI1. Binding was detected by anti-GST antibodies. b) Cleavage of GST-BRI1 by KIFHA conjugated to FeBABE. Arrows mark cleavage products. Lane (1), GST-BRI1 alone; (2) GST-BRI1 + M246C without FeBABE; (4) GST-BRI1 + M246C-FeBABE, with 30 sec of activation; (5) GST-BRI1 + M246C-FeBABE, with 10 min of activation.

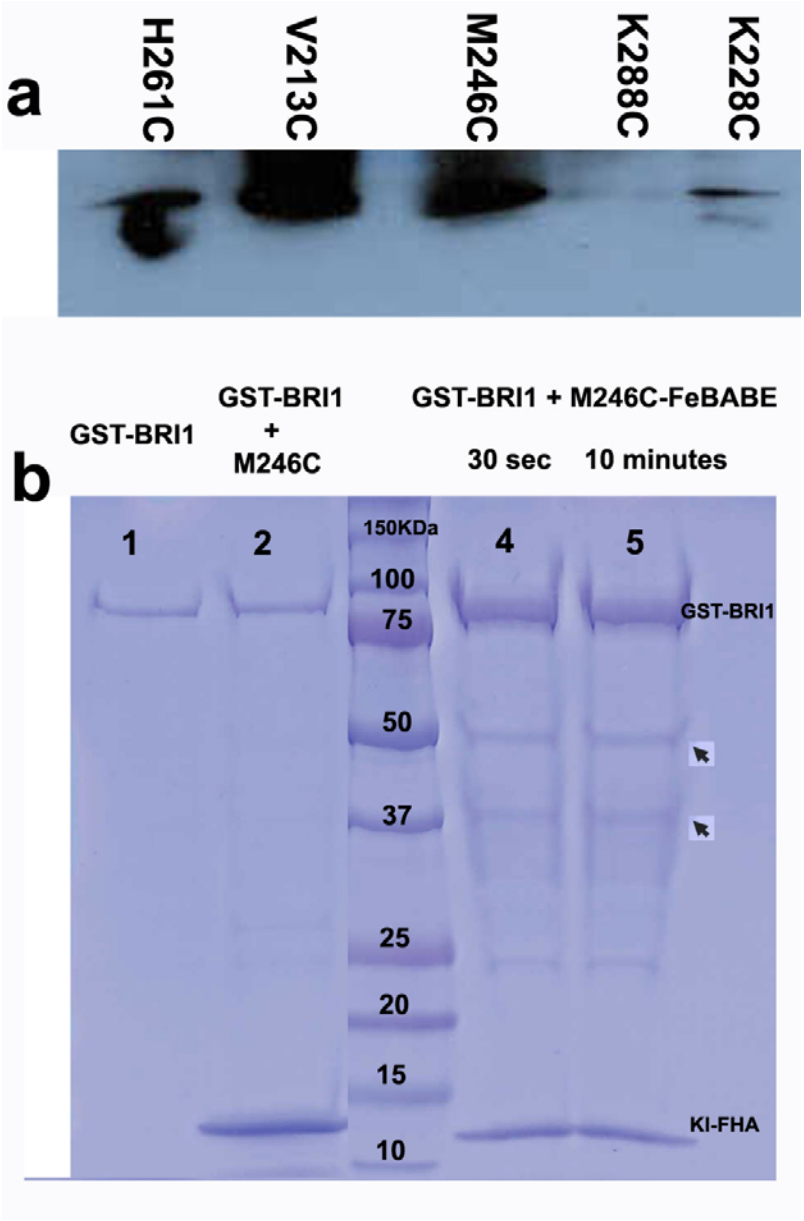

a
H261C
V213C
M246C
K288C
K228C
b
GST-BRI1
GST-BRI1
+
M246C
GST-BRI1 + M246C-FeBABE
30 sec
10 minutes
1
2
150KDa
4
5
100
75
50
37
25
20
15
10
GST-BRI1
KI-FHA

Figure V-6: TROSY NMR spectra of AvrRpt2 for A) free ^{15}N His:AvrRpt2:Flag and B) a 45:1 ratio of ^{15}N His:AvrRpt2:Flag and GST-ROC1 after 9 hours of incubation (*176*).

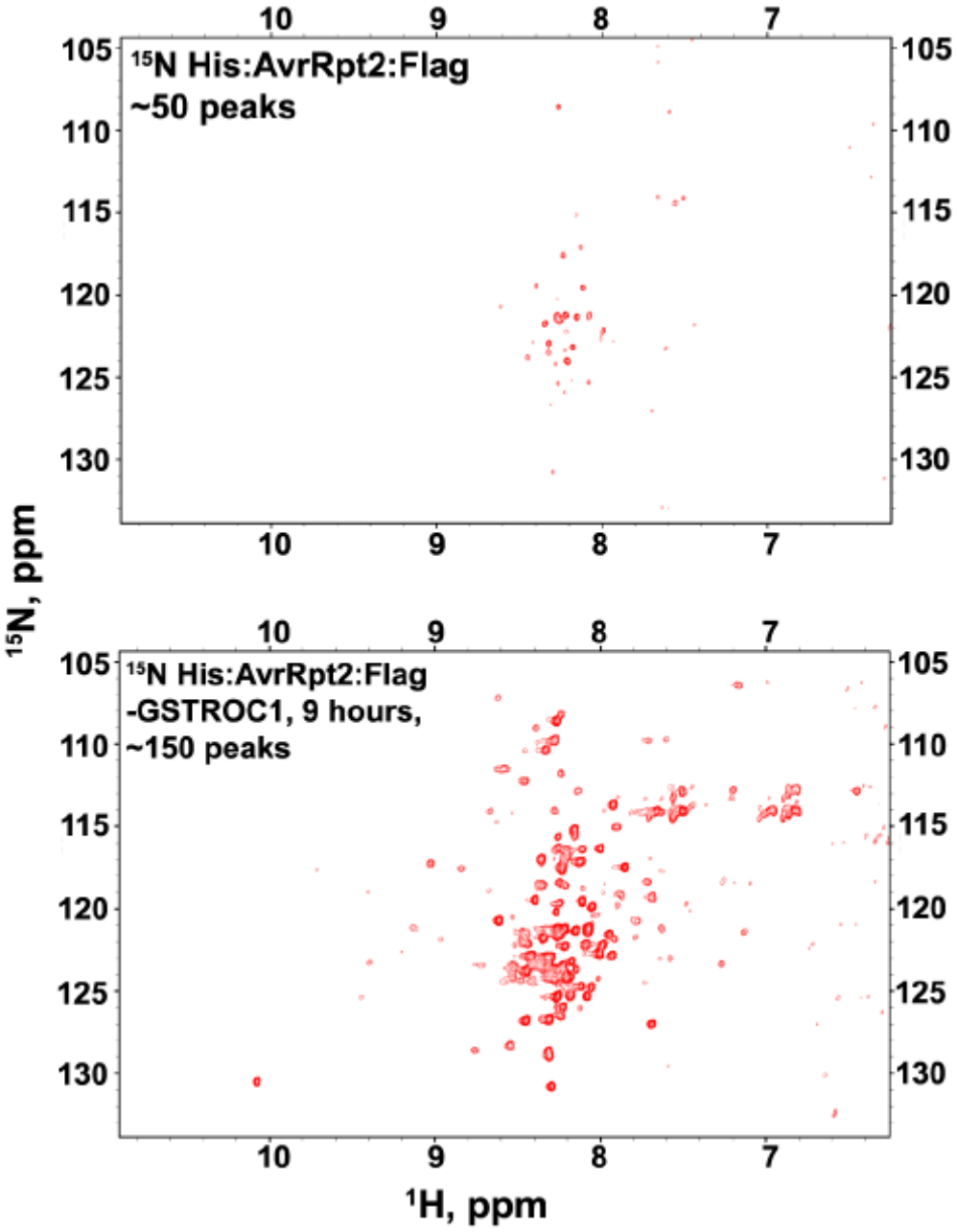
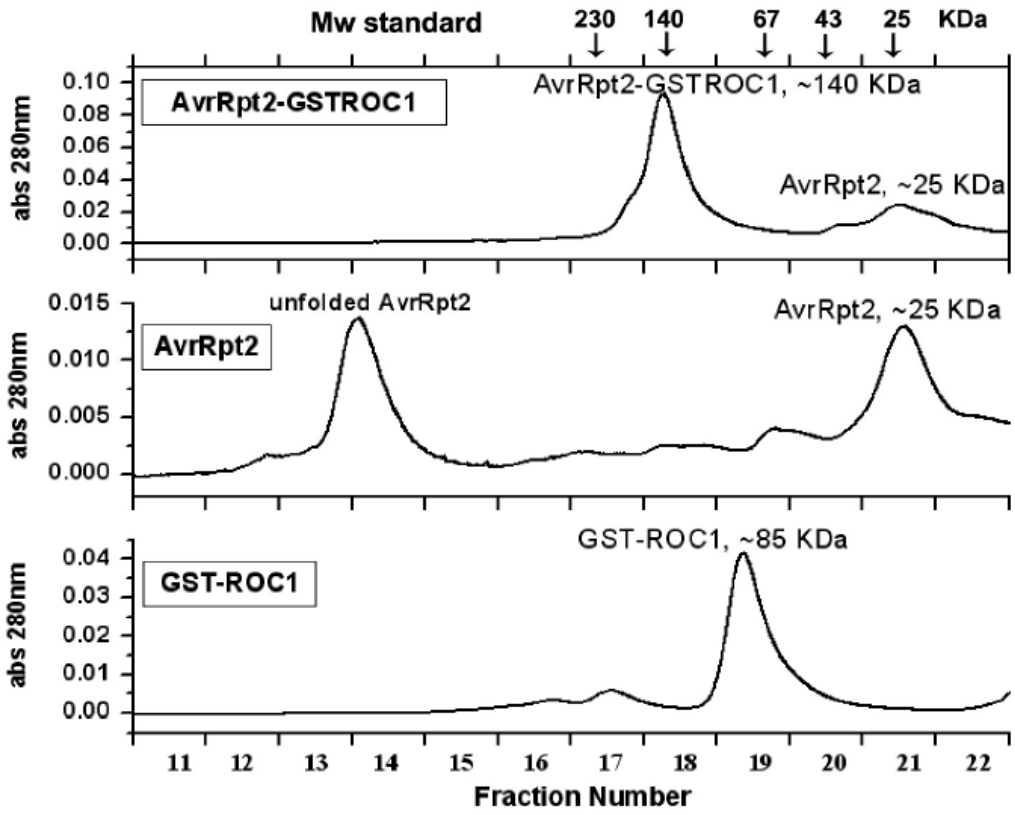


Figure V-7: Evidence of the association of ^{15}N labeled AvrRpt2 and GST-ROC1 resolved by gel filtration chromatography. ^{15}N labeled AvrRpt2 (22 kDa) complexed with GST-ROC1 (47 kDa) (Upper panel), the processed AvrRpt2 (22 kDa) (Middle panel) and GST-ROC1 (Lower panel) were loaded onto a SEC-300 HR gel filtration column equilibrated and eluted with PBS buffer (pH 7.4). The absorbance of the eluted fractions at 280 nm (Abs 280nm) was plotted as a function of the fraction number. The contents of each eluted peaks were labeled. (*175*).



REFERENCES

(1) Holder, N., and Klein, R. (1999) Eph receptors and ephrins: effectors of morphogenesis. *Development 126*, 2033-44.

(2) Becraft, P. W. (1998) Receptor kinases in plant development. *Trends Plant Sci.*, 384-88.

(3) Becraft, P. W. (2002) Receptor kinase signaling in plant development. *Annu. Rev. Cell. Dev. Biol. 18*, 163-92.

(4) Plowman, G. D., Sudarsanam, S., Bingham, J., Whyte, D., and Hunter, T. (1999) The protein kinases of Caenorhabditis elegans: a model for signal transduction in multicellular organisms. *Proc Natl Acad Sci U S A 96*, 13603-10.

(5) Zhu, H., Klemic, J. F., Chang, S., Bertone, P., Casamayor, A., Klemic, K. G., Smith, D., Gerstein, M., Reed, M. A., and Snyder, M. (2000) Analysis of yeast protein kinases using protein chips. *Nat Genet 26*, 283-9.

(6) Smith, C. M. (1999) The protein kinase resource and other bioinformation resources. *Prog Biophys Mol Biol 71*, 525-33.

(7) Shiu, S. H., and Bleecker, A. B. (2001) Receptor-like kinases from Arabidopsis form a monophyletic gene family related to animal receptor kinases. *Proc Natl Acad Sci U S A 98*, 10763-8.

(8) Johnson, K. L., and Ingram, G. C. (2005) Sending the right signals: regulating receptor kinase activity. *Curr Opin Plant Biol 8*, 648-56.

(9) Mu, J. H., Lee, H. S., and Kao, T. H. (1994) Characterization of a pollen-expressed receptor-like kinase gene of Petunia inflata and the activity of its encoded kinase. *Plant Cell 6*, 709-21.

(10) Shah, K., Vervoort, J., and de Vries, S. C. (2001) Role of threonines in the Arabidopsis thaliana somatic embryogenesis receptor kinase 1 activation loop in phosphorylation. *Journal of Biological Chemistry 276*, 41263-9.

(11) Bassett, C. L., Nickerson, M. L., Cohen, R. A., and Rajeevan, M. S. (2000) Alternative transcript initiation and novel post-transcriptional processing of a leucine-rich repeat receptor-like protein kinase gene that responds to short-day photoperiodic floral induction in morning glory (Ipomoea nil).[erratum appears in Plant Mol Biol 2001 Aug;46(6):761-2]. *Plant Molecular Biology 43*, 43-58.

(12) Shiu, S. H., and Bleecker, A. B. (2003) Expansion of the receptor-like kinase/Pelle gene family and receptor-like proteins in Arabidopsis. *Plant Physiol 132*, 530-43.

(13) Shiu, S. H., and Bleecker, A. B. (2001) Plant receptor-like kinase gene family: diversity, function, and signaling. *Sci STKE 2001*, RE22.

(14) Kobe, B., and Deisenhofer, J. (1994) The leucine-rich repeat: a versatile binding motif. *Trends Biochem Sci 19*, 415-21.

(15) Walker, J. C., and Zhang, R. (1990) Relationship of a putative receptor protein kinase from maize to the S-locus glycoproteins of Brassica. *Nature 345*, 743-6.

(16) Dievart, A., and Clark, S. E. (2004) LRR-containing receptors regulating plant development and defense. *Development 131*, 251-61.

(17) Li, J., Lee, G. I., Van Doren, S. R., and Walker, J. C. (2000) Commentary - The FHA domain mediates phosphoprotein interactions. *J Cell Sci. 113*, 4143-4149.

(18) Li, J., and Chory, J. (1997) A putative leucine-rich repeat receptor kinase involved in brassinosteroid signal transduction. *Cell 90*, 929-38.
(19) Friedrichsen, D. M., Joazeiro, C. A., Li, J., Hunter, T., and Chory, J. (2000) Brassinosteroid-insensitive-1 is a ubiquitously expressed leucine-rich repeat receptor serine/threonine kinase. *Plant Physiology 123*, 1247-56.
(20) Li, J., Wen, J. Q., Lease, K. A., Doke, J. T., Tax, F. E., and Walker, J. C. (2002) BAK1, an Arabidopsis LRR receptor-like protein kinase, interacts with BRI1 and modulates brassinosteroid signaling. *Cell 110*, 213-222.
(21) Wang, Z. Y., Wang, Q., Chong, K., Wang, F., Wang, L., Bai, M., and Jia, C. (2006) The brassinosteroid signal transduction pathway. *Cell Res 16*, 427-34.
(22) Li, J., and Nam, K. H. (2002) Regulation of brassinosteroid signaling by a GSK3/SHAGGY-like kinase. *Science 295*, 1299-301.
(23) Ding, Z., Wang, H., Liang, X., Morris, E. R., Pandit, S., Skolnick, J., Walker, J. C., and Van Doren, S. R. (2007) Phosphoprotein and Phosphopeptide Interactions with the FHA Domain from Arabidopsis Kinase-Associated Protein Phosphatase. *Biochemistry*, submitted.
(24) Clark, S. E., Williams, R. W., and Meyerowitz, E. M. (1997) The *CLAVATA1* gene encodes a putative receptor kinase that controls shoot and floral meristem size in Arabidopsis. *Cell 89*, 575-585.
(25) Schoof, H., Lenhard, M., Haecker, A., Mayer, K. F., Jurgens, G., and Laux, T. (2000) The stem cell population of Arabidopsis shoot meristems in maintained by a regulatory loop between the CLAVATA and WUSCHEL genes. *Cell 100*, 635-44.
(26) Jeong, S., Trotochaud, A. E., and Clark, S. E. (1999) The Arabidopsis CLAVATA2 gene encodes a receptor-like protein required for the stability of the CLAVATA1 receptor-like kinase. *Plant Cell 11*, 1925-34.
(27) Fletcher, J. C., Brand, U., Running, M. P., Simon, R., and Meyerowitz, E. M. (1999) Signaling of cell fate decisions by CLAVATA3 in Arabidopsis shoot meristems. *Science 283*, 1911-4.
(28) Trotochaud, A. E., Hao, T., Wu, G., Yang, Z., and Clark, S. E. (1999) The CLAVATA1 Receptor-like kinase requires CLAVATA3 for its assembly into a signaling Complex that includes KAPP and a Rho -related protein. *Plant Cell 11*, 393-405.
(29) Laux, T., Mayer, K. F., Berger, J., and Jurgens, G. (1996) The WUSCHEL gene is required for shoot and floral meristem integrity in Arabidopsis. *Development 122*, 87-96.
(30) Mayer, K. F., Schoof, H., Haecker, A., Lenhard, M., Jurgens, G., and Laux, T. (1998) Role of WUSCHEL in regulating stem cell fate in the Arabidopsis shoot meristem. *Cell 95*, 805-15.
(31) Braun, D. M., Stone, J. M., and Walker, J. C. (1997) Interaction of the maize and Arabidopsis kinase interaction doamins with a subset of receptor-like protein kinases: implication for transmembrane signaling in plants. *Plant J. 12*, 83-95.
(32) Li, J., Smith, G. P., and Walker, J. C. (1999) Kinase interaction domain of kinase-associated protein phosphatase, a phosphoprotein-binding domain. *Proc. Natl. Acad. Sci. USA 96*, 7821-7826.

(33) Yu, L. P., Simon, E. J., Trotochaud, A. E., and Clark, S. E. (2000) POLTERGEIST functions to regulate meristem development downstream of the CLAVATA loci. *Development 127*, 1661-70.

(34) Song, S. K., Lee, M. M., and Clark, S. E. . (2006) POL and PLL1 phosphatases are CLAVATA1 signaling
intermediates required for Arabidopsis shoot and floral stem cells. . *Development* 4691-4698.

(35) Gomez-Gomez, L., and Boller, T. (2000) FLS2: an LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in Arabidopsis. *Mol Cell 5*, 1003-11.

(36) Asai, T., Tena, G., Plotnikova, J., Willmann, M. R., Chiu, W. L., Gomez-Gomez, L., Boller, T., Ausubel, F. M., and Sheen, J. (2002) MAP kinase signalling cascade in Arabidopsis innate immunity. *Nature 415*, 977-83.

(37) Gomez-Gomez, L., Felix, G., and Boller, T. (1999) A single locus determines sensitivity to bacterial flagellin in Arabidopsis thaliana. *Plant Journal 18*, 277-84.

(38) Stone, J. M., Colinge, M. A., Smith, R. D., Horn, M. A., and Walker, J. C. (1994) Interaction of a protein phosphatase with an *Arabidopsis* Serine-Threonine Receptor Kinase. *Science 266*, 793-795.

(39) Kerk, D., Bulgrien, J., Smith, D. W., Barsam, B., Veretnik, S., and Gribskov, M. (2002) The complement of protein phosphatase catalytic subunits encoded in the genome of Arabidopsis. *Plant Physiol 129*, 908-25.

(40) Park, A. R., Cho, S. K., Yun, U. J., Jin, M. Y., Lee, S. H., Sachetto-Martins, G., and Park, O. K. (2001) Interaction of the Arabidopsis receptor protein kinase Wak1 with a glycine-rich protein, AtGRP-3. *J. Biol. Chem. 276*, 26688-26693.

(41) Gómez-Gómez, L., Bauer, Z., and Boller, T. (2001) Both the extracellular leucine-rich repeat domain and the kinase activity of FLS2 are required for flagellin binding and signaling in Arabidopsis. *Plant Cell 13*, 1155-1163.

(42) Shah, K., Russinova, E., Gadella, T. W., Jr., Willemse, J., and De Vries, S. C. (2002) The Arabidopsis kinase-associated protein phosphatase controls internalization of the somatic embryogenesis receptor kinase 1. *Genes Dev 16*, 1707-20.

(43) Stone, J. M., Trotochaud, A. E., Walker, J. C., and Clark, S. E. (1998) Control of meristem development by CLAVATA1 receptor kinase and kinase-associated protein phosphatase interaction. *Plant Physiol. 117*, 1217-1225.

(44) Williams, R. W., Wilson, J. M., and Meyerowitz, E. M. (1997) A possible role for kinase-associated protein phosphatase in the *Arabidopsis* CLAVATA1 signaling pathway. *Proc. Natl. Acad. Sci. USA 94*, 10467-10472.

(45) Hofman, K., and Bucher, P. (1995) The FHA Domain: A Putative Nuclear Signalling Domain Found in Protein Kinases and Transcription Factors. *Trends Biochem. Sci. 20*, 347-349.

(46) Li, J., Smith, G. P., and Walker, J. C. (1999) Kinase interation domain of kinase-associated protein phosphatase, a phosphoptotein-binding domain. *Proc. Natl. Acad. Sci. USA 96*, 7821-7826.

(47) Durocher, D., and Jackson, S. P. (2001) DNA-PK, ATM and ATR as sensors of DNA damage: variations on a theme? *Curr Opin Cell Biol 13*, 225-31.

(48) Liao, H., Yuan, C., Su, M. I., Yongkiettrakul, S., Qin, D., Li, H., Byeon, I. J., Pei, D., and Tsai, M. D. (2000) Structure of the FHA1 domain of yeast Rad53 and identification of binding sites for both FHA1 and its target protein Rad9. *J. Mol. Biol. 304*, 941-951.

(49) Lee, G. I., Ding, Z., Walker, J. C., and Van Doren, S. R. (2003) NMR structure of the forkhead-associated domain from the Arabidopsis receptor kinase-associated protein phosphatase. *Proc Natl Acad Sci U S A 100*, 11261-6.

(50) Wang, P., Byeon, I. J., Liao, H., Beebe, K. D., Yongkiettrakul, S., Pei, D., and Tsai, M. D. (2000) II. Structure and specificity of the interaction between the FHA2 domain of Rad53 and phosphotyrosyl peptides. *J Mol Biol 302*, 927-40.

(51) Stavridi, E. S., Huyen, Y., Loreto, I. R., Scolnick, D. M., Halazonetis, T. D., Pavletich, N. P., and Jeffrey, P. D. (2002) Crystal Structure of the FHA Domain of the Chfr Mitotic Checkpoint Protein and Its Complex with Tungstate. *Structure 10*, 891-899.

(52) Li, H., Byeon, I. J., Ju, Y., and Tsai, M. D. (2004) Structure of human Ki67 FHA domain and its binding to a phosphoprotein fragment from hNIFK reveal unique recognition sites and new views to the structural basis of FHA domain functions. *J Mol Biol 335*, 371-81.

(53) Yuan, C., Yongkiettrakul,S, Byeon, I., J, Zhou, S, Tsai, M.,D. (2001) Solution structures of two FHA1-phosphothreonine peptide complexes provide insight into the structural basis of the ligand specificity of FHA1 from yeast Rad53. *J. Mol. Biol. 314*, 563-575.

(54) Wang, P., Byeon, I.-J. L., Liao, H., Beebe, K. D., Yongkiettrakul, S., Pei, D., and Tsai, M.-D. (2000) Structure and Specificity of the Interaction between the FHA2 Domain of RAD53 and Phosphotyrosyl Peptides. *Journal of Molecular Biology 302*, 927-940.

(55) Li, J., Williams, B. L., Haire, L. F., Goldberg, M., Wilker, E., Durocher, D., Yaffe, M. B., Jackson, S. P., and Smerdon, S. J. (2002) Structural and functional versatility of the FHA domain in DNA-damage signaling by the tumor suppressor kinase Chk2. *Mol. Cell 9*, 1045-54.

(56) Durocher, D., Henckel, J., Fersht, A. R., and Jackson, S. P. (1999) The FHA domain is a modular phosphopeptide recognition motif. *Mol. Cell 4*, 387-394.

(57) Durocher, D., Smerdon, S. J., Yaffe, M. B., and Jackson, S. P. (2000) The FHA domain in DNA repair and checkpoint signaling. *Cold Spring Harb Symp Quant Biol 65*, 423-31.

(58) Durocher, D., and Jackson, S. P. (2002) The FHA domain. *FEBS Lett. 513*, 58-66.

(59) Oh, M. H., Ray, W. K., Huber, S. C., Asara, J. M., Gage, D. A., and Clouse, S. D. (2000) Recombinant brassinosteroid insensitive 1 receptor-like kinase autophosphorylates on serine and threonine residues and phosphorylates a conserved peptide motif in vitro. *Plant Physiol 124*, 751-66.

(60) Byeon, I. J., Li, H., Song, H., Gronenborn, A. M., and Tsai, M. D. (2005) Sequential phosphorylation and multisite interactions characterize specific target recognition by the FHA domain of Ki67. *Nat Struct Mol Biol 12*, 987-93.

(61) Liang, X., Lee, G. I., and Van Doren, S. R. (2006) Partially Unfolded Forms and Non-two-state Folding of a beta-Sandwich: FHA Domain from Arabidopsis Receptor Kinase-associated Protein Phosphatase. *J Mol Biol*.

(62) Durocher, D., Taylor, I. A., Sarbassova, D., Haire, L. F., Westcott, S. L., Jackson, S. P., Smerdon, S. J., and Yaffe, M. B. (2000) The molecular basis of FHA domain:phosphopeptide binding specificity and implications for phospho-dependent signaling mechanisms. *Mol Cell 6*, 1169-82.
(63) Songyang, Z., Shoelson, S. E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W. G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R. J., and et al. (1993) SH2 domains recognize specific phosphopeptide sequences. *Cell 72*, 767-78.
(64) Wang, X., and Chory, J. (2006) Brassinosteroids Regulate Dissociation of BKI1, a Negative Regulator of BRI1 Signaling, from the Plasma Membrane. *Science*.
(65) Hornak, J. P. (1997-1999) *Basics of NMR*, Rochester Institute of Technology, Rochester, NY 14623-5603.
(66) Ishima, R., and Torchia, D. A. (2000) Protein dynamics from NMR. *Nat Struct Biol 7*, 740-3.
(67) Fischer, M. W., Zeng, L., Majumdar, A., and Zuiderweg, E. R. (1998) Characterizing semilocal motions in proteins by NMR relaxation studies. *Proc Natl Acad Sci U S A 95*, 8016-9.
(68) Karplus, M., Brunger, A. T., Elber, R., and Kuriyan, J. (1987) Molecular dynamics: applications to proteins. *Cold Spring Harb Symp Quant Biol 52*, 381-90.
(69) Levitt, M. H. (2002) *Spin dynamics: Basics of nuclear magnetic resonance*, John Wiley & Sons, Ltd, Chichester, England.
(70) Wolf-Watz, M., Thai, V., Henzler-Wildman, K., Hadjipavlou, G., Eisenmesser, E. Z., and Kern, D. (2004) Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. *Nat Struct Mol Biol 11*, 945-9.
(71) Berendsen, H. J., and Hayward, S. (2000) Collective protein dynamics in relation to function. *Curr Opin Struct Biol 10*, 165-9.
(72) Benkovic, S. J., and Hammes-Schiffer, S. (2003) A perspective on enzyme catalysis. *Science 301*, 1196-202.
(73) Agarwal, P. K., Geist, A., and Gorin, A. (2004) Protein dynamics and enzymatic catalysis: investigating the peptidyl-prolyl cis-trans isomerization activity of cyclophilin A. *Biochemistry 43*, 10605-18.
(74) Eisenmesser, E. Z., Millet, O., Labeikovsky, W., Korzhnev, D. M., Wolf-Watz, M., Bosco, D. A., Skalicky, J. J., Kay, L. E., and Kern, D. (2005) Intrinsic dynamics of an enzyme underlies catalysis. *Nature 438*, 117-21.
(75) Arumugam, S., Gao, G., Patton, B. L., Semenchenko, V., Brew, K., and Van Doren, S. R. (2003) Increased backbone mobility in β-barrel enhances entropy gain driving binding of N-TIMP-1 to MMP-3. *Journal of Molecular Biology 327*, 719-734.
(76) Ding, Z., Lee, G.-I., Liang, X., Gallazzi, F., Arunima, A., and Van Doren, S. R. (2005) PhosphoThr Peptide Binding Globally Rigidifies much of the FHA Domain from Arabidopsis Receptor Kinase-Associated Protein Phosphatase. *Biochemistry 44*, 10119-34.
(77) Lipari, G., and Szabo, A. (1982) Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 1. Theory and Range of Validity. *Journal of the American Chemical Society 104*, 4546-4559.

(78) Kay, L. E., Torchia, D. A., and Bax, A. (1989) Backbone Dynamics of Proteins as Studied by ^{15}N Inverse Detected Heteronuclear NMR Spectroscopy: Application to Staphylococcal Nuclease. *Biochemistry 28*, 8972-8979.
(79) Lipari, G., and Szabo, A. (1982) Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 2. Analysis of Experimental Results. *Journal of the American Chemical Society 104*, 4559-4570.
(80) Dosset, P., Hus, J. C., Blackledge, M., and Marion, D. (2000) Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *J Biomol NMR 16*, 23-8.
(81) d'Auvergne, E. J., and Gooley, P. R. (2003) The use of model selection in the model-free analysis of protein dynamics. *Journal of Biomolecular NMR 25*, 25-39.
(82) Clore, G. M., Szabo, A., Bax, A., Kay, L. E., Driscoll, P. C., and Gronenborn, A. M. (1990) Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. *Journal of the American Chemical Society 112*, 4989-4991.
(83) Mulder, F. A., Mittermaier, A., Hon, B., Dahlquist, F. W., and Kay, L. E. (2001) Studying excited states of proteins by NMR spectroscopy. *Nat Struct Biol 8*, 932-5.
(84) Palmer, A. G., 3rd. (2004) NMR characterization of the dynamics of biomacromolecules. *Chem Rev 104*, 3623-40.
(85) Peng, J. W., and Wagner, G. (1995) Frequency Spectrum of NH Bonds in Eglin C from Spectral Density Mapping at Multiple Fields. *Biochemistry 34*, 16733-16752.
(86) Viles, J. H., Donne, D., Kroon, G., Prusiner, S. B., Cohen, F. E., Dyson, H. J., and Wright, P. E. (2001) Local structural plasticity of the prion protein. Analysis of NMR relaxation dynamics. *Biochemistry 40*, 2743-53.
(87) Lefevre, J.-F., Dayie, K. T., Peng, J. W., and Wagner, G. (1996) Internal Mobility in the Partially Folded DNA Binding and Dimerization Domains of GAL4: NMR Analysis of the N-H Spectral Density Functions. *Biochemistry 35*, 2674-2686.
(88) Lian, L., and Roberts, G. C. K. (1993) Effects of chemical exchange on NMR spectra., in *NMR of macromolecules. A practical approach* (Roberts, G. C. K., Ed.) pp 153-182, Oxford University Press, Oxford, UK.
(89) Lichtarge, O., Bourne, H. R., and Cohen, F. E. (1996) An evolutionary trace method defines binding surfaces common to protein families. *J Mol Biol 257*, 342-58.
(90) Sowa, M. E., He, W., Slep, K. C., Kercher, M. A., Lichtarge, O., and Wensel, T. G. (2001) Prediction and confirmation of a site critical for effector regulation of RGS domain activity. *Nat Struct Biol 8*, 234-7.
(91) Sowa, M. E., He, W., Wensel, T. G., and Lichtarge, O. (2000) A regulator of G protein signaling interaction surface linked to effector specificity. *Proc Natl Acad Sci U S A 97*, 1483-8.
(92) Morris, E. R., and Walker, J. C. (2003) Receptor-like protein kinases: the keys to response. *Curr Opin Plant Biol 6*, 339-42.

(93) Clark, S. E., Running, M. P., and Meyerowitz, E. M. (1993) *CLAVATA1*, a regulator of meristem and flower development in *Arabidopsis*. *Development 119*, 397-418.
(94) Wang, Z. Y., Seto, H., Fujioka, S., Yoshida, S., and Chory, J. (2001) BRI1 is a critical component of a plasma-membrane receptor for plant steroids. *Nature 410*, 380-3.
(95) Li, J. (2005) Brassinosteroid signaling: from receptor kinases to transcription factors. *Curr Opin Plant Biol 8*, 526-31.
(96) Nam, K. H., and Li, J. (2002) BRI1/BAK1, a receptor kinase pair mediating brassinosteroid signaling. *Cell 110*, 203-12.
(97) Braun, D., M., and Walker, J., C. (1996) Plant transmembrane receptors: new pieces in the signaling puzzle. *Trends Biochem. Sci. 21*, 70-73.
(98) Li, J., and Chory, J. (1997) A putative leucine-rich repeator kinase involved in brassinosteroid signal transduction. *Cell 90*, 929-938.
(99) Wang, X., Goshe, M. B., Soderblom, E. J., Phinney, B. S., Kuchar, J. A., Li, J., Asami, T., Yoshida, S., Huber, S. C., and Clouse, S. D. (2005) Identification and functional analysis of in vivo phosphorylation sites of the Arabidopsis BRASSINOSTEROID-INSENSITIVE1 receptor kinase. *Plant Cell 17*, 1685-703.
(100) Shah, K., Vervoort, J., and de Vries, S. C. (2001) Role of threonines in the Arabidopsis thaliana somatic embryogenesis receptor kinase 1 activation loop in phosphorylation. *J Biol Chem 276*, 41263-9.
(101) Lee, G., Li, J., Walker, J. C., and Van Doren, S. R. (2003) ^{1}H, ^{13}C and ^{15}N resonance assignments of kinase-interacting FHA domain from Arabidopsis phosphatase KAPP. *J. Biomol. NMR. 25*, 253-254.
(102) Wang X, L. X., Meisenhelder J, Hunter T, Yoshida S, Asami T, Chory J. (2005) Autoregulation and homodimerization are involved in the activation of the plant steroid receptor BRI1. *Dev Cell. 8*, 855-65.
(103) Wiseman, T., Williston, S., Brandts, J. F., and Lin, L. N. (1989) Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal Biochem 179*, 131-7.
(104) Pierce, M. M., Raman, C. S., and Nall, B. T. (1999) Isothermal titration calorimetry of protein-protein interactions. *Methods 19*, 213-21.
(105) Doyle, M. L. (1997) Characterization of binding interactions by isothermal titration calorimetry. *Curr Opin Biotechnol. 8*, 31-35.
(106) Fisher, H. F. a. S., N. (1995) Calorimetric methods for interpreting protein-ligand interactions. *Methods Enzymol. 259*, 194-221.
(107) 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*, 277-93.
(108) Ferreon, J. C. a. H. V. J. (2003) Ligand-induced changes in dynamics in the RT loop of the C-terminal SH3 domain of Sem-5 indicate cooperative conformational coupling. *Protein Science 12*, 982-996.
(109) Yang, D., and Kay, L. E. (1996) Contributions to conformational entropy arising from bond vector fluctuations measured from NMR-derived order parameters: application to protein folding. *J Mol Biol 263*, 369-82.

(110) Horn, M., A., and Walker, J., C. (1994) Biochemical properties of the autophosphorylation of RLK5, a receptor-like protein kinase from *Arabidopsis thaliana*. *Biochemica et Biophysica Acta 1208*, 65-74.
(111) Stone, J. M., Trotochaud, A. E., Walker, J. C., and Clark, S. E. (1998) Control of meristem development by CLAVATA1 receptor kinase and KAPP protein phosphatase interactions. *Plant Physiology 117*, 1217-1225.
(112) Dievart, A., Dalal, M., Tax, F. E., Lacey, A. D., Huttly, A., Li, J., and Clark, S. E. (2003) CLAVATA1 dominant-negative alleles reveal functional overlap between multiple receptor kinases that regulate meristem and organ development. *Plant Cell 15*, 1198-211.
(113) DeYoung, B. J., and Clark, S. E. (2001) Signaling through the CLAVATA1 receptor complex. *Plant Mol Biol 46*, 505-13.
(114) Trotochaud, A. E., Jeong, S., and Clark, S. E. (2000) CLAVATA3, a multimeric ligand for the CLAVATA1 receptor-kinase. *Science 289*, 613-7.
(115) Lichtarge, O., and Sowa, M. E. (2002) Evolutionary predictions of binding surfaces and interactions. *Curr Opin Struct Biol 12*, 21-7.
(116) Hanks, S. K., and Hunter, T. (1995) Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *Faseb J 9*, 576-96.
(117) Nuhse, T. S., Stensballe, A., Jensen, O. N., and Peck, S. C. (2004) Phosphoproteomics of the Arabidopsis plasma membrane and a new phosphorylation site database. *Plant Cell 16*, 2394-405.
(118) Baldwin, R. L. (1986) Temperature dependence of the hydrophobic interaction in protein folding. *Proc Natl Acad Sci U S A 83*, 8069-72.
(119) Baker, B. M., and Murphy, K. P. (1998) Prediction of binding energetics from structure using empirical parameterization. *Methods Enzymol 295*, 294-315.
(120) Taylor, S. S., Haste, N. M., and Ghosh, G. (2005) PKR and eIF2alpha: integration of kinase dimerization, activation, and substrate docking. *Cell 122*, 823-5.
(121) Huse, M., Muir, T. W., Xu, L., Chen, Y. G., Kuriyan, J., and Massague, J. (2001) The TGF beta receptor activation process: an inhibitor- to substrate-binding switch. *Mol. Cell 8*, 671-682.
(122) Kay, L. E., Muhandiram, D. R., Wolf, G., Shoelson, S. E., and Forman-Kay, J. D. (1998) Correlation between binding and dynamics at SH2 domain interfaces. *Nat Struct Biol 5*, 156-63.
(123) Olejniczak, E. T., Zhou, M. M., and Fesik, S. W. (1997) Changes in the NMR-derived motional parameters of the insulin receptor substrate 1 phosphotyrosine binding domain upon binding to an interleukin 4 receptor phosphopeptide. *Biochemistry 36*, 4118-24.
(124) Lichtarge, O., Sowa, M. E., and Philippi, A. (2002) Evolutionary traces of functional surfaces along G protein signaling pathway. *Methods Enzymol 344*, 536-56.
(125) Hanks, S. K., and Quinn, A. M. (1991) Protein kinase catalytic domain sequence database: identification of conserved features of primary structure and classification of family members. *Methods Enzymol 200*, 38-62.

(126) Stone, M. J. (2001) NMR relaxation studies of the role of conformational entropy in protein stability and ligand binding [Review]. *Accounts of Chemical Research 34*, 379-388.
(127) Atkinson, R. A., and Kieffer, B. (2004) The role of protein motions in molecular recognition: insights from heteronuclear NMR relaxation measurements. *Progress in Nuclear Magnetic Resonance Spectroscopy 44*, 141-187.
(128) Wang, C., Pawley, N. H., and Nicholson, L. K. (2001) The role of backbone motions in ligand binding to the c-Src SH3 domain. *J Mol Biol 313*, 873-87.
(129) Wang, P., Byeon, I. J., Liao, H., Beebe, K. D., Yongkiettrakul, S., Pei, D., and Tsai, M. D. (2000) II. Structure and specificity of the interaction between the FHA2 domain of Rad53 and phosphotyrosyl peptides. *J. Mol. Biol. 302*, 927-40.
(130) Byeon, I. J., Yongkiettrakul, S., and Tsai, M. D. (2001) Solution structure of the yeast Rad53 FHA2 complexed with a phosphothreonine peptide pTXXL: comparison with the structures of FHA2-pYXL and FHA1-pTXXD complexes. *J. Mol. Biol. 314*, 577-588.
(131) Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone Dynamics of a Free and a Phosphopeptide-Complexed Src Homology 2 Domain Studied by ^{15}N Relaxation. *Biochemistry 33*, 5984-6003.
(132) Boyd, J., Hommel, U., and Campbell, I. D. (1990) Influence of cross-correlation between dipolar and anisotropic chemical shift relaxation mechanisms upon longitudinal relaxation rates of ^{15}N in macromolecules. *Chemical Physics Letters 175*, 477-482.
(133) Sklenar, V., Piotto, M., Leppik, R., and Saudek, V. (1993) Gradient-Tailored Water Suppression for ^{1}H-^{15}N HSQC Experiments Optimized to Retain Full Sensitivity. *Journal of Magnetic Resonance A 102*, 241-245.
(134) Barbato, G., Ikura, M., Kay, L. E., Pastor, R. W., and Bax, A. (1992) Backbone dynamics of calmodulin studied by 15N relaxation using inverse detected two-dimensional NMR spectroscopy: the central helix is flexible. *Biochemistry 31*, 5269-78.
(135) Loria, J. P., Rance, M., and Palmer, A. G., 3rd. (1999) A Relaxation-Compensated Carr-Purcell-Meiboom-Gill Sequence for Characterizing Chemical Exchange by NMR Spectroscopy. *Journal of the American Chemical Society 121*, 2331-2332.
(136) Goddard, T. D., and Kneller, D. G. (2000), University of California, San Francisco, San Francisco.
(137) Farrow, N. A., Zhang, O., Szabo, A., Torchia, D. A., and Kay, L. E. (1995) Spectral density function mapping using ^{15}N relaxation data exclusively. *Journal of Biomolecular NMR 6*, 153-162.
(138) Ishima, R., and Nagayama, K. (1995) Protein Backbone Dynamics Revealed by Quasi Spectral Density Function Analysis of Amide N-15 Nuclei. *Biochemistry 34*, 3162-3171.
(139) Palmer, A. G., 3rd, Rance, M., and Wright, P. E. (1991) Intramolecular Motions of a Zinc Finger DNA-Binding Domain from Xfin Characterized by Proton-Detected Natural Abundance ^{13}C Heteronuclear NMR Spectroscopy. *Journal of the American Chemical Society 113*, 4371-4380.

(140) Tjandra, N., Szabo, A., and Bax, A. (1996) Protein Backbone Dynamics and ^{15}N Chemical Shift Anisotropy from Quantitative Measurement of Relaxation Interference Effects. *Journal of the American Chemical Society 118*, 6986-6991.
(141) Fushman, D., and Cowburn, D. (1998) Model-Independent Analysis of ^{15}N Chemical Shift Anisotropy from NMR Relaxation Data. Ubiquitin as a Test Example. *Journal of the American Chemical Society 120*, 7109-7110.
(142) Volkman, B. F., Alam, S. L., Satterlee, J. D., and Markley, J. L. (1998) Solution structure and backbone dynamics of component IV of Glycerea dibranchiata monomeric hemoglobin-CO. *Biochemistry 37*, 10906-10919.
(143) Akaike, H. (1973) in *Information theory and an extension of the maximum likelihood principle* (Petrov, B. N., and Csaki, F., Eds.) pp 267-281, Academiai Kiado, Budapest.
(144) Larsson, G., Martinez, G., Schleucher, J., and Wijmenga, S. S. (2003) Detection of nano-second internal motion and determination of overall tumbling times independent of the time scale of internal motion in proteins from NMR relaxation data. *J Biomol NMR 27*, 291-312.
(145) Jin, D., Andrec, M., Montelione, G. T., and Levy, R. M. (1998) Propagation of experimental uncertainties using the Lipari-Szabo model-free analysis of protein dynamics. *Journal of Biomolecular NMR 12*, 471-92.
(146) Tolman, J. R., Flanagan, J. M., Kennedy, M. A., and Prestegard, J. H. (1997) NMR evidence for slow collective motions in cyanometmyoglobin.[comment]. *Nature Structural Biology 4*, 292-7.
(147) Tolman, J. R., Al-Hashimi, H. M., Kay, L. E., and Prestegard, J. H. (2001) Structural and dynamic analysis of residual dipolar coupling data for proteins. *J Am Chem Soc 123*, 1416-24.
(148) Zweckstetter, M., and Bax, A. (2000) Prediction of Sterically Induced Alignment in a Dilute Liquid Crystalline Phase: Aid to Protein Structure Determination by NMR. *J Am Chem Soc 122*, 3791-3792.
(149) Fushman, D., Weisemann, R., Thüring, H., and Rüterjans, H. (1994) Backbone dynamics of ribonuclease T1 and its complex with 2'GMP studied by two-dimensional heteronuclear NMR spectroscopy. *Journal of Biomolecular NMR 4*, 61-78.
(150) Pawley, N. H., Wang, C., Koide, S., and Nicholson, L. K. (2001) An improved method for distinguishing between anisotropic tumbling and chemical exchange in analysis of 15N relaxation parameters. *J Biomol NMR 20*, 149-65.
(151) Blackledge, M., Cordier, F., Dosset, P., and Marion, D. (1998) Precision and Uncertainty in the Characterization of Anisotropic Rotational Diffusion by ^{15}N Relaxation. *J Am Chem Soc 120*, 4538-4539.
(152) Dayie, K. T., Wagner, G., and Lefevre, J. F. (1996) Theory and practice of nuclear spin relaxation in proteins. *Annual Review of Physical Chemistry 47*, 243-82.
(153) Gao, G., Semenchenko, V., Arumugam, S., and Van Doren, S. R. (2000) Tissue inhibitor of metalloproteinases-1 undergoes microsecond to millisecond motions at sites of matrix metalloproteinase-induced fit. *J Mol Biol 301*, 537-52.

(154) Buchberger, A., Howard, M. J., Freund, S. M., Proctor, M., Butler, P. J., Fersht, A. R., and Bycroft, M. (2000) Biophysical characterization of elongin C from Saccharomyces cerevisiae. *Biochemistry 39*, 11137-46.
(155) Tugarinov, V., Liang, Z., Shapiro, Y. E., Freed, J. H., and Meirovitch, E. (2001) A structural mode-coupling approach to 15N NMR relaxation in proteins. *J Am Chem Soc 123*, 3055-63.
(156) Vugmeyster, L., Raleigh, D. P., Palmer, A. G., 3rd, and Vugmeister, B. E. (2003) Beyond the decoupling approximation in the model free approach for the interpretation of NMR relaxation of macromolecules in solution. *J Am Chem Soc 125*, 8400-4.
(157) Yongkiettrakul, S., Byeon, I. J., and Tsai, M. D. (2004) The ligand specificity of yeast Rad53 FHA domains at the +3 position is determined by nonconserved residues. *Biochemistry 43*, 3862-9.
(158) Williams Jr, D. C., Benjamin, D. C., Poljak, R. J., and Rule, G. S. (1996) Global Changes in Amide Hydrogen Exchange Rates for a Protein Antigen in Complex with Three Different Antibodies. *Journal of Molecular Biology 257*, 866-876.
(159) Jinn, T.-L., Stone, J. M., and Walker, J. C. (2000) HAESA, and Arabidopsis leucine-rich repeat receptor kinase, controls floral organ abscission. *Genes & Development 14*, 108-117.
(160) Arkin, M. R., Randal, M., DeLano, W. L., Hyde, J., Luong, T. N., Oslob, J. D., Raphael, D. R., Taylor, L., Wang, J., McDowell, R. S., Wells, J. A., and Braisted, A. C. (2003) Binding of small molecules to an adaptive protein-protein interface. *Proc Natl Acad Sci U S A 100*, 1603-1608.
(161) Wyss, D. F., Dayie, K. T., and Wagner, G. (1997) The counterreceptor binding site of human CD2 exhibits an extended surface patch with multiple conformations fluctuating with millisecond to microsecond motions. *Protein Science 6*, 534-542.
(162) Feher, V. A., and Cavanagh, J. (1999) Millisecond-timescale motions contribute to the function of the bacterial response regulator protein SpoOF. *Nature 400*, 289-293.
(163) Huang, K., Ghose, R., Flanagan, J. M., and Prestegard, J. H. (1999) Backbone dynamics of the N-terminal domain in E. coli DnaJ determined by 15N- and 13CO-relaxation measurements. *Biochemistry 38*, 10567-77.
(164) Berjanskii, M. V., Riley, M. I., and Van Doren, S. R. (2002) Hsc70-interacting HPD loop of the J domain of polyomavirus T antigens fluctuates in psec to nsec and μsec to msec. *J. Mol. Biol. 321*, 503-516.
(165) Fisher, H. F., and Singh, N. (1995) Calorimetric Methods for Interpreting Protein-Ligand Interactions. *Methods in Enzymology 259*, 194-221.
(166) Seeliger, M. A., Young, M., Henderson, M. N., Pellicena, P., King, D. S., Falick, A. M., and Kuriyan, J. (2005) High yield bacterial expression of active c-Abl and c-Src tyrosine kinases. *Protein Sci 14*, 3135-9.
(167) Zhou, P., Lugovskoy, A. A., and Wagner, G. (2001) A solubility-enhancement tag (SET) for NMR studies of poorly behaving proteins. *J Biomol NMR 20*, 11-4.
(168) Hubbell, W. L., Gross, A., Langen, R., and Lietzow, M. A. (1998) Recent advances in site-directed spin labeling of proteins. *Curr Opin Struct Biol 8*, 649-56.

(169) Bolin, K. A., Hanson, P., Wright, S. J., and Millhauser, G. L. (1998) An NMR investigation of the conformational effect of nitroxide spin labels on Ala-rich helical peptides. *J Magn Reson 131*, 248-53.
(170) Dominguez, C., Bonvin, A. M., Winkler, G. S., van Schaik, F. M., Timmers, H. T., and Boelens, R. (2004) Structural model of the UbcH5B/CNOT4 complex revealed by combining NMR, mutagenesis, and docking approaches. *Structure (Camb) 12*, 633-44.
(171) van Drogen-Petit, A., Zwahlen, C., Peter, M., and Bonvin, A. (2004) Insight into molecular interactions between two PB1 domains. *J Mol Biol 336*, 1195-1210.
(172) Axtell, M. J., Chisholm, S. T., Dahlbeck, D., and Staskawicz, B. J. (2003) Genetic and molecular evidence that the Pseudomonas syringae type III effector protein AvrRpt2 is a cysteine protease. *Mol Microbiol 49*, 1537-46.
(173) Coaker, G., Falick, A., and Staskawicz, B. (2005) Activation of a phytopathogenic bacterial effector protein by a eukaryotic cyclophilin. *Science 308*, 548-50.
(174) Day, B., Dahlbeck, D., Huang, J., Chisholm, S. T., Li, D., and Staskawicz, B. J. (2005) Molecular basis for the RIN4 negative regulation of RPS2 disease resistance. *Plant Cell 17*, 1292-305.
(175) Coaker, G., Zhu, G., Ding, Z., Van Doren, S. R., and Staskawicz, B. (2006) Eukaryotic cyclophilin as a molecular switch for effector activation. *Mol Microbiol 61*, 1485-96.
(176) Coaker, G., Zhu, G., Ding, Z., Van Doren, S. R., and Staskawicz, B. (2006) Eukaryotic Cyclophilin as Molecular Switch for Effector Activation. *Molecular Microbiology*, In press.

APPENDIX 1: Purification of KI-FHA

KI-FHA domains were over expressed in *E. coli* in LB media or ^{15}N labeled medium as described in the end. A frozen BL21 (DE3 RIL) bacterial stock with the GST-KI-FHA construct was used to plate an LB (Luria-Bertani) agar. The inoculated bacteria was grown at 37 ^{o}C for overnight. One colony was inoculated into a 50 ml medium and incubated at 37 ^{o}C with shaking for overnight. A 3-10 % inoculum from the overnight culture was used to inoculate a 500 ml medium. KI-FHA was induced when the $O.D._{600}$ reached about 0.8 by addition of IPTG to 200 μM. After protein expression for 5 hours, the cells were harvested at 10,000 rpm for 10 minutes. The pellet was resuspended in about 5-10 ml of PBS buffer per gram of wet cells, and lysed by two passes through a French Press cell. After centrifugation of the lysate at 15,000 rpm for 20 minutes, the supernatant was dialyzed against PBS buffer (pH 7.5) in cold. The glutathoine agarose beads (Sigma G4510) were added to the supernatant of lysate to 5 ml per liter of bacterial culture and incubate for 1 hour at 4 ^{o}C with gentle shaking. The beads were collected from the mixture. The same amount of beads was added to the supernatant and incubated same way for another hour. All beads were combined and washed with an excess amount (> 15 volume) of PBS buffer (pH 7.5). After washing, the beads were equilibrated with chilled 3C protease buffer. Precission protease (GE Healthcare) was added about 10 ul

per 1 ml of glutathione agarose and incubated without shaking for > 40 hours at 4 ^{o}C. Cut KI-FHA was collected by elution with PBS. The eluted fractions were pooled in a tube and dialyzed against the NMR buffer. KI-FHA NMR buffer was concentrated to 0.4-0.6 mM. 1 liter of original bacteria medium can yield 7-15 mgs of KI-FHA.

Media

5x M9 salt per liter

$NaHPO_4$	35.5 g
KH_2PO_4	15.0 g
NaCl	2.5 g
$^{15}NH_4Cl$	5.0 g
H_2O	up to 1000 ml

^{15}N labeled media per 500 ml

5X M9	100 ml
1M $MgSO_4$	1 ml
1M $CaCl_2$	0.050 ml
100X vitamin (Sigma) solution	5 ml
Glucose	5 ml
^{15}N Celtone^{TM}	6 % (v/v)
H_2O	up to 500 ml

Buffer

PBS buffer for bacteria lysis and glutathion binding

8 mg/ml NaCl

0.2 mg/ml KCl

1.44 mg/ml Na_2HPO_4

0.24 mg/ml KH_2PO_4

pH 7.4 - 7.5

3C protease (PreScission™-Pharmacia) buffer

50 mM Tris pH7.0

150 mM NaCl

1 mM EDTA

1 mM DTT

NMR buffer

20 mM sodium phosphate pH 6.3

120 mM NaCl

APPENDIX 2: PDB coordinates of the structural model of CLV1 kinase domain

The structure model was threaded by Shashi Pandit, and Jeff Skolnick at Center for the Systems Biology, Georgia Institute of Technology, Atlanta, GA.

```
REMARK  distant profile second pass of ap3potpred_8casp2long35 run 1agwA
264
ATOM      1   CA  THR     7     -12.380  17.308   7.730     1     4
ATOM      1   CA  ALA     8     -10.996  13.769   8.069     2     4
ATOM      1   CA  PHE     9     -12.836  10.889   6.337     3     4
ATOM      1   CA  GLN    10     -14.274   8.209   8.591     4     4
ATOM      1   CA  LYS    11     -12.802   4.696   8.428     5     0
ATOM      1   CA  LEU    12     -14.948   2.091  10.228     6     0
ATOM      1   CA  ASP    13     -12.078  -0.395   9.923     7     0
ATOM      1   CA  PHE    14      -9.842   1.447  12.390     8     0
ATOM      1   CA  LYS    15     -11.872   4.003  14.307     9     0
ATOM      1   CA  SER    16     -11.973   3.789  18.116    10     0
ATOM      1   CA  GLU    17     -14.164   6.044  20.290    11     0
ATOM      1   CA  LEU    20     -12.138   8.764  22.009    12     0
ATOM      1   CA  GLU    21     -13.476   7.825  25.419    13     0
ATOM      1   CA  CYS    22     -11.759   4.449  25.041    14     0
```

ATOM	1	CA	LEU	23	-8.279	5.938	25.013	15	4
ATOM	1	CA	ILE	28	-6.923	7.003	28.399	16	4
ATOM	1	CA	ILE	29	-4.002	9.301	27.573	17	4
ATOM	1	CA	GLY	30	-0.953	8.796	29.770	18	0
ATOM	1	CA	LYS	31	2.717	9.735	30.091	19	0
ATOM	1	CA	GLY	32	4.585	11.116	27.017	20	0
ATOM	1	CA	ALA	34	7.062	8.598	25.599	21	0
ATOM	1	CA	GLY	35	8.793	11.113	23.359	22	0
ATOM	1	CA	ILE	36	4.513	13.799	20.184	23	0
ATOM	1	CA	VAL	37	3.976	10.123	21.127	24	4
ATOM	1	CA	TYR	38	2.313	9.324	24.446	25	4
ATOM	1	CA	ARG	39	1.660	6.104	26.320	26	4
ATOM	1	CA	GLY	40	-2.018	5.335	26.867	27	4
ATOM	1	CA	SER	41	-4.520	2.627	27.779	28	4
ATOM	1	CA	MET	42	-7.223	1.528	25.357	29	4
ATOM	1	CA	PRO	43	-14.322	-4.765	25.795	33	0
ATOM	1	CA	ASN	44	-13.436	-4.957	29.498	34	0
ATOM	1	CA	ASN	45	-9.662	-5.275	29.164	35	4
ATOM	1	CA	VAL	46	-4.230	-2.564	26.413	37	4
ATOM	1	CA	ASP	47	-1.182	-0.355	26.768	38	4
ATOM	1	CA	VAL	48	-0.677	1.388	23.444	39	4
ATOM	1	CA	ALA	49	1.332	4.253	21.962	40	4
ATOM	1	CA	ILE	50	-0.532	7.265	20.581	41	4
ATOM	1	CA	LYS	51	0.966	9.554	17.920	42	4
ATOM	1	CA	ARG	52	-0.401	13.102	17.762	43	4
ATOM	1	CA	LEU	53	0.401	16.625	16.561	44	4
ATOM	1	CA	VAL	54	1.563	19.598	18.627	45	0
ATOM	1	CA	GLY	55	-0.236	22.931	19.000	46	0
ATOM	1	CA	ARG	56	2.227	24.741	16.737	47	0
ATOM	1	CA	GLY	57	1.729	22.167	13.944	48	0
ATOM	1	CA	THR	58	0.825	23.052	10.364	49	0
ATOM	1	CA	GLY	59	-1.616	21.443	7.936	50	0

ATOM	1	CA	ARG	60	1.463	19.888	6.330	51	0
ATOM	1	CA	SER	61	2.257	18.138	9.586	52	2
ATOM	1	CA	ASP	62	-1.321	16.841	9.598	53	2
ATOM	1	CA	HIS	63	-0.814	15.676	6.006	54	2
ATOM	1	CA	GLY	64	2.264	13.701	6.938	55	2
ATOM	1	CA	PHE	65	0.648	12.099	9.985	56	2
ATOM	1	CA	THR	66	-2.349	11.036	7.889	57	2
ATOM	1	CA	ALA	67	0.055	9.666	5.282	58	2
ATOM	1	CA	GLU	68	1.900	7.503	7.808	59	2
ATOM	1	CA	ILE	69	-1.410	6.030	8.997	60	2
ATOM	1	CA	GLN	70	-2.755	5.291	5.519	61	2
ATOM	1	CA	THR	71	0.598	3.690	4.766	62	2
ATOM	1	CA	LEU	72	0.369	1.425	7.813	63	2
ATOM	1	CA	GLY	73	-3.138	0.362	6.668	64	2
ATOM	1	CA	ARG	74	-1.973	-0.811	3.230	65	2
ATOM	1	CA	ILE	75	1.148	-2.616	4.545	66	4
ATOM	1	CA	ARG	76	0.449	-6.768	10.087	68	4
ATOM	1	CA	HIS	77	3.592	-8.637	11.069	69	0
ATOM	1	CA	ARG	78	5.195	-9.648	14.390	70	0
ATOM	1	CA	HIS	79	8.335	-7.670	13.670	71	0
ATOM	1	CA	ILE	80	7.030	-4.194	12.729	72	0
ATOM	1	CA	VAL	81	5.123	-1.620	14.778	73	4
ATOM	1	CA	ARG	82	1.504	-2.368	13.856	74	4
ATOM	1	CA	LEU	83	-1.450	0.017	13.682	75	4
ATOM	1	CA	LEU	84	-4.065	-0.735	16.347	76	4
ATOM	1	CA	GLY	85	-6.671	1.932	15.623	77	4
ATOM	1	CA	TYR	86	-7.383	5.668	15.393	78	4
ATOM	1	CA	VAL	87	-9.452	8.619	16.716	79	4
ATOM	1	CA	ALA	88	-10.065	10.881	13.711	80	0
ATOM	1	CA	ASN	89	-13.440	12.417	14.357	81	0
ATOM	1	CA	LYS	90	-14.444	15.453	16.423	82	0
ATOM	1	CA	ASP	91	-11.006	16.268	17.782	83	0

ATOM	1	CA	THR	92	-7.204	15.840	17.341	84	0
ATOM	1	CA	ASN	93	-6.182	12.949	15.099	85	4
ATOM	1	CA	LEU	94	-4.856	10.065	17.182	86	4
ATOM	1	CA	LEU	95	-2.878	7.284	15.522	87	4
ATOM	1	CA	LEU	96	-2.832	4.321	17.893	88	4
ATOM	1	CA	TYR	97	0.182	2.029	17.390	89	4
ATOM	1	CA	GLU	98	1.851	-0.892	19.204	90	4
ATOM	1	CA	TYR	99	3.694	-0.120	22.424	91	4
ATOM	1	CA	MET	100	7.318	-1.118	22.991	92	0
ATOM	1	CA	PRO	101	8.025	-0.987	26.731	93	0
ATOM	1	CA	ASN	102	11.779	-1.395	26.432	94	0
ATOM	1	CA	GLY	103	12.519	1.476	24.084	95	0
ATOM	1	CA	SER	104	14.781	1.373	21.039	96	0
ATOM	1	CA	LEU	105	17.304	-1.353	20.251	97	0
ATOM	1	CA	GLY	106	20.404	0.861	20.692	98	2
ATOM	1	CA	GLU	107	19.521	1.763	24.304	99	2
ATOM	1	CA	LEU	108	18.370	-1.793	25.003	100	2
ATOM	1	CA	LEU	109	21.767	-3.064	23.882	101	2
ATOM	1	CA	HIS	110	23.760	-0.421	25.737	102	2
ATOM	1	CA	GLY	111	21.852	-1.014	28.966	103	2
ATOM	1	CA	SER	112	22.948	-4.663	28.914	104	0
ATOM	1	CA	LYS	113	26.665	-4.179	28.411	105	0
ATOM	1	CA	GLY	114	20.517	-4.828	32.430	108	0
ATOM	1	CA	GLY	115	21.472	-8.521	32.380	109	0
ATOM	1	CA	HIS	116	23.918	-9.410	29.611	110	0
ATOM	1	CA	LEU	117	22.738	-10.964	26.365	111	0
ATOM	1	CA	GLN	118	24.090	-14.202	25.001	112	0
ATOM	1	CA	TRP	119	25.583	-14.685	21.571	113	0
ATOM	1	CA	GLU	120	22.391	-16.484	20.634	114	2
ATOM	1	CA	THR	121	20.310	-13.582	21.956	115	2
ATOM	1	CA	ARG	122	22.246	-11.137	19.776	116	2
ATOM	1	CA	HIS	123	21.987	-13.220	16.590	117	2

ATOM	1	CA	ARG	124	18.311	-13.697	17.397	118	2
ATOM	1	CA	VAL	125	17.929	-9.897	17.573	119	2
ATOM	1	CA	ALA	126	19.435	-9.644	14.061	120	2
ATOM	1	CA	VAL	127	17.159	-12.399	12.681	121	2
ATOM	1	CA	GLU	128	13.950	-10.703	13.894	122	2
ATOM	1	CA	ALA	129	15.037	-7.384	12.355	123	2
ATOM	1	CA	ALA	130	15.822	-9.148	9.065	124	2
ATOM	1	CA	LYS	131	12.349	-10.679	9.319	125	2
ATOM	1	CA	GLY	132	10.544	-7.382	9.825	126	2
ATOM	1	CA	LEU	133	12.445	-6.023	6.833	127	2
ATOM	1	CA	CYS	134	11.749	-9.098	4.655	128	2
ATOM	1	CA	TYR	135	8.008	-8.485	5.137	129	2
ATOM	1	CA	LEU	136	8.300	-4.715	4.518	130	2
ATOM	1	CA	HIS	137	10.305	-5.363	1.335	131	2
ATOM	1	CA	SER	141	7.711	-7.908	0.173	132	2
ATOM	1	CA	PRO	142	5.013	-5.250	0.726	133	0
ATOM	1	CA	LEU	143	6.959	-2.905	-1.581	134	4
ATOM	1	CA	ILE	144	8.084	-0.621	1.266	135	4
ATOM	1	CA	LEU	145	11.488	1.101	1.401	136	4
ATOM	1	CA	HIS	146	12.363	2.233	4.942	137	4
ATOM	1	CA	ARG	147	15.078	4.754	4.035	138	4
ATOM	1	CA	ASP	148	16.266	5.058	7.659	139	4
ATOM	1	CA	VAL	149	16.800	1.620	9.149	140	4
ATOM	1	CA	LYS	150	18.942	1.865	12.314	141	0
ATOM	1	CA	SER	151	18.909	0.474	15.855	142	0
ATOM	1	CA	ASN	152	17.330	3.769	16.958	143	0
ATOM	1	CA	ASN	153	14.329	2.933	14.741	144	0
ATOM	1	CA	ILE	154	13.804	-0.622	16.060	145	4
ATOM	1	CA	LEU	155	11.762	-0.914	19.241	146	4
ATOM	1	CA	LEU	156	11.784	-3.699	21.842	147	4
ATOM	1	CA	ASP	157	8.549	-4.925	23.454	148	0
ATOM	1	CA	SER	158	8.160	-6.210	27.015	149	0

ATOM	1	CA	ASP	159	8.583	-9.705	25.569	150	0
ATOM	1	CA	PHE	160	11.848	-8.773	23.828	151	0
ATOM	1	CA	GLU	161	10.289	-8.779	20.363	152	4
ATOM	1	CA	ALA	162	12.113	-6.541	17.897	153	4
ATOM	1	CA	HIS	163	9.733	-4.242	16.011	154	4
ATOM	1	CA	VAL	164	10.733	-1.966	13.141	155	4
ATOM	1	CA	ALA	165	9.372	1.587	13.446	156	4
ATOM	1	CA	ASP	166	8.963	4.682	11.269	157	4
ATOM	1	CA	PHE	167	8.632	2.645	8.076	158	4
ATOM	1	CA	GLY	168	5.569	4.601	6.890	159	4
ATOM	1	CA	LEU	169	7.041	8.013	7.576	160	4
ATOM	1	CA	ALA	170	6.568	10.851	5.082	161	4
ATOM	1	CA	LYS	171	9.859	12.362	3.888	162	4
ATOM	1	CA	PHE	172	14.650	13.065	0.222	164	0
ATOM	1	CA	LEU	173	17.284	16.120	3.998	167	4
ATOM	1	CA	VAL	174	19.661	18.207	6.095	168	4
ATOM	1	CA	ASP	175	22.893	16.206	6.075	169	4
ATOM	1	CA	GLY	176	24.366	18.395	8.828	170	4
ATOM	1	CA	ALA	177	21.758	17.368	11.388	171	4
ATOM	1	CA	ALA	178	22.375	15.137	14.397	172	4
ATOM	1	CA	SER	179	18.484	13.368	19.460	174	0
ATOM	1	CA	GLU	180	21.461	11.397	20.828	175	0
ATOM	1	CA	CYS	181	23.787	13.549	18.738	176	0
ATOM	1	CA	MET	182	24.554	11.227	15.820	177	0
ATOM	1	CA	SER	183	24.663	11.713	12.055	178	0
ATOM	1	CA	SER	184	21.966	9.704	10.186	179	0
ATOM	1	CA	ILE	185	24.096	9.896	7.004	180	0
ATOM	1	CA	ALA	186	26.212	7.172	8.647	181	0
ATOM	1	CA	TYR	191	23.519	4.618	7.905	182	0
ATOM	1	CA	ILE	192	22.946	5.637	4.258	183	0
ATOM	1	CA	ALA	193	24.174	3.670	1.289	184	0
ATOM	1	CA	PRO	194	26.572	5.593	-0.963	185	0

ATOM	1	CA	GLU	195	24.031	5.636	-3.797	186	0
ATOM	1	CA	TYR	196	21.302	6.969	-1.472	187	0
ATOM	1	CA	ALA	197	23.619	9.551	0.087	188	0
ATOM	1	CA	TYR	198	25.321	10.715	-3.131	189	0
ATOM	1	CA	THR	199	22.682	9.964	-5.785	190	0
ATOM	1	CA	LEU	200	19.454	10.018	-3.747	191	0
ATOM	1	CA	LYS	201	18.540	6.567	-5.012	192	0
ATOM	1	CA	VAL	202	16.657	4.614	-2.342	193	0
ATOM	1	CA	ASP	203	15.937	0.931	-2.932	194	0
ATOM	1	CA	GLU	204	15.513	-2.268	-0.923	195	0
ATOM	1	CA	LYS	205	19.280	-2.545	-1.533	196	0
ATOM	1	CA	SER	206	20.204	0.686	0.217	197	2
ATOM	1	CA	ASP	207	18.049	-0.616	3.086	198	2
ATOM	1	CA	VAL	208	20.279	-3.709	3.078	199	2
ATOM	1	CA	TYR	209	23.341	-1.442	3.399	200	2
ATOM	1	CA	SER	210	21.588	0.139	6.434	201	2
ATOM	1	CA	PHE	211	20.909	-3.368	7.854	202	2
ATOM	1	CA	GLY	212	24.658	-4.075	7.760	203	2
ATOM	1	CA	VAL	213	25.261	-0.930	9.841	204	2
ATOM	1	CA	VAL	214	22.423	-2.045	12.168	205	2
ATOM	1	CA	LEU	215	24.080	-5.481	12.478	206	2
ATOM	1	CA	LEU	216	27.315	-3.702	13.467	207	2
ATOM	1	CA	GLU	217	25.445	-1.869	16.265	208	2
ATOM	1	CA	LEU	218	24.089	-5.162	17.599	209	2
ATOM	1	CA	ILE	219	27.479	-6.773	17.825	210	2
ATOM	1	CA	ALA	220	29.004	-3.648	19.411	211	0
ATOM	1	CA	GLY	221	24.998	0.113	20.917	213	4
ATOM	1	CA	LYS	222	28.238	1.198	19.334	214	4
ATOM	1	CA	LYS	223	28.318	4.506	17.489	215	4
ATOM	1	CA	PRO	224	29.330	4.218	13.804	216	4
ATOM	1	CA	TRP	237	32.353	6.438	12.998	217	4
ATOM	1	CA	VAL	238	32.795	8.382	16.288	218	4

ATOM	1	CA	ARG	239	33.661	11.708	11.984	220	4
ATOM	1	CA	ASN	240	31.762	14.688	10.483	221	0
ATOM	1	CA	THR	241	29.647	14.772	7.282	222	0
ATOM	1	CA	GLU	242	32.504	15.980	5.056	223	2
ATOM	1	CA	GLU	243	35.008	13.381	6.245	224	2
ATOM	1	CA	GLU	244	32.398	10.652	5.918	225	2
ATOM	1	CA	ILE	245	31.917	11.452	2.239	226	2
ATOM	1	CA	THR	246	35.650	10.935	1.762	227	2
ATOM	1	CA	GLN	247	35.729	7.620	3.649	228	2
ATOM	1	CA	PRO	248	32.808	6.381	1.537	229	2
ATOM	1	CA	SER	249	34.442	7.586	-1.694	230	2
ATOM	1	CA	ASP	250	37.616	5.732	-0.724	231	0
ATOM	1	CA	ALA	256	36.177	2.476	3.586	233	4
ATOM	1	CA	ILE	257	34.582	-0.265	5.709	234	4
ATOM	1	CA	VAL	258	34.435	-0.391	9.522	235	4
ATOM	1	CA	ASP	259	36.514	-2.974	11.404	236	0
ATOM	1	CA	PRO	260	35.176	-6.258	12.768	237	0
ATOM	1	CA	LEU	262	33.606	-5.920	16.209	238	0
ATOM	1	CA	THR	263	35.331	-7.950	18.920	239	0
ATOM	1	CA	GLY	264	33.278	-11.005	19.864	240	0
ATOM	1	CA	TYR	265	31.741	-11.348	16.403	241	0
ATOM	1	CA	PRO	266	32.321	-14.391	14.198	242	0
ATOM	1	CA	THR	268	33.987	-14.067	10.801	243	0
ATOM	1	CA	SER	269	30.698	-15.374	9.390	244	2
ATOM	1	CA	VAL	270	28.684	-12.509	10.845	245	2
ATOM	1	CA	ILE	271	31.288	-9.893	9.868	246	2
ATOM	1	CA	HIS	272	31.194	-11.289	6.318	247	2
ATOM	1	CA	VAL	273	27.429	-10.740	6.400	248	2
ATOM	1	CA	PHE	274	27.991	-7.062	7.283	249	2
ATOM	1	CA	LYS	275	30.638	-6.779	4.557	250	2
ATOM	1	CA	MET	279	28.223	-8.409	2.069	251	2
ATOM	1	CA	CYS	280	25.527	-5.877	3.057	252	2

ATOM	1	CA	VAL	281	28.047	-3.090	2.364	253	0
ATOM	1	CA	GLU	282	28.797	-4.281	-1.168	254	0
ATOM	1	CA	GLU	283	29.703	-1.320	-3.390	255	0
ATOM	1	CA	GLU	284	27.407	-2.695	-6.130	256	0
ATOM	1	CA	ALA	285	23.714	-2.719	-4.979	257	0
ATOM	1	CA	ALA	286	22.826	-5.758	-7.139	258	0
ATOM	1	CA	ALA	287	25.442	-7.897	-5.413	259	0
ATOM	1	CA	ARG	288	24.149	-7.363	-1.877	260	0
ATOM	1	CA	PRO	289	22.034	-10.170	-0.440	261	0
ATOM	1	CA	THR	290	18.319	-9.683	-0.087	262	0
ATOM	1	CA	MET	291	16.449	-9.711	3.220	263	0
ATOM	1	CA	ARG	292	14.932	-13.020	2.108	264	2
ATOM	1	CA	GLU	293	18.501	-14.438	1.909	265	2
ATOM	1	CA	VAL	294	19.562	-12.761	5.161	266	2
ATOM	1	CA	VAL	295	16.596	-14.317	6.929	267	2
ATOM	1	CA	HIS	296	17.649	-17.767	5.733	268	2
ATOM	1	CA	MET	297	21.251	-17.223	6.761	269	2
ATOM	1	CA	LEU	298	20.450	-15.796	10.197	270	2
ATOM	1	CA	THR	299	17.869	-18.514	10.787	271	2
ATOM	1	CA	ASN	300	20.619	-21.068	10.146	272	2
ATOM	1	CA	PRO	301	23.208	-19.243	12.236	273	2
ATOM	1	CA	ASN	307	20.900	-18.717	15.240	274	2
ATOM	1	CA	LEU	308	20.374	-22.479	15.524	275	2
ATOM	1	CA	ILE	309	24.158	-22.953	15.346	276	0
ATOM	1	CA	ALA	310	24.699	-20.180	17.895	277	0
ATOM	1	CA	PHE	311	25.406	-21.337	21.446	278	0

====

TER 1agwA

APPENDIX 3: The input files of the Modelfree calculations on peptide bound KI-FHA

The output relaxation data of pT868CLV1 bound KI-FHA and free KI-FHA, including T_1, T_2, and ^1H {^{15}N} steady-state NOE at 500MHz and 600 MHz, have been deposited to http://www.bmrb.wisc.edu/, with entry number 6474 and 5841, respectively. Modelfree requires multiple input files, as shown below.

```
modelfree -i mfin -p mfpar -d mfdata -m mfmodel -s .pdb -o mfoutput &
```

The main control file (mfin)

```
optimization tval
```

```
seed -228

search grid

diffusion axial none

algorithm fix

simulations pred 400 0.00

selection none

sim_algorithm fix

fields 2 500.1323    599.69

tm        10.70   1  2  10.7   10.7  10
Dratio    1.209   1  2  1.209  1.209 10
Theta     43.33   1  2  43.33  43.33  10
Phi       -220    1  2  -220   -220.0  10
```

The model file (mfmodel)

```
spin   3
M2 tloc  12    0   2       0.000      18.400 20
M2 Theta 0.0   0   2       0.000      15.000 20
M2 S2f   1.0   0   2       0.000       1.000 20
M2 S2s   1.0   1   2       0.000       1.000 20
```

```
M2 te    0.0   1   2       0.000      2000.000 20
M2 Rex   0.0   0   2       0.000      20.000 20

spin   4
M2 tloc  12    0   2       0.000      18.400 20
M2 Theta 0.0   0   2       0.000      15.000 20
M2 S2f   1.0   0   2       0.000       1.000 20
M2 S2s   1.0   1   2       0.000       1.000 20
M2 te    0.0   1   2       0.000      2000.000 20
M2 Rex   0.0   1   2       0.000      20.000 20

spin   5
M2 tloc  12    0   2       0.000      18.400 20
M2 Theta 0.0   0   2       0.000      15.000 20
M2 S2f   1.0   0   2       0.000       1.000 20
M2 S2s   1.0   1   2       0.000       1.000 20
M2 te    0.0   1   2       0.000      2000.000 20
M2 Rex   0.0   0   2       0.000      20.000 20

spin   6
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   1   2      0.000      2000.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   7
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
```

```
M1 te    0.0   1   2      0.000      2000.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   8
M2 tloc  12    0   2       0.000      18.400 20
M2 Theta 0.0   0   2       0.000      15.000 20
M2 S2f   1.0   0   2       0.000       1.000 20
M2 S2s   1.0   1   2       0.000       1.000 20
M2 te    0.0   1   2       0.000      2000.000 20
M2 Rex   0.0   1   2       0.000      20.000 20

spin   9
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2       0.000       0.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   10
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   11
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
```

```
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   12
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   1   2      0.000       400.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   14
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   15
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   17
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
```

```
M1 te    0.0   1   2      0.000      2000.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   18
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   19
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   1   2      0.000       400.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   20
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   21
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
```

```
M1 te    0.0   1   2      0.000      2000.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   22
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2       0.000       0.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   23
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   24
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   25
M2 tloc  12    0   2       0.000      18.400 20
M2 Theta 0.0   0   2       0.000      15.000 20
M2 S2f   1.0   0   2       0.000       1.000 20
M2 S2s   1.0   1   2       0.000       1.000 20
```

```
M2 te    0.0   1   2       0.000      2000.000 20
M2 Rex   0.0   1   2       0.000      20.000 20

spin   26
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   27
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   1   2      0.000       2000.0 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   28
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   1   2      0.000       2000.0 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   29
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
```

```
M1 te    0.0   1   2      0.000       400.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   31
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   32
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2       0.000       0.000 20
M1 Rex   0.0   0   2       0.000       0.000 20

spin   34
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   35
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
```

```
M1 te    0.0   1   2      0.000      2000.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   37
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   1   2      0.000      2000.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   39
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   42
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   1   2      0.000      2000.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   44
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
```

```
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   45
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   46
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   47
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   50
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
```

```
M1 te    0.0   1   2      0.000      2000.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   51
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   54
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   55
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   1   2      0.000       400.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   56
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
```

```
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   57
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2       0.000       0.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   58
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   59
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2       0.000       0.000 20
M1 Rex   0.0   0   2       0.000       0.000 20

spin   60
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
```

```
M1 te    0.0   0   2       0.000       0.000 20
M1 Rex   0.0   0   2       0.000       0.000 20

spin   61
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   1   2      0.000      2000.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   62
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   1   2      0.000      2000.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   63
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   1   2      0.000       400.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   64
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
```

```
M1 te    0.0   0   2       0.000       0.000 20
M1 Rex   0.0   0   2       0.000       0.000 20

spin   65
M2 tloc  12    0   2       0.000      18.400 20
M2 Theta 0.0   0   2       0.000      15.000 20
M2 S2f   1.0   0   2       0.000       1.000 20
M2 S2s   1.0   1   2       0.000       1.000 20
M2 te    0.0   1   2       0.000      400.000 20
M2 Rex   0.0   0   2       0.000       0.000 20

spin   66
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2       0.000       0.000 20
M1 Rex   0.0   0   2       0.000       0.000 20

spin   67
M2 tloc  12    0   2       0.000      18.400 20
M2 Theta 0.0   0   2       0.000      15.000 20
M2 S2f   1.0   0   2       0.000       1.000 20
M2 S2s   1.0   1   2       0.000       1.000 20
M2 te    0.0   1   2       0.000      2000.000 20
M2 Rex   0.0   0   2       0.000      20.000 20

spin   68
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
```

```
M1 te    0.0   0   2       0.000       0.000 20
M1 Rex   0.0   0   2       0.000       0.000 20

spin   69
M2 tloc  12    0   2       0.000      18.400 20
M2 Theta 0.0   0   2       0.000      15.000 20
M2 S2f   1.0   0   2       0.000       1.000 20
M2 S2s   1.0   1   2       0.000       1.000 20
M2 te    0.0   1   2       0.000      1000.00 20
M2 Rex   0.0   0   2       0.000       0.000 20

spin   70
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   71
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   1   2      0.000       400.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   72
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
```

```
M1 te    0.0   1   2      0.000       400.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   73
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   1   2      0.000       400.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   75
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   1   2      0.000       400.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   76
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   77
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
```

```
M1 te    0.0   1   2      0.000       2000.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   78
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   1   2      0.000       400.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   79
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2       0.000       0.000 20
M1 Rex   0.0   0   2       0.000       0.000 20

spin   80
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   81
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
```

```
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   82
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   1   2      0.000      2000.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   83
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   84
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   1   2      0.000       400.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   85
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
```

```
M1 te    0.0   1   2      0.000       400.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   86
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   1   2      0.000      2000.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   87
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   1   2      0.000       2000.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   89
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   1   2      0.000       400.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   91
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
```

```
M1 te    0.0   1   2      0.000      2000.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   92
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   1   2      0.000      2000.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   93
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   1   2      0.000       2000.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   94
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   1   2      0.000      2000.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   95
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
```

```
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   96
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   97
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2       0.000       0.000 20
M1 Rex   0.0   0   2       0.000       0.000 20

spin   99
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2       0.000       0.000 20
M1 Rex   0.0   0   2       0.000       0.000 20

spin   100
M2 tloc  12    0   2       0.000      18.400 20
M2 Theta 0.0   0   2       0.000      15.000 20
M2 S2f   1.0   0   2       0.000       1.000 20
M2 S2s   1.0   1   2       0.000       1.000 20
```

```
M2 te    0.0   1   2       0.000      400.000 20
M2 Rex   0.0   0   2       0.000      20.000 20

spin   101
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2       0.000       0.000 20
M1 Rex   0.0   0   2       0.000       0.000 20

spin   102
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   103
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   104
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
```

```
M1 te    0.0   0   2       0.000       0.000 20
M1 Rex   0.0   0   2       0.000       0.000 20

spin   105
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   106
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2       0.000       0.000 20
M1 Rex   0.0   0   2       0.000       0.000 20

spin   107
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   108
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
```

```
M1 te    0.0   1   2      0.000      2000.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   109
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   110
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2       0.000       0.000 20
M1 Rex   0.0   0   2       0.000      20.000 20

spin   112
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   1   2      0.000       2000.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   114
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
```

```
M1 te    0.0   0   2       0.000       0.000 20
M1 Rex   0.0   0   2       0.000       0.000 20

spin   115
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   116
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   117
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2      0.000       0.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin   118
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
```

```
M1 te    0.0   0   2       0.000       0.000 20
M1 Rex   0.0   0   2       0.000       0.000 20

spin   120
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   0   2       0.000       0.000 20
M1 Rex   0.0   0   2       0.000       0.000 20

spin 121
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   1   2      0.000       2000.000 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin 122
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   1   2      0.000       2000.0 20
M1 Rex   0.0   1   2       0.000      20.000 20

spin 123
M2 tloc  12    0   2       0.000      18.400 20
M2 Theta 0.0   0   2       0.000      15.000 20
M2 S2f   1.0   0   2       0.000       1.000 20
M2 S2s   1.0   1   2       0.000       1.000 20
```

```
M2 te    0.0   1   2       0.000      2000.000 20
M2 Rex   0.0   1   2       0.000      20.000 20

spin 124
M1 tloc  12    0   2       0.000      18.400 20
M1 Theta 0.0   0   2       0.000      15.000 20
M1 S2f   1.0   0   2       0.000       1.000 20
M1 S2s   1.0   1   2       0.000       1.000 20
M1 te    0.0   1   2      0.000      2000.000 20
M1 Rex   0.0   1   2       0.000      20.000 20
```

The Spin relaxation data file (mfdata)

```
spin  3
R1     500.13 1.755  0.1358 1
R2     500.13 4.924  0.3297 0
NOE    500.13                     0

R1     599.69 1.45    0.0419      1
R2     599.69 5.3677  0.0824      1
NOE    599.69              0

spin  4
R1     500.13 1.969   0.1745      1
R2     500.13 5.942   0.251 0
NOE    500.13              0

R1     599.69 1.54   0.0785 1
R2     599.69 6.6357 0.0894 1
NOE    599.69              0
```

```
spin  5
R1     500.13 2.127  0.2498 1
R2     500.13 10.607 0.3139 0
NOE    500.13 0.371  0.0307 1

R1     599.69 1.598  0.0682 1
R2     599.69 10.570 0.1162 1
NOE    599.69 0.493  0.0171 1

spin  6
R1     500.13 2.01857       0.19191       1
R2     500.13 14.57938      0.22744       0
NOE    500.13 0.6375  0.0353      1

R1     599.69 1.52277       0.05264       1
R2     599.69 15.760 0.2355 1
NOE    599.69 0.81392       0.04   1

spin  7
R1     500.13 1.98807       0.18972       1
R2     500.13 14.27348      0.55415       0
NOE    500.13 0.8012  0.0758      1

R1     599.69 1.42328       0.04517       1
R2     599.69 17.2265       0.1953 1
NOE    599.69 0.81902       0.04   1

spin  8
R1     500.13 1.78891       0.09953       1
R2     500.13 14.02721      0.29711       0
NOE    500.13 0.5448  0.0485      1
```

```
R1     599.69 1.33067       0.03701       1
R2     599.69 16.0849       0.3286 1
NOE    599.69 0.81184       0.04   1

spin  9
R1     500.13 1.89609       0.13446       1
R2     500.13 14.24501      0.99634       0
NOE    500.13 0.8149  0.0595      1

R1     599.69 1.3369 0.02824       1
R2     599.69 18.6811       0.2760 1
NOE    599.69 0.80279       0.04   1

spin  10
R1     500.13 2.14546       0.27388       1
R2     500.13 15.80028      0.74395       0
NOE    500.13 0.6783  0.1082      1

R1     599.69 1.1854 0.04778       1
R2     599.69 22.7118       1.0471 1
NOE    599.69 0.78625       0.04   1

spin  11
R1     500.13 2.11595       0.08417       1
R2     500.13 13.98406      1.43928       0
NOE    500.13 0.4014 0.1133 1

R1     599.69 1.39005       0.03169       1
R2     599.69 23.2234       0.4056 1
NOE    599.69 0.77007       0.04   1
```

```
spin  12
R1     500.13 2.10128       0.33822       1
R2     500.13 16.88619      1.42287       0
NOE    500.13 0.8709  0.0871      1

R1     599.69 1.2219 0.10944       1
R2     599.69 23.3973       0.6405 1
NOE    599.69 0.76876       0.04   1

spin  14
R1     500.13 1.95886       0.1815 1
R2     500.13 14.15829      0.88001       0
NOE    500.13 0.7064 0.0976 1

R1     599.69 1.32837       0.04782       1
R2     599.69 17.3671       0.5580 1
NOE    599.69 0.8085 0.04   1

spin  15
R1     500.13 1.82715       0.14322       1
R2     500.13 14.07658      0.20013       0
NOE    500.13 0.6932 0.04   1

R1     599.69 1.33529       0.02853       1
R2     599.69 17.6118       0.2909 1
NOE    599.69 0.82108       0.04   1

spin  17
R1     500.13 2.26091       0.14415       1
R2     500.13 14.997 0.66348       0
NOE    500.13 0.7567 0.0851 1
```

```
R1     599.69 1.56152       0.07144       1
R2     599.69 17.7809       0.271  1
NOE    599.69 0.83502       0.04   1

spin  18
R1     500.13 1.73641       0.13206       1
R2     500.13 16.68335      0.51213       0
NOE    500.13 0.7374 0.0858 1

R1     599.69 1.12854       0.0661 1
R2     599.69 19.305 0.2229 1
NOE    599.69 0.80462       0.04   1

spin  19
R1     500.13 2.03128       0.13905       1
R2     500.13 17.17033      2.41753       0
NOE    500.13 0.5628 0.0987 1

R1     599.69 1.28386       0.13022       1
R2     599.69 21.0615       0.9715 1
NOE    599.69 0.73456       0.04   1

spin  20
R1     500.13 1.82815       0.0996 1
R2     500.13 14.47597      1.42077       0
NOE    500.13 0.5173 0.0475 1

R1     599.69 1.24301       0.04898       1
R2     599.69 18.0440       0.7163 1
NOE    599.69 0.84483       0.04   1

spin  21
```

```
R1     500.13 2.00521       0.07599       1
R2     500.13 12.31224      0.7716 0
NOE    500.13 0.5091 0.0255 1

R1     599.69 1.34102       0.0838 1
R2     599.69 14.2878       1.2085 1
NOE    599.69 0.73377       0.04   1

spin  22
R1     500.13 1.77054       0.19812       1
R2     500.13 13.2908       0.95565       0
NOE    500.13               0

R1     599.69 1.24782       0.06259       1
R2     599.69 16.6694       0.207  1
NOE    599.69 0.83519       0.04   1

spin  23
R1     500.13 2.06313       0.26646       1
R2     500.13 14.34926      1.6225 0
NOE    500.13 0.8666 0.0867 1

R1     599.69 1.33547       0.0626 1
R2     599.69 20.6526       0.6227 1
NOE    599.69 0.83531       0.04   1

spin  24
R1     500.13 2.166  0.315  1
R2     500.13 16.824 0.631  0
NOE    500.13 0.767  0.06   1

R1     599.69 1.435  0.055  1
```

```
R2     599.69 18.282 0.1598 1
NOE    599.69              0

spin  25
R1     500.13 1.86289      0.07739      1
R2     500.13 12.8123      0.43665      0
NOE    500.13 0.7767 0.0388 1

R1     599.69 1.42126      0.0307 1
R2     599.69 16.3425      0.6276 1
NOE    599.69 0.81667      0.04   1

spin  26
R1     500.13 2.105  0.212  1
R2     500.13 16.215 0.844  0
NOE    500.13 0.750  0.0375 1

R1     599.69 1.384  0.051  1
R2     599.69 19.2753      0.2597 1
NOE    599.69              0

spin  27
R1     500.13 2.41429      0.18361      1
R2     500.13 16.55081     0.30132      0
NOE    500.13 0.7430 0.0372 1

R1     599.69 1.61838      0.06417      1
R2     599.69 17.8891      0.2375 1
NOE    599.69 0.82285      0.04   1

spin  28
R1     500.13 2.12224      0.26753      1
```

```
R2     500.13 14.99475      0.40247       0
NOE    500.13 0.7060 0.0353 1

R1     599.69 1.54727       0.02921       1
R2     599.69 16.3026       0.0747 1
NOE    599.69 0.76354       0.04   1

spin  29
R1     500.13 2.19925       0.20266       1
R2     500.13 15.62988      0.43728       0
NOE    500.13 0.5237 0.0262 1

R1     599.69 1.38083       0.05682       1
R2     599.69 16.9176       0.0407 1
NOE    599.69 0.59246       0.04   1

spin  31
R1     500.13 1.86081       0.07618       1
R2     500.13 16.31854      0.38613       0
NOE    500.13 0.6598  0.0330      1

R1     599.69 1.30293       0.02496       1
R2     599.69 17.8285       0.1329 1
NOE    599.69 0.76637       0.04   1

spin  32
R1     500.13 1.66778       0.13045       1
R2     500.13 14.02918      0.53141       0
NOE    500.13 0.6881 0.0344 1

R1     599.69 1.19062       0.02778       1
R2     599.69 16.6085       0.1291 1
```

```
NOE    599.69 0.7752 0.04   1

spin  34
R1     500.13 1.79953       0.11496       1
R2     500.13 15.23693      0.27628       0
NOE    500.13 0.7979 0.0399 1

R1     599.69 1.35575       0.02261       1
R2     599.69 16.603 0.2293 1
NOE    599.69 0.83169       0.04   1

spin  35
R1     500.13 1.70678       0.07283       1
R2     500.13 12.81723      0.44849       0
NOE    500.13 0.6884 0.0344 1

R1     599.69 1.25439       0.03289       1
R2     599.69 13.560 0.1803 1
NOE    599.69 0.8008 0.04   1

spin  37
R1     500.13 2.1575 0.21878       1
R2     500.13 13.57405      0.17836       0
NOE    500.13               0

R1     599.69 1.53421       0.04425       1
R2     599.69 17.4429       0.1792 1
NOE    599.69 0.81093       0.04   1

spin  39
R1     500.13 2.35793       0.17236       1
R2     500.13 18.37897      2.22601       0
```

```
NOE    500.13              0

R1     599.69 1.37931      0.11301      1
R2     599.69 20.6996      0.4799 1
NOE    599.69 0.79928      0.04   1

spin  42
R1     500.13 2.36967      0.25943      1
R2     500.13 15.50628     0.64439      0
NOE    500.13 0.6316 0.0332 1

R1     599.69 1.65207      0.06851      1
R2     599.69 16.3292      0.1235 1
NOE    599.69 0.70915      0.04   1

spin  44
R1     500.13 1.89681      0.16406      1
R2     500.13 14.67136     0.7792 0
NOE    500.13 0.7319 0.0366 1

R1     599.69 1.33014      0.01929      1
R2     599.69 19.5122      0.5026 1
NOE    599.69 0.79365      0.04   1

spin  45
R1     500.13 1.55618      0.15668      1
R2     500.13 15.38462     0.4355 0
NOE    500.13 0.7013 0.0351 1

R1     599.69 1.22669      0.017  1
R2     599.69 17.7179      0.1545 1
NOE    599.69 0.81651      0.04   1
```

```
spin  46
R1     500.13 1.998  0.18204      1
R2     500.13 15.29988     0.70226      0
NOE    500.13              0

R1     599.69 1.32258      0.02659      1
R2     599.69 16.835 0.3118 1
NOE    599.69 0.78187      0.04   1

spin  47
R1     500.13 2.40906      0.27567      1
R2     500.13 13.72684     0.62369      0
NOE    500.13              0

R1     599.69 1.4537 0.13969      1
R2     599.69 18.1752      0.7796 1
NOE    599.69 0.74937      0.04   1

spin  50
R1     500.13 1.4432 0.2874 1
R2     500.13 15.93625     0.85586      0
NOE    500.13              0

R1     599.69 1.6197 0.08369      1
R2     599.69 16.1812      0.0903 1
NOE    599.69 0.73241      0.04   1

spin  51
R1     500.13 1.74642      0.24064      1
R2     500.13 19.85703     3.26876      0
NOE    500.13              0
```

```
R1     599.69 1.49142       0.10032       1
R2     599.69 17.001 0.2075 1
NOE    599.69 0.83492       0.04   1

spin  54
R1     500.13 1.83251       0.15749       1
R2     500.13 15.70845      0.58234       0
NOE    500.13 0.6490 0.0325 1

R1     599.69 1.32468       0.04176       1
R2     599.69 18.0278       0.1827 1
NOE    599.69 0.77851       0.04   1

spin  55
R1     500.13 1.93013       0.15051       1
R2     500.13 12.98701      0.79609       0
NOE    500.13 0.7824 0.0391 1

R1     599.69 1.25976       0.05348       1
R2     599.69 16.4447       0.2839 1
NOE    599.69 0.76378       0.04   1

spin  56
R1     500.13 1.8005 0.15528       1
R2     500.13 14.14427      0.58018       0
NOE    500.13 0.7599 0.0382 1

R1     599.69 1.37874       0.03802       1
R2     599.69 16.5645       0.1306 1
NOE    599.69 0.80346       0.04   1
```

```
spin  57
R1     500.13 1.85632       0.20538       1
R2     500.13 12.00048      1.03256       0
NOE    500.13 0.6583 0.0461 1

R1     599.69 1.31839       0.01392       1
R2     599.69 16.2232       0.1995 1
NOE    599.69 0.82224       0.04   1

spin  58
R1     500.13 1.77715       0.02767       1
R2     500.13 12.65182      0.62267       0
NOE    500.13 0.7655 0.0491 1

R1     599.69 1.19646       0.06156       1
R2     599.69 16.7392       0.1625 1
NOE    599.69 0.80214       0.04   1

spin  59
R1     500.13 1.84162       0.07733       1
R2     500.13 14.6263       0.43   0
NOE    500.13 0.7834 0.0392 1

R1     599.69 1.3541 0.01482       1
R2     599.69 17.0358       0.1352 1
NOE    599.69 0.83394       0.04   1

spin  60
R1     500.13 1.85632       0.06995       1
R2     500.13 14.55816      0.35606       0
NOE    500.13 0.6921 0.0474 1
```

```
R1     599.69 1.34138       0.02249       1
R2     599.69 16.3345       0.2041 1
NOE    599.69 0.77975       0.04   1

spin  61
R1     500.13 1.89538       0.08227       1
R2     500.13 13.77031      0.24461       0
NOE    500.13 0.7140 0.0419 1

R1     599.69 1.36949       0.02513       1
R2     599.69 15.9286       0.1489 1
NOE    599.69 0.81933       0.04   1

spin  62
R1     500.13 2.04165       0.11755       1
R2     500.13 15.50628      0.22289       0
NOE    500.13 0.6881 0.0344 1

R1     599.69 1.44467       0.03819       1
R2     599.69 16.9005       0.1977 1
NOE    599.69 0.82257       0.04   1

spin  63
R1     500.13 1.74459       0.20118       1
R2     500.13 14.54545      0.2835 0
NOE    500.13 0.6480 0.0324 1

R1     599.69 1.19875       0.05777       1
R2     599.69 16.3720       0.0839 1
NOE    599.69 0.72444       0.04   1

spin  64
```

```
R1     500.13 1.67112       0.06451       1
R2     500.13 14.77978      0.72523       0
NOE    500.13 0.7352 0.0368 1

R1     599.69 1.30005       0.0262 1
R2     599.69 16.787 0.1764 1
NOE    599.69 0.78776       0.04   1

spin  65
R1     500.13 1.72652       0.10254       1
R2     500.13 13.22751      0.19071       0
NOE    500.13 0.6979 0.0349 1

R1     599.69 1.27649       0.03454       1
R2     599.69 17.5408       0.1674 1
NOE    599.69 0.79174       0.04   1

spin  66
R1     500.13 1.84877       0.16406       1
R2     500.13 14.11034      0.48183       0
NOE    500.13 0.7278 0.0494 1

R1     599.69 1.3132 0.0269 1
R2     599.69 16.1891       0.2456 1
NOE    599.69 0.81174       0.04   1

spin  67
R1     500.13 1.8018 0.03928       1
R2     500.13 13.97429      0.47063       0
NOE    500.13 0.7903 0.0395 1

R1     599.69 1.31251       0.03032       1
```

```
R2     599.69 16.1473       0.2287 1
NOE    599.69 0.80794       0.04   1

spin  68
R1     500.13 1.79662       0.14493       1
R2     500.13 14.3947       0.64856       0
NOE    500.13               0

R1     599.69 1.33103       0.01772       1
R2     599.69 17.0999       0.3421 1
NOE    599.69 0.80962       0.04   1

spin  69
R1     500.13 1.65782       0.09427       1
R2     500.13 13.44447      0.60733       0
NOE    500.13 0.6972 0.0686 1

R1     599.69 1.36649       0.01475       1
R2     599.69 15.4847       0.1957 1
NOE    599.69 0.82788       0.04   1

spin  70
R1     500.13 1.93013       0.12219       1
R2     500.13 12.51408      0.3226 0
NOE    500.13 0.5726 0.0286 1

R1     599.69 1.39762       0.03438       1
R2     599.69 17.5101       0.1420 1
NOE    599.69 0.78906       0.04   1

spin  71
R1     500.13 2.15517       0.09475       1
```

```
R2     500.13 14.69724      0.81435       0
NOE    500.13 0.8173 0.0409 1

R1     599.69 1.49477       0.03463       1
R2     599.69 16.3452       0.2207 1
NOE    599.69 0.85246       0.04   1

spin  72
R1     500.13 1.76118       0.15416       1
R2     500.13 12.99376      0.4069 0
NOE    500.13 0.6367 0.0318 1

R1     599.69 1.30993       0.03655       1
R2     599.69 13.7646       0.0926 1
NOE    599.69 0.70935       0.04   1

spin  73
R1     500.13 1.76429       0.18085       1
R2     500.13 14.29184      0.29821       0
NOE    500.13 0.7685 0.0384 1

R1     599.69 1.35943       0.01977       1
R2     599.69 18.1686       0.2205 1
NOE    599.69 0.77432       0.04   1

spin  75
R1     500.13 2.85225       0.57029       1
R2     500.13 14.40507      1.40067       0
NOE    500.13               0

R1     599.69 1.96928       0.18382       1
R2     599.69 18.3318       0.5511 1
```

```
NOE    599.69 0.85332      0.04   1

spin  76
R1     500.13 2.25175      0.26518      1
R2     500.13 21.36296     1.50604      0
NOE    500.13              0

R1     599.69 1.56961      0.09116      1
R2     599.69 17.9727      0.2374 1
NOE    599.69 0.86808      0.04   1

spin  77
R1     500.13 2.42895      0.17758      1
R2     500.13 14.97903     0.50932      0
NOE    500.13 0.7788 0.0389 1

R1     599.69 1.59566      0.04379      1
R2     599.69 17.3762      0.2518 1
NOE    599.69 0.83189      0.04   1

spin  78
R1     500.13 2.15657      0.18185      1
R2     500.13 15.37752     0.41145      0
NOE    500.13 0.7774 0.0579 1

R1     599.69 1.49589      0.03647      1
R2     599.69 16.3559      0.1862 1
NOE    599.69 0.80571      0.04   1

spin  79
R1     500.13 1.7762 0.16405      1
R2     500.13 15.59089     0.26495      0
```

```
NOE    500.13 0.7893 0.0395 1

R1     599.69 1.41924       0.04834       1
R2     599.69 17.4125       0.3942 1
NOE    599.69 0.82753       0.04   1

spin  80
R1     500.13 1.90223       0.20517       1
R2     500.13 14.2288       0.45958       0
NOE    500.13 0.7159 0.0358 1

R1     599.69 1.41303       0.03394       1
R2     599.69 17.2206       0.113  1
NOE    599.69 0.829        0.04    1

spin  81
R1     500.13 2.02347       0.23052       1
R2     500.13 15.72574      0.39568       0
NOE    500.13 0.8046 0.0402 1

R1     599.69 1.37608       0.05018       1
R2     599.69 17.0184       0.2349 1
NOE    599.69 0.82138       0.04   1

spin  82
R1     500.13 2.65604       0.2709 1
R2     500.13 24.2483       2.26372       0
NOE    500.13 0.6986 0.0737 1

R1     599.69 1.78221       0.09974       1
R2     599.69 31.5159       0.8463 1
NOE    599.69 0.8502 0.04   1
```

```
spin  83
R1     500.13 2.02429       0.11802       1
R2     500.13 15.57147      0.53101       0
NOE    500.13 0.07049       0.0352 0

R1     599.69 1.4041 0.03687       1
R2     599.69 19.2308       0.1268 1
NOE    599.69 0.80472       0.04   1

spin  84
R1     500.13 1.6412 0.346  1
R2     500.13 16.70844      0.25014       0
NOE    500.13 0.7812 0.0391 1

R1     599.69 1.70794       0.10268       1
R2     599.69 17.2741       0.2348 1
NOE    599.69 0.8372 0.04   1

spin  85
R1     500.13 2.08986       0.13234       1
R2     500.13 14.88982      0.70059       0
NOE    500.13 0.6728 0.0601 1

R1     599.69 1.53186       0.04599       1
R2     599.69 17.3130       0.2032 1
NOE    599.69 0.84154       0.04   1

spin  86
R1     500.13 2.20751       0.08918       1
R2     500.13 15.6789       0.4892 0
NOE    500.13 0.5233 0.0262 1
```

```
R1     599.69 1.51538       0.05442       1
R2     599.69 16.9952       0.1505 1
NOE    599.69 0.82364       0.04   1

spin  87
R1     500.13 1.5878 0.2546 1
R2     500.13 16.6917       0.12705       0
NOE    500.13 0.7562 0.0378 1

R1     599.69 1.60694       0.05113       1
R2     599.69 16.2681       0.1879 1
NOE    599.69 0.82782       0.04   1

spin  89
R1     500.13 1.79888       0.0974 1
R2     500.13 15.93372      0.55093       0
NOE    500.13               0

R1     599.69 1.5444 0.06726       1
R2     599.69 16.2206       0.1813 1
NOE    599.69 0.73848       0.04   1

spin  91
R1     500.13 1.76772       0.15874       1
R2     500.13 14.27756      0.71755       0
NOE    500.13 0.4025 0.0201 1

R1     599.69 1.38754       0.03388       1
R2     599.69 15.0263       0.1113       1
NOE    599.69 0.5158 0.04   1
```

```
spin  92
R1     500.13 2.35239       0.18095       1
R2     500.13 14.70805      0.29853       0
NOE    500.13 0.4080 0.0829 1

R1     599.69 1.57011       0.06434       1
R2     599.69 15.5376       0.1919 1
NOE    599.69 0.76359       0.04   1

spin  93
R1     500.13 2.31642       0.17385       1
R2     500.13 19.72387      0.70026       0
NOE    500.13               0

R1     599.69 1.76056       0.04339       1
R2     599.69 17.2622       0.1728 1
NOE    599.69 0.77442       0.04   1

spin  94
R1     500.13 2.0247 0.11109       1
R2     500.13 15.2045       0.54327       0
NOE    500.13 0.6187 0.0470 1

R1     599.69 1.44113       0.05649       1
R2     599.69 16.6611       0.1274 1
NOE    599.69 0.75149       0.04   1

spin  95
R1     500.13 1.80668       0.07834       1
R2     500.13 15.18834      0.8743 0
NOE    500.13 0.7151 0.0677 1
```

```
R1     599.69 1.40746       0.05784       1
R2     599.69 17.1880       0.3781 1
NOE    599.69 0.78508       0.04   1

spin  96
R1     500.13 2.10393       0.22664       1
R2     500.13 14.99475      0.37774       0
NOE    500.13 0.7572 0.0379 1

R1     599.69 1.39821       0.05298       1
R2     599.69 17.6211       0.2257 1
NOE    599.69 0.82012       0.04   1

spin  97
R1     500.13 1.96078       0.25106       1
R2     500.13 14.77323      0.42995       0
NOE    500.13 0.6695 0.0378 1

R1     599.69 1.27828       0.06716       1
R2     599.69 14.9165       0.144  1
NOE    599.69 0.81117       0.04   1

spin  99
R1     500.13 1.83385       0.1167 1
R2     500.13 14.06074      0.30249       0
NOE    500.13 0.7634 0.0620 1

R1     599.69 1.32118       0.02636       1
R2     599.69 16.57  0.1076 1
NOE    599.69 0.8073 0.04   1

spin  100
```

```
R1     500.13 1.68947       0.10361       1
R2     500.13 12.66464      0.14836       0
NOE    500.13 0.5826 0.0615 1

R1     599.69 1.19076       0.04977       1
R2     599.69 14.3781       0.1714 1
NOE    599.69 0.7153 0.04   1

spin  101
R1     500.13 1.82548       0.1033 1
R2     500.13 13.90241      0.44647       0
NOE    500.13 0.5544 0.0639 1

R1     599.69 1.24517       0.02171       1
R2     599.69 16.1760       0.1198 1
NOE    599.69 0.81881       0.04   1

spin  102
R1     500.13 1.53941       0.20214       1
R2     500.13 15.71586      0.46928       0
NOE    500.13 0.7641 0.0484 1

R1     599.69 1.13404       0.02816       1
R2     599.69 17.3702       0.2972 1
NOE    599.69 0.77044       0.04   1

spin  103
R1     500.13 1.55183       0.08405       1
R2     500.13 14.66921      0.72087       0
NOE    500.13 0.7698 0.06   1

R1     599.69 1.18371       0.03139       1
```

```
R2     599.69 15.6421       0.1749 1
NOE    599.69 0.80697       0.04   1

spin  104
R1     500.13 2.26757       0.23961       1
R2     500.13 14.39263      0.67737       0
NOE    500.13               0

R1     599.69 1.50128       0.07438       1
R2     599.69 18.6881       0.709  1
NOE    599.69 0.84103       0.04   1

spin  105
R1     500.13 1.81984       0.16427       1
R2     500.13 15.7679       0.53206       0
NOE    500.13 0.8258 0.0413 1

R1     599.69 1.25549       0.03247       1
R2     599.69 19.3874       0.1601 1
NOE    599.69 0.81041       0.04   1

spin  106
R1     500.13 1.80018       0.10208       1
R2     500.13 13.88503      0.43379       0
NOE    500.13 0.7364 0.0658 1

R1     599.69 1.37268       0.02939       1
R2     599.69 16.0901       0.3029 1
NOE    599.69 0.79785       0.04

spin  107
R1     500.13 1.87935       0.20803       1
```

```
R2     500.13 13.34401      1.28383       0
NOE    500.13 0.8275 0.0863 1

R1     599.69 1.31614       0.04296       1
R2     599.69 19.2123       0.4872 1
NOE    599.69 0.85944       0.04   1

spin  108
R1     500.13 1.89143       0.18567       1
R2     500.13 14.54122      0.43558       0
NOE    500.13 0.7972 0.0399 1

R1     599.69 1.38236       0.02561       1
R2     599.69 16.5453       0.2929 1
NOE    599.69 0.81637       0.04   1

spin  109
R1     500.13 2.03293       0.16159       1
R2     500.13 14.50326      0.63314       0
NOE    500.13               0

R1     599.69 1.41723       0.02973       1
R2     599.69 15.9413       0.2821 1
NOE    599.69 0.82437       0.04   1

spin  110
R1     500.13 1.96425       0.08951       1
R2     500.13 15.64945      0.46777       0
NOE    500.13 0.7201 0.0411 1

R1     599.69 1.46199       0.06434       1
R2     599.69 16.4636       0.3930 1
```

```
NOE    599.69 0.80583      0.04   1

spin  112
R1     500.13 2.38607      0.22375      1
R2     500.13 17.31302     1.02811      0
NOE    500.13              0

R1     599.69 1.81686      0.06866      1
R2     599.69 17.4004      0.4542 1
NOE    599.69 0.77127      0.04   1

spin  114
R1     500.13 1.78987      0.08458      1
R2     500.13 13.34935     0.8643 0
NOE    500.13 0.7749 0.0871 1

R1     599.69 1.41623      0.02708      1
R2     599.69 16.2311      0.598  1
NOE    599.69 0.82034      0.04   1

spin  115
R1     500.13 1.93648      0.2355 1
R2     500.13 13.25381     0.74833      0
NOE    500.13 0.7819 0.0722 1

R1     599.69 1.14247      0.06265      1
R2     599.69 18.1752      0.3997 1
NOE    599.69 0.81957      0.04   1

spin  116
R1     500.13 1.94363      0.28106      1
R2     500.13 14.69292     0.58504      0
```

```
NOE    500.13              0

R1     599.69 1.374  0.0489 1
R2     599.69 19.8689      0.4303 1
NOE    599.69 0.82407      0.04   1

spin  117
R1     500.13 2.19346      0.20255      1
R2     500.13 16.56178     1.8405 0
NOE    500.13 0.6928 0.0679 1

R1     599.69 1.32961      0.08609      1
R2     599.69 26.1780      1.8366 1
NOE    599.69 0.83451      0.04   1

spin  118
R1     500.13 2.00803      0.13669      1
R2     500.13 13.41922     1.23532      0
NOE    500.13 0.7183 0.0550 1

R1     599.69 1.40528      0.08452      1
R2     599.69 17.4611      0.6250 1
NOE    599.69 0.80548      0.04   1

spin  120
R1     500.13 1.77368      0.11829      1
R2     500.13 14.54545     0.58393      0
NOE    500.13 0.7140 0.0487 1

R1     599.69 1.35263      0.05031      1
R2     599.69 16.4474      0.3544 1
NOE    599.69 0.7892 0.04   1
```

```
spin  121
R1     500.13 1.4096 0.3636 1
R2     500.13 15.2765       0.39206       0
NOE    500.13 0.6393 0.0797 1

R1     599.69 1.83857       0.08417       1
R2     599.69 14.5645       0.1854 1
NOE    599.69 0.70787       0.04   1

spin  122
R1     500.13 2.54907       0.18779       1
R2     500.13 11.36235      0.40538       0
NOE    500.13 0.3215 0.0606 1

R1     599.69 1.77274       0.03897       1
R2     599.69 14.5328       0.12   1
NOE    599.69 0.61282       0.04   1

spin  123
R1     500.13 2.79174       0.22212       1
R2     500.13 9.92063       0.15058       0
NOE    500.13 0.3172 0.0416 1

R1     599.69 1.88288       0.07622       1
R2     599.69 10.0725       0.2242 1
NOE    599.69 0.729        0.04    1

spin  124
R1     500.13 1.50784       0.113  1
R2     500.13 4.44444       0.17798       0
NOE    500.13               0
```

```
R1     599.69 1.37722      0.022  1
R2     599.69 4.7416 0.1239 1
NOE    599.69 0.00489      0.04   1
```

The molecular parameter file (mfpar)

```
spin  3
constants     3     N15     -2.710     1.010     -170.00
vector N H

spin  4
constants     4     N15     -2.710     1.010     -170.00
vector N H

spin  5
constants     5     N15     -2.710     1.010     -170.00
vector N H

spin  6
constants     6     N15     -2.710     1.010     -170.00
vector N H

spin  7
constants     7     N15     -2.710     1.010     -170.00
vector N H

spin  8
constants     8     N15     -2.710     1.010     -170.00
vector N H
```

```
spin  9
constants     9     N15     -2.710     1.010     -170.00
vector N H

spin  10
constants    10     N15     -2.710     1.010     -170.00
vector N H

spin  11
constants    11     N15     -2.710     1.010     -170.00
vector N H

spin  12
constants    12     N15     -2.710     1.010     -170.00
vector N H

spin  14
constants    14     N15     -2.710     1.010     -170.00
vector N H

spin  15
constants    15     N15     -2.710     1.010     -170.00
vector N H

spin  17
constants    17     N15     -2.710     1.010     -170.00
vector N H

spin  18
constants    18     N15     -2.710     1.010     -170.00
vector N H

spin  19
constants    19     N15     -2.710     1.010     -170.00
```

```
vector N H

spin  20
constants    20     N15     -2.710     1.010     -170.00
vector N H

spin  21
constants    21     N15     -2.710     1.010     -170.00
vector N H

spin  22
constants    22     N15     -2.710     1.010     -170.00
vector N H

spin  23
constants    23     N15     -2.710     1.010     -170.00
vector N H

spin  24
constants    24     N15     -2.710     1.010     -170.00
vector N H

spin  25
constants    25     N15     -2.710     1.010     -170.00
vector N H

spin  26
constants    26     N15     -2.710     1.010     -170.00
vector N H

spin  27
constants    27     N15     -2.710     1.010     -170.00
vector N H
```

```
spin  28
constants    28     N15     -2.710     1.010     -170.00
vector N H

spin  29
constants    29     N15     -2.710     1.010     -170.00
vector N H

spin  30
constants    30     N15     -2.710     1.010     -170.00
vector N H

spin  31
constants    31     N15     -2.710     1.010     -170.00
vector N H

spin  32
constants    32     N15     -2.710     1.010     -170.00
vector N H

spin  34
constants    34     N15     -2.710     1.010     -170.00
vector N H

spin  35
constants    35     N15     -2.710     1.010     -170.00
vector N H

spin  36
constants    36     N15     -2.710     1.010     -170.00
vector N H

spin  37
constants    37     N15     -2.710     1.010     -170.00
```

```
vector N H

spin  38
constants    38     N15     -2.710     1.010     -170.00
vector N H

spin  39
constants    39     N15     -2.710     1.010     -170.00
vector N H

spin  42
constants    42     N15     -2.710     1.010     -170.00
vector N H

spin  43
constants    43     N15     -2.710     1.010     -170.00
vector N H

spin  44
constants    44     N15     -2.710     1.010     -170.00
vector N H

spin  45
constants    45     N15     -2.710     1.010     -170.00
vector N H

spin  46
constants    46     N15     -2.710     1.010     -170.00
vector N H

spin  47
constants    47     N15     -2.710     1.010     -170.00
vector N H
```

```
spin  48
constants    48     N15     -2.710     1.010     -170.00
vector N H

spin  49
constants    49     N15     -2.710     1.010     -170.00
vector N H

spin  50
constants    50     N15     -2.710     1.010     -170.00
vector N H

spin  51
constants    51     N15     -2.710     1.010     -170.00
vector N H

spin  53
constants    53     N15     -2.710     1.010     -170.00
vector N H

spin  54
constants    54     N15     -2.710     1.010     -170.00
vector N H

spin  55
constants    55     N15     -2.710     1.010     -170.00
vector N H

spin  56
constants    56     N15     -2.710     1.010     -170.00
vector N H

spin  57
constants    57     N15     -2.710     1.010     -170.00
```

```
vector N H

spin  58
constants    58     N15     -2.710     1.010     -170.00
vector N H

spin  59
constants    59     N15     -2.710     1.010     -170.00
vector N H

spin  60
constants    60     N15     -2.710     1.010     -170.00
vector N H

spin  61
constants    61     N15     -2.710     1.010     -170.00
vector N H

spin  62
constants    62     N15     -2.710     1.010     -170.00
vector N H

spin  63
constants    63     N15     -2.710     1.010     -170.00
vector N H

spin  64
constants    64     N15     -2.710     1.010     -170.00
vector N H

spin  65
constants    65     N15     -2.710     1.010     -170.00
vector N H
```

```
spin  66
constants    66     N15     -2.710     1.010     -170.00
vector N H

spin  67
constants    67     N15     -2.710     1.010     -170.00
vector N H

spin  68
constants    68     N15     -2.710     1.010     -170.00
vector N H

spin  69
constants    69     N15     -2.710     1.010     -170.00
vector N H

spin  70
constants    70     N15     -2.710     1.010     -170.00
vector N H

spin  71
constants    71     N15     -2.710     1.010     -170.00
vector N H

spin  72
constants    72     N15     -2.710     1.010     -170.00
vector N H

spin  73
constants    73     N15     -2.710     1.010     -170.00
vector N H

spin  74
constants    74     N15     -2.710     1.010     -170.00
```

```
vector N H

spin  75
constants    75     N15     -2.710     1.010     -170.00
vector N H

spin  76
constants    76     N15     -2.710     1.010     -170.00
vector N H

spin  77
constants    77     N15     -2.710     1.010     -170.00
vector N H


spin  78
constants    78     N15     -2.710     1.010     -170.00
vector N H

spin  79
constants    79     N15     -2.710     1.010     -170.00
vector N H

spin  80
constants    80     N15     -2.710     1.010     -170.00
vector N H

spin  81
constants    81     N15     -2.710     1.010     -170.00
vector N H

spin  82
constants    82     N15     -2.710     1.010     -170.00
vector N H
```

```
spin  83
constants    83     N15     -2.710     1.010     -170.00
vector N H

spin  84
constants    84     N15     -2.710     1.010     -170.00
vector N H

spin  85
constants    85     N15     -2.710     1.010     -170.00
vector N H

spin  86
constants    86     N15     -2.710     1.010     -170.00
vector N H

spin  87
constants    87     N15     -2.710     1.010     -170.00
vector N H

spin  89
constants    89     N15     -2.710     1.010     -170.00
vector N H

spin  90
constants    90     N15     -2.710     1.010     -170.00
vector N H

spin  91
constants    91     N15     -2.710     1.010     -170.00
vector N H

spin  92
```

```
constants    92     N15     -2.710     1.010     -170.00
vector N H

spin  93
constants    93     N15     -2.710     1.010     -170.00
vector N H

spin  94
constants    94     N15     -2.710     1.010     -170.00
vector N H

spin  95
constants    95     N15     -2.710     1.010     -170.00
vector N H

spin  96
constants    96     N15     -2.710     1.010     -170.00
vector N H

spin  97
constants    97     N15     -2.710     1.010     -170.00
vector N H

spin  99
constants    99     N15     -2.710     1.010     -170.00
vector N H

spin  100
constants   100     N15     -2.710     1.010     -170.00
vector N H

spin  101
constants   101     N15     -2.710     1.010     -170.00
vector N H
```

```
spin  102
constants   102     N15     -2.710     1.010     -170.00
vector N H

spin  103
constants   103     N15     -2.710     1.010     -170.00
vector N H

spin  104
constants   104     N15     -2.710     1.010     -170.00
vector N H

spin  105
constants   105     N15     -2.710     1.010     -170.00
vector N H

spin  106
constants   106     N15     -2.710     1.010     -170.00
vector N H

spin  107
constants   107     N15     -2.710     1.010     -170.00
vector N H

spin  108
constants   108     N15     -2.710     1.010     -170.00
vector N H

spin  109
constants   109     N15     -2.710     1.010     -170.00
vector N H

spin  110
```

```
constants   110     N15     -2.710     1.010     -170.00
vector N H

spin  111
constants   111     N15     -2.710     1.010     -170.00
vector N H


spin  112
constants   112     N15     -2.710     1.010     -170.00
vector N H

spin  113
constants   113     N15     -2.710     1.010     -170.00
vector N H

spin  114
constants   114     N15     -2.710     1.010     -170.00
vector N H

spin  115
constants   115     N15     -2.710     1.010     -170.00
vector N H

spin  116
constants   116     N15     -2.710     1.010     -170.00
vector N H

spin  117
constants   117     N15     -2.710     1.010     -170.00
vector N H

spin  118
constants   118     N15     -2.710     1.010     -170.00
```

```
vector N H

spin  119
constants   119     N15     -2.710     1.010     -170.00
vector N H

spin  120
constants   120     N15     -2.710     1.010     -170.00
vector N H

spin  121
constants   121     N15     -2.710     1.010     -170.00
vector N H

spin  122
constants   122     N15     -2.710     1.010     -170.00
vector N H

spin  123
constants   123     N15     -2.710     1.010     -170.00
vector N H

spin  124
constants   124     N15     -2.710     1.010     -170.00
vector N H
```

PDB file of KI-FHA

Please refer to www.pdb.org, entry number "1MZK"

APPENDIX 4. Sequences of FHA domains for evolutionary trace analysis

209 FHA domain sequences were obtained from http://www.sanger.ac.uk/cgi-bin/Pfam/getacc?PF00498#, and listed in FastA format as below.

```
>Q9Z7J3/2590
WSIGRDSSANDIPIEDPKLGASQAIINKTDGSYYITNLDDTIPIVVNGVAIQETTQLKNEDTILLG
>Q8DKX4/45108
VRVGRSLHNKIVISYPTISRRHFEIVKTDTGYRIRDLDSKLGLVYGDRPIKEKDLKDKDIFYIG
>CHK2_MOUSE/117196
YWFGRDKSCEYCFDGPLLRRTDKYRTYSKKHFRIFREMGPKNCYIVYIEDHSGNGTFVNTELIGKGKRCPLSNNSEIALS
>Q9R019/116195
YWFGRDKSCEYCFDGPLLKRTDKYRTYSKKHFRIFREMGPKNCYIVYLEDHSGNGTFVNTELIGKGKRCPLSNNSEIALS
>O50389/2689
VVVGSDLRADMRVAHPLIARAHLLLRFDRGNWIAIDNDSQSGMFVDGQRVSEVDIYDGLTINIG
>CHK2_HUMAN/113192
YWFGRDKSCEYCFDEPLLKRTDKYRTYSKKHFRIFREVGPKNSYIAYIEDHSGNGTFVNTELVGKGKRRPLNNNSEIALS
>Q9I8V3/85164
YVFGRDKKCDYTFDIPVLNQTDRYKTYSKRHFRIFQELGHGHSRVANIEDLSGNGTFVNKEIIGKGRTLPLTNNAEIALS
>Q98TW0/85164
YVFGRDKKCDYTFDIPVLNQTDRYKTYSKRHFRIFQELGHGHSRVAYIEDLSGNGTFVNKEIIGKGRTLPLTNNAEIALS
>Q90ZY5/91165
YSFGRDKRCDYSFSNSILKKSPYFNTYSKKHFRIFRDENLVYLEDLSGNGTWVDDEKLGNGKQSLLSNNSVIALA
>Q9XTX3/65143
FVCGRGSHDAPTNFNFSTVAEDVGLYKFISKIQFSIDRDTETRRIYLHDHSRNGTLVNHEMIGKGLSRELMNGDLISIA
```

>Q9GR99/40118

FVCGRGSDDAPTNFNFSQVAKDVGLYRFISKIQFSIDRDTETRRIYLHDHSRNGTLVNQEMIGKGLSRELMNGDLISIG

>LOK_DROME/69146

FTAGRGEANDLILTLNDLPEKILTRISKVHFIIKRANCELTNPVYIQDLSRNGTFVNNEKIGTNRMRILKNDDVISLS

>CDS1_SCHPO/60133

WRFGRHKSCEVVLNGPRVSNFHFEIYQGHRNDSDESENVVFLHDHSSNGTFLNFERLAKNSRTILSNGDEIRIG

>FHA1-Rad53_SPK1_YEAST/66133

WTFGRNPACDYHLGNISRLSNKHFQILLGEDGNLLLNDISTNGTWLNGQKVEKNSNQLLSQGDEITVG

>MEK1_SCHPO/62133

VSVGRSNTCNYQLLQFTASYKHFRVYSVLIDDDMDPLVYCEDQSSNGTFLNHRLIGKGNSVLLSDGDILDVR

>Q9XS53/2792

CLFGRGTECDIRLQLPVVSKQQCKIEFNEQLEAILINLNSANPTQLNGATFYHPVKLEHGDLITIF

>Ki-67 Mouse_Q61769/2791

CLFGRSIECDIRIQLPVVSQRHCPIVVQEQEAILYNFSSTNPTQVNGVTIDEPVRLRHGDIITII

>KI67_HUMAN/2791

CLFGRGIECDIRIQLPVVSKQHCKIEIHEQEAILHNFSSTNPTQVNGSVIDEPVRLKHGDVITII

>MDC1 Human_Q14676/54-124

NVVGRMPDCSVALPFPSISKQHAEIEILAWDKAPILRDCGSLNGTQILRPPKVLSPGVSHRLRDQELILFA

>Q9XA21/217279

LVMGRSTEADVRIDDPGVSRRHCEIRTGTPSTIQDLGSTNGIVVDGQHTTRATLRDGSRIVVG

>Q8G6Q5/156219

TVIGRGSGCDIVIDDPGISRRHLEIDITDNGVIARDLGSTNGTYVEGHQVPAATLLDGNTITIG

>Q50190/391454

NIVGRGQDAQFRLPDTGVSRRHLEIRWDGQVALLSDLNSTNGTTVNNAPVQEWQLADGDVIRLG

>P71590/455518

NIIGRGQDAQFRLPDTGVSRRHLEIRWDGQVALLADLNSTNGTTVNNAPVQEWQLADGDVIRLG

>Q8FUH9/224287

NIIGRSNDADLRLPDTGVSRQHAEITWDGRDAILVDLKSTNGTTVNDTPVENWLLADGDVITVG

>Q8NU92/214277

NIIGRSNDADLRLPDTGVSRQHVEITWDGRDAILVDLKSTNGTTVNDTPVDNWLLADGDVITVG

>Q8D6T0/170

MKFGRSESCDWVLPDPERIISGVHGEITKFGNDYLLRDLSTNGIFVNKSVSPVGNGVEVALNDKDVINFG

>EspA_Myxococcus xanthus_Q9S434/62126

HIIGRGSDVTVRIDDHGVSRKHARVVRAGDGACHVTDLDSTNGTLLNGVPVSTAELMEGDRLQIG

>Q9L060/130196

IRLGRSADADVALDDPDVSRMHCAVTVGPDARVSVADLGSTNGTTLDGTRVGDRPVRFPPGALLRIG

>Q9RP36/76139

TSAGRHPDSDIFLDDVTVSRRHAEFRLEGGEFQVVDVGSLNGTYVNREPVDSAVLANGDEVQIG

>YI27_MYCTU/77140

TSAGRHPDSDIFLDDVTVSRRHAEFRLENNEFNVVDVGSLNGTYVNREPVDSAVLANGDEVQIG

>O32919/77140

TSAGRHPDSDIFLDDVTVSRRHAEFRLEGNEFHVVDVGSLNGTYVNREPVDSAVLANGDEVQIG

>Q8G9M6/74137

TSAGRHPDSDIFLDDVTVSRRHAEFRQDEDDFQVVDVGSLNGTYVNREPVDSAVLANGDEVQIG

>Q8FTJ5/67130

TTAGRHPESDIFLDDVTVSRRHAEFRINEGEFEVVDVGSLNGTYVNREPRNSQVLQTGDEIQIG

>Q8NQJ3/68131

TTAGRHPESDIFLDDVTVSRRHAEFRINEGEFEVVDVGSLNGTYVNREPRNAQVMQTGDEIQIG

>Q9KZP8/190254

TTAGRHPQSDIFLDDVTVSRRHVEFRRSPDGSFTVADVGSLNGTYVNRERIDQVALSNGDEVQIG

>Q8G6A0/69132

ITVGRDPRADILLDDSTVSRQHAVFRRVNGQFFVVDAGSLNGTYVNRQRVDQATLKNGDEIMIG

>KIFHA_KAPP_Arabidopsis/208-284

VKLGRVSPSDLALKDSEVSGKHAQITWNSTKFKWELVDMGSLNGTLVNSHSISHPDLGSRKWGNPVELASDDIITLG

>O81444/208-284_KAPP_Rice

ITLGRVPPSDLVLKDSEVSGKHARINWNAKTLKWEIVDMGSLNGTFVNSRAVHHPNVGSRHWGEPAELADGDIITLG

>O49973/202278_KAPP_Maize

ITLGRVPPSDLVLKDSEVSGKHAQINWNGKTLKWELVDMGSLNGTFLNSQAVHHPSAGSRHWGEPAELAHGDIITLG

>Q8DJ88/2591

LKVGRAPDNDIVLNDISVSRHHAFFEWREGQVHLVDLGSKAGTHLNGQAVVPNTPIPLEDGDLITLG

>Q9ADI4/110174

LRIGRDPASGLRLSHETVSRVHAELSRQGGMWVLRDLGSTNGTTVNGRRVIGAAVVHEGDQIGFG

>Q8DJ88/260324

FTIGRDPNNDLVIGHPTVSRHHAKIERRNGDFLLTDLGSSNGTFVNGREVEEPTLLRVGDSIRIG

>Q8DJ99/248314

VRIGRDPSNDMVLDHPVVSRFHARLYLQEGQWYLVDLESANGTFVNNRRLEPRKPVVLPTGALVRIG

>Q8DJH2/151219

LTLGRQASNSYRHWQLDAPTVSLHHASIDRTSAGLYVLRDLGSSNGTYVNAKLLKGPYTLRNGDQIRIG

>Q8YW88/35108

VVIGRDPSCQVVLDAMMYRMVSRRHAVVRPVASSVDSKFSWVLCDLNSANGTYLNGQRLYGCQELHAGDRISLG

>P94184/2489

FTIGRLPECNLYLPFAGVSRKHAQLVKKADGKWIIEDLGSKNGTQVNQSIVSHPRQLQHGDVIWLG

>Q8DIP4/36108

TIIGRDPNTCHIVLDAHIYTSVSRHHAQLTCQKRGAIPVWAIADLGSVNGTYVNQQRVETLTVLNAGDRIQLG

>Q9AJX1/219286

MRIGRALENDLVVSDLQVSRNHAEFHSTPDGRMEIRDLGSHNGTYVNGQPIAKGGTQLLGPTDVVGVG

>O65934/230293

VRIGRANDNDIVIPEVLASRHHATLVPTPGGTEIRDNRSINGTFVNGARVDAALLHDGDVVTIG

>YC67_MYCTU/308372

TRIGRLHDNDIVLDSANVSRHHAVIVDTGTNYVINDLRSSNGVHVQHERIRSAVTLNDGDHIRIC

>P71484/309373

TRIGRLDDNEIVLDDTEVSRHHAVIIDTGTDFLITDMKSTNGVQVRGRRVQRSATLVDGDHIRIG

>O53145/302366

LRIGRSKSNDMVLPDGKVSPYHAVIVNTGESFMITDLRSVNGVYVRGRRIATTATLNDGDHIRIG

>Q98IK6/2894

FEIGRENCDWTLSDPDKFISGRHCEIRYQAGTFWLHDVSRNGTFVNGSSQRMNAPHRLTQGDRLLIG

>DUN1_YEAST/56128

TTIGRSRSCDVILSEPDISTFHAEFHLLQMDVDNFQRNLINVIDKSRNGTFINGNRLVKKDYILKNGDRIVFG

>P73823/2994

WTVGRSQDNDLVIRDNCISRNHAILQATEEDTFLLIDLGSRNGTFVNGRRVSVPIAIQDQDKITFG

>Q8DGB0/92157

WTIGRSEDNSIVLSDRWMSRNHAMIQRMGRGEFYLIDLGSRNGTFVNGRRVSIPIALHNGDRITFG

>Q8DG53/62127

WQIGRSKDCDIVLPDRWCSRHHIRIERQGDHRYYLTDLKSMNGTFVGNKRIQAPYILRHGDRITLG

>Q8DKL2/265

IGRGKDCTIVLADRWASRRHAMIQRLDNDDYYLIDLGSRNGSVLNGQMIQVPTLLKTGDRFIIG

>Q8DKF7/33101

WKIGRASDCDIIIHDPYVSRHQATIELRPRGNALVYLIRDNNSRNCTLLNGTPLSDERVLHHGDIIMMG

>Q8DMF0/3199

WTIGRQLDCSIRLTDAYVSRLHAVINAFLFRGQPLYFIRDAHSRNGTFLNGFPLQHSTLLHHEDVIGVG

>Q9XA20/99164

ITLGRAHDSTIVLDDDYASSRHARIYPDQNGQWIVEDLGSTNGTYLDRARLTTPTPIGPGAPIRIG

>Q50189/83147

VLIGRADDSTLVLTDDYASNRHARLSQRGSEWYVEDLGSTNGTYLDRAKVTTAVRVPLGTPVRIG

>P71589/83147

VLIGRADDSTLVLTDDYASTRHARLSMRGSEWYVEDLGSTNGTYLDRAKVTTAVRVPIGTPVRIG

>Q8NU93/82146

MTMGRSPECTFVVGDDYASGMHARVFKRGSEWFVEDLDSRNGTFVGGTRIDQPEQIAVGTDIRIG

>Q8FUI0/83147

MVIGRSPDCTFVVGDDYASGRHARIFKRGAEWFVEDLDSRNGTFVSGVRIDQPEQIELGTDIRVG

>Q8G6Q4/103169

ITLGRAASNTVVLDDEFVSSHHARVYRDTRSGQWAIEDLNSTNGTVVNQQRINRPTILPARIPVRIG

>Q8ZYE3/74141

VTIGRAVENHVVVTDPAVSRRHLRLIVTSNGIVVEDVGSSNGTFLLEEGKERQIKVENIGKKGLIRIG

>Q9AJX1/2588

YALGRDPQGELVFDDARVSWRHATISFNGRGWVVEDHGSTNGTFVHGQRVQQMELGAGTVLNLG

>Q97Z47/131199

ISIGRSPENVIVIPDPEVSRKHAIISFENNELYLEDLNSTNGTYVYDGKVFQPVKGKIKVQHNSIIKLG

>Q973R9/160228

ISVGRSPENVIIVPDPEVSRRHALISLEGGELYIEDLNSTNGTYIYDGKLFQPVKGKQKISPNTIIKLG

>Q8R8M6/61124

TSIGRSDECDIVIESPYVSARHALIKKRGKRFYIEDLNSTNGTFVNGKRVKVARIKNGDIITLG

>Q97LQ0/87151

ITIGRKDDNSIMLNEGYVSGHHARVYLRNNQYILEDLNSTNGTVLNGQKIKSKAYIKSGDEIKIG

>Q8XNI4/72136

ISIGRREDNSIVLNDQFVSSYHAKIYVKNNDFYLEDLASTNGTFINDSKVEGRVRLKVNDQIRMG

>Q8YXN7/31105

YSIGRHTSNAIVLHSRSVSRQHAILLRVTLPETDQCSFRIIDGNFKGQGSTNGLLVNGTRCFSHNLRHGDVINFG

>Q8DHW2/30104

YSIGRDPTNSIVLHSKMVSRQHAILFRVTSPETNSYLFRLIDGDLQGKRSTNGTVVNGQRIVAHDLRTGDMIVFG

>Q8DIG5/28100

YSIGRDVTNAIVIHGHGVSRQHALLLRIPSPNGYRYRVVDGNADGKPSANGVLVNGQRVQSHELKDGDEINFG

>Q8DJI8/29103

YSLGRDESNAIVIDFETVSRQHAILLRVPVPGTTSYRYRLVDGNANGKPSTNGTFVNGKRISSHELQHGDVILFG

>Q8DL70/31105

YTIGRHPSNRIVLKSPMISRQHAVLLRMLDPGSGNFFFRLIDGDLQGRRSINGLTVNGKPCQSHILQHGDVIVFG

>P72747/29104

YSIGRESSNDIVIYDQVVSRHHATLIRIKPTSQNEGYSYRILDGDLEGNRSTNGLIINGQDSESHDLQHGDVILFG

>P72746/32106

YDIGRDPSSAIPIYDRQVSRHHATLIRVNDYQNHQYTYRLIDGNLQGKRSTNGVVVNGQYCLSHELEHGDVIRFG

>YI95_SYNY3/2396

YSIGRHPSNTIVIPSPQISRRHATLIKKINPNLDISFHIIDGDLEGHRSRNGVWVNGESHLDYELVHGDVIALS

>P72748/2498

YSIGRHSGTSIQIHSRQASRHHATLMRKTNSKTNEECFWIIDGDLEGNKSQNGVFVNGEKRLIHELKNGDLINFG

>Q8YQP4/42114

YSIGRDKESNIRLVSQFVSRRHATLVRLPKNNSYYYRIVDGDGKGRASANGLMINGRKLPAHDLKNEDEIVFG

>P74513/32105

YTIGRSPRADIRIKSQFVSRIHAVLVRKSSDDVQAAYRIIDGDEDGQSSVNGLMINGKKVQEHIIQTGDEIVMG

>Q8DJ99/155218

LILGRDPKADIILDHPFVSFHHAKLVRSGDHYKIIDLGSKNGLHWRGTTISEHILAPGDVIHVG

>Q8YNI6/64132

YVLGRSSKSSDIVIRNPVVSQIHLSLSRDSSQRTPVFIIKDENSTNGIYRGKRRVNTLELRHGDILTLG

>P73218/61127

YFLGRSSSCDIQLQNPLVSQTHCSLRRDPDHPNQFFIRDEGSSNGIYLQRRRLKSYRLQHGDEITLG

>Q8DIU0/61127

YILGRSQSRADIMIDAEIVSGTHARLKRLTPTGVFEIQDLDSTNGIYRQRQRLQMGELHHGDVITLG

>Q8PF63/3196

GSIGRSDESDWVLAAPGVSRTHAVVRFLNGMYFIEDRSTNGMSLNGTALTRADPSALSDGDRLQLD

>Q8DJ99/41107

LIIGTDPQSQVILVGEGISRHHALILATETGYQVVDLGSASGTFLNGTKLQPRTPVVLKAGDRLKIG

>Q55902/125198

IHLGKPNDKVPPDIDVSGFANSDVVSRVHADIRLEGNEYFLEDVGSANGTYVNHTALLPGNRHLLRPGDRLALG

>FRAH_ANASP/204277

VHIGKPNDRIPPDVDVSGFANSEIVSRVHADIRLEGDAHYIEDVGSSNGTYINNLPLLPGNRHRLRPGDRISLG

>Q8DIU3/128201

IHIGKPNDRIPPDIDVAGFPNSEIVSRIHADLRIEGNDHYIEDVGSANGTYINNTPLYPGNRHRLQSGDRIALG

>Q9D768/144218

ATIGREKDMEHTVRIPEVAVSKFHAEVYFDHDLQSYVLVDQGSQNGTIVNGKQILQPKTKCDPYVLEHGDEVKIG

>Q97LZ2/115178

FTIGRGKFNDIVFDDIKVSEKHAEIFEDNGKYVLVDLNSTNKTYLNGEMITRAILSEDDQINIC

>Q8F4E7/414478

TTLGNGEFSDVVVSDPGVNKQHARIRKIKNRFVLYDLISDGGTYLNGKKILRPRILYDFDEISLG

>Q8MLA6/343410

GSLGREGAHDVIIPDVNVSKCHLKFKYENKLGIYQCLDLGSRNGTILNGSPMSSDAMDLVHGSVITLG

>Q9SJS6/131208

VTIGRLPEKADVVIPVATESNTPSLLTSVSGVHATINTNEKNLLVTDMNSTNGTFIEDKRLIPGVAAPAFPGTRITFG

>Q9I751/2897

LTIGRGPDNDWVLPDPERLVSSRHCTILNRDGVYYLTDTSTNGVLLVNAGHRLRRGNSEPLQDGETVRLG

>Q8YW86/435511

TRIGRTKDNDIVIPELSVSKRHAEILCRNNFTGNQARTYYLQDFSTYGTTWFLGSNGWQQLLREEVPLKSGMQLKFG

>Q98IL9/2695

LVIGRSADAGWQIDDPDMFVSRAHCKIRGDRDGYFVTDTSSSGLFVDDSDSPLGPGRSMRLQSGMRLRMG

>Q9I360/2196

KTFGRNGGLIGRGGECDWAIPDRKRHLSKQHARVSYRNGAFYLTDTSSNGIRAGNGSRLPHGEPQRIEQDSVFLLG

>Q93EC5/2895

RTLGRAPDCDWRLPEDRRSVSKLHCIIERDREGFLLRDQSANGSRVDGVAVHEGQIVRLADRSRLELG

>Q8U7W4/2693

RAIGRSRDCDWQIDDNERRVSKLHCTLSRDGEGFIILDQSANGTLVDGRLLLEGESARLRDGSQINIG

>Q8YPV6/3199

VRIGRDPLRCDIVLTNPTVSGLHVEIFFHSQQQNFYIRNLRSQNPPLVDGQQLIQGEKPLNQGSIIYLG

>Q9L252/516594

ITIGRRRHSTGDTPDIDLAVPPEDPGVSHQHAVLVQQPDGSWAVVDQNSTNGTTVNGGEEPIQPFVPVPLQDGDRVHVG

>Q8YS22/91156

WTIGRDRHNGICTYDKLLSRHHAAIKYVENQGFLLIDFQSTNGSFVNGEPVYQPIILKDGDRVRLG

>Chfr_Human_Q9NRT4/38106

WTIGRRRGCDLSFPSNKLVSGDHCRIVVDEKSGQVTLEDTSTSGTVINKLKVVKKQTCPLQTGDVIYLV

>Q8BJZ9/38106

WTIGRRRGCDLSFPSNKLVSGDHCKLTVDEISGEVTLEDTSTNGTVINKLQVVKKQTYPLHSGDIIYLV

>Q8BWH4/38106

WTIGRRRGCDLSFPSNKLVSGDHCKLTVDEISGEVTLEDTSTNGTVINKLQVVKKQTYPLQSGDIIYLV

>O52176/315378

LTIGLAHCDLSFPGDEGLAGRHCELSPTPTGALLRDLSGGLGTYVRLASGVERPLRPGDRVRLG

>Q8W382/25110

YKVGRKDCDVIVQTDTSISRVHAEIVVEKMVAWDPQSGAPANPSYVRVVDRSKYGTFFNKVQGTQGSRLHKDEDAMLADG
DTVTFG

>Q9RYH9/226305

LMVGRFDASSGPVDIDLSSLPGAEHISRHHAELYREGSQWFVRDLGSTNGVFVRKGGQSAFSPRLQEPTALSDGDELAFG

>Q9KBZ7/2897

WSIGRLGKSWKPDIAFDNVFISRKHALLYVEEGQVFVKDLDSKHGTYVNDQRLAPHAPERLSHGDRLSLA

>Q8ZXQ4/2691

RVFGRSDFLWHPRHQFISRRHFAVGYEEGVYFVEDLGSTNGTFLNGVDIRGKGRARLQPGDVINVA

>YKI5_CAEEL/39117

KSFGRATTNDVIFLDGTEHVELEPQFISRCHARVHHTNQDGVEEYLVEDISENGTYINDRRLSKDKREILKSGDTIKFG

>Q9W5R5/235307

CKVGRLIAKSKAGEGNAIFDCKVLSRNHAILWYTPDGKFWVKDTKSSNGTFINDNKLGSDPAELHYGDIVKFG

>Q8ITW1/31109

VKIGRAVARIQATAENAVFDCKVLSRNHAILWFRDGEFWLKDTKSSNGTFVNNEKLQQTVSGRDTDVRRVFSGDIIQLG

>Q06001/99192

LKLGRPVANSNSSSSSSLRGGKRVDSHTFSQVRSDNGNFDSRVLSRNHALLSCDPLTGKVYIRDLKSSNGTFINGQRIGS
NDVEIKVGDVIDLG

>Q03944/185273

LKLGRPVTNSVNKNNSGSKRDLFSQQVRPDNGNFDSRVLSRNHACLSCDPTSGKIYIRDLKSSNGTFVNGVKIRQNDVEL
KVGDTVDLG

>O74388/27106

YKIGRHTNKSTSPSPSNLFFNSKVLSRQHAELWLDKDTLSVYIRDVKSSNGTFVNETRLSPENKPSAPCKLNSGDIVDFG

>Q9VT40/34111

YTIGRLATDLIVAQDLSISRNHAQLLIQTEADGDDTLHIEDLGSRYGTFIFPKNSQKPRKVPAKTSTPLPVGTRLRFG

>Nbs1_Rattus norvegicus_Q9JIL9/24-100

YIVGRKNCAILIENDQSISRNHAVLRVNFPVTSLSQTDEIPTLIIKDNSKYGTFINEEKMQNGLSSTLKTGDRVTFG

>O88981/24100

YVVGRKNCGILIENDQSISRNHAVLTVNFPVTSLSQTDEIPTLTIKDNSKYGTFVNEEKMQTGLSCTLKTGDRVTFG

>O60934/24100

YVVGRKNCAILIENDQSISRNHAVLTANFSVTNLSQTDEIPVLTLKDNSKYGTFVNEEKMQNGFSRTLKSGDGITFG

>Q9DE07/2297

YVVGRKNCAFLIQDDQSISRSHAVLTVSRPETTHSQSVSVPVLTIKDTSKYGTFVNGSKLSGASRSLQSGDRVNFG

>YGA9_SCHPO/23108

YIVGRNVSDDSSHIQVISKSISKRHARFTILTPSEKDYLICAQFTGGPCEFEVKDLDTKFGTKVNEKVVGQNGDSYKEKD
LKIQLG

>O23305/2895

IRVGRIVRGNEIAIKDAGISTKHLRIESDSGNWVIQDLGSSNGTLLNSNALDPETSVNLGDGDVIKLG

>DMA1_SCHPO/60137

IYIGRYTERYNGGDVSAIVFRSKVVSRRHAQIFYENNTWYIQDMGSSSGTFLNHVRLSPPSKTSKPYPISNNDILQLG

>YNL6_YEAST/295379

LVIGRYTERVRDAISKIPEQYHPVVFKSKVVSRTHGCFKVDSQGNWYIKDVKSSSGTFLNHQRLSPASSLSKDTPLRDGD
ILQLG

>YHR5_YEAST/189273

IIIGRYTERVREAISKIPDQYHPVVFKSKVISRTHGCFKVDDQGNWFLKDVKSSSGTFLNHQRLSSASTTSKDYLLHDGD
IIQLG

>RNF8_MOUSE/38109

VTIGRGLSVTYQLISKVCPLMISRSHCVLKQNPEGQWTIMDNKSLNGVWLNRERLAPLQGYCIRKGDHIQLG

>RNF8_HUMAN/38109

VTVGRGFGVTYQLVSKICPLMISRNHCVLKQNPEGQWTIMDNKSLNGVWLNRARLEPLRVYSIHQGDYIQLG

>Q9W5Q1/303383

FLVGRDRKVVDLAVDHPSCSKQHAALQYRLVPFEREDGSHGKRVRLYLIDLDSANGTFLNNKKIDARKYYELIEKDVIKF
G

>P91125/188268

YLIGRDHKIADIPVDHPSCSKQHAVLQFRSMPFTRDDGTKARRIMPYIIDLGSGNGTFLNEKKIEPQRYIELQEKDMLKF
G

>SNIP1_Human_Q8TAD8/281361

YLLGRHRRIADIPIDHPSCSKQHAVFQYRLVEYTRADGTVGRRVKPYIIDLGSGNGTFLNNKRIEPQRYYELKEKDVLKF
G

>Q9LJU3/403473

YLFGRERRIADIPTDHPSCSKQHAVIQYREMEKEKPDGMMGKQVKPYIMDLGSTNKTYINPRVRTVAREFC

>Q8IEN7/474545

YLIGKEQLAVDIQLNNISISKQHAVIQFKKHESKILPFLLDLNSTNGTYINNEKIQPNKYYELRETDIIRFG

>YL16_YEAST/104189

YLVGRELGHSLDTDLDDRTEIVVADIGIPEETSSKQHCVIQFRNVRGILKCYVMDLDSSNGTCLNNVVIPGARYIELRSG
DVLTLS

>KAB1_Human_Q9UQ09/2390

IFVGRDDCELMLQSRSVDKQHAVINYDASTDEHLVKDLGSLNGTFVNDVRIPEQTYITLKLEDKLRFG

>Q9VE47/88181

WTFGRLPENDVPAAHPTISRFHVVLQYKPKAPPKPETAKEDDEMEEEDEEPKNDQPEGWYIYDMGSTHGTFLNKQRVPPK
VYIRMRVGHMLKLG

>Q9FNZ7/124191

YMFGRIDLCDFVLEHPTISRFHAVLQFRNDGEVFLYDLGSTHGSFINKTQVKKKIYVEIHVGDVIRFG

>Q9BWU0/189266

CLFGRLSGCDVCLEHPSVSRYHAVLQHRASGPDGECDSNGPGFYLYDLGSTHGTFLNKTRIPPRTYCRVHVGHVVRFG

>YOT2_CAEEL/108182

VVIGRIKPGCDLLMEHPSISRYHCILQYGNDKMSKTGKGWHIFELGSTHGSRMNKKRLPPKQYIRTRVGFIFQFG

>Q9V7W9/40109

YLFGRNSQMNDFCIDHASCSRVHSAFVYHKHLNIAYLVDLGSTHGTFIGTLRLEAHKPTQLQINSTFHFG

>O61817/45114

YYFGRNNKQVDFAVEHASCSRVHALLLYHGLLQRFALIDMDSSHGTFLGNVRLRPLEVVFMDPGTQFHLG

>Q8H4Y5/124192

HIFGRQVPACDFVLDHQSVSRQHAAVVPHRNGSIYVIDLGSVHGTFVANERLTKDNPVELEVGQSLRFA

>P90527/33102

TVFGRSSEVASVLLDHPSVSRRHAALVYHGANNRFYLIDLQSATGTQVNDEQVKPLTPTTVKENFTFKFG

>Q97N03/392459

FKIGRISGQADYISDNKAVGKLHAEIRKQNEKYYLIDLTSRNGTFVNGQKINSDELYEIRNGDTIMFA

>Q97LZ4/440507

FKIGRLTGSVDYVSDNRAIGKMHAEIRKINSEYYLMDLDSKNGTFINDRRLESNELYKISENDILKFA

>Q96TX9/44124

LKIGRSSKLAAKGFVPSSQNGWYDSPVMSRQHAEIIADFTQKRLKLRDLGSLHGTYINKSDQRLEKDEQVEIKDGDNIRF
G

>Q8YMH9/27111

CIMGRSPEANIQLPDDAEHKTISRYHCLLDIMPPNIRIRDFGSKNGTYVNGQKIGQREAHQTPEEGAKINFPEYDLQSDD
EIELG

>Q9VSE1/2392

NLMGRHSRCRIILGGQYQFVSREHANIIVSEKGVEVESLNPLNGLFINGGKLGDKINRLAVAEGSTISLG

>Q8YPV7/27101

IAFGREFVRLPAELDGKRVSRMLLNSDQVSRYHALIVWENNQLVVIDQDSVNGIFINGQQQKRSVLTNGDTLQIG

>Q9U3Q2/2687

VNFGREKKVCHITFDPHAARVSRIHASIEWGDEGLFFTDKSKEGTEINGTRLKQSSQELHEG

>Q55956/30103

HILGRDSQFADLIVPTDWSIISRCQATLVEDQGNYFIYDGDRLNPSSNKLFINHHLITHDIGYPLQNGDVIKIG

>O65934/2992

VVVGRDLRADVRVAHPLISRAHLLLRFDQGRWVAIDNGSLNGLYLNNRRVPVVDIYDAQRVHIG

>Q23717/342431

WDLGRHSKSTIQVSHSSVSTNHLLIVNIKEHDQVAQELKYKSNTSHKFDESSPIEKDWGIGVIDLGSSIGTMLKIQPKHR
LATDEVIHLA

>Q8I5Z0/678740

WDGGRHAKSLIHMPHSSVSLSHLLFVKAENNNIGVVDCGSTIGTMIKVNNYHVLKEGDVIHIG

>Q97N00/130195

LSIGRDEDNNISIRDDLIDRKHCEIKCDSNNKFYVTDLKSKYGVYIDGKKVEDKVYLKDNNFITIC

>Q55956/138202

VVIGRDETANLQLLSSAVSRQHAVLDYLEPNTYLLHDYSTNGVYVNGTKVNGKILINTAAIIKIV

>FHA2-Rad53_SPK1_YEAST/601681

FFIGRSEDCNCKIEDNRLSRVHCFIFKKRHAVGKSMYESPAQGLDDIWYCHTGTNVSYLNNNRMIQGTKFLLQDGDEIKI
I

>YGI1_YEAST/22109

KTIGRSSSFDQNSLCKPYNLYFDEPELSRQHAVLCIKTPIPKIEGVPSIEQLRICIRDLNNKTGTVNLVSDGPNDEIDLK
NGDAFGLI

>Q8MYD6/38111

NKIGRASSNNIYLKGDTSVSNYHAQLVFNGAFFLIKDLNSLNKTMVQYRPNEDTTVLKPNLEYPIINGSKITFG

>Q8IR47/2188

HLLGKSGACNVVLTYDNMSDIHAVVVVRNGRIFIKDLESIHGIYLNFKEHRLGSEFYEIFVGDDVAFG

>MEK1_YEAST/47119

VKVGRNDKECQLVLTNPSISSVHCVFWCVFFDEDSIPMFYVKDCSLNGTYLNGLLLKRDKTYLLKHCDVIELS

>Q8II28/310377

ICFGKNETNDIVCMNPSISRFHCVLYMCEDFQVYLIDVGSKSGTKLNNCICEIHKKYKVSNNDVISLG

>Q9VZJ3/477550

ITFGRDAKDCVVDVDLGLEGPAAKISRRQGTIKLRSNGDFFIANEGKRAIFIDGTPLLSANKARLGHNCTVEIS

>TOJ3_ Coturnix coturnix_Q9DEV2/431504

ITLGRATKDNQIDVDLALEGPAWKISRKQGVIKLKNNGDFFIANEGRRPIYIDGRPVLGGNKWKLNNNSVVEIA

>MCS1_HUMAN/363436

ITLGRATKDNQIDVDLSLEGPAWKISRKQGVIKLKNNGDFFIANEGRRPIYIDGRPVLCGSKWRLSNNSVVEIA

>Q9LQZ6/460533

VLLGRATGEYPVDIDLGRSGSETRFSRRQALIKLKQDGSFEIKNLGKFSIWMNDEEINHGEVVILKNNCLIQVN

>Q8GXA6/32107

IILGRNSKKATVDVDLSSLGGGMSISRNHARIFYDFTRRRFSLEVLGKNGCLVEGVLHLPGNPNVKLDSQDLLQIG

>NtFHA1_Tobacco_Q945P0/30-105

IILGRNSKKSTVDVDLSSLGGGMNISRHHARIFYDFQRRRFNLEVLGKNGCFVEGVLHLPGNPPIKLDSQDLLQIG

>YKY8_SCHPO/39111

VTMGRKASNSSDCDVHLGDTKAISRQHAKIFYSFPNQRFEISVMGKNGAFVDGEFVERGKSVPLRSGTRVQIG

>FHL1_YEAST/300374

AIIGRRSENDFSHKVDVNGPSKSISRRHAQIFYNFGTGRFELSIIGKNGAFVDDIFVEKGNTVPLRNKTKIQIG

>O60129/96177

IILGREPANPSPKGKNEDLEVIDMNFGPSKVVSRKHAVVEYDLDDQTWNCSVYGRNGIKVDGKLFKNGETVKLTSGSILE
VA

>FKH2_YEAST/83172

VSIGRNTDPLNSALQENSDGVKNSYRVNIDLGPAKVVSRKHAIIKYNMNIGGWELHILGRNGAKVNFQRTHNGPNNPPIR
LSSGTLLDIG

>FKH1_YEAST/76162

VTIGRNTDSLNLNAVPGTVVKKNIDIDLGPAKIVSRKHAAIRFNLESGSWELQIFGRNGAKVNFRRIPTGPDSPPTVLQS
GCIIDIG

>Q8YZC9/2692

IRIGRAADNHVILSDNLVSRHHLEIRQVSSGGGGSWQVVSKGTNGTFLNGVLVIQDALPNNALLQLA

>Q8DJU4/2693

IRIGRAADNDVILNDILVSRYHAELSCYRDPENLGRWFLKSLGAMGTFVDGRLVNESKLASGSLIQLG

>P74282/2890

IRVGRAADNDVVLYSAVVSRHHLELRREEEAWVAINLGANGTYMEAEMVEKLVLKDGMVLRLA

>Q8DIU4/2688

IRIGRAVDNHVVLYSAVVSRHHVELRRHGLQWEVVNLGTNGTYLDGKRIQQATLTDGGILRLA

>Q9NIN1/341404

AIIGRGHESDVRLSDISVSRMHASLELDGGKVVIHDQQSKFGTLVRAKAPFSMPIKGPICLQVS

>Q8I355/717782

VRMGRGHDSDVRVNDISVSRFHALIKFHNGNFYIEDCKSKFGTLIQIRKPVFFNIRRNKFIALQIG

>FXK1_MOUSE/107177

VTIGRNSSQGSVDLSMGLSSFISRRHLQLSFQEPHFYLRCLGKNGVFVDGAFQRRGAPALQLPQQCTFRFP

>Q8IQE9/171243

TVIGRNSSTSLVHFNVAENNLVSRKHFQVLYDVELRAFFVQCLSKNGIFVDDFLQRRNVDPLRLPQRCYFRFP

>Q23717/573662

MVFGRGKNGALGLRKVEVTESNGYISREHCIFYYSSIRNREYEGLDYLSRSNGCLSNWFVKDVSTSGTFLRLKPFSYPVR
VLPGMVLKVG

>Q8I5Z0/833908

VVLGRGPYAQSSYKKISVTNSNGYVSREHCLIYYDGSKDAGERWLLRDTSTLGTFLKIKPFSNPIPLPIGSIFKAG

>CpABA1_Q9FS22/507-583

FIIGSAPAEDHPGTSVTIPSPQVSPRHARINYKDGAFFLIDLRSEHGTWIIDNEGKQYRVPPNYPARIRPSEAIQFG

>ABA2_PRUAR/558634

CIIGSAPHGDVSGISIAIPKPQVSEMHARISYKDGAFYLTDLRSEHGTWIADIEGKRYRVPPNFPARFRPSDAIEIG

>LOS/ABA1_Arath_Q9FS21/558-634

CIVGSEPDQDFPGMRIVIPSSQVSKMHARVIYKDGAFFLMDLRSEHGTYVTDNEGRRYRATPNFPARFRSSDIIEFG

>Q8W3L2/556632

YLIGSESHEDFPRTSIVIPSAQVSKMHARISYKDGAFYLIDLQSEHGTYVTDNEGRRYRVSSNFPARFRPSDTIEFG

>ABA2_CAPAN/555631

CTIGSVSHANIPGKSVVIPLSQVSDMHARISYNGGAFLGTAFRSDHGTWFIDNEGRRYRVSPNFPMRFHSSDVIVFG

>ABA2_NICPL/557633

CNIGSVSHANIPGKSVVIPLPQVSEMHARISYKGGAFFVTDLRSEHGTWITDNEGRRYRASPNFPTRFHPSDIIEFG

>ABA2_LYCES/563639

CTVGSISHTNIPGKSIVLPLPQVSEMHARISCKDGAFFVTDLRSEHGTWVTDNEGRRYRTSPNFPTRFHPSDVIEFG

>Q9AVE7/553624

LSIGSRSDPSNSTASLALPLPQISENHATITCKNKAFYVTDNGSEHGTWITDNEGRRYRRTSELPCPFPSLG

>Q8T7Z6/79154

NTIGRSSTSSISLKSSKLDTSVTSRKHCTIYYDNFGLRITSNSLNGLTKIYRNGCNIYLRHNQISPLYVGETIQIG

>At5g38840

YLFGRDGICDFALEHPSISRFHAVIQYKRSGAAYIFDLGSTHGTTVNKNKVDKKVFVDLNVGDVIRFG

>At3g20550

YLFGRERRIADIPTDHPSCSKQHAVIQYREMEKEKPDGMMGKQVKPYIMDLGSTNKTYINES
PIEPQRYYELFEKDTIKFG

>At3g02400

IRIGRIVRGNEIAIKDAGISTKHLRIVSDSENWIIHDLGSSNGTILNSDTIDSDTPVNLSHGDEIKLG

>At1g34350

LVVGRHPDCDILLTHPSISRFHLEIRSISSRQKLFVTDLSSVHGTWVRDLRIEPHGCVEVEEGDTIR IG

>At3g07260

IILGRNSKKSTVDVDLSSLGGGMNISRNHARIFYDFTRRRFSLEVLGKNGCFVEGVLHLPGNPNVKLDS QDLLQIG

>At5g47790

HIFGRQHQTCDFVLDHQSVSRQHAAVVPHKNGSIFVIDLGSAHGTFVANERLTKDTPVELEVGQSLRFA

>At5g07400

YTIGRSSSDGFCDFVIDHSSISRKHCQILFDSQSHKLYIFDGVIHLPSGSFSQVYDEFRRRLVGVEDLGNL
KFRASLNGVYVNRVRVRKSKVQEVSIDDEVLFF

>At3g54350

VLVGRSTEDLAVDIDLGREKRGSKISRRQAIIRLGDDGSFHIKNLGKYSISVNEKEVDPGQSLILKSDCLVEI R

>At1g60700

VIIGRSSGGLNVDI DLGKYNYGSKISRRQALVKLENYGSFSLKNLGKQHILVNGGKLDRGQIVT LTSCSSINVS

>At3g07220

IILGRNSKKATVDVDLSSLGGGMNISRNHARIFYDFTRRRFSLEVLGKNGCLVEGVLHLPGNPNVKLDS QDLLQIG

>At1g75530

VLLGRATGEYPVDIDLGRSGSETRFSRRQALIKLKQDGSFEIKNLGKFSIWMNDEEINHGEVVILKNNCLI QIR

>NIPP1_Human_Q12972

YL FGRNPDLCDFTIDHQSCSRV HAALVYHKHL KRVFLIDLNS THGTFLGHIR LEPHKPQQIP
IDSTVSFGASTRAYTL

APPENDIX 5. Sequences of the kinase domains of receptor-like kinases for evolutionary trace analysis

61 kinase domain sequences of receptor-like kinases in Arabidopsis were used for evolutionary trace analysis and listed in FastA format as below.

```
>At1g08590.catpro
WIYTRWDLYSNFAREYIFCKKPREEWPWRLVAFQRLCFTAGDILSHIKESNIIGMGAIGIVYKAEVMRRPLLTVAVKKLW
RSPSPQNDIEDHHQEEDEEDDILREVNLLGGLRHRNIVKILGYVHNEREVMMVYEYMPNGNLGTALHSKDEKFLLRDWLS
RYNVAVGVVQGLNYLHNDCYPPIIHRDIKSNNILLDSNLEARIADFGLAKMMLHKNETVSMVAGSYGYIAPEYGYTLKID
EKSDIYSLGVVLLELVTGKMPIDPSFEDSIDVVEWIRRKVKKNESLEEVIDASIAGDCKHVIEEMLLALRIALLCTAKLP
KDRPSIRDVITMLAEAKPRRKSVCQVAGDLPIFRNSPVVGLI
>At5g65710.catpro
FKRKPKRTNKITIFQRVGFTEEDIYPQLTEDNIIGSGGSGLVYRVKLKSGQTLAVKKLWGETGQKTESESVFRSEVETLG
RVRHGNIVKLLMCCNGEEFRFLVYEFMENGSLGDVLHSEKEHRAVSPLDWTTRFSIAVGAAQGLSYLHHDSVPPIVHRDV
KSNNILLDHEMKPRVADFGLAKPLKREDNDGVSDVSMSCVAGSYGYIAPEYGYTSKVNEKSDVYSFGVVLLELITGKRPN
DSSFGENKDIVKFAMEAALCYPSPSAEDGAMNQDSLGNYRDLSKLVDPKMKLSTREYEEIEKVLDVALLCTSSFPINRPT
MRKVVELLKEKKSLE
>At5g65700.catpro
LKKASESRAWRLTAFQRLDFTCDDVLDSLKEDNIIGKGGAGIVYKGVMPNGDLVAVKRLAAMSRGSSHDHGFNAEIQTLG
RIRHRHIVRLLGFCSNHETNLLVYEYMPNGSLGEVLHGKKGGHLHWDTRYKIALEAAKGLCYLHHDCSPLIVHRDVKSNN
ILLDSNFEAHVADFGLAKFLQDSGTSECMSAIAGSYGYIAPEYAYTLKVDEKSDVYSFGVVLLELVTGRKPVGEFGDGVD
IVQWVRKMTDSNKDSVLKVLDPRLSSIPIHEVTHVFYVAMLCVEEQAVERPTMREVVQILTEIPKLPPSKDQPMTESAPE
SELSPKSGVQSPPDLLNL
>AT5G63930.catpro
```

VRTVASSAQDGQPSEMSLDIYFPPKEGFTFQDLVAATDNFDESFVVGRGACGTVYKAVLPAGYTLAVKKLASNHEGGNNN
NVDNSFRAEILTLGNIRHRNIVKLHGFCNHQGSNLLLYEYMPKGSLGEILHDPSCNLDWSKRFKIALGAAQGLAYLHHDC
KPRIFHRDIKSNNILLDDKFEAHVGDFGLAKVIDMPHSKSMSAIAGSYGYIAPEYAYTMKVTEKSDIYSYGVVLLELLTG
KAPVQPIDQGGDVVNWVRSYIRRDALSSGVLDARLTLEDERIVSHMLTVLKIALLCTSVSPVARPSMRQVVLMLIESERS
EGEQEHLDTEELTQTTTP

>At5g62710.catpro_

LITFHGDLPYSSTELIEKLESLDEEDIVGSGGFGTVYRMVMNDLGTFAVKKIDRSRQGSDRVFEREVEILGSVKHINLVN
LRGYCRLPSSRLLIYDYLTLGSLDDLLHERAQEDGLLNWNARLKIALGSARGLAYLHHDCSPKIVHRDIKSSNILLNDKL
EPRVSDFGLAKLLVDEDAHVTTVVAGTFGYLAPEYLQNGRATEKSDVYSFGVLLLELVTGKRPTDPIFVKRGLNVVGWCY
FSQMNTVLKENRLEDVIDKRCTDVDEESVEALLEIAERCTDANPENRPAMNQVAQLLEQEVMSPSSGIDYYDDSHSDYC

>At5g62230.catpro

MQQKKILQGSSKQAEGLTKLVILHMDMAIHTFDDIMRVTENLNEKFIIGYGASSTVYKCALKSSRPIAIKRLYNQYPHNL
REFETELETIGSIRHRNIVSLHGYALSPTGNLLFYDYMENGSLWDLLHGSLKKVKLDWETRLKIAVGAAQGLAYLHHDCT
PRIIHRDIKSSNILLDENFEAHLSDFGIAKSIPASKTHASTYVLGTIGYIDPEYARTSRINEKSDIYSFGIVLLELLTGK
KAVDNEANLHQLILSKADDNTVMEAVDPEVTVTCMDLGHIRKTFQLALLCTKRNPLERPTMLEVSRVLLSLVPSLQVAKK
LPSLDHSTKKLQQENEVRNPDAEASQWFVQFREVISKSSI

>At5g61480.catpro

YGNRVDGGGRNGGDIGPWKLTAFQRLNFTADDVVECLSKTDNILGMGSTGTVYKAEMPNGEIIAVKKLWGKNKENGKIRR
RKSGVLAEVDVLGNVRHRNIVRLLGCCTNRDCTMLLYEYMPNGSLDDLLHGGDKTMTAAAEWTALYQIAIGVAQGICYLH
HDCDPVIVHRDLKPSNILLDADFEARVADFGVAKLIQTDESMSVVAGSYGYIAPEYAYTLQVDKKSDIYSYGVILLEIIT
GKRSVEPEFGEGNSIVDWVRSKLKTKEDVEEVLDKSMGRSCSLIREEMKQMLRIALLCTSRSPTDRPPMRDVLLILQEAK
PKRKTVGDNVIVVGDVNDVNFEDVCSVDVGHDVKCQRIGV

>AT5G56040.catpro

RITGKQEELDSWEVTLYQKLDFSIDDIVKNLTSANVIGTGSSGVVYRVTIPSGETLAVKKMWSKEENRAFNSEINTLGSI
RHRNIIRLLGWCSNRNLKLLFYDYLPNGSLSSLLHGAGKGSGGADWEARYDVVLGVAHALAYLHHDCLPPILHGDVKAMN
VLLGSRFESYLADFGLAKIVSGEGVTDGDSSKLSNRPPLAGSYGYMAPEHASMQHITEKSDVYSYGVVLLEVLTGKHPLD
PDLPGGAHLVQWVRDHLAGKKDPREILDPRLRGRADPIMHEMLQTLAVSFLCVSNKASDRPMMKDIVAMLKEIRQFDMDR
SESDMIKGGKCEKWQPQPLPPEKIVSTPRGSSNCSFAYSDESV

>At5g53890.catpro

SRKDVDDRINDVDEETISGVSKALGPSKIVLFHSCGCKDLSVEELLKSTNNFSQANIIGCGGFGLVYKANFPDGSKAAVK
RLSGDCGQMEREFQAEVEALSRAEHKNLVSLQGYCKHGNDRLLIYSFMENGSLDYWLHERVDGNMTLIWDVRLKIAQGAA
RGLAYLHKVCEPNVIHRDVKSSNILLDEKFEAHLADFGLARLLRPYDTHVTTDLVGTLGYIPPEYSQSLIATCRGDVYSF
GVVLLELVTGRRPVEVCKGKSCRDLVSRVFQMKAEKREAELIDTTIRENVNERTVLEMLEIACKCIDHEPRRRPLIEEVV
TWLEDLPMESVQQQ

>At5g49660.catpro

RMSKNRAVIEQDETLASSFFSYDVKSFHRISFDQREILESLVDKNIVGHGGSGTVYRVELKSGEVVAVKKLWSQSNKDSA
SEDKMHLNKELKTEVETLGSIRHKNIVKLFSYFSSLDCSLLVYEYMPNGNLWDALHKGFVHLEWRTRHQIAVGVAQGLAY
LHHDLSPPIIHRDIKSTNILLDVNYQPKVADFGIAKVLQARGKDSTTTVMAGTYGYLAPEYAYSSKATIKCDVYSFGVVL
MELITGKKPVDSCFGENKNIVNWVSTKIDTKEGLIETLDKRLSESSKADMINALRVAIRCTSRTPTIRPTMNEVVQLLID
ATPQGGPDMTSKPTTKIKDSIVSDHLTQTRL

>AT5G48940.catpro

MIRDDNDSETGENLWTWQFTPFQKLNFTVEHVLKCLVEGNVIGKGCSGIVYKAEMPNREVIAVKKLWPVTVPNLNEKTKS
SGVRDSFSAEVKTLGSIRHKNIVRFLGCCWNKNTRLLMYDYMSNGSLGSLLHERSGVCSLGWEVRDIKANNILIGPDFEP
YIGDFGLAKLVDDGDFARSSNTIAGSYGYIAPEYGYSMKITEKSDVYSYGVVVLEVLTGKQPIDPTIPDGLHIVDWVKKI
RDIQVIDQGLQARPESEVEEMMQTLGVALLCINPIPEDRPTMKDVAAMLSEICQEREESMKVDGCSGSCNNGRERGKDDS
TSSVMQQTAKYLRSSSTSFSASSLLYSSSSSATSNVRPNLK

>At5g44700.catpro

DLFKKVRGGNSAFSSNSSSSQAPLFSNGGAKSDIKWDDIMEATHYLNEEFMIGSGGSGKVYKAELKNGETIAVKKILWKD
DLMSNKSFNREVKTLGTIRHRHLVKLMGYCSSKADGLNLLIYEYMANGSVWDWLHANENTKKKEVLGWETRLKIALGLAQ
GVEYLHYDCVPPIVHRDIKSSNVLLDSNIEAHLGDFGLAKILTGNYDTNTESNTMFAGSYGYIAPEYAYSLKATEKSDVY
SMGIVLMEIVTGKMPTEAMFDEETDMVRWVETVLDTPPGSEAREKLIDSELKSLLPCEEEAAYQVLEIALQCTKSYPQER
PSSRQASEYLLNVFNNRAASYREMQTDTDK

>At5g42440.catpro

RRRPPIQNHPRRNRNFPDPDPDLNTETVTESFDPSICEISMAELTIATKNFSSDLIVGDGSFGLVYRAQLSNGVVVAVKK
LDHDALQGFREFAAEMDTLGRLNHPNIVRILGYCISGSDRILIYEFLEKSSLDYWLHETDEENSPLTWSTRVNITRDVAK
GLAYLHGLPKPIIHRDIKSSNVLLDSDFVAHIADFGLARRIDASRSHVSTQVAGTMGYMPPEYWEGNTAATVKADVYSFG
VLMLELATRRRPNLTVVVDEKEVGLAQWAVIMVEQNRCYEMLDFGGVCGSEKGVEEYFRIACLCIKESTRERPTMVQVVE
LLEELCRFM

>At5g25930.catpro

RKQRRRGLETWKLTSFHRVDFAESDIVSNLMEHYVIGSGGSGKVYKIFVESSGQCVAVKRIWDSKKLDQKLEKEFIAEVE
ILGTIRHSNIVKLLCCISREDSKLLVYEYLEKRSLDQWLHGKKKGGTVEANNLTWSQRLNIAVGAAQGLCYMHHDCTPAI
IHRDVKSSNILLDSEFNAKIADFGLAKLLIKQNQEPHTMSAVAGSFGYIAPEYAYTSKVDEKIDVYSFGVVLLELVTGRE
GNNGDEHTNLADWSWKHYQSGKPTAEAFDEDIKEASTTEAMTTVFKLGLMCTNTLPSHRPSMKEVLYVLRQQGLEATKKT
ATEAYEAPLLVSLSGRRTSKRVEDEDLGFV

>At5g07280.catpro

AMTKRVKQRDDPERMEESRLKGFVDQNLYFLSGSRSREPLSINIAMFEQPLLKVRLGDIVEATDHFSKKNIIGDGGFGTV
YKACLPGEKTVAVKKLSEAKTQGNREFMAEMETLGKVKHPNLVSLLGYCSFSEEKLLVYEYMVNGSLDHWLRNQTGMLEV
LDWSKRLKIAVGAARGLAFLHHGFIPHIIHRDIKASNILLDGDFEPKVADFGLARLISACESHVSTVIAGTFGYIPPEYG

QSARATTKGDVYSFGVILLELVTGKEPTGPDFKESEGGNLVGWAIQKINQGKAVDVIDPLLVSVALKNSQLRLLQIAMLC
LAETPAKRPNMLDVLKALKEI

>AT5G07180.catpro

KQQKPVLKGSSKQPEGSTKLVILHMDMAIHTFDDIMRVTENLDEKYIIGYGASSTVYKCTSKTSRPIAIKRIYNQYPSNF
REFETELETIGSIRHRNIVSLHGYALSPFGNLLFYDYMENGSLWDLLHGPGKKVKLDWETRLKIAVGAAQGLAYLHHDCT
PRIIHRDIKSSNILLDGNFEARLSDFGIAKSIPATKTYASTYVLGTIGYIDPEYARTSRLNEKSDIYSFGIVLLELLTGK
KAVDNEANLHQMILSKADDNTVMEAVDAEVSVTCMDSGHIKKTFQLALLCTKRNPLERPTMQEVSRVLLSLVPSPPPKKL
PSPAKVQEGEERRESHSSDTTTPQWFVQFREDISKSSL

>BRI1

EMRKRRRKKEAELEMYAEGHGNSGDRTANNTNWKLTGVKEALSINLAAFEKPLRKLTFADLLQATNGFHNDSLIGSGGFG
DVYKAILKDGSAVAIKKLIHVSGQGDREFMAEMETIGKIKHRNLVPLLGYCKVGDERLLVYEFMKYGSLEDVLHDPKKAG
VKLNWSTRRKIAIGSARGLAFLHHNCSPHIIHRDMKSSNVLLDENLEARVSDFGMARLMSAMDTHLSVSTLAGTPGYVPP
EYYQSFRCSTKGDVYSYGVVLLELLTGKRPTDSPDFGDNNLVGWVKQHAKLRISDVFDPELMKEDPALEIELLQHLKVAV
ACLDDRAWRRPTMVQVMAMFKEIQAGSGIDSQSTIRSIEDGGFSTIEMVDMSIKEVPEGKL

>AT4G28650.catpro

KWYSNGFCGDETASKGEWPWRLMAFHRLGFTASDILACIKESNMIGMGATGIVYKAEMSRSSTVLAVKKLWRSAADIEDG
TTGDFVGEVNLLGKLRHRNIVRLLGFLYNDKNMMIVYEFMLNGNLGDAIHGKNAAGRLLVDWVSRYNIALGVAHGLAYLH
HDCHPPVIHRDIKSNNILLDANLDARIADFGLARMMARKKETVSMVAGSYGYIAPEYGYTLKVDEKIDIYSYGVVLLELL
TGRRPLEPEFGESVDIVEWVRRKIRDNISLEEALDPNVGNCRYVQEEMLLVLQIALLCTTKLPKDRPSMRDVISMLGEAK
PRRKSNSNEENTSRSLAEKHSSVFSTSPVNGLL

>HAESA

KLRALKSSTLAASKWRSFHKLHFSEHEIADCLDEKNVIGFGSSGKVYKVELRGGEVVAVKKLNKSVKGGDDEYSSDSLNR
DVFAAEVETLGTIRHKSIVRLWCCCSSGDCKLLVYEYMPNGSLADVLHGDRKGGVVLGWPERLRIALDAAEGLSYLHHDC
VPPIVHRDVKSSNILLDSDYGAKVADFGIAKVGQMSGSKTPEAMSGIAGSCGYIAPEYVYTLRVNEKSDIYSFGVVLLEL
VTGKQPTDSELGDKDMAKWVCTALDKCGLEPVIDPKLDLKFKEEISKVIHIGLLCTSPLPLNRPSMRKVVIMLQEVSGAV
PCSSPNTSKRSKTGGKLSPYYTEDLNSV

>AT4G26540.catpro

AAGKQLLGEEIDSWEVTLYQKLDFSIDDIVKNLTSANVIGTGSSGVVYRITIPSGESLAVKKMWSKEESGAFNSEIKTLG
SIRHRNIVRLLGWCSNRNLKLLFYDYLPNGSLSSRLHGAGKGGCVDWEARYDVVLGVAHALAYLHHDCLPTIIHGDVKAM
NVLLGPHFEPYLADFGLARTISGYPNTGIDLAKPTNRPPMAGSLWLHGSSFDFDLFCLLGFTEHASMQRITEKSDVYSYG
VVLLEVLTGKHPLDPDLPGGAHLVKWVRDHLAEKKDPSRLLDPRLDGRTDSIMHEMLQTLAVAFLCVSNKANERPLMKDV
VAMLTEIRHIDVGRSETEKIKAGGCGSKEPQQFMSNEKIINSHGSSNCSFAFSDDSV

>At4g20270.catpro

RMRKNNPNLWKLIGFQKLGFRSEHILECVKENHVIGKGGRGIVYKGVMPNGEEVAVKKLLTITKGSSHDNGLAAEIQTLG

RIRHRNIVRLLAFCSNKDVNLLVYEYMPNGSLGEVLHGKAGVFLKWETRLQIALEAAKGLCYLHHDCSPLIIHRDVKSNN
ILLGPEFEAHVADFGLAKFMMQDNGASECMSSIAGSYGYIAPEYAYTLRIDEKSDVYSFGVVLLELITGRKPVDNFGEEG
IDIVQWSKIQTNCNRQGVVKIIDQRLSNIPLAEAMELFFVAMLCVQEHSVERPTMREVVQMISQAKQPNTF

>At4g20140.catpro

FKKVGHGSTAYTSSSSSSQATHKPLFRNGASKSDIRWEDIMEATHNLSEEFMIGSGGSGKVYKAELENGETVAVKKILWK
DDLMSNKSFSREVKTLGRIRHRHLVKLMGYCSSKSEGLNLLIYEYMKNGSIWDWLHEDKPVLEKKKKLLDWEARLRIAVG
LAQGVEYLHHDCVPPIVHRDIKSSNVLLDSNMEAHLGDFGLAKVLTENCDTNTDSNTWFACSYGYIAPEYAYSLKATEKS
DVYSMGIVLMEIVTGKMPTDSVFGAEMDMVRWVETHLEVAGSARDKLIDPKLKPLLPFEEDAACQVLEIALQCTKTSPQE
RPSSRQACDSLLHVYNNRTAGYKKL

>At3g49670.catpro

SEAKAWRLTAFQRLDFTCDDVLDSLKEDNIIGKGGAGIVYKGTMPKGDLVAVKRLATMSHGSSHDHGFNAEIQTLGRIRH
RHIVRLLGFCSNHETNLLVYEYMPNGSLGEVLHGKKGGHLHWNTRYKIALEAAKGLCYLHHDCSPLIVHRDVKSNNILLD
SNFEAHVADFGLAKFLQDSGTSECMSAIAGSYGYIAPEYAYTLKVDEKSDVYSFGVVLLELITGKKPVGEFGDGVDIVQW
VRSMTDSNKDCVLKVIDLRLSSVPVHEVTHVFYVALLCVEEQAVERPTMREVVQILTEIPKIPLSKQQAAESDVTEKAPA
INESSPDSGSPPDLLSN

>AT3G24240.catpro

RRNIDNERDSELGETYKWQFTPFQKLNFSVDQIIRCLVEPNVIGKGCSGVVYRADVDNGEVIAVKKLWPAMVNGGHDEKT
KNVRDSFSAEVKTLGTIRHKNIVRFLGCCWNRNTRLLMYDYMPNGSLGSLLHERRGSSLDWDLRYRILLGAAQGLAYLHH
DCLPPIVHRDIKANNILIGLDFEPYIADFGLAKLVDEGDIGRCSNTVAGSYGYIAPEYGYSMKITEKSDVYSYGVVVLEV
LTGKQPIDPTVPEGIHLVDWVRQNRGSLEVLDSTLRSRTEAEADEMMQVLGTALLCVNSSPDERPTMKDVAAMLKEIKQE
REEYAKVDLLLKKSPPPTTTMQEECRKNEMMMIPAAAASSSKEMRREERLLKSNNTSFSASSLLYSSSSSIE

>At3g19700.catpro

KLNKTVQKKNDWQVSSFRLLNFNEMEIIDEIKSENIIGRGGQGNVYKVSLRSGETLAVKHIWCPESSHESFRSSTAMLSD
GNNRSNNGEFEAEVATLSNIKHINVVKLFCSITCEDSKLLVYEYMPNGSLWEQLHERRGEQEIGWRVRQALALGAAKGLE
YLHHGLDRPVIHRDVKSSNILLDEEWRPRIADFGLAKIIQADSVQRDFSAPLVKGTLGYIAPEYAYTTKVNEKSDVYSFG
VVLMELVTGKKPLETDFGENNDIVMWVWSVSKETNREMMMKLIDTSIEDEYKEDALKVLTIALLCTDKSPQARPFMKSVV
SMLEKIEPSYNKNSGEASYGESANDEITKVV

>At3g13380.catpro

VQKKEKQREKYIESLPTSGSSSWKLSSVHEPLSINVATFEKPLRKLTFAHLLEATNGFSADSMIGSGGFGDVYKAKLADG
SVVAIKKLIQVTGQGDREFMAEMETIGKIKHRNLVPLLGYCKIGEERLLVYEYMKYGSLETVLHEKTKKGGIFLDWSARK
KIAIGAARGLAFLHHSCIPHIIHRDMKSSNVLLDQDFVARVSDFGMARLVSALDTHLSVSTLAGTPGYVPPEYYQSFRCT
AKGDVYSYGVILLELLSGKKPIDPEEFGEDNNLVGWAKQLYREKRGAEILDPELVTDKSGDVELLHYLKIASQCLDDRPF
KRPTMIQVMTMFKELVQVDTENDSLDEFLLKETPLVEESRDKEP

>At2g35620.catpro

LGRVESKSLVIDVGGDLPYASKDIIKKLESLNEEHIIGCGGFGTVYKLSMDDGNVFALKRIVKLNEGFDRFFERELEILG
SIKHRYLVNLRGYCNSPTSKLLLYDYLPGGSLDEALHKRGEQLDWDSRVNIIIGAAKGLAYLHHDCSPRIIHRDIKSSNI
LLDGNLEARVSDFGLAKLLEDEESHITTIVAGTFGYLAPEYMQSGRATEKTDVYSFGVLVLEVLSGKLPTDASFIEKGFN
IVGWLNFLISENRAKEIVDLSCEGVERESLDALLSIATKCVSSSPDERPTMHRVVQLLESEVMTPCPSDFYDSSSD

>At2g33170.catpro

TAPYVHDKEPFFQESDIYFVPKERFTVKDILEATKGFHDSYIVGRGACGTVYKAVMPSGKTIAVKKLESNREGNNNNSNN
TDNSFRAEILTLGKIRHRNIVRLYSFCYHQGSNSNLLLYEYMSRGSLGELLHGGKSHSMDWPTRFAIALGAAEGLAYLHH
DCKPRIIHRDIKSNNILIDENFEAHVGDFGLAKVIDMPLSKSVSAVAGSYGYIAPEYAYTMKVTEKCDIYSFGVVLLELL
TGKAPVQPLEQGGDLATWTRNHIRDHSLTSEILDPYLTKVEDDVILNHMITVTKIAVLCTKSSPSDRPTMREVVLMLIES
GERAGKVIVSTTCSDLPPPAPP

>At2g31880.catpro

RGSEKPPGPSIFSPLIKKAEDLAFLENEEALASLEIIGRGGCGEVFKAELPGSNGKIIAVKKVIQPPKDADELTDEDSKF
LNKKMRQIRSEINTVGHIRHRNLLPLLAHVSRPECHYLVYEYMEKGSLQDILTDVQAGNQELMWPARHKIALGIAAGLEY
LHMDHNPRIIHRDLKPANVLLDDDMEARISDFGLAKAMPDAVTHITTSHVAGTVGYIAPEFYQTHKFTDKCDIYSFGVIL
GILVIGKLPSDEFFQHTDEMSLIKWMRNIITSENPSLAIDPKLMDQGFDEQMLLVLKIACYCTLDDPKQRPNSKDVRTML
SQIKH

>At2g26330.catpro

PPFLDGSLDKPVTYSTPKLVILHMNMALHVYEDIMRMTENLSEKYIIGHGASSTVYKCVLKNCKPVAIKRLYSHNPQSMK
QFETELEMLSSIKHRNLVSLQAYSLSHLGSLLFYDYLENGSLWDLLHGPTKKKTLDWDTRLKIAYGAAQGLAYLHHDCSP
RIIHRDVKSSNILLDKDLEARLTDFGIAKSLCVSKSHTSTYVMGTIGYIDPEYARTSRLTEKSDVYSYGIVLLELLTRRK
AVDDESNLHHLIMSKTGNNEVMEMADPDITSTCKDLGVVKKVFQLALLCTKRQPNDRPTMHQVTRVLGSFMLSEQPPAAT
DTSATLAGSCYVDEYANLKTPHSVNCSSMSASDAQLFLRFGQVISQNSE

>At2g02220.catpro

RRSGEVDPEIEESESMNRKELGEIGSKLVVLFQSNDKELSYDDLLDSTNSFDQANIIGCGGFGMVYKATLPDGKKVAIKK
LSGDCGQIEREFEAEVETLSRAQHPNLVLLRGFCFYKNDRLLIYSYMENGSLDYWLHERNDGPALLKWKTRLRIAQGAAK
GLLYLHEGCDPHILHRDIKSSNILLDENFNSHLADFGLARLMSPYETHVSTDLVGTLGYIPPEYGQASVATYKGDVYSFG
VVLLELLTDKRPVDMCKPKGCRDLISWVVKMKHESRASEVFDPLIYSKENDKEMFRVLEIACLCLSENPKQRPTTQQLVS
WLDDV

>At2g01950.catpro

RDADDAKMLHSLQAVNSATTWKIEKEKEPLSINVATFQRQLRKLKFSQLIEATNGFSAASMIGHGGFGEVFKATLKDGSS
VAIKKLIRLSCQGDREFMAEMETLGKIKHRNLVPLLGYCKIGEERLLVYEFMQYGSLEEVLHGPRTGEKRRILGWEERKK
IAKGAAKGLCFLHHNCIPHIIHRDMKSSNVLLDQDMEARVSDFGMARLISALDTHLSVSTLAGTPGYVPPEYYQSFRCTA
KGDVYSIGVVMLEILSGKRPTDKEEFGDTNLVGWSKMKAREGKHMEVIDEDLLKEGSSESLNEKEGFEGGVIVKEMLRYL
EIALRCVDDFPSKRPNMLQVVASLRELRGSENNSHSHSNSL

>At1g78530.catpro

YKRWKRKHTIHENGFPVKGGGKMVMFRSQLLNSVSSDMFMKKTHKLSNKDILGSGGFGTVYRLVIDDSTTFAVKRLNRGT

SERDRGFHRELEAMADIKHRNIVTLHGYFTSPHYNLLIYELMPNGSLDSFLHGRKALDWASRYRIAVGAARGISYLHHDC

IPHIIHRDIKSSNILLDHNMEARVSDFGLATLMEPDKTHVSTFVAGTFGYLAPEYFDTGKATMKGDVYSFGVVLLELLTG

RKPTDDEFFEEGTKLVTWVKGVVRDQREEVVIDNRLRGSSVQENEEMNDVFGIAMMCLEPEPAIRPAMTEVVKLLEYIKL

STRSSF

>CLV1

NKKKNQKSLAWKLTAFQKLDFKSEDVLECLKEENIIGKGGAGIVYRGSMPNNVDVAIKRLVGRGTGRSDHGFTAEIQTLG

RIRHRHIVRLLGYVANKDTNLLLYEYMPNGSLGELLHGSKGGHLQWETRHRVAVEAAKGLCYLHHDCSPLILHRDVKSNN

ILLDSDFEAHVADFGLAKFLVDGAASECMSSIAGSYGYIAPEYAYTLKVDEKSDVYSFGVVLLELIAGKKPVGEFGEGVD

IVRWVRNTEEEITQPSDAAIVVAIVDPRLTGYPLTSVIHVFKIAMMCVEEEAAARPTMREVVHMLTNPPKSVANLIAF

>At1g74360.catpro

AEIDLLDGSKTRHDMTSSSGGSSPWLSGKIKVIRLDKSTFTYADILKATSNFSEERVVGRGGYGTVYRGVLPDGREVAVK

KLQREGTEAEKEFRAEMEVLSANAFGDWAHPNLVRLYGWCLDGSEKILVHEYMGGGSLEELITDKTKLQWKKRIDIATDV

ARGLVFLHHECYPSIVHRDVKASNVLLDKHGNARVTDFGLARLLNVGDSHVSTVIAGTIGYVAPEYGQTWQATTRGDVYS

YGVLTMELATGRRAVDGGEECLVEWARRVMTGNMTAKGSPITLSGTKPGNGAEQMTELLKIGVKCTADHPQARPNMKEVL

AMLVKISGKAELFNGLSSQGYIEM

>At1g73080.catpro

RPEKDAYVFTQEEGPSLLLNKVLAATDNLNEKYTIGRGAHGIVYRASLGSGKVYAVKRLVFASHIRANQSMMREIDTIGK

VRHRNLIKLEGFWLRKDDGLMLYRYMPKGSLYDVLHGVSPKENVLDWSARYNVALGVAHGLAYLHYDCHPPIVHRDIKPE

NILMDSDLEPHIGDFGLARLLDDSTVSTATVTGTTGYIAPENAFKTVRGRESDVYSYGVVLLELVTRKRAVDKSFPESTD

IVSWVRSALSSSNNNVEDMVTTIVDPILVDELLDSSLREQVMQVTELALSCTQQDPAMRPTMRDAVKLLEDVKHLARSCS

SDSVR

>At1g72300.catpro

NPGDSENAELEINSNGSYSEVPPGSDKDISLVLLFGNSRYEVKDLTIFELLKATDNFSQANIIGCGGFGLVYKATLDNGT

KLAVKKLTGDYGMMEKEFKAEVEVLSRAKHENLVALQGYCVHDSARILIYSFMENGSLDYWLHENPEGPAQLDWPKRLNI

MRGASSGLAYMHQICEPHIVHRDIKSSNILLDGNFKAYVADFGLSRLILPYRTHVTTELVGTLGYIPPEYGQAWVATLRG

DVYSFGVVMLELLTGKRPMEVFRPKMSRELVAWVHTMKRDGKPEEVFDTLLRESGNEEAMLRVLDIACMCVNQNPMKRPN

IQQVVDWLKNIEAEKNQNNREEPEEEEET

>At1g72180.catpro

VKIRELDSENRDINKADAKWKIASFHQMELDVDEICRLDEDHVIGSGSAGKVYRVDLKKGGGTVAVKWLKRGGGEEGDGT

EVSVAEMEILGKIRHRNVLKLYACLVGRGSRYLVFEFMENGNLYQALGNNIKGGLPELDWLKRYKIAVGAAKGIAYLHHD

CCPPIIHRDIKSSNILLDGDYESKIADFGVAKVADKGYEWSCVAGTHGYMAPELAYSFKATEKSDVYSFGVVLLELVTGL

RPMEDEFGEGKDIVDYVYSQIQQDPRNLQNVLDKQVLSTYIEESMIRVLKMGLLCTTKLPNLRPSMREVVRKLDDADPCV

SNSQDTTGKITV

>At1g55610.catpro

VQKKEQKREKYIESLPTSGSCSWKLSSVPEPLSINVATFEKPLRKLTFAHLLEATNGFSAETMVGSGGFGEVYKAQLRDG

SVVAIKKLIRITGQGDREFMAEMETIGKIKHRNLVPLLGYCKVGEERLLVYEYMKWGSLETVLHEKSSKKGGIYLNWAAR

KKIAIGAARGLAFLHHSCIPHIIHRDMKSSNVLLDEDFEARVSDFGMARLVSALDTHLSVSTLAGTPGYVPPEYYQSFRC

TAKGDVYSYGVILLELLSGKKPIDPGEFGEDNNLVGWAKQLYREKRGAEILDPELVTDKSGDVELFHYLKIASQCLDDRP

FKRPTMIQLMAMFKEMKADTEEDESLDEFSLKETPLVEESRDKEP

>At1g34110.catpro

RNNHLYKTSQNSSSSPSTAEDFSYPWTFIPFQKLGITVNNIVTSLTDENVIGKGCSGIVYKAEIPNGDIVAVKKLWKTKD

NNEEGESTIDSFAAEIQILGNIRHRNIVKLLGYCSNKSVKLLLYNYFPNGNLQQLLQGNRNLDWETRYKIAIGAAQGLAY

LHHDCVPAILHRDVKCNNILLDSKYEAILADFGLAKLMMNSPNYHNAMSRVAEYGYTMNITEKSDVYSYGVVLLEILSGR

SAVEPQIGDGLHIVEWVKKKMGTFEPALSVLDVKLQGLPDQIVQEMLQTLGIAMFCVNPSPVERPTMKEVVTLLMEVKCS

PEEWGQIWLSGFDFLYIKYHC

>At1g31420.catpro

KKLGKVEIKSLAKDVGGGASIVMFHGDLPYSSKDIIKKLEMLNEEHIIGCGGFGTVYKLAMDDGKVFALKRILKLNEGFD

RFFERELEILGSIKHRYLVNLRGYCNSPTSKLLLYDYLPGGSLDEALHERGEQLDWDSRVNIIIGAAKGLSYLHHDCSPR

IIHRDIKSSNILLDGNLEARVSDFGLAKLLEDEESHITTIVAGTFGYLAPEYMQSGRATEKTDVYSFGVLVLEVLSGKRP

TDASFIEKGLNVVGWLKFLISEKRPRDIVDPNCEGMQMESLDALLSIATQCVSPSPEERPTMHRVVQLLESEVMTPCPSE

FYDSSSD

>At1g28440.catpro

RTFKKARAMERSKWTLMSFHKLGFSEHEILESLDEDNVIGAGASGKVYKVVLTNGETVAVKRLWTGSVKETGDCDPEKGY

KPGVQDEAFEAEVETLGKIRHKNIVKLWCCCSTRDCKLLVYEYMPNGSLGDLLHSSKGGMLGWQTRFKIILDAAEGLSYL

HHDSVPPIVHRDIKSNNILIDGDYGARVADFGVAKAVDLTGKAPKSMSVIAGSCGYIAPEYAYTLRVNEKSDIYSFGVVI

LEIVTRKRPVDPELGEKDLVKWVCSTLDQKGIEHVIDPKLDSCFKEEISKILNVGLLCTSPLPINRPSMRRVVKMLQEIG

GGDEDSLHKIRDDKDGKLTPYYNEDTSDQGSIA

>At1g17750.catpro

RGTKTEDANILAEEGLSLLLNKVLAATDNLDDKYIIGRGAHGVVYRASLGSGEEYAVKKLIFAEHIRANQNMKREIETIG

LVRHRNLIRLERFWMRKEDGLMLYQYMPNGSLHDVLHRGNQGEAVLDWSARFNIALGISHGLAYLHHDCHPPIIHRDIKP

ENILMDSDMEPHIGDFGLARILDDSTVSTATVTGTTGYIAPENAYKTVRSKESDVYSYGVVLLELVTGKRALDRSFPEDI

NIVSWVRSVLSSYEDEDDTAGPIVDPKLVDELLDTKLREQAIQVTDLALRCTDKRPENRPSMRDVVKDLTDLESFVRSTS

GSVH

>At1g17230.catpro

PAFVALEDQTKPDVMDSYYFPKKGFTYQGLVDATRNFSEDVVLGRGACGTVYKAEMSGGEVIAVKKLNSRGEGASSDNSF

RAEISTLGKIRHRNIVKLYGFCYHQNSNLLLYEYMSKGSLGEQLQRGEKNCLLDWNARYRIALGAAEGLCYLHHDCRPQI

VHRDIKSNNILLDERFQAHVGDFGLAKLIDLSYSKSMSAVAGSYGYIAPEYAYTMKVTEKCDIYSFGVVLLELITGKPPV

QPLEQGGDLVNWVRRSIRNMIPTIEMFDARLDTNDKRTVHEMSLVLKIALFCTSNSPASRPTMREVVAMITEARGSSSLS

SSSITSETPLEEANSSKGMYLHIHTHTHTLLCNFRTFCDSHMNISCYSAEI

>At1g09970.catpro

YLKKTEKKEGRSLKHESWSIKSFRKMSFTEDDIIDSIKEENLIGRGGCGDVYRVVLGDGKEVAVKHIRCSSTQKNFSSAM

PILTEREGRSKEFETEVQTLSSIRHLNVVKLYCSITSDDSSLLVYEYLPNGSLWDMLHSCKKSNLGWETRYDIALGAAKG

LEYLHHGYERPVIHRDVKSSNILLDEFLKPRIADFGLAKILQASNGGPESTHVVAGTYGYIAPGKKPIEAEFGESKDIVN

WVSNNLKSKESVMEIVDKKIGEMYREDAVKMLRIAIICTARLPGLRPTMRSVVQMIEDAEPCRLMGIVISKESDVKVKEI

>At4g08850

IKATGEFDPKYLIGTGGHGKVYKAKLPNAIMAVKKLNETTDSSISNPSTKQEFLNEIRALTEIRHRNVVKLFGFCSHRRN
TFLVYEYMERGSLRKVLENDDEAKKLDWGKRINVVKGVAHALSYMHHDRSPAIVHRDISSGNILLGEDYE

AKISDFGTAKLLKPDSSNWSAVAGTYGYVAPELAYAMKVTEKCDVYSFGVLTLEVIKGEHPGDLVSTLSSSPPDATLSLK
SISDHRLPEPTPEIKEEVLEILKVALLCLHSDPQARPTML

SISTAFS

>At1g35710

FDPTHLIGTGGYSKVYRANLQDTIIAVKRLHDTIDEEISKPVVKQEFLNEVKALTEIRHRNVVKLFGFCSHRRHTFLIYE
YMEKGSLNKLLANDEEAKRLTWTKRINVVKGVAHALSYMHHDRITPIVHRDISSGNILLDNDYTAKISDFGTAKLLKTDS
SNWSAVAGTYGYVAPEFAYTMKVTEKCDVYSFGVLILELIIGKHPGDLVSSLSSSPGEALSLRSISDERVLEPRGQNREK
LLKMVEMALLCLQANPESRPTMLSISTTFS

>At1g34210

FSNKNILGRGGFGKVYKGRLADGTLVAVKRLKEERTPGGELQFQTEVEMISMAVHRNLLRLRGFCMTPTERLLVYPYMAN
GSVASCLRERPPSQLPLAWSIRQQIALGSARGLSYLHDHCDPKIIHRDVKAANILLDEEFEAVVGDFGLARLMDYKDTHV
TTAVRGTIGHIAPEYLSTGKSSEKTDVFGYGIMLLELITGQRAFDLARLANDDDVMLLDWVKGLLKEKKLEMLVDPDLQS
NYTEAEVEQLIQVALLCTQSSPMERPKMSEVVRMLEGDGLAEKWDEWQKVEVLRQEVELSSHPTSDWILDSTDNLHAMEL
SGPR

>At1g71830

FSNKNILGRGGFGKVYKGRLADGTLVAVKRLKEERTPGGELQFQTEVEMISMAVHRNLLRLRGFCMTPTERLLVYPYMAN
GSVASCLRERPPSQPPLDWPTRKRIALGSARGLSYLHDHCDPKIIHRDVKAANILLDEEFEAVVGDFGLAKLMDYKDTHV
TTAVRGTIGHIAPEYLSTGKSSEKTDVFGYGIMLLELITGQRAFDLARLANDDDVMLLDWVKGLLKEKKLEMLVDPDLQT
NYEERELEQVIQVALLCTQGSPMERPKMSEVVRMLEGDGLAEKWDEWQKVEILREEIDLSPNPNSDWILDSTYNLHAVEL
SGPR

>BAK1

FSNKNILGRGGFGKVYKGRLADGTLVAVKRLKEERTQGGELQFQTEVEMISMAVHRNLLRLRGFCMTPTERLLVYPYMAN
GSVASCLRERPESQPPLDWPKRQRIALGSARGLAYLHDHCDPKIIHRDVKAANILLDEEFEAVVGDFGLAKLMDYKDTHV
TTAVRGTIGHIAPEYLSTGKSSEKTDVFGYGVMLLELITGQRAFDLARLANDDDVMLLDWVKGLLKEKKLEALVDVDLQG
NYKDEEVEQLIQVALLCTQSSPMERPKMSEVVRMLEGDGLAERWEEWQKEEMFRQDFNYPTHHPAVSGWIIGDSTSQIEN
EYPSGPR

>At2g13790

FSNKNVLGRGGFGKVYKGRLADGNLVAVKRLKEERTKGGELQFQTEVEMISMAVHRNLLRLRGFCMTPTERLLVYPYMAN
GSVASCLRERPEGNPALDWPKRKHIALGSARGLAYLHDHCDQKIIHRDVKAANILLDEEFEAVVGDFGLAKLMNYNDSHV
TTAVRGTIGHIAPEYLSTGKSSEKTDVFGYGVMLLELITGQKAFDLARLANDDDIMLLDWVKEVLKEKKLESLVDAELEG
KYVETEVEQLIQMALLCTQSSAMERPKMSEVVRMLEGDGLAERWEEWQKEEMPIHDFNYQAYPHAGTDWLIPYSNSLIEN
DYPSGPR

>At2g13800

FSKRNVLGKGRFGILYKGRLADDTLVAVKRLNEERTKGGELQFQTEVEMISMAVHRNLLRLRGFCMTPTERLLVYPYMAN
GSVASCLRERPEGNPALDWPKRKHIALGSARGLAYLHDHCDQKIIHLDVKAANILLDEEFEAVVGDFGLAKLMNYNDSHV

TTAVRGTIGHIAPEYLSTGKSSEKTDVFGYGVMLLELITGQKAFDLARLANDDDIMLLDWVKEVLKEKKLESLVDAELEG
KYVETEVEQLIQMALLCTQSSAMERPKMSEVVRMLEGDGLAERWEEWQKEEMPIHDFNYQAYPHAGTDWLIPYSNSLIEN
DYPSGPR

>At5g65240

FSEKNVLGQGGFGKVYKGLLSDGTKVAVKRLTDFERPGGDEAFQREVEMISVAVHRNLLRLIGFCTTQTERLLVYPFMQN
LSVAYCLREIKPGDPVLDWFRRKQIALGAARGLEYLHEHCNPKIIHRDVKAANVLLDEDFEAVVGDFGLAKLVDVRRTNV
TTQVRGTMGHIAPECISTGKSSEKTDVFGYGIMLLELVTGQRAIDFSRLEEEDDVLLLDHVKKLEREKRLEDIVDKKLDE
DYIKEEVEMMIQVALLCTQAAPEERPAMSEVVRMLEGEGLAERWEEWQNLEVTRQEEFQRLQRRFDWGEDSINNQDAIEL
SGGR

>At5g10290

FSEKNVLGQGGFGKVYKGVLPDNTKVAVKRLTDFESPGGDAAFQREVEMISVAVHRNLLRLIGFCTTQTERLLVYPFMQN
LSLAHRLREIKAGDPVLDWETRKRIALGAARGFEYLHEHCNPKIIHRDVKAANVLLDEDFEAVVGDFGLAKLVDVRRTNV
TTQVRGTMGHIAPEYLSTGKSSERTDVFGYGIMLLELVTGQRAIDFSRLEEEDDVLLLDHVKKLEREKRLGAIVDKNLDG
EYIKEEVEMMIQVALLCTQGSPEDRPVMSEVVRMLEGEGLAERWEEWQNVEVTRRHEFERLQRRFDWGEDSMHNQDAIEL
SGGR

>At5g63710

FNESNLIGQGGFGKVYRGLLPDKTKVAVKRLADYFSPGGEAAFQREIQLISVAVHKNLLRLIGFCTTSSERILVYPYMEN
LSVAYRLRDLKAGEEGLDWPTRKRVAFGSAHGLEYLHEHCNPKIIHRDLKAANILLDNNFEPVLGDFGLAKLVDTSLTHV
TTQVRGTMGHIAPEYLCTGKSSEKTDVFGYGITLLELVTGQRAIDFSRLEEEENILLLDHIKKLLREQRLRDIVDSNLTT
YDSKEVETIVQVALLCTQGSPEDRPAMSEVVKMLQGTGGLAEKWTEWEQLEEVRNKEALLLPTLPATWDEEETTVDQESI
RLSTAR

>At2g23950

FSSKSILGAGGFGNVYRGKFGDGTVVAVKRLKDVNGTSGNSQFRTELEMISLAVHRNLLRLIGYCASSSERLLVYPYMSN
GSVASRLKAKPALDWNTRKKIAIGAARGLFYLHEQCDPKIIHRDVKAANILLDEYFEAVVGDFGLAKLLNHEDSHVTTAV
RGTVGHIAPEYLSTGQSSEKTDVFGFGILLLELITGMRALEFGKSVSQKGAMLEWVRKLHKEMKVEELVDRELGTTYDRI
EVGEMLQVALLCTQFLPAHRPKMSEVVQMLEGDGLAERWAASHDHSHFYHANMSYRTITSTDGNNQTKHLFGSSGFEDED
DNQALDSFA

MELSGPR

>At4g30520

FSSKNILGAGGFGNVYRGKLGDGTMVAVKRLKDINGTSGDSQFRMELEMISLAVHKNLLRLIGYCATSGERLLVYPYMPN
GSVASKLKSKPALDWNMRKRIAIGAARGLLYLHEQCDPKIIHRDVKAANILLDECFEAVVGDFGLAKLLNHADSHVTTAV
RGTVGHIAPEYLSTGQSSEKTDVFGFGILLLELITGLRALEFGKTVSQKGAMLEWVRKLHEEMKVEELLDRELGTNYDKI
EVGEMLQVALLCTQYLPAHRPKMSEVVLMLEGDGLAERWAASHNHSHFYHANISFKTISSLSTTSVSRLDAHCNDPTYQM
FGSSAFDDDDDHQPLDSFAMELSGPR

>At3g25560

FSSKNLVGKGGFGNVYKGCLHDGSIIAVKRLKDINNGGGEVQFQTELEMISLAVHRNLLRLYGFCTTSSERLLVYPYMSN
GSVASRLKAKPVLDWGTRKRIALGAGRGLLYLHEQCDPKIIHRDVKAANILLDDYFEAVVGDFGLAKLLDHEESHVTTAV
RGTVGHIAPEYLSTGQSSEKTDVFGFGILLLELITGLRALEFGKAANQRGAILDWVKKLQQEKKLEQIVDKDLKSNYDRI
EVEEMVQVALLCTQYLPIHRPKMSEVVRMLEGDGLVEKWEASSQRAETNRSYSKPNEFSSSERYSDLTDDSSVLVQAMEL
SGPR

>At5g16000

FSSKNLLGKGGYGNVYKGILGDSTVVAVKRLKDGGALGGEIQFQTEVEMISLAVHRNLLRLYGFCITQTEKLLVYPYMSN
GSVASRMKAKPVLDWSIRKRIAIGAARGLVYLHEQCDPKIIHRDVKAANILLDDYCEAVVGDFGLAKLLDHQDSHVTTAV
RGTVGHIAPEYLSTGQSSEKTDVFGFGILLLELVTGQRAFEFGKAANQKGVMLDWVKKIHQEKKLELLVDKELLKKKSYD
EIELDEMVRVALLCTQYLPGHRPKMSEVVRMLEGDGLAEKWEASQRSDSVSKCSNRINELMSSSDRYSDLTDDSSLLVQA
MELSGPR

>At1g60800

FNSKNILGRGGYGIVYKGHLNDGTLVAVKRLKDCNIAGGEVQFQTEVETISLALHRNLLRLRGFCSSNQERILVYPYMPN
GSVASRLKDNIRGEPALDWSRRKKIAVGTARGLVYLHEQCDPKIIHRDVKAANILLDEDFEAVVGDFGLAKLLDHRDSHV
TTAVRGTVGHIAPEYLSTGQSSEKTDVFGFGILLLELITGQKALDFGRSAHQKGVMLDWVKKLHQEGKLKQLIDKDLNDK

FDRVELEEIVQVALLCTQFNPSHRPKMSEVMKMLEGDGLAERWEATQNGTGEHQPPPLPPGMVSSSPRVRYYSDYIQESS
LVVEAIELSGPR

>At5g45780

FSPKNILGQGGFGMVYKGYLPNGTVVAVKRLKDPIYTGEVQFQTEVEMIGLAVHRNLLRLFGFCMTPEERMLVYPYMPNG
SVADRLRDNYGEKPSLDWNRRISIALGAARGLVYLHEQCNPKIIHRDVKAANILLDESFEAIVGDFGLAKLLDQRDSHVT
TAVRGTIGHIAPEYLSTGQSSEKTDVFGFGVLILELITGHKMIDQGNGQVRKGMILSWVRTLKAEKRFAEMVDRDLKGEF
DDLVLEEVVELALLCTQPHPNLRPRMSQVLKVLEGLVEQCEGGYEARAPSVSRNYSNGHEEQSFIIEAIELSGPR

VITA

Zhaofeng Ding was born on August 12th, 1978, at Chongqing, China. After attending public schools from 1986 to 1997 in Chongqing, he received his B.S. degree from Biochemistry department of Nanjing University of China in 2001. He is currently finishing his Ph.D. degree in Biochemistry at University of Missouri-Columbia.