FUNCTIONAL SIGNIFICANCE OF EXTRACELLULAR SIGNAL REGULATED KINASE (ERK2) PHOSPHORYLATION STATES: IMPLICATIONS FOR DNA BINDING

APPROVED BY SUPERVISORY COMMITTEE
(14 spaces down)

Melanie H. Cobb Ph.D.

Luke Rice, Ph.D.

Gang Yu, Ph.D.

Joseph Albanesi, Ph.D.

DEDICATION

Simply, in memory of my grandmother, Frances Jean McReynolds.

FUNCTIONAL SIGNIFICANCE OF EXTRACELLULAR SIGNAL REGULATED KINASE (ERK2) PHOSPHORYLATION STATES: IMPLICATIONS FOR DNA BINDING

by

Andrea Christine McReynolds

DISSERTATION / THESIS

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

December 2013

FUNCTIONAL SIGNIFICANCE OF EXTRACELLULAR SIGNAL REGULATED KINASE (ERK2) PHOSPHORYLATION STATES: IMPLICATIONS FOR DNA BINDING

Publication No. _____

Andrea Christine McReynolds, Ph.D.

The University of Texas Southwestern Medical Center at Dallas, 2013

Melanie H. Cobb, 2013

The protein kinase extracellular signal-regulated kinase 2 (ERK2) has been well understood structurally for nearly twenty years. New insight is emerging about its structure and function. A novel autophosphorylation has recently been found to occur on a critical active site residue, threonine 188. Autophosphorylation of this residue has been suggested to occur as a result of the confluence of receptor tyrosine kinase (RTK) and G protein-coupled receptor (GPCR) signaling pathways. The possibility of autophosphorylation on threonine 188 seemed to be inconsistent with what was known about the structure and function of protein kinases generally and ERK2 in particular; T188 and the comparable residue in other protein kinases is required for catalytic activity. I found ERK2 phosphorylated on T188 *in vitro* in partially purified preparations of the recombinant protein purified following expression in bacteria. This suggests that phosphorylation of ERK2 on T188 can occur without the input of upstream RTK or GPCR signaling. Through mutagenesis experiments, I found mutation of T188 sharply reduced activity toward substrates in vitro. Protein fractions containing pT188 and purified ERK2 T188D and T188E mutants, that may act as phosphomimetics, appear to have an increased affinity for DNA binding. I examined critical residues in the activation loop important for phosphorylation

and activation of ERK2 and found that perturbation of these residues influences DNA binding specificity. Although the MAPK pathway and role of ERK2 is well understood, our data suggest that previously unrecognized, higher-order signaling mechanisms that arise from additional phosphorylation events may be involved in less well characterized properties of ERK2. Second, our finding that T188 autophosphorylation can occur in recombinant ERK2 independently from upstream GPCR signaling illustrates the need for a reinvestigation of the regulation of ERK2 autophosphorylation. Finally, direct and specific DNA binding may be driven by differences in phosphorylation. The result of this work serves to redefine a most important signaling molecule in terms of its structural modifications and relationship to overall function.

TABLE OF CONTENTS

PRIOR PUBLICATIONS

1. Thomas, V.L., **McReynolds A.C**., Shoichet BK. (2010) Structural bases for stability function trade-offs in antibiotic resistance. J Mol. Bio. 396 (1): 47-59. PMID 19913034

2. Chen, Y., **McReynolds A.C**., Shoichet, B.K. (2009). Re-examining the role of Lys 67 in class C- βeta lactamase catalysis. Protein Science 18: 662-669. PMID 19241376

3. Mobley D.L., Graves A.P., Chodera J.D., **McReynolds A.C**., Shoichet B.K., Dill K.A. (2007) Predicting absolute ligand binding free energies to a simple model binding site. J. Mol Biol. Aug 24; 371(4):1118-34. PMID 17599350

4. Arnold, L.A., Estébanez-Perpiñá, E., Togashi, M., Shelat, A., Ocasio, C.A., **McReynolds, A.C.**, Phuong Nguyen, Baxter, J.D., Fletterick, R.J., Webb, P., Guy, R.K., (2006). A High-Throughput Screening method to Identify Small Molecule Inhibitors of Thyroid Hormone Receptor Co-activator Binding. Sci. STKE, Vol. 2006, Issue 341, pp. pl3, 27 June 2006.
PMID 16804159

5. Arnold L.A., Estebanez-Perpina E., Togashi M., Jouravel N, Shelat A, **McReynolds A.C**., Mar E., Nguyen P, Baxter J.D., Fletterick R.J., Webb P, Guy R.K. (2005). Discovery of small molecule inhibitors of the interaction of the thyroid hormone receptor with transcriptional co-regulators. J Biol. Chem. 280 (52): 43048-55. PMID 16263725

6. Geistlinger T.R., **McReynolds A.C**., Guy R.K. (2004) Ligand selective inhibition of the interaction of steroid receptor co-activator and estrogen receptor isoforms. Chem Biol. 11, 273-81. PMID 15123288

7. Moore, J.M., Galicia, S.J., **McReynolds, A.C**., Nguyen, N., Scanlan, T.S., and Guy, R.K.(2004) Quantitative Proteomics of the Thyroid Hormone Receptor-Coregulator Interactions. J. Biol. Chem. 279: 19832-19838. PMID 15100213

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

LIST OF DEFINITIONS

ppERK – doubly phosphorylated ERK

R4F MEK1 – Constitutively active MEK1

IEX – Ion Exchange Chromatography (Mono Q Anion used in this work)

pT188 – ERK2 phosphorylated on Threonine 188

pY185 – ERK2 phosphorylated on the activation loop residue, Tyrosine 185

pT183 – ERK2 phosphorylated on the activation loop residue, Threonine 183

MBP – Myelin Basic protein (substrate of ERK1/2)

GPCR – G Protein Coupled Receptor

Peak 1 – first eluting peak from a Mono Q Anion exchange experiment

Peak 2 – second eluting peak from a Mono Q Anion exchange experiment

ChIP– Chromatin Immunoprecipitation

ChIP-seq – Chromatin Immunoprecipitation followed by sequencing

ChIP-Chip – Chromatin Immunoprecipitation followed by microarray analysis

ITC – Isothermal Titration Calorimetry

(PTX) – Pertussis toxin

EGF – Epidermal growth Factor (stimulates MAPK pathway

Chapter 1
Introduction

1.1 Overview of the mitogen-activated protein kinase (MAPK) signaling pathway

The MAPK ERK1/2 pathway responds to cellular need for growth, differentiation, and metabolism. Extracellular ligands such as epidermal growth factor (EGF), lysophosphatidic acid (LPA) and hormones such as the insulin secretagogue, glucagon-like peptide 1 (GLP1), bind to trans-membrane receptors and activate the MAPK pathway. Among many types of receptors that can activate MAPK cascades, receptor tyrosine kinases (RTKs) and G-protein coupled receptors (GPCRs) activate the pathways and are by far the most thoroughly studied molecular mechanisms. For purposes that will become clear later in this dissertation, the convergence of these pathways on the MAPK cascade is of particular interest.

The mechanism outlined for the EGF receptor will be used as an example of how the MAPK pathway is activated following ligand binding to an RTK. Once ligand binds to the receptor, the receptor dimer becomes activated by an intra-molecular allosteric mechanism (Zhang et al., 2006). The receptor auto-phosphorylates on tyrosine residues within its cytoplasmic domain and creates a phosphotyrosine binding site to recruit adaptor proteins and guanine nucleotide exchange factors. Once recruited to the membrane, this set of effectors binds to the small, monomeric G protein Ras and initiates nucleotide exchange of GDP for GTP. GTP loaded Ras is now able to initiate the phosphorylation cascade by recruiting the MAP kinase kinase kinase (MAP3K), Raf-1 to the plasma membrane to facilitate its activation by phosphorylation. Activated Raf, in turn, phosphorylates the downstream MAP kinase kinase (MAP2K), MEK1/2, by phosphorylation. The relay continues with MEK1/2 phosphorylating their only known substrates, ERK1/2. Phosphorylation and activation of ERK1/2 results in the ability to phosphorylate a myriad of

effector proteins important for cellular regulation events such as cytoskeleton dynamics, nuclear import, nuclear hormone receptor gene regulation, stem cell pluripotency, translation and, of course, cellular proliferation. It is currently estimated that ERK1/2 has over 500 substrates.

1.2 MAPK pathways and signaling by G proteins

Concomitant with MAPK pathway activation by RTK inputs, is the activation of MAPK pathways by G Protein Coupled Receptors (GPCR). G proteins are coupled to seven transmembrane receptors through isoprenylation (Mumby et al., 1990). G proteins are characteristically composed of three subunits, α, β and γ. β and γ subunits are extensively linked through multiple WD motifs effectively existing as an irreversible dimer (Wall et al., 1995) and the α subunit makes contacts with membrane receptors (Conklin and Bourne, 1993). Activation of receptors, and therefore their conformational change, results in subsequent activation of associated G proteins. Activation of G proteins is characterized by the rapid exchange of GTP for GDP facilitated by guanine exchange factors (GEFs) in G protein α subunits. The energy-liberating hydrolysis of GTP by associated G protein associated proteins (GAPs) also aids in their activation. Once G proteins are activated, the trimeric $\alpha\beta\gamma$ subunits dissociate to α or $\beta\gamma$ subunits. The α or $\beta\gamma$ subunits can then function as singular proteins (α) or dimeric complexes ($\beta\gamma$) which feed into signaling pathways to continue signal transduction.

The confluence of G protein and MAPK pathway signaling has been appreciated for some time (Crespo and Gutkind, 2004; Crespo et al., 1994). G protein signaling into the ERK1/2 MAPK pathway was first characterized when pertussis toxin (PTX) mediated inhibition of M2 muscarinic receptors coupled to G_i proteins caused a dose-dependent decrease in the ERK1/2 MAPK pathway activation (Crespo et al., 1994). Dose dependent activation of the ERK1/2 pathway was also seen in response to increasing carbacol

concentrations which activate Gq coupled receptors. In this same work, overexpression of βγ subunits, but not α subunits, was sufficient to activate ERK1/2 suggesting that G proteins serve to potentiate MAPK signaling in addition to activation by RTK signaling.

The mechanism for how G proteins signaling feeds into the MAPK pathway derived from experiments which determined that signaling pathways can connect through Protein kinase C (PKC). Raf-1, the MAP3K in the ERK1/2 pathway, is phosphorylated by PKC (Kolch et al., 1993). Over-expression studies of PKCα were shown to be sufficient for induction of ERK1/2 activity in vitro by targeting c-Raf-1 and activating the pathway (Schonwasser et al., 1998).

Interestingly, GPCRs can couple to more than one G protein and affect downstream signaling. Bradykinin receptors (B_2R) have been found to couple to Gαq/11, Gαi/0, Gα and Gα12/13 proteins. Stimulation of B2R with Bradykinin agonist in HEK293T cells showed MEK1/2 activation in a dose responsive pattern (Blaukat et al., 2000). However, when Raf mutants were transfected into these cells and given 1uM agonist, loss of ERK1/2 pathway activation correlated linearly. The authors also suspected that muscarinic receptors (isoforms, M2 or M3) coupled to Gi and Gq fed into the ERK1/2 pathway as well. To test the hypothesis that more than one G protein pathway can activate ERK1/2, M2 or M3 receptors and an HA tagged form of ERK2 were overexpressed. When M2/M3 receptors were inhibited by PTX and cells were stimulated with carbachol, an agonist of Gq associated receptors, activation of ERK1/2 as measured by MBP phosphorylation was abolished. These experiments would demonstrate that, in some cellular contexts, additive G protein signaling exists to mediate ERK1/2 pathway activation.

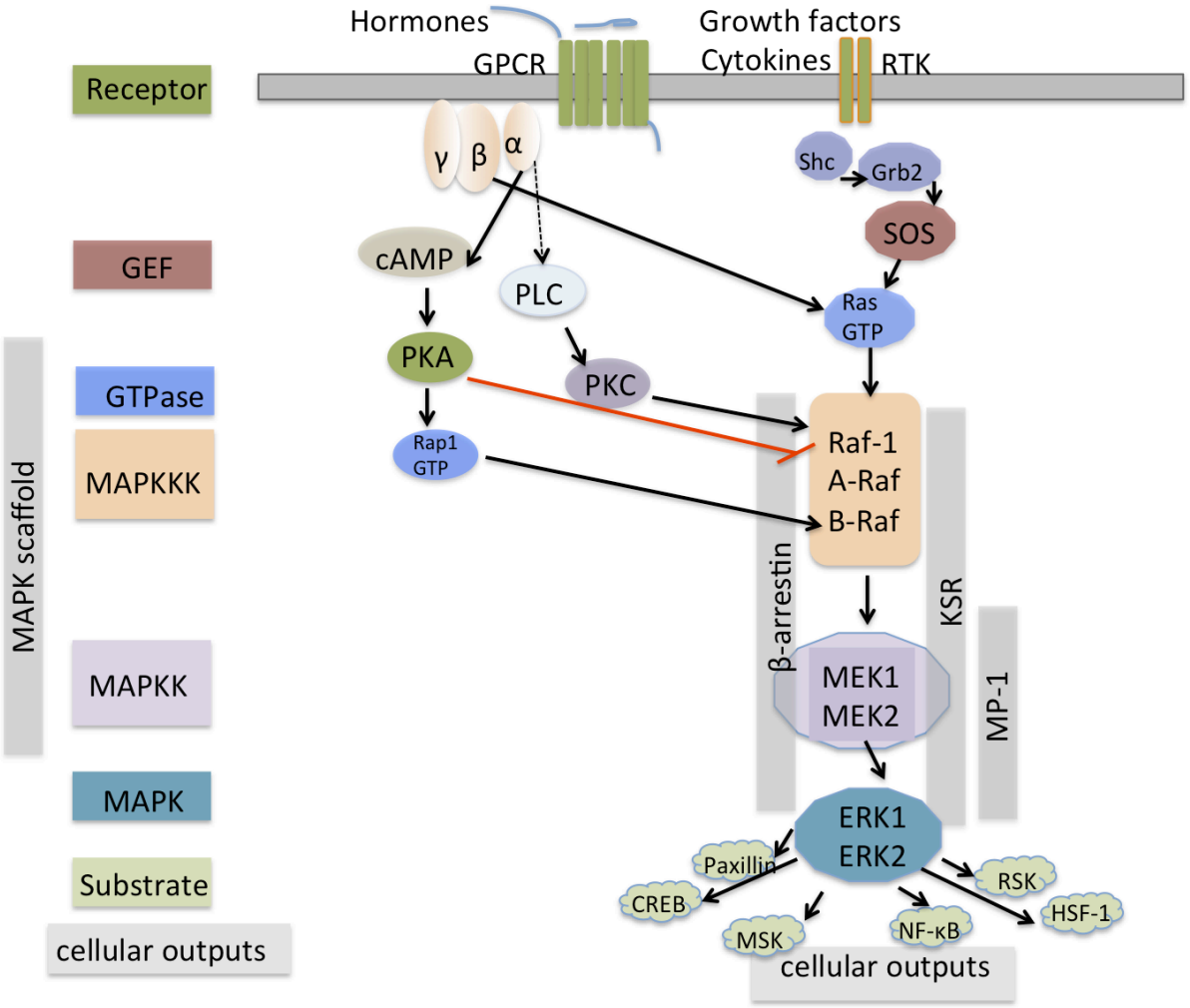


Figure 1.1 The MAPK ERK1/2 pathway is potentiated by RTK and GPCR signaling. Adapted from Qi and Elion, (Qi and Elion, 2005).

1.3 General Protein Kinase Structure and Function

Phosphorylation is a major mediator of cellular signal transduction. A substantial number of proteins expressed in eukaryotic cells, more than 500 in mammals, are protein kinases. Within a eukaryotic genome, it is now estimated that kinases represent about 1-2% of genes (Manning et al., 2002a). Prokaryotic protein kinases do not share sequence identity with eukaryotic protein kinases and will not be considered here. Prior to the 1980s, primary sequences of protein kinases were not available, and the earliest sequences were obtained by protein sequencing methods. With the development of cloning methods, protein kinase sequences began to accumulate. Sequence alignments of protein kinases from many eukaryotic species showed that protein kinases possessed conserved sequence elements and the large number of enzymes in the protein kinase superfamily began to be identified. Protein kinases are enzymes that 1) are subject to regulation, 2) prefer substrates based on their sequence specificity and 3) upon their activation carry out enzymatic phosphorylation on bound substrates. Protein kinases have varying modes of regulation and cellular functions. Early classification of protein kinases sought to reflect these differences. For example, the eukaryotic protein kinase classification scheme of Hanks and Hunter subdivides the family into branches based on sequence similarity. Initially, classification used five groups. The AGC group contains calcium and cAMP dependent kinases such as PKA and PKC. The CaMK group contains protein kinases regulated by calcium and calmodulin and similar protein kinases. The CMGC group consists of proteins within the cyclin-dependent kinase (CDK), MAP kinase, Glycogen synthase, and Clk families (Hanks and Hunter, 1995). Their associated functions defined, the number of groups within the classification scheme increased by four to include: STE family of kinases, which includes MEKs, a number of MAP3Ks, PAKs, etc.; CK1, TKL (tyrosine kinase like), and the RGC (receptor guanylate-cyclase) group which contain kinases similar to the TKL group (Manning et al., 2002b) .

Sequence alignments of protein kinases illustrated the catalytic domain similarity even

before crystallographic structures were available. However, allosteric changes upon binding to substrate and the subsequent mechanism of catalysis for each individual enzyme are best examined by a complete representation of their structures—preferably in low and high activity forms. Crystallization of protein kinases from several groups within the kinase superfamily, shows that a characteristic kinase fold important for overall structure and catalytic activity. Seminal crystallographic studies began with the elucidation of the cyclic AMP dependent protein kinase (cAPK) catalytic domain structure (Knighton et al., 1991). Subsequent to the elucidation of the active cAPK structure, the structures of the low activity ERK2 [pdb:1ERK] and cyclin-dependent protein kinase (CDK2) [pdb:1B38] followed. The structures of cAPK, ERK2 and CDK2 showed that protein kinases possess an overall characteristic fold consisting of two major domains: an amino and a carboxy terminal domain. Taking a closer look at all three structures, the similarity of each active site is further appreciated. The importance of both the amino and carboxy terminal domains in each structure are now considered in detail to illustrate the structural similarity between three different protein kinases and to further draw the parallels between similar structure and function.

Within the protein kinases, the amino terminal domain is composed of five, anti-parallel β-sheet strands and at least one α-helix (helix C). α-helix C forms a major part of the active site and has two major functions. First, α-helix C binds to a residue critical for catalysis, the catalytic lysine, possibly to neutralize charge and aid lysine in its important role in coordinating the α and β phosphates of ATP. α-helix C is also important for regulating the open and closed forms of the active site in each protein kinase. α-helix C can make contacts with the activation loop of each enzyme in the closed conformation and is repositioned to interact with the catalytic lysine in the open conformation (in ERK2 α-helix C makes contacts with a triad of residues, DFG, which perpetuates the open conformation) (Zhang et al., 1994).

Comparison of cAPK, ERK2 and CDK2 structures show several characteristic

residues (Asp 184 (cAPK), Asp 145 (CDK2), and Asp 165 (ERK)) all act to coordinate one of two $Mg2^{+}$ ions bonded to the β and γ phosphates of ATP within the active site. The active site of each molecule contains a lysine that is critical for kinase activity (Lys 33 (CDK2); Lys72 (cAPK); Lys 52 (ERK2)). The catalytic lysine coordinates the α and β phosphates of ATP. Experiments mutating the catalytic lysine in ERK2 (K52R) show that ATP is misaligned within the active site, its associated Mg^{2+} ion is prevented from the necessary coordination by chelating residues (Robinson et al., 1996). The glycine rich loop consists of roughly six residues which are situated in the active site and serve to 'anchor' the α and β phosphates of ATP which consequently poise the γ- phosphate for phosphotransfer.

Positioned between the amino and carboxy termini of the molecule is the site of activation. Phosphorylation of key residues (Y185 and Thr 183 in ERK2; Thr160 in CDK2) must occur for full activation of the molecule. Each protein kinase possesses an activation lip, or loop, which contains residues phosphorylated by an upstream kinase resulting in activation. This activation loop sits at the interface of the amino and carboxy termini. Apart from this, the carboxy terminal domain also contains substrate-binding regions important for the selective binding of client proteins and/or peptides. The overall kinase fold can be seen from the structural similarities in these protein kinases. The remainder of this thesis will focus on the effects of the perturbation of ERK2 structure and the resulting function.

1.4 MAPK ERK2 Structure and Activation

As previously described, the structure of ERK2 has been solved and has the characteristic fold of the kinase superfamily possessing a two-domain structure. Crystallographic data has shown that ERK2 can exist in a low-activity (non-phosphorylated) conformation (Zhang et al., 1994) and high-activity (doubly phosphorylated) conformation (Canagarajah et al., 1997). The low-activity structure is characterized by an open conformation in which the amino and carboxy terminal domains are held apart by α helix C. The un-phosphorylated activation lip in the low activity structure makes contacts with the small α-helical structure unique to MAP kinases—the MAPK insert region.

A three amino acid motif within the phosphorylation lip, TEY, serves as the site for activating phosphorylation. Phosphorylation of threonine and tyrosine are required for full activation (Prowse and Lew, 2001; Rossomando et al., 1992). Phospho-peptide mapping experiments found Tyr185, although buried within a hydrophobic groove in the un-phosphorylated, low-activity conformation, is the first residue phosphorylated by MEK1/2 (Robbins et al., 1993). Thus, upon binding the upstream activating kinase, MEK1/2, ERK becomes first phosphorylated on Tyr185 causing loop rearrangement and liberation of Thr183. Thr183 then becomes subsequently phosphorylated after MEK1/2 rebinds to ERK through a distributive mechanism resulting in the fully phosphorylated form of the enzyme (herein designated as ppERK) (Ferrell and Bhatt, 1997). Unlike CDK2 and cAPK, ERK2 requires dual-phosphorylation resulting in full activation of the enzyme. Phosphorylation of Y185 results in liberation from its buried position and causes a reorientation such that it is able to bind substrate and make contacts within the substrate recognition motif (discussed below).

In addition to dual phosphorylation, activation is characterized by a rotation of the amino and carboxy terminal domains rotating toward each other. This rotation functions to reorganize the active site such that alignment of the catalytic lysine occurs with the active site residues D147 and D165. Hydrogen deuterium exchange experiments comparing the

rate of deuterium incorporation into ERK2 between its phosphorylated and un-phosphorylated states, showed that there was a significant decrease in deuterium incorporation upon the binding of a non-hydrolyzable nucleotide analog, AMP-PNP. The reduction of deuterium incorporation was mapped to a motif critical for binding to nucleotide (or analog) and facilitating inter-domain closure (Lee et al., 2005). This would suggest that upon binding ATP, or its associated analog, the conformational flexibility of the enzyme is lost which is characteristic of a closed, immobile conformation of the active protein. Using molecular dynamics simulations, domain closure was seen for active ERK2 (including wt and auto-activated mutants) but not for the inactive form of the enzyme further showing that rotation of the two domains is critical for activation (Barr et al., 2011) although small changes in conformation exist between ERK1 and ERK2 active and inactive forms (Ring, 2011). Yet another feature of ppERK is the loss of contacts between the activation loop residues and the MAPK insert, a region of the molecule suggested to make contacts with DNA, the implications of which will be discussed in chapter 4.

ERK2 exists in populations of monomers and dimers within the cell. The mechanism of dimerization was studied using crystallographic and other biochemical experiments, which suggest that the dimer interface includes the C-terminal extension containing a lengthy, leucine-rich loop that spans the molecule named L16 and the activation loop. L16 contains three hydrophobic residues (L333, L336, L344) and two charged residues (H176, E343) shown by mutagenesis experiments to be required for dimerization (Wilsbacher et al., 2006).

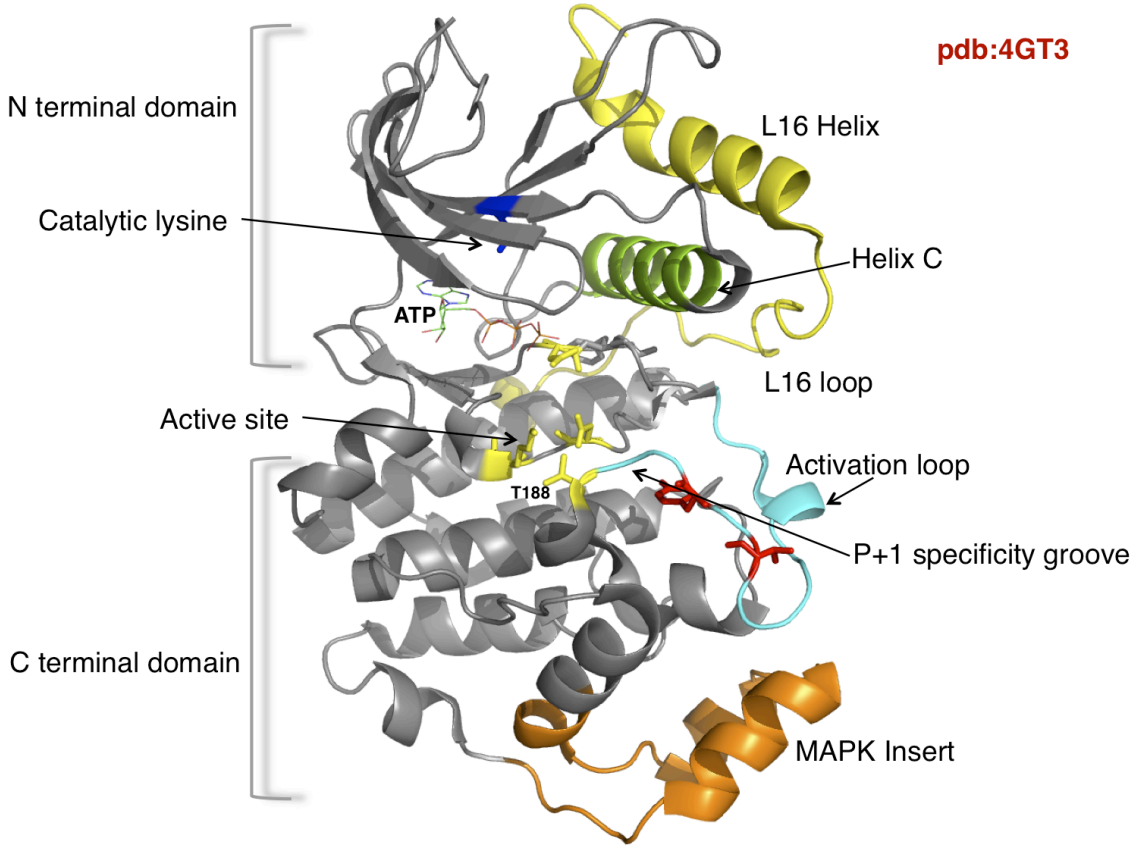


Figure 1.2. Low activity structure of rat ERK2 complexed with ATP (pdb:4GT3). Un-phosphorylated ERK2 bound to ATP (pdb:4GT3) is annotated to show critical regions of the molecule. The Map kinase insert is colored gold and is important for binding to MEK1/2 (through an FXF motifs) and for making contacts to DNA. T183 and Y185, residues phosphorylated by MEK1/2, are important for full activation and are shown in red. The P+1 groove and its critical residue, Thr188, is shown in yellow depicting its proximity to the active site residues. Helix C, important for inter-domain interactions, which mediate closure of the enzyme and contacts with ATP, is shown in green. The catalytic lysine is colored blue and the full activation loop is colored cyan. The L16 loop, which facilitates dimerization is show in yellow in the distal part of the molecule.

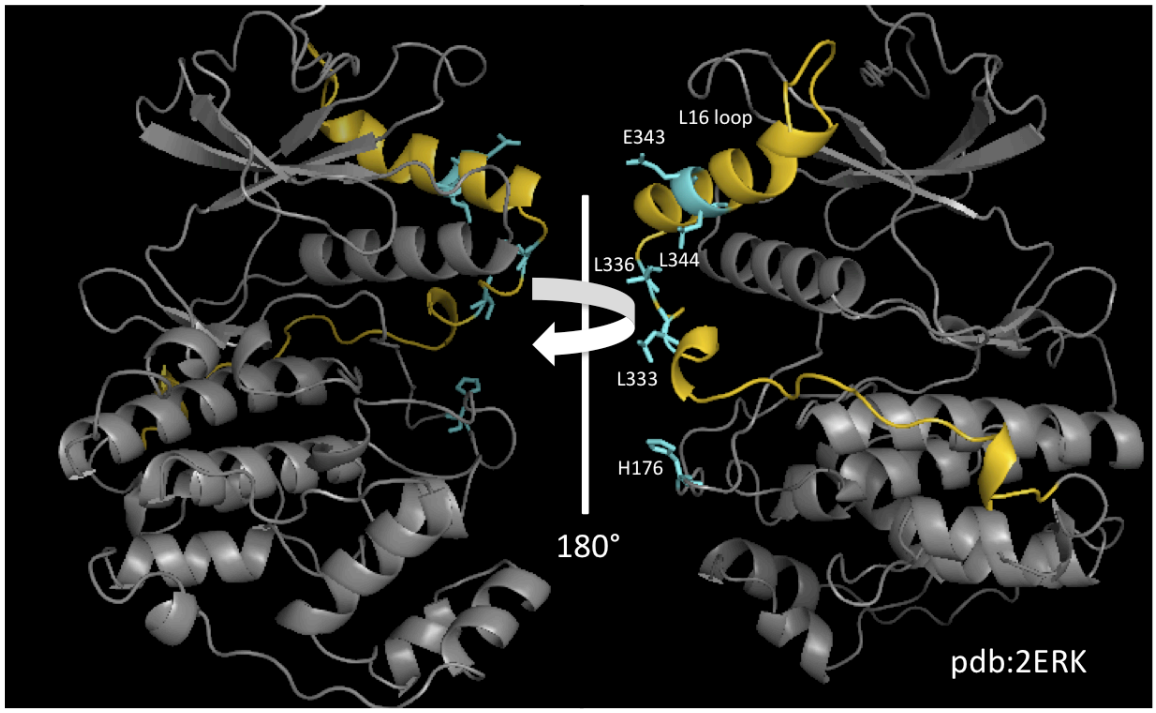


Figure 1.3. Dimerization of ERK2. Phosphorylated ERK2 (pdb:2ERK) dimers are mediated by hydrophobic residues (colored cyan) in the C-terminal extension, which contains the L16 loop. Two charged residues E343 and H176 act as charge clamps which contribute to the dimerization of the molecule. The figure shows a monomer rotated 180° to show the backside of the molecule for illustration. L16 loop in total is colored in gold.

1.5 Importance of the P+1 loop

The activation loop lies adjacent to the P+1 specificity groove, or the P+1 specificity pocket. The P+1 groove is a surface depression where substrates are recognized and positioned for phosphorylation. The P+1 loop contains an invariant Thr residue that is necessary for activity. Mutagenesis of the Thr at a comparable position in many protein kinases results in loss of catalytic activity (Goldsmith et al., 2007). The EPH family of RTKs possesses either a Met or Thr as the critical residue within the loop. Mutagenesis experiments with the RTK EphA5, demonstrates that the mutation T845M results in the loss of its phosphorylation of STAT3 (Yuan et al., 2004). Additionally, MARK2, which phosphorylates the microtubule-associated tau, is inactive when its critical P+1 loop residue, S212 is phosphorylated (Timm et al., 2003). There are exceptions. The mammalian mitotic checkpoint protein, Mps1, was found to be auto-phosphorylated on T686 for which phosphorylation was required for its phosphorylation of Smad2. When phosphomimics were introduced at this position (T686D, T686E), activity was not restored, suggesting that phosphorylation of the Thr residue only will suffice for activation of this enzyme subsequent to its auto-phosphorylation (Wang et al., 2009). Although Thr is mainly found to be the favored residue within the P+1 loop, few other kinases have a different residue at this position. In any event, perturbation of key residues within the activation loop drastically alters activity. Within ERK2 the critical P+1 residue is Thr188. As part of the P+1 specificity groove it is positioned to make critical hydrogen bonds to several residues within the active site including the catalytic Asp 147 (Fig. 1.4). Therefore, structural and cellular data mainly support the hypothesis that Thr188 is a required residue within the active site of ERK2, and its ability to make critical hydrogen bonds is required for function of ERK2 as well as many other protein kinases.

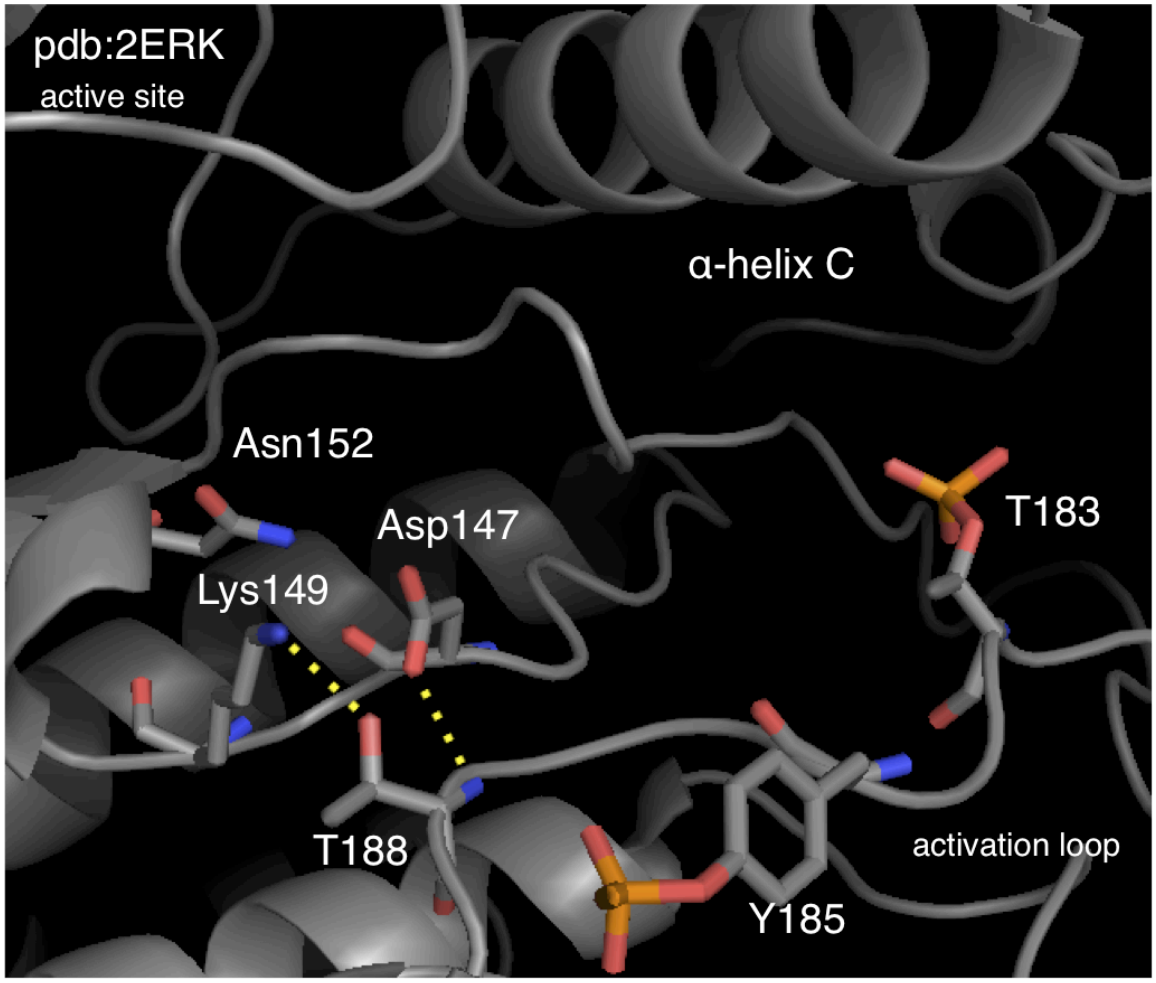


Figure 1.4. T188 is a critical residue. T188, as part of the activation loop and by extension part of the substrate recognition motif (P+1 groove), makes critical contacts to residues within the active site and thus is part of a rich hydrogen-bonding network. Upon phosphorylation, T183 makes contacts to regions of helix C which functions to bring the amino and carboxy domains closer together resulting in proper arrangement of active site residues and full activation of the molecule.

1.6 ERK2 in vitro kinetics

ERK1/2 has over 500 substrates and the kinetics of phosphorylation may differ for each substrate. Quantitative measurements of its activity have been mostly characterized with the substrate, myelin basic protein (MBP). The activity of ERK1/2 toward substrate following activation by MEK has been measured in vitro. In the case of ERK2, its phosphorylation state coupled with correct docking with substrates directly regulates its activity. Un-phosphorylated ERK2 has nearly undetectable activity, turning over less than 1/1000 of the total number of molecules every second. Upon dual phosphorylation within its activation loop by MEK1/2, the rate of catalysis of ERK2 increases nearly 50,000X for a Kcat of 10 s^{-1}. The overall efficiency of the enzyme toward MBP is $2.4x10^{6}$ $M^{-1}s^{-1}$. Active ERK2 can effectively phosphorylate MBP with a specific activity of 2000 nmol/min/mg of enzyme (Table 1). Measurements of activity have been made when specific residues of ERK2 were mutated. For example, single mutations in the activation loop on critical residues T183 and Y185 greatly diminish activity even in the presence of MEK1/2 (Robbins et al., 1993).

Table 1. Specific activities of phosphorylated forms of ERK2 activated by MEK1 in-vitro (Robbins, et al 1993.)

ERK2 variant (after activation with MEK1R4F)	Specific Activity (S.A) nmol/min/mg
ppERK (co-expressed or activated with MEK1R4F)	2200 nmol/min/mg
ppERK K52R	100 nmol/min/mg
ppERK T183E	200 nmol/min/mg
ppERK Y185F	Negligible
ppERK Y191F	1400 nmol/min/mg
ppERK S39D	2100 nmol/min/mg
ERKwt (w/o MEK activation)	1.7 nmol/min/mg

1.7 ERK2 ATP binding pocket and the importance of the 'gatekeeper' residue

Many protein kinases are phosphorylated and therefore activated by upstream kinases in response to a specific stimulus. However, many protein kinases are also subject to auto-phosphorylation even in the absence of an activated signaling cascade. It is therefore appropriate to ask what inherent mechanisms, if any, are present to prevent rogue auto-phosphorylation or mechanisms that allow an activated auto-phosphorylated protein kinase.

The answer to this question has been approached from two different directions. The first approach involved structural characterization of the protein kinase ATP binding site, which showed that individual mutations within could perturb the size of the ATP pocket and encourage auto-phosphorylation. The second approach involved systematic mutational analysis of residues within protein kinases that could synergistically lead to auto-phosphorylation.

The structure of the MAP kinase, p38 (Wilson et al., 1996) allowed the characterization of pyridinyl imidazole inhibitors specific to its ATP binding pocket. Addition of the imidazole inhibitors to human monocytes compromised the production of pro-inflammatory tumor necrosis factor and interleukin-1 (Lee et al). SB203580, a pyridinyl imidazole, was later shown to be a competitive inhibitor of the p38 ATP binding site. Several homologs of p38 were tested with SB203580 and it was discovered that some homologs were sensitive to SB203580, while others were not (Gum et al., 1998). These data suggested that differences in the ATP binding site might account for the selectivity to the pyridinyl compounds. The ATP pockets from each homolog were compared and the result was that there were three residues that differed in each homologue. To probe which of the residues were responsible for the inhibitor selectivity, several residues from the ATP pocket of each p38 homolog were compared. The mutants were expressed in yeast and tested in crosslinking experiments for binding to radiolabeled SB203580. Three residues

within the p38 ATP site (T106, H107, L108) were found to perturb binding to inhibitor upon mutation. Most significant was T106M that rendered p38 insensitive to the compound; remarkably, this mutation did not affect its kinase activity. Interestingly, when ATP-pocket residues of the homologs of p38 not sensitive to compound were mutagenized to those in SB203580-sensitve wild type p38, the homologs lost their insensitivity and were bound to inhibitor. This same phenomenon was seen for the stress activated protein kinase, c-Jun N terminal kinase, JNK1 (Eyers et al., 1998). These data suggest that there are several residues within the active site which mediate specificity toward inhibitor by significantly affecting the size of the ATP binding pocket, and that the residues in question, depending on their R group, might act to occlude the pocket effectively shielding it from inhibitor binding. Residues that confer this specificity are characteristically large residues in the active site known as 'gatekeepers'. Mutagenesis experiments have shown that when bulky gatekeepers are mutated to smaller glycine or alanine residues, the resultant active site becomes larger and able to accept a battery of chemically designed ATP analogs capable of fitting into the active site and thereby attenuating activity or being used for phosphoryl transfer in place of ATP. The gatekeeper residue has become the subject of a number of chemical genetics approaches that design specific protein kinase inhibitors for therapeutic and investigational purposes (Shokat and Velleca, 2002; Zhang et al., 2005). The general theme of a 'gatekeeper' residue was extrapolated to the ATP pocket of ERK2. Wild-type ERK2 is not susceptible to inhibition by p38 specific inhibitors. However, when the equivalent mutations that render p38 sensitive to pyridinyl inhibitors where mutated in ERK2, binding to SB203580 and SB202190 (under conditions of multiple ATP pocket mutations) were seen in the low nanomolar range (Fox et al., 1998). The ERK2 gatekeeper, Q103 (Q105 human numbering) was found to bind inhibitor concentrations at significant levels over wild type.

The importance of gatekeeper residues on protein kinase active sites and the resulting permissive binding to polycyclic inhibitor molecules upon their mutagenesis was a major advance in inhibitor design. However, the gatekeeper residues had not until this

point been enzymatically characterized. Thus, gatekeeper mutations ERK2 were characterized and found to increase overall basal enzymatic activity. Experiments conducted on the ERK2 gatekeeper Q103, showed that when Q103 was mutated to small, non-hydrophobic residues (Q103A and Q103G) auto-phosphorylation of the activation loop occurred intra-molecularly and basal activity of ERK2 increased relative to wt (with Q103A possessing more basal activity that Q103G) (Emrick et al., 2006).

Prior to the investigation of the effects of Q103 on auto-phosphorylation, additional work on ERK2 showed that other residues within ERK2, when mutagenized, allowed spontaneous intra-molecular phosphorylation and increases in basal activity (Emrick et al., 2001). Ten residues were found within the overall ERK2 structure that perturbed phosphorylation and activity. However, three key residues acted synergistically to increase activation and activity to significant levels relative to wt. Residues L73 and S151 map to the active site of ERK2 and D319 is a key residue in binding to substrates. Double mutations of L73P/D319N resulted in activation of the molecule with a 50-fold increase over wt. The triple mutation, L73P/ S151D/D319N increased the activity of ERK2 90 fold over wt ERK2 in vitro. Compiling the data from experiments investigating the gatekeeper mutation and the additional mutations in ERK2 not located in the ATP binding pocket revealed an interesting connection. The gatekeeper residue and the three residues responsible for the high levels of intra-molecular phosphorylation and increases in activation were connected in a close, hydrophobic spine (Fig.1.6) traversing the center of the molecule and ending at the activation loop, such that changes in the gatekeeper residue were relayed to the activation loop ultimately causing residues Y185 and, T183 to a lesser degree, to become subject to auto-phosphorylation. Additional auto-phosphorylation mutants have been investigated in S. *cerevisiae* (Levin-Salomon et al., 2009) and in D. *melanogaster* (Levin-Salomon et al., 2008) and shown to have increased activity and auto-phosphorylation presumably through the same connected network as the well characterized Q103 gatekeeper residue in ERK2.

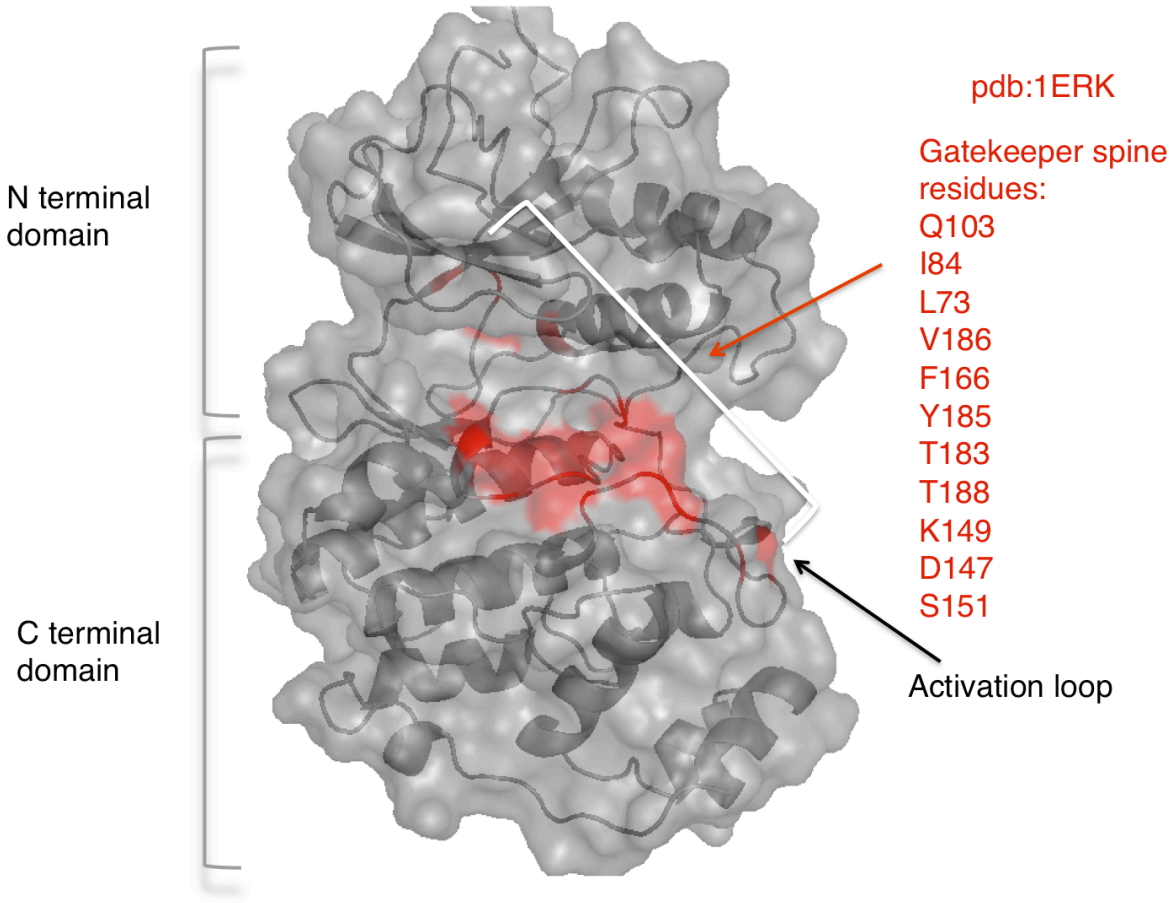


Figure 1.5. An interconnected network of residues starts at the gatekeeper Q103 within the amino terminal domain, and ends at the activation loop. The gatekeeper residue Q103 is linked to the activation loop by a series of residues that traverse the molecule from the amino to the carboxy terminus. When residues Q103,L73, D319 or S151 are mutated, autophosphorylation results.

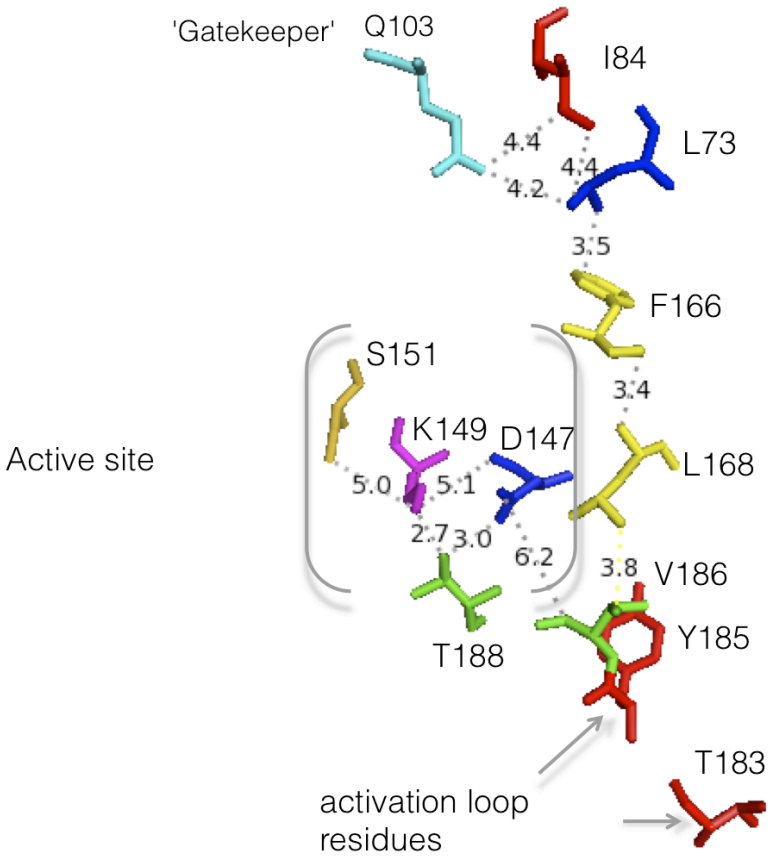


Figure 1.6. Low activity ERK2 hydrophobic spine in detail. [pdb:1ERK]. Figure amended from Emrick et al showing the connectivity of the hydrophobic spine, which traverses the center of ERK. The connectivity starts from the 'gatekeeper' residue in the amino terminal domain through the activation loop and into the active site. ***Emrick et al, PNAS (2006) vol. 103 (48) ,18101–18106.***

1.8 **ERK and the nucleus: Implications for transcription**

ERK1/2 have been estimated to have cellular concentrations greater than 100nM. In addition to high concentrations, ERK1/2 have hundreds of cellular substrates. ERK1/2 are found in both the nucleus and the cytoplasm. Regulation of ERK1/2 cellular location can be demonstrated with experiments, which sought to determine the specific criteria that determined the cellular trafficking of ERK1/2. Immunofluorescence and labeling experiments of ERK2 in which cells were starved or exposed to serum, rich with mitogens and other associated growth factors, showed that the identity of ligand that activated the MAPK pathway was important in determining the localization of ERK1/2 (Chen et al., 1992). Additional work showed that ERK1/2 is not inherently nuclear and is, in fact, imported into the nucleus reflecting the amount of stimulatory ligand. In experiments where an increasing dose of EGF was given to cells, signal from GFP labeled, ERK2 showed a linear increase correlating to EGF concentration (Whitehurst et al., 2004a). These experiments demonstrate that nuclear localization is tunable according the amount of stimulatory ligand.

Several mechanisms control the nuclear entry and exit of a variety of proteins, nucleic acids and metabolites. Small proteins, peptides and nucleic acid molecules can migrate into and out of the nuclear pore through passive diffusion and this was observed for ERK1/2. Conversely, most large proteins need active facilitation of transport across the nuclear pore. A variety of karyopherins (importins and exportins) in addition to cellular energy (Ran GTP) facilitate this translocation by interacting directly with nuclear pore proteins, with entry dependent on the Ran-GTP gradient. ERK2 has been shown to be actively translocated across the nuclear membrane. In experiments where ERK2 was fused to β-galactosidase effectively increasing its molecular weight and therefore excluding the possibility of passive diffusion, ERK2 was translocated into the nucleus suggesting that ERK1/2 uses a mechanism of active diffusion (Adachi et al., 1999). Upon phosphorylation, ERK1/2 can utilize certain importins for nuclear entry (Seger) interacting directly with nuclear pore proteins (Plotnikov et al., 2011). ERK2 can bind the nucleoporins (Nup153,

Nup214) through FG repeats (Whitehurst et al., 2002). Interestingly, a number of nuclear pore proteins including Nup153 are also phosphorylated by ERK1/2. Additional experiments showed in nuclear import assays, GFP tagged ERK2 was determined to localize to the nucleus in the absence or presence of cellular energy. However, blocking the nuclear pore complex with wheat germ agglutinin prevented ERK1/2 passage through it (Whitehurst, et al, 2002). Phosphorylation of ERK1/2 is also important in nuclear translocation. Phosphorylation of ERK1/2 can drive dimer formation and experiments with dimerized ERK2 showed the kinetics of nuclear retention were more than six times as long and non-dimerized ERK2 (Khokhlatchev et al., 1998). Once inside the nucleus ERK1/2 can be readily exported out of the nucleus, by binding to MEK1/2, which harbors a nuclear export signal. ERK1/2 can then be sequestered outside of the nucleus by binding cytoplasmic proteins such as PEA-15 (Whitehurst et al., 2004b). These data suggest that there may be multiple points of regulation for ERK1/2 nuclear entry and exit.

Nuclear localized ERK1/2 can have a variety of functions. Phosphorylation of nuclear transcription factors is a major activity of ERK1/2. ERK1/2 was found to phosphorylate Elk-1, a transcription factor that activates immediate early genes in response to serum in many cells (Gille et al., 1995). Numerous additional transcription factors are substrates for ERK1/2 such as PPARγ and c-Jun.

An additional role exists for ERK1/2 in nuclear compartments: binding to chromatin. ERK1/2 has been found to bind to promoter regions of genes and intergenic sites mainly through ChIP-seq experiments. Computational and proteomics screens have also identified ERK1/2 as important for interacting with DNA. In a screen to determine novel DNA binding proteins, ERK2 was uncovered as a direct DNA binding protein. It was suggested that direct binding of ERK2 to the interferon -γ promoter through its MAPK insert region caused suppression of target genes. Interestingly, binding of ERK2 was not dependent on its activation as the kinase dead form (K54R, human numbering) bound directly to DNA (Hu et al., 2009). Work in this thesis specifically explores conditions

necessary for ERK2 / DNA binding.

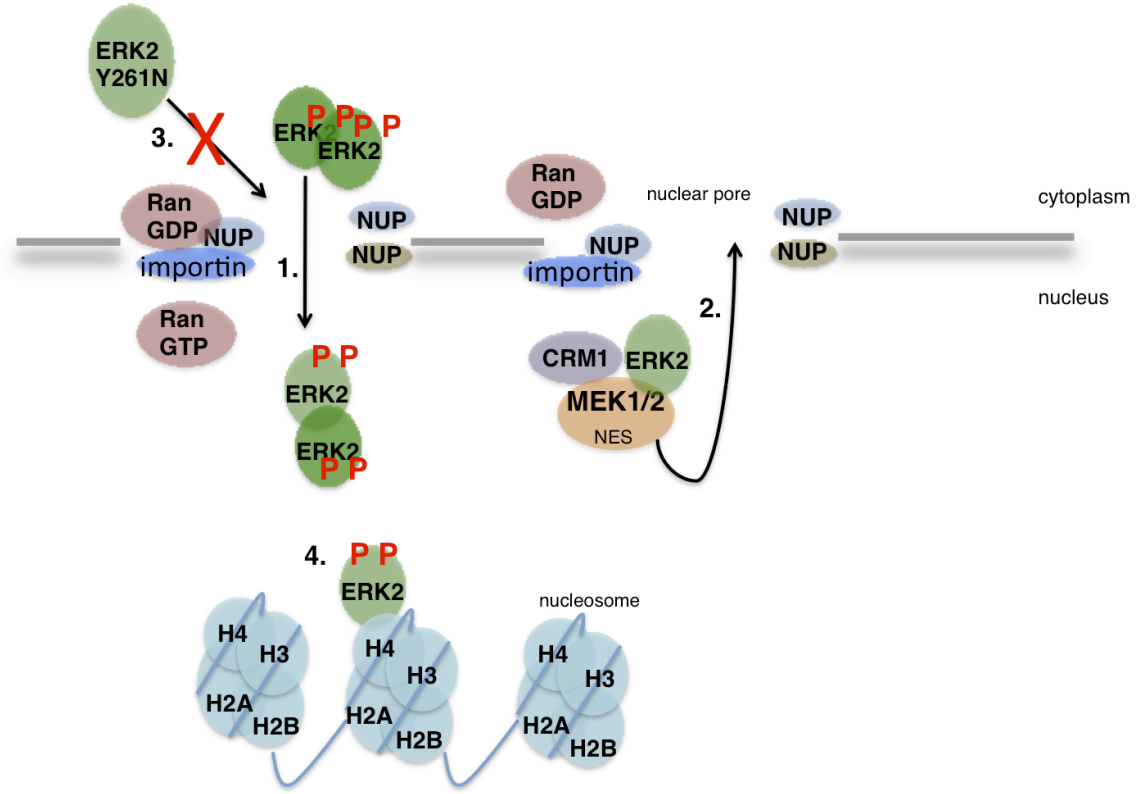


Figure 1.7. Relationship between ERK1/2, the nucleus and chromatin. 1) Diffusion of ERK1/2 can occur by a carrier-independent mechanism but diffusion is often related to its phosphorylation and dimerization. 2) MEK1/2 export of ERK1/2 occurs through its NES. 3) Mutations in the ERK2 Map kinase insert region (aa 241-272) prevent nuclear localization. 4.) Phosphorylated ERK binds to chromatin and is able to bind DNA directly.

Chapter 2.

Phosphorylation at T188 occurs in vitro within a heterologous bacterial expression system without G protein-mediated signaling inputs

Introduction

Recent identification of a novel autophosphorylation site within the activation loop of the MAPK ERK2

Until 2008, few questions were thought to remain about the mechanism of phosphorylation and activation of ERK1/2 and the identity of the residues important for this process. The standing paradigm for ERK activation is phosphorylation of Y185 within the activation loop by MEK1/2 followed by the phosphorylation of T183 to yield fully activated ERK1/2.

ERK2 autophosphorylation has recently been proposed to occur on a site Thr188, which resides in the P+1 surface depression groove at the end of the activation loop (Lorenz et al., 2009). The proposed mechanism involves the activation of the ERK1/2 MAP kinase pathway by ligand bound RTKs in combination with activated GPCRs (M1 muscarinic and α_1 adrenergic). The GPCR associated G proteins liberate $\beta\gamma$ subunits that bind to components of the activated MAPK pathway complex (Raf,MEK,ERK). $\beta\gamma$ binding to Raf/MEK/ERK causes activated ERK1/2 to autophosphorylate on Thr188. These events were reported to occur in hypertrophic cardiac myocytes, stressed human hearts, and in mice in which Gq-coupled receptors were activated followed by cardiac challenge. The proposed effect of Thr188 autophosphorylation is the trafficking of ERK1/2 into the nucleus to stimulate the transcription of genes involved in cardiac hypertrophy.

The authors initially found the novel autophosphorylation site by observing that biochemical pull-down of over-expressed constitutively active MEK1 or ERK2 immunoprecipitated $\beta\gamma$ protein. The authors performed immunoblotting analysis with a

phosphorylated ERK1/2 (ppERK) specific antibody and noticed that the pattern of phosphorylation differed between that of purified ERK2 incubated with MEK only and ERK2 incubated with βγ only. ERK2 incubated with βγ had diminished reactivity with ppERK specific antibody as compared to ERK2 incubated with a constitutively active MEK1 (R4F MEK1) that phosphorylates ERK on the canonical pT183 and pY185 residues.

Next, when purified ERK2 was incubated with βγ, a novel phosphopeptide was seen that was not identical to any peptide with ERK2 incubated with MEK or alone in the presence of [γ-^{32}P]ATP; however, all peptides were seen with purified ERK incubated with MEK and βγ. Subsequent mutational analysis showed that of the threonine residues available for phosphorylation, loss of the novel phosphopeptide species was only detected when T188 was mutated to alanine.

The phenotype for this auto-phosphorylation event is purported to be up-regulation of hypertrophic specific genes in cardiac myocytes. The authors suggest this occurs by the translocation of triply phosphorylated ERK into the nucleus to possibly phosphorylate factors necessary for transcriptional activation of hypertrophy. Experiments from this paper demonstrate that pT188-containing ERK2 represented by the T188D phosphomimic is readily translocated into the nucleus.

T188 is critical for activity within ERK2. Therefore this paper is of interest because it raises the question of how T188, which makes critical hydrogen bonding contacts in the active site, can sustain autophosphorylation and still allow the molecule to retain its activity. It is also unclear how specific interaction with βγ facilitates the autophosphorylation event. Furthermore, although the novel T188 autophosphorylation (pT188) was found in a cellular context, the authors claim pT188 autophosphorylation can be detected in vitro. Interestingly, previous work describing gatekeeper mutations of ERK2 and their in vitro activity did not report autophosphorylation of T188.

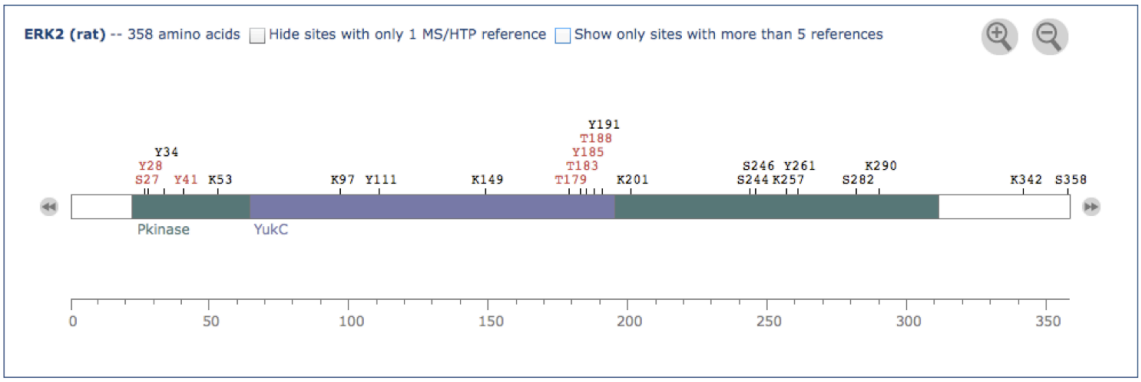


Figure 2.1. Reported phosphorylation sites in rat ERK2. A schematic of phosphorylation sites within ERK2 as annotated in Phosphosite. T188 is reported. S294 and Y326, novel phosphorylation sites found in our analysis, are not. Hornbeck PV. et al. (2012) Nucleic Acids Res. 2012 40:D261-70.

Hypothesis and approach

The MAPK signaling pathway is modular in that several protein kinases need to come together for full activation of the pathway. This same modularity can be recreated in a heterologous bacterial expression system in vitro (Heise and Cobb, 2006). Individual kinases can be expressed alone or in the presence of additional upstream activating kinases. In the case of the MAPK ERK2, the rat isoform, which possesses a 6XHis tag for purification, can be expressed alone to generate the enzyme, which is not fully phosphorylated nor displays full enzymatic activity. The second approach is to express ERK2 in the presence of a constitutively active form of the upstream MAP/ERK kinase MEK1/2, R4FMEK1 (Mansour et al., 1996). Co-expression of ERK2 and R4FMEK1 results in the fully phosphorylated, activated form of ERK2 (Khokhlatchev, et al, Robbins, et.al).

Initial purification of 6X His tagged ERK2 (without co-expression of constitutive active MEK1) is by immobilized metal affinity column (nickel) chromatography and results in a partially purified form of the enzyme. ERK2 can

be more highly purified on anion exchange chromatography (IEX). It has been previously demonstrated that IEX resolves ERK2 into two distinctly eluting peaks when a gradient of increasing salt (NaCL) concentration is used to elute the protein (Prowse and Lew, 2001). Initial enzymatic activity assays on ERK2 in fractions from IEX have demonstrated that ERK2 with distinct activity states can be separated using this method. Purification data from my experiments corroborate this finding, and shows a partially active form, which results from phosphorylation of only Y185, can occur in both peaks.

IEX purification can separate ERK2 phosphorylation states and modifications that occur during expression in bacteria. Early characterization of the phosphorylated species in fractions of IEX-purified ERK2 showed that catalytic activity was concentrated in the last fractions of the first eluting peak from a Mono Q IEX column and I detected this same phenomenon in my bacterial preparations (Fig. 2.2). This would suggest that within these fractions of IEX-purified ERK2, there exist a number of differing charge states, phosphorylation states and possibly differing conformations.

In this portion of my work, I report the finding of phosphorylated T188 as a result of heterologous, bacterial protein expression. pT188 was identified in recombinant ERK2 preparations without the input of G protein signaling molecules or other upstream kinases in the MAPK ERK1/2 cascade. The remainder of this section will focus on the purification method and changes I instituted which may have led to the generation of pT188. Furthermore, in collaboration with Dr. Yan Li (NIH), we demonstrate that pT188 ERK2 is detectable by mass spectrometry. Therefore, I conclude that pT188 is an autophosphorylation event that occurs without the input of G proteins opposite to earlier published results, and that pT188 is easily identifiable in bacterial preparations.

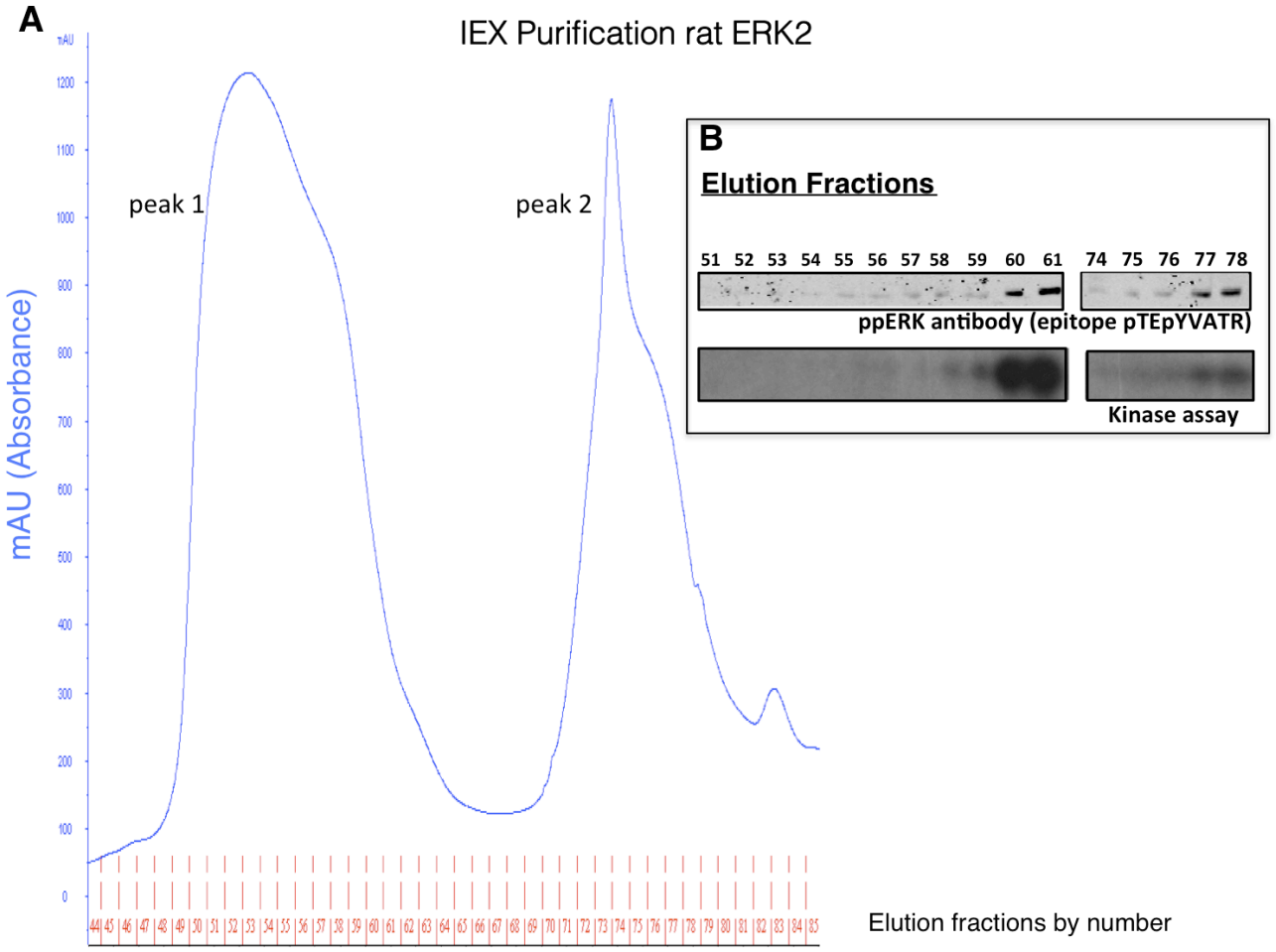


Figure 2.2 IEX chromatography of rat ERK2 wt protein results in several phosphorylation and activity states. A.) Nickel purified 6XHis ERK2 wt protein was purified using a NaCL gradient and a Mono Q Anion exchange resin. B.) Individual fractions were taken from portions of the eluting peaks from the purification in panel A and were normalized and analyzed by in vitro kinase assays and immunoblotting using antibodies specific for doubly phosphorylated ERK2.

Results

I serendipitously detected phosphorylation of T188, a critical residue necessary for the hydrogen-bonding network within the ERK2 active site, contained within a preparation of recombinant rat ERK2 wt protein. Prior to this detection, I was performing electrophoretic mobility shift assays (EMSAs) to investigate direct binding of ERK2 to synthetic oligonucleotides (chapter 3). As such, I purified ERK2 for my studies using a Mono Q anion exchange column (IEX), which routinely yielded two separated peaks (herein referred to as peak #1 or peak#2). I knew that activity and charge states were distinct between eluting peaks. However, I did not know which peak (or eluting fractions thereof) was optimal to use in EMSA binding experiments. I decided to pool fractions within each of the two peaks together and test both peak #1 and peak #2 in an EMSA experiment to determine which peak of protein would optimally bind to oligonucleotide. I tried several amounts of protein from each peak and found that protein from the second-eluting peak bound better to a synthetic oligonucleotide, which represented a region of the mouse insulin gene promoter (Fig. 4.4 and see chapter 3). These data suggested to me that inherent differences between the peaks would influence binding to oligonucleotides. The phosphorylation state(s) of ERK2 contained in the eluted fractions were also unknown. At that time, I did not know if the differential binding was due to the conformation differences in the protein, or a tightly associated co-factor that may co-elute with ERK2 or simply the charge state of the protein.

As a result, we analyzed the protein peaks by mass spectrometry. Dr. Yan Li, (formerly with the UTSW Mass Spectrometry core, now at the NIH) performed mass spectrometry on each of the two eluting peaks using digestion with trypsin and AspN proteases. As such, she was successful in detecting several

phosphorylated species: pY185 alone or with pT183, and pY185pT188. pY185 was by far the most common phosphorylation state. The doubly activating pT183pY185 occurred, but was rare. pThr188 was detected, but only seemed to occur with a signal from pY185 in the same fraction and its abundance appears low. In almost every instance in which ERK2 phosphorylation was detected, pY185 was one of the phosphorylated residues. Therefore, my goal was to purify a homogenous fraction of pT188 using MonoQ IEX to characterize its properties.

Purification of pThr188

In order to purify pThr188 to homogeneity, I considered several possibilities for biochemical purification. The first was an ATP affinity column. Knowing that the position of T188 makes it intimately associated with the active site which is adjacent to the ATP binding pocket, and knowing that T188 is required for activity, I assumed that phosphorylation of T188 might cause egregious reorganization of the active site and by extension the ATP binding pocket and thus abolish its ability to bind ATP or an ATP conjugated affinity column. If pT188 ERK2 could not bind to an ATP affinity column, but non-ATP ERK2 could, I could use this method to separate phosphospecies. The second approach was to alter the elution protocol for ERK2 when using the anion exchange column. In this set of experiments, I hypothesized that using a variety of chaotropic salts would affect the elution of protein in a manner that was different than elution with a gradient of NaCL and therefore result in additional, better resolved peaks further separated by charge and/or phosphorylation state. I decided to pursue IEX with differential elution b salts. I added the second peak of ERK2 from a subsequent protein prep to the same anion exchange column and eluted with a shallow gradient of $NaHCO_3$. Unfortunately, I did not resolve any additional peaks (Fig. 2.3). I also tried other salts such as KCL

also without further resolution.

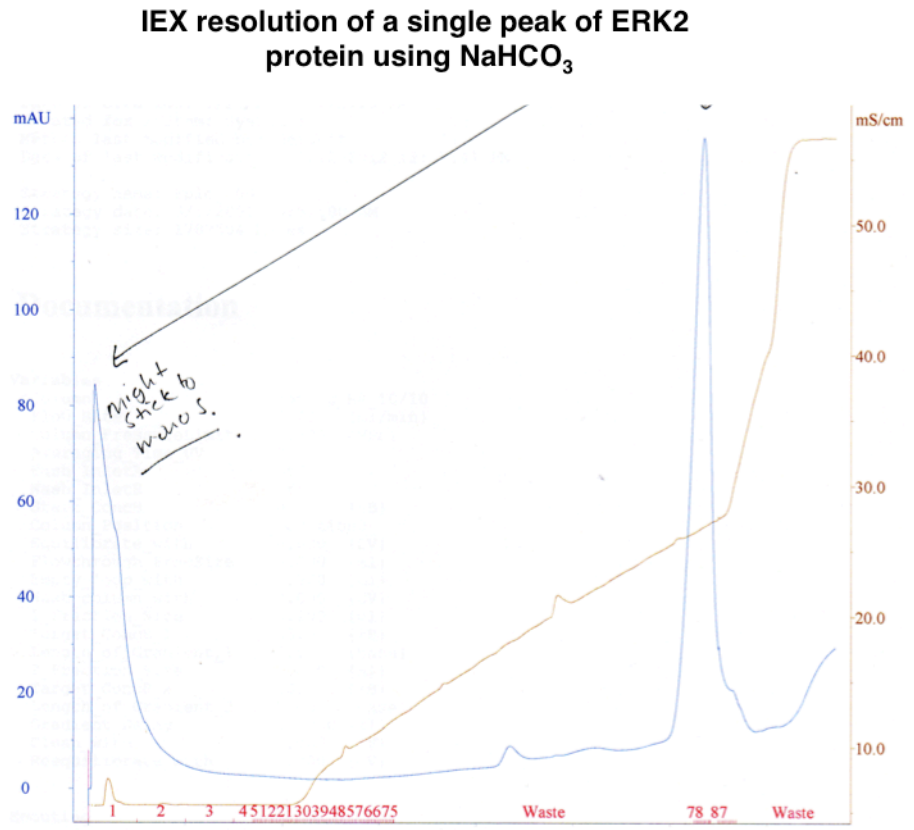


Figure 2.3 Multiple species of phosphorylated ERK2 are not resolved with additional IEX salt gradients. In an attempt to resolve multiple phosphoforms of ERK2 contained within previously eluted IEX peaks (previously eluted with an increasing NaCL gradient) the second eluting peak of an ERK2 wt purification was again added to a mono Q IEX column with the exception that several different salts were used in an attempt to improve resolution. These data show an analytical run in which $NaHCO_3$ was used. This method was not successful, yielding only a single, unresolved peak.

I generated many additional bacterial preparations that were similarly purified. Since I did not succeed in resolving additional peaks when I changed conditions for elution within IEX, I decided to analyze each individual fraction of my purifications instead of pooled fractions represented by each peak. Interestingly, mass spectra determined that certain fractions harbored pT188. In fact, one bacterial preparation yielded ERK2 phosphorylated on T188 in almost all fractions tested. This would appear that resolution of pT188 was not successful, however in a few fractions only associated with peak2, there was an obvious absence of pT188 with pY185 as the only phosphorylated species. pT188 seemed to consistently occur in the presence of pY185, never presenting as a singly phosphorylated species. Additionally, pT188 occurred in low abundance. These data suggested that pT188 was present in this preparation, it was contained in most fractions and it likely occurred in the presence of pY185. The data further showed that the IEX method I used initially could indeed separate differentially phosphorylated fractions. Interestingly, mass spectrometry analysis detected novel phosphorylation sites within my preparation of ERK2 not previously described (see Appendix A).

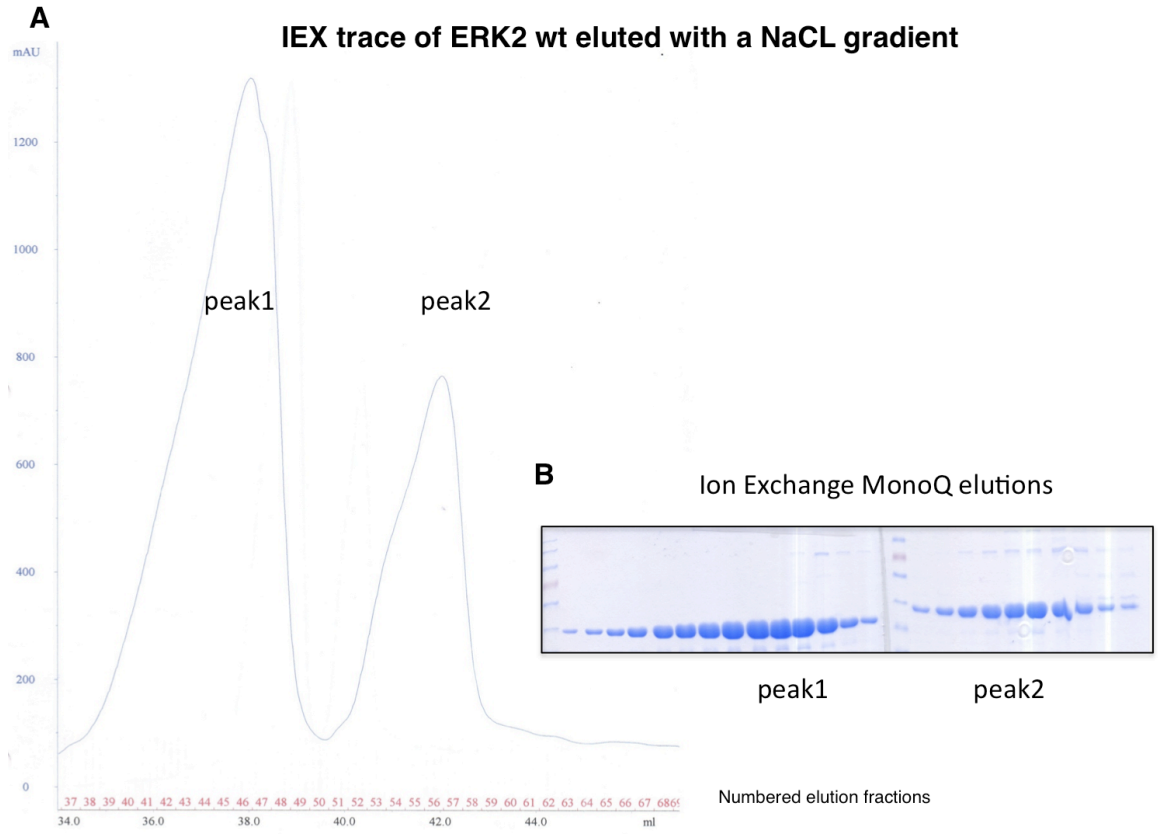
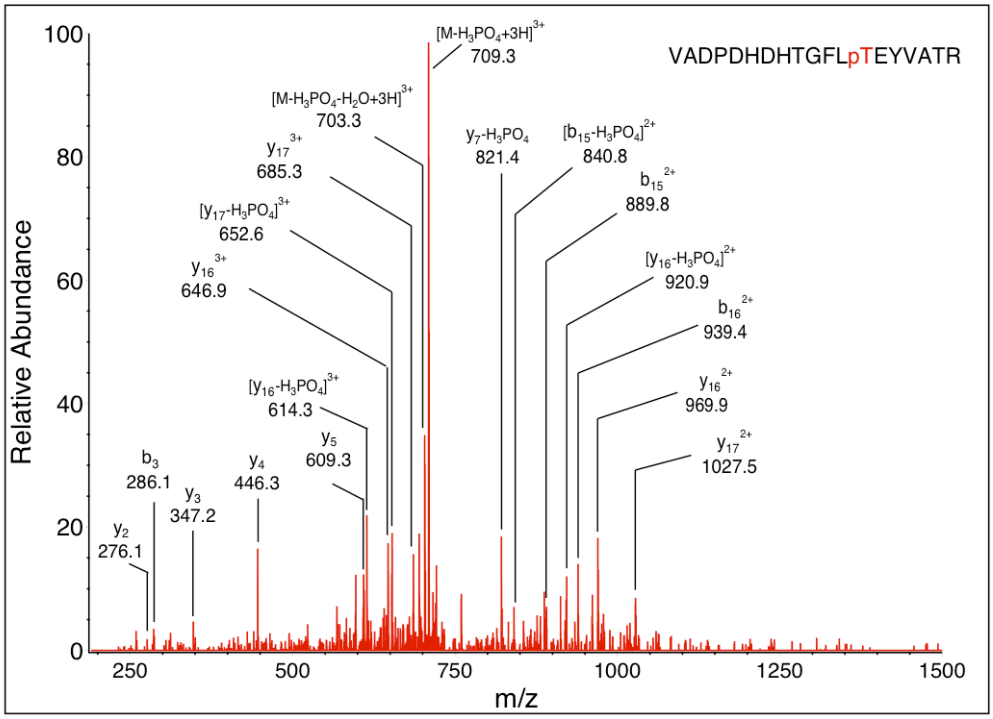
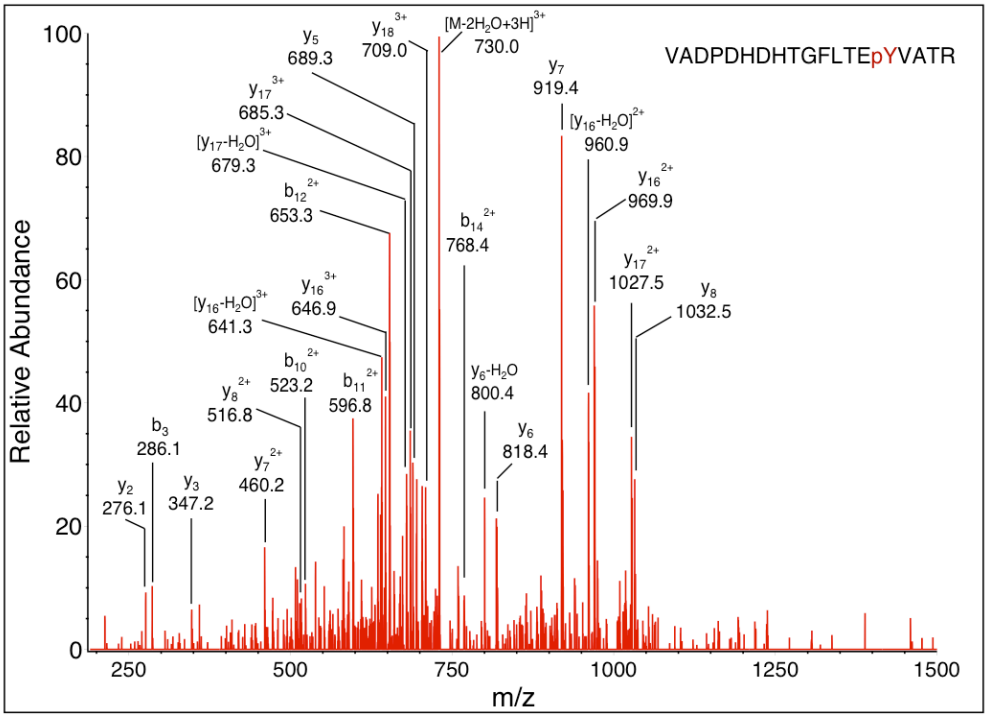


Figure 2.4. IEX trace of ERK2wt purification identifies pT188 in multiple fractions. A) ERK2wt was purified with a monoQ Anion exchange column using a gradient of 10mM – 1M NaCL. B) Individual fractions were analyzed by mass spectrometry after fractions were resolved on a 12% SDS-PAGE gel and doubly digested with Trypsin and AspN (flavastacin) proteases. See Appendix A for mass spectrometry analysis for each fraction.

A


VADPDHDHTGFLpTEYVATR
[M-H3PO4+3H]3+ 709.3
[M-H3PO4-H2O+3H]3+ 703.3
y17 3+ 685.3
[y17-H3PO4]3+ 652.6
y16 3+ 646.9
[y16-H3PO4]3+ 614.3
y5 609.3
y4 446.3
b3 286.1
y3 347.2
y2 276.1
y7-H3PO4 821.4
[b15-H3PO4]2+ 840.8
b15 2+ 889.8
[y16-H3PO4]2+ 920.9
b16 2+ 939.4
y16 2+ 969.9
y17 2+ 1027.5
Relative Abundance
m/z
0
20
40
60
80
100
250
500
750
1000
1250
1500

B


VADPDHDHTGFLTEpYVATR
y5 689.3
y18 3+ 709.0
[M-2H2O+3H]3+ 730.0
y17 3+ 685.3
[y17-H2O]3+ 679.3
b12 2+ 653.3
y16 3+ 646.9
[y16-H2O]3+ 641.3
b10 2+ 523.2
y8 2+ 516.8
b11 2+ 596.8
b3 286.1
y7 2+ 460.2
y2 276.1
y3 347.2
y7 919.4
[y16-H2O]2+ 960.9
y16 2+ 969.9
b14 2+ 768.4
y17 2+ 1027.5
y8 1032.5
y6-H2O 800.4
y6 818.4
Relative Abundance
m/z
0
20
40
60
80
100
250
500
750
1000
1250
1500

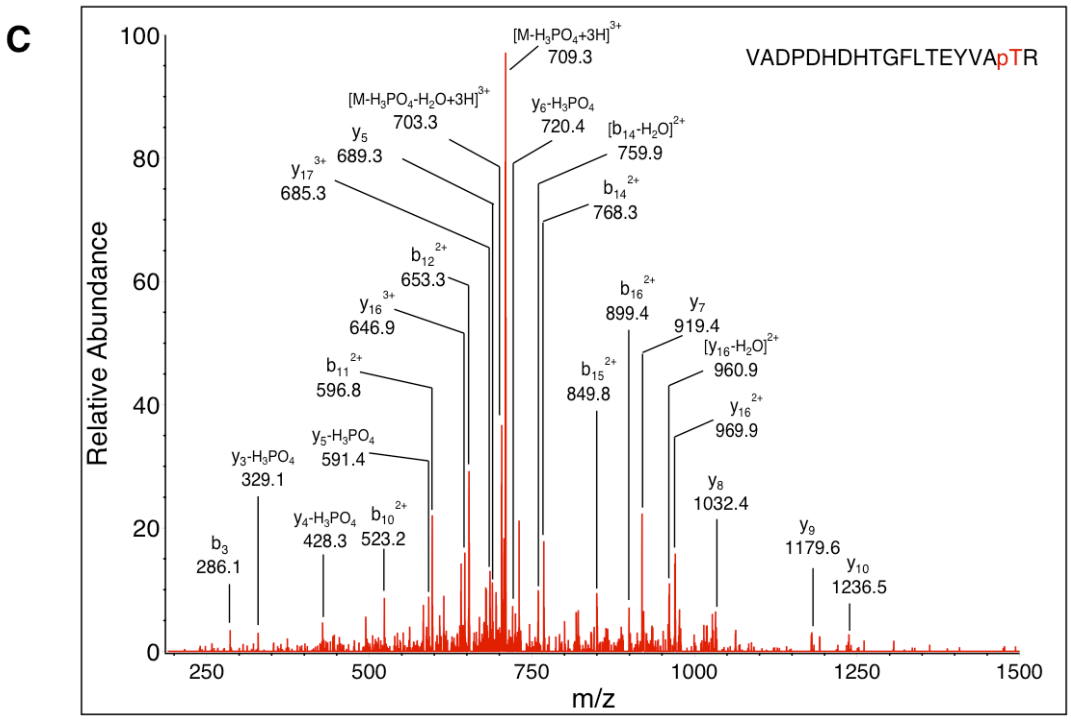


Figure 2.5. Mass spectrometry of individual ERK2 wt fractions.Individual fractions from ERK2 wt purifications (not expressed with constitutively active MEK1) were subjected to proteolytic cleavage (Trypsin followed by AspN) and further analyzed by LC/MS/MS. Three residues in the activation loop were found to be auto-phosphorylated: pT183 (A), pY185 (B) and pT188 (C) (see page 33 for spectra A and B).

The fractionation and purification of the various ERK2 autophosphorylation states allowed me to test each fraction for activity and compare that activity to its phosphorylation state determined by mass spec analysis. I performed in vitro kinase assays on selected fractions from ion exchange purified ERK2. Upon activation of each fraction with a constitutively active MEK (R4FMEK1) in vitro, each representative fraction was able to activate ERK as judged by its phosphorylation of the substrate MBP in in vitro kinase assays. This was a surprise given the fact that several of the fractions I assayed contained pThr188, which, based on ERK2 structural details noted earlier, I expected would be inactive. It must be stated that I could not determine unequivocally what percentage of phosphorylated Thr188 was present within each fraction (as compared to singly phosphorylated Y185 which appeared to be the major species). The main conclusion from these data is that the fractions containing pT188 are activated in vitro. Further experiments are being performed to determine if the percentage of pT188 in each fraction might be related to small changes in activity (presumably reduction of activity). This will help us ascribe the implications of pT188 phosphorylation to overall activity within each sample.

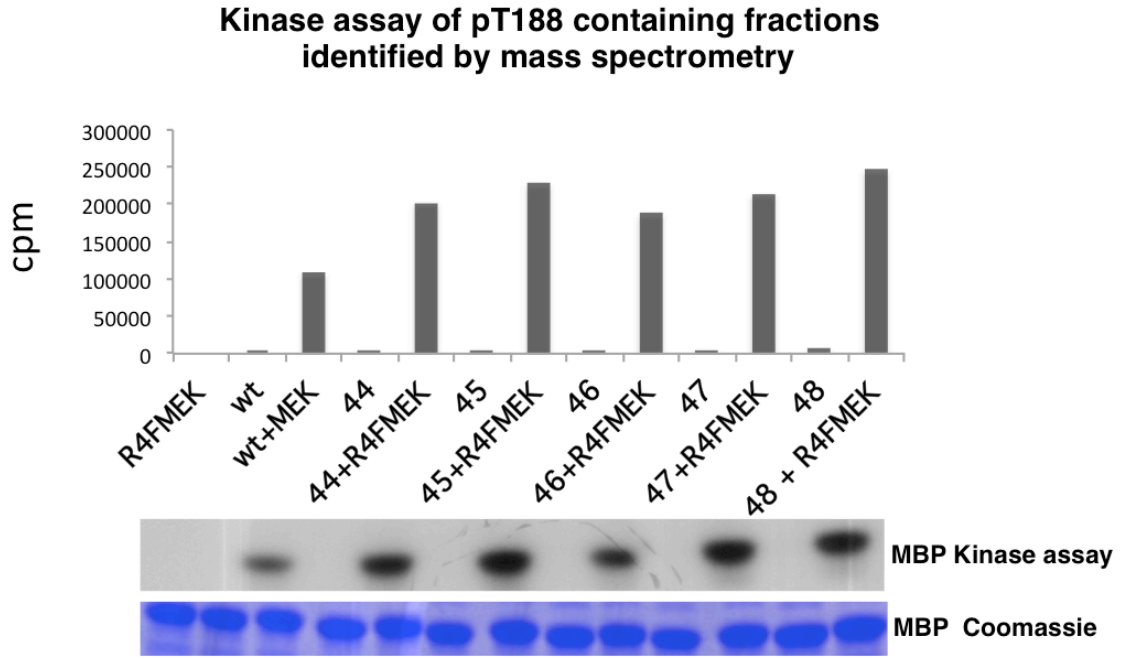


Figure 2.6. pT188 containing fractions have enzymatic activity. pT188 containing fractions identified by mass spectrometry were found to have activity when ERK2 fractions were activated by a constitutively active MEK1 (R4FMEK1) and added to in-vitro kinase assays using MBP as a substrate. CPM refers to counts per minute as measure by scintillation counting. Experiment was performed in triplicate.

Discussion

In this study, I find that Thr188 autophosphorylation occurs without the input of upstream G protein subunits and without the full complement of additional MAPK cascade kinases—contrary to the published model of T188 phosphorylation. Mass spectrometry data confirms the presence of pT188 even though this species occurs as a nominal, fractional amount of total bacterial ERK2 wt protein expression. We have communicated to the authors who originally discovered this phosphorylation state and we have received a pT188 specific antibody such as was used in their publication. I will perform immunoblotting experiments to ask several questions. Original work was performed on protein immune-precipitated from cells or on a phosphomimetic protein. Experiments will be performed that test the antibody for detection of pT188 in bacterial preparations and whether this is a reliable method for the determination of the percentage of pT188 contained in each fraction. These experiments will be useful in developing our own phosphorylation-site-specific antibodies to distinguish between ERK2 phosphorylation states.

If upstream signals from G proteins in conjunction with MAPK components must be present for phosphorylation on T188, how might this occur otherwise? I modified bacterial expressions to produce protein in excess of 15 mg/mL by inducing at a later time point (OD_{600} =0.9). I speculate that the high concentration of ERK2 within bacteria may enhance the likelihood of autophosphorylation by an intermolecular mechanism. It is not clear whether pT188 auto-phosphorylation occurs as a uni- or bimolecular reaction as ERK2 auto-phosphorylation has previously been found to occur intra-molecularly (Robbins et al., 1993). I expect that ERK2 phosphorylated on pY185 can then auto-phosphorylate on pT188, based on the finding that pT188 was not found without pY185. This might suggest that prior to pT188 phosphorylation, the protein is correctly folded. The effect of the cellular concentration of ERK2 and the rate of phosphorylation by MEK has been explored using mathematical models, which attempt to explain how a high local concentration of ERK2 and MEK1/2 affect ERK phosphorylation (Aoki et al., 2013). For example, by using

polyethylene glycol (PEG) in in vitro phosphorylation experiments to recreate high solvent viscosity, the authors found that the processivity of ERK phosphorylation by MEK increases in a dose-responsive manner. Increasing PEG concentrations led to doubly phosphorylated ERK2 instead of mono phosphorylated ERK (pY185) followed by phosphorylation on T183. Of course, the effects of ADP dissociation rates were studied, but the idea is that as PEG has an effect on solution viscosity, and this in turn has an effect on how ERK2 is phosphorylated due to increasing ERK2 local concentration. This model does not deal with auto-phosphorylation per se, but an important biophysical characterization of solution dynamics and their effects on phosphorylation, or of all protein modifications for example, is made and may be an important explanation about how autophosphorylation occurs in our high-level recombinant protein expression system.

Incorrect folding may influence autophosphorylation. High-level expression may induce incorrect folding of a fraction of protein, which may be amenable to auto-phosphorylation. If this was this case, it did not result in precipitation of the protein. I did not see a significant amount of insoluble protein in the form of aggregates during purification and protein that was subsequently used in kinetic analysis was active toward substrate.

Our data yielded a fraction that had novel phosphorylation sites in addition to T188 (see Appendix A). S294 and Y326 were phosphorylated. Additional post-translational modifications with biological meaning are being identified for ERK1/2. Interestingly, S-nitrosylation was found to occur on ERK1 in MCF-7 breast cancer cells exposed to nitric oxide (Feng et al., 2013). Results suggested that phosphorylation of ERK2 was ablated when ERK1 was S-nitrosylated. These studies point to the implications of previously underappreciated modifications of ERK1/2 and their potential biological importance.

The physiological relevance of phosphorylation of T188 is not known apart from its proposed role in cardiac hypertrophy. If we speculate that purified ERK2 pT188, occurring alone or in combination with another site of phosphorylation acts to disrupt the active site, pT188-containing species would be inactive. Inactive ERK2 may have important biological

functions. For example, in situations in which ERK1/2 activity needs to be rapidly down-regulated, auto-phosphorylation of T188 might be used as an autoinhibitory mechanism. Therefore, further biochemical characterization of novel phosphorylation sites of ERK2 is necessary and relevant.

Chapter 3:
Characterization of T188 Phosphomimics T188D and T188E

Introduction

Phosphorylation alters the local charge and conformation of a protein, sometimes causing global changes. Phosphorylation of a protein is often activating, generating an active form of the enzyme from an inactive form. There are also many instances in which phosphorylation is inhibitory.

Addition of phosphates to a residue within a protein can serve several functions. The most obvious is that the addition of phosphate can alter local chemistry within a molecule thereby introducing charge-charge repulsion and change of the local static interactions. Additionally, the somewhat bulky size of phosphate can serve as a covalent appendage that can sterically block access to residues, pockets or active sites from other enzymes, cofactors or small molecules.

Substitution of normally phosphorylated amino acids by negatively charged amino acids can be useful to mimic in vivo phosphorylation when upstream kinases are not known or available.

This idea is illustrated by experiments involving the E. *coli* metabolic enzyme, isocitrate dehydrogenase. Phosphorylation of isocitrate dehydrogenase by isocitrate dehydrogenase kinase abolishes its activity. Mutagenesis experiments showing that substitution of the regulatory S113 at this position drastically altered activity (Thorsness and Koshland, 1987). Changing S113 to threonine reduced activity by 34%. Substitution with tyrosine or alanine further reduced activity. Aspartic acid substitution on the other hand blocked activity, effectively rendering the enzyme inactive. Therefore, the negatively charged

aspartate in this system was not merely inactivating, but faithfully replicated the function of phosphorylation on S113. Glutamate is also widely utilized to recreate the effect of phosphorylation by substitution.

Hypothesis and approach

In the previous chapter I described how I found ERK2 pT188 and my subsequent efforts to purify ERK2 pT188 to homogeneity. Finding ERK2 phosphorylated on T188 was fortuitous and occurred after I had started to work on mimicking pT188 with acidic substitutions. In an attempt to characterize the properties of ERK2 phosphorylated on T188, I initially made the ERK2 T188D and T188A mutants, and later T188E. The importance of Thr188 to the hydrogen-bonding network within the active site of ERK2 led us to question early on how the proposed auto-phosphorylation of Thr188 would not egregiously perturb active site contacts and allow ERK2 to retain its kinase activity toward substrate. Apart from the active site, perturbation of the P+1 substrate specificity groove in which Thr188 is resident, might at the very least affect how protein substrates could be accommodated in the active site.

Our hypothesis was simply that phosphorylation on the invariant Thr would lead to an inactive enzyme, not an enzyme with kinase activity such as that which has been previously reported. Herein, I describe biochemical experiments characterizing T188 with phosphomimics.

Results

Purification of T188D, T188A and T188E

I started with purification of T188D and T188A. Expression of T188D was performed according to the published protocols for the expression of ERK2 wt except that induction of bacteria occurred late in the log phase of bacterial growth (OD_{600} 0.8-0.9). The degree of expression and solubility of T188D paralleled that of wt ERK2 and expression of ERK2 T188D was perhaps slightly better than wt ERK2. The yield of ERK2 T188D routinely approached 15-20 mg/L. ERK2 T188A, however, proved to be largely insoluble and difficult to express in bacteria. I attempted to improve solubility of ERK2 T188A by lowering the induction temperature to16°C from 30°C and by adding sorbitol and betaine to the growth medium. Neither effort improved the overall production of ERK2 T188A in my hands. Therefore, ERK2 T188A was not included in further analyzed.

I began kinetic studies investigating the phosphorylation of ERK2 T188D and T188E compared to wt. I utilized purified ERK2 wt, T188D and T188E to test the activity toward the standard model substrate MBP. ERK2 wt readily phosphorylated MBP, whereas T188D and T188E did not. This result suggests that changes to Thr188 in the form of phosphomimetic residues result in a inactive enzyme (Fig 3.1). I also tested several additional substrates comparing ERK2 wt with ERK2 T188D activity and found that the identity of the substrate was not a factor in the loss of activity of T188D (Fig. 3.2). T188E was not tested in this set of experiments. In preliminary experiments, I tested the ability of ERK2 wt, T188D and T188E to phosphorylate a short peptide, ERKtide, based on the sequence of tyrosine hydroxylase (Prowse et al., 2000). I reasoned that T188D and T188E may be inefficient at recognizing full-length proteins, but may perform better

enzymatically if challenged with a smaller substrate. The data suggested that ERK wt, but not ERK T188D or T188E, was able to phosphorylate ERKtide and therefore neither small peptides nor protein substrates were phosphorylated by phosphomimics of ERK T188 (Fig. 3.3).

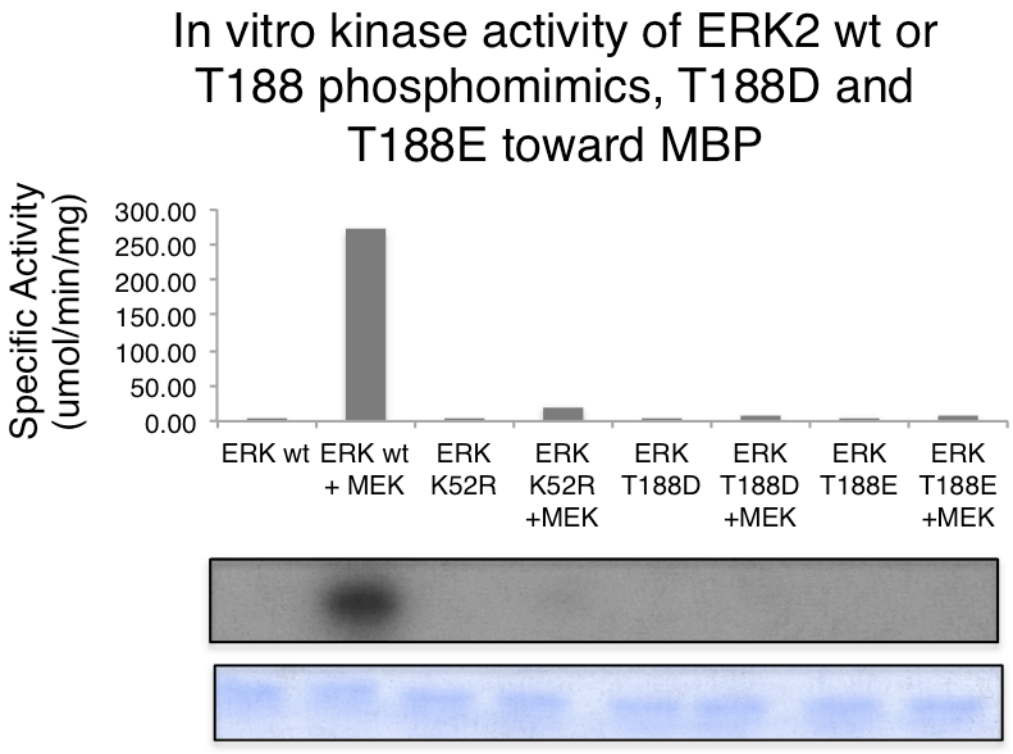


Figure 3.1. ERK2 T188 mutants (phosphomimics) have reduced activity toward substrate MBP. Autoradiogram showing ERK2 phosphomimetic mutations at T188 severely decrease activity toward substrate. Coomassie stain of MBP added to reactions are shown to demonstrate equal loading. Representative of 3 independent experiments.

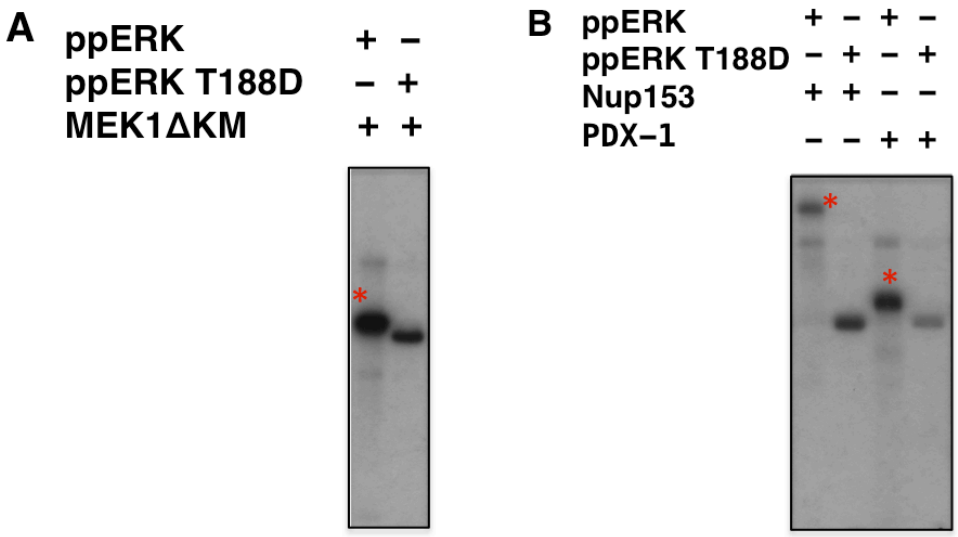


Figure 3.2. ERK T188D global loss of activity. ERK wt and T188D were phosphorylated with MEK1 and assayed for activity in an in vitro kinase assay against several known substrates (a kinase dead MEK, the transcription factor PDX-1 and the nuclear pore protein, Nup153. ERKwt can phosphorylate all substrates, T188D cannot. loses activity toward additional substrates. Red asterisk denotes substrates, which are phosphorylated by ERKwt. Signal in the T188D lanes are most likely ERKT188D bound to radiolabeled ATP, yet unable to carry out catalysis.

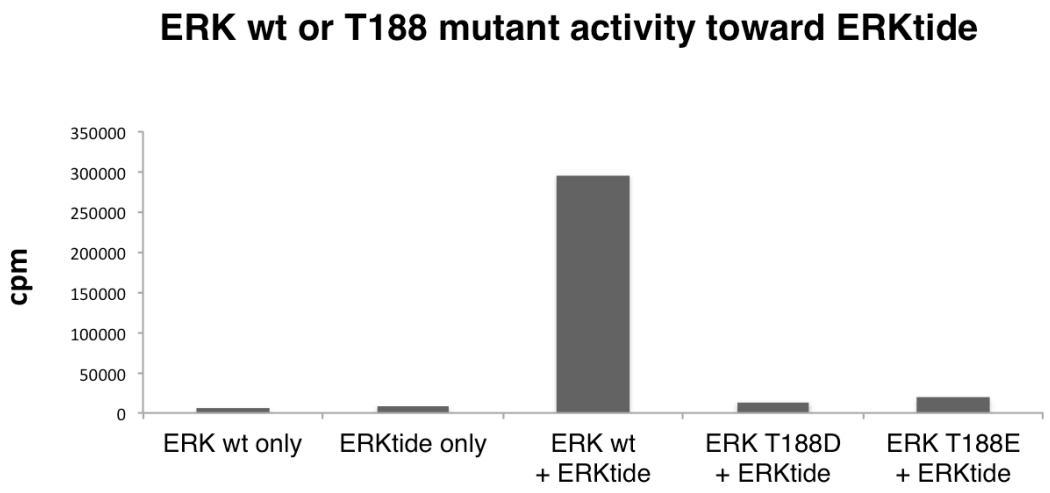


Figure 3.3. T188 phosphomimics do not have activity toward small peptides. Preliminary experiments in which ERKwt,ERK T188D and T188E were tested for their ability to phosphorylate small ERK peptide substrate, ERKtide (=ATGPLSPGPFGRR) in vitro kinase assays.

Although ERK2 T188D and T188E behave as expected, I needed to rule out that inactivity was due solely to the perturbation of T188 and its ability to make hydrogen bonding contacts and not an overall misfolding of the protein. I decided to test recognition of ERK2 proteins by R4FMEK1. MEK1/2, the upstream kinase of ERK1/2, has been shown to bind to at least two regions on ERK2. One of the two binding sites is in the MAPK insert region (Robinson, 2002) and MEK1 does not recognize ERK2 without the MAPK insert. The second binding site occurs in the common docking (CD) domain used by many substrates of ERK2 (Tanoue et al., 2000; Tanoue et al., 2002; Xu et al., 1999). The rationale is as follows: if R4FMEK1 recognizes ERK2 T188D or T188E and phosphorylates it to the same degree as wt ERK2, then the overall fold of the mutants is likely to be near that of the wild type protein even though T188 is mutated or phosphorylated. I performed this experiment by activating ERK2 wt or T188 mutants with R4FMEK1 in the presence of unlabeled ATP. Protein was probed with an antibody specific for ppERK and an antibody, which recognized ERK2 independent of its phosphorylation state. Control lanes were ppERK (positive control) and an autophosphorylation mutant, TEF (Y185F). ERK TEF is devoid of activity because it lacks the major MEK phosphorylation site, Y185, which is mutated to phenylalanine. The result of this experiment shows that T188D and T188E both are phosphorylated to an equal degree matching that of ERK wt. This suggests that, by this criterion, the overall fold of the mutants is not changed and facilitates MEK recognition and phosphorylation. In addition to immunodetection of phosphorylation, I performed an in vitro kinase assays with ERK as the substrate to measure the molar ratio of incorporated phosphate into both ERK2 wt and T188 mutants by MEK. In vitro kinase assays were performed as previously stated and phosphorylation of ERK2 was detectable by this method as well.

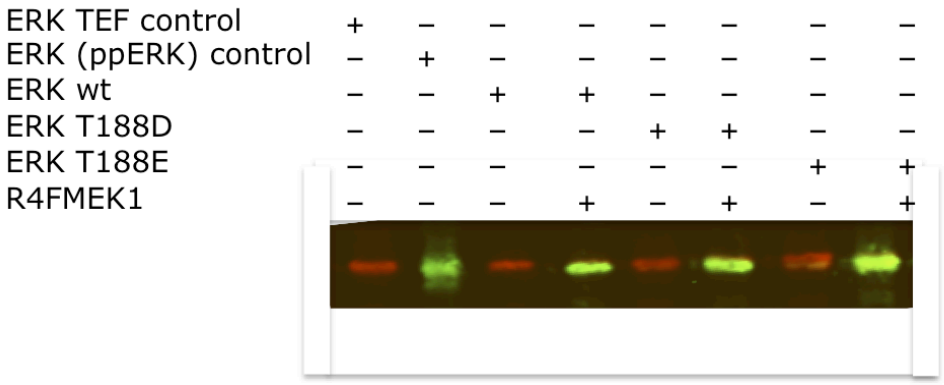


Figure 3.4. T188D and T188E are phosphorylated. A constitutively active MEK1 was used in vitro to phosphorylate ERK2 wt, the activation loop mutant,Y185F (which does not have any intrinsic auto-phosphorylation activity), or T188 mutants, T188D or T188E. Phosphorylation signal was detected by immunoreactivity to an antibody specific for doubly phosphorylated ERK2 (green) or an antibody, which detects ERK2 independent of its phosphorylation. Signal was detected using a Licor Odyssey infrared imager. Samples were run on 12% SDS-PAGE gels so gel shift is not seen

In order to determine the consequences of T188 phosphorylation as evidenced by phosphomimics, I began initial crystallography experiments with T188D. I performed a screen of conditions necessary for crystallizing T188D based on previous conditions of ERK2 and I was successful in determining conditions that yielded small crystals (Fig). Although I was not able to solve the structure due to time constraints, I am confident that my purification protocol and information derived from my initial screening trials will aid in the determination of the crystal structure.

Figure 3.5. Crystals of ERK2 T188D. Small crystals of ERK2 T188D were found when the protein (4mg/mL) was incubated with 100mM Bicine pH 8.5, 20% w/v Polyethylene glycol 10,000 20°C (hanging drop).

Discussion

Our in vitro kinetic analysis suggest that phosphomimics of ERK2 Thr188, namely T188D and T188E, lose activity toward substrates--both peptide and full-length proteins. My experiments suggest that despite changes at Thr188, ERK2 is still recognized by its upstream kinase MEK and becomes phosphorylated. Despite phosphorylation, T188D and T188E fail to become activated. These experiments confirm earlier conclusions, which suggest T188 is required for ERK2 activity. In separate experiments, ERK wt, T188D and T188E were also challenged with a short peptide harboring a phosphorylation site from tyrosine hydroxylase (ERKtide). Activated wt ERK2 phosphorylated ERKtide, whereas ERK2 T188D and T188E were not activated and failed to phosphorylated ERKtide. T188 resides in the interface between the active site and the P+1 loop. The inability of ERK2 perturbed at T188 to phosphorylate short peptides in place of full-length proteins further suggests that the substrate recognition groove as well as critical contacts in the active site may be altered. However, although this work suggests that perturbation of T188 results in inactivity toward substrates overall, it remains possible that specific subsets of proteins are substrates or become substrates for pT188 forms of ERK2. Some substrates may be able to bind ERK2 through docking sites even with T188 phosphorylated and become phosphorylated under some as yet undetermined circumstances. Specific effectors falling within this context would have to be individually analyzed. However, even if phosphorylation were to occur on select substrates in the presence of T188 perturbation, the overall rate of enzymatic phosphorylation may not compare to ERK2 wt rates for that same substrate which would still perturb overall signaling in a biological context. The fact that active site residues are more than likely misaligned and the hydrogen bonding network is disordered would suggest that inactivity would be seen with any substrate.

It also remains possible that ERK2 T188 phosphomimics do not adequately reproduce the conformation of ERK2 pT188. In a cellular setting, then phosphorylation of ERK2 T188 may serve a novel purpose. For example, in a cellular context where ERK2 activity needs to be attenuated quickly in the absence of phosphatases, ERK may autophosphorylate on T188 to rapidly attenuate activity and downstream signaling. It is unknown whether the effects of pT188 are reversible. For example, if after phosphorylation of T188 the hydrogen bonding network is broken, whether the protein can properly reorganize its active site to be regulated normally or if instead pT188 may target the protein for degradation. Analysis of the auto-phosphorylation mutant, ERK2 T183E may provide some insight. Because ERK2 T183E has some basal activity, during purification we might attempt to find fractions with pT188 but not pT183. Fractions containing pT188 could then be used in phosphatase assays with serine/threonine specific phosphatases to determine if ERK molecules previously phosphorylated on T188 can be dephosphorylated and amenable to restored basal level activity. I am currently planning final cell-based experiments to determine the correlation between in-vitro activity when T188 is perturbed and signaling in cells.

Chapter 4

ERK2 Direct Binding to DNA –A study of the Insulin gene promoter and phosphorylation states of ERK2

Introduction

Fluctuations in glucose concentration in the circulation are detected and modulated by pancreatic beta cells (β cells) within the islets of Langerhans in the endocrine pancreas. These cells respond to glucose by secreting the hormone insulin. Insulin promotes glucose uptake, utilization and storage by tissues to restore its circulating concentration to resting values. An acute response to glucose is mediated by insulin secretion. To maintain insulin for future secretion, existing pre-proinsulin is processed, insulin RNA is translated and the insulin gene is transcribed to maintain a high level of insulin transcript (German et al., 1990; German and Wang, 1994; Odagiri et al., 1996).

ERK1/2 involvement in modulation of glucose -induced insulin gene transcription has been characterized (Arnette et al., 2003). A number of transcription factors have been identified that are glucose responsive, several of which are ERK1/2 substrates (Khoo et al., 2004; Khoo et al., 2003). Experiments challenging pancreatic β-cells to conditions of elevated glucose above 5mM, result in the activation of the MAPK ERK pathway and increased association of the glucose-regulated factors as well as ERK1/2 with the insulin promoter region (Lawrence et al., 2005). For example, chromatin immunoprecipitation (ChIP) experiments suggest that ERK1/2 is recruited to specific regions of the promoter under conditions of elevated glucose and other stimulatory conditions such as stimulation by the augmenting hormone glucagon-like peptide (GLP-1), depolarization of K+ channels and activation of the MAPK pathway with mitogenic growth factors (Lawrence et al., 2008). Studies in a rat insulinoma cell line (Ins1) revealed that under conditions of high glucose (11mM) a sharp recruitment of β-cell specific transcription factors that were sensitive to

inhibition of ERK1/2 also resulted, revealing an aspect of the mechanism of activation relative to transcription of insulin mRNA.

Further evidence suggested that the presence of ERK1/2 at the promoter is enhanced by its phosphorylation and activation. Experiments with the addition of one of several MEK1/2 inhibitors (PD0325901 or U0126 or PD98059) or expression of a kinase-dead ERK2 (K52R) within β-cells interfered with insulin gene transcription. By extension and with evidence from mutagenesis experiments, active ERK1/2 serves as the kinase that phosphorylates a number of transcription factors residing at the promoter including PDX-1, Beta-2 (aka, NeuroD1) and the bHLH protein, E47.

ERK1/2 promoter residency is not limited to the insulin gene promoter, however. Among estrogen-regulated genes, ERK1/2 was recruited to estrogen response elements concomitantly with the nuclear hormone receptor, estrogen receptor α (Madak-Erdogan et al., 2011). Therefore, the importance of ERK1/2 at the promoter regions of genes is clearly demonstrated.

The insulin gene has been studied in a number of organisms, but the best characterized are those of rat, mouse and human. Sequence alignments show that the rat and mouse insulin promoter regions have close to 66% identity to the human promoter (Hay and Docherty, 2006). The insulin gene contains a number of cis-acting sequences just upstream of the promoter that serve to recruit various transcription factors, ERK1/2 and RNA polymerase II for commencement of transcription of insulin in response to glucose. In transcriptional assays using the rat insulin gene II promoter, incremental deletion of the upstream region preceding the gene aided in determining that the glucose-responsive elements of the promoter extend approximately 400 base pairs (bp) upstream of the transcriptional start site (tss) (Whelan et al., 1989), and that the most critical region of the promoter for responding to insulin production are the first 117 bp upstream of the promoter (Odagiri et al., 1996). Nevertheless, the insulin gene is modular in that distinct regions, or regulatory elements, characterize its promoter. Among the prominent regulatory elements

distributed throughout are A and E boxes, cAMP response elements (CRE), and enhancer regions. A boxes have a characteristic TAAT motif and bind the β-cell specific PDX-1 transcription factor. E boxes bind basic **h**elix-**l**oop **h**elix (bhlh) transcription factors and contain the consensus sequence (5'CANNTG).

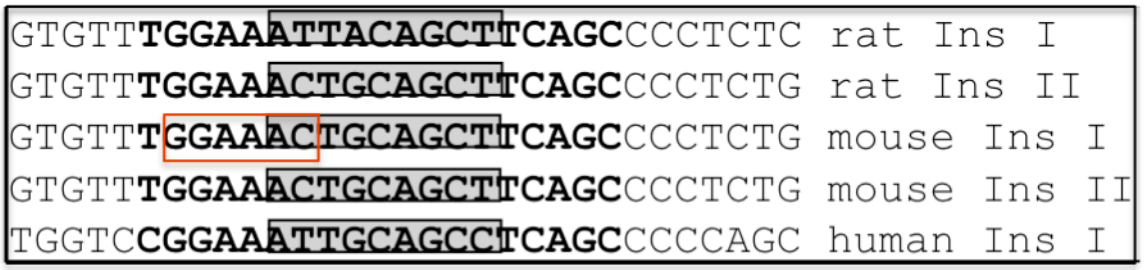


Figure 4.1. Alignment of a region of the mouse insulin promoter from rat, mouse and human. Alignment shows a high degree of sequence similarity. A consensus region in the mouse insulin promoter, outlined in a red box is a proposed MAPK (ERK1/2) binding site.

The presence of ERK1/2 at the insulin promoter region may be necessary for the phosphorylation of many associated transcription factors. However, not much is known about the direct interactions of ERK1/2 with DNA itself and the mechanism by which such an intimate association of ERK1/2 and DNA might be accomplished. To build upon initial data suggesting that ERK1/2 bound directly to DNA, our lab performed chromatin immunoprecipitation-sequencing experiments (ChIp-seq) to determine novel genes to which ERK was bound.

Chip-seq experiments were performed by immunoprecipitating crosslinked ERK1 and ERK2 from two lung specific cells lines: human bronchial epithelial (HBEC) or a K-Ras dependent adenocarcinoma cell line (H358) and from MIN6 beta cells which were treated with an ERK1/2 stimulus, epidermal growth factor (EGF) (glucose for MIN6 cells) or the MEK1/2 inhibitors PD0325901 or U0126. A library was created by amplification and

sequencing of enriched regions of interacting DNA. Hits were aligned to a human database (GRCh37; UCSC genome browser) determine novel DNA sites for ERK2 by specifically looking at the 5' and 3' un-translated regions (UTR) of genes. Many genes were identified from this initial screen. A group was selected for further study including several involved in autophagy (Unc-like kinase 1, Activating molecule in Beclin-1 Regulated Autophagy –AMBRA1) and chromatin regulation (Lysine demethylase 2a- KDM2a) as well as phosphatases (Eyes absent 1- EYA-1).

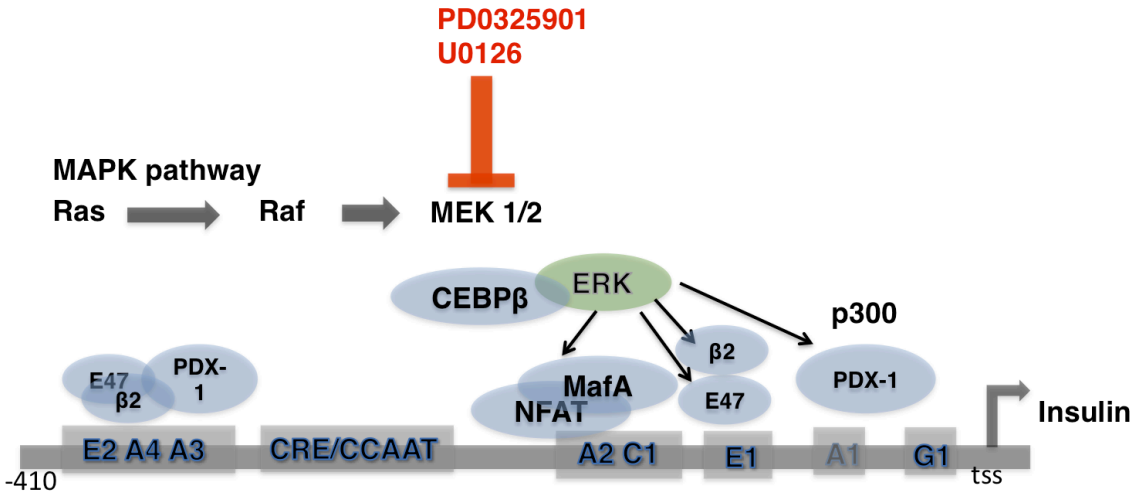


Figure 4.2. rat Insulin gene Promoter. Modular elements of the insulin promoter are shown and arrows denote that they are substrates of ERK1/2. MEK1/2 specific inhibitors are indicated in red. A, E and CRE boxes are labeled are shown relative to base pair numbering. Adapted from Lawrence et al.

Hypothesis and approach

My studies investigated ERK2 direct binding to DNA. I hypothesized that the various phosphorylation states of ERK2 conferred differential binding to DNA. First, using our well-characterized system of the insulin promoter, I began by establishing that ERK2 binds directly to the mouse insulin gene promoter using in vitro gel shift assays. I then explored the importance of phosphorylation to direct DNA binding with phosphomimics of activation loop or P+1 pocket residues within ERK2. Finally, I investigated differential phosphorylation and conformational states of ERK2 and their effects on selective binding to various promoter regions of ERK1/2-associated genes validated from the Chip-seq screen. Overall, I sought to establish patterns of ERK2:DNA binding to ultimately establish characteristics of kinase involvement in high-order regulation of genes.

Results

ERK2 (ppERK) binds directly to the mouse insulin promoter region

I began testing direct ERK2 DNA binding by using the standard electrophoretic gel mobility shift assay (EMSA). In standard gel shift assays, radioactively labeled oligonucleotides are incubated with nuclear extract or purified protein and resolved by electrophoresis on a low percentage, sodium dodecyl sulfate polyacrylamide gel (SDS-PAGE). If oligonucleotides are bound by any number of proteins within a nuclear extract or by a purified protein, a shift of the mobility of the oligonucleotide will occur which runs more slowly than free oligonucleotide on a gel. Initially, I invested a great deal of effort in optimizing the oligonucleotide binding assay with radiolabeled oligonucleotides. However, in my hands this approach was not consistently reproducible. Therefore, I modified the standard gel shift assays for use with our Licor Odyssey Infrared (IR) scanner. Oligonucleotides were purchased with an infrared dye attached at the 5' end purified by high performance liquid chromatography (HPLC) to separate labeled oligonucleotide from free, uncoupled dye label. IR labeled oligonucleotides were incubated with purified ERK2,

resolved on a non-denaturing polyacrylamide gel and imaged (700 nm).

I used the doubly phosphorylated form of ERK2 (referred to as ppERK), and a synthetic oligonucleotide representing a 19 bp region of the mouse insulin promoter previously designed in our lab to assay direct binding. I used the mouse insulin gene promoter in these assays, because alignment of rat, mouse, and human insulin genes shows a high degree of similarity (Fig. 4.1). Furthermore, the oligonucleotide used in these studies contained the sequence G/GAAAG/C found to be a widely utilized consensus sequence in yeast MAPK transcription and found to be a sequence recognized by ERK2 in other contexts of transcriptional recruitment.

I began by reproducing the experiment testing ppERK binding to the 19bp oligonucleotide with the same incubation conditions reported by Hu et. al that tested for direct binding of ERK2 to an oligonucleotide harboring the consensus sequence 5'G/GAAAG/C3' found to be important in yeast MAPK transcription. . I found that visualization of gel shifts was more consistently reproducible using the Licor method. Titration of an increasing concentration of ppERK with a constant oligonucleotide concentration shows that ppERK binds directly to the 19bp oligonucleotide in a concentration-dependent manner (Fig. 4.3). Binding to the labeled oligonucleotide can be specifically competed with unlabeled oligonucleotide.

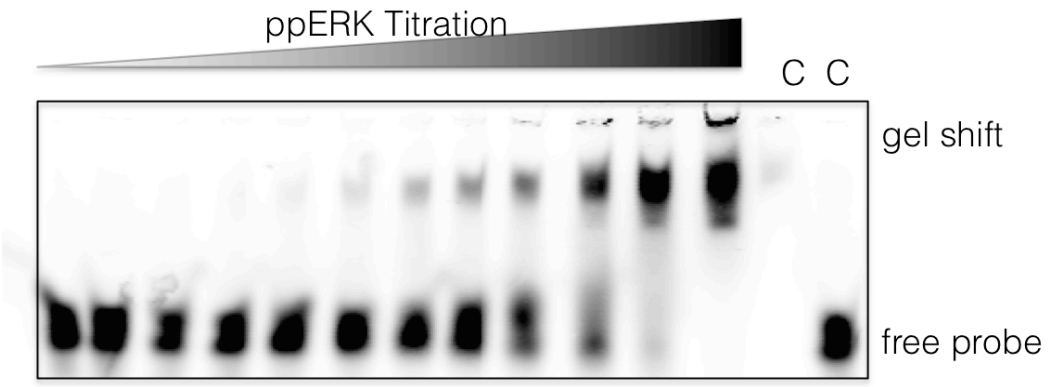


FIgure 4.3. ppERK binds to insulin gene promoter DNA. Electrophoretic mobility shift assay (EMSA) showing ppERK binds directly to a synthetic oligonucleotide representing a section of the mouse Insulin promoter region in a dose dependent manner. ppERK is titrated with constant19bp mouse insulin oligonucleotide.
C=control oligonucleotides minus ppERK.

ERK2 differentially binds oligonucleotides derived from genes found in Chip-seq promoter regions depending on its activation / phosphorylation state.

I collaborated with Dr. Elhadji Dioum, a former postdoc in our laboratory, to investigate the requirements for ERK2:DNA binding. In early experiments, I purified ERK2 wt protein that was not co-expressed with MEK1/2. This protein probably had some perhaps heterogeneous phosphorylation (see chapter 1). I found that different fractions of ERK2 from the ion exchange purification scheme differentially bound to the 19bp insulin oligonucleotide (Fig. 4.4). In an attempt to resolve factors contained in each peak that conferred discriminatory binding, eluted protein was subjected to mass spectrometry and found to have different auto-phosphorylation states (see chapter 1).

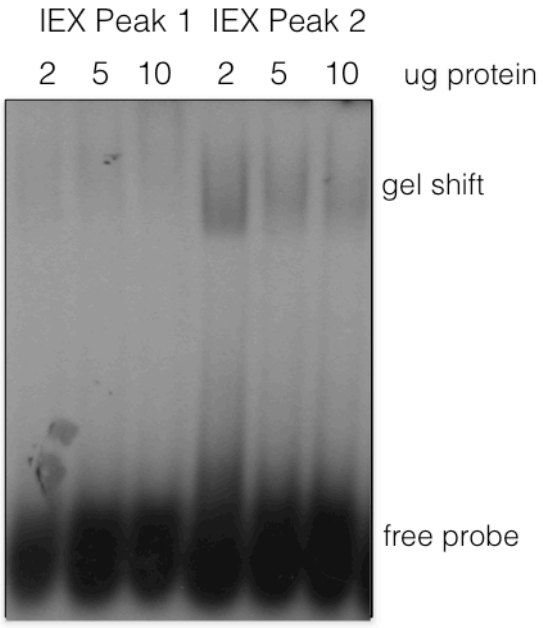


Figure 4.4. Purified recombinant ERK2 (non ppERK2) binds directly to DNA. Electrophoretic mobility shift assay (EMSA) was performed on fractions from each eluting peak of an IEX purification of recombinant ERK2. Fractions were pooled and microgram amounts of protein were incubated with a radiolabeled, synthetic oligonucleotide representing a 19bp region of the mouse insulin oligonucleotide. The result is direct DNA binding.

The detection of several phosphorylation states within the fractions recovered from the purification of ERK2 led me to hypothesize that various phosphorylation states may affect binding specificity to DNA. In order to test this unambiguously, I decided to take advantage of two ERK auto-phosphorylation mutants previously characterized by our lab (Robbins et al., 1993). T183E and Y185F are mutations of ERK2 activation loop residues that must be phosphorylated by MEK to activate ERK. T183E results in a decrease in activity relative to wt and Y185F ERK2 fail to be activated by MEK; neither has much activity. I decided to utilize these mutations in comparison with fully active ppERK to probe for differential binding.

I tested oligonucleotides representing the following genes from the ChIP-seq screen performed in our laboratory: Eya-1 phosphatase, Unc-like kinase 1 (ULK-1), lysine demethylase 2a along with mouse insulin (Table 4.1). Oligonucleotides were synthesized based on the 5' UTR region of each of the genes. All except for KDM2a contained the consensus G/CAAAG/C sequence that had been defined previously.
ppERK bound strongly to the mouse insulin promoter region while T183E and Y185F did not. The oligonucleotide from KDM2a associated poorly with all ERK2 forms tested. The ULK-1 oligonucleotide bound ppERK best, and T183 ERK2 bound to a lesser degree. Finally, the oligonucleotide from the EYA-1 gene seemed to bind equally to ppERK and Y185F ERK2. This result is interesting because it raises the question of yet another higher-order type of regulation: recruitment of heterogeneous dimers.

Table 4.1 Oligonucleotides representing ChIP-seq identified ERK1/2 associated gene promoters

Gene name	Sequence written 5' to 3'
mouse Insulin promoter 48bp REV	GTCAGCAGATGGCCAGAGGGGCTGAAGCTGCAGTTTCCAAACACTTCC
mouse Insulin promoter 48bp FWD	GGAAGTGTTTGGAAACTGCAGCTTCAGCCCCTCTGGCCATCTGCTGAC
mouse Insulin short 19bp FWD	GGAAACTGCAGCTTCAGCC
mouse Insulin short 19bp REV	GGCTGAAGCTGCAGTTTCC
mouse Insulin promoter 40bp FWD	AGATGTGTTGACGTCCAATGAGTGCTTTCTGCAGACCTAG
mouse Insulin promoter 40bp REV	CTAGGTCTGCAGAAAGCACTCATTGGACGTCAACACATCT
human Lysine Demthylase 2a FWD (KDM2a)	GAGGCGGGGCCAGGCCGCTGCACCGAGGAGCCCCCGGGCCCAGGGCAGACGGGCCCCGTG
human Lysine Demthylase 2a REV	CACGGGGCCCGTCTGCCCTGGGCCCGGGGGCTCCTCGGTGCAGCGGCCTGGCCCCGCCTC
Unc-like Kinase-1 FWD (ULK-1)	GGAAAAGGAGGGAGGAGGACGGAGAGAGGGGAAAGG
Unc-like Kinase-1 REV	CCTTTCCCCTCTCTCCGTCCTCCTCCCTCCTTTTCC
Eyes absent-1 FWD (EYA-1)	GGTTGAATGAAAATGCACATCAAAGAT
Eyes absent-1 REV	CCAACTTACTTTTACGTGTAGTTTCTA

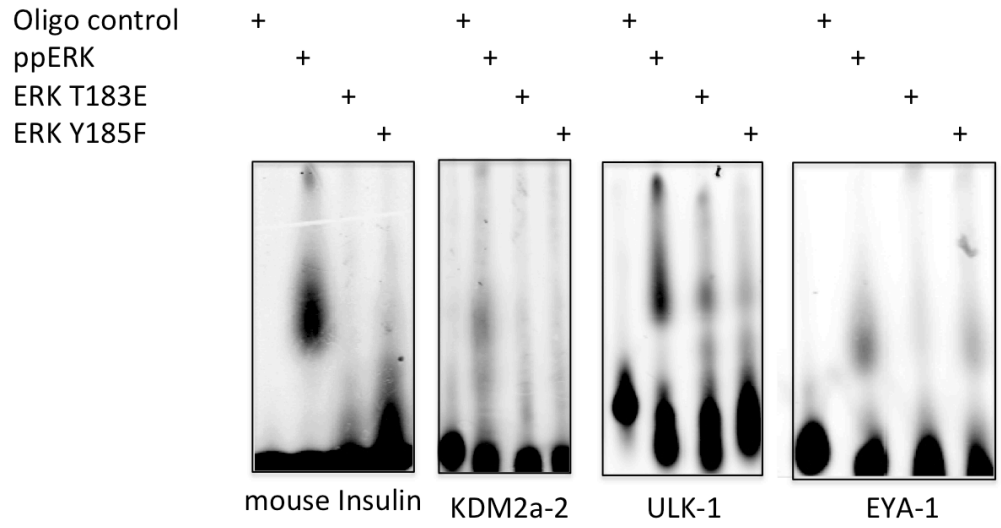


Figure 4.5. Differential binding to Chip seq oligonucleotides representing portions of various gene promoter regions by recombinant ERK2 or ERK2 activation loop mutants. Various forms of ERK2 represented by mutations in the phosphorylation (activation) loop display differential binding to oligonucleotides representing enriched promoter sequences identified in Chip-seq screen (unpublished).

The structure of low activity ERK2 (pdb:1ERK) shows the activation loop makes critical contacts to the MAPK insert. The MAPK insert has been demonstrated to be an important contact point for direct binding to DNA. In addition to ERK2 direct DNA contacts, direct kinase:DNA interactions have also been described in bacterial systems. The bacterial persistence factor, HipA, possesses an overall eukaryotic kinase structure similar to ERK2 in that its amino terminal domain is a mixture of α-helices and β-sheets while its carboxy terminal domain is mainly α-helical (Evdokimov et al., 2009). Interestingly, HipA has kinase activity because it contains important catalytic residues found in other kinases including the catalytic Asp (Asp309). HipA, usually found bound to the bacterial neutralizing protein HipB, was found to phosphorylate the bacterial protein EF-Tu. More importantly, upon bindng to ATP, HipA was found to undergo a conformation rotation of its two domains and bind directly to DNA through a portion of the molecule located in the same region of the molecule as the MAPK insert in ERK2 (Schumacher et al., 2009).

Returning to my T188 phosphomimics, I hypothesized that perturbation of ERK2 T188 by mutagenesis and, therefore, perturbation of the activation loop as a result, might influence contacts to DNA via the MAPK insert. I tested the T188 phosphomimics for direct DNA binding. Based on my previous evidence that doubly phosphorylated ERK2 bound directly to DNA, I wished to test the requirement of phosphorylation of T188D and T188E for binding to DNA. I hypothesized that changes at T188 would affect the activation loop without the need for activation loop phosphorylation by MEK and that these mutants would be able to bind to DNA as a result. Indeed, data show that un-phosphorylated ERK T188D and T188E bind to the 19bp mouse insulin oligonucleotide, but un-phosphorylated ERKwt does not (Fig. 4.6).

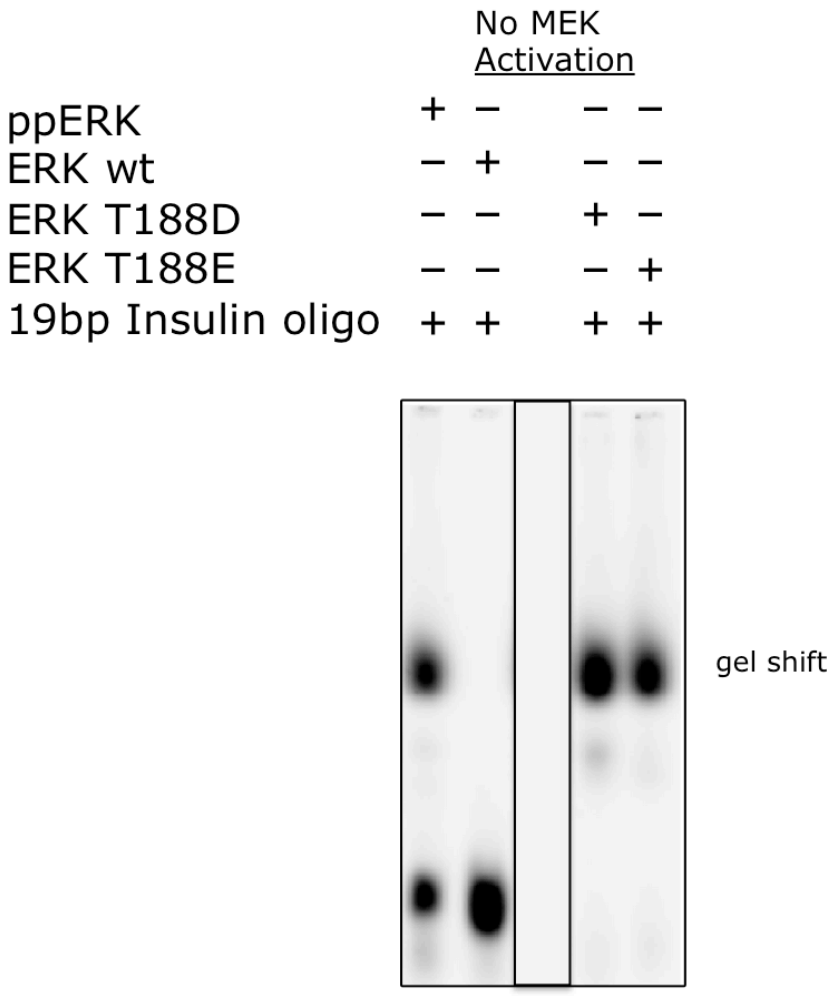


Figure 4.6. ERK2 T188D/E binds to DNA apart from activation loop phosphorylation. Recombinant ERK2 (co-expressed with constitutively active MEK1 which served as a positive control), or recombinant ERK2 not expressed with MEK1 was incubated with a 19bp oligonucleotide representing the mouse Insulin gene promoter region. Binding is seen with inactive T188D or T188E mutations independent of upstream activation by MEK1.

Discussion

ERK2 has an important role in regulating insulin gene transcription. Thus, the direct association of ERK2 with insulin gene sequences may contribute to this important function. Recruitment of ERK1/2 to the promoter site of the insulin gene is critical for the phosphorylation of transcription factors. Existing data from our lab suggested that ERK1/2 could make contacts with DNA directly. Based on mutagenesis of the MAPK insert region of ERK2, direct binding of ERK1/2 with an oligonucleotide containing the canonical MAPK consensus DNA binding sequence showed it was dependent on the insert. Indirect luciferase expression experiments using portions of an interferon γ promoter element suggested that ERK2 acts as a transcriptional repressor (Hu et al., 2009). Data from this characterized interaction suggests that activity is not necessary for interaction, because a kinase-dead form of ERK2 bound to the interferon-γ promoter derived oligonucleotide. Although provocative, many questions remained about the mechanism and function of binding of ERK2 to DNA elements.

First, is direct binding to DNA dependent upon the presence of transcription factors or other associated molecules? Considering the amount of ERK2 substrates resident at the insulin promoter, it was surprising to find that ERK2 bound directly to our synthetic oligonucleotide in the absence of other proteins. Based on other data, one assumption that could have been made is that ERK2 binds cooperatively to DNA in a fashion dependent on the presence of cofactors or substrates. Because our system tested a purified oligonucleotide binding to a purified ERK2, our data suggest that pre-recruitment of ERK substrates is not required for ERK2 binding. When ERK2 was characterized for binding to estrogen response elements (ERE) in the presence of ERα, it was demonstrated that ERα presence at the ERE was required to recruit ERK2; however, ERK2 presence at the promoter was not required for the recruitment of ERα (Madak-Erdogan et al., 2011).

Clearly, protein kinase binding to DNA is complex and often combinatorial.

Another question needed to illuminate the mechanism for direct DNA binding is the requirement for ERK2 activation. Data from the interferon-γ promoter suggests that activation is not needed for DNA binding because kinase-dead ERK2 can bind directly to oligonucleotides. ChIP-seq data from our lab shows that for certain genes, ERK2 is recruited to DNA even when cells were treated with the MEK inhibitor PD0325901 before stimulation with EGF. However, this demonstration of inactive ERK binding to DNA may not be so clear cut. Although MEK1/2 inhibition would suggest that a pool of non-phosphorylated inactive ERK relative to activated ERK may exist in the promoter region, inhibition of ERK2 phosphorylation is not absolute. Blots of cell fractions suggest that the inhibitor is less effective in reducing nuclear pERK2 relative to cytosolic pERK2; therefore, the fraction of phosphorylated ERK2 forms in the ChIP-Seq studies is ill-defined. Inhibition of MEK1/2 does not automatically guarantee that ERK2 will be un-phosphorylated or inactive ERK2. Nevertheless, QPCR data from our lab shows that several genes in H358 cells experiences ERK2 recruitment and an increase in message in the presence of MEK1/2 inhibition. Therefore, inactive ERK2 may be recruited to promoters and inactivation does not necessarily down-regulate transcripts. Inactive ERK2 represented by T188 phosphomimics showed direct binding to DNA which would support this hypothesis. Finally, the total catalog of ERK2 phosphorylation or other modification states are not known. As methods to measure modifications improve, the relationships between modifications and functions will become clearer.

A future direction hopefully will be quantitative measurement of the affinities for ERK2 to the tested oligonucleotides. Dissociation constants will provide unambiguous information on the differences between ERK2 phosphorylation states and the abilities of different ERK2 phosphoforms to differentially bind specific oligonucleotide sequences. Another dimension to direct binding is the ability of the different states of ERK2 to bind at different places within the same promoter region. For example, sequence specificity on either side of the G/CAAAG/C consensus sequence is most likely important for binding to

ERK2. It would be interesting to quantitatively measure which phosphorylation state(s) are recruited to specific sequences. Quantitative data reflecting these differences might serve two very useful purposes. First, our knowledge of how a single gene is regulated at multiple intragenic regions would be enhanced, and second, quantitative data such as these would help to validate often confounding chip-sequencing data. However, organically, this body of work testing the various phosphorylation states and structural rearrangements of ERK2 as they relate to direct DNA binding, suggests that a higher-order mechanism of transcriptional regulation by ERK2 exists apart from its residency at gene promoters and it role in phosphorylation of transcription factors.

Future Directions

Investigating the biological relevance of pT188

The functional relevance of ERK2 pT188 has been reported as being important to the transcriptional regulation of genes important in cardiac hypertrophy. This work hypothesized, although it did not show biochemically, that ERK pT188 acts to phosphorylate transcription factors important for the regulation of cardiac hypertrophic genes. Further work suggested that the abolition of ERK2pT188 species reduced pathological cardiac hypertrophy (Ruppert et al., 2013) suggesting the presence of pT188 ERK2 acts egregiously in cardiac tissue. Based on the in depth structural knowledge of eukaryotic kinases and ERK2 in particular, the likelihood that pT188 retains it catalytic activity is low because T188, the invariant threonine present in most protein kinases, is required for activity. Therefore, this raises the question as to how pT188 ERK2 functions in vivo catalytically and what biological purpose it serves, if any. In this body of work, I found that perturbation of T188 results in the complete loss of catalytic activity, which is in opposition to initial work characterizing ERK2 pT188 autophosphorylation. However, I hypothesize that autophosphorylation of T188 and the loss of ERK2 catalytic activity may have a biological function.

Among its many substrates, ERK2 was found to phosphorylate the embryonic stem cell factor, Kruppel-like factor 4 (Klf-4) in mice (Kim et al., 2012). Klf-4 is a zinc finger transcription factor important in somatic cell reprogramming and negative regulation of apoptosis. Klf-4 induces the self-renewal of stems cells. Phosphorylation of Klf-4 on S123 was shown to cause its degradation. Therefore, inhibiting ERK1/2 pathway would inhibit Klf-4 phosphorylation. Inhibition of Klf-4 phosphorylation leads to its stabilization. Persistent levels of Klf-4 would support self-renewal of stems cells and the ability to bypass differentiation. Therefore pT188 could play a part in a developmental context. For example, temporally, when self-renewal must trump differentiation, the presence of pT188

(non-functional ERK2) could mimic the turning off of the ERK1/2 pathway in absence of phosphatases to down-regulate its own activity and prevent phosphorylation of substrates in a biologically controlled, temporally specific manner. Therefore, in the presence of pT188, which would equal the absence of a functional ERK1/2 pathway overall, stems cells persist. This would be an interesting phenomenon to study in the context of cancer stems cells. For example, with improved antibodies to pT188, cancerous tissues in which Klf-4 is found in higher than normal levels, can be probed for ERK pT188 to ascertain if expression of this ERK2 autophosphorylated form coincides with a cancer stem cell phenotype.

Determination of the DNA binding patterns of ERK2 based on its phosphorylation state.

In this body of work, I demonstrated that changes in the ERK2 activation loop affected binding to DNA. I showed that, doubly phosphorylated ERK2 directly bound a short oligonucleotide representing a region of the mouse insulin promoter. In experiments where unphosphorylated ERK2 was incubated with the short oligonucleotide, binding was abolished. However, ERK2 mutated at T188 bound oligonucleotide independent of its phosphorylation by MEK. This suggested that's changes within the P+1 pocket in which T188 resides, and therefore the active site may have an effect on mediating DNA binding. Therefore, in an in vivo context, the presence of pT188 could indeed bind directly to DNA without additional phosphorylation and because it is assumed pT188 species differs from that of the canonical pT183pY185 ERK2 active species, the regions of DNA, chromatin and ultimately ERK2 responsive genes may be different for various ERK2 phosphorylated species.

A long-term goal for this project would be to characterize this proposed higher-order ERK2 mediated transcriptional regulation. One approach would be to test various tagged activation loop and T188 mutants for DNA binding in vitro. Using ChIP-chip methodology, tagged ERK2 mutants would be expressed and chromatin

immunoprecipitation would be carried out to enrich for specifically bound DNA sequences based on ERK2 structure. Micro arrays would then follow by probing gene chips with enriched, ERK2 structure specific sequences and overlapping genes would be discarded. This method may yield information into which genes are responsive to a particular conformation of ERK2 and may allow the segregation of ERK2 responsive genes into categories based on ERK2 phosphorylation.

A short-term goal for characterizing ERK2 specific DNA binding would be the employment of biophysical methods to measure the affinity of structurally dissimilar ERK2 proteins to binding of DNA sequences. We have demonstrated that ERK2 mutations in the activation loop can discriminate between different oligonucleotides. I attempted to develop a fluorescence polarization assay to measure the dissociation constants of ERK2 binding to DNA by this method. I fully appreciated that binding of ERK2 to DNA in a strictly in vitro setting (i.e. without cofactors, etc.) may result in low affinity measurements, however binding was seen in gel shift assays. I would like to explore the quantitative characterization of ERK binding to DNA to generate reliable dissociation constants. Several questions could be answered from these studies. First, what is the order of magnitude in terms of the differences in binding between different forms of ERK2 and a single oligonucleotide. Are differences in binding twofold or tenfold? Furthermore, a single form of ERK2 (for example, doubly phosphorylated ERK) can be used to measure quantitative differences as they relate to sequence specificity. For example, mutagenesis of our 19bp mouse insulin oligonucleotide tested with ppERK2 could help us to determine sequence specificity apart form the known preferred consensus 5'G/CAAAG/C3' motif. These experiments could be performed using Isothermal Titration Calorimetry (ITC) in addition to fluorescence polarization.

Crystallographic analysis of T188 mutants and binding to DNA

Ultimately, the structural determination of ERK2 binding to DNA using crystallography would be ideal because it would show exactly how DNA is bound to ERK2, but also what long-range interactions are introduced into ERK2 itself when bound to DNA. With binding constants from polarization or ITC measurements, ERK2:DNA complexes, which have a relatively tight binding constant, could be chosen for crystallography screens. However, even if binding constants are not close to those found in vivo for most transcription factors (and they most likely would be higher due to the simplistic system of purified DNA challenged with short oligonucleotides), there is a possibility to obtain co-crystals. From personal communication with Dr. Maria Schumacher at Duke university, generating protein:DNA crystals starts with trials where recombinant protein is mixed with an equimolar concentration of oligonucleotide (in some cases a 1.2:1 oligonucleotide:protein mixture was shown to yield co-crystals). Therefore, it is plausible that the structural determination of ERK2 bound directly to DNA could be determined.

Materials and Methods

Rat ERK2 recombinant protein expression and histidine tag purification

Rat ERK2 harboring an amino terminal 6X Histidine tag was expressed from the NpT7 promoter construct in BL21 DE 3 bacteria. Constructs were transformed into BL21 DE 3 cells and grown in Lennox broth to an OD_{600} =0.8-0.9 at 37°C. When cells reached this OD_{600}, induction was facilitated by the addition of Isopropyl thiogalatopyranoside (IPTG) to a final concentration of 0.5mM and the temperature was lowered to 30°C where induction was carried out overnight. Cell were harvested in a sonication buffer (50mM NaHPO4, 300mM NaCl, with the addition of protease inhibitors) and pellets were flash frozen in liquid nitrogen and stored at -80°C until purification. Upon purification, pellets were thawed rapidly in tepid water and lysozyme (1mg/mL) stock was added and cells and incubated on ice. Cell lysate was sonicated with a Fisher model 500 sonicator, duty equal to 60% for 1 minute with 2 sec pulses for a total of three times on ice. The lysate was then centrifuged using a Ti-45 ultracentrifuge rotor at 105,000xg, for 1 hour at 4°C and in some instances, DNASE I was added to further clear DNA debris. The clarified lysate was added to nickel resin (Clontech) and allowed to batch bind overnight at 4°C with stirring. Bound 6XHis ERK protein was washed three times with ice cold sonication buffer plus protease inhibitors followed by three washes with a low imidazole containing wash buffer (50mM Tris, pH 7.4, 150mM NaCL, 20mM Imidazole, pH 7.5). The washed resin was then subjected to an imidazole gradient (20-200mM) and eluting fractions were analyzed by SDS-PAGE gels, which were coomassie stained. Fractions containing recombinant ERK2 as judged by SDS-PAGE were pooled and immediately diluted with a low salt buffer used for Ion Exchange purification (50mM Tris, pH8.0, 20mM NaCL, 1mM DTT, 0.5mM EDTA). Buffer exchange was facilitated by the addition of diluted protein to centricon protein concentrators where buffer was exchanged by centrifuging the concentrators at 3500xg, twice. ERK2 protein in the low salt buffer was then added to an Anion Exchange Mono Q

column for further purification.

Ion Exchange purification of 6XHis tagged ERK2

ERK2 (exchanged into a low salt ion exchange buffer (20mM NaCL)) was added to the Mono Q column with a 10mL superloop. Purification was facilitated by a linear gradient of NaCL (20mM-1M) with a flow rate of 0.4mL/min for 7.5 column volumes with fractions collected every 0.4mL. Fractions were collected and analyzed by SDS-PAGE and/or immunoblotting to determine the purest fractions. Individual fractions (or pooled fractions in some instances) were subjected to mass spectroscopy. ERK2 protein not analyzed by mass spectrometry was exchanged into ERK storage buffer (____) with the addition of 10% glycerol final, aliquotted, flash frozen with liquid nitrogen and stored at -80°C until further use.

In vitro Kinase assays

ERK2wt or mutants thereof were activated at 3uM with a constitutively active MEK1 (R4FMEK1) at a concentration of 40-50nM in vitro in activation buffer (10mM Tris, 1mM DTT, 1mM Benzamidine, 10X ATP mix [100uM ATP, 150mM $MgCl_2$ containing $^{32}P\gamma ATP$ with a specific activity of 200-300 cpm/pmol]). Reactions were allowed to proceed at 30°C for 30 minutes after which reactions were stored on ice. To assess phosphorylation of substrate, Myelin Basic protein (MBP) was added at 50ug/mL final to activation buffer plus the addition of phosphorylated ERK2 at a final concentration of 0.3uM. Reactions were allowed to proceed for 10 minutes 20°C and reactions were stopped with the addition of 6X SDS-PAGE loading dye (Tris, glycerol, bromophoenol blue) and heated to 90°C to denature proteins. Reactions were run on SDS PAGE gels, coomassie stained and destained (5% Methanol, 7% glycerol) and dried under vacuum. Dried gels were exposed to film and bands were excised and scintillation counted with scintillation fluid and

experimental controls serving as a background control during counting. Counts per minute (CPMs) were averaged as reactions were performed in triplicate.

Peptide phosphorylation assays

Peptide phosphorylation assays were performed similarly to MBP phosphorylation assays except reactions were spotted onto P81 paper and washed three times with 75mM H_3PO_4 for 10 minutes, then once with acetone for five minutes. Paper strips were dried overnight and counted with a scintillation counter to get counts per minute.

Electrophoretic mobility shift assays

EMSA assays were performed by incubating ERK2 (2ug) in binding buffer (10mM HEPES pH 7.4, 1mM DTT, 1mM Benzamidine, 15mM $MgCl_2$, 0.5ug/mL poly didc) and infrared labeled oligonucleotide (Alexa fluor 700) for 30 minutes at 30°, followed by an overnight incubation at 4°C. Reactions were run on a 5% non-denaturing gel at 20°C in an EMSA running buffer (10mM Tris pH 7.4, 400mM Glycine) and imaged on a Licor Odyssey imager.

Mass spectrometry analysis

(performed by Yan Li, Ph. D. (NIH) as relayed to me via written communication) "Extracted peptides were analyzed by a nano-LC/MS/MS system with an Ultimate 3000 HPLC (Thermo-Dionex) connected to an Orbitrap Elite mass spectrometer (Thermo Scientific) via an Easy-Spray ion source (Thermo Scientific). Peptides were separated on ES800 Easy-Spray column (75 µm inner diameter, 15 cm length, 3 µm C18 beads; Thermo Scientific) at a flow rate of 300 nl/min with a 35 min linear gradient of 2–24% mobile phase B (mobile phase A: 2% acetonitrile, 0.1% formic acid; mobile phase B: 98% acetonitrile, 0.1% formic acid). Thermo Scientific Orbitrap Elite mass spectrometer was operated in positive nano-electrospray mode. MS data were acquired in both profile and data

dependent modes. The resolution of the survey scan was set at 60k at m/z 400 with a target value of 1×10^6 ions. The m/z range for MS scans was 300–1600. The isolation window for MS/MS fragmentation was set to 1.9 and the top two most abundant ions were selected for product ion analysis. Ion trap enhanced scan rate was used for the MS/MS data acquisition with decision tree procedure activated. Dynamic exclusion was 12 s and early expiration was abled. Xcalibur RAW files were converted to peak list files in mgf format using Mascot Distiller (version 2.4.3.3). Database search was performed using Mascot Daemon (2.4.0) against a house-built database containing NCBI human sequences and Tag+Erk2 sequence."

Appendix A: Phosphorylation states found within an ERKwt bacterial expression purification as detected by mass spectrometry analysis.

1. Results of 39

Protein coverage: 98%. Matched peptides shown in **red**.

```
  1 GGLTMAHHHH HHMAAAAAAG PEMVRGQVFD VGPRYTNLSY IGEGAYGMVC
 51 SAYDNLNKVR VAIKKISPFE HQTYCQRTLR EIKILLRFRH ENIIGINDII
101 RAPTIEQMKD VYIVQDLMET DLYKLLKTQH LSNDHICYFL YQILRGLKYI
151 HSANVLHRDL KPSNLLLNTT CDLKICDFGL ARVADPDHDH TGFLTEYVAT
201 RWYRAPEIML NSKGYTKSID IWSVGCILAE MLSNRPIFPG KHYLDQLNHI
251 LGILGSPSQE DLNCIINLKA RNYLLSLPHK NKVPWNRLFP NADSKALDLL
301 DKMLTFNPHK RIEVEQALAH PYLEQYYDPS DEPIAEAPFK FDMELDDLPK
351 EKLKELIFEE TARFQPGYRS
```

Phosphopeptide/s detected:

Ions detected	Corresponding sequence	Phosphorylation site/region
m/z 1112.49 z = 2 m/z 741.99 z = 3	VADPDHDHTGFLTEYVATR with 1 Phospho (STY)	The MS/MS spectra showed that the phosphorylation site is Y15. .

2. Results of 40

Protein coverage: 97%.

Matched peptides shown in **red**.

```
  1 GGLTMAHHHH HHMAAAAAAG PEMVRGQVFD VGPRYTNLSY IGEGAYGMVC
 51 SAYDNLNKVR VAIKKISPFE HQTYCQRTLR EIKILLRFRH ENIIGINDII
101 RAPTIEQMKD VYIVQDLMET DLYKLLKTQH LSNDHICYFL YQILRGLKYI
151 HSANVLHRDL KPSNLLLNTT CDLKICDFGL ARVADPDHDH TGFLTEYVAT
201 RWYRAPEIML NSKGYTKSID IWSVGCILAE MLSNRPIFPG KHYLDQLNHI
251 LGILGSPSQE DLNCIINLKA RNYLLSLPHK NKVPWNRLFP NADSKALDLL
301 DKMLTFNPHK RIEVEQALAH PYLEQYYDPS DEPIAEAPFK FDMELDDLPK
351 EKLKELIFEE TARFQPGYRS
```

Phosphopeptide/s detected:

Ions detected	Corresponding sequence	Phosphorylation site/region
i) m/z 741.99 z = 3 ii) m/z 741.99 z = 3	VADPDHDHTGFLTEYVATR with 1 Phospho (STY) and VADPDHDHTGFLTEYVATR with 1 Phospho (STY)	Two forms of the phosphopeptide were found: peptide phosphorylated on T18 (VADPDHDHTGFLTEYVATR) and Y15 (VADPDHDHTGFLTEYVATR) The signal corresponds to VADPDHDHTGFLTEYVATR is stronger than that of VADPDHDHTGFLTEYVATR

3. Results of 41

Protein coverage: 97%.

Matched peptides shown in **red**.

```
  1 GGLTMAHHHH HHMAAAAAAG PEMVRGQVFD VGPRYTNLSY IGEGAYGMVC
 51 SAYDNLNKVR VAIKKISPFE HQTYCQRTLR EIKILLRFRH ENIIGINDII
101 RAPTIEQMKD VYIVQDLMET DLYKLLKTQH LSNDHICYFL YQILRGLKYI
151 HSANVLHRDL KPSNLLLNTT CDLKICDFGL ARVADPDHDH TGFLTEYVAT
201 RWYRAPEIML NSKGYTKSID IWSVGCILAE MLSNRPIFPG KHYLDQLNHI
251 LGILGSPSQE DLNCIINLKA RNYLLSLPHK NKVPWNRLFP NADSKALDLL
301 DKMLTFNPHK RIEVEQALAH PYLEQYYDPS DEPIAEAPFK FDMELDDLPK
351 EKLKELIFEE TARFQPGYRS
```

Phosphopeptide/s detected:

Ions detected	Corresponding sequence	Phosphorylation site/region
i) m/z 741.99 z = 3 ii) m/z 741.99 z = 3	VADPDHDHTGFLTEYVATR with 1 Phospho (STY) and VADPDHDHTGFLTEYVATR with 1 Phospho (STY)	Two forms of the phosphopeptide were found: peptide phosphorylated on T18 (VADPDHDHTGFLTEYVATR) and Y15 (VADPDHDHTGFLTEYVATR) The signal corresponds to VADPDHDHTGFLTEYVATR is stronger than that of VADPDHDHTGFLTEYVATR

4. Results of 42

Protein coverage: 97%.

Matched peptides shown in **red**.

```
  1 GGLTMAHHHH HHMAAAAAAG PEMVRGQVFD VGPRYTNLSY IGEGAYGMVC
 51 SAYDNLNKVR VAIKKISPFE HQTYCQRTLR EIKILLRFRH ENIIGINDII
101 RAPTIEQMKD VYIVQDLMET DLYKLLKTQH LSNDHICYFL YQILRGLKYI
151 HSANVLHRDL KPSNLLLNTT CDLKICDFGL ARVADPDHDH TGFLTEYVAT
201 RWYRAPEIML NSKGYTKSID IWSVGCILAE MLSNRPIFPG KHYLDQLNHI
251 LGILGSPSQE DLNCIINLKA RNYLLSLPHK NKVPWNRLFP NADSKALDLL
301 DKMLTFNPHK RIEVEQALAH PYLEQYYDPS DEPIAEAPFK FDMELDDLPK
351 EKLKELIFEE TARFQPGYRS
```

Phosphopeptide/s detected:

Ions detected	Corresponding sequence	Phosphorylation site/region
m/z 741.99 z = 3	VADPDHDHTGFLTEYVATR with 1 Phospho (STY)	The MS/MS spectra showed that the phosphorylation site is Y15.
m/z 768.65 z =3	VADPDHDHTGFLTEYVATR With 2 Phospho (STY)	From the MS/MS spectrum, it is likely that Y15 and T18 are the phosphorylation sites.

5. Results of 43

Protein coverage: 97%.

Matched peptides shown in **red**.

```
  1 GGLTMAHHHH HHMAAAAAAG PEMVRGQVFD VGPRYTNLSY IGEGAYGMVC
 51 SAYDNLNKVR VAIKKISPFE HQTYCQRTLR EIKILLRFRH ENIIGINDII
101 RAPTIEQMKD VYIVQDLMET DLYKLLKTQH LSNDHICYFL YQILRGLKYI
151 HSANVLHRDL KPSNLLLNTT CDLKICDFGL ARVADPDHDH TGFLTEYVAT
201 RWYRAPEIML NSKGYTKSID IWSVGCILAE MLSNRPIFPG KHYLDQLNHI
251 LGILGSPSQE DLNCIINLKA RNYLLSLPHK NKVPWNRLFP NADSKALDLL
301 DKMLTFNPHK RIEVEQALAH PYLEQYYDPS DEPIAEAPFK FDMELDDLPK
351 EKLKELIFEE TARFQPGYRS
```

Phosphopeptide/s detected:

Ions detected	Corresponding sequence	Phosphorylation site/region
i) m/z 741.99 z = 3 ii) m/z 741.99 z = 3	VADPDHDHTGFLTEYVATR with 1 Phospho (STY) and VADPDHDHTGFLTEYVATR with 1 Phospho (STY)	Two forms of the phosphopeptide were found: peptide phosphorylated on T18 (VADPDHDHTGFLTEYVATR) and Y15 (VADPDHDHTGFLTEYVATR) The signal corresponds to VADPDHDHTGFLTEYVATR is stronger than that of VADPDHDHTGFLTEYVATR
m/z 768.65 z =3	VADPDHDHTGFLTEYVATR With 2 Phospho (STY)	From the MS/MS spectrum, it is likely that Y15 and T18 are the phosphorylation sites.

1. Results of 44

Protein coverage: 95%. Matched peptides shown in **red**.

```
  1 GGLTMAHHHH HHMAAAAAAG PEMVRGQVFD VGPRYTNLSY IGEGAYGMVC
 51 SAYDNLNKVR VAIKKISPFE HQTYCQRTLR EIKILLRFRH ENIIGINDII
101 RAPTIEQMKD VYIVQDLMET DLYKLLKTQH LSNDHICYFL YQILRGLKYI
151 HSANVLHRDL KPSNLLLNTT CDLKICDFGL ARVADPDHDH TGFLTEYVAT
201 RWYRAPEIML NSKGYTKSID IWSVGCILAE MLSNRPIFPG KHYLDQLNHI
251 LGILGSPSQE DLNCIINLKA RNYLLSLPHK NKVPWNRLFP NADSKALDLL
301 DKMLTFNPHK RIEVEQALAH PYLEQYYDPS DEPIAEAPFK FDMELDDLPK
351 EKLKELIFEE TARFQPGYRS
```

Two phosphopeptides were detected:

Ions detected	Corresponding sequence	Phosphorylation site/region
i) m/z 741.99 z = 3 ii) m/z 741.99 z = 3	VADPDHDHTGFLTEYVATR with 1 Phospho (STY) and VADPDHDHTGFLTEYVATR with 1 Phospho (STY)	Two forms of the phosphopeptide were found: peptide phosphorylated on T18 (VADPDHDHTGFLTEYVATR) and Y15 (VADPDHDHTGFLTEYVATR)
m/z 768.65 z =3	VADPDHDHTGFLTEYVATR With 2 Phospho (STY)	The MS/MS spectrum showed that the phosphorylation sites are Y15 and T18.

2. Results of 45

Protein coverage: 95%.

Matched peptides shown in **red**.

```
  1 GGLTMAHHHH HHMAAAAAAG PEMVRGQVFD VGPRYTNLSY IGEGAYGMVC
 51 SAYDNLNKVR VAIKKISPFE HQTYCQRTLR EIKILLRFRH ENIIGINDII
101 RAPTIEQMKD VYIVQDLMET DLYKLLKTQH LSNDHICYFL YQILRGLKYI
151 HSANVLHRDL KPSNLLLNTT CDLKICDFGL ARVADPDHDH TGFLTEYVAT
201 RWYRAPEIML NSKGYTKSID IWSVGCILAE MLSNRPIFPG KHYLDQLNHI
251 LGILGSPSQE DLNCIINLKA RNYLLSLPHK NKVPWNRLFP NADSKALDLL
301 DKMLTFNPHK RIEVEQALAH PYLEQYYDPS DEPIAEAPFK FDMELDDLPK
351 EKLKELIFEE TARFQPGYRS
```

Two phosphopeptides were detected:

Ions detected	Corresponding sequence	Phosphorylation site/region
i) m/z 741.99 z = 3 ii) m/z 741.99 z = 3	VADPDHDHTGFLTEYVATR with 1 Phospho (STY) and VADPDHDHTGFLTEYVATR with 1 Phospho (STY)	Two forms of the phosphopeptide were found: peptide phosphorylated on T18 (VADPDHDHTGFLTEYVATR) and Y15 (VADPDHDHTGFLTEYVATR)
m/z 768.65 z =3	VADPDHDHTGFLTEYVATR With 2 Phospho (STY)	The MS/MS spectrum showed that the phosphorylation sites are Y15 and T18.

3. Results of 46

Protein coverage: 94%.

Matched peptides shown in **red**.

```
  1 GGLTMAHHHH HHMAAAAAAG PEMVRGQVFD VGPRYTNLSY IGEGAYGMVC
 51 SAYDNLNKVR VAIKKISPFE HQTYCQRTLR EIKILLRFRH ENIIGINDII
101 RAPTIEQMKD VYIVQDLMET DLYKLLKTQH LSNDHICYFL YQILRGLKYI
151 HSANVLHRDL KPSNLLLNTT CDLKICDFGL ARVADPDHDH TGFLTEYVAT
201 RWYRAPEIML NSKGYTKSID IWSVGCILAE MLSNRPIFPG KHYLDQLNHI
251 LGILGSPSQE DLNCIINLKA RNYLLSLPHK NKVPWNRLFP NADSKALDLL
301 DKMLTFNPHK RIEVEQALAH PYLEQYYDPS DEPIAEAPFK FDMELDDLPK

351 EKLKELIFEE TARFQPGYRS
```

One phosphopeptide was detected:

Ions detected	**Corresponding sequence**	**Phosphorylation site/region**
i) m/z 741.99 z = 3 ii) m/z 741.99 z = 3	VADPDHDHTGFLTEYVATR with 1 Phospho (STY) and VADPDHDHTGFLTEYVATR with 1 Phospho (STY)	Two forms of the phosphopeptide were found: peptide phosphorylated on T18 (VADPDHDHTGFLTEYVATR) and Y15 (VADPDHDHTGFLTEYVATR) .

4. Results of 47

Protein coverage: 97%.

Matched peptides shown in **red**.

```
  1 GGLTMAHHHH HHMAAAAAAG PEMVRGQVFD VGPRYTNLSY IGEGAYGMVC
 51 SAYDNLNKVR VAIKKISPFE HQTYCQRTLR EIKILLRFRH ENIIGINDII
101 RAPTIEQMKD VYIVQDLMET DLYKLLKTQH LSNDHICYFL YQILRGLKYI
151 HSANVLHRDL KPSNLLLNTT CDLKICDFGL ARVADPDHDH TGFLTEYVAT
201 RWYRAPEIML NSKGYTKSID IWSVGCILAE MLSNRPIFPG KHYLDQLNHI
251 LGILGSPSQE DLNCIINLKA RNYLLSLPHK NKVPWNRLFP NADSKALDLL
301 DKMLTFNPHK RIEVEQALAH PYLEQYYDPS DEPIAEAPFK FDMELDDLPK
351 EKLKELIFEE TARFQPGYRS
```

One phosphopeptide was found

Ions detected	**Corresponding sequence**	**Phosphorylation site/region**
i) m/z 741.99 z = 3 ii) m/z 741.99 z = 3	VADPDHDHTGFLTEYVATR with 1 Phospho (STY) and VADPDHDHTGFLTEYVATR with 1 Phospho (STY)	Two forms of the phosphopeptide were found: peptide phosphorylated on T18 (VADPDHDHTGFLTEYVATR) and Y15 (VADPDHDHTGFLTEYVATR) .

6. Results of 51

Protein coverage: 97%.

Matched peptides shown in **red**.

```
  1 GGLTMAHHHH HHMAAAAAAG PEMVRGQVFD VGPRYTNLSY IGEGAYGMVC
 51 SAYDNLNKVR VAIKKISPFE HQTYCQRTLR EIKILLRFRH ENIIGINDII
101 RAPTIEQMKD VYIVQDLMET DLYKLLKTQH LSNDHICYFL YQILRGLKYI
151 HSANVLHRDL KPSNLLLNTT CDLKICDFGL ARVADPDHDH TGFLTEYVAT
201 RWYRAPEIML NSKGYTKSID IWSVGCILAE MLSNRPIFPG KHYLDQLNHI
251 LGILGSPSQE DLNCIINLKA RNYLLSLPHK NKVPWNRLFP NADSKALDLL
301 DKMLTFNPHK RIEVEQALAH PYLEQYYDPS DEPIAEAPFK FDMELDDLPK
351 EKLKELIFEE TARFQPGYRS
```

Phosphopeptide/s detected:

Ions detected	Corresponding sequence	Phosphorylation site/region
i) m/z 741.99 z = 3 ii) m/z 741.99 z = 3 iii) m/z 741.99 z = 3	VADPDHDHTGFLTEYVATR with 1 Phospho (STY) and VADPDHDHTGFLTEYVATR with 1 Phospho (STY) and VADPDHDHTGFLTEYVATR with 1 Phospho (STY)	Three forms of the phosphopeptide were found: peptide phosphorylated on T18 (VADPDHDHTGFLTEYVATR) and Y15 (VADPDHDHTGFLTEYVATR) and T13 (VADPDHDHTGFLTEYVATR) The signal corresponds to VADPDHDHTGFLTEYVATR is stronger than those of VADPDHDHTGFLTEYVATR and VADPDHDHTGFLTEYVATR.
i) m/z 768.65 z =3 ii) m/z 768.65 z =3	VADPDHDHTGFLTEYVATR With 2 Phospho (STY) VADPDHDHTGFLTEYVATR With 2 Phospho (STY)	i) Y15 and T18 are the phosphorylation sites. ii) T13 and Y15 are the phosphorylation sites.
m/z 870.44 z = 2	LFPNADSKALDLLDK with 1 Phospho (STY)	S7 is the phosphorylation site.
m/z 1148.2	IEVEQALAHPYLEQYYDPSDEPIAEAPFK	Y15 is the phosphorylation site.

7. Results of 52

Protein coverage: 98%.

Matched peptides shown in **red**.

```
  1 GGLTMAHHHH HHMAAAAAAG PEMVRGQVFD VGPRYTNLSY IGEGAYGMVC
 51 SAYDNLNKVR VAIKKISPFE HQTYCQRTLR EIKILLRFRH ENIIGINDII
101 RAPTIEQMKD VYIVQDLMET DLYKLLKTQH LSNDHICYFL YQILRGLKYI
151 HSANVLHRDL KPSNLLLNTT CDLKICDFGL ARVADPDHDH TGFLTEYVAT
201 RWYRAPEIML NSKGYTKSID IWSVGCILAE MLSNRPIFPG KHYLDQLNHI
251 LGILGSPSQE DLNCIINLKA RNYLLSLPHK NKVPWNRLFP NADSKALDLL
301 DKMLTFNPHK RIEVEQALAH PYLEQYYDPS DEPIAEAPFK FDMELDDLPK
351 EKLKELIFEE TARFQPGYRS
```

Phosphopeptide/s detected:

Ions detected	Corresponding sequence	Phosphorylation site/region
i) m/z 741.99 z = 3 ii) m/z 741.99 z = 3 iii) m/z 741.99 z = 3	VADPDHDHTGFLTEYVATR with 1 Phospho (STY) and VADPDHDHTGFLTEYVATR with 1 Phospho (STY) and VADPDHDHTGFLTEYVATR with 1 Phospho (STY)	Three forms of the phosphopeptide were found: peptide phosphorylated on T18 (VADPDHDHTGFLTEYVATR) and Y15 (VADPDHDHTGFLTEYVATR) and T13 (VADPDHDHTGFLTEYVATR) The signal corresponds to VADPDHDHTGFLTEYVATR is stronger than those of VADPDHDHTGFLTEYVATR and VADPDHDHTGFLTEYVATR.
m/z 768.65 z =3	VADPDHDHTGFLTEYVATR With 2 Phospho (STY)	Y15 and T18 are the phosphorylation sites.

8. Results of 53

Protein coverage: 97%.

Matched peptides shown in **red**.

```
  1 GGLTMAHHHH HHMAAAAAAG PEMVRGQVFD VGPRYTNLSY IGEGAYGMVC
 51 SAYDNLNKVR VAIKKISPFE HQTYCQRTLR EIKILLRFRH ENIIGINDII
101 RAPTIEQMKD VYIVQDLMET DLYKLLKTQH LSNDHICYFL YQILRGLKYI
151 HSANVLHRDL KPSNLLLNTT CDLKICDFGL ARVADPDHDH TGFLTEYVAT
201 RWYRAPEIML NSKGYTKSID IWSVGCILAE MLSNRPIFPG KHYLDQLNHI
251 LGILGSPSQE DLNCIINLKA RNYLLSLPHK NKVPWNRLFP NADSKALDLL
301 DKMLTFNPHK RIEVEQALAH PYLEQYYDPS DEPIAEAPFK FDMELDDLPK
351 EKLKELIFEE TARFQPGYRS
```

Phosphopeptide/s detected:

Ions detected	Corresponding sequence	Phosphorylation site/region
m/z 741.99 z = 3	VADPDHDHTGFLTEYVATR with 1 Phospho (STY)	Y15 is the phosphorylation site.
m/z 768.65 z =3	VADPDHDHTGFLTEYVATR With 2 Phospho (STY)	Y15 and T18 are the phosphorylation sites.

9. Results of 54

Protein coverage: 97%.

Matched peptides shown in **red**.

```
  1 GGLTMAHHHH HHMAAAAAAG PEMVRGQVFD VGPRYTNLSY IGEGAYGMVC
 51 SAYDNLNKVR VAIKKISPFE HQTYCQRTLR EIKILLRFRH ENIIGINDII
101 RAPTIEQMKD VYIVQDLMET DLYKLLKTQH LSNDHICYFL YQILRGLKYI
151 HSANVLHRDL KPSNLLLNTT CDLKICDFGL ARVADPDHDH TGFLTEYVAT
201 RWYRAPEIML NSKGYTKSID IWSVGCILAE MLSNRPIFPG KHYLDQLNHI
251 LGILGSPSQE DLNCIINLKA RNYLLSLPHK NKVPWNRLFP NADSKALDLL
301 DKMLTFNPHK RIEVEQALAH PYLEQYYDPS DEPIAEAPFK FDMELDDLPK
351 EKLKELIFEE TARFQPGYRS
```

Phosphopeptide/s detected:

Ions detected	Corresponding sequence	Phosphorylation site/region
m/z 741.99 z = 3	VADPDHDHTGFLTEYVATR with 1 Phospho (STY)	Y15 is the phosphorylation site.
m/z 768.65 z =3	VADPDHDHTGFLTEYVATR With 2 Phospho (STY)	Y15 and T13 are the phosphorylation sites.

10. Results of 55

Protein coverage: 97%.

Matched peptides shown in **red**.

```
  1 GGLTMAHHHH HHMAAAAAAG PEMVRGQVFD VGPRYTNLSY IGEGAYGMVC
 51 SAYDNLNKVR VAIKKISPFE HQTYCQRTLR EIKILLRFRH ENIIGINDII
101 RAPTIEQMKD VYIVQDLMET DLYKLLKTQH LSNDHICYFL YQILRGLKYI
151 HSANVLHRDL KPSNLLLNTT CDLKICDFGL ARVADPDHDH TGFLTEYVAT
201 RWYRAPEIML NSKGYTKSID IWSVGCILAE MLSNRPIFPG KHYLDQLNHI
251 LGILGSPSQE DLNCIINLKA RNYLLSLPHK NKVPWNRLFP NADSKALDLL
301 DKMLTFNPHK RIEVEQALAH PYLEQYYDPS DEPIAEAPFK FDMELDDLPK
351 EKLKELIFEE TARFQPGYRS
```

Phosphopeptide/s detected:

Ions detected	Corresponding sequence	Phosphorylation site/region
m/z 741.99 z = 3	VADPDHDHTGFLTEYVATR with 1 Phospho (STY)	Y15 is the phosphorylation site.
m/z 768.65 z =3	VADPDHDHTGFLTEYVATR With 2 Phospho (STY)	Can't assign the phosphorylation sites.

11. Results of 56

Protein coverage: 97%.

Matched peptides shown in **red**.

```
  1 GGLTMAHHHH HHMAAAAAAG PEMVRGQVFD VGPRYTNLSY IGEGAYGMVC
 51 SAYDNLNKVR VAIKKISPFE HQTYCQRTLR EIKILLRFRH ENIIGINDII
101 RAPTIEQMKD VYIVQDLMET DLYKLLKTQH LSNDHICYFL YQILRGLKYI
151 HSANVLHRDL KPSNLLLNTT CDLKICDFGL ARVADPDHDH TGFLTEYVAT
201 RWYRAPEIML NSKGYTKSID IWSVGCILAE MLSNRPIFPG KHYLDQLNHI
251 LGILGSPSQE DLNCIINLKA RNYLLSLPHK NKVPWNRLFP NADSKALDLL
301 DKMLTFNPHK RIEVEQALAH PYLEQYYDPS DEPIAEAPFK FDMELDDLPK
351 EKLKELIFEE TARFQPGYRS
```

Phosphopeptide/s detected:

Ions detected	Corresponding sequence	Phosphorylation site/region
m/z 741.99 z = 3	VADPDHDHTGFLTEYVATR with 1 Phospho (STY)	Y15 is the phosphorylation site.
m/z 768.65 z =3	VADPDHDHTGFLTEYVATR With 2 Phospho (STY)	Y15 and T18 are the phosphorylation sites.

5. Results of 57

Protein coverage: 97%.

Matched peptides shown in **red**.

```
  1 GGLTMAHHHH HHMAAAAAAG PEMVRGQVFD VGPRYTNLSY IGEGAYGMVC
 51 SAYDNLNKVR VAIKKISPFE HQTYCQRTLR EIKILLRFRH ENIIGINDII
101 RAPTIEQMKD VYIVQDLMET DLYKLLKTQH LSNDHICYFL YQILRGLKYI
151 HSANVLHRDL KPSNLLLNTT CDLKICDFGL ARVADPDHDH TGFLTEYVAT
201 RWYRAPEIML NSKGYTKSID IWSVGCILAE MLSNRPIFPG KHYLDQLNHI
251 LGILGSPSQE DLNCIINLKA RNYLLSLPHK NKVPWNRLFP NADSKALDLL
301 DKMLTFNPHK RIEVEQALAH PYLEQYYDPS DEPIAEAPFK FDMELDDLPK
351 EKLKELIFEE TARFQPGYRS
```

One phosphopeptide was found

Ions detected	Corresponding sequence	Phosphorylation site/region
i) m/z 741.99 z = 3	VADPDHDHTGFLTEYVATR with 1 Phospho (STY)	MS/MS spectrum showed that Y15 is the phosphorylation site. .

6. Results of 58

Protein coverage: 97%.

Matched peptides shown in **red**.

```
  1 GGLTMAHHHH HHMAAAAAAG PEMVRGQVFD VGPRYTNLSY IGEGAYGMVC
 51 SAYDNLNKVR VAIKKISPFE HQTYCQRTLR EIKILLRFRH ENIIGINDII
101 RAPTIEQMKD VYIVQDLMET DLYKLLKTQH LSNDHICYFL YQILRGLKYI
151 HSANVLHRDL KPSNLLLNTT CDLKICDFGL ARVADPDHDH TGFLTEYVAT
201 RWYRAPEIML NSKGYTKSID IWSVGCILAE MLSNRPIFPG KHYLDQLNHI
251 LGILGSPSQE DLNCIINLKA RNYLLSLPHK NKVPWNRLFP NADSKALDLL
301 DKMLTFNPHK RIEVEQALAH PYLEQYYDPS DEPIAEAPFK FDMELDDLPK
351 EKLKELIFEE TARFQPGYRS
```

One phosphopeptide was found

Ions detected	Corresponding sequence	Phosphorylation site/region
i) m/z 741.99 z = 3	VADPDHDHTGFLTEYVATR with 1 Phospho (STY)	MS/MS spectrum showed that Y15 is the phosphorylation site. .

Bibliography

Adachi, M., M. Fukuda, and E. Nishida. 1999. Two co-existing mechanisms for nuclear import of MAP kinase: passive diffusion of a monomer and active transport of a dimer. *The EMBO journal*. 18:5347-5358.

Aoki, K., K. Takahashi, K. Kaizu, and M. Matsuda. 2013. A quantitative model of ERK MAP kinase phosphorylation in crowded media. *Scientific reports*. 3:1541.

Arnette, D., T.B. Gibson, M.C. Lawrence, B. January, S. Khoo, K. McGlynn, C.A. Vanderbilt, and M.H. Cobb. 2003. Regulation of ERK1 and ERK2 by glucose and peptide hormones in pancreatic beta cells. *The Journal of biological chemistry*. 278:32517-32525.

Barr, D., T. Oashi, K. Burkhard, S. Lucius, R. Samadani, J. Zhang, P. Shapiro, A.D. MacKerell, and A. van der Vaart. 2011. Importance of domain closure for the autoactivation of ERK2. *Biochemistry*. 50:8038-8048.

Blaukat, A., A. Barac, M.J. Cross, S. Offermanns, and I. Dikic. 2000. G protein-coupled receptor-mediated mitogen-activated protein kinase activation through cooperation of Galpha(q) and Galpha(i) signals. *Molecular and cellular biology*. 20:6837-6848.

Canagarajah, B.J., A. Khokhlatchev, M.H. Cobb, and E.J. Goldsmith. 1997. Activation mechanism of the MAP kinase ERK2 by dual phosphorylation. *Cell*. 90:859-869.

Chen, R.H., C. Sarnecki, and J. Blenis. 1992. Nuclear localization and regulation of erk- and rsk-encoded protein kinases. *Molecular and cellular biology*. 12:915-927.

Conklin, B.R., and H.R. Bourne. 1993. Structural elements of G alpha subunits that interact with G beta gamma, receptors, and effectors. *Cell*. 73:631-641.

Crespo, P., and J.S. Gutkind. 2004. Activation of MAPKs by G protein-coupled receptors. *Methods Mol Biol*. 250:203-210.

Crespo, P., N. Xu, W.F. Simonds, and J.S. Gutkind. 1994. Ras-dependent activation of MAP kinase pathway mediated by G-protein beta gamma subunits. *Nature*. 369:418-420.

Emrick, M.A., A.N. Hoofnagle, A.S. Miller, L.F. Ten Eyck, and N.G. Ahn. 2001. Constitutive activation of extracellular signal-regulated kinase 2 by synergistic point mutations. *The Journal of biological chemistry*. 276:46469-46479.

Emrick, M.A., T. Lee, P.J. Starkey, M.C. Mumby, K.A. Resing, and N.G. Ahn. 2006. The gatekeeper residue controls autoactivation of ERK2 via a pathway of intramolecular connectivity. *Proceedings of the National Academy of Sciences of the United States of America*. 103:18101-18106.

Evdokimov, A., I. Voznesensky, K. Fennell, M. Anderson, J.F. Smith, and D.A. Fisher. 2009. New kinase regulation mechanism found in HipBA: a bacterial persistence switch. *Acta crystallographica. Section D, Biological crystallography*. 65:875-879.

Eyers, P.A., M. Craxton, N. Morrice, P. Cohen, and M. Goedert. 1998. Conversion of SB 203580-insensitive MAP kinase family members to drug-sensitive forms by a single amino-acid substitution. *Chemistry & biology*. 5:321-328.

Feng, X., T. Sun, Y. Bei, S. Ding, W. Zheng, Y. Lu, and P. Shen. 2013. S-nitrosylation of ERK inhibits ERK phosphorylation and induces apoptosis. *Scientific reports*. 3:1814.

Ferrell, J.E., Jr., and R.R. Bhatt. 1997. Mechanistic studies of the dual phosphorylation of mitogen-activated protein kinase. *The Journal of biological chemistry*. 272:19008-19016.

Fox, T., J.T. Coll, X. Xie, P.J. Ford, U.A. Germann, M.D. Porter, S. Pazhanisamy, M.A. Fleming, V. Galullo, M.S. Su, and K.P. Wilson. 1998. A single amino acid substitution makes ERK2 susceptible to pyridinyl imidazole inhibitors of p38 MAP kinase. *Protein science : a publication of the Protein Society*. 7:2249-2255.

German, M.S., L.G. Moss, and W.J. Rutter. 1990. Regulation of insulin gene expression by glucose and calcium in transfected primary islet cultures. *The Journal of biological chemistry*. 265:22063-22066.

German, M.S., and J. Wang. 1994. The insulin gene contains multiple transcriptional elements that respond to glucose. *Molecular and cellular biology*. 14:4067-4075.

Gille, H., M. Kortenjann, O. Thomae, C. Moomaw, C. Slaughter, M.H. Cobb, and P.E. Shaw. 1995. ERK phosphorylation potentiates Elk-1-mediated ternary complex formation and transactivation. *The EMBO journal*. 14:951-962.

Goldsmith, E.J., R. Akella, X. Min, T. Zhou, and J.M. Humphreys. 2007. Substrate and docking interactions in serine/threonine protein kinases. *Chemical reviews*. 107:5065-5081.

Gum, R.J., M.M. McLaughlin, S. Kumar, Z. Wang, M.J. Bower, J.C. Lee, J.L. Adams, G.P. Livi, E.J. Goldsmith, and P.R. Young. 1998. Acquisition of sensitivity of stress-activated protein kinases to the p38 inhibitor, SB 203580, by alteration of one or more amino acids within the ATP binding pocket. *The Journal of biological chemistry*. 273:15605-15610.

Hanks, S.K., and T. Hunter. 1995. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*. 9:576-596.

Hay, C.W., and K. Docherty. 2006. Comparative analysis of insulin gene promoters: implications for diabetes research. *Diabetes*. 55:3201-3213.

Heise, C.J., and M.H. Cobb. 2006. Expression and characterization of MAP kinases in bacteria. *Methods (San Diego, Calif.)*. 40:209-212.

Hu, S., Z. Xie, A. Onishi, X. Yu, L. Jiang, J. Lin, H.S. Rho, C. Woodard, H. Wang, J.S. Jeong, S. Long, X. He, H. Wade, S. Blackshaw, J. Qian, and H. Zhu. 2009. Profiling the human protein-DNA interactome reveals ERK2 as a transcriptional repressor of interferon signaling. *Cell*. 139:610-622.

Khokhlatchev, A.V., B. Canagarajah, J. Wilsbacher, M. Robinson, M. Atkinson, E. Goldsmith, and M.H. Cobb. 1998. Phosphorylation of the MAP kinase ERK2 promotes its homodimerization and nuclear translocation. *Cell*. 93:605-615.

Khoo, S., T.B. Gibson, D. Arnette, M. Lawrence, B. January, K. McGlynn, C.A. Vanderbilt, S.C. Griffen, M.S. German, and M.H. Cobb. 2004. MAP kinases and their roles in pancreatic beta-cells. *Cell biochemistry and biophysics*. 40:191-200.

Khoo, S., S.C. Griffen, Y. Xia, R.J. Baer, M.S. German, and M.H. Cobb. 2003. Regulation of insulin gene transcription by ERK1 and ERK2 in pancreatic beta cells. *The Journal of biological chemistry*. 278:32969-32977.

Kim, M.O., S.H. Kim, Y.Y. Cho, J. Nadas, C.H. Jeong, K. Yao, D.J. Kim, D.H. Yu, Y.S. Keum, K.Y. Lee, Z. Huang, A.M. Bode, and Z. Dong. 2012. ERK1 and ERK2 regulate embryonic stem cell self-renewal through phosphorylation of Klf4. *Nature structural & molecular biology*. 19:283-290.

Knighton, D.R., J.H. Zheng, L.F. Ten Eyck, V.A. Ashford, N.H. Xuong, S.S. Taylor, and J.M. Sowadski. 1991. Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science*. 253:407-414.

Kolch, W., G. Heidecker, G. Kochs, R. Hummel, H. Vahidi, H. Mischak, G. Finkenzeller, D. Marme, and U.R. Rapp. 1993. Protein kinase C alpha activates RAF-1 by direct phosphorylation. *Nature*. 364:249-252.

Lawrence, M.C., K. McGlynn, B.H. Park, and M.H. Cobb. 2005. ERK1/2-dependent activation of transcription factors required for acute and chronic effects of glucose on the insulin gene promoter. *The Journal of biological chemistry*. 280:26751-26759.

Lawrence, M.C., K. McGlynn, C. Shao, L. Duan, B. Naziruddin, M.F. Levy, and M.H. Cobb. 2008. Chromatin-bound mitogen-activated protein kinases transmit dynamic signals in transcription complexes in beta-cells. *Proceedings of the National Academy of Sciences of the United States of America*. 105:13315-13320.

Lee, T., A.N. Hoofnagle, K.A. Resing, and N.G. Ahn. 2005. Hydrogen exchange solvent protection by an ATP analogue reveals conformational changes in ERK2 upon activation. *Journal of molecular biology*. 353:600-612.

Levin-Salomon, V., K. Kogan, N.G. Ahn, O. Livnah, and D. Engelberg. 2008. Isolation of intrinsically active (MEK-independent) variants of the ERK family of mitogen-activated protein (MAP) kinases. *The Journal of biological chemistry*. 283:34500-34510.

Levin-Salomon, V., I. Maayan, L. Avrahami-Moyal, I. Marbach, O. Livnah, and D. Engelberg. 2009. When expressed in yeast, mammalian mitogen-activated protein kinases lose proper regulation and become spontaneously phosphorylated. *The Biochemical journal*. 417:331-340.

Lorenz, K., J.P. Schmitt, E.M. Schmitteckert, and M.J. Lohse. 2009. A new type of ERK1/2 autophosphorylation causes cardiac hypertrophy. *Nature medicine*. 15:75-83.

Madak-Erdogan, Z., M. Lupien, F. Stossi, M. Brown, and B.S. Katzenellenbogen. 2011. Genomic collaboration of estrogen receptor alpha and extracellular signal-regulated kinase 2 in regulating gene and proliferation programs. *Molecular and cellular biology*. 31:226-236.

Manning, G., G.D. Plowman, T. Hunter, and S. Sudarsanam. 2002a. Evolution of protein kinase signaling from yeast to man. *Trends in biochemical sciences*. 27:514-520.

Manning, G., D.B. Whyte, R. Martinez, T. Hunter, and S. Sudarsanam. 2002b. The protein kinase complement of the human genome. *Science*. 298:1912-1934.

Mansour, S.J., J.M. Candia, K.K. Gloor, and N.G. Ahn. 1996. Constitutively active mitogen-activated protein kinase kinase 1 (MAPKK1) and MAPKK2 mediate similar transcriptional and morphological responses. *Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research*. 7:243-250.

Mumby, S.M., P.J. Casey, A.G. Gilman, S. Gutowski, and P.C. Sternweis. 1990. G protein gamma subunits contain a 20-carbon isoprenoid. *Proceedings of the National Academy of Sciences of the United States of America*. 87:5873-5877.

Odagiri, H., J. Wang, and M.S. German. 1996. Function of the human insulin promoter in primary cultured islet cells. *The Journal of biological chemistry*. 271:1909-1915.

Plotnikov, A., D. Chuderland, Y. Karamansha, O. Livnah, and R. Seger. 2011. Nuclear extracellular signal-regulated kinase 1 and 2 translocation is mediated by casein kinase 2 and accelerated by autophosphorylation. *Molecular and cellular biology*. 31:3515-3530.

Prowse, C.N., J.C. Hagopian, M.H. Cobb, N.G. Ahn, and J. Lew. 2000. Catalytic reaction pathway for the mitogen-activated protein kinase ERK2. *Biochemistry*. 39:6258-6266.

Prowse, C.N., and J. Lew. 2001. Mechanism of activation of ERK2 by dual phosphorylation. *The Journal of biological chemistry*. 276:99-103.

Qi, M., and E.A. Elion. 2005. MAP kinase pathways. *Journal of cell science*. 118:3569-3572.

Ring, A.Y., K.M Sours, Lee, T. Ahn, G. 2011. Distinct patterns of activation-dependent changes in conformational mobility between ERK1 and ERK2. *International Journal of Mass Spectrometry*. 302:101-109.

Robbins, D.J., E. Zhen, H. Owaki, C.A. Vanderbilt, D. Ebert, T.D. Geppert, and M.H. Cobb. 1993. Regulation and properties of extracellular signal-regulated protein kinases 1 and 2 in vitro. *The Journal of biological chemistry*. 268:5097-5106.

Robinson, F., Whitehurst, AW., Raman, M., Cobb, MH. 2002. Identification of Novel Point Mutations in ERK2 That Selectively Disrupt Binding to MEK1. *Journal Biological Chemistry*. 277:14044-14852.

Robinson, M.J., P.C. Harkins, J. Zhang, R. Baer, J.W. Haycock, M.H. Cobb, and E.J. Goldsmith. 1996. Mutation of position 52 in ERK2 creates a nonproductive binding mode for adenosine 5'-triphosphate. *Biochemistry*. 35:5641-5646.

Rossomando, A.J., J. Wu, H. Michel, J. Shabanowitz, D.F. Hunt, M.J. Weber, and T.W. Sturgill. 1992. Identification of Tyr-185 as the site of tyrosine autophosphorylation of recombinant mitogen-activated protein kinase p42mapk. *Proceedings of the National Academy of Sciences of the United States of America*. 89:5779-5783.

Ruppert, C., K. Deiss, S. Herrmann, M. Vidal, M. Oezkur, A. Gorski, F. Weidemann, M.J. Lohse, and K. Lorenz. 2013. Interference with ERK(Thr188) phosphorylation impairs pathological but not physiological cardiac hypertrophy. *Proceedings of the National Academy of Sciences of the United States of America*. 110:7440-7445.

Schonwasser, D.C., R.M. Marais, C.J. Marshall, and P.J. Parker. 1998. Activation of the mitogen-activated protein kinase/extracellular signal-regulated kinase pathway by conventional, novel, and atypical protein kinase C isotypes. *Molecular and cellular biology*. 18:790-798.

Schumacher, M.A., K.M. Piro, W. Xu, S. Hansen, K. Lewis, and R.G. Brennan. 2009. Molecular mechanisms of HipA-mediated multidrug tolerance and its neutralization by HipB. *Science*. 323:396-401.

Shokat, K., and M. Velleca. 2002. Novel chemical genetic approaches to the discovery of signal transduction inhibitors. *Drug discovery today*. 7:872-879.

Tanoue, T., M. Adachi, T. Moriguchi, and E. Nishida. 2000. A conserved docking motif in MAP kinases common to substrates, activators and regulators. *Nature cell biology*. 2:110-116.

Tanoue, T., T. Yamamoto, and E. Nishida. 2002. Modular structure of a docking surface on MAPK phosphatases. *The Journal of biological chemistry*. 277:22942-22949.

Thorsness, P.E., and D.E. Koshland, Jr. 1987. Inactivation of isocitrate dehydrogenase by phosphorylation is mediated by the negative charge of the phosphate. *The Journal of biological chemistry*. 262:10422-10425.

Timm, T., X.Y. Li, J. Biernat, J. Jiao, E. Mandelkow, J. Vandekerckhove, and E.M. Mandelkow. 2003. MARKK, a Ste20-like kinase, activates the polarity-inducing kinase MARK/PAR-1. *The EMBO journal*. 22:5090-5101.

Wall, M.A., D.E. Coleman, E. Lee, J.A. Iniguez-Lluhi, B.A. Posner, A.G. Gilman, and S.R. Sprang. 1995. The structure of the G protein heterotrimer Gi alpha 1 beta 1 gamma 2. *Cell*. 83:1047-1058.

Wang, W., Y. Yang, Y. Gao, Q. Xu, F. Wang, S. Zhu, W. Old, K. Resing, N. Ahn, M. Lei, and X. Liu. 2009. Structural and mechanistic insights into Mps1 kinase activation. *Journal of cellular and molecular medicine*. 13:1679-1694.

Whelan, J., D. Poon, P.A. Weil, and R. Stein. 1989. Pancreatic beta-cell-type-specific expression of the rat insulin II gene is controlled by positive and negative cellular transcriptional elements. *Molecular and cellular biology*. 9:3253-3259.

Whitehurst, A., M.H. Cobb, and M.A. White. 2004a. Stimulus-coupled spatial restriction of extracellular signal-regulated kinase 1/2 activity contributes to the specificity of signal-response pathways. *Molecular and cellular biology*. 24:10145-10150.

Whitehurst, A.W., F.L. Robinson, M.S. Moore, and M.H. Cobb. 2004b. The death effector domain protein PEA-15 prevents nuclear entry of ERK2 by inhibiting required interactions. *The Journal of biological chemistry*. 279:12840-12847.

Whitehurst, A.W., J.L. Wilsbacher, Y. You, K. Luby-Phelps, M.S. Moore, and M.H. Cobb. 2002. ERK2 enters the nucleus by a carrier-independent mechanism. *Proceedings of the National Academy of Sciences of the United States of America*. 99:7496-7501.

Wilsbacher, J.L., Y.C. Juang, A.V. Khokhlatchev, E. Gallagher, D. Binns, E.J. Goldsmith, and M.H. Cobb. 2006. Characterization of mitogen-activated protein kinase (MAPK) dimers. *Biochemistry*. 45:13175-13182.

Wilson, K.P., M.J. Fitzgibbon, P.R. Caron, J.P. Griffith, W. Chen, P.G. McCaffrey, S.P. Chambers, and M.S. Su. 1996. Crystal structure of p38 mitogen-activated protein kinase. *The Journal of biological chemistry*. 271:27696-27700.

Xu, B., J.L. Wilsbacher, T. Collisson, and M.H. Cobb. 1999. The N-terminal ERK-binding site of MEK1 is required for efficient feedback phosphorylation by ERK2 in vitro and ERK activation in vivo. *The Journal of biological chemistry*. 274:34029-34035.

Yuan, Z.L., Y.J. Guan, L. Wang, W. Wei, A.B. Kane, and Y.E. Chin. 2004. Central role of the threonine residue within the p+1 loop of receptor tyrosine kinase in STAT3 constitutive phosphorylation in metastatic cancer cells. *Molecular and cellular biology*. 24:9390-9400.

Zhang, C., D.M. Kenski, J.L. Paulson, A. Bonshtien, G. Sessa, J.V. Cross, D.J. Templeton, and K.M. Shokat. 2005. A second-site suppressor strategy for chemical genetic analysis of diverse protein kinases. *Nature methods*. 2:435-441.

Zhang, F., A. Strand, D. Robbins, M.H. Cobb, and E.J. Goldsmith. 1994. Atomic structure of the MAP kinase ERK2 at 2.3 A resolution. *Nature*. 367:704-711.

Zhang, X., J. Gureasko, K. Shen, P.A. Cole, and J. Kuriyan. 2006. An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. *Cell*. 125:1137-1149.