Key words: Phosphorylation, Prokaryote, Archaea, and Biochemistry.

Protein O-Kinases in the Archaeon *Sulfolobus solfataricus*

Brian Howard Lower

A Dissertation Submitted to the Faculty of the Virginia Polytechnic Institute and State University in Partial Fulfillment of the Requirements for the Degree of:

Doctor of Philosophy

in

Biochemistry

Peter J. Kennelly, Chair

Timothy J. Larson

Shirley Luckhart

David L. Popham

Malcolm Potts

July 23, 2001

Blacksburg, Virginia

Protein O-Kinases in the Archaeon *Sulfolobus solfataricus*

Brian Howard Lower

ABSTRACT

For many years, it has been understood that protein phosphorylation-dephosphorylation constitutes one of the most ubiquitous mechanisms for controlling the functional properties of proteins. Although originally believed to be a eukaryotic phenomenon, protein phosphorylation is now known to occur in all three domains of life *Eukarya*, *Bacteria*, and *Archaea*. Very little is known, however, concerning the origins and evolution of protein phosphorylation-dephosphorylation. Knowledge of the structure and properties of the protein kinases resident in the members of the *Archaea* represents a key piece of this puzzle.

The extreme acidothermophilic archaeon, *Sulfolobus solfataricus*, exhibits a membrane-associated protein kinase activity. Solubilization of the kinase activity requires the presence of detergent such as Triton X-100 or octyl glucoside, indicating its activity reside in an integral membrane protein. This protein kinase utilizes purine nucleotides as phosphoryl donors *in vitro* with a requirement for a divalent metal ion cofactor, favoring Mn^{+2}. A preference for NTPs over NDPs and for adenyl nucleotides over the analogous guanyl nucleotides was observed. The enzyme appears to be a glycoprotein that displays catalytic activity on SDS-PAGE corresponding to a molecular mass of ≈67 kDa, as well as an apparent molecular mass of ≈125 kDa on a gel filtration column. Challenged with several exogenous substrates revealed the protein kinase to be relatively selective. Only casein, reduced carboxyamidomethylated and maleylated lysozyme (RCM lysozyme), histone H4 proved, and a peptide modeled after myosin light chains (KKRAARATSNVFA) were phosphorylated to appreciable levels *in vitro*. All of the aforementioned substrates were phosphorylated on threonine, while histone H4 was phosphorylated on serine as well. When the phosphoacceptor threonine in the MLC peptide was substituted with serine an appreciable decrease in phosphorylation was noted. The protein kinase underwent autophosphorylation on threonine and was relatively insensitive to several known "eukaryotic" protein kinase inhibitors.

Primary sequence motifs based on known conserved subdomains of eukaryotic protein kinases were used to search the genome of *S. solfataricus* for eukaryotic-like protein kinase sequences. Six hypothetical proteins were identified from *S. solfataricus* whose primary sequence exhibited noticeable similarities to eukaryotic protein kinases. The hypothetical protein encoded

by *ORF sso0197* contained 7 putative subdomains, *ORFs sso0433, sso2291*, *sso2387,* and *sso3207* contained 8 putative subdomains, and *ORF sso3182*, contained 9 putative subdomains of the 12 characteristically conserved subdomains found within eukaryotic protein kinases.

ORF sso2387 was cloned and expressed in *Escherichia coli*. The expressed protein, SsPK2, was solubilized from inclusion bodies using 5 M urea. SsPK2 was able to phosphorylate casein, BSA, RCM lysozyme, and mixed histones *in vitro*. Phosphoamino acid analysis of casein, BSA, and mixed histones revealed that they were all phosphorylated on serine. SsPK2 underwent autophosphorylation on serine at elevated temperature using both purine nucleotide triphosphates as phosphoryl donors *in vitro*, but exhibited a noticeable preference for ATP. Autophosphorylate of SsPK2 also occurred at elevated temperature using a variety of divalent metals cofactors in order of $Mn^{2+} > Mg^{2+} >> Ca^{2+} \approx Zn^{2+}$. Polycations such as polyLys stimulated the phosphorylation of exogenous substrates while polyanions such as poly(Glu:Tyr) were shown to inhibit the phosphorylation of exogenous substrates. Of the "eukaryotic" protein kinases inhibitors tested, only tamoxifen had any noticeable effect of the catalytic activity of SsPK2 towards itself and exogenous substrates. A truncated form of SsPK2 containing the perceived catalytic domain also exhibited protein kinase activity towards itself and exogenous substrates. The observed protein kinase activity for SsPK2trunk was similar to that observed for SsPK2.

Proteins from the membrane fraction of *S. solfataricus* subject to phosphorylation *in vitro* on serine or threonine residues were identified using MALDI-MS / peptide fingerprinting techniques. Nine phosphoproteins were assigned a tentative identification using the ProFound protein search engine from Rockefeller University. The identity of two of nine phosphoproteins, a translational endoplasmic reticulum ATPase and an ≈ 42 kDa hypothetical protein, were determined with a relatively high degree of confidence. Collectively the results suggested MALDI-MS peptide mapping coupled with [^{32}P] labeling *in vivo* will have a tremendous potential for mapping out a major portion of the phosphoproteome of *S. solfataricus*.

ACKNOWLEDGEMENTS

I would like to thank my advisor Peter J. Kennelly for his guidance and support during my doctoral studies, and for allowing me the freedom to pursue my own ideas and interests. I would especially like to thank him for tolerating my peculiar work hours as this allowed me the opportunity to help raise my girls while my wife earned her Masters degree. In addition, I would like to thank him for all the hours he spent correcting my poor grammar and English, "he right done a good job of it!" I would also like to thank the members of my committee Timothy Larson, Shirley Luckhart, David Popham, and Malcolm Potts, Eugene Gregory who chaired my preliminary examination, Tom Sitz under whom I was a teaching assistant, John Hess, Dennis Dean, and Jiann-Shin Chen.

Thank you to Rich Helm and Kim Harich for helping teach me about MALDI-MS/peptide fingerprinting. Thanks to Ken Bischoff who taught me about protein purification and various techniques in molecular biology, and perhaps most importantly not to take myself too seriously. Thank you to Liang Shi who taught me about cloning and Ben Potters who always questioned everything I did, and for all those games of "pig". Many thanks Bettina Siebers for teaching me about two-dimensional electrophoresis and various techniques in molecular biology, helping me study for my prelim, and of course, the many dance steps of clogging. Much thanks to my brother, Steven, for encouraging me to come to Virginia Tech in the first place and for all the fun we've had together these past few decades, may it continue for many more. I would also like to acknowledge the people that I worked with and shared lab space including Brian, Tom, Tim, Jim, Songhai, Archana, Rick, Hong, Farzana, Ali, Sue, Debbie, and Janet. Thanks also to the Biochemistry secretaries Mary Jo, Karen, Sheila, and Peggy.

I would especially like to thank my family and friend for their love and support through my life. Thank you to my in laws Wayne and Sue, to my parents Jim and Patty, my sister Beth, my brothers Jon, Kevin, and Michael, my grandparents G&G Frank, G&G Lower, G&G Kelly, G&G Pardee, George and Peg Kilbride, and Ben Foote. I would most especially like to thank my girls Marsh, Brit, Peanut, and Nanna whom I live for every day. I was truly blessed four times by four wonderful girls, I ♥ U all ∞!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ADP	adenosine diphosphate
Amp	ampicillin
AMP	adenosine monophosphate
ATP	adenosine triphosphate
BLAST	basic local alignment search tool
bp	base pair
BSA	bovine serum albumin
cAMP	3', 5'-cyclic adenosine monophosphate
CAPS	3-(cyclohexylamion)-1-propane sulfonic acid
CIAP	calf intestinal alkaline phosphatase
cGMP	3', 5'-cyclic guanosine monophosphate
CMC	critical micelle concentration
CPM	counts per minute
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
dGTP	deoxyguanosine triphosphate
dTTP	deoxythymidine triphosphate
dNTP	deoxynucleoside triphosphate
Da	dalton
DAS	Dense Alignment Surface Analysis
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EGTA	ethylenedioxydiethylenedinitrolo tetraacetic acid
EtBr	ethidium bromide
FPLC	fast performance liquid chromatography
GDP	guanosine diphosphate
GMP	guanosine monophosphate
GTP	guanosine triphosphate
IEF	isoelectric focusing
IPG	immobilized polyacrylamide gel
IPTG	isopropyl-β-D-thiogalactopyranoside
kbp	kilobase pair
kDa	kilodalton
LB	Luria broth
MALDI-MS	matrix-assisted laser desorption ionization mass spectrometry
Mb	megabases
MBP	myelin basic protein
MES	2-(N-morpholino) ethane sulfonic acid
MOPS	3-(N-morpholino) propane sulfonic acid
M_r	relative molecular mass
$NaPP_i$	sodium pyrophosphate
ORF	open reading frame

PAA	Phosphoamino acid
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
pI	isoelectric point
PMSF	phenylmethanesulphonyl fluoride
PVDF	polyvinylidene difluoride
RCM	reduced carboxyamidomethylated and maleylated
SDS	sodium dodecyl sulfate
TAE	Tris-acetate / EDTA
TBS	Tris-Buffered Saline
TBST	Tris-Buffered Saline + Tween 20
TCA	trichloroacetic acid
TFA	trifluoroacetic acid
TFMS	trifluoromethanesulfonic acid
TOF	time of flight
TLC	thin layer chromatography
Tris	tris(hydroxymethyl) aminomethane

CHAPTER I

Introduction

Protein Phosphorylation-Dephosphorylation: a Ubiquitous Molecular Regulatory Mechanism

Living organisms evolved a diverse collection of mechanisms for adapting to ever changing environmental conditions. Indeed survival itself depends upon the ability to monitor and respond to a wide range of internal and external variables, or signals. It follows that signal processing is a universal feature of every cell for the purpose of controlling cellular machinery. Both extracellular and intracellular signal processing depends upon the specific interactions between proteins. In principle, any protein having the ability to transform an input signal into an output signal can act as a computational or information carrying element. Among the events involved in the communication between proteins, protein phosphorylation-dephosphorylation occupies a central position in that it is perhaps the most versatile mechanism for changing the functional state of a protein in a rapid and reversible manner (Hunter, 1995).

The covalent attachment of an electron dense phosphate group onto an amino acid side-chain of a protein triggers a conformational change that alters its functional properties (Johnson and Bradford, 1993). Such covalent bonds, i.e. phosphomonoesters of serine, threonine, and tyrosine, are stable in a typical cellular environment of moderate temperature, neutral pH, and aqueous milieu thus requiring the intervention of a protein phosphatase to restore it to its former state. Through the concerted actions of protein kinases and protein phosphatases, which themselves sometimes undergo phosphorylation, proteins can be interchanged on command between functionally and informationally distinguishable states.

Interconnected sets of protein kinases, protein phosphatases, and phosphoproteins they target, provide the basis for constructing regulatory networks of great versatility and sophistication. Processes targeted by these networks include metabolic pathways (Cohen, 1988; Hoekstra, et al., 1991; Kennelly, 1991), cell division and cell growth (Cyert and Thorner, 1989; Mendenhall and Hodge, 1998), gene transcription and translation

(Erikson, 1991; Kaufman, 1997; Pain 1994), hormonal responses (Cohen, 1982; Cohen, 1988), neuronal signal transduction (Walaas and Greengard, 1991; Cohen, 1982), muscle contraction (De Lanerolle and Paul, 1991; Sellers and Adelstein, 1987), pathogenic virulence factors (Cartier et al., 1999; Grose 1990; Kennelly and Potts, 1999; Mathews, 1990), and learning and memory (Soderling, 1993). Dysfunctions in protein phosphorylation-dephosphorylation networks have been implicated in many human diseases such as cancer (Selkirk, et al., 1996), diabetes (Sjoholm, 1998), Alzheimer's Disease (Julien and Mushynski, 1998), and cystic fibrosis (Sheppard and Welsh, 1999).

Protein Kinases

Protein kinases catalyze the transfer of the gamma phosphoryl group of a nucleotide triphosphate, usually ATP, to a nucleophilic acceptor group on an amino acid side chain of a protein. The acceptor amino acid forms one of several covalent bonds with the phosphoryl group. These linkages include phosphomonoesters with the hydroxyl groups of serine, threonine, and tyrosine; acyl phosphates with the carboxylate groups of aspartic and glutamic acid; and phosphoramides with histidine, lysine, and arginine. By convention protein kinases are classified on the basis of the type of amino acid residue they modify, e.g. a protein-tyrosine kinase catalyzes the transfer of a phosphate group from ATP to a tyrosine residue of the protein substrate.

Despite the fact that no standard approach currently exists for naming protein kinases, the vast majority falls into two major superfamilies based both on primary amino acid sequence and mechanistic similarities. Histidine kinases, the first conserved unit of the so-called two-component regulatory system, contain a conserved catalytic domain of approximately 250 amino acids (Cozzone, 1998). Upon activation via an associated sensor domain or protein, the histidine kinase undergoes autophosphorylation on nitrogen of a conserved His residue. The phosphoryl group is transferred to an Asp residue in the response regulator domain of its downstream target, completing the transfer reaction (Parkinson, 1993). Both domains may be localized on separate molecules or on one single polypeptide chain.

The second superfamily is defined by its homology to the catalytic subunit of cAMP-dependent protein kinase (cAPK) and ability to phosphorylate serine and/or

threonine, or tyrosine residues (Taylor et al., 1992). These cAPK-like protein kinases consist of a 250-300 amino acid catalytic domain that is subdivided into 12 highly conserved subdomains. The cAMP-kinases catalyze the direct, one-step transfer of a phosphate group from ATP to a serine, threonine, or tyrosine residue (Hanks and Hunter, 1995). The catalytic domain itself is bilobal with a smaller N-terminal nucleotide-binding lobe (subdomains I-IV) hinged to a C-terminal substrate-binding lobe (Zheng et al., 1993). Subdomain I contains the consensus motif Gly-X-Gly-X-X-Gly-X, known as the "glycine-rich motif," that serves to anchor the ATP to the kinase. Subdomain II contains an important invariant lysine residue, usually Lys72, which forms a salt bridge with the gamma phosphate of ATP. It is believed that an invariant glutamate residue in subdomain III, usually Glu 91, forms a salt bridge with Lys 72, stabilizes the interaction of Lys 72 with the alpha and beta phosphates of ATP. No invariant residues are present in subdomain IV so its role seems to be structural. Subdomain V serves as the bridge linking the two lobes of the protein kinase. It consists of a hydrophobic beta strand in the N-terminal lobe connected to an alpha helix in the C-terminal lobe by a chain of highly conserved amino acids. These amino acids include Glu121, Val 123 and Glu 127 that help to anchor the MgATP by hydrogen bonding with the adenine or ribose rings and residues Met120, Tyr122 and Val123 that form a hydrophobic pocket that surrounds the adenine ring. Subdomain VIa contains no invariant residues so its function also appears to be structural.

Subdomain VIb, known as the "catalytic loop," contains the invariant residues Asp166 and Asn171, which are members of the larger consensus motif of His-Arg-Asp-Leu-Lys-X-X-Asn. It is believed that Asp 166 acts as a base accepting the proton from the substrate's hydroxyl group at the site of phosphorylation leaving an oxyanion that performs a nucleophilic attack of the gamma phosphate from MgATP. Subdomain VII, usually occupying positions 184-186, contains a very well conserved motif of Asp-Phe-Gly. The Asp184 is an invariant residue that functions to orient the gamma phosphate of MgATP for transfer to the phosphoprotein substrate. It does this by chelating the Mg^{2+}ion bridging the beta and gamma phosphates of the MgATP. Subdomain VIII contains a highly conserved triplet of Ala-Phe-Gly, occupying positions 206-208, which plays a major role in substrate recognition. Subdomain IX contains an invariant Asp residue at

position 220 that acts to stabilize the catalytic loop through hydrogen bonding. Subdomain X does not have any well-conserved motifs. Subdomain XI extends to about 300 amino acids of the protein and defines the C-terminal boundary of its catalytic domain. This subdomain contains a nearly invariant Arg residue at position 280, and for protein-serine/threonine kinases contains a consensus motif of His-X-Aromatic-Hydrophobic found 9-13 positions down from Arg 280 (Smith et al., 1997).

Recently several novel protein kinases were discovered that do not conform to either of the two major kinase superfamilies. These include myosin heavy chain kinase A from *Dictyostelium discoideum* (Futey et.al., 1995) and the mammalian elongation factor-2 kinase (Ryazanov et.al., 1997), which together have defined a new eukaryotic protein kinase family. Computer database searches using the various algorithms failed to detect any homology between the proposed catalytic domain of these novel protein kinases and members of either the histidine kinase or cAMP-dependent protein kinase family. By eye, however, these novel protein kinases do display some similarity to "eukaryotic-like" cAMP-dependent protein kinases (Futey et. al., 1995; Pavur et. al., 2000). A possible nucleotide-binding motif is observed and candidate residues for the invariant lysine of subdomain II and for the conserved glutamate of subdomain III are also present. Beyond this restricted similarity, no significant resemblance is evident. Likewise, a novel protein kinase, bacterial HPr (heat-stable protein) kinase, was also discovered in Gram-positive bacteria (Galinier et. al., 1998). A recent report suggests HPr kinase may function as a protein kinase / protein phosphatase in a manner reminiscent of the isocitrate dehydrogenase kinase / phosphatase (Kravanja et. al., 1999).

Isocitrate dehydrogenase kinase / phosphatase (AceK) is a unique bifunctional enzyme with the ability to both phosphorylate and dephosphorylate isocitrate dehydrogenase (IDH) (LaPorte et. al., 1982; LaPorte et. al., 1985). AceK bears little resemblance to other known protein kinases and no resemblance to other known protein phosphatases (Cortay et. al., 1988; Klumpp et. al., 1988). Site-directed mutagenesis and enzyme kinetic studies indicate that the phosphorylation and dephosphorylation of IDH occurs at the same active site, however the product of the protein phosphatase reaction is P_i and not MgATP, suggesting that dephosphorylation does not proceed via a simple reversal of the phosphorylation reaction (Miller et. al., 1996). It is postulated that ADP

acts as a phosphate acceptor forming a transient intermediate MgATP that is quickly hydrolyzed to ADP and P_i during the dephosphorylation reaction (Miller et. al., 1996). Thus dephosphorylation and phosphorylation of IDH by AceK occurs via two thermodynamically independent mechanisms, as is the case for other "conventional" phosphoproteins that are phosphorylated and dephosphorylated by two separate enzymes.

Isolated reports have appeared describing other protein kinases in eukaryotes whose sequences cannot be accommodated within established protein kinase superfamilies. These include the product of the *BCR* gene (Maru and Witte, 1991), which has an apparent nucleotide binding domain (subdomain I) and phosphotransferase domain (subdomain VIb) (Laurent et.al. 2001); Dyrk-related protein kinases of several eukaryotic organisms (Becker and Joost, 1999), which are intriguing in their ability to catalyze tyrosine-directed autophosphorylation as well as serine/threonine-directed phosphorylation of exogenous substrates; Goodpasture antigen protein kinase (Raya et. al., 1999), a protein-serine/threonine protein kinase which specifically phosphorylates the Human Goodpasture antigen; and actin fragment protein kinase (Eichinger et. al., 1996), whose C-terminal domain exhibits homology to the cAMP-dependent protein kinases. However, it is difficult to establish the significance of these examples until corroborating evidence appears that firmly establishes the functional properties for these protein kinases and demonstrates the existence of additional families members.

Protein Phosphatases

Most protein phosphatases can be grouped into three superfamilies: the PPP family of serine/threonine phosphatases (PP1/2A/2B), the PPM family of serine/threonine phosphatases (PP2C), and the protein-tyrosine phosphatases (PTPs). The original classification system used to group protein-serine/threonine phosphatases was first introduced in the 1980s to distinguish the various protein phosphatases based on their enzymatic properties (Cohen P., 1991). These enzymes were divided into two groups by their ability to phosphorylate either the α subunit or β subunit of phosphorylase kinase and their sensitivity to nanomolar amounts of two heat/acid-stable inhibitor proteins, inhibitor 1 (I-1) and inhibitor 2 (I-2). Type 1 protein phosphatases (PP1) preferentially phosphorylate the β subunit of phosphorylase kinase and are inhibited by I-1 and I-2.

Conversely, Type 2 protein phosphatases (PP2) preferentially phosphorylate the α subunit of phosphorylase kinase and are insensitive to I-1 and I-2. Type 2 protein phosphatases can be subdivided into three main groups PP2A, PP2B, and PP2C via their divalent metal ion requirement. PP2A is metal ion independent, while PP2B is Ca^{2+} dependent and PP2C is Mg^{2+} dependent. Members of the PPP superfamily (PP1/2A/2B) share a highly conserved catalytic domain of roughly 220 amino acids characterized by three highly conserved motifs: Gly-Asp-X-His-Gly (motif I), Gly-Asp-X-X-Asp-Arg-Gly (motif II), and Gly-Asn-His-Glu (motif III) (Barton et al., 1994). Members of the PPM superfamily, whose defining member is PP2C, are characterized by a catalytic domain spanning approximately 290 amino acids containing 11 highly conserved motifs of which seven contain the absolutely conserved residues Asp 38 (motif I), Asp 60 (motif II), Thr 128 (motif IV), Gly 145 (motif V), Gly 198 (motif VI), Asp 238 and Gly 240 (motif VIII), and Asp 282 (motif XI) (Kennelly et. al., 1998; Barford, 1996). It is believed that members of the PPP as well as PPM superfamilies catalyze the one-step hydrolysis of a phosphoserine or phosphothreonine residue (Kennelly, 1999).

Protein-tyrosine phosphatases are the most abundant family of protein phosphatases and are divided into two main subfamilies: conventional PTPs and low molecular weight PTPs. Some conventional PTPs are more promiscuous in their enzymatic activities. These enzymes are referred to as dual-specific protein phosphatases (DSPs) because of their ability to hydrolyze all three phosphoester linkages: phosphoserine, phosphothreonine, and phosphotyrosine (Kennelly and Potts, 1999). Both conventional PTPs and low molecular weight PTPs share a characteristic active site signature sequence motif (HAT motif): Cys-X_5-Arg, as well as a two-step mechanism for the hydrolysis of phosphotyrosine residues (Guan and Dixon, 1991). In the first step, the deprotonated thiol group of the active site Cys residue carries out a nucleophilic attack on the protein bound phosphate group, resulting in the transfer of the phosphate to the enzyme and subsequent release of dephosphorylated protein. Water then enters the active site and hydrolyzes the cysteinyl phosphate group to generate inorganic phosphate and free enzyme (Zhang et. al., 1994). The most prominent difference between the conventional PTPs and low molecular weight PTPs is the position of the HAT motif within the catalytic domain and its location relative to a conserved aspartic acid residue,

which serves as a general acid/base (Zhang, et. al., 1994). In the conventional PTPs the HAT motif is located within the central segment of the catalytic domain, which is approximately 250 amino acids in length, and the conserved aspartic acid residue is 25-50 amino acids to the N-terminal end of the active site cysteine (Kennelly and Potts, 1999). For low molecular weight PTPs the HAT motif is located near the extreme N-terminus of the catalytic domain, which is considerably smaller in length, approximately 140 amino acids, with the essential aspartic acid residue located 80-110 residues to the C-terminal end of the active site cysteine (Kennelly et. al., 1998). While the two catalytic domains are extremely different in structure, the spatial relationships of the catalytic residues within the primary sequence are remarkably similar (Eckstein et. al., 1996).

A Brief History of Protein Phosphorylation-Dephosphorylation

E. G. Krebs and E. H. Fisher first demonstrated the regulation of enzyme activity by protein phosphorylation in their seminal studies of glycogen metabolism in skeletal muscle (Krebs and Fisher, 1956). They discovered that phosphorylase kinase phosphorylates glycogen phosphorylase, thereby converting it to its constitutively active form. Because of this early association with the neuroendocrine system, protein phosphorylation was largely regarded as a refined regulatory mechanism which emerged late in evolutionary time to meet the specific needs of the multicellular organisms (Greengard, 1978; Kennelly and Potts, 1996). Not until the early 1980s did it, become apparent that protein phosphorylation was likewise a phenomenon of the "lower" eukaryotes and bacteria (Wang and Koshland, 1978; Garnak and Reeves, 1978; Manai and Cozzone, 1979). However, as research on protein phosphorylation progressed, a new dichotomy emerged in which it was presumed eukaryotic organisms almost exclusively targeted the hydroxyl amino acids serine, threonine, and tyrosine, while prokaryotes preferred to target histidine and carboxyl amino acids (Shi, et al., 1998). It is only within the last decade that many "eukaryotic-like" and "prokaryotic-like" protein kinases and protein phosphatases were discovered outside their respective phylogenetic domains (Kennelly and Potts, 1996). These reports strongly challenged the notion that genuinely "eukaryotic" and "prokaryotic" protein kinases and protein phosphatases exist, and opened up new perspectives for the study of signal transduction in these organisms.

Signal Transduction: Eukaryotes Versus Prokaryotes

Because many cellular processes are regulated by protein phosphorylation-dephosphorylation, it is one of the most extensively studied and understood molecular regulatory mechanisms. An organism's protein phosphorylation network, whether it is prokaryote or eukaryote, is comprised of three interrelated parts: protein kinases, protein phosphatases, and the phosphoprotein substrates the former two enzymes target. The interplay between these three components as they respond to a multitude of internal and external signals dictates the phosphorylation status of proteins within a cell. These phosphorylation events result in modifications of gene expression or enzymatic reactions, which enable cells to generate an appropriate response to any given signal (Zhang, 1996).

In eukaryotes, a signal is often transmitted through a cascade of protein phosphorylation-dephosphorylation events that alter the activity of many target proteins. This is best illustrated by studies on mitogen-activated protein kinase (MAPK) pathways in yeast (Herskowitz, 1995). These pathways generally contain three protein kinases that act in succession: a MAP kinase kinase kinase (MAPKKK or MEKK), which phosphorylates a MAP kinase kinase (MAPKK or MEK), which in turn phosphorylates a MAP kinase (MAPK) (Gustin et. al., 1998). Also critical to these pathways are the various protein-serine/threonine and protein-tyrosine phosphatases that dephosphorylate the protein kinases of the MAPK cascades and the downstream MAPK phosphoprotein substrates. MAPK cascades are found in animals, plants, and fungi where they regulate transcription factors by MAPK-mediated phosphorylation (Gustin et. al., 1998). Extensive genetic and biochemical analysis as well as the complete sequencing of the genome of *Saccharomyces cerevisiae* has revealed five distinct MAP kinase pathways, which are involved in the regulation of several processes including mating, cell integrity, sporulation, and osmosensing (Gustin et. al., 1998; Levin and Errede, 1995). Such elaborate mechanisms are thought to have evolved to accommodate the level of cellular and molecular complexity found in eukaryotic organisms (Zhang, 1996).

Most bacterial signal-transduction pathways are composed of two proteins (two-component regulatory pathways): a histidine kinase and a response regulator (Stock et al., 1992; Zhang, 1996). The mechanism itself appears rather primitive in as much as the

phosphoramide linkage resembles a short-lived high-energy transition state linking the phosphoryl transfer from the donor ATP to an acceptor protein (Stock et. al., 1992). In proteins, the hydrolysis of the phosphoramide linkage provides sufficient energy to bring about the subsequent phosphorylation of an Asp residue on the response regulator, therefore it seems unlikely that the phosphohistidine residue on the kinase is dephosphorylated by a protein phosphatase (Stock and Stock, 1990). Acyl phosphate linkages, on the other hand, are comparably more stable and require enzymatic hydrolysis of their phospho-aspartate bonds. Several protein-aspartate phosphatases responsible for dephosphorylating the response regulator protein of two-component systems have recently been discovered (Perego et. al., 1994; Perego and Hoch, 1996).

A good example of the two-component regulatory pathway is demonstrated by bacterial chemotaxis (Stock et al. 1990). Here, the sensor histidine kinase (CheA) autophosphorylates in response to the conformation of the membrane receptor protein and subsequently transfers the phosphoryl group to an aspartic acid residue on CheY. In this active phosphorylated form, CheY exhibits enhanced binding to a switch component, FliM, at the flagellar motor (Schuster et. al., 2001). This induces a change from counterclockwise to clockwise flagellar rotation causing the bacterium to tumble. Ultimately, the phosphorylation status of CheY, determined by the sensor kinase CheA and the protein phosphatase CheZ, controls the swimming behavior of the bacterium (Schuster et. al., 2001).

The past decade has witnessed the discovery of surprising parallels between the protein kinases and protein phosphatases of eukaryotic and prokaryotic organisms. Several laboratories have reported “bacterial-like” two-component modules in eukaryotic organisms such as yeast (Ota and Varshavsky, 1993; Maeda et al., 1994; Tao et al., 1999; Janiak-Spens et al., 1999); fungi *Neurospora crassa* and *Candida albicans* (Alex et al. 1996; Calera et al., 1998); plants *Arabidopsis thaliana* (Chang et al., 1993; Kakimoto, 1996); and slime molds *Dictyostelium discoideum* (Brown and Firtel 1998; Thomason et al., 1999; Wang et al., 1999). Conversely, homologs of “eukaryotic” O-kinases and O-phosphatases have been discovered in many prokaryotic organisms. These include a protein-tyrosine phosphatase, Ptp, from *Acinetobacter johnsonii* (Grangeasse et al., 1998) and a protein-tyrosine phosphatase, Wzb, from *Escherichia coli* (Vincent et al., 1999).

Both were shown to specifically dephosphorylate an endogenous protein-tyrosine kinase, Ptk and Wzc respectively, which autophosphorylates on tyrosine residues. These findings suggest the occurrence of a regulatory mechanism in prokaryotes connected with reversible protein phosphorylation on tyrosine. Other examples include the discovery of a protein-serine/threonine kinase, PpkA, from *Pseudomonas aeruginosa*, which appears to be important in regulating the expression of virulence factors (Motley and Lory, 1999) and protein phosphorylation on tyrosine as well as serine and threonine residues in the multicellular prokaryote *Myxococcus xanthus* (Frasch and Dworkin, 1996; Udo et al., 1997).

In Cyanobacteria, several "eukaryotic-like" protein phosphatases and protein kinases have similarly been encountered. These include: a protein-tyrosine/serine phosphatase, IphP, from *Nostoc commune* UTEX 584, which represented the first protein-tyrosine phosphatase encoded by the chromosomal DNA of any prokaryote (Potts et al., 1993); *the* glnB gene product, PII protein, from *Synechococcus* sp. PCC 7942 which signals the cellular state of nitrogen assimilation relative to CO_2 fixation by being phosphorylated at a seryl residue (Forchhammer and Tandeau de, 1995); a protein-serine/threonine kinase, SpkA from *Synechococcus* sp. PCC 6803, which the authors suggest regulates cellular motility via the phosphorylation of several membrane proteins (Kamei et. al., 2001); and finally, four genes from the *Anabaena* sp. PCC 7120, which encode three "eukaryotic-like" protein kinases, PknC, PknD and PknE, and one a protein phosphatase PP1/2A/2B homolog, PrpA, all of which are hypothesized to regulate similar biological processes such as heterocyst structure formation and nitrogen fixation by governing the level of phosphorylation of targets phosphoproteins (Zhang and Libs, 1998; Zhang et al., 1998; Gonzales et. al., 2001).

Several newly identified "eukaryotic-like" protein kinases from the *Bacteria* include: PknB from *Mycobacterium tuberculosis* which was shown to autophosphorylate and phosphorylate the substrate myelin basic protein on serine/threonine residues (Av-Gay et. al., 1999); PutA from *Salmonella typhimurium* which autophosphorylates on several threonine, serine, and tyrosine residues (Ostrovsky and Maloy, 1995); RsbT, a protein-serine/threonine kinase from *Bacillus subtilis* that is involved in environmental stress signaling pathways (Gaidenko et al., 1999); enfin, an ATP-dependent protein-

serine kinase, HPr(Ser), from *Streptococcus salivarius* (Brochu and Vadeboncoeur, 1999) and *Enterococcus faecalis* (Kravanja et al., 1999) which is involved in the regulation of carbohydrate metabolism.

The recent sequencing of numerous bacterial genomes provides a comprehensive picture of the number of "eukaryotic-like" protein kinases and protein phosphatases in members of the *Bacteria* by utilizing various genome analysis techniques. *Synechocystis* sp. strain PCC 6803 was found to possess seven protein-serine/threonine kinases, seven protein-serine/threonine and protein-tyrosine phosphatases, one protein kinase interacting protein, and one protein kinase regulatory subunit (Zhang et. al., 1998). Interestingly, in several cases, genes of two-component regulatory pathways were found within the same gene cluster as those encoding a protein serine/threonine kinase or a protein-serine/threonine phosphatase. This may suggest the possibility of the coupling of some protein-serine/threonine kinases and phosphatases to two-component systems, with sensors of two-component systems serving as membrane receptors that act upstream of a cascade of serine/threonine kinases in the same signal transduction pathway, as is the case for ethylene response in *Arabidopsis thaliana* (Chang et. al., 1993) and for high osmolarity adaptation in *S. cerevisiae* (Maeda et. al., 1994). Inspection of the genomes *Bacillus subtilis* 168, *Borrelia burgdorferi* B31, *Escherichia coli* K-12, *Haemophilus influenzae* KW20, *Helicobacter pylori* 26695, *Mycoplasma genitalium* G-37, and *Synechocystis* sp PCC 6803 revealed that each contains at least one ORF whose predicted product displayed sequence features characteristic of protein-serine/threonine and protein-tyrosine kinases, and protein-serine/threonine and protein-tyrosine phosphatases (Shi et. al., 1998). The intracellular pathogen *Mycobacterium tuberculosis* appears to contain eleven "eukaryotic-like" protein kinases (Av-Gay and Everett, 2000), and lastly, the gram-negative bacterium *Myxococcus xanthus* was shown to contain at least thirteen "eukaryotic-like" protein-serine/threonine kinases (Pkn1 to Pkn13) which share many of the conserved subdomains of cAMP-dependent protein kinases (Inouye et. al., 2000). The C-terminal regions of eleven of these protein kinases were shown contain at least one transmembrane domain, suggesting that they function as transmembrane sensor kinases.

Likewise, examination of several eukaryotes genomes has revealed the presence of ORFs whose predicted products display sequence features characteristic of "bacterial-

like" histidine-aspartate phosphorelay proteins (Thomason and Kay, 2000). These include 20 phosphorelay genes from *Dictyostelium discoideum*; 10 phosphorelay genes from the yeast *Saccharomyces cerevisiae* and *Saccharomyces pombe*; 8 phosphorelay genes from the fungi *Candida albicans*, *Neurospora crassa*, and *Aspergillus nidulans*; and 41 phosphorelay genes from *Arabidopsis thaliana* (Thomason and Kay, 2000). It is interesting to note that all these identified genes are from the genomes of lower eukaryotes and plants. To date no such genes have been identified from any member of the animal kingdom including the published genomes of *Caenorhabditis elegans* and *Drosophila melanogaster*, as well as the Human Genome Project (Thomason and Kay, 2000). Since these pathways are seemingly absent in animals and several phosphorelay systems have been implicated in pathogenic virulence (Uhl and Miller, 1996; Locht, 1999) and antibiotic resistance (Hakenbeck et. al., 1999; Novak et. al., 2000), current attention has focused on the development of therapeutics that target these pathways (Thomason and Kay).

Although it is still evident that eukaryotic and prokaryotic protein phosphorylation networks retain certain fundamental distinctions, there is a plethora of evidence suggesting that these systems and their corresponding enzymes share a number of structural and functional similarities. Commonality suggests that reversible phosphorylation arose early in evolution as a dynamic and versatile regulatory mechanism. It is also apparent that "eukaryotic-like" protein kinases and protein phosphatases existed before the divergence of prokaryotes and eukaryotes (Kennelly and Potts, 1996; Leonard et. al., 1998, Ponting et. al.1999, 1999, Zhang, 1996).

The Three Domains of Life

More than two decades after its original proposal, the archaeal hypothesis endures as the best explanation for the extraordinary diversity of molecular and biochemical features found in prokaryotic organisms (Whitman et al., 1999). The hypothesis simply avers that the prokaryotes are not monophyletic, but rather are a bifurcated group consisting of two very ancient phylogenetic lineages (Woese and Fox, 1977). The former bipartite classification, i.e. the Prokaryotes and the Eukaryotes, was forlorn for a more natural tripartite organization consisting of three domains: the *Eukarya*, the *Bacteria*, and

the *Archaea* (Woese et al., 1990). Using this classification, the designation of "eukaryote" and "prokaryote" takes on a morphological designation rather than a phylogenetic one. Eukaryote simply refers to organisms, which possess an internal nuclear membrane, that functions to separate their genomic material away from the cytoplasm (*Eukarya*), while prokaryote refers to organisms, which do not (*Bacteria* and *Archaea*). The most universally accepted "Tree of Life" to date is presented in Figure 1-1. This phylogeny depicts members of the *Archaea* and *Eukarya* having a more neoteric common ancestor than either has had with members of the *Bacteria* (Doolittle, 1995).

The *Archaea* are diverse collection of organisms comprised of three main groups: halophiles, methanogens, and hyperthermophiles. The halophiles inhabit highly saline environments such as solar salt evaporation ponds or natural salt lakes (i.e. Great Salt Lake in Utah), most species requiring 1.5 – 5.5 M NaCl for optimal growth (Madigan et. al., 1997). All extremely halophilic *Archaea* are chemoorganotrophs (obtain their energy from the oxidation of organic compounds) and most are obligate aerobes. The methanogens are all obligate anaerobes that, as the name implies, convert such compounds such as carbon dioxide, methanol, and acetate to methane in order to generate ATP (Madigan et. al., 1997). They are found in geothermal hot springs, deep-sea vents, and places where decomposing organic matter is found. The hyperthermophilic *Archaea* are a phylogenetically distinct group of organisms found in geothermally heated soils and waters or deep-sea hydrothermal vents (Madigan et. al., 1997). Many hyperthermophilic Archaea can grow chemolithotrophically (obtain their energy from the oxidation of inorganic compounds) using H_2, elemental sulfur, or ferrous iron as the energy source. Nearly all members of this group require elemental sulfur, which they use either as an electron donor for chemolithotrophic metabolism or as an electron acceptor for anaerobic respiration (Madigan et. al., 1997). All hyperthermophiles have a temperature optima above 80^{o}C and several members are capable of growth at temperatures above 100^{o}C.

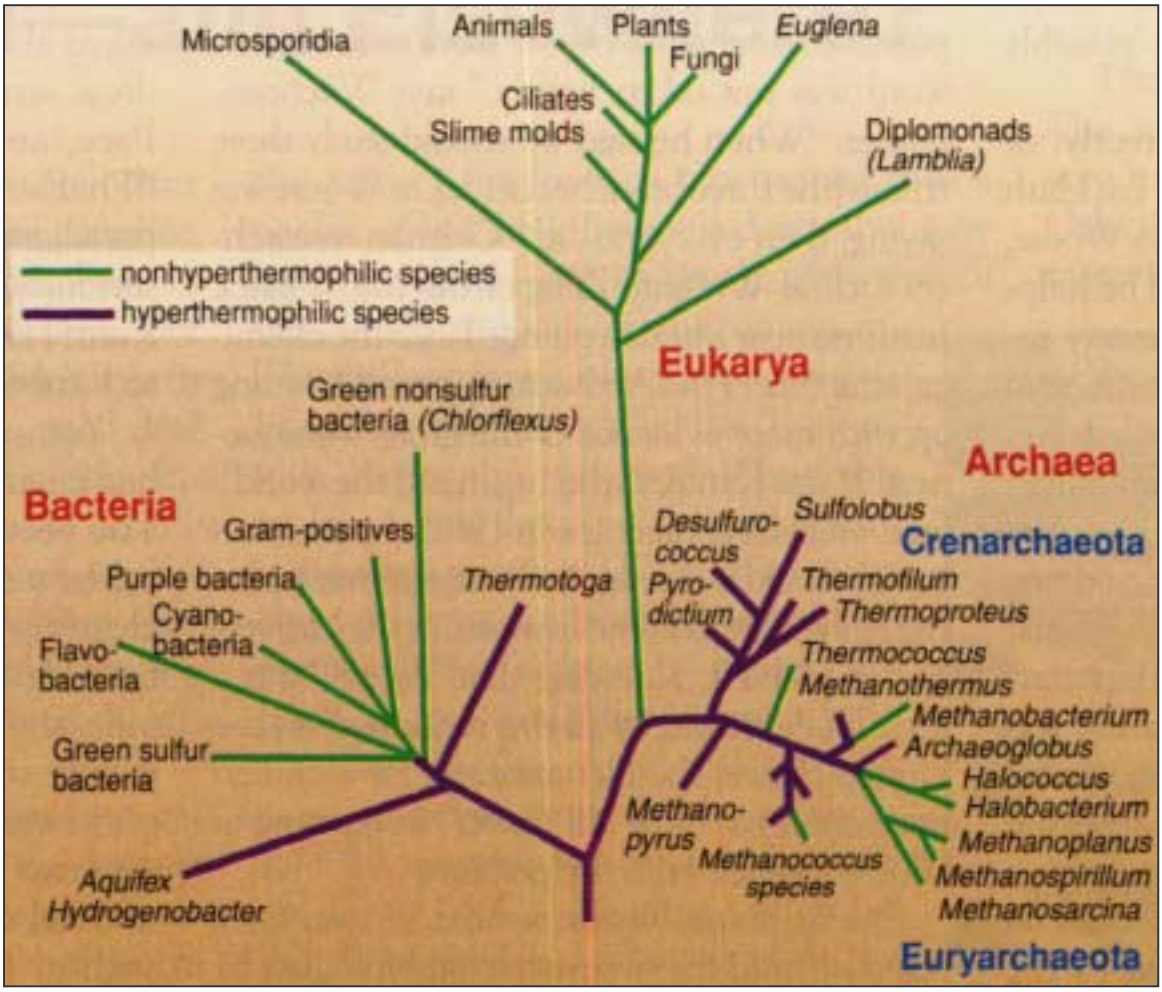


Figure 1-1. The Woese Tree of Life.

This phylogenetic tree shows the *Archaea* and the *Eukarya* being more closely related than either is to members of the *Bacteria*. This tree, which was originally based on rRNA analyses and now confirmed by other genomic sequences, such as *Methanococcus jannaschii*, separates all living organisms into three domains: *Archaea, Bacteria,* and *Eukarya*. The tree shows that most life is unicellular, and that the oldest cells were hyperthermophiles (Morell,1997).

Characteristics of the Archaea

Although classified morphologically as prokaryotes, members of the *Archaea* share many common characteristics with the *Eukarya* and are best described as "honorary eukaryotes with prokaryotic cellular organization" (Robb and Place, 1995). Their genomic organization is similar to *Bacteria,* as they possess a single circular genome, which is similar in size to typical bacterial genomes ranging from approximately 1.6 – 4.0 Mb in size. DNA binding proteins, homologous to eukaryal histones, have been identified in many archaeons and shown to stabilize duplex DNA against thermal denaturation in hyperthermophiles (Reddy and Suryanarayana, 1989). The majority of the proteins involved in replicating the archaeal chromosome are homologs of proteins required for eukaryotic DNA replication (Cann and Ishino, 1999). These include DNA polymerases, some which possess 3' – 5' exonuclease activity, topoisomerases, ATP-dependent DNA ligases, helicases, primases, and gyrases. Archaeal transcription more closely resembles that of the *Eukarya* as well. Their RNA polymerases contain up to 14 subunits, and sequence comparisons suggest that archaeal RNA polymerases more closely resemble eukaryal than bacterial ones (Werner et. al., 2000). Unlike *E. coli* RNA polymerase, archaeal RNA polymerase shows a weak ability to bind to promoter DNA and/or to initiate transcription *in vitro* (Thomm, 1996). Active transcription *in vitro* and *in vivo* depends strictly on the presence of a TATA box resembling the TATA box of eukaryal polymerase II promoters, which is recognized by a protein related to eucaryal TATA-binding protein (Thomm, 1996). It has also been demonstrated that archaeal transcription factors, which are homologous to eukaryal transcription factors, are essential for transcription as well (Hanzelka et. al., 2001). In contrast to eukaryotes, archaeal mRNAs are not polyadenylated and can be polycistronic in structure (Madigan et. al., 1997). Introns have been found in archaeal 23S rRNA, 16S rRNA, and tRNA genes (Itoh et. al., 1998; Armbruster et. al., 1997; Dalgaard and Garrett, 1992). The *Archaea* have 70S ribosomes, but unlike the *Bacteria*, translation is not inhibited by tetracycline, erythromycin, chloramphenicol, streptomycin, and kanamycin suggesting a mechanism homologous to that found in the *Eukarya* (Madigan et. al., 1997). The *Archaea*, like the *Eukarya*, use methionyl tRNA, instead of N-formyl-methionyl tRNA, to initiate translation (Lee et. al., 1999).

Genome vs. genome comparison of several archaeal and bacterial organisms indicate that genes related to basic metabolism and catabolic functions are shared among prokaryotes (Doolittle, 1999). It is believed this phenomena resulted from extensive horizontal gene transfer between the two domains following their divergence, however, it was also shown that informational genes (i.e. those involved in transcription, translation, DNA replication) were seldomly transferred between the domains, corroborating the previously cited biochemical and genetic evidence (Jain et. al., 1999; Feng et. al., 1997; Rivera et. al., 1998). Additionally, archaeal protein O-linked and N-linked glycosylation has been observed on surface-layer (S-layer) proteins and flagellins (Moens and Vanderleyden, 1997). The eukaryotic N-glycosylation consensus motif, Ser/Thr-X-Asn where X is any amino acid except proline, has been reported for *Halobacterium halobium* (Lechner and Wieland, 1989) and *Methanothermus fervidus* (Bröckl et. al., 1991). However, as a consequence of their lack of intracellular compartmentalization, (e.g. no golgi apparatus and endoplasmic reticulum), the *Archaea* utilize a mechanism distinct from that of eukaryotes to glycosylate their proteins (Moens and Vanderleyden, 1997).

The *Archaea* display a number of unique characteristics. Comparison of the archaeal genomes sequenced to date with the completed genomes of other organisms reveals substantial genetic diversity (Klenk et. al., 1997; Bult et. al., 1996; She et. al., 2000). Archaeal membranes are constructed of chemically unique lipids. In contrast to the lipids in *Bacteria* and *Eukarya* in which fatty acid side chains are linked to glycerol via ester linkages, archaeal lipids consist of a glycerol molecule bonded to isoprene molecules via an ether linkage (Madigan et. al., 1997). The two main classes of lipids present in the *Archaea* are glycerol diethers and glycerol tetraethers (Sprott, 1992). Thermophilic and extremely acidophilic Archaea construct their membranes of tetraether lipids, forming a lipid monolayer. These tetraether membranes are highly stable and have a limited permeability for protons even at high temperatures making it possible for these organisms to maintain a near neutral cytosolic pH and a viable proton motive force even under extreme conditions (Albers et. al., 2000; Schafer G., 1999). Cell walls of the *Archaea* are constructed of pseudopeptidoglycan, a polysaccharide similar to peptidoglycan consisting of alternating N-acetylglucosamine and N-acetyltalosaminuronic acid repeats, polysaccharides, glycoproteins, or proteins. The most

common type of cell wall is a surface layer (S-layer), consisting of glycoprotein and/or protein (Madigan et. al., 1997). Halophiles contain negatively charged amino acids in their cell walls, which act to neutralize the positive charges of the high Na^+ environment preventing cell lysis (Madigan et. al., 1997).

In the halophilic and thermophilic *Archaea*, catabolism of glucose proceeds via a modified Entner-Doudoroff pathway (Madigan et. al. 1997). In thermophiles, glucose-6-phosphate is oxidized to 6-phosphogluconate and then dehydrated to form 2-keto-3-deoxy-6-phosphogluconate (KDGP). In halophiles, glucose is oxidized to gluconate, dehydrated to form 2-keto-3-deoxygluconate, and then phosphorylated to form KDGP (Buchanan et. al., 1999). Metabolism then proceeds by the normal Entner-Doudoroff pathway. In sulfur-dependent *Archaea*, such as *Sulfolobus* and *Thermoplasm*a, 2-deto-3-deoxygluconate is cleaved directly to pyruvate and glyceraldehydes with no net ATP yield (Buchanan et. al., 1999).

Protein Phosphorylation in the Archaea

Currently, very little is known concerning protein phosphorylation-dephosphorylation in the *Archaea*. The phenomenon was first examined in 1980 in the halophilic archaeon *Halobacterium halobium* (Spudich and Stoeckenius, 1980). Two light-regulated phosphoproteins were shown to have both acid and hydroxylamine resistant phosphate linkages, suggesting the presence of phosphoserine or phosphothreonine residues. Several years later, Skorko reported that several proteins of the sulphur-dependent archaeon *Sulfolobus acidocaldarius* were phosphorylated *in vivo* on both serine and threonine residues (Skorko, 1984). However, no protein-kinase activity was detected when partially purified cell extracts were incubated with eukaryal phospho-acceptor proteins.

The past several years, however, have witnessed the discovery of unambiguous protein phosphorylation events in the *Archaea*. First, an archaeal protein-serine/threonine phosphatase designated PP1-arch was isolated and its gene cloned from *Sulfolobus solfataricus*. This enzyme displayed 29-31% identity to eukaryal protein phosphatases of the PPP superfamily PP1/2A/2B (Kennelly et. al., 1993; Leng et al., 1995). Two years later a second protein phosphatase, PP1-arch 2, that exhibited 30% sequence identity with

eukaryotic members of this same superfamily of protein phosphatases, was cloned from the methanogenic archaeon *Methanosarcina thermophila* (Solow et al., 1997). More recently, a third protein-serine/threonine phosphatase, Py-PP1, was cloned from the hyperthermophilic archaeon *Pyrodictium abyssi* (Mai et al., 1998). Its primary structure had approximately 40% homology to both PP1-arch and PP1-arch2, as well as 31-34% identity to the PPP phosphatases from the *Eukarya*.

Open reading frames (ORFs) encoding potential "eukaryotic-like" protein kinases and protein phosphatases have also been found in several members of the *Archaea* (King and Smith, 1995; Leonard et. al., 1998; Shi et al., 1998). These comprise three homologous ORFs encoding a potential protein kinase from the methanogenic archaeons *Methanococcus vannielii*, *M. thermolithotrophicus* and *M. voltae* (Smith and King, 1995). Likewise, three ORFs potentially encoding a protein-tyrosine phosphatase were found in *M. jannaschii* and *Pyrococcus horikoshii* (Stravopodis and Kyrpides, 1999). Whether any of these gene homologs encode a functional protein kinase or protein phosphatase is still unknown.

A two-component regulatory system similar to the CheA/CheY chemotaxis system in *Bacteria* has also been discovered in the *Archaea*. Rudolph and Oesterhelt reported a histidine kinase, CheA, from the archaeon *Halobacterium salinarium* that has 34% identity with CheA from *Bacillus subtilis* (Rudolph and Oesterhelt, 1995). The same year they cloned two genes from the same organism, *cheY* and *cheB*, whose protein products displayed 31% and 38% identity, respectively, to the known signal transduction proteins CheY and Che B from *Escherichia coli* (Rudolph et al., 1995). These proteins were implicated in the mechanism of chemo- and phototactic signal transduction in *H. salinarium*, suggesting this system is similar to the two-component signaling system known from chemotaxis in the eubacterium *E. coli*. Additionally, ORFs whose predicted products display sequence similarities to protein-histidine kinases have also been identified from several archaeal genomes (Kim and Forst, 2001). This study found that while many bacterial genomes contained several different types of protein-histidine kinases, archaeal genomes either contained one specific type or lacked any protein-histidine kinase altogether. From these findings the authors suggest that the different

types of protein-histidine kinases originated in *Bacteria* and that specific protein-histidine kinase types were acquired in *Archaea* by horizontal gene transfer events.

Why Sulfolobus solfataricus?

Selecting a member of the *Archaea* to unearth fundamental principles governing protein phosphorylation-dephosphorylation networks in every organism, would seem rather precarious. At first glance, members of this domain appear to be specialized for life at the extremes those anaerobic, geothermal, salty, acidic habitats that would cook other organisms to a crisp. One main branch of the *Archaea*, the *Crenarchaeota* to which *S. solfataricus* (Figure 1-2) belongs, is especially notorious for growth at high temperatures. Recently, however, novel crenarchaea have been reported from freezing cold marine environments, temperate rice paddies, terrestrial soils, freshwater lake sediments, permafrost layers, and even deep-sea oil reservoirs growing on superheated crude oil 3,000 meters below the bed of the North Sea (DeLong, 1998; Forterre, 1997; Stetter et al., 1993). Such natural microbial diversity testifies that members of the *Crenarchaeota*, and *Archaea* in general, are not an odd collection of beasts, but rather constitute a major element of the global biomass.

S. solfataricus itself has become one of the best-studied hyperthermophiles for several reasons: its complete genome was recently sequenced (She et. al., 2000); its relative ease of growth in a laboratory setting, with an optimum pH of 1.0 - 5.0, temperature ranges from 60^{o} – 90^{o}C, and aerobic growth conditions (Stetter and Zillig, 1985); the more than 30 plasmids and viruses that have been identified in *Sulfolobus* isolates (Zillig et. al, 1998); and perhaps most prominently, the fact that *S. solfataricus* is considered among the most "eukaryotic-like" members of the *Archaea* (Woese and Wolfe, 1985).

Given that contemporary phylogenetic trees group the archaeal and eukaryotic domains on the same branch, the *Archaea* appeared to offer the greater long-term promise of the two prokaryotic domains in a search for protein kinases and protein phosphatases that exhibited recognizable homology in sequence and possibly even biological function to those in eukaryotes. Among the prokaryotes, *S. solfataricus* represents perhaps the closest living relative to eukaryotic organisms, which can most readily allow for the study

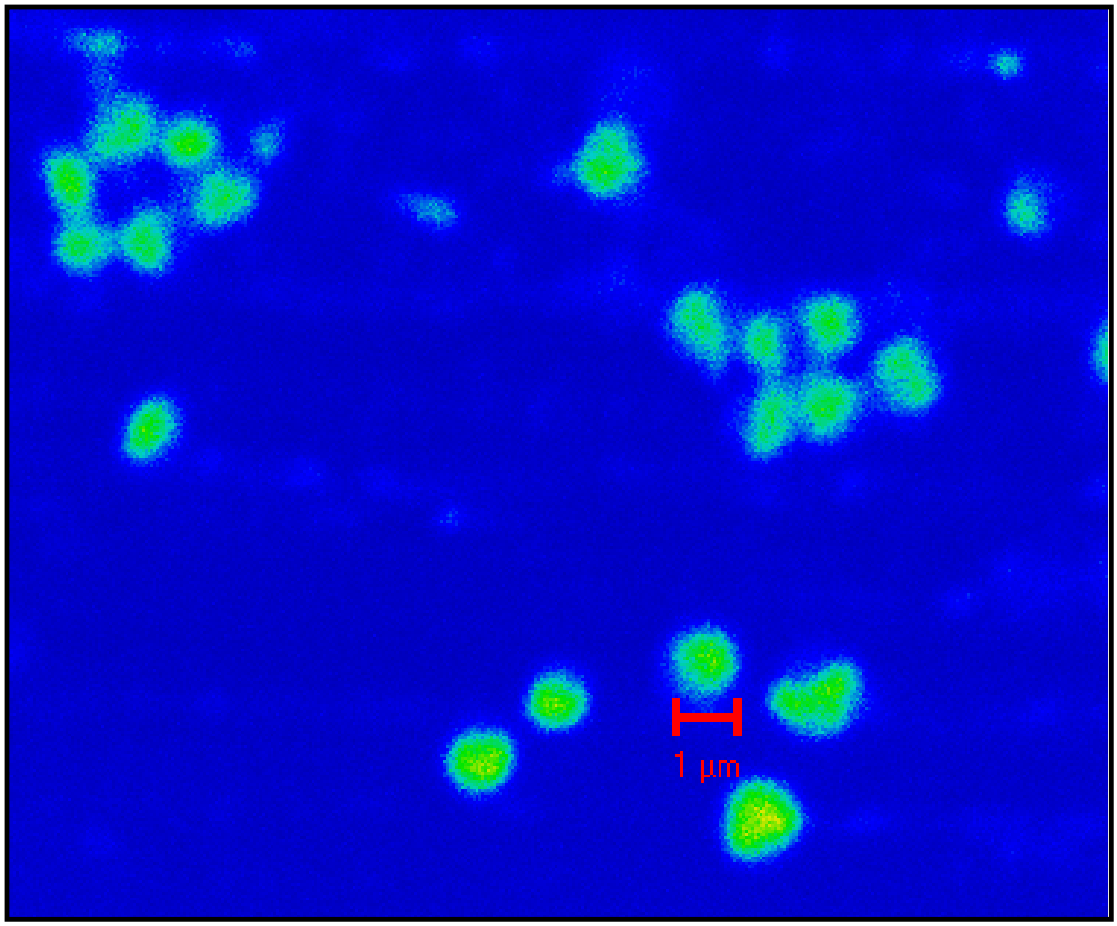


Figure 1-2. Scanning Electron Micrograph of *Sulfolobus solfataricus* Stained With Acrydine Orange.

S. solfataricus cells typically are 1.0 µm in diameter and spherical in shape, sometimes forming distinct lobes. Thomas Brock first discovered *Sulfolobus* in 1970 in the sulfur-rich hot springs of Yellowstone National Park (Brock et. al., 1972). It is an obligate aerobe capable of growth on H_2S, FeS_2 or S^o as well as on a variety of organic compounds. Its pH range for growth is 1.0 - 5.0 with a temperature range of 60^o – 90^oC.

of integrated protein phosphorylation-dephosphorylation networks as they may pertain to medically relevant organisms like ourselves. Hanks and Hunter estimated that the typical mammalian genome encodes approximately one thousand protein kinases and one thousand protein phosphatases, which in turn target several thousand phosphoproteins (Hanks and Hunter, 1995). The sheer quantitative complexity and the increasingly daunting prospect of elucidating the net effects of such networks demands the development of a much smaller model which would offer insight into the fundamental principles of these networks. Genome sequence analysis on several related species of *Archaea* indicates that these organisms contain approximately 100-fold fewer protein O-kinases and O-phosphatases than the mammalian cell (Kennelly, et al., 1998). Even if one is to take into account the fact that the genome size for *S. solfataricus* (2.99 Mb) is two-three times larger than the *Archaea* surveyed in Kennelly et. al., 1998, the number of O-phosphoproteins contained in *S. solfataricus* still pales in comparison to higher eukaryotes.

Finally, while *S. solfataricus* may represent a simplistic organism when weighed against the phosphorylation networks of higher eukaryotes such as ourselves, in actuality it represents a remarkably complex creature capable of diverse range of biochemical activities. It exhibits a broad range of metabolic activities, which have enabled it to inhabit a geographically diverse range of environments characterized by their extreme temperatures and acidity (Madigan et. al., 1997; Brock et. al., 1972). It also possesses various refined sensor and regulatory systems (Klenk et. al., 1997; Smith et. al., 1997; Taylor and Zhulin, 1998), that has accustomed it to living life "on the edge," and should provide valuable insight into the molecular regulatory mechanisms, which they possess (Figure 1-3).

Specific Aims and Significance of this Project

The long-term objective of this project is to develop an extensive functional model characterizing the complete network of protein O-phosphorylation processes (i.e. those phosphorylation events, which target serine, threonine, and tyrosine residues) in a single organism, the archaeon *Sulfolobus solfataricus*. Through the study of protein phosphorylation-dephosphorylation in these "protoeukaryotes," it is our intent to

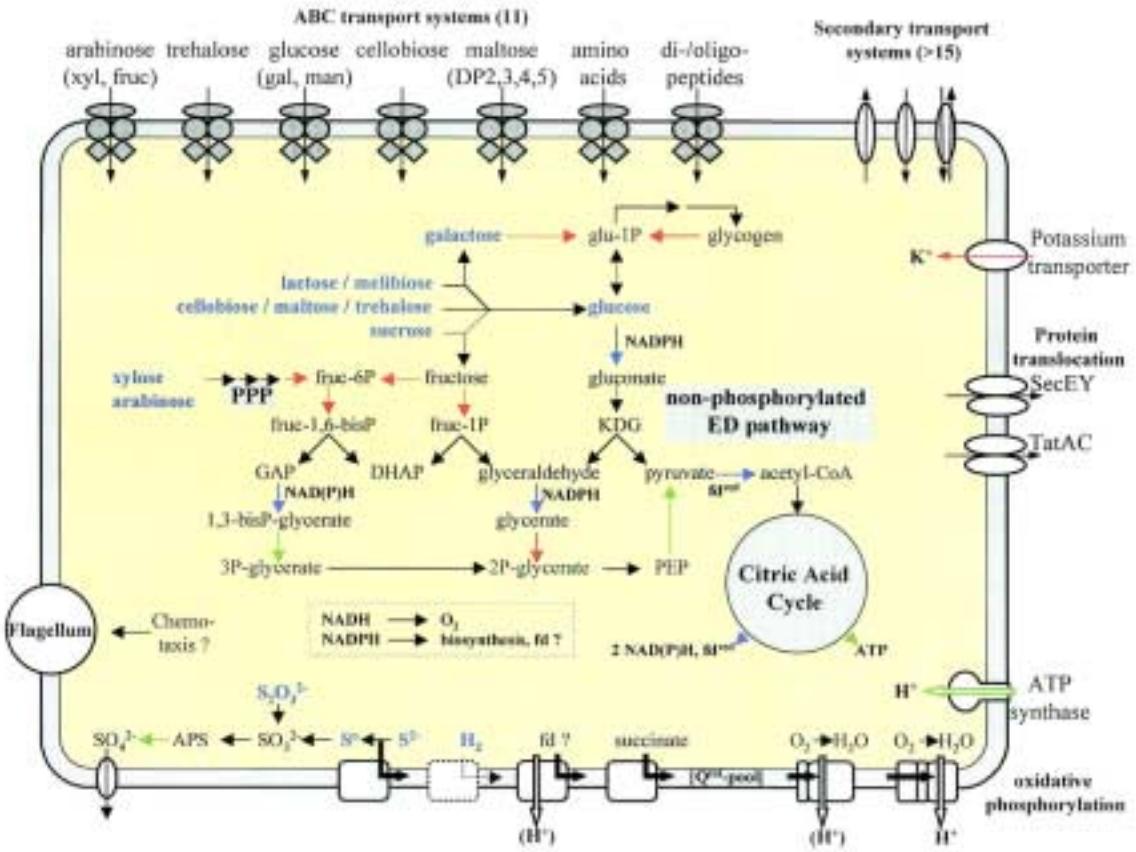


Figure 1-3. Overview of Metabolism and Transport in *Sulfolobus solfataricus*.

Pathways for energy production and carbohydrate catabolism are shown, and extracellular enzymes that hydrolyze polymers (proteases, glycosyl hydrolases) are not shown. Arrows denote the following reactions: chemical conversion (black), energy consuming (red), energy yielding (green), redox (blue), respiratory electron transfer (solid black), proton export (black, open), and import (green, open). Conversions that were anticipated but for which no gene was detected are shown as broken arrows. Eleven operons encoding ABC transport systems are present, and those with established substrate specificity are depicted (xyl, xylose; fruc, fructose; glu, glucose; gal, galactose; man, mannose. DP, degree of polymerization) (17). At least 15 secondary transporters (permeases) are present (symport and antiport). Carbohydrates that are imported and/or support growth of *S. solfataricus* are in blue. All but one of the genes encoding enzymes of the nonphosphorylated Entner-Doudoroff (ED) pathway was identified (KDG, 2-keto-3-deoxygluconate; GAP, glyceraldehyde-3P; DHAP, dihydroxyacetone-P; PEP, phosphoenolpyruvate). Only two genes involved in the pentose phosphate pathway (PPP) were identified (see text). All citric acid cycle enzymes are encoded. Several components of the aerobic respiratory network are identified that are involved in (*i*) reduction of the caldariella-quinone (Q^{cal}) pool: a putative ferredoxin dehydrogenase (see text), succinate dehydrogenase, and (*ii*) oxidation of the Q^{cal} pool: SoxABCD and SoxM terminal oxidases; an ATP synthase converts the proton gradient into ATP; alternative electron donors (in blue) are hydrogen (H_2) and sulfide (S^{2-}), reducing the Q^{cal} pool via hydrogenase and sulfide reductase, respectively. Elemental sulfur and thiosulphate are completely converted to sulfate (APS, adenylylsulfate); some flagellar components are present (see text). Both Sec/signal recognition particle-type and Tat-type protein translocation systems are present (This figure taken directly from She et. al., 2000).

reconstruct the first phosphorylation networks and trace their evolution into the many-fold more complex phosphorylation labyrinths observed in higher eukaryotes such as ourselves. Not only will this lead to a fuller understanding of the protein phosphorylation strategies employed by members of the *Archaea* as a whole, but should provide us with a library of phylogenetically diverse protein kinases and protein phosphatases with which to advance our understanding of the structure-function relationships of these enzymes. By tracing the evolutionary history of these enzymes, we will be able to address some intriguing evolutionary questions concerning the origin and development of protein phosphorylation as a cellular regulatory process. This proposal outlines the first phase of the project, the identification of one of the individual components of the complete protein phosphorylation network. If successful not only will an important piece of the evolutionary puzzle fall into place but we will describe for the first time a protein serine/threonine kinase of archaeal origins in detail. The specific aims of the research delineated herein consist of the following:

1. To purify and characterize an apparent protein-serine/threonine kinase from the membrane-fraction of *Sulfolobus solfataricus*.
2. To identify ORFs from *S. solfataricus* whose predicted products exhibit homology to cAMP dependent protein kinases, including the ORF for the protein kinase in Specific Aim 1.
3. To clone a gene encoding a potential "eukaryotic" protein kinase from *S. solfataricus*. The ORF selected will be that whose predicted product most closely matches the protein examined in Specific Aim 1.
4. To express the potential protein kinase in *Escherichia coli* and evaluate its functional properties *in vitro*.
5. To use MALDI-MS peptide mapping to identify proteins that undergo phosphorylation *in vitro* on serine and threonine from the membrane-fraction of *S. solfataricus*.

CHAPTER II

Materials and Instrumentation

Unless otherwise stated, all chemicals were of analytical or HPLC grade and purchased from Fisher (Pittsburgh, Pennsylvania) or Sigma (St. Louis, Missouri). All radioisotopes were purchased from NEN Research Products (Boston, Massachusetts), except [β-^{32}P] GDP, which was purchased from ICN (Costa Mesa, California). Electronic autoradiography was preformed on a Packard Instant Imager (Meriden, Connecticut). Liquid scintillation counting was preformed using a Beckman Model LS 5801 or LS 6500 Multipurpose liquid scintillation counter, and ultracentrifugation was performed on a Beckman Class H Ultracentrifuge (Fullerton, California). Spectrophotometric measurements were performed on a Hitachi UV-2000 Spectrophotometer. An Accumet Model 10 pH meter was used to measure the pH of solutions. PCR was performed on a PE Applied Biosystems GeneAmp PCR System 9700 with the dual 384-well sample block module (Foster City, California). FPLC was performed with a Pharmacia P-500 pump and LCC-500 Liquid Chromatography Controller (Uppsala, Sweden). A Sorvall Superspeed RC2-B centrifuge with SS-34 rotor or GSA rotor (Newton, Connecticut), or Fisher Scientific Marathon 12Kbr centrifuge (Pittsburgh, Pennsylvania) was used to pellet cells. A Pellicon Cassette System was used to harvest *S. solfataricus* cells. A Power Laboratory Press from American Instruments (Silver Spring, Maryland) or Sonic Dismembrator Model 3000 from Fisher (Pittsburgh, Pennsylvania) was used to lyse cells. Fuji Medical X-Ray film was purchased from Fisher Scientific (Pittsburgh, Pennsylvania). EC4000P Series 90 Programmable power supply from E-C Apparatus Corp. (Holbrook, New York) was used for SDS-PAGE and phosphoamino acid analysis. Pharmacia MultiDrive XL power supply and LKB Multi Temp II thermostatic circulator was used for isoelectric focusing and phosphoamino acid analysis (Uppsala, Sweden). Power Pac 200 from Bio-Rad (Hercules, California), and Electrophoretic Blotting Unit Model EBU-1000 from American BioNuclear (Emeryville, California) were used for western blotting. Electronic Scans were made on a flatbed UMAX Astra 1220S Scanner (Fremont, California). MALDI-MS was performed on a

KOMPACT SEQ MALDI-TOF MS instrument equipped with a nitrogen UV laser (337 nm) from Kratos Analytical (Chestnut Ridge, New York).

Materials for Protein Purification and Enzyme Characterization

Protein assay dye reagent, Mini-PROTEAN II xi electrophoresis cell and preparatory electrophoresis cell Model 491 (16 x 16 cm gels) used for SDS-PAGE, prestained SDS-PAGE low molecular weight protein standards, and prestained SDS-PAGE high molecular weight protein standards were from Bio-Rad (Hercules, California). DE-52 cellulose, P81 phosphocellulose filter paper, and 3MM filter paper were from Whatman (Hillsboro. Oregon). Trichloroacetic acid (10% w/v) was from LabChem Inc (Pittsburgh, Pennsylvania). Myosin light chain (MLC) peptide (KKRAARATSNVFA) and T8S peptide (KKRAARASSNVFA) was synthesized by Genosys (The Woodlands, Texas). Mono-P HR5/20 column (4 ml) and Polybuffer 74 were from Pharmacia (Uppsala, Sweden). Immobiline dry strips, Sephacryl S-200, and chelating Sepharose were from Pharmacia (Piscataway, New Jersey). Reduced carboxyamidomethylated and maleylated lysozyme (RCML) was prepared as described by Tonks et al., 1988. Immobilon P and Centricon YM-10 Centrifugal Filter Devices were from Millipore (Bedford, Massachusetts). Problott Immobilized Membranes were purchased from Applied Biosystems (Rockville, Maryland). Cellulose thin layer plates, 20 x 20 cm, without fluorescent indicator were from Fisher (Pittsburgh, Pennsylvania). ScintiSafe Plus 50% scintillation mixture was from Fisher (Pittsburgh, Pennsylvania). Gel Code glycoprotein stain was from Pierce (Rockford, Illinois). The DIG-Glycan Differentiation Kit was from Roche (Mannheim, Germany). The Glyco-Free Deglycosylation Kit was purchased from Glyko Inc. (Novato, California). Concanavalin A, agarose conjugate column and Genistein were from Calbiochem (San Diego, California). *Galanthus nivalis* lectin insolubilized on 4% cross-linked beaded agarose for column chromatography was from Sigma (St. Louis, Missouri). Sequencing Grade Modified Trypsin was from Promega (Madison, Wisconsin). The protein kinase inhibitors staurosporine, dihydrochloride H7, tamoxifen, hydrochloride ML-9, dihydrochloride, and PKI were purchased from Sigma (St. Louis, Missouri). Molecular weight standards for gel filtration chromatography (blue dextran, δ-globulin, bovine

serum albumin, carbonic anhydrase, and soybean trypsin inhibitor) were purchased from Sigma (St. Louis, Missouri). Anti-Xpress antibody was from Invitrogen (Carlsbad, California) and mouse anti-goat antibody was from Sigma (St. Louis, Missouri). MALDI-MS was performed at the Biochemistry Department, Virginia Tech (Blacksburg, Virginia).

Materials for Molecular Biology

The agarose gel electrophoresis apparatus was from Owl Scientific Inc. (Portsmouth, New Hampshire). *Taq* polymerase PCR kits, restriction enzymes, isopropyl-β-thiogalactopyranoside (IPTG), Wizard PCR preps, Wizard Minipreps, and Wizard DNA Clean-Up Kits were from Promega (Madison, Wisconsin). *Escherichia coli* strain BL21 (DE3) pLysS genotype: F- *ompT hsd*SB (rB - mB -) *gal dcm* (DE3) pLysS (Cam^R), strain TOP 10 F' genotype: F′ {*lacI*^{q} , T*n10*(Tet^R)} *mcrA (mr*r-*hs*dRMS-*mc*rBC) F 8*0lac*Z M15 *.lac* 74 *recA1 deoR araD139 (ara-le*u)7697 *galU galK rpsL* (Str^R) *endA1 nupG*, and pCR T7/NT TOPO linearized plasmid were from Invitrogen (Carlsbad, California). Oligonucleotides for PCR were synthesized by Life Technologies (Rockville, Maryland). *Sulfolobus solfataricus* (ATCC 35092) genomic DNA was purchased from American Type Culture Collection (Manassas, Virginia). DNA sequencing was performed at the DNA Sequencing Facility, Virginia Tech (Blacksburg, Virginia). Miller LB agar and Miller LB broth were purchased from Fisher (Pittsburgh, Pennsylvania), and SOC media was purchased from Life Technologies (Rockville, Maryland). DNase I, Ribonuclease A, PMSF, and egg white lysozyme were purchased from Sigma (St. Louis, Missouri).

General Procedures

Growth of Organisms

Escherichia coli was grown on LB media (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl) with or without antibiotic (Sambrook et. al., 1989). For daily use, *E. coli* was streaked on LB plates with appropriate antibiotic. For long-term storage, an overnight

culture of *E. coli*, 800 µl, was mixed with 200 µl of sterile glycerol (final concentration 20% (v/v) glycerol) and stored at –70°C.

S. solfataricus (ATCC 35091) was grown on media consisting of 1 g/L yeast extract, 1 g/L casamino acids, 3.1 g/L KH_2PO_4, 2.5 g/L $(NH_4)_2SO_4$, 0.25 g/L $CaCl_2 \cdot 2H_2O$, 0.2 g/L $MgSO_4 \cdot 7H_2O$ dissolved in tap water (de Rosa et. al., 1975). The pH of the media was adjusted to 4.0 by adding 2 M H_2SO_4. The organisms were grown in an 8L flask with vigorous aeration at 65-70°C. Cells were harvested every other day, at which point the O.D. $_{600\ nm}$ of the culture generally ranged from 0.5 - 0.8, by concentration using a Pellicon Cassette System followed by centrifugation at 1000 x g, at 4°C for 20 minutes. Cell pellets were stored at –20°C. For long-term storage glycerol permanents of *S. solfataricus* were made by mixing 800 µl of cell culture, grown to an O.D. $_{600nm} \approx$ 0.6-0.8, with 200 µl sterile glycerol (final concentration of 20% v/v) and stored at –70°C.

Solutions for Molecular Biology

The directions for making solutions Luria Broth (LB), LB plates, TE pH 8.0, 0.5M EDTA, stock solutions of ampicillin, stock solutions of chloramphenicol, stock solutions of IPTG, and 6X gel-loading buffer all appear in the appendices of Molecular Cloning, A Laboratory Manual (Sambrook et. al., 1989).

Agarose Gel Electrophoresis of DNA Samples

Samples of genomic or plasmid DNA were routinely resolved using agarose gel electrophoresis as described in Molecular Cloning, A Laboratory Manual (Sambrook et. al., 1989). High melt agarose gels, 0.8 – 1.0% (w/v) depending upon the size range of DNA to be separated, were prepared in TAE buffer (0.04 M Tris-acetate, 1 mM EDTA) containing 0.4% EtBr. DNA samples were prepared by mixing 6 volumes with one volume 6X gel loading buffer. Electrophoresis was conducted at 100 volts in TAE buffer.

Protein Assays

Protein concentration was determined by the method of Bradford assay using Coomassie Protein Assay Reagent from Pierce (Bradford, 1976). Bovine serum albumin

(0-20 μg) in ddH_2O was used to generate a standard curve. The protein concentration in a sample was obtained by diluting the sample such that its absorbance at 595_{nm} fell within the linear range of the standard curve and triplicate spectrometric measurements were taken for each protein sample and averaged.

SDS-PAGE

SDS-polyacrylamide resolving gels, 8%, 10%, 12.5%, and 15% (w/v) acrylamide, 5% (w/v) acrylamide stacking gels, and running buffer (25mM Tris, 192mM glycine, 0.1% (w/v) SDS) were prepared as described by Laemmli (Laemmli, 1970). Protein samples were prepared by mixing 4 volumes with 1 volume 4x-SDS loading buffer consisting of 200 mM Tris pH 6.8, 400 mM DTT, 8% (w/v) SDS (electrophoresis grade), 0.4% (w/v) bromophenol blue, and 40% (v/v) glycerol (Sambrook et. al., 1989). The mixture was heated for 5 minutes at 100°C, briefly centrifuged to collect the sample, and loaded into the wells of the polymerized gel. Bio-Rad prestained low range or high range molecular weight markers were loaded in one well of the gel to facilitate the estimation of the approximate molecular weights of the proteins. Gels were run at a constant current of ≤ 25 mAmp/gel until all prestained standards had entered the running gel, following which current could be increased to ≤ 50 mAmp/gel. Gels were stained with 0.075% (w/v) Coomassie R-250 Brilliant Blue and destained by several washes in 10% (v/v) acetic acid, 20% (v/v) methanol, or 10% (v/v) acetic acid, 50% (v/v) methanol (Fairbanks et. al., 1971). The gels were wrapped in saran wrap, scanned, and stored at room temperature. For protein samples prepared for amino acid sequencing the following exceptions were made: i) proteins were electrophoresed at 15°C on larger 16 x 16 cm SDS-polyacrylamide gels using the preparatory electrophoresis cell Model 491 and ii) the protein band of interest was excised from the gel using a clean razor blade and stored at -20°C.

Western Blotting of Protein onto PVDF Membrane

Following SDS-PAGE or 2-D Electrophoresis, proteins within the SDS-polyacrylamide gel were routinely transferred to Immobilon P or ProBlott PVDF membrane using a wet transfer electroblotting system (Electrophoretic Blotting Unit

Model EBU-1000 from American BioNuclear). The PVDF membrane was wetted with 100% methanol, rinsed 5 times with ddH_2O, and equilibrated for at least 5 minutes in blotting buffer consisting of 10 mM CAPS, pH 11, 10% (v/v) methanol. The polyacrylamide gel containing the electrophoresed proteins was also equilibrated for 1-5 minutes in blotting buffer. Next, the transblotting sandwich was assembled according to the manufacturer's instructions and the protein electroblotted at 40 volts (current limit at 1.0 Amp), at 4°C overnight or 100 volts (current limit set at 1.0 Amp), at 4°C for 1 - 2 hours. Following, the PVDF membrane was removed from the transblotting sandwich and either rinsed extensively with ddH_2O or stained with Coomassie R-250. The SDS-polyacrylamide gel was also removed and stained with Coomassie R-250 to ensure that most or all of the protein had transferred to the PVDF membrane.

Detection of Proteins in SDS-polyacrylamide Gels and on PVDF Membrane

Following SDS-PAGE, 2-D Electrophoresis, or Western Blotting proteins were detected with a conventional staining technique using Coomassie R-250. For SDS-polyacrylamide gels, the gels were stained for 2 hours – overnight following electrophoresis in 0.075% (w/v) Coomassie R-250 in 50% methanol, 10% glacial acetic acid. Next, the gels were destained with 20% (v/v) methanol, 10% (v/v) glacial acetic acid or 50% (v/v) methanol, 10% (v/v) glacial acetic acid until no background Coomassie remained. The gels were then wrapped in Saran wrap, scanned, and stored at room temperature or 4°C. For PVDF membrane, following electroblotting the membrane was immersed in 100% methanol for 10 seconds and then stained for 1 minute with 0.1% (w/v) Coomassie R-250, 20% (v/v) methanol, and 1% (v/v) glacial acetic acid. Next, the membrane was destained with 50% (v/v) methanol until little/no background Coomassie remained, then washed extensively with ddH_2O, dried, scanned, and stored at room temperature.

Procedures for Purifying and Characterizing a Membrane-Associated Threonine Specific Protein Kinase from *Sulfolobus solfataricus*

Two-Dimensional Electrophoresis

Two-dimensional electrophoresis was performed essentially as described by Gorg et. al., 1985, with all material and equipment purchased from Pharmacia (Uppsala, Sweden). Proteins were generally precipitated with acetone to remove salts and detergents, prior to IEF. This was accomplished by adding 3 volumes of ice-cold acetone to a protein sample, incubating at –20^{o}C for 2 - 15 hours, centrifuging at 14000 rpm for 10 minutes in a microcentrifuge, and air drying the protein pellet. The protein was then solubilized by adding 125 ul (for 7 cm IPG strips) rehydration solution consisting of 9M urea, 4% (w/v) CHAPS detergent, 100 mM DTT, trace amounts of bromophenol blue, and appropriate 1x IPG buffer for 1-2 hours at room temperature. This solution was applied to one slot of a Immobiline DryStrip Reswelling Tray, overlaid with an Immobiline IPG dry strips (7 cm pH 4-7 linear or 7 cm pH 3-10 linear) followed by IPG Cover Fluid, and the IPG strip was allowed to rehydrate for a minimum of 10 hours. Next, the rehydrated IPG strip was rinsed with ddH_2O, transferred to the aligner in the Immobiline DryStrip tray with the acidic end (pointed end) towards the red anode. After aligning each electrode across the aligned IPG strip(s), IPG Cover Fluid was applied to the tray to completely cover the IPG strip(s). Isoelectric focusing was conducted on a Multiphor II system using a cooling plate connected to a thermostatic circulator set at 20^{o}C. For IEF of 7cm pH 4-7 linear IPG strips the following protocol was used to program the Pharmacia MultiDrive XL power supply: phase 1, 200v, 2mA, 5W, 1 minute at 1Vh; phase 2, 3500v (voltage was programmed to ramped from 1 to 3500 gradually), 2mA, 5W, 1 hour 30 minutes at 2800Vh; phase 3, 3500v, 2mA, 5W, 1 hour 30 minutes at 5200Vh; total duration of 3 hours 1 minute and 8000Vh. For IEF of 7cm pH 3-10 linear IPG strips the following protocol was used: phase 1, 200v, 2mA, 5W, 1 minute at 1Vh; phase2, 3500v (voltage ramped from 1 to 3500 gradually), 2mA, 5W, 1 hour 30 minutes at 2800Vh; phase 3, 3500v, 2mA, 5W, 1 hour 5 minutes at 3700Vh; total duration of 2 hours 36 minute and 6500Vh. After IEF, the IPG strips were stored in a plastic tube at –70^{o}C if the second-dimension separation did not immediately follow.

Prior to the second-dimension SDS-PAGE separation, the IPG strips were equilibrated for 15 minutes, on a shaker, at room temperature in 5 mL SDS equilibration buffer consisting of 50 mM Tris-Cl, pH 8.8, 6 M urea, 30% (v/v) glycerol, 2% (w/v) SDS, 50 mg DTT, and trace amounts of bromophenol blue. A second equilibration was performed on a shaker for 15 minutes at room temperature in 5mL SDS equilibration buffer (without DTT) containing 250 mg iodoacetamide. Following, equilibration buffer was drained and the IPG strip was applied to the top of a SDS-polyacrylamide resolving gel (8-12.5% (w/v) acrylamide), sealed in place with 0.5 % (w/v) agarose, and electrophoresed according to the SDS-PAGE procedure described in Material and Methods.

In Gel Assays of Protein Kinase Activity, Nucleotide Specificity, and Autophosphorylation

The protein kinase activities of individual polypeptides were assayed in gel following their resolution by SDS-PAGE or 2-D electrophoresis as described by Bischoff and Kennelly (Bischoff and Kennelly, 1999). The running of the SDS-polyacrylamide gel was prepared as described previously with the exception that 0.5 mg/ml casein was added prior to initiation of polymerization. Following electrophoresis, SDS was removed by washing the gel two times in 40 ml of 20% (v/v) isopropanol, 20 mM MES, pH 6.5 for 1 hour each. Next, the gel was washed one time in 40 ml of 20 mM MES, pH 6.5 for 1 hour to remove the isopropanol. The conformation of proteins within the gel were randomized by soaking the gel in 40 ml 6 M guanidine hydrochloride, 20 mM MES, pH 6.5 for 2-4 hours. Next, the proteins within were allowed to renature by incubating the gel with 5 changes in 40 ml 20 mM MES, pH 6.5, containing 0.1% (v/v) Triton X-100, 5 mM $MnCl_2$, 5 mM $MgCl_2$, 1 mM DTT for 1 hour each for the first 4 washes, and overnight for the final wash. An "in gel" assay was then performed by incubating the gels for 1 hour at 25^{o}C in renaturation buffer containing, 50 μM ATP (15 μCi of [γ-^{32}P] ATP / ml), or 50 μM GTP (15 μCi of [γ-^{32}P] GTP / ml), or 50 μM GDP (15 μCi of [β-^{32}P] GDP / ml). The gels were washed extensively in 2% (w/v) sodium pyrophosphate, 20 mM MES, pH 6.5 to removed excess NTP, wrapped in saran wrap, and the locations of regions in which phosphorylation of immobilized casein occurred were determined by

autoradiography with X-ray film and/or by electronic autoradiography. The quantity of [^{32}P] phosphate transferred to the casein was quantified by electronic autoradiography. Autophosphorylation of polypeptides "in gel" was determined as described above with the exceptions that no exogenous phosphoacceptor protein (e. g. casein) was copolymerized into the SDS-polyacrylamide, the "in gel" assay was performed at 65^{o}C instead of 25^{o}C, and the SDS-polyacrylamide gel was stained with Coomassie R-250 following the "in gel" assay.

Nucleotide specificity was determined by preincubating the gels with nonlabeled nucleotides following the renaturation step (Bischoff and Kennelly, 1999). This was done by incubating the gels overnight at 25^{o}C in 40 ml of renaturation buffer containing one of following unlabeled phosphate donors, each at a final concentration of 50 μM: ATP, ADP, AMP, GTP, GDP, GMP, CTP, UTP and NaPPi as a negative control. After this overnight preincubation, the gels were washed 5 times for 30 minutes each in 40 ml of renaturation buffer to remove unlabeled nucleotide donor. The gels were then incubated for 1 hour at 25^{o}C in renaturation buffer containing, 50 μM ATP (15 μCi of [γ-^{32}P] ATP / ml) as described above for "in gel" assays. The quantity of radioactivity incorporated into casein was then determined by electronic autoradiography.

Solution Assay of Protein Kinase Activity

Protein kinase activity was routinely assayed in solution using a filter paper method (Walsh et. al., 1972). Samples containing the protein kinase were incubated at 25^{o}C, 37^{o}C, or 65^{o}C in a volume of 50 μl or 100 μl of 20 mM MES, pH 6.5, containing 50 μM ATP (300-1000 μCi of [γ-^{32}P] ATP / ml), 2 mM DTT, 5 mM $MnCl_2$, 5 mM $MgCl_2$, either 0.1% (v/v) Triton X-100 or 12.5 mM octyl glucoside, and a phosphoacceptor substrate such as casein or MLC peptide. Reactions were initiated by addition of ATP. Following incubation for periods up to 60 minutes, the reaction was terminated in one of two manners. When a protein or amino acid copolymer (e.g. poly Glu:Tyr) substrate was employed, a 33% portion of each reaction was spotted onto a 2 x 2 cm square of Whatman 3MM paper and immediately washed in 10% (w/v) TCA containing 4% (w/v) NaPPi for 20 minutes, followed by 4 washes in 10% (w/v) TCA containing 2% (w/v) NaPPi each for periods of 20 minutes – overnight. When a peptide

substrate (e.g. MLC peptide or T8S peptide) was employed, a 33% portion of each reaction was spotted onto a 2 x 2 cm square of Whatman P81 phosphocellulose paper and immediately washed in 150 mM H_3PO_4 5 times each for periods of 20 minutes – overnight. Following the wash steps, the filter paper was dried and the quantity of [^{32}P] phosphate immobilized determined by liquid scintillation counting in 1 ml of ScintiSafe Plus 50% scintillation fluid using a liquid scintillation counter.

Phosphoamino Acid Analysis

Phosphoamino acid analysis was performed essentially as described by Kamps and Sefton (Kamps and Sefton, 1989). Radiolabeled proteins were isolated by SDS-PAGE or 2-D electrophoresis and transferred to Immobilon P membrane as described above. The ^{32}P-labeled proteins were located using autoradiography, excised, and cut into approximately 1 mm x 1mm squares. These membrane pieces were wet with 100% methanol and rinsed 5 times with ddH_2O, making sure to remove all the ddH_2O before proceeding. Next, the wet membrane squares were incubated in 6.05 N HCl for 1 hour at 100^{o}C. Following acid hydrolysis, the supernatant fluid was removed and the liquid evaporated with heat (approximately 60^o – 70^o C) under a gentle stream of N_2 gas. The hydrolysate was then dissolved in 5 – 10 µl pH 1.9 (50 ml 88% formic acid, 156 ml glacial acetic acid, adjusted to 2 L with ddH2O) or pH 3.5 (100 ml glacial acetic acid, 10 ml pyridine, adjusted to 2 L with ddH_2O) buffer by agitating on vortex mixer.

Two-dimensional thin layer electrophoresis was used to resolve the phosphoamino acids in the hydrolysate. Xylene cyanole, 0.5 µl of 1 mg/ml stock solution, was applied to a 20 x 20 x 0.1 cm cellulose thin layer chromatography plate without fluorescent indicator. By dividing the plate into 4 10 x 10 cm sections, up to 4 samples were accommodated at one time. One spot was placed 2 cm from the bottom and 6 cm from the left side and a second spot was placed 2 cm from the bottom and 8 cm from the left side for sample number 1; one spot was placed 2 cm from the bottom and 12 cm from the left side and a second spot was placed 2 cm from the bottom and 14 cm from the left side for sample number 2; one spot was placed 10 cm from the bottom and 6 cm from the left side and a second spot was placed 10 cm from the bottom and 8 cm from the left side for sample number 3; and one spot was placed 10 cm from the bottom and 12 cm

from the left side and a second spot was placed 10 cm from the bottom and 14 cm from the left side for sample number 4. A 1 µl mixture of 5 mM each phosphoserine, phosphothreonine, and phosphotyrosine was mixed with each sample and then this was applied in 0.5 µl aliquots 2 cm from the bottom and 7 cm from the left side for sample number 1; 2 cm from the bottom and 13 cm from the left side for sample number 2; 10 cm from the bottom and 7 cm from the left side for sample number 3; and 10 cm from the bottom and 13 cm from the left side for sample number 4. Between each application, the spot was dried with an air gun. Just prior to electrophoresis, samples were focused on the TLC plate by wetting the plate with damp Whatman 3MM filter paper equilibrated in either pH 1.9 buffer (1st dimension) or pH 3.5 buffer (2nd dimension).

Two-dimensional thin layer electrophoresis was conducted on a Multiphor II system using a cooling plate connected to a thermostatic circulator set at 7^{o}C. The phosphoamino acids were separated in the first dimension in pH 1.9 buffer at 30 volts for 5 minutes, 60 volts for 2 minutes, 125 volts for 2 minutes, and 250 volts for 176 minutes. The current limit was set at 100 mAmps and the power limit set at 200 watts. The TLC plate then was removed, dried, turned 90^{o} clockwise and placed back on the Multiphor II plate. The phosphoamino acids were separated in the second dimension in pH 3.5 buffer at 30 volts for 5 minutes, 60 volts for 2 minutes, 125 volts for 2 minutes, and 250 volts for 76 minutes. The current limit was set at 100 mAmps and the power limit set at 200 watts. In order to visualize the three phosphoamino acid standards, the plate was dried with an air gun, sprayed with 0.2% (w/v) ninhydrin in acetone, and baked for 5 minutes at 80^{o}C. The locations of the ^{32}P-labeled phosphoamino acids from the protein samples were visualized by autoradiography.

Glycoprotein Staining Using GelCode Glycoprotein Stain

Glycoprotein was detected following SDS-PAGE using a GelCode Glycoprotein Staining Kit specifically designed to detect the sugar moieties of glycoproteins. Glycoprotein staining reagents and protocol were provided by the manufacturer (Pierce). Each step was carried out at room temperature and with gentle agitation. Following electrophoresis, proteins within the SDS-polyacrylamide gel were fixed in 100 ml of 50% (v/v) methanol for 30 minutes. The gel was washed two times in 100 ml 3% (v/v) acetic

acid for 10 minutes each and then immersed in 25 ml of Oxidation Solution for 15 minutes. The gel was then washed three times in 100 ml 3% (v/v) acetic acid for 5 minutes. Next, the gel was immersed in 25 ml of GelCode Glycoprotein Staining Reagent for 15 minutes and then transferred to 25 ml of Reduction Solution for 5 minutes. Finally, the gel was washed extensively in 3% (v/v) acetic acid and then with ddH_2O. Glycoproteins were seen as magenta bands. Gels were wrapped in Saran wrap, scanned, and stored at room temperature.

Characterization of Carbohydrate Moieties of Glycoproteins

A DIG Glycan Differentiation Kit was used to characterize carbohydrate moieties of glycoproteins bound to PVDF membrane. This kit uses lectins, conjugated with steroid hapten digoxigenin, that bind to specific carbohydrate moieties of glycoproteins allowing for the immunological identification of the carbohydrate structure. DIG Glycan reagents and protocol was provided by the manufacturer, Boehringer Mannheim. All steps were carried out at room temperature and with gentle agitation, except where noted. Following Western Blotting, the PVDF membrane was incubated in 20 ml Blocking Solution for 30 minutes or overnight at 4^oC. The membrane was then washed twice for 10 minutes each in 50 ml TBS (0.05 M Tris HCl, 0.15 M NaCl, pH 7.5) and once for 10 minutes in TBS containing 1 mM $MgCl_2$, 1 mM $MnCl_2$, and 1 mM $CaCl_2$. Next, the membrane was divided into 5 sections (each containing one lane of prestained low range molecular weight markers and one lane into which the protein sample had been electrophoresed) and each section incubated with one of five different lectin solutions (10 µl Galanthus nivalis agglutinin, GNA; 10 µl Sambucus nigra agglutinin, SNA; 50 µl Maackia amurensis agglutinin, MAA; 100 µl Peanut agglutinin (PNA); or 10 µl Datura stramonium agglutinin, DSA) in 10 ml TBS containing 1 mM $MgCl_2$, 1 mM $MnCl_2$, and 1 mM $CaCl_2$ for 1 hour. The membranes were washed 3 times for 10 minutes each in 50 ml TBS and then incubated with 10 µl anti-digoxigenin-AP in 10 ml TBS for 1 hour. Glycoproteins were visualized as a dark/light purple bands by incubating the membranes in 10 ml Tris buffer, pH 9.5 containing 200 µl NBT/X-phosphate solution without shaking for 1 – 5 minutes. Washing the membrane extensively with several changes of

ddH_2O stopped color development. The membranes were then dried, scanned, and stored in the dark at room temperature.

Chemical Deglycosylation of Glycoproteins

Sulfolobus solfataricus glycoproteins were chemically deglycosylated using a GlycoFree Deglycosylation Kit from Glyko, which employs anhydrous trifluoromethanesulfonic acid (TFMS) to chemically cleave both N- and O-linked glycans in a non-selective manner. The manufacturer provided the protocol and reagents. Proteins were precipitated using ice cold acetone to remove salts and detergents, prior to deglycosylation. This was accomplished by adding 3 volumes of ice-cold acetone to a protein sample, incubating at –20°C for 2 - 15 hours, centrifuging at 14000 rpm for 10 minutes in a microcentrifuge, and air drying the protein pellet. Proteins were then resuspended in 0.25% (w/v) SDS with heat (65°C) and agitation on a vortex mixer, and stored at –20°C for ≥ 2 hours. Next, proteins were lyophilized at ≤ 0.5 milliTorr for ≥ 18 hours. The lyophilized protein was transferred to a 1 ml glass vial capped with a Teflon-faced seal and placed in a dry ice/ethanol cold bath for 20 seconds. Immediately prior to use, 60 µl of Reagent B (toluene) was added to 1 vial of Reagent A (TFMS). Using a glass syringe, 50 µl of this TFMS / toluene reagent was added to the lyophilized protein sample by piercing the Teflon-lined septum of the vial with the syringe needle and allowing the reagent to run slowly down the side of the reaction vessel over a period of 15 – 20 seconds. The needle was withdrawn and the vial was left in the cold bath for an additional 10 seconds, then placed in a –20°C freezer for 4 hours. At 5 and 10 minutes, the contents of the vial were briefly agitated and examined to ensure that the protein mixture had melted and was homogenous throughout. Following this step the reaction vial was removed from the freezer, its teflon septum removed, and placed in a dry ice/ethanol cold bath for 20 seconds. Using a micropipette, 150 µl of Reagent C was added to the vial by allowing the reagent to flow down the side of the vial over a period of 15 – 20 seconds. The vial was left in the cold bath for an additional 20 seconds, then transferred to dry ice for 5 minutes, then to wet ice for a further 15 minutes. Finally, 400 ul of 0.5% (w/v) ammonium bicarbonate was added to the vial to neutralize excess TFMS.

Procedure for Identifying Proteins from the Membrane Fraction of *Sulfolobus solfataricus* that Underwent Phosphorylation on Threonine *In Vitro*

In Gel Trypsin Digestion for MALDI-MS Analysis

The Coomassie stained protein band/spot of interest was excised from a SDS-polyacrylamide gel with a clean razor blade, chopped into approximately 1 x 1 mm^2 pieces, and placed into a 1.5 ml eppendorff tube. A gel slice from a protein free region of the gel was also excised for a control digestion and processed in parallel to facilitate the identification of trypsin autoproteolysis fragments. The total gel volume was estimated for each excised section. Coomassie R-250 dye was removed from the stained proteins by incubating the gel slices in 100 µl of 25 mM ammonium bicarbonate, containing 50% (v/v) acetonitrile and agitated on a vortex mixer for 10 minutes. This step was repeated up to 4 times, removing the wash solution each time with a gel loading pipet tip until the gel slices appeared shrunken and white/light blue in color. The gel slices were then dried for approximately 30 minutes in a vacuum centrifuge.

The protein was reduced by adding a sufficient volume of 10 mM DTT in 25 mM ammonium bicarbonate to cover the gel slices, then incubating for 1 hour at 56^{o}C. The solution was cooled to room temperature and the supernatant liquid removed with a gel loading pipet tip. Next, the protein was alkylated by adding roughly the same volume of 55 mM iodoacetamide in 25 mM ammonium bicarbonate to the eppendorff tube and incubating in the dark at room temperature for 45 minutes with occasional agitation using a vortex mixer. The iodoacetamide solution then was removed with a gel loading pipet tip and the gel slices washed with approximately 100 µl of 25 mM ammonium bicarbonate, pH 8 for 10 minutes with continuous agitation on a vortex mixer. This solution was removed with a gel loading pipet tip and the gel slices dehydrated with approximately 100 µl 25 mM ammonium bicarbonate, and 50% (v/v) acetonitrile for 10 minutes with continuous mixing on a vortex mixer. This rehydration/dehydration step was repeated one additional time. The supernatant liquid was removed with a gel loading pipet tip and the gel pieces were dried for approximately 30 minutes in a vacuum centrifuge.

The dried gel slices were then rehydrated in 1 gel volume of 0.1 mg/ml sequencing grade modified trypsin dissolved in 25 mM ammonium bicarbonate, pH 8 and then agitated for 5 minutes using a vortex mixer. This mixture was incubated for 12 – 16 hours at 37°C. Next, 2 gel volumes of ddH_2O were added and the mixture agitated for 5 minutes on a vortex mixer. A gel loading pipet tip was used to remove the supernatant liquid, which was transferred to a new 1.5 ml eppendorph tube. Two additional peptide extractions were performed on the gel slices by adding 2 gel volumes of 5% (v/v) TFA, 50% (v/v) acetonitrile, and agitating as before for 5 minutes. Each time the supernatant liquid was removed with a gel loading pipet tip and pooled with the previous washings. The pooled washings, which contained the extracted tryptic peptides, were concentrated to a final volume of approximately 10 µl in a vacuum centrifuge. The final volume was brought up to approximately 25 µl by adding 5% (v/v) TFA, and 50% (v/v) acetonitrile, and the solution stored at –20°C. For MALDI-MS analysis, 0.5 µl of the peptide solution was mixed with 0.5 µl α-hydroxycinnamic acid in 0.1% (v/v) TFA, 50% (v/v) acetonitrile, spotted on the sample target, and allowed to air dry.

Procedures for Cloning and Expressing the Gene Encoding a Protein Serine Kinase From *Sulfolobus solfataricus*

Polymerase Chain Reaction

The reaction mixture for DNA amplification via PCR contained 550 ng of *S. solfataricus* genomic DNA or 20 ng of plasmid DNA, 6 pmol of each primer, 200 µM each dATP, dCTP, dGTP, and dTTP, 2.5 mM $MgCl_2$, 2 µl 10x reaction buffer (Promega), and 2.5 units *Taq* polymerase in a total volume of 20 µl. The thermocycler used was a PE Applied Biosystems GeneAmp PCR System 9700 with the dual 384-well sample block module. For amplification of genomic DNA to obtain the entire *S. solfataricus* protein kinase gene, the thermocycler was programmed for an initial 2 minute 30 second incubation period at 94°C to denature the genomic DNA. This was followed by a cycle of denaturing at 92°C for 15 seconds, annealing at 45°C for 30 seconds, and elongation at 72°C for 30 seconds. For the next 29 cycles the elongation time was increased by 1 second / cycle. A final elongation step at 72°C for 7 minutes was done to ensure that all PCR products were full length and 3' adenylated. To obtain the approximately 850 bp

fragment of the *S. solfataricus* protein kinase gene, plasmid DNA (pCR T7/NT TOPO plasmid containing the full length *S. solfataricus* protein kinase gene) was used as template rather than genomic DNA. The thermocycler was programmed for an initial 2 minute 30 second incubation period at 94°C to denature the plasmid DNA. This was followed by a cycle of denaturing at 92°C for 15 seconds, annealing at 50°C for 30 seconds, and elongation at 72°C for 30 seconds. For the next 29 cycles the elongation time was increased by 1 second / cycle. A final elongation step at 72°C for 7 minutes was done to ensure that all PCR products were full length and 3' adenylated.

pCR T7/NT TOPO Cloning Reactions and Transformations

Cloning and transformation reactions were performed using a TOPO TA cloning kit as described by the manufacturer, Invitrogen. These kits provide a highly efficient, one-step strategy for the direct insertion of *Taq* polymerase-amplified PCR products into a linearized plasmid. Briefly, Taq polymerase has a nontemplate-dependent terminal transferase activity that adds a single deoxyadenosine to the 3' ends of PCR products. The linearized vector, pCR T7/NT TOPO, has overhanging 3' deoxythymidine residues. When the two are mixed at room temperature the PCR product ligates efficiently with the vector (Figure 2-1).

Aliquots of fresh PCR reaction mixture (see above), 2 – 4 µl, 1 µl of pCR T7/NT TOPO linearized vector, and 1 µl salt solution (1.2 M NaCl, 0.06 M $MgCl_2$) were brought to a final volume of 6 µl with sterile ddH_2O (ligation mixture) and incubated for 10 – 30 minutes at room temperature. Immediately thereafter, transformations were performed by adding 2 µl of the ligation mixture to one vial of thawed One Shot Chemically Competent *E. coli* cells on ice. The mixture was incubated on ice for 30 minutes, then the cells were heat-shocked for 30 seconds at 42°C without shaking. Next, 250 µl of room temperature SOC media was added, the tubes capped tightly, and the cells were agitated in a horizontal position at 200 rpm, 37°C, for 30 minutes. The transformed cells, 10 – 50 µl, were plated onto prewarmed LB Amp (100 µg/ml) plates and incubated overnight at 37°C.

DNA Sequencing

DNA Sequencing was done at the Virginia Tech DNA Sequencing Facility (Blacksburg, Virginia). Plasmids used as templates for DNA Sequencing were isolated from 3 – 5 ml overnight cultures of *E. coli* grown in LB Amp (100 μg/ml) media. Plasmids were purified using Promega Wizard Minipreps according to the manufacturers instructions. For each sequencing reaction, 500 ng of plasmid DNA and 4 pmol of primer were provided to the DNA Sequencing Facility. The gene encoding the full-length *S. solfataricus* protein kinase and the 850 bp piece of the *S. solfataricus* protein kinase gene were sequenced in the forward direction using the T7 promoter primer and in the reverse direction using the pRSET reverse primer (Figure 2-1).

Expression of Recombinant Protein in E. coli

Expression of proteins from their cloned DNA in *E. coli* was carried out according to the manufacturer, Invitrogen. Briefly, 1 μl (approximately 50 - 100 ng) of pCR T7/NT TOPO plasmid containing the cloned protein kinase gene from *S. solfataricus* was added to 1 vial of BL21(DE3)pLysS One Shot Chemically Component *E. coli* cells on ice. Following a 30-minute incubation on ice, the cells were heat-shocked for 30 seconds at 42^{o}C without shaking. Next, 250 μl of room temperature SOC media was added, tubes capped tightly, and the cells were agitated in a horizontal position at 200 rpm, 37^{o}C, for 30 minutes. The entire transformation reaction was added to a 125 ml flask containing 10 ml LB, Amp (100 μg/ml), and chloramphenicol (34 μg/ml) and grown overnight at 37^{o}C with shaking (225 rpm). The following day glycerol permanents were made from the overnight cultures and stored at –70^{o}C for future use. In a 1 L flask, 4 ml of overnight culture was used to inoculate 200 ml of LB containing 100 μg/ml Amp and 34 μg/ml chloramphenicol. The culture was grown at 37^{o}C and 225 rpm until it reached an $O.D._{600nm} \approx 0.5 – 0.8$ (approximately 3 hours). Overexpression of the recombinant protein was induced by addition of IPTG, 200 μl of 1 M stock solution, to a final concentration of 1 mM. After 4 hours of induction, a 100 μl aliquot of the culture was centrifuged at 14,000 rpm for 2 minutes in a microcentrifuge. The cell pellet was then lysed by adding 100 μl 4X SDS-loading buffer and the protein composition analyzed on either a 10% (w/v) or 12% (w/v) SDS-polyacrylamide gel. The remainder of the cells

were also collected by centrifugation at 1000 x g, 4°C for 10 minutes, and stored at –20°C.

Detection of Recombinant Proteins Expressed with pCR T7/NT TOPO

Recombinant protein expressed from pCR T7/NT TOPO plasmid was detected immunologically using Anti-Xpress antibody (Figure 2-1). Following electrophoretic transfer, the Immobilon-P membrane was soaked in 100% methanol for 10 seconds and then allowed to dry at room temperature for at least 4 hours. The membrane was then placed in a tray containing Anti-Xpress antibody diluted 1:5000 in 10 ml TBST + BSA (20 mM Tris base, 140 mM NaCl, 0.05% (w/v) Tween-20, 1% (w/v) BSA, pH 7.5) with gentle agitation for 1 - 12 hours. Next, the membrane was rinsed two times for several seconds each in 10 ml TBST (20 mM Tris base, 140 mM NaCl, 0.05% (w/v) Tween-20, pH 7.5), and then placed in a tray containing Mouse Anti-Goat alkaline phosphatase-conjugated secondary antibody diluted 1:2500 in 10 ml TBST + BSA and gently agitated for 1 hour. Next, the membrane was washed one time in TBST for 2 minutes, twice for 2 minutes each in TBS (20 mM Tris base, 140 mM NaCl, pH 7.5), and once briefly in alkaline phosphatase buffer (100 mM diethanolamine, 100 mM NaCl, 5 mM $MgCl_2$, pH 9). Finally, recombinant protein was visualized as a dark/light purple band by incubating the membrane in 10 ml alkaline phosphatase buffer containing 66 µl nitro blue tetrazolium stock solution (50 mg/ml), and 33 µl bromochloroindolyl phosphate stock solution (50 mg/ml) until the color developed (1 – 5 minutes). Washing the membrane extensively with several changes of ddH_2O stopped color development. The membrane was then dried, scanned, and stored in the dark at room temperature.

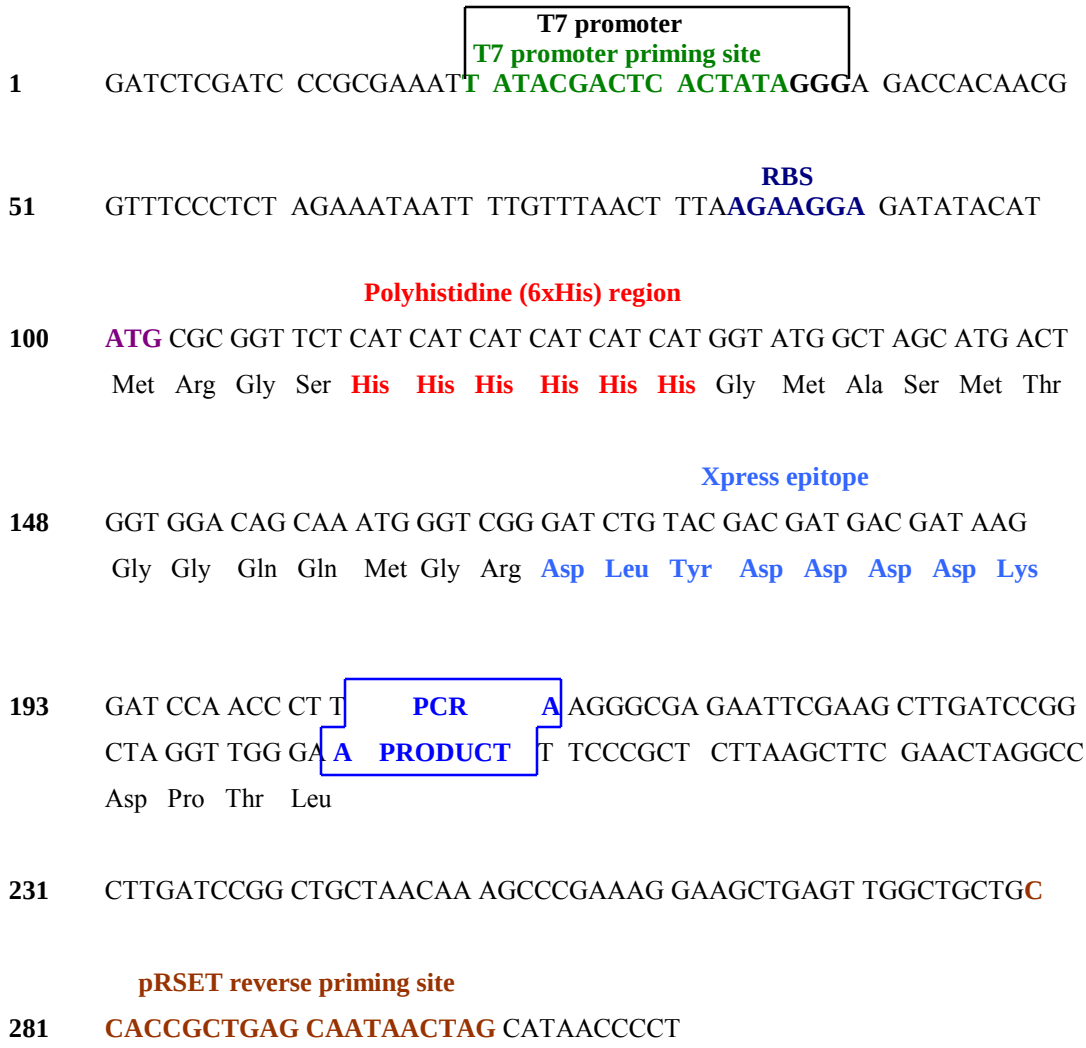


Figure 2-1. pCR T7/NT TOPO Vector.

Shown is a diagram of the pCR T7/NT TOPO vector complete with **T7 priming site**, **pRSET reverse priming site**, **polyhistidine (6xHis) region**, **Xpress epitope**, **ATG start codon**, **ribosomal binding site** (**RBS**), **T7 promoter**, and ***Taq* polymerase-amplified PCR product ligation site**.

CHAPTER III

Identification, Partial Purification and Characterization of a Membrane-Associated Protein Kinase From *Sulfolobus solfataricus*

Identification of a Membrane-Associated Protein Kinase From Sulfolobus solfataricus

Not long ago the first archaeal protein-serine/threonine phosphatase, PP1-arch 1, was characterized and its gene cloned from *S. solfataricus* (Kennelly et al., 1993; Leng et al., 1995). This was an important discovery. It suggested that regulatory protein phosphorylation occurs in *S. solfataricus* and established the presence of "eukaryotic" signal transduction proteins in the *Archaea*. Subsequently, two other homologous protein phosphatases, PP1-arch2 from *Methanosarcina thermophila* (Solow et al., 1997) and Py-PP1 from *Pyrodictium abyssi* (Mai et al., 1998) were cloned. That *S. solfataricus* and other members of the *Archaea* possess "eukaryote-like" protein phosphatases, and the recent discovery of ORFs encoding potential "eukaryote-like" protein kinases and protein phosphatases in several archaeons (Leonard et al., 1998; Shi et al., 1998; Smith and King, 1995; Stravopodis and Kyrpides, 1999) predicates the existence of similar protein kinases in *S. solfataricus*.

Dr. Kenneth M. Bischoff originally addressed this premise by attempting to detect protein kinase activity from whole-cell lysate of *S. solfataricus* using an "in gel" protein kinase assay as described in Methods. In brief, whole-cell lysate was electrophoresed in SDS-polyacrylamide gels copolymerized with casein, the proteins within denatured-renatured with guanidine hydrochloride, the gel incubated with [γ-^{32}P] ATP and potential divalent metal cofactors Mn^{2+} and Mg^{2+}, and the remaining radiolabeled ATP thoroughly washed away. Autoradiography revealed the presence of one prominent ^{32}P-labeled band. This ^{32}P-labeled band presumably arose from the covalent phosphorylation of the copolymerized casein accessible to a protein kinase from *S. solfataricus* that had been electrophoresed into the SDS-polyacrylamide gel. The estimated molecular mass of the putative protein kinase was ≈ 85 kDa (Lower et. al., 2000).

In order to further characterize this protein kinase activity and determine its cellular location, Dr. Bischoff performed a second "in gel" protein kinase assay on

partially purified protein fractions from *S. solfataricus*. The extract and detergent soluble extract of the membrane fraction were electrophoresed into a SDS-polyacrylamide gel copolymerized with either casein or BSA. The gel copolymerized with BSA was used to distinguish phosphotransfer to casein from other mechanisms of phosphate incorporation, such as autophosphorylation of the protein kinase itself or the trapping of a phosphoenzyme intermediate. As observed in Figure 3-1, a conspicuous band of phosphorylation at ≈ 85 kDa was evident for the detergent solubilized membrane fraction in the gel copolymerized with casein, while little if any phosphorylation is observed for the soluble extract. For the gel copolymerized with BSA, little if any phosphorylation is detected for either of the *S. solfataricus* protein fractions. These results suggested that *S. solfataricus* harbors a protein kinase activity and that the enzyme responsible appears to be associated with the cellular membrane. In later "in gel" assay experiments, using more highly purified enzyme fractions from *S. solfataricus*, the polypeptide species responsible for the protein kinase activity migrated at a molecular mass of ≈ 67 kDa rather that ≈ 85 kDa. We hypothesized that this discrepancy in molecular mass was due to the presence of large quantities of nonionic detergent (Triton X-100) required to solubilize the polypeptide in the early experiments, thus precluding complete coating of the polypeptide with SDS thereby retarding migration during SDS-PAGE.

A defining characteristic of protein kinases is that they catalyze the formation of a covalent protein-phosphate bond. In order to determine whether a covalent casein-phosphate bond was being formed, Dr. Bischoff incubated a detergent extract from *S. solfataricus* with casein and [γ-^{32}P] ATP for 1 hour. Next, the mixture was subjected to SDS-PAGE, transferred to a PVDF membrane, and the [^{32}P] phosphorylated species visualized by autoradiography. A [^{32}P] phosphorylated band whose apparent M_r corresponded to that predicted for casein was observed, indicating that the [^{32}P] phosphate had been transferred to the phosphoacceptor substrate. Subsequent PAA analysis on the [^{32}P] phosphorylated casein revealed that the [^{32}P] phosphate was bound to threonine verifying that the factor detected in the detergent soluble membrane extract of *S. solfataricus* displayed the essential characteristic of a protein kinase (Lower et. al., 2000).

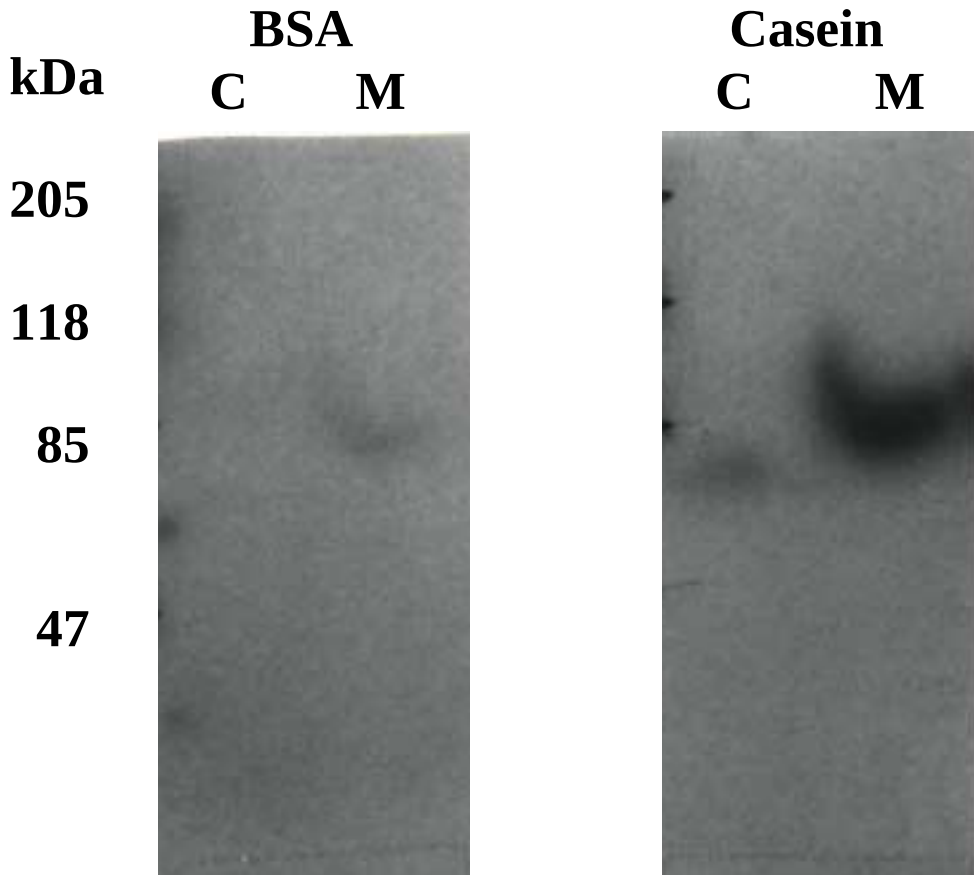


Figure 3-1. Detection of Protein Kinase Activity in the Membrane Fraction of *Sulfolobus solfataricus*.

Shown are the results of an in gel assay of *S. solfataricus* protein kinase activity following SDS-PAGE. Portions, containing 20 μg of protein each, of the cytosolic fraction (**C**) and the detergent extract of the membrane fraction (**M**) from *S. solfataricus* were electrophoresed into an SDS-10% (w/v) polyacrylamide gel in which either BSA (left) or casein (right) had been copolymerized into the gel matrix at a concentration of 1 mg/ml. Following electrophoresis, the gel was washed free of SDS and the proteins within it were renatured using guanine hydrochloride. The gels were each then incubated for 60 minutes in 30 ml of 20 mM MES, pH 6.5, containing 0.5 mM EDTA, 5 mM Mg^{2+}, 5 mM Mn^{2+}, and 50 μM [γ-^{32}P] ATP (specific activity, 30 Ci/mol). The gels then were washed to remove free ATP, and regions of each gel in which [^{32}P] phosphate was incorporated into protein were visualized by autoradiography using X-ray film. Shown are the autoradiograms of each gel with the positions of the molecular weight markers of the protein standards indicated at the left (Lower et. al., 2000).

Solubilization and Partial Purification of the Sulfolobus solfataricus Protein Kinase

Attempts to purify the *S. solfataricus* protein kinase met with limited success. Efforts to extract the protein kinase activity using high salt (0.5 M NaCl) and/or moderately acidic conditions (pH 5.0) failed. Solubilization of protein kinase activity required the presence of nonionic detergents such as Triton X-100 or octyl glucoside. Following solubilization, detergent also was required to maintain catalytic activity. We therefore concluded that the *S. solfataricus* protein kinase was an integral membrane protein.

Purification of the *S.* solfataricus protein kinase was performed as described in Lower et. al., 2000 (Table 3-1). Frozen *S. solfataricus*, 20 g wet weight, was thawed and resuspended in 40ml of 20 mM MES, pH 6.5, containing 0.5 mM EDTA, 1 mM PMSF, and 10 μg / ml DNase I. PMSF was made as a 100X stock solution in isopropanol and added on the day of use. The cells were lysed at 4°C by two passages through a French pressure cell at 12,000 lb/in^2. The lysate was centrifuged at 1,000 x g for 10 minutes at 4°C to remove debris and any remaining whole cells.

The cell extract was centrifuged at 100,000 x g for 75 minutes at 4°C to separate the membrane fraction from the soluble fraction. The resulting pellet was washed, and integral membrane proteins were solubilized by a modification of the procedure of Ramwani and Mishra (Ramwani and Mishra, 1986). First, the membrane fraction was resuspended by pipeting and agitation on a vortex mixer in 40 ml solution of 20 mM sodium acetate, pH 5.0, containing 0.5 M NaCl, in order to liberate peripherally associated proteins. The particulate was collected by centrifugation at 100,000 x g for 60 minutes at 4°C and the supernatant liquid (peripheral fraction) discarded. The pellet was then resuspended in 20 mM MES, pH 6.5, containing 125 mM NaCl and 0.4% (v/v) Triton X-100 (or 25 mM octyl glucoside). Use of octyl glucoside was preferred because of its higher critical micelle concentration (CMC). The mixture was centrifuged at 100,000 x g for 60 minutes at 4°C, and the supernatant solution retained as the detergent extract.

The detergent extract was diluted approximately 5-fold by the addition of 20 mM MES, pH 6.5. The diluted extract was applied onto a 1.5 x 18 cm column of DE-52

Fraction	Vol (ml)	Amount protein (mg)	Activity (pmol/min)	Specific (pmol/min/ mg)	Recovery (%)	Enrich-ment (fold)
Detergent Extract	25	40	1,375	34.4	100	1.0
DE-52	22	7.5	1,632	218	119	6.3
Mono-P	2.6	0.065	156	2,400	11	70

Table 3-1. Partial Purification of the *Sulfolobus solfataricus* Protein Kinase.

Shown is a summary of a typical preparation of the source of the protein kinase activity from a detergent extract (Triton X-100) of the membrane fraction from 20g (wet weight) of *S. solfataricus*. All steps were performed as described in the text. Protein kinase activity was measured using casein as the phosphoacceptor during a solution assay conducted at 25^{o}C. Protein concentrations were determined using a Bradford assay as described in Methods. This table appears in Lower et. al., 2000.

cellulose, which had been equilibrated with 20 mM MES pH 6.5, containing 25 mM NaCl, and either 0.1% Triton X-100 or 12.5 mM octyl glucoside. The column was washed with three column volumes of the same equilibration buffer, and then developed with a linear gradient of 25-500 mM NaCl in 150 ml (total volume) of equilibration buffer. Fractions, approximately 3 ml, were collected and assayed for protein kinase activity using a solution assay as described in Methods. The fractions containing protein kinase activity, which eluted at approximately 200 mM NaCl, were pooled and retained as the DE-52 fraction (Figure 3-2). Three peaks of protein kinase activity were observed eluting from the DE-52 column when Triton X-100 was used as the solubilizing detergent. When an "in gel" protein kinase assay was preformed on these three fractions, one band of protein kinase activity at ≈ 67 kDa was observed in all three, indicating that each contained the same active polypeptide species. When octyl glucoside was used as the solubilizing detergent instead of Triton X-100, only one peak of activity having a molecular mass of ≈ 67 kDa was observed eluting from the DE-52 column. Since Triton X-100 (CMC ≈ 0.3 mM) has a lower CMC than does octyl glucoside (CMC ≈ 25 mM), we hypothesized that the three fractions of activity observed when Triton X-100 were artifacts arising from the presence of detergent micelles.

The DE-52 fraction (approximately 10 ml) was dialyzed at 4^{o}C against 1L of 20 mM piperazine, pH 5.5, containing 0.05% (v/v) Triton X-100 overnight, and then again against fresh buffer for 2 - 4 hours. This solution was applied to a Mono-P HR5/20 chromatofocusing column that had been equilibrated with dialysis buffer using a Pharmacia FPLC system at a flow rate of 0.5 ml/min. The column was washed with 10 ml of 20 mM piperazine, pH 5.5, containing 0.05% (v/v) Triton X-100 at a flow rate of 0.5 ml/min. Finally, the column was developed isocratically (pH 7.0 – 4.0) at a flow rate of 0.5 ml/min with a solution consisting of a 1:8 dilution of polybuffer 74, pH 4.0, containing 0.05% (v/v) Triton X-100. Fractions, approximately 1 ml, were collected and assayed for protein kinase activity using a solution assay as described in Methods (Figure 3-3). The fractions containing protein kinase activity, which eluted at a pH of ≈ 4.75, were pooled and retained as the Mono-P fraction. As was the case with ion exchange chromatography, multiple peaks of protein kinase activity were observed eluting from the Mono-P column when Triton X-100 was used as the solubilizing

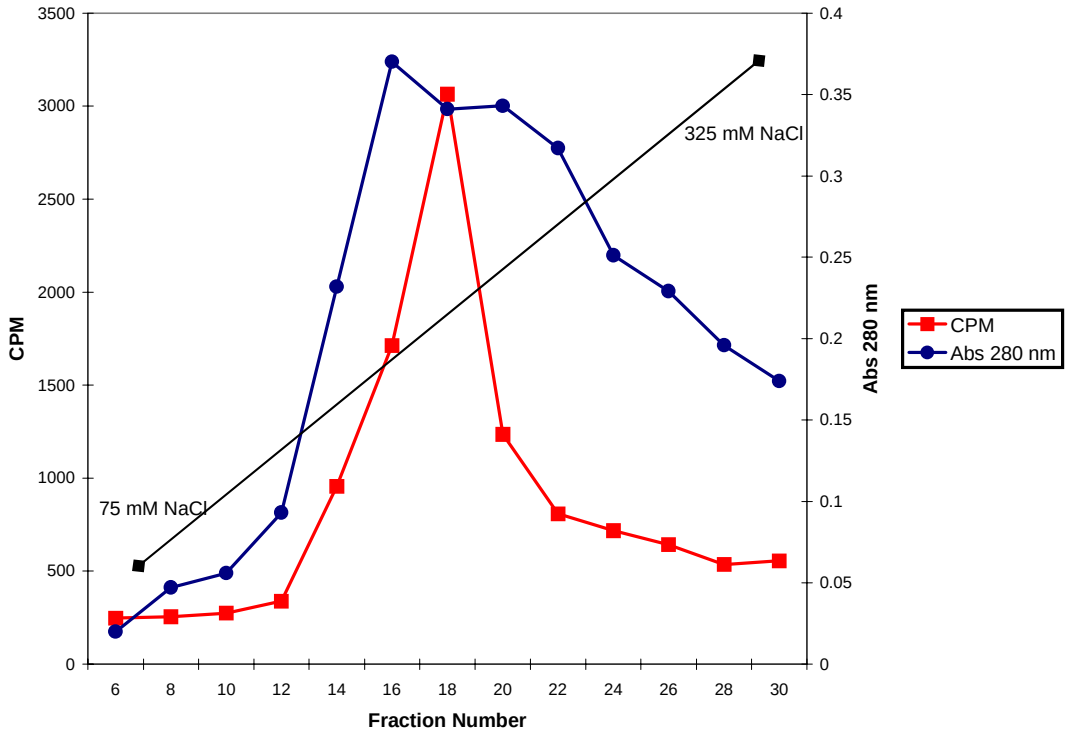


Figure 3-2. Anion Exchange Chromatography of the *Sulfolobus solfataricus* Protein Kinase on DE-52 Cellulose.

Proteins from the detergent soluble membrane extract were separated on a DE-52 column as described in the text. Shown here is a typical preparation of a detergent extract prepared using octyl glucoside of the protein kinase using a DE-52 column. Fractions, 3 ml, were collected and analyzed for protein kinase activity using a solution assay as described in Methods. Briefly, 40 µl aliquots of even numbered elution fractions were incubated with casein and [γ-^{32}P] ATP for 1 hour at 25^{o}C, precipitated onto filter paper, washed extensively with TCA/NaPP$_i$, dried, and radioactivity incorporated into the casein determined by liquid scintillation counting. Shown on the first y-axis is the [^{32}P] incorporated into casein during the assay, and on the second y-axis is the Abs 280 nm as a function of the elution fraction. Also, note that only one peak of protein kinase activity is observed eluting from the column at ≈ 200 mM NaCl.

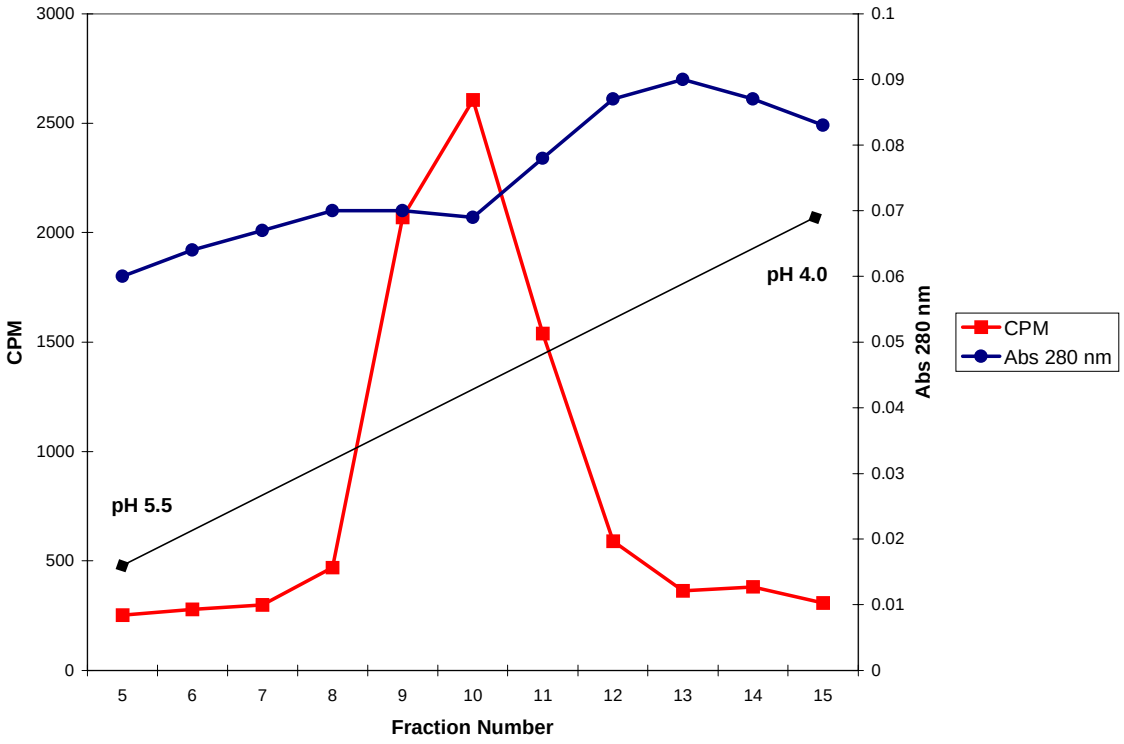


Figure 3-3. Chromatofocusing Chromatography of the *Sulfolobus solfataricus* Protein Kinase on Mono-P.

Proteins from the DE-52 (octyl glucoside) fraction were separated on a Mono-P column as described in the text. Shown here is a typical preparation of the protein kinase using a Mono-P column. Fractions, 1 ml, were collected and analyzed for protein kinase activity using a protein kinase solution assay as described in Methods. Briefly, 20 µl aliquots of the elution fractions were incubated with casein and [γ-^{32}P] ATP for 1 hour at 25^{o}C, precipitated onto filter paper, washed extensively with TCA/NaPP$_i$, dried, and radioactivity incorporated into the casein determined by liquid scintillation counting. Shown on the first y-axis is the of [^{32}P] incorporated into casein during the assay, and on the second y-axis is the relative protein concentration determined by Abs 280 nm as a function of the elution fraction. Also, note that only one peak of protein kinase activity is observed eluting from the column at a pH of $\approx$ 4.75.

detergent. When an "in gel" protein kinase assay was preformed on these two fractions, one band of protein kinase activity at ≈ 67 kDa was observed for both fractions indicating the fractions contained the same active polypeptide species. When octyl glucoside was used as the solubilizing detergent instead of Triton X-100, only one peak of activity having a molecular mass of ≈ 67 kDa was observed eluting from the Mono-P column. Again, we reasoned that because Triton X-100 has a lower CMC than octyl glucoside, the two fractions of activity observed eluting off the column when Triton X-100 was employed represented the same active polypeptide species.

As is noted in Table 3-1, while significant purification was achieved using chromatofocusing, yields were poor. Thus, many of the experiments described in this chapter were performed with the DE-52 fraction. It should be noted that, where reported in the figures and tables, the performance of the DE-52 fraction paralleled that observed with the Mono-P fraction. Also, the degree of purification of the DE-52 and Mono-P fractions over cell lysate could not be determined as we were not able to perform quantitative solution assays until the protein kinase had been solubilized with detergent. Thus, the significant enrichment that undoubtedly resulted from isolation and purification of the protein kinase remains unaccounted for.

Gel Filtration Chromatography was performed on a portion of the DE-52 fraction in order to determine the apparent molecular mass of the *S. solfataricus* protein kinase. The DE-52 fraction, 3 ml, was concentrated to a volume of ≈ 300 µl using a Centricon YM-10 Centrifugal Filter Device (Millipore) centrifuged at 5,000 x g, 4^{o}C, for 60 minutes. The concentrated DE-52 fraction was loaded onto a 2 x 90 cm Sephacryl S-200 column equilibrated in 20 mM MES, pH 6.5, containing 200 mM NaCl and 15 mM octyl glucoside. Protein was eluted with the equilibration buffer and 1 ml fractions were collected and analyzed for protein kinase activity using a solution assay as described in Methods. The fractions containing protein kinase activity eluted from the column with an apparent molecular mass of ≈ 125 kDa, as compared to molecular weight standards (Figure 3-4). This suggests that the protein kinase exists as a heteroligomer, perhaps, a homodimer. The physiological significance of the ≈ 125 kDa oligomer must be taken with caution, as the protein kinase activity present in the crude detergent soluble (octyl glucoside) membrane extract eluted in the void volume (> 250 kDa) of the S-200 column.

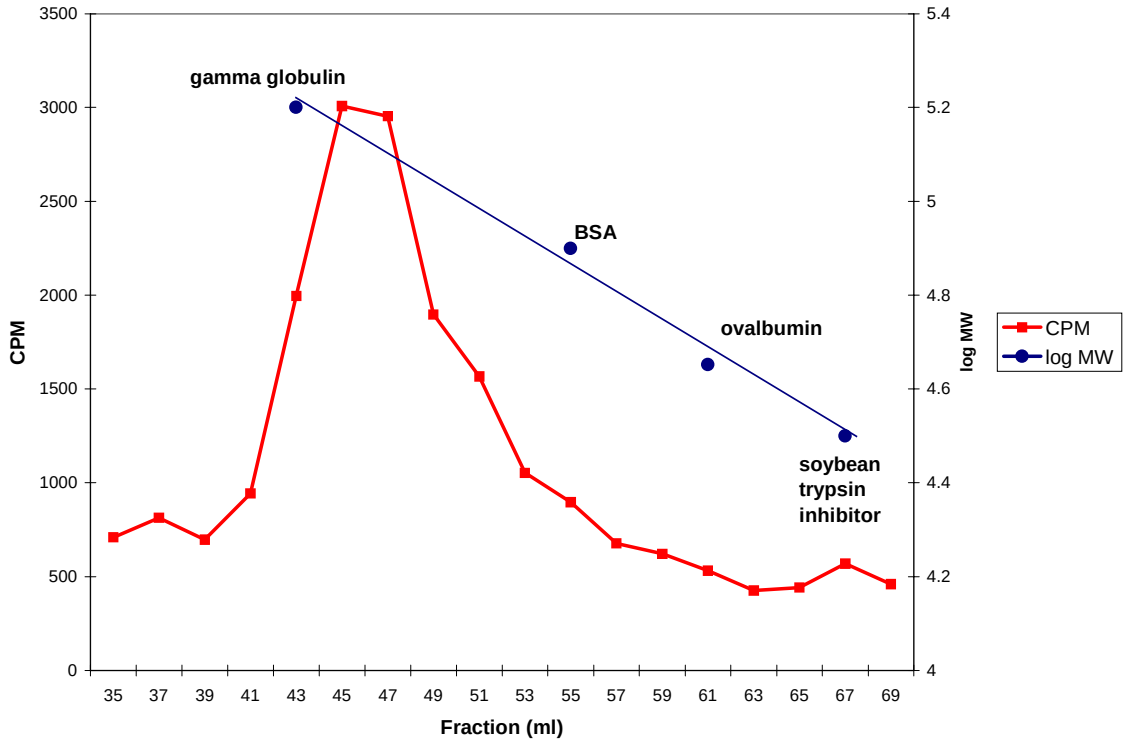


Figure 3-4. Gel Filtration Chromatography of the *Sulfolobus solfataricus* Protein Kinase on Sephacryl S-200.

Proteins from the concentrated DE-52 fraction containing the *S. solfataricus* protein kinase activity were separated on a Sephacryl S-200 column as described in the text. Fractions, 1 ml, were collected and analyzed for protein kinase activity using a protein kinase solution assay as described in Methods. Briefly, 30 µl aliquots of the elution fractions were incubated with casein and [γ-^{32}P] ATP for 1 hour at 25^{o}C, precipitated onto filter paper, washed extensively with TCA/NaPPi, dried, and radioactivity incorporated into the casein determined by liquid scintillation counting. Shown on the first y-axis is the relative protein kinase activity in CPMs of ^{32}P-incorporated into casein during the assay as a function of the elution volume. Shown on the second y-axis is the log of the molecular weight of the protein standards (gamma globulin, 160 kDa; BSA, 68 kDa; ovalbumin, 45 kDa; soybean trypsin inhibitor, 30 kDa) used to calibrate the column as a function of elution volume. Note that only one peak of protein kinase activity is observed eluting from the column with an apparent molecular mass of $\approx$ 125 kDa.

None the less, the ability to recover protein kinase activity following SDS-PAGE (seen in Figure 3-1) indicates that if the enzyme does exist as a oligomer, the essential determinants for its catalytic function reside within a single subunit of ≈ 67 kDa.

Phosphorylation of Exogenous Protein and Peptide Substrates by the Sulfolobus solfataricus Protein Kinase.

In order to describe the *S. solfataricus* protein kinase in more detail, the enzyme was challenged with a variety of protein, polymer, and peptide substrates (Table 3-2). Protein kinase activity was measured *in* vitro using a solution assay as described in Methods, using several eukaryotic phosphoacceptor proteins as possible substrates. Of these, casein proved to be the most effective exogenous substrate, while RCM lysozyme, mixed histones, histone H4, and MLC peptide (KKRAARATSNVFA), a short peptide modeled after the phosphorylation site sequence of myosin light chains (Kennelly et. al., 1987), also were phosphorylated by the protein kinase. No phosphorylation could be detected with poly(Glu_4:Tyr), poly(Glu:Tyr), BSA, or myelin basic protein

Phosphoamino acid analysis was performed on casein, RCM-lysozyme, histone H4, and MLC peptide to determine the type of amino acid(s) that were phosphorylated by the protein kinase (Figure 3-5). The results indicate that the protein kinase has a preference for phosphorylating threonine residues *in vitro* as casein, RCM-lysozyme, and MLC peptide were phosphorylated exclusively on threonine residues. Histone H4 was phosphorylated on serine in addition to threonine.

The MLC peptide allowed us to further explore the observed tendency of the protein kinase to phosphorylate threonine residues *in vitro*. A variant of the MLC peptide, named T8S peptide (KKRAARASSNVFA), in which the threonine at position 8 was changed to a serine, was synthesized and utilized as a potential phosphoacceptor substrate. No phosphorylation of the T8S peptide could be detected when the solution assay for protein kinase activity was conducted at 25^{o}C, suggesting the protein kinase had a preference for phosphorylating threonine residues *in vitro*. However, when the temperature was raised to 65^{o}C or polyanionic species such as heparin or poly(Glu_4:Tyr) added to the reaction mixture, detectable levels of phosphorylation of T8S peptide occurred (Table 3-3). Poly(Glu_4:Tyr) was observed to stimulate the protein kinase

Protein or Peptide	Concentration (mg/ml)	Activity (pmol ^{32}P/min/mg)		Relative Activity (% casein)	
		DE-52	Mono-P	DE-52	Mono-P
Casein	0.60	44.0 ± 6.5	78.8 ± 3.9	100	100
RCM lysozyme	0.80	7.5 ± 1.2	13.9 ± 0.4	17	18
Mixed Histones	0.20	1.8 ± 0.5	17.5 ± 0.5	4	22
Histone H4	0.40	6.6 ± 1.5	20.4 ± 1.3	15	26
MLC Peptide	0.25 mM	4.2 ± 0.4	1.9 ± 0.3	10	2
Myelin Basic Protein	1.0	n.d. *	-	n.d.	-
Poly (glu4:tyr)	1.0	n.d. *	-	n.d.	-
Poly (glu: tyr)	1.0	n.d. *	-	n.d.	-
BSA	1.0	n.d. *	-	n.d.	-

*n.d. = not detectable, < 1.0 pmol ^{32}P / min / mg

Table 3-2. Activity of *Sulfolobus solfataricus* Protein Kinase Toward Exogenous Proteins, Amino Acid Copolymers, and Peptides.

The activity of the *S. solfataricus* protein kinase, 3.2 - 5.1 μg of DE-52 fraction or 0.14 μg of Mono-P fraction, toward a range of exogenous proteins, amino acid copolymers, and peptides was determined using the solution assay technique describes in Methods. Substrate concentrations were selected as those that yielded the highest activity in preliminary experiments. In those instances where no activity was detected in preliminary experiments, a concentration of 1.0 mg/ml was selected. Shown are the averages of triplicate determinations plus or minus standard error (Lower et. al., 2000).

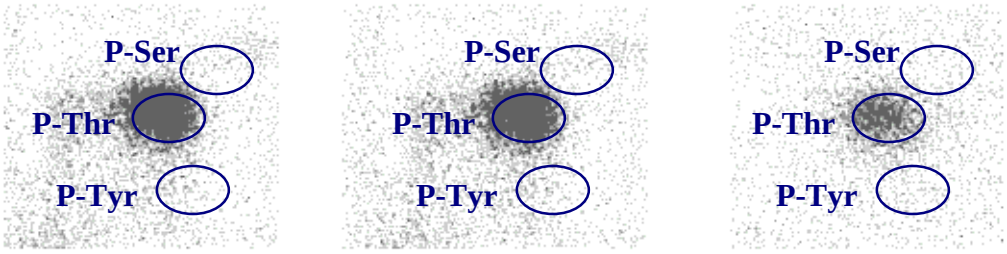


Figure 3-5. The *Sulfolobus solfataricus* Protein Kinase Exhibits a Tendency to Phosphorylate Threonine Residues *in vitro* Using ATP as the Phosphoryl Donor.

The indicated proteins or peptides were incubated with the *Sulfolobus solfataricus* protein kinase during a solution protein kinase assay as described in Methods. Following the assay the radiolabeled proteins were isolated by SDS-PAGE while the peptide was separated by passage through a G-25 column. The phosphoproteins were transferred to PVDF membranes and the portion of the membrane containing labeled, or aliquot of the gel-filtered peptide, was incubated in 6N HCl for 1 hour at 95^{o}C. The supernatant liquid was then applied to a 20 x 20 cm silica gel TLC plate, along with standards of P-Ser, P-Thr and P-Tyr, and then subjected to 2-dimensional electrophoresis. The first dimension (left to right) was performed at pH 1.9 and the second dimension (bottom to top) at pH 3.5. Standards were visualized by ninhydrin staining, while phosphorylated amino acids were visualized by electronic autoradiography. Shown are the electronic autoradiograms of the TLC plates from the analysis of casein (left), RCM lysozyme (center), MLC peptide (right). Histone H4 (not shown) contained P-Ser and P-Thr. Circles indicate the positions of P-Ser, P-Thr and P-Tyr (Lower et. al., 2000).

towards MLC peptide and T8S peptide, but not towards macromolecular substrates such as casein and myelin basic protein.

A determination of the kinetic parameters of the protein kinase towards the phosphorylation of MLC and T8S peptides was done using a solution assay performed at either 25°C or 65°C (Table 3-4), the latter being the highest temperature at which enzyme stability could be maintained for significant periods of time. This temperature also falls within the optimum temperature of growth for *S. solfataricus* (60° – 90°C). The results showed that while the K_m for the two peptides were similar with that of T8S peptide in fact being lower, the V_{max} for phosphorylation of MLC peptide was over 20-fold higher (Figure 3-6). Similarly, the effect of poly(Glu_4:Tyr) on the phosphorylation of MLC peptide was solely characterized by an increase in V_{max} (25-fold), while K_m remained relatively constant. As expected, calculations of substrate specificity values (V_{max} / K_m) for MLC peptide, T8S peptide, and ATP, suggest that the protein kinase has a preference for ATP (especially at 65°C) and for MLC peptide > T8S peptide.

Metal Ion Preference of the Sulfolobus solfataricus Protein Kinase.

Protein kinases utilize divalent metal ion cofactors as Lewis Acids to aid in catalysis. In most kinases, the divalent metal forms a salt bridge with the oxyanions of the beta and gamma phosphates of ATP. This facilitates the nucleophilic attack on the gamma phosphate of ATP by the protein substrate. Solution assays for protein kinase activity were performed as described in Methods with the exception that, in place of the standard divalent metal ion cofactor, Mn^{2+}, divalent metals Ca^{2+}, Mg^{2+}, Mn^{2+}, Ni^{2+}, and Zn^{2+} were added to a final concentration of 5 mM. Control assays, in which EDTA was added to the reaction mixture to a final concentration of 5 mM in place of a divalent metal, were performed in parallel. Analysis of the cofactor requirements at 25°C revealed a strong preference for Mn^{2+} over other divalent metal ions. However, at temperatures within the physiological range of *S. solfataricus*, 65°C, Mg^{2+} proved equally effective as a cofactor (Figure 3-7).

Peptide Substrate	Polyanion (1 mg/ml)	Specific Activity (pmol ^{32}P/min/mg)	Relative Activity	Relative Activity
MLC	none	4.8	1.0	
MLC	Glu:Tyr	6.4	1.3	
MLC	Glu_4:Tyr	58.0	12.0	
T8S	none	n.d.*		—
T8S	Glu:Tyr	n.d.*		—
T8S	Glu_4:Tyr	1.3		—

*n.d. = not detectable, < 1.0 pmol ^{32}P / min / mg

Table 3-3. Stimulation of the *Sulfolobus solfataricus* Protein Kinase Towards Peptide Substrates by the Polyanions Poly (Glu_4:Tyr) and Poly (Glu:Tyr).

The activity of the *S.* solfataricus protein kinase, 3.2 - 5.1 μg of DE-52 fraction, towards MLC peptide (KKRAARATSNVFA) or T8S peptide (KKRAARASSNVFA)was assayed in the presence of the indicated polyanion at 25°C. Conditions for the solution assays were as described in Methods with the exception that either poly(Glu:Tyr) or poly(Glu_4:Tyr) was added to a final concentration of 1.0 mg/ml. MLC peptide or T8S peptide was added to a final concentration of 0.25 mM.

Substrate (polyanion)	Temp. (^{O}C)	K_m (μM)	V_{max} (pmol^{32}P/min/mg)	V_{max} / K_m (pmol^{32}P/ μM min/mg)
MLC peptide	25^{o}	30 ± 1	3.88 ± 0.09	2.3
MLC peptide	65^{o}	45 ± 5	333 ± 11	7.4
MLC peptide (Glu_4:Tyr)	25^{o}	13 ± 2	102 ± 13	7.8
T8S peptide*	65^{o}	20 ± 2	13.9 ± 1.4	1.4
ATP	25^{o}	0.74 ± 0.19	5.43 ± 0.94	7.3
ATP	65^{o}	15.7 ± 1.8	345 ± 23	22
GTP	65^{o}	63.6 ± 10.6	140 ± 11	2.2

* Phosphorylation of T8S peptide was undetectable at 25^{o} C.

Table 3-4. Kinetic constants of the *Sulfolobus solfataricus* Protein Kinase Toward MLC Peptide, T8S Peptide, and ATP Under Various Conditions.

The kinetic constants for ATP, MLC peptide, and T8S peptide were determined under a variety of conditions by the Lineweaver-Burk method. Duplicate rate measurements at various substrate concentrations were determined using the procedure described in Methods the exceptions listed in Table IV. All assays were performed using 3.2 μg DE-52 fraction and substrate concentration ranges of 0 – 200 μM ATP, 0 – 0.7 mM MLC peptide and 0 – 0.6 mM T8S peptide. Shown are the average kinetic constants plus or minus standard error (Lower et. al., 2000).

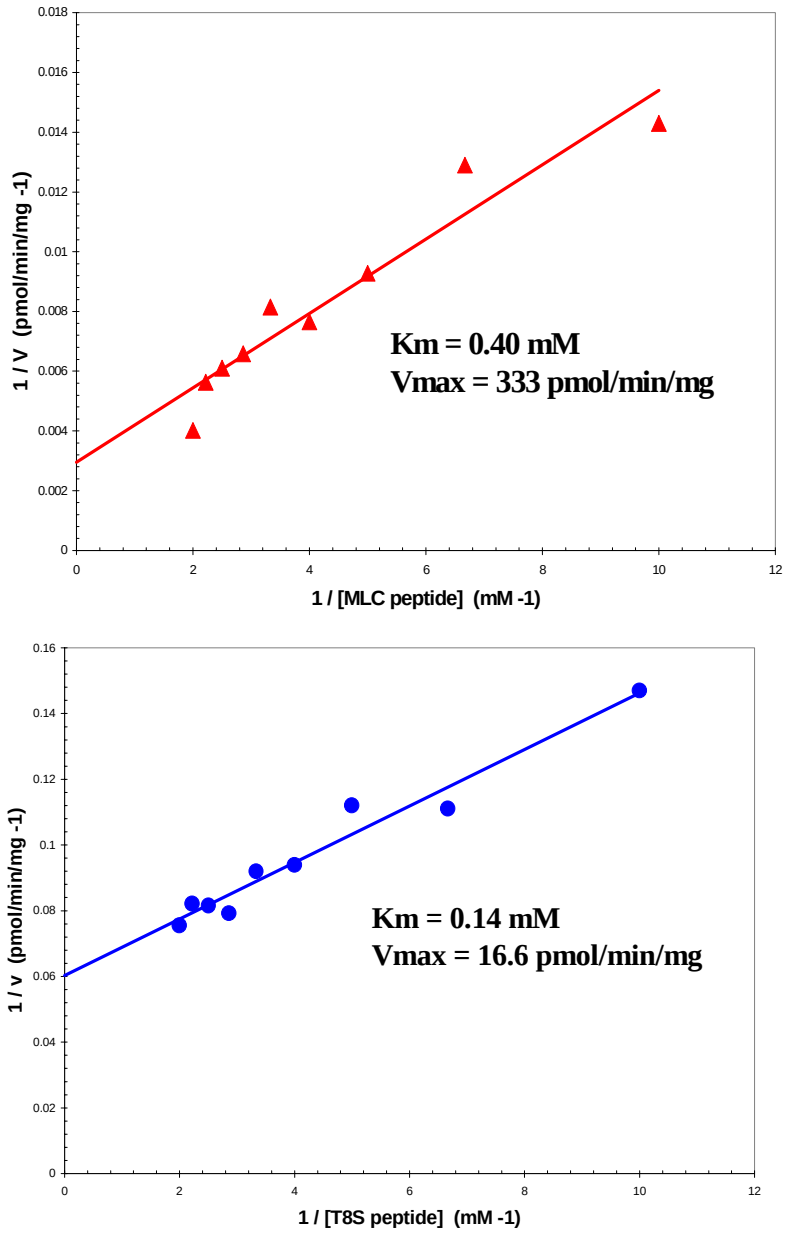


Figure 3-6. Comparison of the Rates of Phosphorylation of MLC Peptide and T8S Peptide by the *Sulfolobus solfataricus* Protein Kinase.

The kinetic constants for **MLC peptide** (▲), and **T8S peptide** (●) were determined at 65°C. Duplicate rate measurements at the indicated substrate concentrations were determined using a solution protein kinase assay as described in Methods. The assays were performed using 3.2 µg DE-52 fraction and substrate concentration ranges from 0 – 0.7 mM MLC peptide and 0 – 0.6 mM T8S peptide. Shown, are two Lineweaver-Burk reciprocal plots of protein kinase activity as a function of peptide concentration.

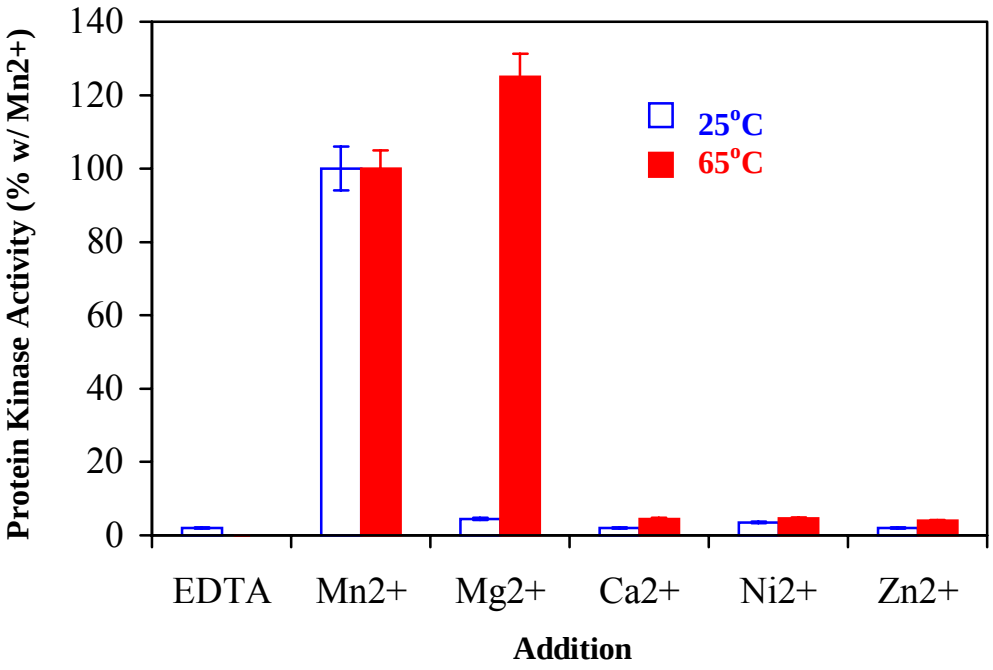


Figure 3-7. Survey of Potential Metal Ion Cofactors for the *Sulfolobus solfataricus* Protein Kinase.

The activity of the *S. solfataricus* protein kinase was assayed in the presence of the indicated divalent metal ions at both 25°C and 65°C. Conditions were as described in Methods with the exception that, in the place of the standard divalent metal, Mn^{2+}, the compounds listed were added at a final concentration of 5 mM. For the assays conducted at 25°C, Mono-P fraction, 0.25 –0.50 mg, was used as the source of the protein kinase activity and casein was used as the phosphoacceptor substrate at a final concentration of 0.8 mg /ml. For the assays conducted at 65°C, DE-52 fraction, 3.0 µg, was used as the source of protein kinase activity and MLC peptide was used as the phosphoacceptor substrate at a final concentration of 0.5 mM. Shown is the average of duplicate determinations ± standard error (Lower et. al., 2000).

Nucleotide Preference of the Sulfolobus solfataricus Protein Kinase

A variation of the "in gel" assay technique was used to examine the nucleotide specificity of the *S. solfataricus* protein kinase as described in Methods. Briefly, samples of the protein kinase (DE-52 fraction) were subjected to SDS-PAGE and renatured by standard procedures as described in Methods. Next, the gel was divided into individual lanes and each section was incubated overnight in assay buffer containing 50 μM of one potential nucleotide donor ATP, ADP, AMP, GTP, GDP, GMP, CTP, UTP or $NaPP_i$. These compounds were not radiolabeled and hence any phosphate transferred to the casein during this period would be undetectable by autoradiography. The purity of all nucleotide solutions was verified by thin-layer chromatography on polyethyleneimine cellulose to determine that the nucleotide diphosphates were free of contamination by nucleotide triphosphates. After overnight preincubation, the gels were washed extensively to remove unlabeled nucleotide donor and a standard "in gel" assay was performed using [γ-^{32}P] ATP. If substantial phosphotransfer occurred during the overnight preincubation then transfer of [^{32}P] phosphate from [γ-^{32}P] ATP during the subsequent "in gel" assay should be attenuated for lack of available phosphorylation sites on the immobilized casein. As is observed in Table 3-5, purine nucleotide di- and triphosphates served as phosphodonor substrates *in vitro*, with an apparent order of preference for ATP > GTP > ADP > GDP. Neither of the pyrimidine nucleotide triphosphates tested, CTP and UTP, nor the other potential phosphate donors, AMP, GMP, and PP_i served as substrates for the protein kinase.

To establish the validity of this method for evaluating nucleotide specificity, the ability of the protein kinase (contained within the DE-52 fraction or the Mono-P fraction) to catalyze the transfer of [^{32}P] phosphate from [γ-^{32}P] GTP or [β-^{32}P] GDP was assayed in solution. Conditions for the assay were as described in Methods with the exception that either [γ-^{32}P] GTP or [β-^{32}P] GDP was substituted for [γ-^{32}P] ATP. As shown in Table 3-6, both sources of *S. solfataricus* protein kinase catalyzed the phosphorylation of casein or MLC peptide using either guanine nucleotide donor, as predicted. Subsequent phosphoamino acid analysis revealed that both casein and MLC peptide were

Nucleotide	Phosphate incorporation into casein by the ≈ 67 kDa protein kinase (% control)
ATP	32 ± 5 (3)
ADP	70 ± 17 (3)
AMP	96 ± 4 (3)
GTP	50 ± 12 (3)
GDP	85 ± 14 (2)
PPi	96 ± 4 (2)

Table 3-5. Nucleotide Specificity of the *Sulfolobus solfataricus* Protein Kinase.

Samples of DE-52 fraction, 12.5 µg each, were subjected to SDS-PAGE in gels copolymerized with casein. The protein kinase was renatured within the gel, the gels sectioned into individual lanes, and the lanes preincubated overnight with the indicated nucleotides at a concentration of 50 mM, Afterwards the gels were washed to remove nonlabeled nucleotides, and an "in gel" assay, as described in Methods, was performed using [γ-^{32}P] ATP. The gels were then washed free of [γ-^{32}P] ATP and the incorporation of [^{32}P] was determined by electronic autoradiography. Phosphorylation of casein by unlabeled phosphate during the preincubation should be manifested as an attenuation of [^{32}P] during subsequent incubation with [γ-^{32}P] ATP. Shown are the levels of [^{32}P] radioactivity incorporated into casein by the ≈ 67 kDa protein kinase relative to those for controls in which no nucleotide was present during preincubation. All values are averages of multiple determinations shown in parentheses ± standard errors (Lower et. al., 2000).

Protein Kinase Fraction	Phospho-acceptor	Activity (pmol^{32}P/min/mg) (% that with ATP) with nucleotide donor		
		ATP	GTP	GDP
DE-52	Casein	36.9 ± 1.8 (100)	24.1 ± 0.1 (65)	3.0 ± 0.1 (8)
DE-52	MLC peptide	45.9 ± 0.1 (100)	20.0 ± 2.8 (44)	1.4 ± 0.3 (3)
Mono-P	Casein	57.0 ± 5.1 (100)	40.2 ± 2.1 (70)	0.6 ± 0.5 (1)
Mono-P	MLC peptide	55.4 ± 2.2 (100)	89.2 ± 8.0 (161)	23.0 ± 4.3 (26)

Table 3-6. Phosphoryl Transfer by the *Sulfolobus solfataricus* Protein Kinase from GTP or GDP *in vitro*.

The activity of the *S. solfataricus* protein kinase, 4.0 μg of DE-52 fraction or 0.28 μg of Mono-P fraction, toward casein and MLC peptide was determined using the solution assay technique described in Methods. Where indicated [γ-^{32}P] GTP or [β-^{32}P] GDP was substituted for the normal phosphoryl donor, [γ-^{32}P] ATP. Phosphorylation of casein was measured at 25^{o}C and MLC peptide at 65^{o}C. Shown are the averages, ± standard error, of duplicate (DE-52 fraction) or triplicate (Mono-P fraction) analyses (Lower et., al., 2000).

phosphorylated exclusively on threonine residues, establishing that a covalent protein-phosphate bond had been formed (Figure 3-8).

Additionally, kinetic parameters for the protein kinase towards GTP were determined at 65°C using MLC peptide as the exogenous phosphoacceptor substrate (Table 3-4). The analysis reveals that the K_m for ATP is approximately 5-fold lower and the V_{max} for ATP was about 2-fold higher. Furthermore, the specificity constant (V_{max} / K_m) for ATP >> GTP assayed at 25°C, indicating that while the protein kinase can utilize GTP as a phosphoryl donor *in vitro*, it prefers to use ATP.

In addition, transfer of [^{32}P] phosphate to casein by the *S. solfataricus* protein kinase could be detected using an "in gel" assay in which [γ-^{32}P] ATP, [γ-^{32}P] GTP, or [β-^{32}P] GDP was employed as the phosphodonor substrate. Conditions for the "in gel" assay were as described in Methods with the exception that [γ-^{32}P] ATP, [γ-^{32}P] GTP, or [β-^{32}P] GDP was added to the reaction mixture at a final concentration of 50 μM. As seen in Figure 3-9, each nucleotide donor displayed a single catalytically active polypeptide species of ≈ 67 kDa, with the degree of phosphoryl transfer exhibiting an order of preference of ATP > GTP >> GDP.

Sensitivity of the Sulfolobus solfataricus Protein Kinase to Inhibitors of Eukaryotic Protein Kinases

Because the *S. solfataricus* protein kinase displayed an ability to phosphorylate serine and threonine residues *in vitro*, a characteristic of the so-called "eukaryotic" protein kinases, we challenged the enzyme with several known eukaryotic protein kinase inhibitors. These included PKI peptide, a highly specific inhibitor of the cAMP-dependent protein kinase; genistein, an inhibitor of protein-tyrosine kinases; tamoxifen, an inhibitor of protein kinase C; H7, a specific inhibitor of cAMP-dependent protein kinases; H89, a selective inhibitor of protein kinase A; ML-9, a selective inhibitor of MLC kinase; and staurosporine, a broad range protein kinase inhibitor (reviewed in Hemmings, 1997). A solution assay for determining protein kinase activity was

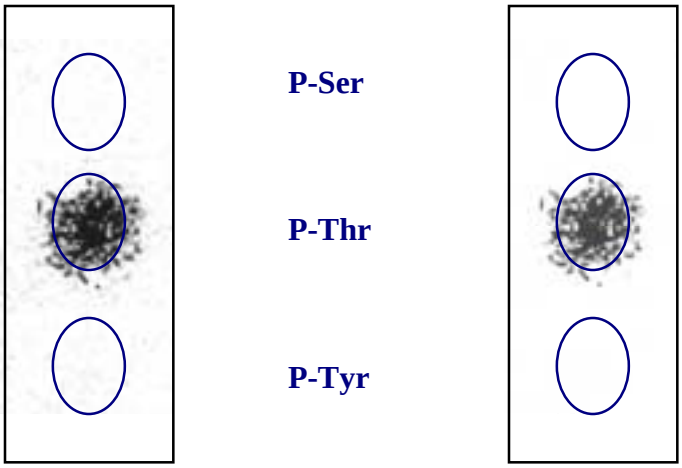


Figure 3-8. The *Sulfolobus solfataricus* Protein Kinase Exhibits a Tendency to Phosphorylate Threonine Residues *in vitro* Using GTP as the Phosphoryl Donor.

Casein (1 mg/ml) or MLC peptide (1.0 mM) was incubated with the *Sulfolobus solfataricus* protein kinase (3.0 μg DE-52 fraction) during a solution assay. Conditions were as described in text with the exception that the assay using MLC peptide was performed at 65°C, while the assay using casein was performed at 25°C. Afterwards, radiolabeled casein was isolated on a 10% (w/v) SDS-polyacrylamide gel using a SDS-PAGE, while radiolabeled MLC peptide was isolated on a 15% (w/v) SDS-polyacrylamide gel using SDS-PAGE. The phosphoproteins were transferred to PVDF membranes and the portion of the membrane containing labeled substrate was incubated in 6N HCl for 1 hour at 95°C. The supernatant liquid was then applied to a 20 x 20 cm silica gel TLC plate, along with standards of P-Ser, P-Thr and P-Tyr, and then subjected to electrophoresis (bottom to top) at pH 3.5 as described in Methods. Standards were visualized by ninhydrin staining, while phosphorylated amino acids were visualized by electronic autoradiography. Shown are the electronic autoradiograms of the TLC plates from the analysis of casein (left) and MLC peptide (right). Circles indicate the positions of P-Ser, P-Thr and P-Tyr.

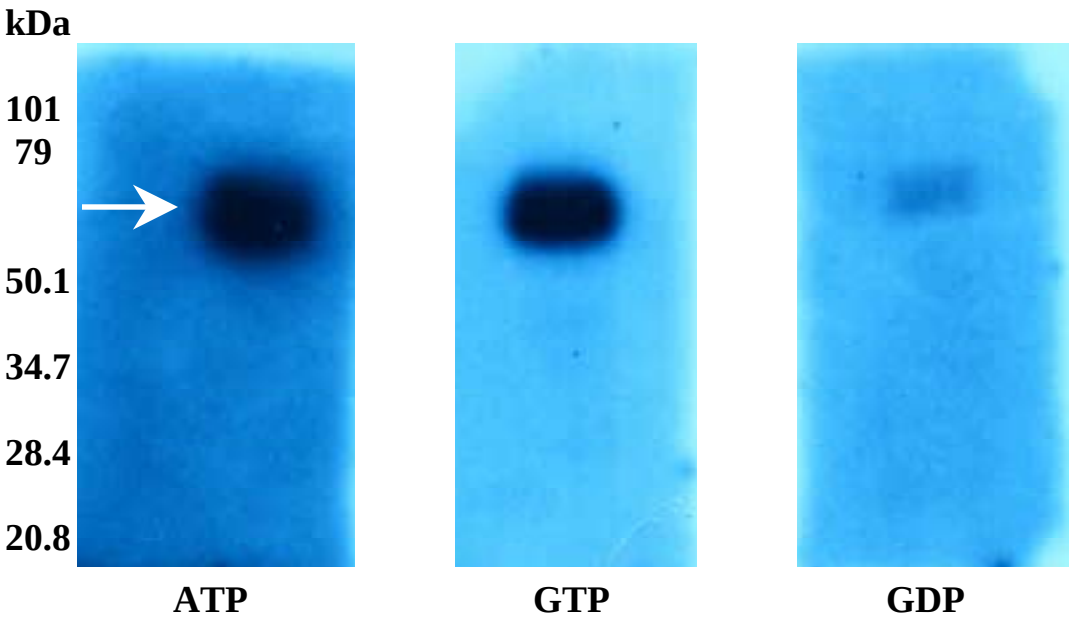


Figure 3-9. The *Sulfolobus solfataricus* Protein Kinase Exhibits an Ability to Utilize ATP, GTP, or GDP as the Phosphoryl Donor During an "in gel" Assay.

Shown are the results of an "in gel" assay of *S. solfataricus* protein kinase activity following SDS-PAGE. Portions 20 µg, of DE-52 fraction were electrophoresed into a 10% (w/v) SDS-polyacrylamide gel in which casein had been copolymerized into the gel matrix at a concentration of 1 mg/ml. Following electrophoresis, the gels were washed free of SDS and the proteins within renatured using guanidine hydrochloride. The gels were each then washed extensively in a renaturation solution consisting of 20 mM MES, pH 6.5, containing 0.1% (v/v) Triton X-100, 5 mM $MnCl_2$, 5 mM $MgCl_2$, and 1 mM DTT. Following, the each gel was incubated for 60 minutes at 25°C in renaturation solution containing either 50 µM [γ-^{32}P] ATP (specific activity 30 Ci/mole), 50 µM [γ-^{32}P] GTP, or 50 µM [β-^{32}P] GDP. The gels were washed free of excess nucleotides and regions of each gel in which [^{32}P] phosphate was incorporated into protein was visualized by autoradiography using X-ray film. Shown are the autoradiograms of each gel with molecular weight standards indicated at left and an arrow ($\rightarrow$) indicating the predicted position of the $\approx$ 67 kDa protein kinase. Subsequent phosphoamino acid analysis of casein labeled with [γ-^{32}P] ATP or [γ-^{32}P] GTP indicated that phosphorylation took place on threonine.

performed at 25°C using MLC peptide as the phosphoacceptor substrate as described in Methods, with the exception that protein kinase inhibitors were also added to the reaction mixture. All the compounds were tested in millimolar concentrations, well above the concentrations for which they are known to inhibit their eukaryotic target enzymes.

Only three of the compounds tested, H7, Genistein, and ML-9, inhibited the *S. solfataricus* protein kinase in a concentration-dependent manner (Figure 3-10). Of these three inhibitors, ML-9 was clearly the most effective at inhibiting protein kinase activity. Even so, the apparent 50% inhibitory concentration (IC_{50}) exhibited by ML-9 was between 0.5 – 1.0 mM, a range well above the concentrations known to inhibit eukaryotic protein kinases such as MLC kinase (K_i = 3.8 μM), protein kinase C (K_i = 54 μM), and cAMP-dependent protein kinase (K_i = 32 μM) (Hidaka and Kobayashi, 1992).

Autophosphorylation of the Sulfolobus solfataricus Protein Kinase

When the catalytic efficiency of the *S. solfataricus* protein kinase towards peptide substrates was examined at elevated temperatures (65°C), it was observed that the quantity of radioactivity incorporated into controls lacking any exogenous phosphoacceptor substrate were much higher than expected. This behavior indicated that the protein kinase was phosphorylating some other component within the DE-52 fraction, or that the protein kinase was undergoing autophosphorylation at the elevated temperature. To examine this further an "in gel" assay was performed on the DE-52 fraction as described in Methods, with the exception that no exogenous substrate was copolymerized into the gel and the assay was conducted at 65°C instead of 25°C. As a control, a second gel into which casein had been copolymerized was ran in parallel. As expected, one band of [^{32}P] incorporation at ≈ 67 kDa was observed in the gel containing copolymerized casein as the exogenous phosphoacceptor. Similarly, one band of [^{32}P] incorporation at ≈ 67 kDa was observed in the gel in which no exogenous substrate had been copolymerized. While these results suggested an autophosphorylation event, the possibility remained that the protein kinase had phosphorylated a polypeptide from *S. solfataricus* of similar M_r that was also present in the DE-52 fraction and accessible to it during the "in gel" assay. To disqualify this notion, a subsequent "in gel" assay was

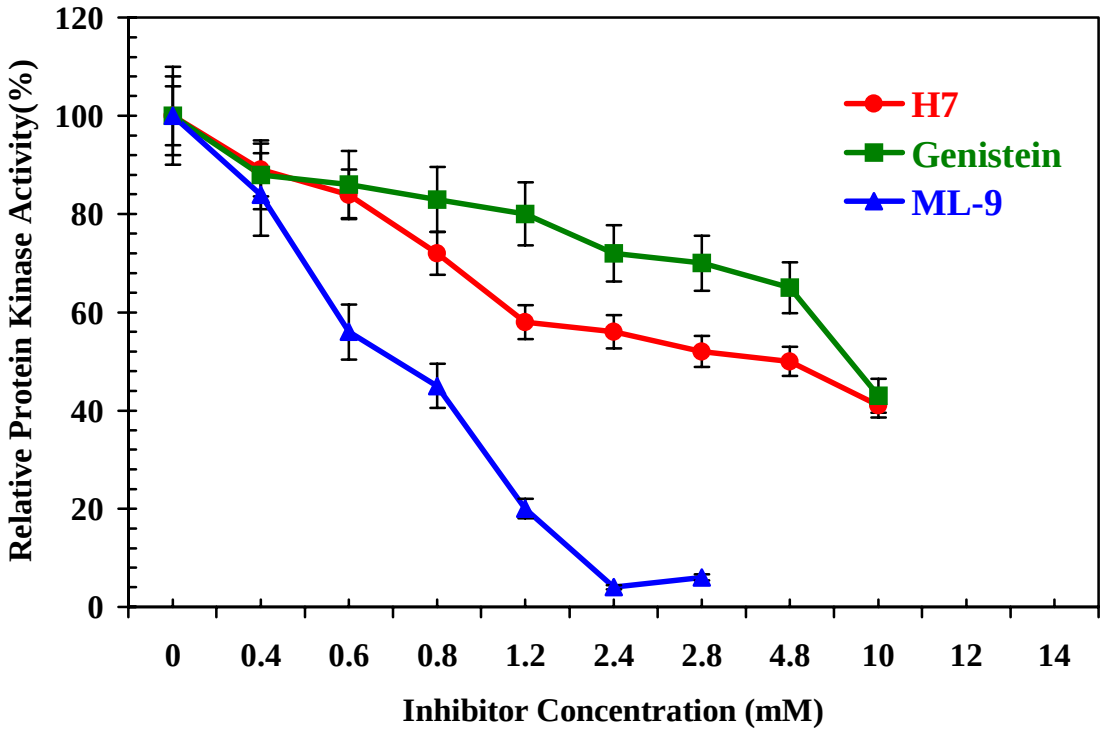


Figure 3-10. Sensitivity of *Sulfolobus solfataricus* Protein Kinase to Inhibitors of Eukaryotic Protein Kinases.

DE-52 fraction, 3.2 µg of protein, was assayed for protein kinase activity toward MLC peptide using the solution assay procedure described in Methods with the exception that, where indicated, the following compounds were present at the concentrations indicated on the ordinate: H7 (•), genistein (❷), and ML-9 (▲). Enzyme activity measured in the absence of inhibitors was 4.0 pmol ^{32}P/min/mg, corresponds to 100% relative activity. Other known eukaryotic protein kinase inhibitors including H89 (1µM - 2 mM), PKI peptide (1 - 200 µg/ml), staurosporine (0.01 - 4.0 mM), and tamoxifen (0.25 - 2.0 mM) were also tested but exhibited little to no effect on protein kinase activity. Shown are the averages of triplicate determinations, ± standard error (Lower et. al., 2000).

performed using two-dimensional electrophoresis (IEF/SDS-PAGE) instead of SDS-PAGE to completely resolve the proteins within the DE-52 fraction. Again, two SDS-polyacrylamide gels were run in parallel, one copolymerized with casein and one without. As anticipated, the presence of a diffuse band of [^{32}P] incorporation at ≈ 67 kDa was observed in both gels (Figure 3-11). It was therefore concluded that the protein kinase itself was the phosphorylated polypeptide species and that autophosphorylation had taken place. Subsequent PAA analysis on the ≈ 67 kDa [^{32}P]-phosphorylated polypeptide indicated that autophosphorylation took place on threonine, as expected (Figure 3-12). Addition of high concentrations of exogenous substrate, MLC peptide, reduced the degree of phosphate incorporation into the protein kinase in a concentration dependent manner (Table 3-7), consistent with catalysis of peptide phosphorylation by the same protein that underwent autophosphorylation. When protein kinase that had been autophosphorylated "in gel" was subsequently in the presence of the diffusible phosphoacceptor substrate MLC peptide, the level of [^{32}P] phosphate associated with the protein kinase remained unaltered. This behavior suggests that the protein bound phosphate group is not a catalytic intermediate, as it could not be transferred to exogenous substrate.

The ability of the *S. solfataricus* protein kinase to autophosphorylate itself using [γ-^{32}P] GTP was also examined. Using a solution assay as described in Methods, DE-52 fraction was incubated with [γ-^{32}P] GTP at 65^{o}C for 1 hour, then resolved using SDS-PAGE. The presence of one band of [^{32}P] phosphorylation was detected at ≈ 67 kDa, indicating that the protein kinase could also utilize GTP as a phosphodonor substrate during autophosphorylation. Subsequent PAA analysis of the [^{32}P]-phosphorylated band indicated the phosphate was bound as phosphothreonine (Figure 3-12).

By contrast to the results obtained from the "in gel" assay using SDS-PAGE, the area of [^{32}P] phosphorylation detected within the two-dimensional gels were observed to be comparatively diffuse, particularly in the first dimension. This phenomenon was observed regardless of whether casein was or was not copolymerized within the SDS-polyacrylamide gel. Such behavior might indicate the diffusion of enzyme through the gel matrix during the extended wash periods, or the existence of multiple active polypeptide species of similar M_r but slightly different pI. To explore this possibility, the

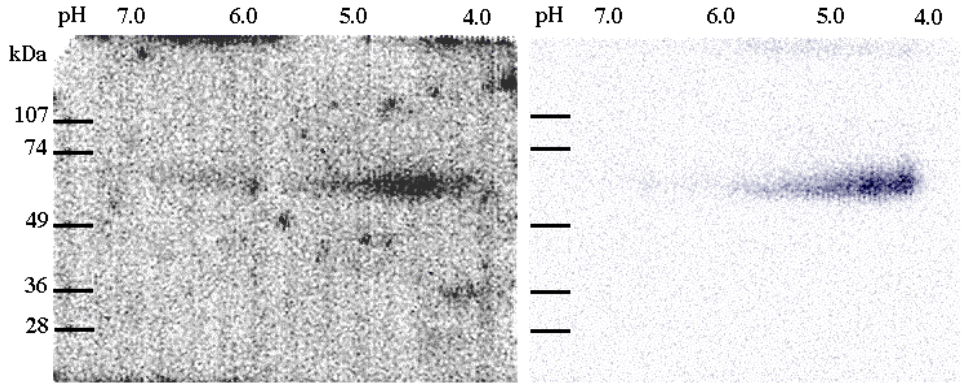
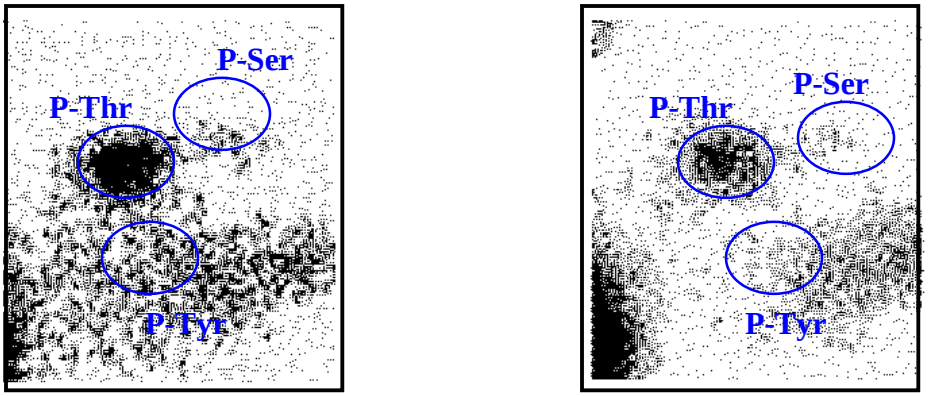


Figure 3-11. The *Sulfolobus solfataricus* Protein Kinase Undergoes Autophosphorylation at Elevated Temperatures.

Shown are the results of an "in gel" assay of *Sulfolobus solfataricus* protein kinase following 2-D Electrophoresis. Portions, containing 35 – 70 μg total protein, of the DE-52 fraction were resolved in the first-dimension by isoelectric focusing in the pH range of 4.0 – 7.0, and in the second-dimension on an 8.0% (w/v) SDS-Polyacrylamide gel in which casein had (left) and had not (right) been copolymerized into the gel matrix at a concentration of 0.5 mg/ml. Following SDS-PAGE, an "in gel" assay was performed as described in Methods. Shown are the autoradiograms of each gel. A single band of ≈ 67 kDa and pI ≈ 4.2 – 5.2 appears in both gels, suggesting that this protein kinase has the ability to both phosphorylate exogenous substrates as well as autophosphorylate itself at elevated temperature. Previous experiments (data not shown) demonstrated that a temperature of at least 40^{o}C is required for autophosphorylation. The molecular weight standards are indicated at left and pH range is indicated across the top of each gel.



ATP **GTP**

Figure 3-12. Phosphoamino Acid Analysis of the *Sulfolobus solfataricus* Protein Kinase Following Autophosphorylation Using Either [γ-^{32}P] ATP or [γ-^{32}P] GTP.

Briefly, 20 μg of DE-52 fraction was incubated with [γ-^{32}P] ATP or [γ-^{32}P] GTP at 65^{o}C during a solution assay as described in Methods. Following the assay, the radiolabeled proteins were isolated by SDS-PAGE and transferred to PVDF membranes. The portion of the membrane containing labeled protein was incubated in 6N HCl for 1 hour at 95^{o}C. The supernatant liquid was then applied to a 20 x 20 cm silica gel TLC plate, along with standards of P-Ser, P-Thr and P-Tyr, then subjected to two-dimensional electrophoresis. The first dimension (left to right) was performed at pH 1.9 and the second dimension (bottom to top) at pH 3.5. Standards were visualized by ninhydrin staining, while phosphorylated amino acids were visualized by electronic autoradiography. Shown are the electronic autoradiograms of the TLC plates from the analysis of the protein kinase autophosphorylated using ATP (left) or GTP (right). Circles indicate the positions of P-Ser, P-Thr and P-Tyr. These results indicate that the protein kinase autophosphorylates exclusively on threonine residues.

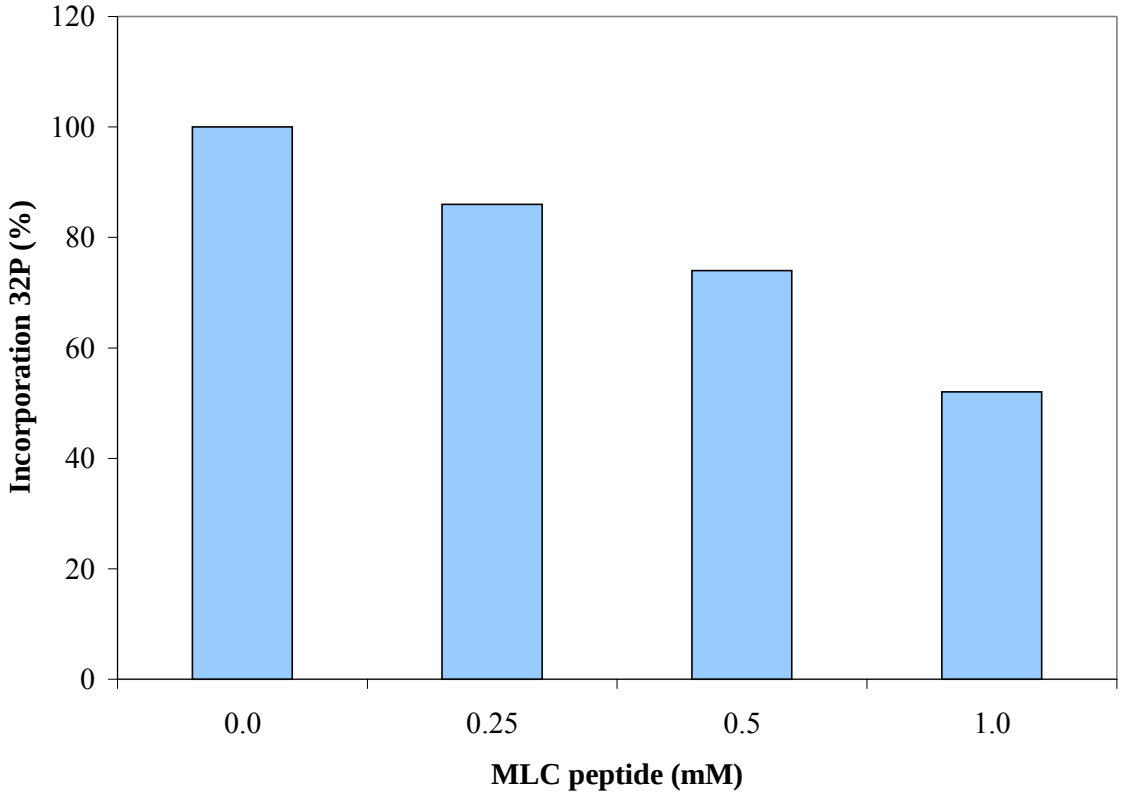


Table 3-7. The Presence of a Substrate Peptide Inhibits Autophosphorylation of the *Sulfolobus solfataricus* Protein Kinase.

Autophosphorylation of the *S. solfataricus* protein kinase was inhibited in a concentration dependent manner by the addition of MLC peptide to the reaction mixture. The DE-52 fraction, 20 μg, was incubated with [γ-^{32}P] ATP at 65^{o}C for 60 minutes as described in the text, with the exception that, where indicated, MLC peptide was added to a final concentration of 0.25, 0.5, or 1.0 mM. The reaction was stopped by precipitating the proteins with 3 volumes of ice-cold acetone. The precipitated proteins were collected by centrifugation and resuspended in SDS-sample buffer and analyzed by SDS-PAGE using an 8% (w/v) SDS-polyacrylamide gel. The amount of [^{32}P]-incorporated into the ≈ 67 kDa protein kinase was quantified by electronic autoradiography. Shown is the percent incorporation of [^{32}P] phosphate by the ≈ 67 kDa as it autophosphorylated itself during the solution assay. A value of 100% corresponds to 1370 CPM.

DE-52 fraction was incubated with [γ-^{32}P] ATP at 65^{o}C to allow the protein kinase to autophosphorylate itself with [^{32}P] phosphate. Next, the radiolabeled proteins were resolved using 2-D electrophoresis and then immediately stained with Coomassie R-250 to minimize the protein diffusion within the gel. As seen in Figure 3-13, a cluster of two to three [^{32}P] phosphorylated polypeptides that migrated to the position corresponding to the previously observed M_r ($\approx$ 67 kDa) and pI (4.2 – 5.2) of the *S. solfataricus* protein kinase were visible as distinctly stained polypeptides. In addition, two other sets of [^{32}P] phosphorylated phosphoproteins with apparent molecular masses of 45 – 50 kDa and $\approx$ 85 kDa were detected. These may represent trapped phosphoenzyme intermediates, autophosphorylated polypeptides that do not become apparent during "in gel" assays because they do not renature in an active form following SDS-PAGE, or perhaps proteins have been phosphorylated by the *S. solfataricus* protein kinase prior to electrophoresis.

In Vivo Glycosylation of the Sulfolobus solfataricus Protein Kinase

The resolution of the *S. solfataricus* protein kinase into multiple [^{32}P] phosphorylated spots by 2-D electrophoresis may, as previously described, represent multiple active polypeptides species or proteolytically nicked forms of the protein kinase. Alternatively, it may indicate the presence of a post-translational modification such as phosphorylation and/or glycosylation. An argument in favor of the former explanation was the appearance of a single major spot of Coomassie-stained protein on a 2-D gel at the position predicted for the *S. solfataricus* protein kinase when prior incubation with ATP was omitted.

To determine whether the *S. solfataricus* protein kinase was glycosylated, SDS-PAGE gels in which the *S. solfataricus* protein kinase was electrophoresed were stained with a GelCode Glycoprotein Staining Reagent (Pierce). Briefly, 20 μg of the DE-52 fraction was incubated with [γ-^{32}P] ATP at 65^{o}C for 60 minutes as described in the Methods to label the protein kinase via autophosphorylation, then proteins resolved using SDS-PAGE. The SDS-polyacrylamide gel was then stained for glycoproteins as described in Methods, using a GelCode Glycoprotein Staining Kit (Pierce). The [^{32}P] phosphorylated polypeptide representing the $\approx$ 67 kDa autophosphorylated protein kinase

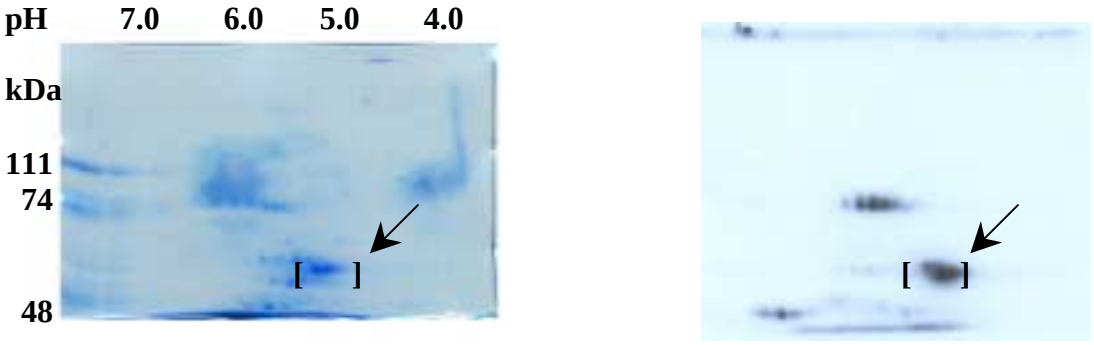


Figure 3-13. Analysis of Autophosphorylated *Sulfolobus solfataricus* Protein Kinase by Two-Dimensional Electrophoresis.

The DE-52 fraction, 75 μg total protein, was incubated at 65°C with [γ-^{32}P] ATP for 60 minutes as described in the Methods. The reaction was quenched by adding 3 volumes of ice-cold acetone. The protein mixture was analyzed by two-dimensional electrophoresis, as described in Methods, using an Immobiline IPG dry strip with a linear pH gradient of 4.0 to 7.0 for the first dimension followed by SDS-PAGE through a 10% (w/v) polyacrylamide gel for the second dimension. The gel was stained for protein, and the [^{32}P] phosphorylated species were visualized by autoradiography using X-ray film. Shown are a picture of the Coomassie-stained gel (left) and an autoradiogram of the gel (right). The brackets [↙] mark the position of the *S. solfataricus* protein kinase. If autophosphorylation was performed "in gel" (Figure 3-11) instead of "pre gel" (seen here), only a single spot of protein and radioactivity appeared at this position [↙], and the other [^{32}P] phosphorylated species observed on this gel do not appear. The latter may represent proteins phosphorylated by the protein kinase in solution, phosphoprotein intermediates, or autophosphorylated proteins that do not renature following electrophoresis, and hence do not become phosphorylated "in gel" (Lower et. al., 2000).

was visualized as a magenta band on the gel, thus indicating that the enzyme was glycosylated. Additionally, a magenta stained polypeptide band coincident with the ≈ 67 kDa [^{32}P] phosphorylated band was also evident following "in gel" assays.

To validate these results, we attempted to chemically remove the carbohydrate moiety from the polypeptide portion of the glycoprotein using trifluoromethanesulfonic acid (TFMS), and then resolve the proteins using SDS-PAGE. If the protein kinase was in fact glycosylated, then its deglycosylated form should be relatively more mobile than the native form on a SDS-polyacrylamide gel. The DE-52 fraction, 75 μg total protein, was incubated with [γ-^{32}P] ATP at 65°C for 60 minutes, treated with TFMS to deglycosylate the glycoproteins contained within, and resolved on a 10% (w/v) SDS-polyacrylamide gel as described in Methods. Next, the gel was then stained with Glycoprotein Staining Reagent (Pierce), as described in Methods, to verify that the TFMS treatment deglycosylated the proteins in the mixture (Figure 3-14). As expected, a [^{32}P] phosphorylated band of ≈ 67 kDa, stained with the Glycoprotein Staining Reagent, was apparent for the sample that had not been treated with TFMS. However, in the sample treated with TFMS, no [^{32}P] phosphorylated species of this apparent mass was observed. Instead, a [^{32}P] phosphorylated band of significantly smaller M_r (≈ 62 kDa), presumably the *S. solfataricus* protein kinase, was observed. Furthermore, no sugar moieties were detected by the Glycoprotein Staining Reagent, indicating that the TFMS treatment resulted in the complete deglycosylation of all glycoproteins present in the sample. It was therefore concluded that the protein kinase did indeed undergo glycosylation *in vivo*. Rather unexpectedly, five other glycoproteins with apparent molecular masses of > 107 kDa, ≈ 100 kDa, ≈ 58 kDa, ≈ 48 kDa, and ≈ 28 kDa were detected in the DE-52 fraction by the Glycoprotein Staining Reagent. This suggests that glycosylation is a common phenomenon among membrane-associated proteins of *S. solfataricus*.

Characterization of the Glycosylation Structure of the Sulfolobus solfataricus Protein Kinase

We used a DIG Glycan Differentiation Kit (Roche) to identify and characterize the carbohydrate structure of the *S. solfataricus* protein kinase. This kit uses lectins,

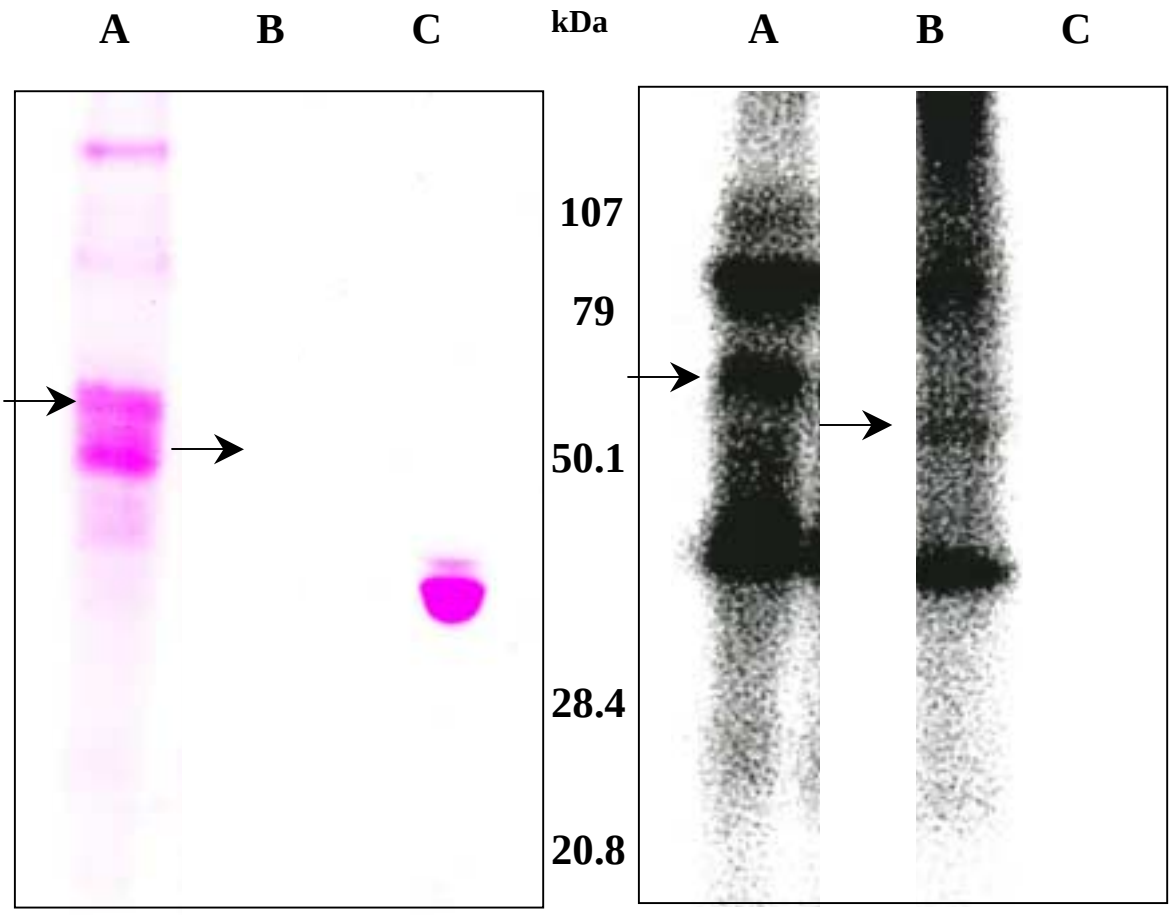


Figure 3-14. The *Sulfolobus solfataricus* Protein Kinase Appears to be Glycosylated.

The DE-52 fraction, 75 µg total protein, was incubated with [γ-^{32}P] ATP at 65^{o}C for 60 during a solution assay as described in Methods. The reaction was quenched by the addition of 3 volumes of ice-cold acetone. The precipitated proteins were collected by centrifugation, resuspended in 100 µl 0.25% (w/v) SDS, and lyophilized for 18 hours. The dry protein sample was then treated with anhydrous trifluoromethanesulfonic acid (TFMS) to cleave both N- and O-linked glycans (lane B) as described in Methods. Next, the proteins were resolved on a 10% (w/v) polyacrylamide gel and then glycoproteins were detected using GelCode Glycoprotein Staining Reagent as described in Methods. The [^{32}P]-phosphorylated species were visualized by electronic autoradiography. Shown is a picture of the GelCode-stained gel (left) and an autoradiogram of the gel (right). Lane A contains phosphorylated DE-52 fraction only, Lane B contains phosphorylated DE-52 fraction treated with TFMS, and Lane C contains Horseradish Peroxidase (Positive Control for GelCode Glycoprotein Stain). The arrow (➔) marks the position of the *S. solfataricus* protein kinase. Note that following treatment with TFMS (lane B), the *S. solfataricus* protein kinase has a higher mobility (smaller M_r), and that the Gel Code Glycoprotein Staining Reagent does not detect the presence of glycosylation.

conjugated with steroid hapten digoxigenin, that bind to specific carbohydrate moieties of glycoproteins allowing for the immunological identification of the carbohydrate structure. The DE-52 fraction, 1.0 μg total protein, was incubated with [γ-^{32}P] ATP at 65^{o}C for 60 minutes to radiolabel the protein kinase via autophosphorylation. The phosphorylated proteins, control glycoproteins (carboxypeptidase Y was used as the control glycoprotein for *Galanthus nivalis* agglutinin (GNA); transferrin was used as the control glycoprotein for *Sambucus nigra* agglutinin (SNA); fetuin was used as the control glycoprotein for SNA, *Maackia amurensis* agglutinin (MAA), and *Datura stramonium* agglutinin (DSA); and asialofetuin was used as the control glycoprotein for Peanut agglutinin (PNA) and DSA and molecular weight marker, were then separated using SDS-PAGE and electroblotted onto Immobilon-P as described in Methods. The membrane was separated into sections (each section containing phosphorylated DE-52 fraction, molecular weight marker, and appropriate control glycoproteins) and incubated with the appropriate digoxigenin conjugated lectin as described in Methods. As seen in Figure 3-15, the only lectin that bound to the ≈ 67 kDa [^{32}P] autophosphorylated protein kinase was the GNA (Galanthus nivalis agglutinin) lectin. This lectin recognizes terminal mannose, α(1-3), α(1-6), or α(1-2) linked to mannose, and thus is suitable for identifying "high mannose" N-glycan chains or O-glycosidically linked mannoses in glycoproteins. The GNA lectin also detected the presence of several other glycoproteins contained within the DE-52 fraction.

Attempts to Determine a Partial Amino Acid Sequence for the Sulfolobus solfataricus Protein Kinase

Our goal during the purification of the *S. solfataricus* protein kinase was to obtain enough highly purified protein kinase that a partial amino acid sequence of the enzyme could be determined. Oligonucleotide primers modeled on this partial sequence could then be used to clone its gene using PCR, or used to probe a genomic library to identify clones containing the gene. Once a DNA clone was obtained, its oligonucleotide sequence could be determined and the primary sequence of the encoded protein kinase deduced there from.

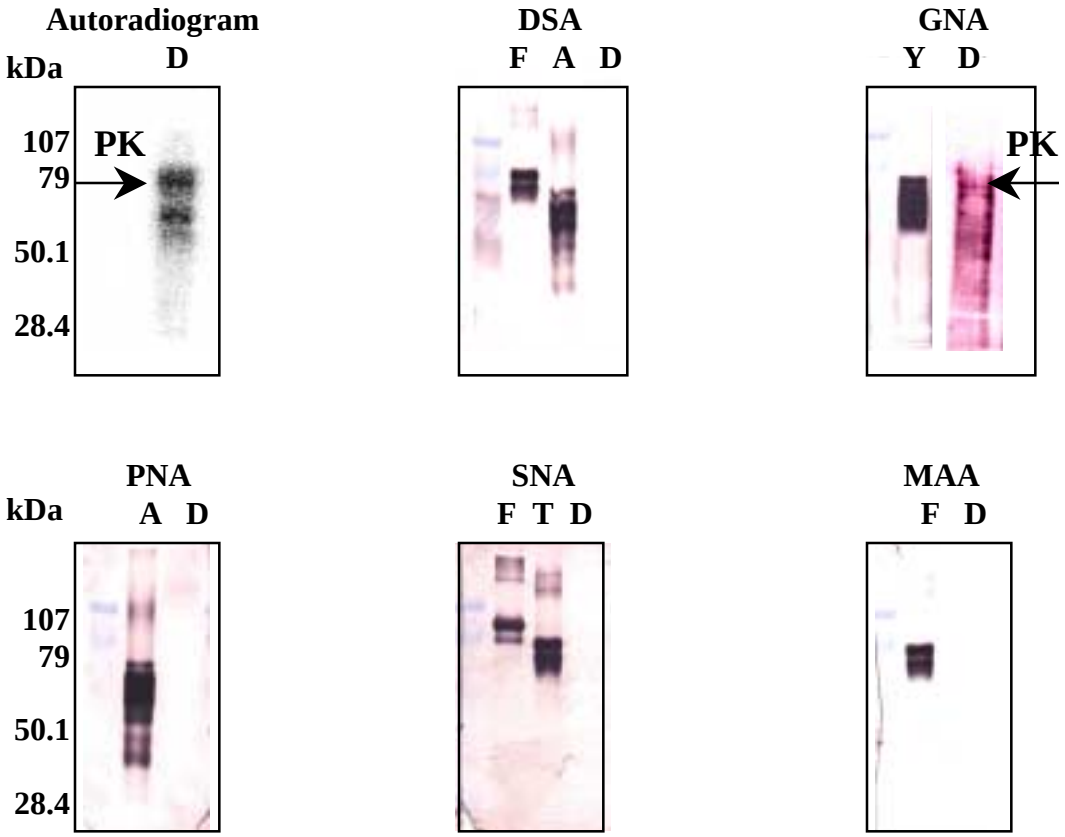


Figure 3-15. The *Sulfolobus solfataricus* Protein Kinase has Glycosidically Linked Terminal Mannose Sugars.

The DE-52 fraction (**D**) , 1.0 µg total protein, was incubated with [γ-^{32}P] ATP at 65^{o}C for 60 during a solution assay as described in Methods. The phosphorylated proteins along with control glycoproteins carboxypeptidase Y (**Y**), transferrin (**T**), fetuin (**F**), and asialofetuin (**A**) were then resolved on a 10% (w/v) SDS-polyacrylamide gel and transferred to Immobilon-P membrane as described in Methods. The membrane was then cut into sections and each section incubated in 10 ml Tris-Buffered Saline, 1 mM $MgCl_2$, 1 mM $MnCl_2$, and 1 mM $CaCl_2$ containing the appropriate digoxigenin conjugated lectin for 1 hour as described in Methods. Glycoproteins were visualized as a dark/light brown bands by incubating the membranes in NBT/X-phosphate solution as described in Methods. Shown is an autoradiogram with an arrow (➔) indicating the position of the *S. solfataricus* protein kinase (**PK**) and each of the five membranes labeled with either **DSA**, **GNA**, **MAA**, **PNA**, or **SNA** lectin. Molecular weight markers are indicated at left. Note that the only lectin to bind to the glycoproteins contained within the DE-52 fraction was GNA, which selectively recognizes terminal mannose sugars.

Attempts to purify the *S. solfataricus* protein kinase met with limited success (Table 3-1). We experimented with a variety of purification techniques including ion-exchange chromatography, size exclusion chromatography, chromatofocusing chromatography (using FPLC), hydrophobic interaction chromatography, lectin affinity chromatography (Concanavalin A or *Galanthus nivalis*), and heparin affinity chromatography until we finally settled on the purification scheme outlined in Table 3-1, as this was the one which resulted in the highest yield of catalytically active enzyme. Efforts to purify the protein kinase contained within either the DE-52 or Mono-P fraction using one-dimensional electrophoresis (SDS-PAGE) were unsuccessful. This technique was limited in its ability to resolve the enzyme from other polypeptides contained within either fraction, and thus did not permit the unambiguous identification of the polypeptide species responsible for the protein kinase activity. Therefore, we attempted to use two-dimensional electrophoresis (IEF followed by SDS-PAGE) to purify the protein kinase, as this technique offers approximately 100 – 200 fold greater resolving power than SDS-PAGE alone (Link, 1999). The DE-52 fraction, 75 – 125 μg, was incubated with [γ-^{32}P] ATP at 65^{o}C for 1 hour to allow the protein kinase to autophosphorylate. Next, the [^{32}P] phosphorylated proteins were resolved using 2-D electrophoresis and stained with Coomassie R-250 as described in Methods. As seen in Figure 3-16, two [^{32}P] phosphorylated polypeptides that migrated to the position corresponding to the previously observed M_r (≈ 67 kDa) and pI (4.2 – 5.2) of the *S. solfataricus* protein kinase were visible as distinctly stained polypeptides. These spots were excised using a clean razor blade and sent to the W. M. Keck Foundation at Yale University for "in gel" trypsin digestion followed by LC/MS or MALDI-MS protein identification. Two additional 2-D gels were run in parallel and the proteins within electroblotted onto an Immobilon-P membrane. The two [^{32}P] phosphorylated polypeptides corresponding to the protein kinase were excised with a clean razor bland and sent to the W. M. Keck Foundation at Yale University and to Dr. Mark Lively at Wake Forest University School of Medicine for amino terminal sequencing. In each case, no amino acid sequence data was obtained for the polypeptide. Possible explanations given for why the MS analysis failed include insufficient protein in the gel slice, failure of the protein to digest well with trypsin, or poor ionization of peptides MS analysis. Possible explanations given for why the amino

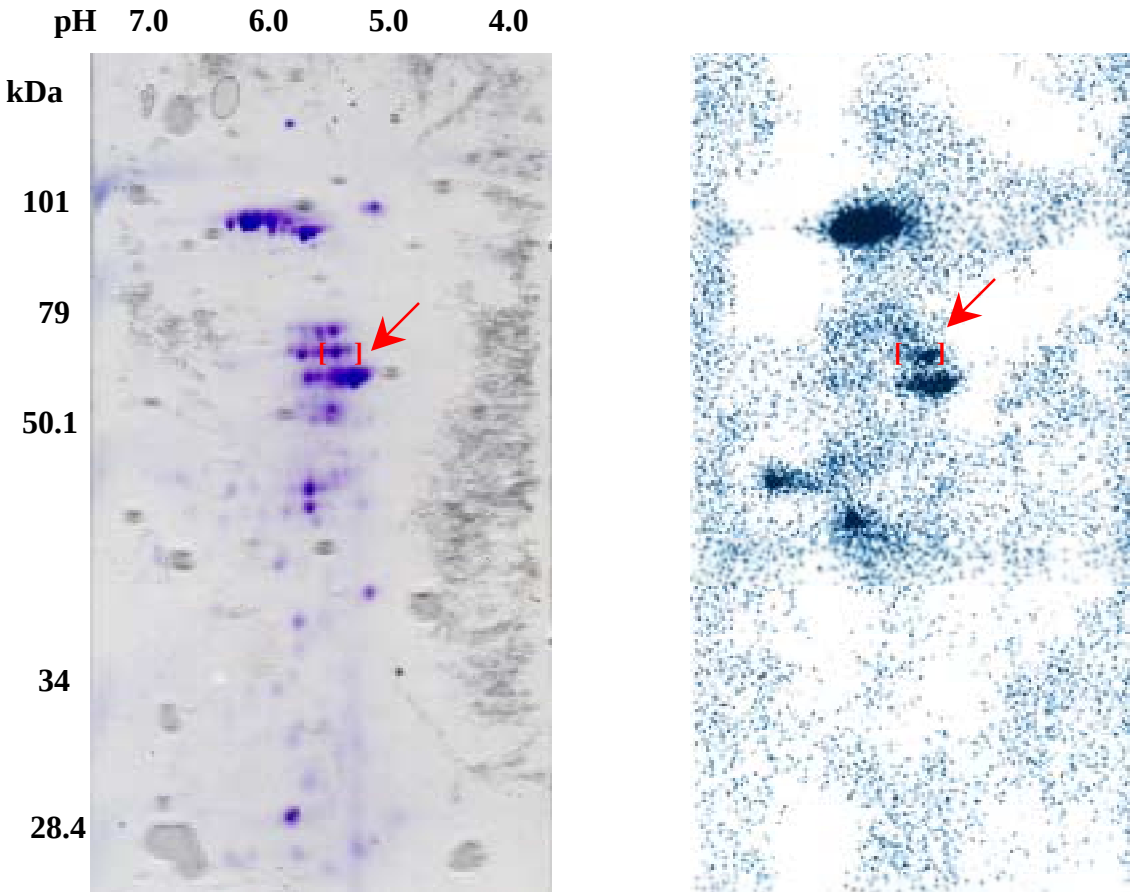


Figure 3-16. Analysis of Autophosphorylated *Sulfolobus solfataricus* Protein Kinase by Two-Dimensional Electrophoresis.

The DE-52 fraction, 75 µg total protein, was incubated with [γ-^{32}P] ATP at 65^{o}C for 60 minutes to allow the *S. solfataricus* protein kinase to autophosphorylate itself with [^{32}P] phosphate. The reaction was quenched by adding 3 volumes of ice-cold acetone. The precipitated protein was collected by centrifugation and resuspended in rehydration solution as described in Methods. The protein mixture was analyzed by two-dimensional electrophoresis using a 7 cm Immobiline gel with a linear pH gradient of 4.0 to 7.0 for the first dimension followed by SDS-PAGE through a 16 x 16 cm 10% (w/v) SDS-polyacrylamide gel for the second dimension. The gel was stained for protein with Coomassie R-250, and the [^{32}P] phosphorylated species were visualized by electronic autoradiography. Shown are a picture of the Coomassie-stained gel (left) and an autoradiogram of the gel (right). The brackets mark the position of the *S. solfataricus* protein kinase.

terminal sequencing failed was that the N-terminal end of the protein was blocked, or that insufficient protein transferred to the PVDF membrane for analysis.

CHAPTER IV

Identification of Open Reading Frames in *Sulfolobus solfataricus* Whose Predicted Products Exhibit Homology to Eukaryotic Protein Kinases.

Since biochemical efforts to reveal the identify the *S. solfataricus* protein kinase failed, we decided to take a genomic approach to identify the gene encoding this enzyme. Our working hypothesis was that the source of the protein kinase activity was a member of the "eukaryotic" family of protein kinases. We employed two methods to search the partial genome (complete genome finished in June, 2000) of *Sulfolobus solfataricus* for hypothetical proteins whose primary sequence displayed sequence features characteristic of "eukaryotic-like" protein-serine/threonine/tyrosine kinases. The first method, employed two eukaryotic protein kinase "signature motifs" as templates for the initial identification of any hypothetical proteins kinases. The first motif (Gly-X-Gly-X-X-Gly-X) represents the glycine-rich ATP binding domain (subdomain I) of eukaryotic protein kinases (Smith and King, 1995). The second motif (Asp-X_4-Asn) represents the catalytic loop (subdomain VIb) of cAMP-dependent protein kinases. The second method, utilized the sequence of the catalytic subunit of eukaryotic protein kinase from *Candida albicans* (gi|11596395|) to identify *ORFs* encoding hypothetical protein kinases. Its sequence was obtained electronically from GenBank, and selected as template because it contained all twelve characteristic subdomains of cAMP-dependent protein kinases.

Candidate hypothetical proteins were identified using a web-based BLAST (basic local alignment search tool) server from the *Sulfolobus* MAGPIE project (http://niji.imb.nrc.ca/sulfolobus/). Protein sequences from the *Sulfolobus* project were aligned against either subdomain I (Gly-X-Gly-X-X-Gly-X), or subdomain VIb (Asp-X_4-Asn), or the catalytic subunit of cAMP-dependent protein kinase from *C. albicans* using a blastp alignment tool. Where necessary, alignments were adjusted manually to maximize apparent similarity in the areas of signature motifs conserved among known eukaryotic protein kinases (described in Introduction). Visual inspection or inspection using the ExPASy (Expert Protein Analysis System) ScanProsite Program (http://www.expasy.ch/) was also performed to eliminate those hypothetical proteins lacking subdomain VIb as

this subdomains is considered absolutely essential to constitute a plausibly functional enzyme. Hypothetical transmembrane domains were identified using the web-based DAS (Dense Alignment Surface Method) Transmembrane Prediction Server (http://www.sbc.su.se/~miklos/DAS/maindas.html) from the Stockholm Bioinformatic Center in Stockholm, Sweden. This algorithm uses low-stringency dot-plots of the query sequence against a collection of non-homologous membrane proteins using a previously derived, special scoring matrix to identify putative transmembrane helices contained within a protein's primary sequence (Cserzo et. al., 1997).

Six hypothetical proteins were identified from *Sulfolobus solfataricus* whose primary sequence exhibited noticeable similarities to eukaryotic protein kinases. The *ORFs* for the hypothetical proteins were *sso0197, sso0433, sso2291, sso2387, sso3182,* and *sso3207*. Five of the six hypothetical proteins (*ORFs sso0197, sso2291, sso2387, sso3182, sso3207*) have both the nucleotide binding subdomain and the catalytic subdomain, while the remaining hypothetical protein (*ORFs sso0433*) has the essential catalytic subdomain but lacks an apparent nucleotide binding motif. Although *ORF sso0433* must be viewed with caution, functional exceptions having substitutions in the nucleotide binding domain are known (Ryazanov et. al., 1997).

ORF sso0197 encodes a 287 amino acid hypothetical protein having a MW of ≈ 33.1 kDa, and pI of ≈ 9.5 (Figure 4-1), which contains two hypothetical transmembrane domain as identified using DAS (Figure 4-2). This hypothetical protein contains what appears to be Gly_{50} and Gly_{52} of subdomain I, as well as the highly conserved lysine residue of subdomain II (at position 71), and glutamate of subdomain III (at position 90). It also possesses the completely conserved catalytic domain (subdomain VI_b) and plausible candidates for subdomain V, the glutamate of subdomain VII, and the arginine of subdomain XI.

ORF sso0433 encodes a 223 amino acid hypothetical protein having a MW of ≈ 26 kDa, and pI of ≈ 6.4 (Figure 4-3), but contains no hypothetical transmembrane domains as defined using DAS (Figure 4-4). However, it does contain two putative myristylation sites. Its primary sequence of 223 amino acids, while being considerably

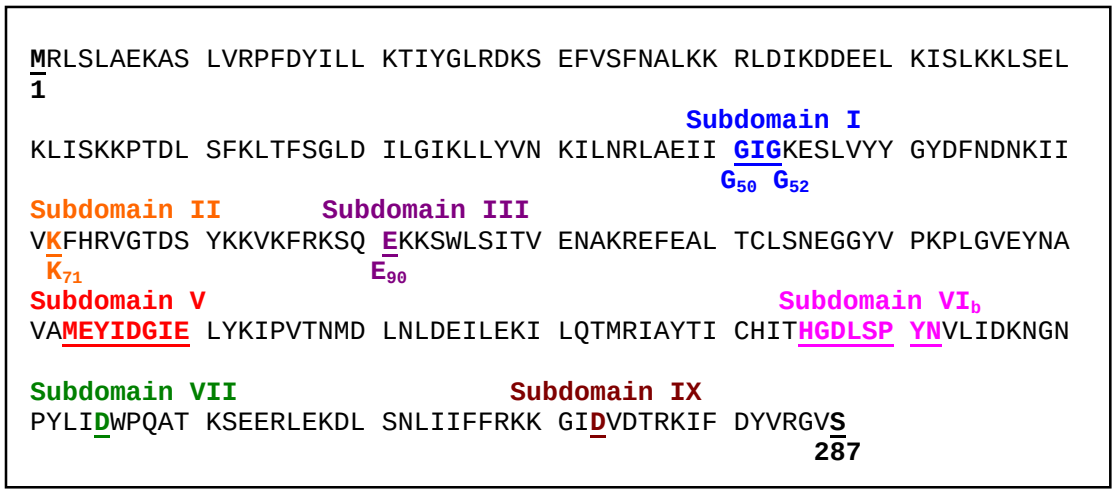
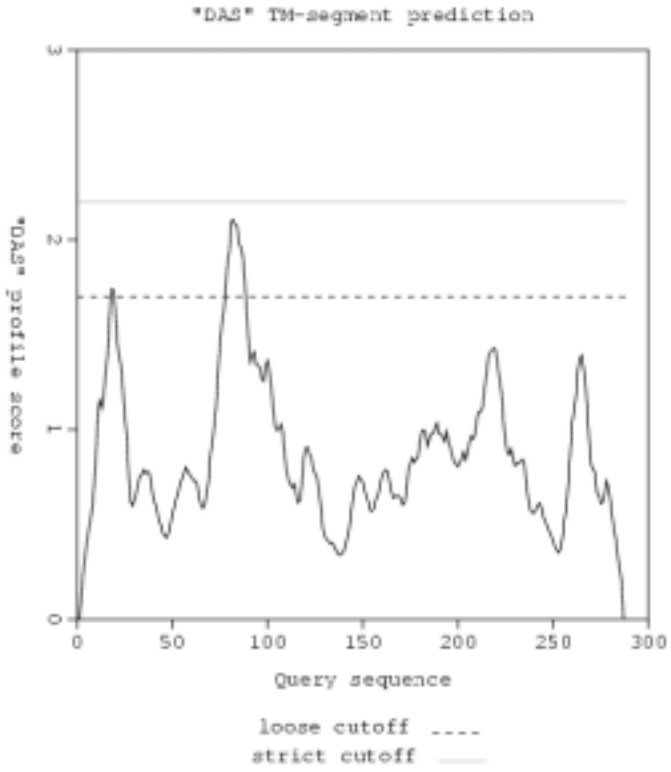


Figure 4-1. The Amino Acid Sequence of the ≈ 33 kDa Hypothetical Protein Encoded by *Sulfolobus solfataricus ORF sso0197* Exhibits Similarity to Eukaryotic Protein Kinases.

Inspection of the partial genome of *S. solfataricus* for homologs of eukaryotic protein kinases, as described in the text, revealed the presence of *ORF sso0197* whose predicted product displays sequence features characteristic of eukaryotic protein-serine/threonine/tyrosine kinases. *S. solfataricus ORF sso0197* encodes a hypothetical protein consisting of 287 amino acids, with a MW of ≈ 33.1 kDa, and pI of ≈ 9.5. Shown is its deduced amino acid sequence and prospective candidates for the subdomains or essential residues characteristic of eukaryotic protein kinases (described in Introduction).



Start	Stop	Length	Cutoff
18	19	2	1.7
78	88	11	1.7

Figure 4-2. Potential Transmembrane Helices of the ≈ 33 kDa Hypothetical Protein Encoded by *Sulfolobus solfataricus ORF sso0197*.

Hypothetical transmembrane domains were identified using the web-based DAS (Dense Alignment Surface Method) Transmembrane Prediction Server (http://www.sbc.su.se/~miklos/DAS/maindas.html) from the Stockholm Bioinformatic Center in Stockholm, Sweden. This algorithm uses low-stringency dot-plots of the query sequence against a collection of non-homologous membrane proteins using a previously derived, special scoring matrix to identify putative transmembrane helices contained within a protein's primary sequence. Shown is the DAS curve for the hypothetical protein encoded by *S. solfataricus ORF sso0197* (DAS profile score as a function of the amino acid sequence). A DAS profile score of 1.7 characterizes a loose cutoff and a DAS profile score of 2.2 characterizes a strict cutoff for defining putative transmembrane domains. Also shown are the start and stop amino acids, as well as the length of the potential transmembrane segment. As seen in the graph two putative transmembrane domains are defined using DAS.

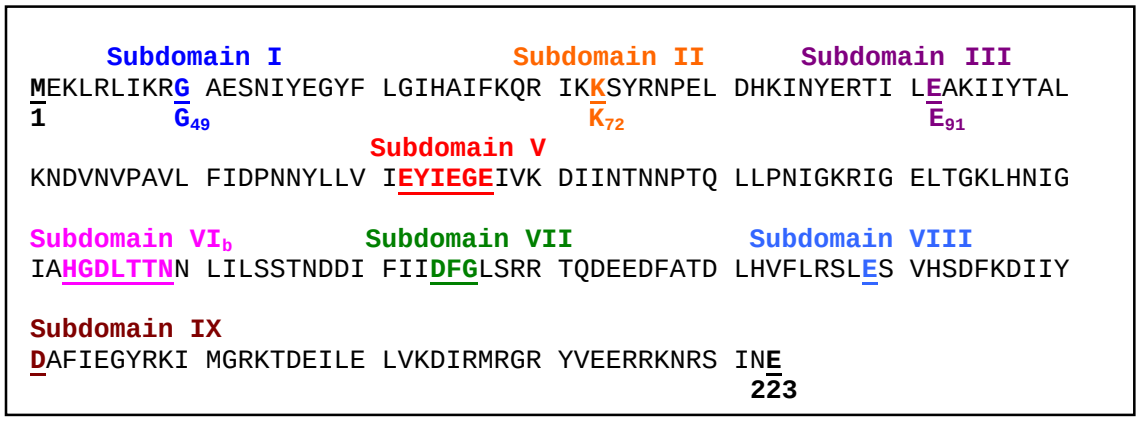


Figure 4-3. The Amino Acid Sequence of the ≈ 26 kDa Hypothetical Protein Encoded by *Sulfolobus solfataricus ORF sso0433* Exhibits Similarity to Eukaryotic Protein Kinases.

Inspection of the partial genome of *S. solfataricus* for homologs of eukaryotic protein kinases, as described in the text, revealed the presence of *ORF sso0433* whose predicted product displays sequence features characteristic of eukaryotic protein-serine/threonine/tyrosine kinases. *S. solfataricus ORF sso0433* encodes a hypothetical protein consisting of 223 amino acids, with a MW of ≈ 26 kDa, and pI of ≈ 6.4. Shown is its deduced amino acid sequence and prospective candidates for the conserved subdomains or essential residues characteristic of eukaryotic protein kinases (described in Introduction).

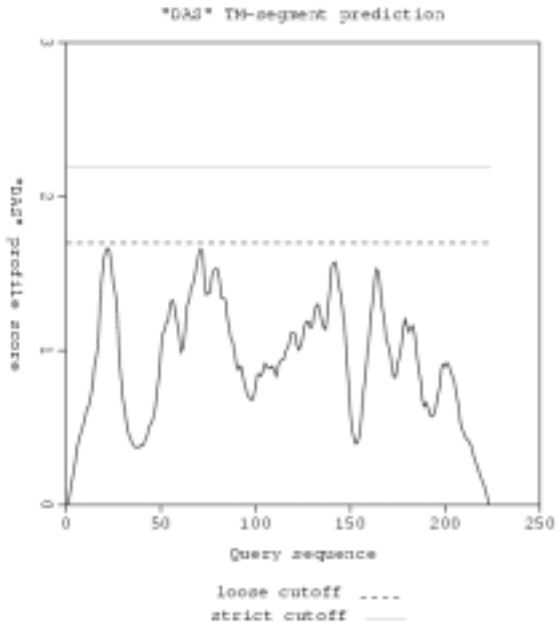


Figure 4-4. Potential Transmembrane Helices of the ≈ 26 kDa Hypothetical Protein Encoded by *Sulfolobus solfataricus ORF sso0433*.

Hypothetical transmembrane domains were identified using the web-based DAS (Dense Alignment Surface Method) Transmembrane Prediction Server (http://www.sbc.su.se/~miklos/DAS/maindas.html) from the Stockholm Bioinformatic Center in Stockholm, Sweden. This algorithm uses low-stringency dot-plots of the query sequence against a collection of non-homologous membrane proteins using a previously derived, special scoring matrix to identify putative transmembrane helices contained within a protein's primary sequence. Shown is the DAS curve for the hypothetical protein encoded by *S. solfataricus ORF sso0433* (DAS profile score as a function of the amino acid sequence). A DAS profile score of 1.7 characterizes a loose cutoff and a DAS profile score of 2.2 characterizes a strict cutoff for defining putative transmembrane domains. As is seen in the graph no obvious transmembrane domains are defined using DAS.

smaller than a "typical" catalytic core, which is usually ≈ 280 amino acids in length (Hanks and Hunter, 1995), contains the two highly conserved amino acid residues Lys_{72} (subdomain II) and Glu_{91} (subdomain III), as well as subdomains V, VI_b, and VII, and plausible candidates for subdomains I, IX, and XI. No obvious ATP binding motif was apparent within its primary sequence however.

ORF sso2291 encodes a 554 amino acid hypothetical protein having a MW of ≈ 62.9 kDa and pI of ≈ 9.7 (Figure 4-5), that contains seven hypothetical transmembrane domains as identified using DAS (Figure 4-6). Its primary sequence contains eight of the twelve highly conserved active-site subdomains including subdomain I, II, III, V, VI_b, VII, IX, and XI (Figure 4-7).

ORF sso3182 encodes a 537 amino acid hypothetical protein having a MW of ≈ 60.6 kDa and pI of ≈ 9.5 (Figure 4-8), that contains five hypothetical transmembrane domains as identified using DAS (Figure 4-9). This hypothetical protein possesses nine of the twelve highly conserved active-site subdomains including I, II, III, V, VI_b, VII, VIII, IX, and XI (Figure 4-10), as well as a putative eukaryotic RNA recognition motif spanning residues 57 – 64.

ORF sso3207 encodes a 669 amino acid hypothetical protein having a MW of ≈ 77357 kDa and pI of ≈ 5.6 (Figure 4-11), that contains six hypothetical transmembrane domains as identified using DAS (Figure 4-12). It possesses eight of the twelve highly conserved active-site subdomains including I, II, III, V, VI_b, VII, IX, and XI (Figure 4-13).

ORF sso2387 encodes a 583 amino acids hypothetical protein having a MW of ≈ 66.5 kDa and pI of ≈ 6.7 (Figure 4-14), and contains no apparent transmembrane domain as defined using DAS (Figure 4-15). It also contains four putative myristylation sites and a putative cell attachment sequence from residues 105 – 152. This hypothetical protein contains subdomain I as well as the highly conserved lysine residue of subdomain II (at position 72), and glutamate of subdomain III (at position 81) as well as a completely conserved catalytic domain (subdomain VI_b) and plausible candidates for subdomains V, VII, VIII, and XI.

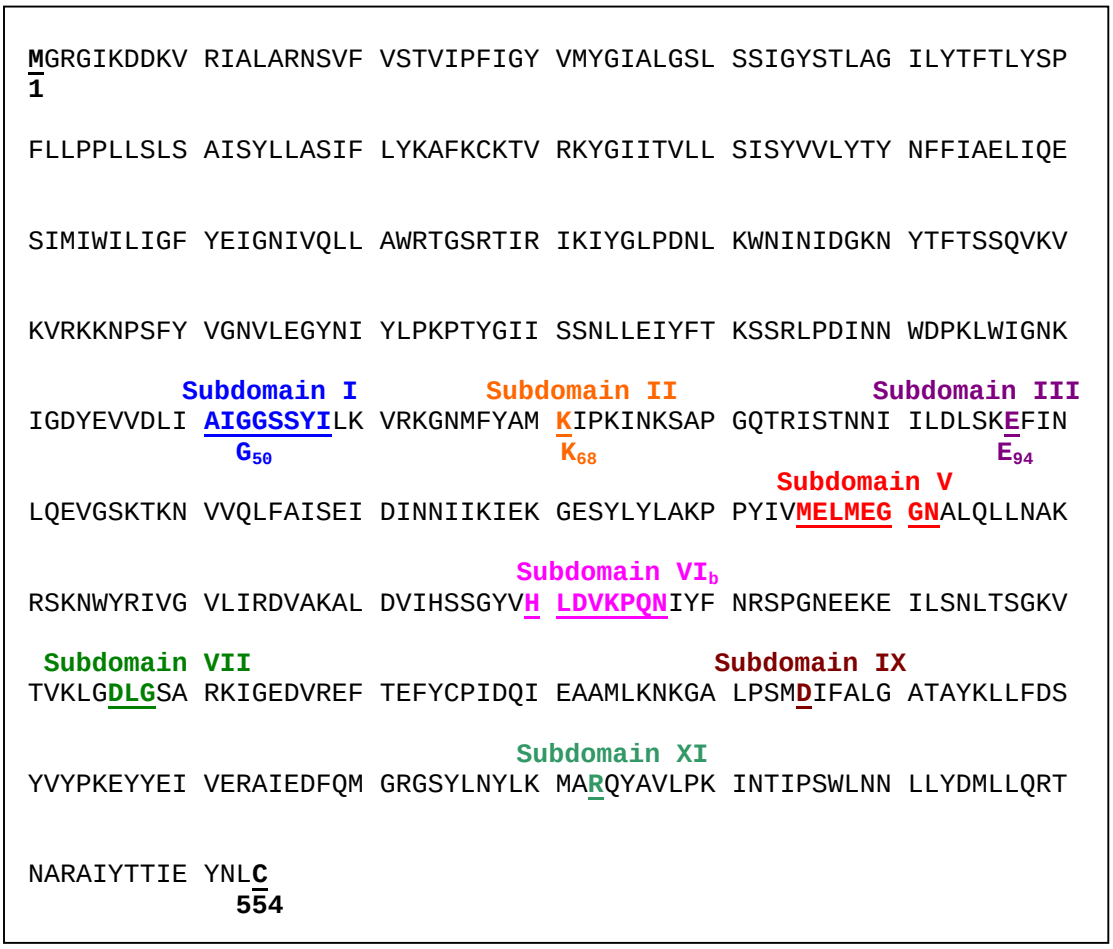
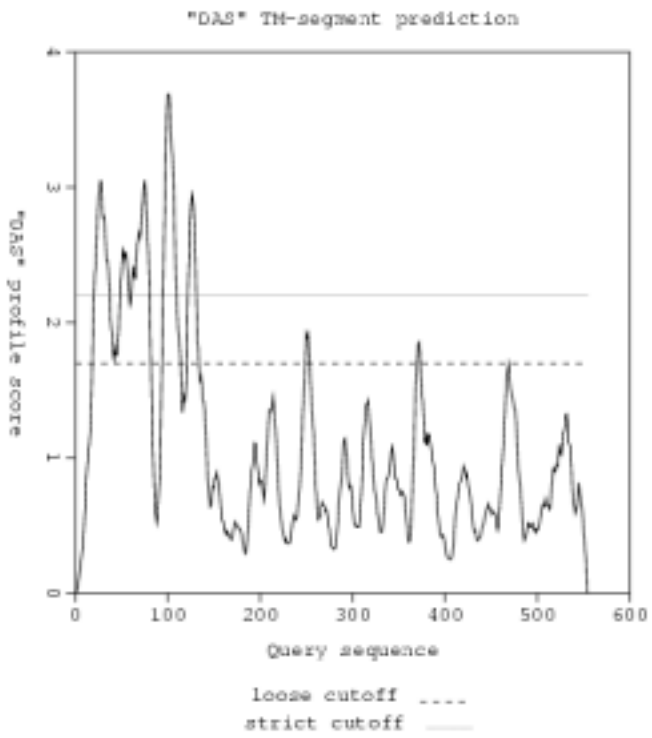


Figure 4-5. The Amino Acid Sequence of the ≈ 63 kDa Hypothetical Protein Encoded by *Sulfolobus solfataricus ORF sso2291* Exhibits Similarity to Eukaryotic Protein Kinases.

Inspection of the partial genome of *S. solfataricus* for homologs of eukaryotic protein kinases, as described in the text, revealed the presence of *ORF sso2291* whose predicted product displays sequence features characteristic of eukaryotic protein-serine/threonine/tyrosine kinases. *S. solfataricus ORF sso2291* encodes a hypothetical protein consisting of 554 amino acids, with a MW of ≈ 63 kDa, and pI of ≈ 9.7. Shown is its deduced amino acid sequence and prospective candidates for conserved subdomains or essential residues characteristic of eukaryotic protein kinases (described in Introduction).



Start	Stop	Length	Cutoff	Start	Stop	Length	Cutoff
19	81	63	1.7	95	109	15	2.2
21	37	17	2.2	121	133	13	1.7
49	58	10	2.2	122	130	9	2.2
61	80	20	2.2	249	253	5	1.7
94	114	21	1.7	370	373	4	1.7

Figure 4-6. Potential Transmembrane Helices of the ≈ 63 kDa Hypothetical Protein Encoded by *Sulfolobus solfataricus ORF sso2291*.

Hypothetical transmembrane domains were identified using the web-based DAS (Dense Alignment Surface Method) Transmembrane Prediction Server (http://www.sbc.su.se/~miklos/DAS/maindas.html) from the Stockholm Bioinformatic Center in Stockholm, Sweden. This algorithm uses low-stringency dot-plots of the query sequence against a collection of non-homologous membrane proteins using a previously derived, special scoring matrix to identify putative transmembrane helices contained within a protein's primary sequence. Shown is the DAS curve for the hypothetical protein encoded by *S. solfataricus ORF sso2291* (DAS profile score as a function of the amino acid sequence). A DAS profile score of 1.7 characterizes a loose cutoff and a DAS profile score of 2.2 characterizes a strict cutoff for defining putative transmembrane domains. As is seen in the graph ten putative transmembrane domains are defined using DAS.

```
                                10        20        30        40        50        60
Subdomain                     .    | I    .    |    .  II    .    |    .    |    .    |
1JNK     Human     64    YQNLKPIGSGAQGIVCAAYDAVLDRNVAIKKLSRPFQ-----NQTHAK--------RAYR  110
sso2291  S.solfat  244   YEVVDLIAIGGSSYILKVR--KGNMFYAMKIPKINKS-----APGQTRISTNNIILDLSK  296
DCAMKL1  Rattus    183   YKVGRTIGDGNFAVVKECIERSTAREYALKIIKKSKC------RGKEH--------MIQN  128
STPK     Yeast     6     YRDLQLIGQGSFGSVFRAQDVESSKIVALKVVDLD--A----TKDQIE--------TLTQ  51
DCEK1    Yeast     589   YKILKPISKGAFGSVYLAQKRTTGDYFAIKILKKSNM------KNQVI--------NVRA  636
PCKu     Human     583   IFPDEVLGSGQFGIVYGGKHRKTGRDVAIKIIDKLRF--------QES--------QLRN  629
MDPK     Human     81    FEILKVIGRGAFSEVAVVKMKQTGQVYAMKIMNKWDM--------EVS--------CFRE  128
ROCK1    Mouse     76    YEVVKVIGRGAFGEVQLVRHKSTRKVYAMKLLSKFEMI----KRSDSA--------FFWE  123
COT1     Neurospo  214   YQTIKIIGKGAFGEVKLVQKKADGKVYAMKSLIKTEM--------QLA--------HVRA  261
CDS1     Yeast     167   YEIIRTLGSGTFAVVKLAVEVNSGKWYAIKIINKRKILLTSSEKRATE--------MFQR  218

                     70        80        90       100       110       120
               III  .    |    .    |    .    |    .    |    .    |  V   |
1JNK     111   ELVLMKCV-NHK--NIISLLNVFTPQKTLE-----EFQDVY------LVMELMDA-NLCQ  155
query    297   EFINLQEV-GSKTKNVVQLFAISEIDINNIIKIEKGESYLYLAKPPYIVMELMEGGNALQ  355
6225243  129   EVSILRRV-KHP--NIVLLIEEMDVP-----------TELY------LVMELVKGGDLFD  168
7493389  52    EINFLIDL-NSV--HITKYYASFVDG-----------FRLW------ITMEYCDGGSCLD  91
2507198  637   ERAILMSQGESP--FVAKLYYTFQSK-----------DYLY------LVMEYLNGGDCGS  677
2499575  630   EVAILQNL-HHP--GVVNLECMFETP-----------ERVF------VVMEKLHGDMLEM  669
1706450  129   ERDVLVNG-DRR--WITQLHFAFQDE-----------NYLY------LVMEYYVGGDLLT  168
6677759  124   ERDIMAFA-NSP--WVVQLFYAFQDD-----------RYLY------MVMEYMPGGDLVN  163
6166017  262   ERDILAES-DSP--WVVKLYTTFQDA-----------NFLY------MLMEFLPGGDLMT  301
1168875  219   EIDILKSL-HHP--GVVQCHEICEND-----------DELF------IVMEYVEGGDLMD  258

                    130       140       150       160       170       180
                  .    |    .    |    .    |    .  VIb   .    |    .    |
1JNK     156   VIQME---LDHERMS---YLLYQMLCGIKHLHSAGIIHRDLKPSNIVVKSD----C----  201
query    356   LLNAK---RSKNWYRIVGVLIRDVAKALDVIHSSGYVHLDVKPQNIYFNRS----PGNEE  408
6225243  169   AITSTS-KYTERDAS---GMLYNLASAIKYLHSLNIVHRDIKPENLLVYEHQDGSK----  220
7493389  92    LLKLSG-TFSERVIA---EVMRQVLEALVYLHGQGKMHRDIKAANILTMKD-----G---  139
2507198  678   LLKTMG-VLDLDWIR---TYIAETVLCLGDLHDRGIIHRDIKPENLLISQN----G----  725
2499575  670   ILSSEKGRLPEHITK---FLITQILVALRHLHFKNIVHCDLKPENVLLASA-----DP--  719
1706450  169   LLSKFGERIPAEMAR---FYLAEIVMAIDSVHRLGYVHRDIKPDNILLDRC-----G---  217
6677759  164   LMSNYD--VPEKWAR---FYTAEVVLALDAIHSMGFIHRDVKPDNMLLDKS----G----  210
6166017  302   MLIKYE-IFSEDITR---FYIAEIVLAIDAVHKLGFIHRDIKPDNILLDRG-----G---  349
1168875  259   FLIANG-SIDEQDCK---PLLKQLLETLLHLHKQGVTHRDIKPENILITND----F----  306

                    190       200       210       220       230       240
                  .    |    . VII   .    |    .    |  VIII  |    .    |
1JNK     202   -----------TLKILDFGLARTAGTS--FMMTPYVVTRYYRAPEVIL-------GMGY  240
query    409   KEILSNLTSGKVTVKLGDLGSARKIGEDVREFTEFYCPIDQIEAAMLK--------NKGA  460
6225243  221   -----------SLKLGDFGLATIVDGP----LYTVCGTPTYVAPEIIA-------ETGY  257
7493389  140   -----------LVKLADFGVSGQLESLR-DKNDDFVGTPFWMAPEVVK-------QTGY  179
2507198  726   -----------HLKLTDFGLSRVGYMK---PSKRFIGTPDYIAPEVILG------NPGI  875
2499575  720   ----------FPQVKLCDFGFARIIGEK---FRRSVVGTPAYLAPEVL-R------NKGY  760
1706450  218   -----------HIRLADFGSCLKLRAD---RSLVAVGTPDYLSPEILQAVGGGPGTGSY  265
6677759  211   -----------HLKLADFGTCMKMNKEGMVRCDTAVGTPDYISPEVLKSQ---GGDGYY  255
6166017  350   -----------HVKLTDFGLSTGFHKL---MAYSTVGTPDYIAPEIF-T------GHGY  435
1168875  307   -----------HLKISDFGLAKVIHGT---FLETFCGTMGYLAPEVLKSKN-VNLDGGY  352

                    250       260       270       280       290       300
                 IX   |    .    |    .    |    .    |    .    |    .    |
1JNK     241   KENVDIWSVGCIMGEMVRHKILFPG----RDYIDQWNKVIEQLGTPCPEFMKKLQPTVRN  296
query    461   LPSMDIFALGATAYKLLFDSYVYPK----EYYEIVERAIEDFQMGRGSYLNYLKMA----  512
6225243  258   GLKVDIWAAGVITYILLCGFPPFRGS--GDDQEVLFDQILMGQVDFPSPYWDN-------  308
7493389  180   NYKADIWSLGITAYELATGEPPYS-------GIHPMKVLLLIPKHSPPSLERSK------  226
2507198  876   K-ASDWWSLGCVVFEFLFGYPPFNA----ETPDQVFQNILARRINWPAEV----------  920
2499575  761   NRSLDMWSVGVIIYVSLSGTFPFN-----EDED-IHDQIQNA-AF-MYPPNPWKE-----  807
1706450  266   GPECDWWALGVFAYEMFYGQTPFYA----DSTAETYGKIVHYKEHLSLPLVDEG------  315
6677759  256   GRECDWWSVGVFLYEMLVGDTPFYAD---SLVGTYSKIMNHKNSLTFPDDND--------  304
6166017  436   SFDCDWWSLGTIMFECLVGWPPFCA----EDSHDTYRKIVNWRHSLYFPDDIT-------  484
1168875  353   DDKVDIWSLGCVLYVMLTASIPFASSSQAKCIELISKGAYPIEPLLENE-----------  401

                    310       320       330       340       350       360
                  .    |    .    |    .    |    .    |    .   XI  .    |
1JNK     297   YVENRPKYAGLTFPKLFPDSLFPADSEHNKLKASQARDLLSKMLVIDPAKRISVDDALQ-  355
query    513   ----------------------------------------------------RQYAVLPKIN  522
6225243  309   ----------------------------VSDSAKE-LINMMLLVNVDQRFSAVQVLE-  336
7493389  227   ----------------------------FSRAFCD-FVSNCLKKNPKDRATAEYLSK-  254
```

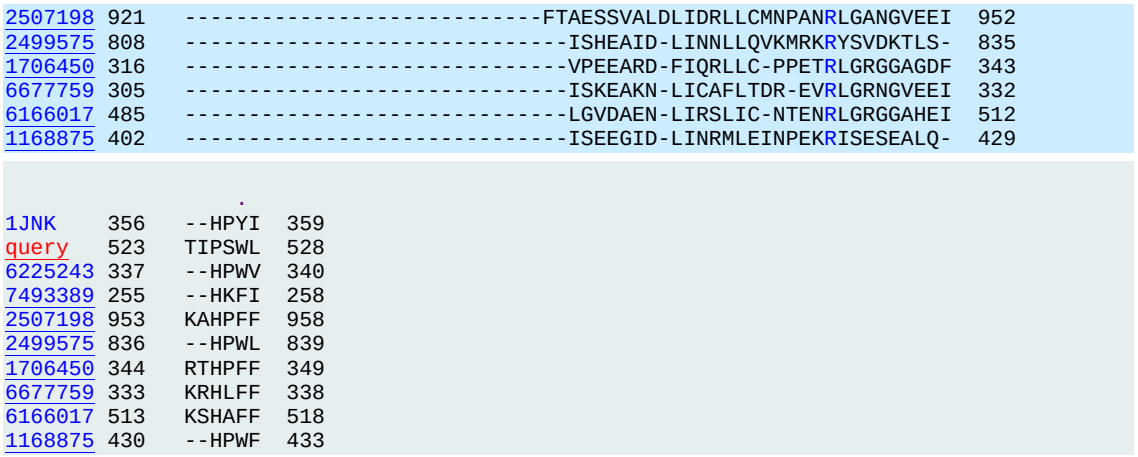


Figure 4-7. Protein Alignment of the Amino Acid Sequence for *Sulfolobus solfataricus ORF sso2291*, Which Encodes an ≈ 63 kDa Hypothetical Protein.

Shown is an alignment of an ≈ 63 kDa hypothetical protein from *S. solfataricus* encoded by *ORF sso2291* (**red**) with other known eukaryotic protein kinases derived using a web-based RPS-BLAST (Reversed Position Specific Blast) program found at the NCBI CD-search database (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Alignments are restricted to those areas encompassing the 12 conserved subdomains characteristic of eukaryotic protein kinases. Highly conserved amino acid residues are shown in **bold**. Dashes indicate gaps introduced during the alignment process. Abbreviations used along with protein accession numbers (underlined) include: 1JNK, a member of the mitogen-activated protein (MAP) kinase family from *Homo sapiens*; DCAMKL1, a Ser/Thr protein kinase from *Rattus norvegicus*; STPK, a Ser/Thr protein kinase from *Schizosaccharomyces pombe*; DCEK1, a Ser/Thr protein kinase from *Schizosaccharomyces pombe*; PCKu, a Mu type protein kinase C from *Homo sapiens*; MDPK, a myotonic dystrophy protein kinase from *Homo sapiens*; ROCK1, a Rho-associated coiled-coil forming Ser/Thr protein kinase of mice; COT1, a Ser/Thr protein kinase from *Neurospora crassa*; and CDS1, a Ser/Thr protein kinase from *Schizosaccharomyces pombe*.

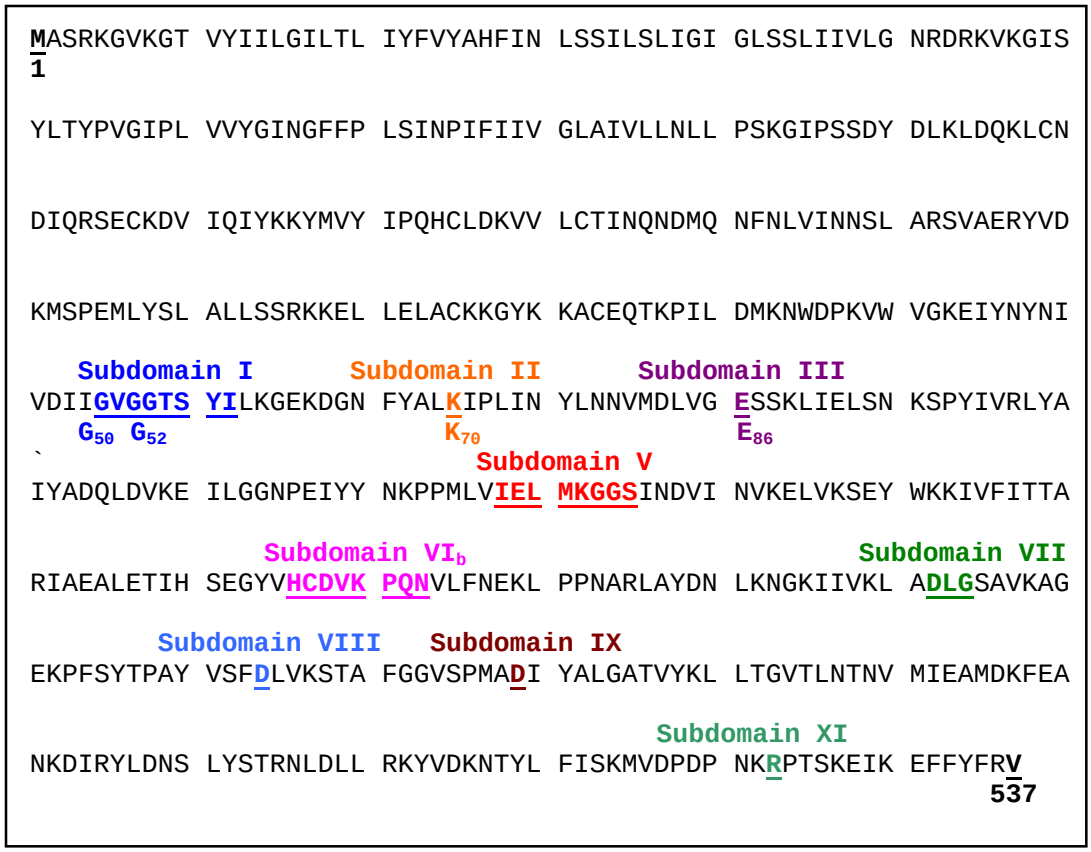
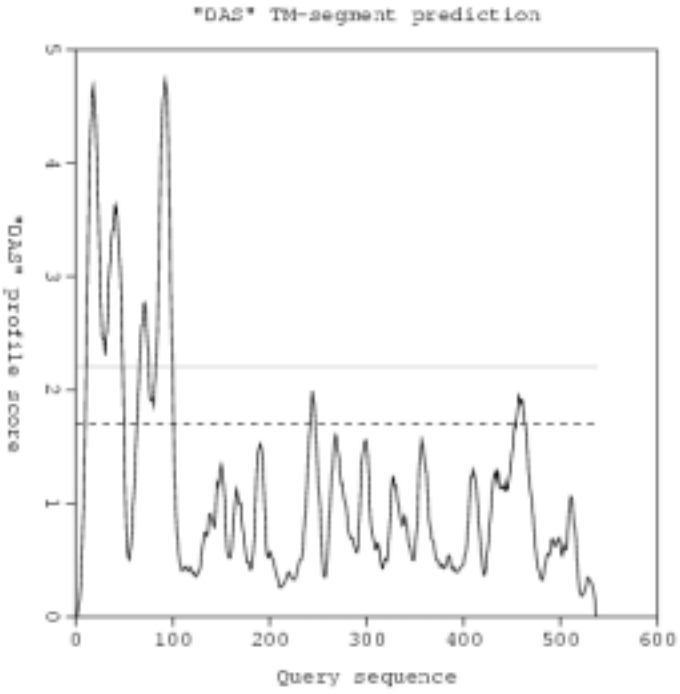


Figure 4-8. The Amino Acid Sequence of the ≈ 61 kDa Hypothetical Protein Encoded by *Sulfolobus solfataricus ORF sso3182* Exhibits Similarity to Eukaryotic Protein Kinases.

Inspection of the partial genome of *S. solfataricus* for homologs of eukaryotic protein kinases, as described in the text, revealed the presence of *ORF sso3182* whose predicted product displays sequence features characteristic of "eukaryotic-like" protein-serine/threonine/tyrosine kinases. *S. solfataricus ORF sso3182* encodes a hypothetical protein consisting of 537 amino acids, with a MW of ≈ 61 kDa, and pI of ≈ 9.5. Shown is its deduced amino acid sequence and prospective candidates for conserved subdomains or essential residues characteristic of eukaryotic protein kinases (described in Introduction).



Start	Stop	Length	Cutoff
10	49	40	1.7
11	48	38	2.2
64	100	37	1.7
66	75	10	2.2
83	99	17	1.7
243	247	5	1.7
455	468	10	1.7

Figure 4-9. Potential Transmembrane Helices of the ≈ 61 kDa Hypothetical Protein Encoded by *Sulfolobus solfataricus ORF sso3182*.

Hypothetical transmembrane domains were identified using the web-based DAS (Dense Alignment Surface Method) Transmembrane Prediction Server (http://www.sbc.su.se/~miklos/DAS/maindas.html) from the Stockholm Bioinformatic Center in Stockholm, Sweden. This algorithm uses low-stringency dot-plots of the query sequence against a collection of non-homologous membrane proteins using a previously derived, special scoring matrix to identify putative transmembrane helices contained within a protein's primary sequence. Shown is the DAS curve for the hypothetical protein encoded by *S.* solfataricus *ORF sso3182* (DAS profile score as a function of the amino acid sequence). A DAS profile score of 1.7 characterizes a loose cutoff and a DAS profile score of 2.2 characterizes a strict cutoff for defining putative transmembrane domains. As is seen in the graph seven putative transmembrane domains are defined using DAS.

```
                                  10        20        30        40        50        60
Subdomain                      .    | I  .    |    .    |    . II |    .    |    .    |
CK2       Zea mays   34   YEVVRKVGRGKYSEVFEGINVN--------NNEKCIIKILK----PVKKKK------IKR  75
sso3182   S.solfa    238  YNIVDIIGVGGTSYILKGE--K--------DGNFYALKIPL----INYLNN------VMD  277
TyrPK2    Dictyodis  108  IQFIQKVGEGAFSEVWEGWWK---------GIHVAIKKLKIIGDEEQFKER-----FIR  152
cGMPPK    Drosophi   457  LEVVSTLGIGGFGRVELVKAHHQ------DRVDIFALKCLK----KRHIVDTKQEEHIFS  506
CaPK      Soybean    34   YEVGRKLGQGQFGTTFECTRR-A-------SGGKFACKSIP----KRKLLCKEDYEDVWR  81
CaMPKIV   Mouse      42   FEVESELGRGATSIVYRCKQK-G-------TQKPYALKVL-----KKTVD-K---KIVRT  84
RSK-1     Mouse      407  YVVKETIGVGSYSVCKRCVHK-A-------TNMEYAVKVID----KSKRD-------PSE  447
PfPK2     Plasmodi   111  YVLNKKIGKGSFSTAYIGTNI-L-------YGNRVVVKEVD----KSKVK-E---SNVYT  154
SPK1      S.cerevi   198  SIIDEVVGQGAFATVKKAIER-T-------TGKTFAVKIIS----KRKVIGNM--DGVTR  243
MEK1/MRE4 S.cerevi   162  EITNRIVGNGTFGHVLITHNSKERDEDVCYHPENYAVKIIK----LKPNK-------FDK  210

                                  70        80        90       100       110       120
                             III  |    .    |    .    |    .    |    .    |    V    |
1A6O      76   ---EIKILQNLC-GGPNIVKLLDIVRDQ-------------HSKTPSLIFEYVNNTDFK  117
query     278  LVGESSKLIELSNKSPYIVRLYAIYADQLDVKEILGGNPEIYYNKPPMLVIELMKGGSIN  337
125874    153  ---EVQNLKKGN--HQNIVMFIGACY-----------------KPACIITEYMAGGSLY  189
400128    507  ---ERHIMLSSR--SPFICRLYRTFR---------------DEKYVYMLLEACMGGEIW  545
116054    82   ---EIQIMHHLS-EHANVVRIEGTYE---------------DSTAVHLVMELCEGGELF  121
266411    85   ---EIGVLLRLS--HPNIIKLKEIFE---------------TPTEISLVLELVTGGELF  123
125690    448  ---EIEILLRYG-QHPNIITLKDVYD---------------DGKHVYLVTELMRGGELL  487
400138    155  ---EIEVLRKVM--HKYIIKLISAYE---------------QEGFVYLVLEYLKGGELF  193
134835    244  ---ELEVLQKLN--HPRIVRLKGFYE---------------DTESYYMVMEFVSGGDLM  282
127299    211  ---EARILLRLD--HPNIIKVYHTFCD--------------RNNHLYIFQDLIPGGDLF  250

                                 130       140       150       160       170       180
                               .    |    .    |    .    |    .    |   VIb   |    .    |
1A6O      118  VLYPTLT---------DYDIRYYIY-ELLKALDYCHSQGIMHRDVKPHNVMIDHE-----  162
query     338  DVINVKELVK---SEYWKKIVFITTARIAEALETIHSEGYVHCDVKPQNVLFNEKLPPNA  394
125874    190  NILHNPNSSTPKVKYSFPLVLKMAT-DMALGLLHLHSITIVHRDLTSQNILLDEL-----  243
400128    546  TMLRDRGS------FEDNAAQFIIG-CVLQAFEYLHARGIIYRDLKPENLMLDER-----  593
116054    122  DRIVQKGH------YSERQAARLIK-TIVEVVEACHSLGVMHRDLKPENFLFDTI-----  169
266411    124  DRIVEKGY------YSERDARDAVK-QILEAVAYLHENGIVHRDLKPENLLYATP-----  171
125690    488  DKILRQKF------FSEREASFVLH-TISKTVEYLHSQGVVHRDLKPSNILYVDE-----  535
400138    194  EYLNNNGP------YTEQVAKKAMK-RVLIALEALHSNGVVHRDLKMENLMLENP-----  241
134835    283  DFVAAHGA------VGEDAGREISR-QILTAIKYIHSMGISHRDLKPDNILIEQD-----  330
127299    251  SYLAKGDCLT---SMSETESLLIVF-QILQALNYLHDQDIVHRDLKLDNILLCTP-----  301

                                 190       200       210       220       230       240
                               .    |    .    |   VII .    |    .    |    . VIII   .    |
1A6O      163  ----------L---RKLRLIDWGLAEFYHP---GKEYNVRVASRYFKGPEL---------  197
query     395  RLAYDNLKNGK---IIVKLADLGSAVKAGE-------KPFSYTPAYVSFDL---------  435
125874    244  --------------GNIKISDFGLSAEKSREGSMTMTNGGICNPRWRPPEL---------  280
400128    594  --------------GYVKIVDFGFAKQIG-T--SSKTWTFCGTPEYVAPEI---------  627
116054    170  ----------DE-DAKLKATDFGLSVFYK-P--GESFCDVVGSPYYVAPEV---------  206
266411    172  -----------APDAPLKIADFGLSKIV--E-HQVLMKTVCGTPGYCAPEI---------  208
125690    536  ----------SGNPECLRICDFGFAKQLR-A-ENGLLMTPCYTANFVAPEV---------  574
400138    242  ----------ND-PSSLKIIDFGLASFLN----SPSMNMRCGSPGYVAPEI---------  277
134835    331  -----------D-PVLVKITDFGLAKVQG-N--GSFMKTFCGTLAYVAPEVIRG------  369
127299    302  -----------EPCTRIVLADFGIAKDLN-S-NKERMHTVVGTPEYCAPEVGFRANRKAY  348

                                 250       260       270       280       290       300
                               .    |    .    |   IX    |    .    |    .    |    .    |
1A6O      198  --------LVDL-----QDYDYSLDMWSLGCMFAGMIFRKEPFFYGHDNHDQLVKIAKVL  244
query     436  --------VKSTA---FGGVSPMADIYALGATVYKLLTGVTLNT-----------NVMIE  473
125874    281  --------TKN-L----GHYSEKVDVYCFSLVVWEILTG---------------------  306
400128    628  --------ILN------KGHDRAVDYWALGILIHELLNG---------------------  652
116054    207  --------LR-------KLYGPESDVWSAGVILYILLS------GVPP-----------F  234
266411    209  --------LRG------CAYGPEVDMWSVGIITYILL---------CG------------  233
125690    575  --------LKR------QGYDEGCDIWSLGILLYTMLAG---YTPFAN------------  605
400138    278  --------LKC------ASYGTKVDIFSLGVILFNIL---------CG------------  302
134835    370  ----KDTSVSPDEYEERNEYSSLVDMWSMGCLVYVILTG---------------------  404
127299    349  QSFSRAATLEQ------RGYDSKCDLWSLGVITHIMLTG---------------------  381

                                 310       320       330       340       350       360
                               .    |    .    |    .    |    .    |    .    |    .    |
1A6O      245  GTDGLNVYLNKYRIELDPQLEALVGRHSRKPWLKFMNAD-NQHLVSPEAIDFLDKLLRYD  303
query     474  AMDKFEANKDIRYLDNS-------------LYSTRNLDLLRKYVDKNTYLFISKMVDPD  519
125874    307  -EIPFSD------LDGSQRSAQVAYAG--------LRPP-IPEYCDPELKLLLTQCWEAD  350
400128    653  -TPPFSAP------DPMQTYNLILKG--------IDMIA-FPKHISRWAVQLIKRLCRDV  696
```

```
116054  235  WAESEP--------GIFRQILLGKLD---------FHSE-PWPSISDSAKDLIRKMLDQN  276
266411  234  -FEPFYD-------ERGDQFMFRRILNC----EYYFISP-WWDEVSLNAKDLVKKLIVLD  280
125690  606  GPSDTP-------EEILTRIGSGKFT---------LSGG-NWNTVSETAKDLVSKMLHVD  648
400138  303  -YPPFRG-------NNVKEIFKKNMR-C----HISFNTK-HWINKSESVKEIILWMCCKN  348
134835  405  -HLPFSG---STQDQLYKQIGRGSYH---------EGPL-KDFRISEEARDFIDSLLQVD  450
127299  382  -ISPFYG-------DGSERSIIQNAKIG----KLNFKLK-QWDIVSDNAKSFVKDLLQTD  428

                    370       380
                 XI   |    .    |
1A6O    304  HQERL-----TALEAMTHPYF  319
query   520  PNKRP-----TSKEIKEFFYF  535
125874  351  PNDRP-----PFTYIVNKL--  364
400128  697  PSERLGYQTGGIQDIKKHKWF  717
116054  277  PKTRL-----TAHEVLRHPWI  292
266411  281  PKKRL-----TTFQALQHPWV  296
125690  649  PHQRL-----TAKQVLQHPWI  664
400138  349  PDDRC-----TALQALGHQWF  364
134835  451  PNNRS-----TAAKALNHPWI  466
127299  429  VVKRL-----NSKQGLKHIWI  444
```

Figure 4-10. Protein Alignment of the Amino Acid Sequence for *Sulfolobus solfataricus* ORF *sso3182*, Which Encodes an ≈ 61 kDa Hypothetical Protein.

Shown is an alignment of an ≈ 61 kDa hypothetical protein from *S. solfataricus* encoded by *ORF sso3182* (**red**) with other known eukaryotic protein kinases derived using a web-based RPS-BLAST (Reversed Position Specific Blast) program found at the NCBI CD-search database (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Alignments are restricted to those areas encompassing the 12 conserved subdomains characteristic of eukaryotic protein kinases. Highly conserved amino acid residues are shown in **bold**. Dashes indicate gaps introduced during the alignment process. Abbreviations used along with protein accession numbers (underlined) include: CK2, a Ser/Thr protein kinase involved in cell proliferation in *Zea mays*, Tyr PK2, a Tyr-protein kinase in *Dictyostelium discoideum*, cGMPPK, a cGMP-dependent protein kinase from *Drosophila melanogaster*; CaPK, a Ca^{2+} dependent protein kinase SK5 from *Glycine max*, CaMPKIV, the catalytic chain of a calcium/calmodulin dependent protein kinase type IV from *Mus musculus*; RSK-1, ribosomal protein S6 alpha kinase 1 from *Mus musculus*; PfPK2, a probable Ser/Thr protein kinase from *Plasmodium falciparum*; SPK1, a Ser protein kinase from *Saccharomyces cerevisiae*; MEK1/MRE4, a Ser/Thr protein kinase from *Saccharomyces cerevisiae*.

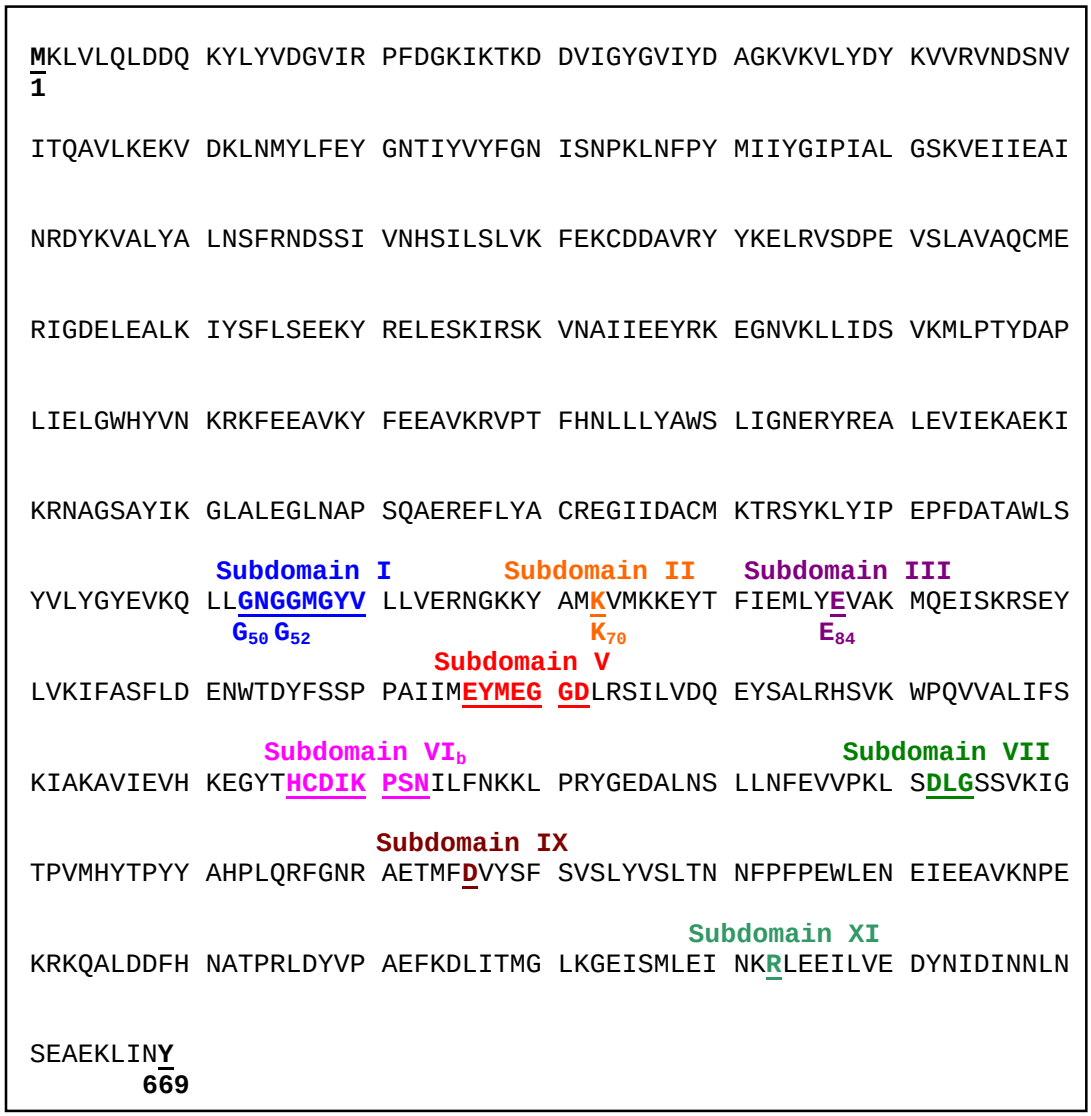
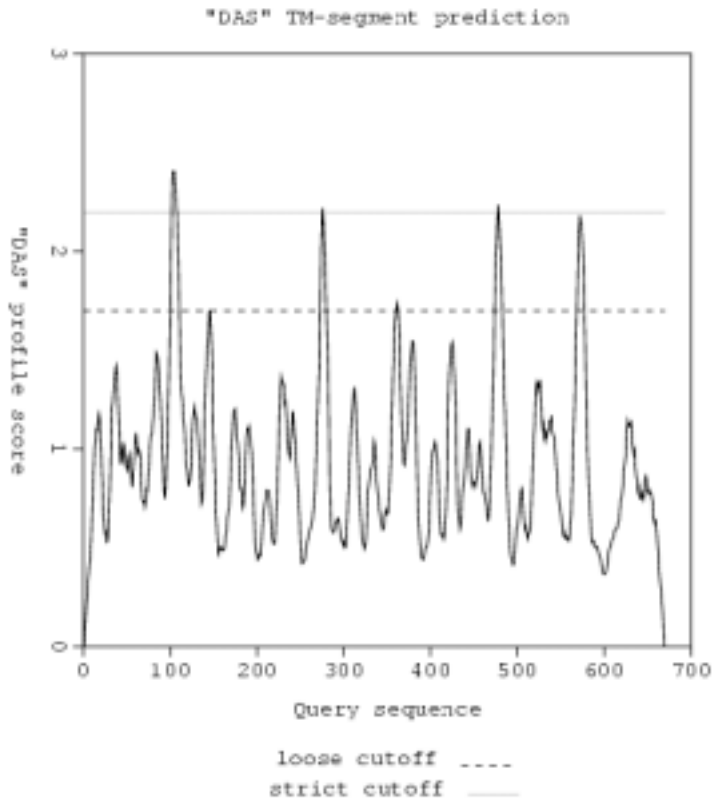


Figure 4-11. The Amino Acid Sequence of the ≈ 77 kDa Hypothetical Protein Encoded by *Sulfolobus solfataricus ORF sso3207* Exhibits Similarity to Eukaryotic Protein Kinases.

Inspection of the partial genome of *S. solfataricus* for homologs of eukaryotic protein kinases, as described in the text, revealed the presence of *ORF sso3207* whose predicted product displays sequence features characteristic of eukaryotic protein-serine/threonine/tyrosine kinases. *S. solfataricus ORF sso3207* encodes a hypothetical protein consisting of 669 amino acids, with a MW of ≈ 77 kDa, and pI of ≈ 5.6. Shown is its deduced amino acid sequence and perspective candidates for conserved subdomains or essential residues characteristic of eukaryotic protein kinases (described in Introduction).



Start	Stop	Length	Cutoff	Start	Stop	Length	Cutoff
101	110	10	1.7	361	362	2	1.7
102	107	6	2.2	474	482	9	1.7
273	280	8	1.7	478	478	1	2.2
276	276	1	2.2	569	577	9	1.7

Figure 4-12. Potential Transmembrane Helices of the ≈ 77 kDa Hypothetical Protein Encoded by *Sulfolobus solfataricus ORF sso3207*.

Hypothetical transmembrane domains were identified using the web-based DAS (Dense Alignment Surface Method) Transmembrane Prediction Server (http://www.sbc.su.se/~miklos/DAS/maindas.html) from the Stockholm Bioinformatic Center in Stockholm, Sweden. This algorithm uses low-stringency dot-plots of the query sequence against a collection of non-homologous membrane proteins using a previously derived, special scoring matrix to identify putative transmembrane helices contained within a protein's primary sequence. Shown is the DAS curve for the hypothetical protein encoded by *S. solfataricus ORF sso3207* (DAS profile score as a function of the amino acid sequence). A DAS profile score of 1.7 characterizes a loose cutoff and a DAS profile score of 2.2 characterizes a strict cutoff for defining putative transmembrane domains. As is seen in the graph eight putative transmembrane domains are defined using DAS.

```
                                10        20        30        40        50        60
                              .    | I  .    |    .    II   .    |    . III    .    |
1JNK H.sapein 64    YQNLKPIGSGAQGIVCAAYDAVLDRNVAIKKLSRPFQ-NQTHAKRAYRELVLMKCV--NH  120
sso3207 S.sof 366   YEVKQLLGNGGMGYVLLVE--RNGKKYAMKVMKKEYT-FIEMLY----EVAKMQEISKRS  418
CDC7 S.pombe  9     ITLGDCLGKGAFGAVYRGLNIKNGETVAVKKVKLSKM--KSDLSVIKMEIDLLKNL--DH  65
Pkh1p  S.cerv 125   FKFGEQLGDGSYSSVVLATARDSGKKYAVKVLSKEY-LRQKKVKYVTVEKLALQKLN-GT  183
MDPK Human    81    FEILKVIGRGAFSEVAVVKMKQTGQVYAMKIMNKWDM----EVSCFREERDVLVNG--DR  138
PCaPKbParamec 56    VQVVRYVKLGAFGEVRLVIHKSSGYKRAMKQIRKDKI-IKEDEENMFSEVNTLKEL--DH  112
CDPK Soybean  34    YEVGRKLGQGQFGTTFECTRRASGGKFACKSIPKRKL----DYEDVWREIQIMHHLS-EH  92
ROCK Mouse    76    YEVVKVIGRGAFGEVQLVRHKSTRKVYAMKLLSKFEMIKRSDSAFFWEERDIMAFA--NS  133
Cot1 N.crassa 214   YQTIKIIGKGAFGEVKLVQKKADGKVYAMKSLIKTEM----QLAHVRAERDILAES--DS  271
MKK1 S.cerevi 214   ITTLGILGEGAGGSVAKCRLKNGKKVFALKTINTMNT--PEYQKQIFRELQFNKSF--KS  270

                          70        80        90       100       110       120
                        .    |    .    |    . V|    .    |    .    |    .    |
5542282 121   KNIISLLNVFTPQKTLEEFQDVY-LVMELMDA----NLCQVIQME---LD--------HE  164
query   419   EYLVKIFASFLDENWTDYFSSPPAIIMEYMEGG---DLRSILVDQE--YSALRHSVKWPQ  473
1168817 66    PNIVKYRGSYQTN------DSLC-IILEYCENG---SLRSICKNFG-KIP--------EN  106
6320698 184   KGIFKLFFTFQDE------ASLY-FLLEYAPHG---DFLGLIKKYG-SLN--------ET  224
1706450 139   RWITQLHFAFQDE------NYLY-LVMEYYVGG---DLLTLLSKFGERIP--------AE  180
2271461 113   PNIVKLHELFQDA------KNYY-LVTEYLEGG---ELFQKITDMK-HFS--------EK  153
116054  93    ANVVRIEGTYEDS------TAVH-LVMELCEGG---ELFDRIVQKG-HYS--------ER  133
6677759 134   PWVVQLFYAFQDD------RYLY-MVMEYMPGG---DLVNLMSNYD--VP--------EK  173
6166017 272   PWVVKLYTTFQDA------NFLY-MLMEFLPGG---DLMTMLIKYE-IFS--------ED  312
1709042 271   DYIVQYYGMFTDE-----SSSIY-IAMEYMGGKSLEATYKNLLKRGGRIS--------ER  317

                        130       140       150       160       170       180
                      .    |    .    |    VIb  |    .    |    .    |    .    |
5542282 165   RMSYLLYQMLCGIKHLHSAGIIHRDLKPSNIVVKSD---CT---------------LKI  205
query   474   VVALIFSKIAKAVIEVHKEGYTHCDIKPSNILFNKK---LPRYGEDALNSLLNFEVVPKL  530
1168817 107   LVALYTFQVLQGLLYLHNQGVIHRDIKGANILTTKD---GT---------------IKL  147
6320698 225   CARYYASQIIDAVDSLHNIGIIHRDIKPENILLDKN---MK---------------VKL  265
1706450 181   MARFYLAEIVMAIDSVHRLGYVHRDIKPDNILLDRC----H---------------IRL  221
2271461 154   MAADIMKQILAGVVHCHEKKVVHRDLKPENILFENKKPNSN---------------LKI  197
116054  134   QAARLIKTIVEVVEACHSLGVMHRDLKPENFLFDTI----K---------------LKA  177
6677759 174   WARFYTAEVVLALDAIHSMGFIHRDVKPDNMLLDKS---GH---------------LKL  214
6166017 313   ITRFYIAEIVLAIDAVHKLGFIHRDIKPDNILLDRG----H---------------VKL  353
1709042 318   VIGKIAESVLRGLSYLHERKVIHRDIKPQNILLNEK---GE---------------IKL  358

                        190       200       210       220       230       240
                    VII     |    .    |    .    |    .  VIII .    |    .    |
5542282 206   LDFGLARTAGTS-------------FMMTPYVVTRYYRAPEVIL-------GMGYKE-N  243
query   531   SDLGSSVKIGTP-------------VMHY----TPYYAHP--LQRFGNRAETM------  553
1168817 148   ADFGVATKINAL-------------EDHSVVGSPYWMAPEVIE-------LVGATT-A  184
6320698 266   TDFGTAKILPEEPSNTADGKPYFDLYAKSKSFVGTAEYVSPELLN--------DNYTDSR  317
1706450 222   ADFGSCLKLRAD-------------RSLVAVGTPDYLSPEILQAVGGGPGTGSYGP-E  268
2271461 198   IDFGTSRKMETN-------------QNLTKRLGTPYYIAPEVLK--------RNYNE-K  234
116054  178   TDFGLSVFYKPG-------------SFCDVVGSPYYVAPEVL-R------K-LYGP-E  214
6677759 215   ADFGTCMKMNKEG-----------MVRCDTAVGTPDYISPEVLKSQ---GGDGYYGR-E  258
6166017 354   TDFGLSTGFHKL-------------MAYSTVGTPDYIAPEIF-T------GHGYSF-D  438
1709042 359   CDFGVSGEAVNS-------------LAMTFTGTSFYMAPERIQ-------GQPYSV-T  395

                        250       260       270       280       290       300
                    IX  .    |    .    |    .    |    .    |    .    |    .    |
5542282 244   VDIWSVGCIMGEMVRHKILFPG-RDYIDQWNKVIEQLGTPCPEFMKKLQPTVRNYVENRP  302
query   554   FDVYSFSVSLYVSLTNNFPFPEWLENEIEEAVKNPEKRKQALDDFHNATPRLDYVPAEFK  608
1168817 185   SDIWSVGCTVIELLDGNPPYYDLDPTSALFRMVKDEHPPLPSN----------------  227
6320698 318   CDIWAFGCILYQMLAGKPPFKAANEYLTFQKVMKIQYAFTAG----------------  359
1706450 269   CDWWALGVFAYEMFYGQTPFYA-DSTAETYGKIVHYKEHLSLPLVDEG-----------  315
2271461 235   CDVWSCGVILYIMLCGYPPFGG--QDQQILQNIELGKYEFDPEDWNK-----------  279
116054  215   SDVWSAGVILYILLSGVPPFWA-ESEPGIFRQILLGKLDFHSEPWPS-----------  260
6677759 259   CDWWSVGVFLYEMLVGDTPFYADSLVGTYSKIMNHKNSLTFPDDND-----------  304
6166017 439   CDWWSLGTIMFECLVGWPPFCA-EDSHDTYRKIVNWRHSLYFPDDIT-----------  484
1709042 396   CDVWSLGLTLLEVAGGRFPFESDKITQNVAPIELLTMILTFSPQLKDEPELDIS------  449

                        310       320       330       340       350       360
                      .    |    .    |    .    |    .    |   XI    |    .    |
5542282 303   KYAGLTFPKLFPDSLFPADSEHNKLKASQARDLLSKMLVIDPAKRISVDDALQ---HPYI  359
query   609   DLLITMGLKGEISML------------------------EINKRLEEI-LVE---DYNI  665
1168817 228   ----------------------ISSAAKS-FLMQCFQKDPNLRIKTRKLLK---HPWV  259
6320698 360   ----------------------FPQIVKD-LVKKLLVRDPNDRLTIKQIKA---HLFF  391
```

```
1706450 316   ----------------------VPEEARD-FIQRLLC-PPETRLGRGGAGDFRTHPFF  349
2271461 280   ----------------------ISEDAKN-LIKRMLTKDYQLRISAQEAYN---DPWI  311
116054  261   ----------------------ISDSAKD-LIRKMLDQNPKTRLTAHEVLR---HPWI  292
6677759 305   ----------------------ISKEAKN-LICAFLTDR-EVRLGRNGVEEIKRHLFF  338
6166017 485   ----------------------LGVDAEN-LIRSLIC-NTENRLGRGGAHEIKSHAFF  518
1709042 450   ----------------------WSKTFRS-FIDYCLKKDARERPSPRQMLK---HPWI  481
```

Figure 4-13. Protein Alignment of the Amino Acid Sequence for *Sulfolobus solfataricus ORF sso3207*, Which Encodes an ≈ 77 kDa Hypothetical Protein.

Shown is an alignment of an ≈ 77 kDa hypothetical protein from *S. solfataricus* encoded by ORF sso3207 (**red**) with other known eukaryotic protein kinases derived using a web-based RPS-BLAST (Reversed Position Specific Blast) program found at the NCBI CD-search database (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Alignments are restricted to those areas encompassing the 12 conserved subdomains characteristic of eukaryotic protein kinases. Highly conserved amino acid residues are shown in **bold**. Dashes indicate gaps introduced during the alignment process. Abbreviations used along with protein accession numbers (underlined) include: 1JNK, a member of the mitogen-activated protein (MAP) kinase family from *Homo sapiens*; CDC7, a Ser/Thr protein kinase from *Schizosaccharomyces pombe* that is involved in cell division; Pkh1p a Ser/Thr protein kinase from *Saccharomyces cerevisiae*; MDPK, a myotonic dystrophy protein kinase from *Homo sapiens*; PCaPKb, a Ca^{2+} dependent protein kinase from *Paramecium*; CDPK, a Ca^{2+} dependent protein kinase from soybean; ROCK, a Ser/Thr protein kinase implicated in the Rho-mediated signalling pathways of mice; Cot 1, a protein kinase whose sequence resembles that of the cAMP-dependent protein kinase family found to be essential for hyphal elongation in *Neurospora crassa*; and MKK1 a Ser/Thr protein kinase found in *Saccharomyces cerevisiae*.

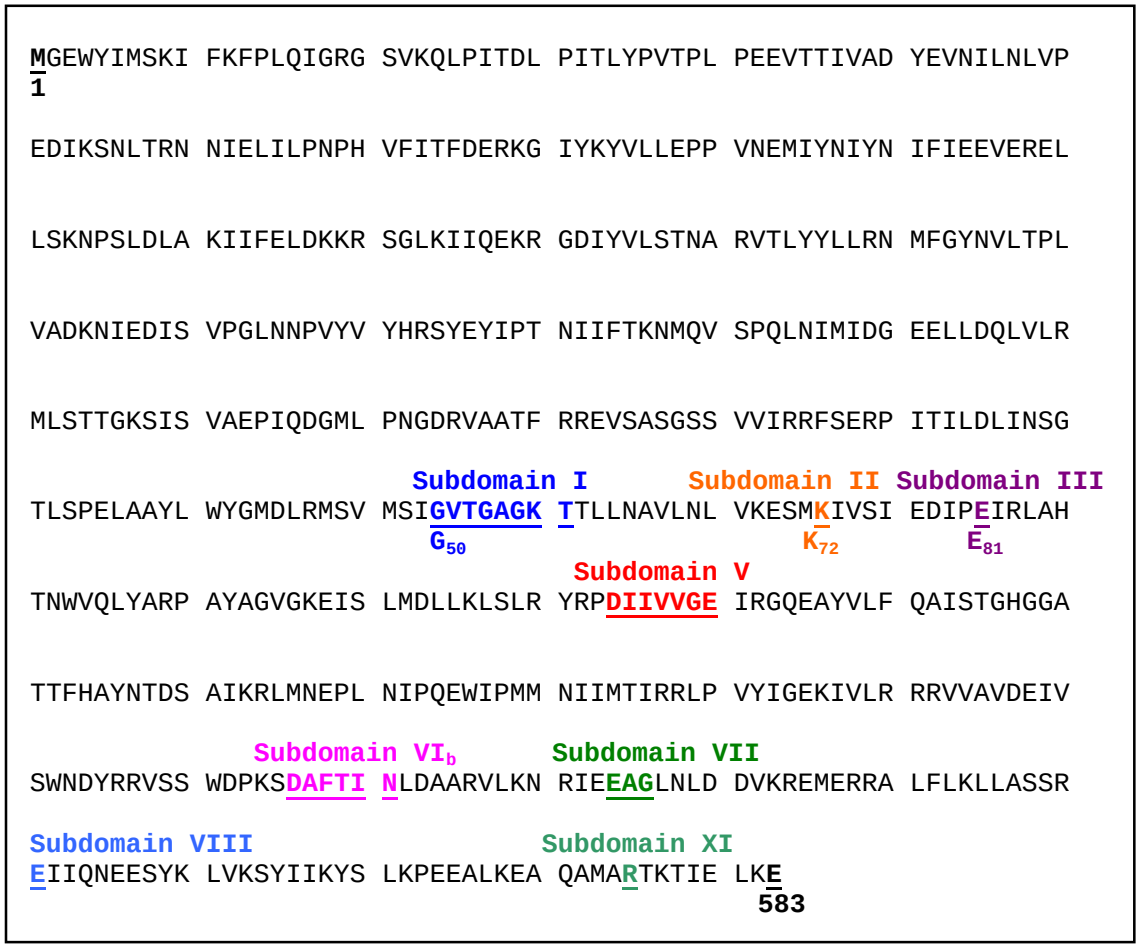


Figure 4-14. The Amino Acid Sequence of the ≈ 67 kDa Hypothetical Protein Encoded by *Sulfolobus solfataricus ORF sso2387* Exhibits Similarity to Eukaryotic Protein Kinases.

Inspection of the partial genome of *S. solfataricus* for homologs of eukaryotic protein kinases, as described in the text, revealed the presence of *ORF sso2387* whose predicted product displays sequence features characteristic of eukaryotic protein-serine/threonine/tyrosine kinases. *S. solfataricus ORF sso2387* encodes a hypothetical protein consisting of 583 amino acids, with a MW of ≈ 66.5 kDa, and pI of ≈ 6.7. Shown is its deduced amino acid sequence and perspective candidates for the conserved subdomains or essential residues characteristic of eukaryotic protein kinases (described in Introduction).

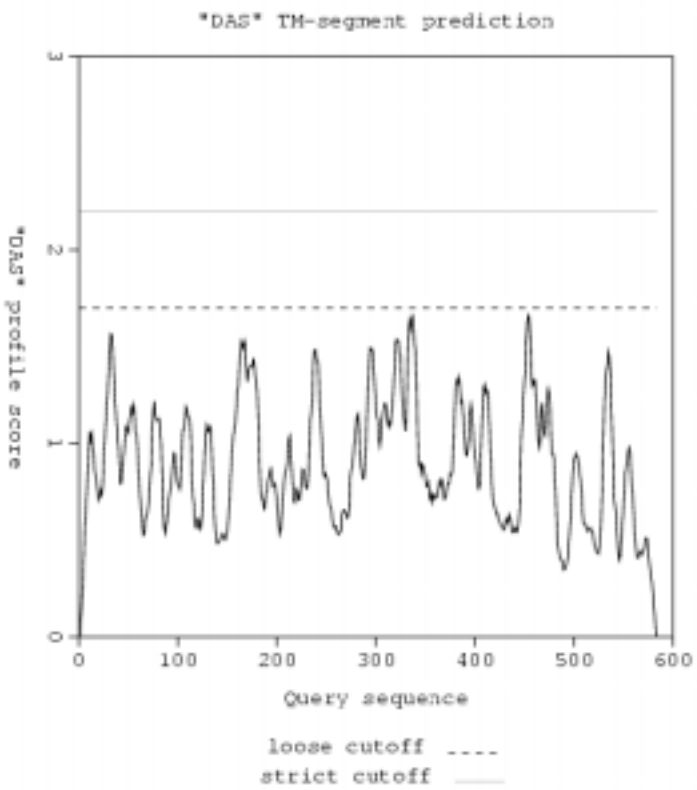


Figure 4-15. Potential Transmembrane Helices of the ≈ 67 kDa Hypothetical Protein Encoded by *Sulfolobus solfataricus ORF sso2387*.

Hypothetical transmembrane domains were identified using the web-based DAS (Dense Alignment Surface Method) Transmembrane Prediction Server (http://www.sbc.su.se/~miklos/DAS/maindas.html) from the Stockholm Bioinformatic Center in Stockholm, Sweden. This algorithm uses low-stringency dot-plots of the query sequence against a collection of non-homologous membrane proteins using a previously derived, special scoring matrix to identify putative transmembrane helices contained within a protein's primary sequence. Shown is the DAS curve for the hypothetical protein encoded by *S*. solfataricus *ORF sso2387* (DAS profile score as a function of the amino acid sequence). A DAS profile score of 1.7 characterizes a loose cutoff and a DAS profile score of 2.2 characterizes a strict cutoff for defining putative transmembrane domains. As is seen in the graph no obvious transmembrane domains are defined using DAS.

Identification of Candidate ORFs Encoding the Membrane-Associated Protein Kinase From Sulfolobus solfataricus

Of these six hypothetical proteins, three (encoded by *ORFs: sso2291*, *sso2387*, and *sso3182*) exhibit noticeable similarities to the previously characterized membrane-associated protein kinase described from *S. solfataricus* (Chapter III). These similarities include molecular weight (all three hypothetical proteins are between ≈ 60 – 70 kDa), consensus sequence motifs for potential N-glycosylation sites (Asn-X-Ser/Thr), as well numerous potential threonine phosphorylation sites as identified using the ExPASy ScanProsite Program (http://www.expasy.ch/). Furthermore, the hypothetical proteins encoded by *ORFs sso2291* and *sso3182* contain potential transmembrane domains as identified using DAS (http://www.sbc.su.se/~miklos/DAS/maindas.html). While DAS analysis failed to identify any potential transmembrane helices in the hypothetical protein encoded by *ORF sso2387*, the ExPASy ScanProsite Program identified four consensus sequence motifs of potential myristylation sites (amino acid residues 258-263, 324-329, 327-332, 418-423), a modification that could potentially act to anchor the enzyme to the membrane.

Of the six identified *ORFs* from *S. solfataricus* encoding potential "eukaryotic-like" protein kinases, *ORF sso2387* was chosen to be cloned for two main reasons: i) this *ORF* was discovered well in advance of the other *ORFs* from *S. solfataricus* and ii) this *ORF* encodes a hypothetical protein whose primary sequence shares characteristics with the membrane-associated protein kinase from *S. solfataricus* described in Chapter III.

CHAPTER V

Cloning and Expression of a Gene From *Sulfolobus solfataricus* (*sso2387*), Purification and Characterization of its Encoded Protein Kinase (SsPK2)

Analysis of Sulfolobus solfataricus ORF sso2387and its Protein Product SsPK2

In order to determine if *ORF sso2387* encoded a functional protein kinase, possibly the *S. solfataricus* protein kinase described in Chapter III, the gene was cloned into an expression vector and expressed in *E. coli* as described in this Chapter. From here on, I will refer to the hypothetical protein encoded by *ORF sso2387* as SsPK2. The nucleotide and DNA derived amino acid sequences of *ORF sso2387* are shown in Figure 5-1. Analysis of nucleotide composition of *ORF sso2387*, as well as amino acid composition and molecular weight of SsPK2 are shown in Table 5-1. As seen in Figure 5-1, translation of SsPK2 appears to start at the nucleotide sequence ATG, and at a position –36 of the sequence there is a potential promoter (cttctata), which matches the consensus promoter sequence known for archaeal genes (Bell et. al., 1999). Identification of potential "eukaryotic" protein kinase subdomains contained within the primary sequence of SsPK2 are displayed in Figure 4-14.

Cloning ORF sso2387 into TOPO T7/NT Vector and Expression of SsPK2 in E. coli

The gene for SsPK2 was amplified in a PCR reaction in which *S. solfataricus* genomic DNA was used as template and two oligonucleotide primers corresponding to the region encoding the predicted amino terminal sequence (Ss67F01), and a sequence modeled after the region encoding the predicted carboxyl terminus (Ss67R01) were used as primers (Figure 5-1). Both primers were designed using WWW Primer Picker Primer 3, a web-based primer design program developed at the Whitehead Institute, Cambridge, Massachusetts. Amplification of genomic DNA by PCR using these two oligonucleotides as primers (described in Methods) yielded a ≈ 1.8 kbp piece of DNA (Figure 5-2). The PCR product was then ligated into linearized pCR T7/NT TOPO vector (Invitrogen, Figure 2-1) and transformed into TOP 10 F' cells (Invitrogen) as described in Methods. Transcription of *sso2387* in pCR T7/NT TOPO vector is

tagttattaccgcagtgctggtaatcatttcattacatatatcactaatac**cttctata**t
ctcctcctcgacatctttataacctag

Ss67F01 →

atgggggagtggtatataatga

atggggggagtggtatataatgagtaaaattttcaagttccccttacagatcggtagaggt
 M$_1$ G E W Y I M S K I F K F P L Q I G R G

agcgttaagcaattacctatcacagatcttcctataacgctttacccagttacacctttg
 S V K Q L P I T D L P I T L Y P V T P L

cccgaagaggttacaacgattgtcgcggattatgaggttaatatcctaaatttagtcccg
 P E E V T T I V A D Y E V N I L N L V P

gaggatattaagtcaaacctaactcgaaataatattgaactgatattaccaaatcctcat
 E D I K S N L T R N N I E L I L P N P H

gtcttcattacttttgatgagagaaaaggtatctacaaatatgttttattagaaccaccg
 V F I T F D E R K G I Y K Y V L L E P P

gttaatgaaatgatctataatatctacaatatatttatagaggaagtggagagagaactg
 V N E M I Y N I Y N I F I E E V E R E L

ctttctaagaatccctctttagatcttgcaaaaattatattcgaactggataagaaaagg
 L S K N P S L D L A K I I F E L D K K R

tcaggtcttaaaattatccaagagaagagaggagatatctacgttttgagtacaaatgct
 S G L K I I Q E K R G D I Y V L S T N A

agagttactttgtactatttattaagaaacatgttcggatacaacgtattaaccccactt
 R V T L Y Y L L R N M F G Y N V L T P L

gtagctgataaaaatatagaagatatttcggttcctggtctaaataatccagtctatgta
 V A D K N I E D I S V P G L N N P V Y V

tatcatagaagttatgaatatattccaactaatattatatttactaagaacatgcaagta
 Y H R S Y E Y I P T N I I F T K N M Q V

tctccacaacttaatataatgatagatggtgaggaactgctagatcaattggttctaaga
 S P Q L N I M I D G E E L L D Q L V L R

atgctttctactacaggtaagtcaatttctgttgctgaaccaatacaagacggtatgtta
 M L S T T G K S I S V A E P I Q D G M L

ccaaatggtgatagggttgccgcaacatttaggcgcgaggtatcagccagtggttcttca
 P N G D R V A A T F R R E V S A S G S S

gtagtaataagaagatttagcgaaaggcctatcacaatactagatttaattaattctggt
 V V I R R F S E R P I T I L D L I N S G

accctatctccagaactagcagcatatctatggtatggaatggatctgagaatgagtgtc
 T L S P E L A A Y L W Y G M D L R M S V

atgtcaataggagttaccggggccggaaagaccactttacttaatgcagttctaaatcta
 M S I G V T G A G K T T L L N A V L N L

gtaaaagaaagcatgaagatcgtctccatagaagatattccagaaattagattagcccat
 V K E S M K I V S I E D I P E I R L A H

actaattgggttcagctatacgctaggccagcatatgcaggagtaggtaaagagatttca
 T N W V Q L Y A R P A Y A G V G K E I S

ttaatggatctgctaaaattatccctcagatacaggccagatataatagttgtaggtgag
 L M D L L K L S L R Y R P D I I V V G E

ataagagggcaagaggcttacgtattattccaagcgatatcaactggacatggaggtgct
 I R G Q E A Y V L F Q A I S T G H G G A

acgacattccacgcgtataataccgactctgcaataaagaggctcatgaatgagccccta
 T T F H A Y N T D S A I K R L M N E P L

aatattccacaagaatggatacctatgatgaacataataatgacaattaggaggttacca
 N I P Q E W I P M M N I I M T I R R L P

gtatatataggagaaaagatagtcctaagaagacgtgttgtagcagttgatgaaatagtt
 V Y I G E K I V L R R R V V A V D E I V

agttggaacgactatagaagggtctcgagctgggatccaaaaagtgatgcgtttacaatt
 S W N D Y R R V S S W D P K S D A F T I

aatctagatgctgccagagtgttaaaaaatagaatagaggaagctggtcttaatctagat
 N L D A A R V L K N R I E E A G L N L D

gacgtgaaaagagaaatggagagaagagcattattcctaaagttgttagcgtcttccaga
 D V K R E M E R R A L F L K L L A S S R

gagataatacaaaatgaggagagttataagcttgtgaagagctatataataaaatacagc
 E I I Q N E E S Y K L V K S Y I I K Y S

ttaaaacccgaagaagctctaaaagaggctcaagcaatggctaggacaaaaactatagag
 L K P E E A L K E A Q A M A R T K T I E

ttaaaagaataagtgaaaacaaaattattaataaagatgaaattgtgttattgaaaagac
L K E_{583}stop

tattaatatatgccattatcgcaagaataatccctaatatttttatactcaagatttctt
       atacggtaatagcgttcttatt
            ← Ss67R01

Figure 5-1. DNA Derived Amino Acid Sequence of SsPK2

Shown is the DNA sequence for *S. solfataricus ORF sso2387* and the protein it encodes, SsPK2. The predicted **atg** start codon is underlined and red in color; the predicted **taa** stop codon is underlined and orange in color; and the potential promoter, **cttctata**, is underlined and purple in color. The first Met residue encoded by *sso2387* is red in color and labeled M_1, while the last amino acid encoded by *sso2387*, E_{583}, is blue in color. Also shown are the oligonucleotide primers used to amplify *ORF sso2387* using PCR, as well as the region of the DNA sequence where they anneal. The forward primer (**Ss67FRD**) is shown in red and the reverse PCR primer (**Ss67REV**) is shown in blue.

Predicted Molecular Weight

66,524 Da

Nucleotide Base Composition

Adenine	=	625 (35.7%)	Thymine	=	488 (27.9%)
Guanine	=	362 (20.7%)	Cytosine	=	277 (15.8%)

Adenine + Thymine	=	1113 (63.5%)
Guanine + Cytosine	=	639 (36.5%)

Amino Acid Composition

Ala	32 (5.5%)	Gln	13 (2.2%)	Leu	66 (11.3%)	Ser	37 (6.3%)
Arg	37 (6.3%)	Glu	45 (7.7%)	Lys	33 (5.7%)	Thr	29 (5.0%)
Asn	33 (5.7%)	Gly	27 (4.6%)	Met	19 (3.3%)	Trp	6 (1.0%)
Asp	26 (4.5%)	His	5 (0.9%)	Phe	14 (2.4%)	Tyr	27 (4.6%)
Cys	0 (0.0%)	Ile	60 (10.3%)	Pro	30 (5.1%)	Val	44 (7.5%)

Table 5-1. Nucleotide Base Composition of ORF sso2387 and Predicted Amino Acid Composition and Molecular Weight of SsPK2.

Shown is the predicted molecular weight of SsPK2 (top), nucleotide base composition of sso2387 (middle), and amino acid composition of SsPK2 (bottom). For the nucleotide composition the total number of each base is indicated and the percent total of ORF sso2387 given in parentheses. For the amino acid composition the total number of each amino acid is indicated and the percent total of SsPK2 given in parentheses.

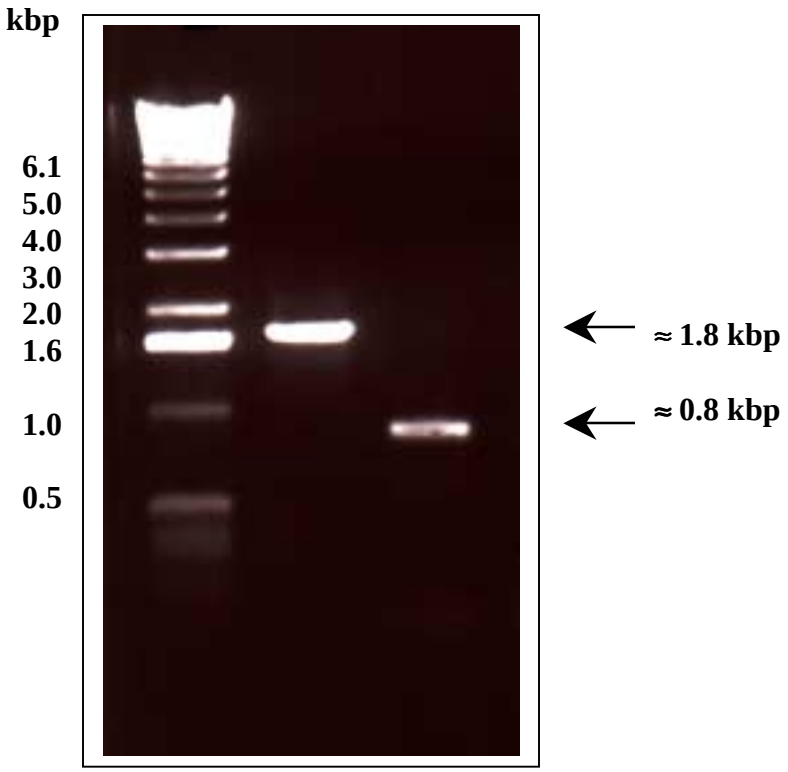


Figure 5-2. Amplification of *ORF sso2387* and the Amino Terminal Deletion Mutant of *OFR sso2387* from *Sulfolobus solfataricus* Using the Polymerase Chain Reaction

The Polymerase Chain Reaction was performed to amplify *OFR sso2387* using 6 pmol each of the forward primer (Ss67F01) and the reverse primer (Ss67R01) and 550 ng of genomic DNA (lane 1) or to amplify the amino terminal deletion mutant of *OFR sso2387* using 6 pmol of each of the forward primer (Ss67F02) and the reverse primer (Ss67RO1) and 20 ng of plasmid DNA (lane 2) as described in Methods. The DNA produced in each of the reactions was resolved on a 0.8% (w/v) agarose gel containing 0.5 μg / ml EtBr in TAE buffer. Shown is a picture of the gel illuminated with UV light. Lane 1 contains 1.0 μg of 1 kb DNA ladder with the approximate size (kbp) of the standards indicated to the left of the gel. Lane 2 contains the DNA amplified during the PCR reaction used to amplify *ORF sso2387* and its approximate size (kbp) indicated to the right. Lane 3 contains the DNA amplified during the PCR reaction used to amplify the amino terminal deletion mutant of *ORF sso2387* and its approximate size (kbp) indicated to the right Note that only one DNA band is observed in the PCR reactions corresponding to the expected size of *ORF sso2387* (≈ 1.8 kbp) in lane 2 and the amino terminal deletion mutant of *OFR sso2387* (≈ 0.8 kbp) in lane 3.

controlled by the T7 promoter. The pCR T7/NT TOPO vector also adds a sequence of DNA encoding six successive histidine residues 5' to the cloned *sso2387* gene. The resulting N-terminal leader sequence of six histidine residues should facilitate the purification of the recombinant protein by metal affinity chromatography.

Recombinant plasmids were purified from TOP 10 F' cells using Wizard Minipreps (Promega) and sequenced using both forward (T7 primer) and reverse primers (pRSET primer), as described in Methods, to verify that the PCR amplified genomic DNA was indeed OFR sso2387 and that it had ligated into the vector correctly. Next, the plasmids were transformed into *E. coli* strain BL21(DE3)pLysS as described in Methods. This strain of cells has a chromosomal copy of the T7 RNA Polymerase gene under the control of the *lacUV5* promoter, which is induced when isopropyl β-D-thiogalactoside (IPTG) is added to the growth medium. Contained within this strain of *E. coli* is a plasmid, pLysS, which encodes T7 lysozyme and confers chloramphenicol resistance. The T7 lysozyme inhibits T7 RNA Polymerase and thereby reduces the basal level expression of the recombinant protein prior to IPTG induction resulting from any leaky transcription of the T7 polymerase gene. *E. coli* BL21(DE3)pLysS cells transformed with the pCR T7/NT TOPO plasmid containing *S. solfataricus ORF sso2387* were designated BL21SsPK2.

Overnight cultures of BL21SsPK2 grown in 4 ml of LB broth containing 100 μg/ml ampicillin and 34 μg/ml chloramphenicol were used to inoculate 200 ml of LB broth containing 100 μg/ml ampicillin, 34 μg/ml chloramphenicol in a 1 L flask. The culture was grown at 37°C with vigorous shaking (225rpm) until it reached an $O.D._{600nm}$ of ≈ 0.5 - 0.8. Isopropyl β-D-thiogalactoside (1.0 M) then was added to a final concentration of 1 mM to induce production of SsPK2, and the culture was incubated for an additional 4 hours at 37°C with vigorous shaking (225 rpm). The cells were harvested by centrifugation at 4°C, 1,000 x g, for 15 minutes and then stored at –20°C. Cells were thawed and resuspended in 4 ml of 50 mM MOPS, pH 7.0, containing 1 mM PMSF, 100 μg/ml lysozyme, 5μg/ml RNaseA, and 2 U/ml DNaseI and incubated on ice for 10 minutes. Next, the cells were lysed on ice by sonication with 3 x 30 second bursts at 30% intensity. Lysed cells were centrifuged at 3,000 x g, 4°C for 15 minutes to remove insoluble debris and the supernatant liquid, soluble cell extract, was stored at 4°C. The

cell pellet was resuspended in 50 mM MOPS, pH 7.0, containing 5 M Urea to solubilize any protein aggregates, e. g. inclusion bodies. Insoluble debris was removed by centrifugation at 3,000 x g, 4°C, for 15 minutes and the supernatant liquid, which consisted of solubilized inclusion bodies, was stored at 4°C. This fraction was designated the urea soluble fraction.

In order to determine if SsPK2 was soluble, the proteins contained within the soluble cell extract and urea soluble fractions were resolved using SDS-PAGE, then recombinant fusion proteins visualized by Western Blot Analysis as described in Methods. The results indicated that the majority (≥ 90%) of SsPK2, when overexpressed in BL21SsPK2, was present in inclusion bodies. Attempts to produce soluble SsPK2 by alternative methods such as growing BL21SsPK2 cells at 37°C following IPTG induction also resulted in the formation of inclusion bodies. Furthermore, attempts to solubilize SsPK2 from inclusion bodies using buffer containing concentrations of urea ≤ 5 M or nonionic detergents such as Triton X-100 or octyl glucoside were only marginally effective (< 20% solubilized). Consequently, it was concluded that the most effective method to solubilize SsPK2 would be to use a buffer containing 5M urea.

It was asked if SsPK2 exhibited protein kinase activity even in the presence of 5 M urea (Urea is more compatible with protein refolding than other denaturing agents such as SDS) (Doonan, 1996). Since it was previously shown that the *S. solfataricus* protein kinase (described in Chapter III) could autophosphorylate on threonine residues we decided to determine if SsPK2 could autophosphorylate as well. The urea soluble fraction containing SsPK2 was incubated with [γ-^{32}P] ATP, at 65°C, for 1 hour then the proteins within resolved using SDS-PAGE as described in Methods. As seen in Figure 5-3, one [^{32}P] phosphorylated band with a M_r of ≈ 72 kDa (SsPK2 is ≈ 66.5 kDa and its N-terminal His tag adds an additional ≈ 5 kDa) was observed. This [^{32}P] phosphorylated species comigrated with that identified immunologically using Anti-Xpress antibody (Invitrogen) as the recombinant protein SsPK2. This suggested that either an endogenous protein kinase activity contained within the urea soluble fraction was responsible for phosphorylating SsPK2, or that SsPK2 was able to autophosphorylate. Curiously, the Anti-Xpress antibody also immunoreacted with several polypeptides of smaller M_r than the recombinant SsPK2 protein. Such behavior suggested that SsPK2 was being

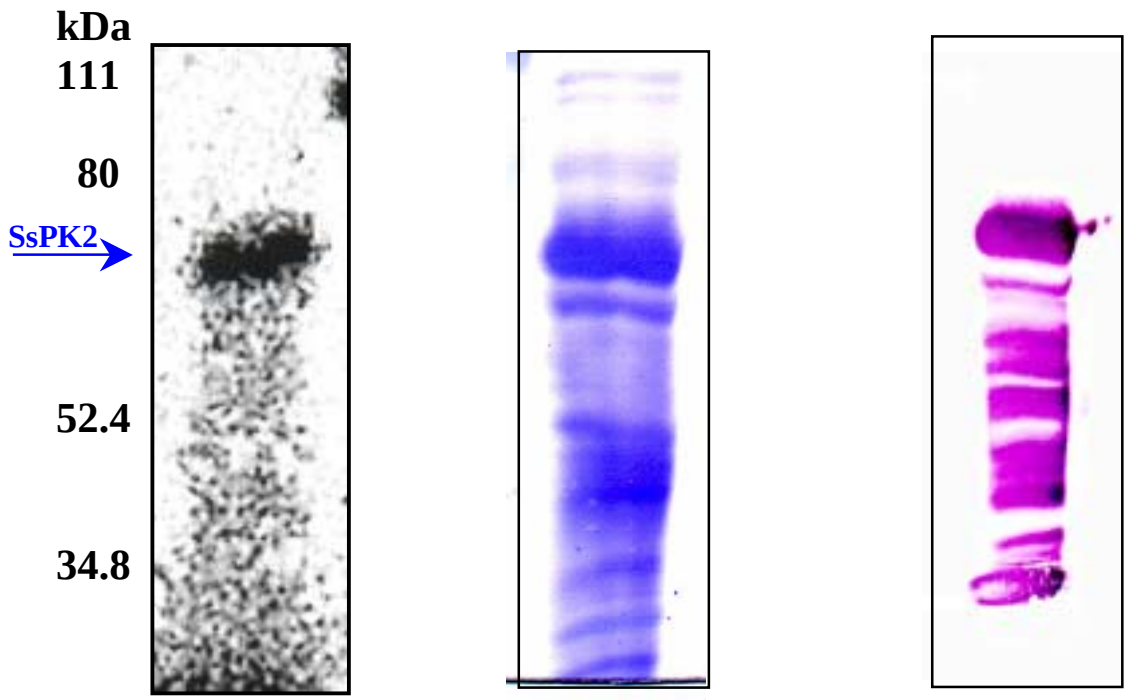


Figure 5-3. Recombinant Protein SsPK2 Appears to Autophosphorylate.

The urea-solubilized fraction, ≈ 50 µg, from induced BL21SsPK2 cells was incubated with [γ^{32}P] ATP, at 65°C, for 1 hour as described in Methods. Next, the protein was resolved on a 10% (w/v) SDS-polyacrylamide gel and stained with Coomassie R-250 as described in Methods. Shown is a picture of the Autoradiogram (left), Coomassie R-250 stained proteins (middle), and the Western Blot (right) of the gel. Molecular weight markers are indicated to the left and the position of SsPK2, at ≈ 72 kDa, is indicated with the blue arrow (**SsPK2 →**).

proteolytically degradated either during its purification, subsequent to translation inside BL21SsPK2 or during the assay. The same pattern of degradation was observed when the protein content of whole BL21SsPK2 cells was analyzed following induction. It therefore was concluded that a majority of the proteolysis occurred intracellularly and not during the purification of SsPK2. Since this was considered an "acceptable loss," we did not explore the use of other strains of *E. coli* genetically manipulated to be deficient in proteases.

The identity of SsPK2 was confirmed using MALDI- MS analysis. The portion of the gel containing the ≈ 72 kDa Coomassie stained species in Figure 5-3 was excised and digested with trypsin as described in Methods. The masses of the resulting peptides were determined using MALDI-MS and compared to the theoretical peptide masses of the recombinant SsPK2 protein, which were derived using the web-based Peptide Mass Program provided by ExPASy Proteomics Server (http://www.expasy.ch/tools/peptide-mass.html). The "option settings" for the Peptide Mass Program were programmed for: one maximum missed trypsin cleavage, all cysteines in reduced form, and all methionines have not been oxidized. Fourteen peptides derived from the trypsin digest of the recombinant protein SsPK2 matched the hypothetical peptide masses, within ± 0.5 Da, predicted using the Peptide Mass Program. These peptides covered ≈ 40% of the recombinant protein's primary sequence. It therefore was concluded that the ≈ 72 kDa polypeptide was indeed SsPK2.

Purification and Characterization of SsPK2

In order to determine if the source of the protein kinase activity detected in the urea soluble fraction was SsPK2, the recombinant protein was purified using metal affinity chromatography. A chelating Sepharose Fast Flow (Pharmacia) column (1 ml) was charged with one volume of 0.1 M $NiSO_4$ as indicated by the manufacturer. The column was then washed with five volumes ddH_2O and equilibrated with 5 volumes of 50 mM MOPS, pH 7.0, containing 5 M urea. A portion of the urea soluble fraction (1.0 ml) was applied to the column, the flow thru collected, and then reapplied to the column. This was repeated three times. Next, the column was washed with five volumes of 50 mM MOPS, pH 7.0, containing 5 M urea, followed by five volumes of 50 mM MOPS,

pH 7.0, containing 5 M urea and 50 mM imidazole. Adherent proteins then were eluted from the column with 5 volumes of 50 mM MOPS, pH 7.0, containing 5 M urea and 250 mM imidazole. Fractions, 1 ml, were collected and analyzed for total protein using SDS-PAGE. Those fractions containing SsPK2 alone, as identified by Coomassie R-250 staining and Western Blot Analysis (usually fractions 2 – 4), were pooled and retained as the Ni fraction (Figure 5-4). Next, the Ni fraction was incubated with [γ-^{32}P] ATP, at 65^{o}C, for 1 hour to determine if recombinant SsPK2 could autophosphorylate. One [^{32}P] phosphorylated band at ≈ 72 kDa corresponding to SsPK2 (as determined by Western Blot Analysis) was observed. It was therefore concluded that SsPK2 itself was the phosphorylated species and that autophosphorylation had taken place.

Next, it was asked whether SsPK2 was the threonine-specific protein kinase from *S. solfataricus* described in Chapter III. Since it already had been established that the *S. solfataricus* protein kinase could autophosphorylate on threonine residues, the recombinant protein should likewise autophosphorylate on threonine residues. Phosphoamino acid analysis was therefore performed on [^{32}P] phosphorylated SsPK2 to determine which amino acid residue(s) were being modified. The analysis revealed that the phosphate was bound as phosphoserine (Figure 5-5), establishing that a covalent protein-phosphate bond had been formed. The results also indicated that this *ORF* was not the gene encoding the threonine-specific protein kinase described in Chapter III, and thus it appeared that we had cloned a gene (*ORF sso2387*) from *S. solfataricus* that encoded a previously unknown protein kinase, SsPK2. However, since the threonine-specific protein kinase was solubilized using detergent (25 mM octyl glucoside) and the recombinant SsPK2 protein was solubilized using urea (5.0 M), there remained a possibility that under these different conditions, SsPK2 autophosphorylated on serine residues, while the threonine-specific protein kinase autophosphorylated on threonine residues. Therefore, we solubilized the recombinant SsPK2 protein using 25 mM octyl glucoside, and the threonine-specific protein kinase using 1 M urea (it was determined in preliminary experiments that a concentration of urea > 1.0 M resulted in a loss of protein kinase activity) and allowed the protein kinases to autophosphorylate with [γ-^{32}P] ATP, at 65^{o}C, for 1 hour. Next, phosphoamino acid analysis was performed on the [^{32}P] autophosphorylated proteins as described in Methods. The results indicated that

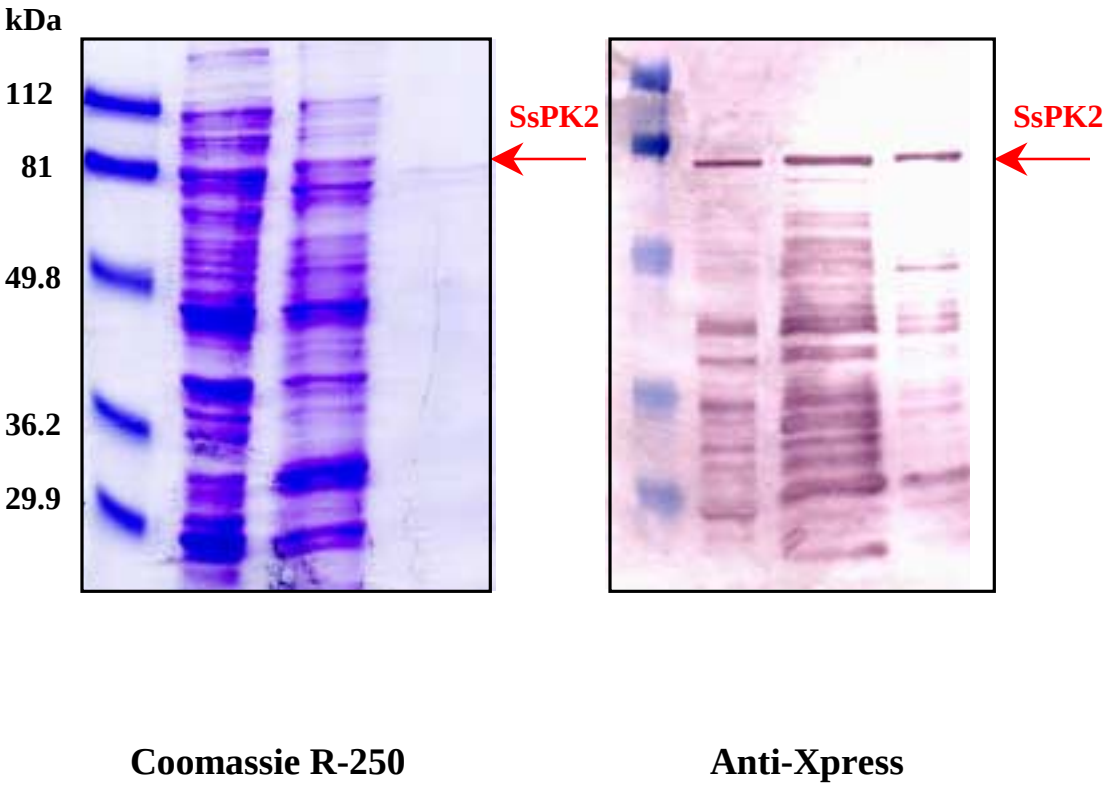


Figure 5-4. SDS-PAGE of SsPK2 Stained with Coomassie R-250 or Detected Immunologically Using Anti-Xpress.

Portions of the cell extract, urea soluble fraction, and Nickel-chelating column fraction were resolved on a 10% (w/v) SDS-polyacrylamide gel. Lane 1 contains 8 µl of SDS-PAGE low range molecular weight marker, lane 2 contains 25 µg of the cell extract, lane 3 contains 15 µg of the urea soluble fraction, and lane 4 contains 1.0 µg (50 µl) of the Nickel fraction. The SDS-polyacrylamide gel was stained with Coomassie R-250 as described in Methods (left), or electroblotted to PVDF membrane and then the membrane incubated with Anti-Xpress Antibody as described in Methods to detect the recombinant protein (right). **SsPK2** is labeled in red (←).

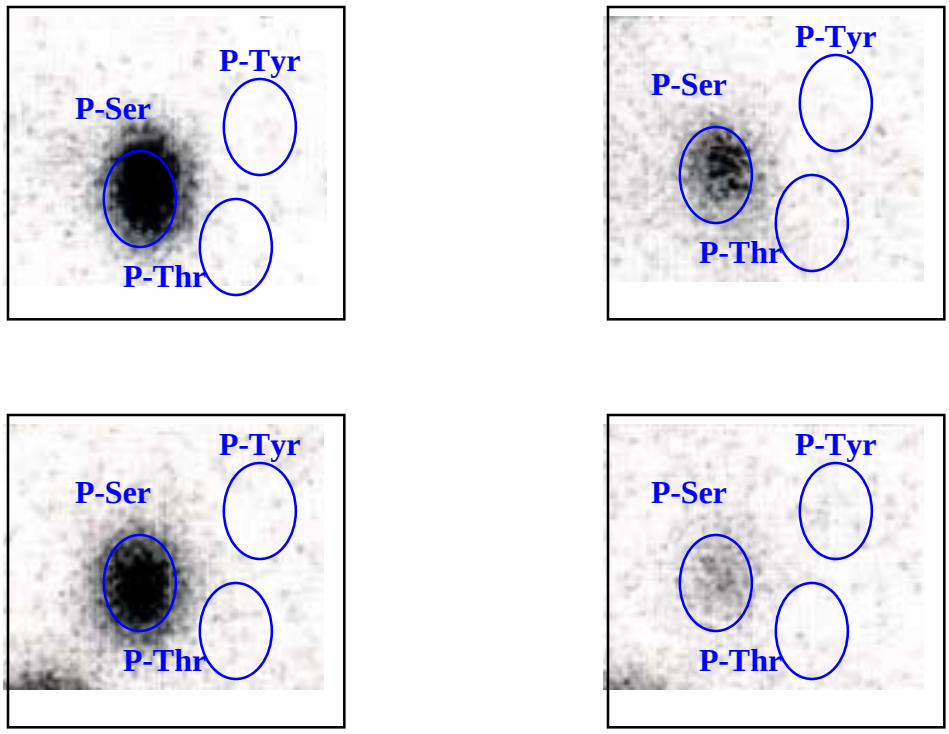


Figure 5-5. The Recombinant Protein, SsPK2, Encoded by *Sulfolobus solfataricus ORF sso2387* Exhibits a Tendency to Phosphorylate Serine Residues *In Vitro*.

The recombinant protein SsPK2, ≈ 0.05 µg of Ni fraction, was incubated with casein during a solution protein kinase assay as described in Methods, with the exception that the assay was conducted in 50 mM MOPS, pH 7.0, containing 5 M urea at 37°C; or SsPK2, ≈ 5 µg, alone was incubated with [γ^{32}P] ATP at 65°C in a buffer consisting of 50 mM MOPS, pH 7.0, containing 5.0 M urea. Following the assay the radiolabeled proteins were isolated by SDS-PAGE, transferred to PVDF membranes, and the portion of the membrane containing labeled protein was incubated in 6N HCl for 1 hour at 95°C. The supernatant liquid was then applied to a 20 x 20 cm silica gel TLC plate, along with standards of P-Ser, P-Thr and P-Tyr, and then subjected to 2-dimensional electrophoresis. The first dimension (bottom to top) was performed at pH 1.9 and the second dimension (right to left) at pH 3.5. Standards were visualized by ninhydrin staining, while phosphorylated amino acids were visualized by electronic autoradiography. Shown are the electronic autoradiograms of the TLC plates from the analysis of SsPK2 (top left), Casein (top right), BSA (bottom left), mixed histones (bottom right). Circles indicate the positions of P-Ser, P-Thr and P-Tyr.

under these condition autophosphorylation of the threonine-specific protein kinase occurred on threonine, while the autophosphorylation of the recombinant SsPK2 protein occurred on serine. Therefore, we concluded that SsPK2 was indeed a novel protein kinase from *S. solfataricus* and not the threonine-specific protein kinase described in Chapter III.

Next, SsPK2 was assayed using a conventional solution assay and casein as a potential phosphoacceptor substrate to determine if SsPK2 could phosphorylate an exogenous protein. The Ni fraction, ≈ 0.05 μg, was incubated with casein (1.0 mg/ml) and [γ^{32}P] ATP at 37^{o}C for 30 minutes, then the assay mixture subjected to SDS-PAGE and transferred to Immobilon P as described in Methods. When the [^{32}P] species were visualized using autoradiography, it was observed that [^{32}P] phosphate indeed had been transferred to the casein. Subsequent phosphoamino amino acid analysis revealed that the phosphate was bound to serine (Figure 5-5). This established that a covalent protein-phosphate bond had been formed and that the activity detected displayed the essential features of a protein kinase.

SsPK2 subsequently was challenged with several other exogenous protein and polymer substrates using a solution assay as described in Methods (Table 5-1). As seen in Table 5-1, SsPK2 was able to phosphorylate a variety of proteins including mixed histones, casein, reduced carboxyamidomethylated and maleylated (RCM) lysozyme, and BSA; while no phosphorylation could be detected with poly(Glu:Tyr). Phosphoamino acid analysis revealed that BSA and mixed histones were phosphorylated exclusively on serine residues (Figure 5-5).

When the catalytic performance of SsPK2 towards myelin basic protein (MBP), MLC peptide, or T8S peptide was examined using a solution assay, the results indicated that SsPK2 was capable of phosphorylating all three substrates. However, when the phosphorylated proteins or peptides were resolved on SDS-PAGE, little/no [^{32}P] was bound to the exogenous substrates, while SsPK2 had ≈ 2-times higher levels of [^{32}P] incorporation than controls lacking MBP (Figure 5-6), MLC peptide or T8S peptide, or in assays in which BSA was employed as exogenous substrate instead. Such behavior suggested that the former three species stimulated autophosphorylation of SsPK2. It was hypothesized that the basic residues contained within the primary sequences

Protein or Copolymer	Activity (pmol ^{32}P/min/mg)
BSA	578 ± 26.5
Casein	372 ± 16.7
Mixed Histones	1932 ± 216
RCM lysozyme	664 ± 22.3
Poly (Glu:Tyr)	n.d.*
Poly (Glu_4:Tyr)	n.d.*

*n.d. = not detectable, < 1.0 pmol ^{32}P / min / mg

Table 5-2. The Activity of SsPK2 Towards Exogenous Protein or Amino Acid Copolymer Substrates.

The activity of the recombinant protein SsPK2, 0.05 – 0.10 μg of Ni fraction, toward a range of exogenous proteins and amino acid copolymers was determined using the solution assay technique describes in Methods, with the exception that the assays were performed in a buffer of 50 mM, pH 7.0, containing 5.0 M urea at 37^{o}C. Substrate concentrations were 1.0 mg/ml for each assay. Shown are the averages of triplicate determinations plus or minus standard error.

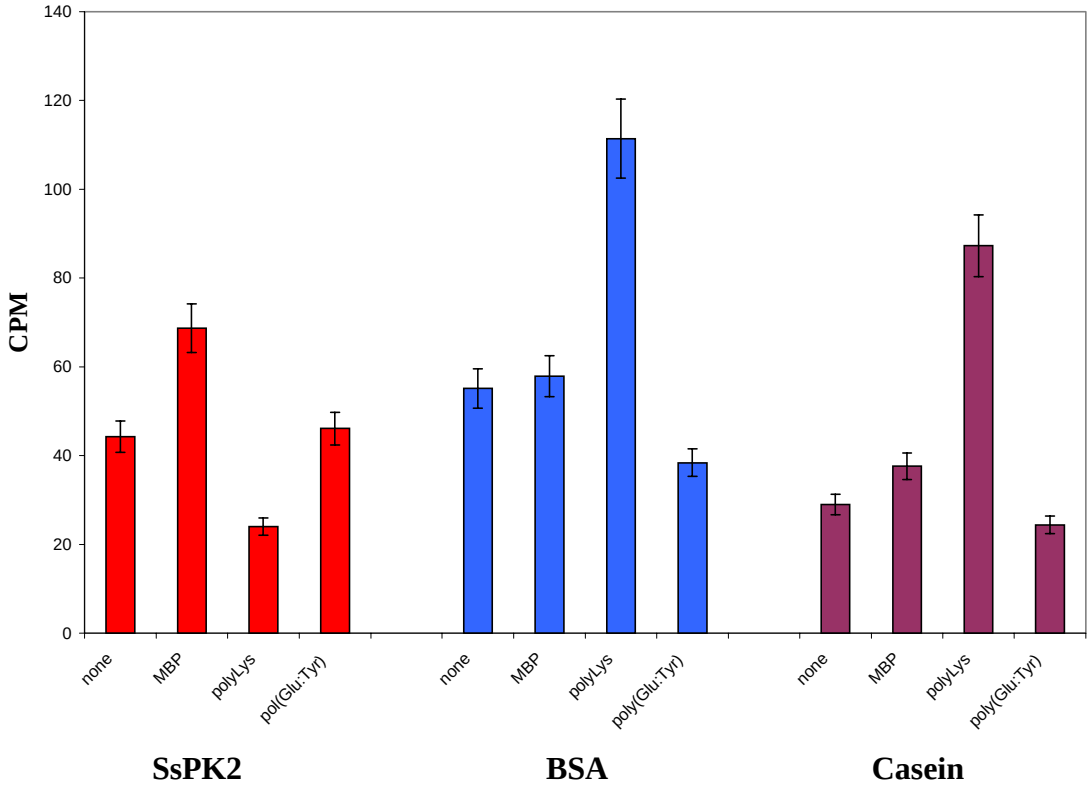


Figure 5-6. Stimulation of SsPK2 Autophosphorylation by MBP, and Phosphorylation of Exogenous Substrates by PolyLys.

The activity of the SsPK2 towards itself (**red**), BSA (**blue**) or casein (**purple**) was assayed in the presence of the indicated polycations, polyanion, or basic protein as described in Methods with the exception that assays were performed in 50 mM MOPS, pH 7.0, containing 5 M urea. Where indicated polyLys, MBP, or poly(Glu:Tyr) was added to a final concentration of 1.0 mg/ml. BSA or casein was added to a final concentration of 1 mg/ml. For autophosphorylation of SsPK2, 25 µg the urea soluble fraction was incubated with [γ^{32}-P] ATP at 65^{o}C for 1 hour. For phosphorylation of BSA or casein, 0.1 µg of the Ni fraction was incubated with the exogenous substrate at 37^{o} for 1 hour. Following the solution assay, proteins were resolved on a 10% (w/v) SDS-polyacrylamide gel and the amount of [^{32}P] incorporated into SsPK2, BSA, or casein was quantified using electronic autoradiography. Shown are the averages of duplicate determinations ± standard errors.

of these substrates were responsible for this stimulatory effect. Therefore, a solution assay in which polyLys was added to an assay mixture containing SsPK2 and [γ-^{32}P] ATP was performed at 65^{o}C. To determine whether polycations stimulated SsPK2 phosphorylation of exogenous substrates as well, a solution assay was performed at 37^{o}C in which polyLys was added to an assay mixture containing SsPK2, BSA or casein, and [γ-^{32}P] ATP. Following the assay, the proteins were resolved using SDS-PAGE and the quantity of [^{32}P] incorporated into SsPK2, BSA, or casein was quantified using electronic autoradiography. While polyLys was actually observed to inhibit autophosphorylation of SsPK2, polyLys was observed to stimulate SsPK2 activity towards both BSA and casein (Figure 5-6). Interestingly, in assays in which poly(Glu:Tyr) was substituted for polyLys, little/no change in protein kinase activity was observed.

Autophosphorylation of SsPK2

MS analysis was employed to identify the site(s) of autophosphorylation in SsPK2 we used MALDI-MS. The ≈ 72 kDa phosphorylated polypeptide corresponding to SsPK2, as seen in Figure 5-3, was excised and incubated with trypsin as described in Methods. The masses of the resulting peptides were determined using MALDI-MS and compared to the theoretical peptide masses of the recombinant SsPK2 protein, which were derived using the web-based Peptide Mass Program provided by ExPASy Proteomics Server (http://www.expasy.ch/tools/peptide-mass.html). The "option settings" for the Peptide Mass Program were programmed for: four maximum missed trypsin cleavages, all cysteines in reduced form, and all methionines have not been oxidized. Three peptides containing serine residues and having positive mass shifts, corresponding to the addition of one or more PO_3 groups (one PO_3 group has a mass of + 80 Da), relative to the calculated mass of a hypothetical peptide were observed in the MS spectrum (Table 5-3).

To determine whether autophosphorylation had any effect on SsPK2's ability to phosphorylate exogenous substrates, the enzyme was allowed to autophosphorylate by incubating it with γ[^{32}P] ATP at 65^{o}C for 1 hour (it was previously determined that time periods < 1 hour resulted in less than 100% autophosphorylation). Next, the mixture was cooled and incubated with BSA and [γ-^{32}P] ATP at 37^{o}C for 1 hour during a solution

Peptide	Residues	Mass of MH^+		Corresponding Number of PO_4 (Added Mass)
		Observed	Calculated	
VAATFRREV**S**-A**S**G**SS**VVIRR	266-285	2308.08	2148.20	2 (160 Da)
LLA**SS**REIIQN-EE**S**YKLVK	535-553	2300.40	2220.22	1 (80 Da)
REMERRALFL-KLLA**SS**R	523-540	2236.27	2076.18	2 (160 Da)

Table 5-3. Identification of Possible Autophosphorylation Sites on SsPK2 using MALDI-MS.

Autophosphorylated SsPK2 was digested with trypsin and the digest was analyzed by MALDI-MS as described in Methods and the text. The mass values observed were compared to theoretical masses calculated from the known SsPK2 sequence as described in the text. Shown are the observed and calculated peptide masses (MH^+) as well as the corresponding peptides and possible serine phosphorylation sites in bold (**S**). Also shown are the corresponding numbers of PO_4 groups and the masses of each (Da).

assay as described in Methods. These results were compared with data obtained when i) no prior autophosphorylation of SsPK2 occurred or ii) SsPK2 was incubated at 65°C for 1 hour but no [γ-^{32}P] ATP was added. As seen in Figure 5-7, autophosphorylated SsPK2 displayed a noticeable decrease in activity towards BSA relative to the control (no preheat treatment).

However, the control in which SsPK2 was preheated at 65°C for 1 hour without the addition of [γ-^{32}P] ATP also displayed a decrease in activity. Consequently, it was difficult to determine if the decrease in enzymatic activity was due to autophosphorylation or preheating SsPK2.

Metal Ion Preference of SsPK2

In order to determine the divalent metal ion requirement of SsPK2 for both autophosphorylation and the phosphorylation of exogenous substrates, solution assays were performed as described in Methods with the exception that, in place of the standard divalent metal ion cofactor, Mn^{2+}, Ca^{2+}, Mg^{2+}, Mn^{2+}, Ni^{2+}, and Zn^{2+} were added to a final concentration of 5 mM, and assays were performed in 50 mM MOPS, pH 7.0, containing 5 M urea. Control assays, in which EDTA was added to the reaction mixture to a final concentration of 5 mM in place of a divalent metal, were performed in parallel. Analysis of the cofactor requirements at 37°C (a temperature necessitated by the lower thermal stability of substrate proteins) revealed a strong preference for Mn^{2+} over other divalent metal ions for the phosphorylation of exogenous substrates (Figure 5-8). However, when the enzyme's ability to autophosphorylate was measured at temperatures within the physiological range of *S. solfataricus*, 65°C, Mg^{2+} proved equally effective as a cofactor (Figure 5-8). Interestingly, Ca^{2+}or Zn^{2+}, also served as a divalent metal ion cofactor for autophosphorylation of SsPK2 at 65°C. To validate the autophosphorylation results observed using Ca^{2+}, a second assay was performed in which EGTA was added to the assay mixture to a final concentration of 20 mM. In these assays, little/no protein kinase activity was observed.

Nucleotide Preference of SsPK2

To evaluate the nucleotide specificity of SsPK2, the ability of protein kinase

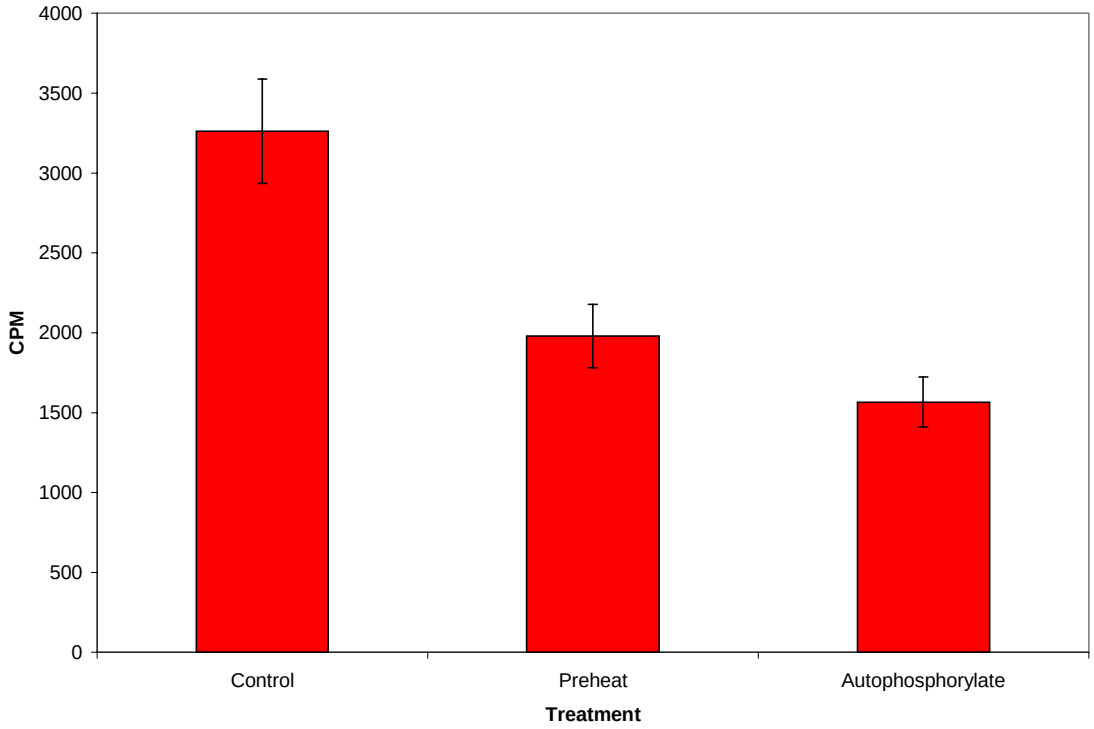


Figure 5-7. Effects of Autophosphorylation of SsPK2 on Enzymatic Activity.

The effect of autophosphorylation of SsPK2 towards phosphorylation of BSA was determined using a solution assay as described in Methods, with the exception that the assay was conducted at 37°C in a solution of 50 mM MOPS, pH 7.0, containing 5 M urea. SsPK2, 20 µg of the urea soluble fraction, was either preheated at 65°C for 1 hour, or preheated at 65°C in a solution containing [γ-^{32}P] ATP, 5 mM Mn^{2+} and 5 mM Mg^{2+} for 1 hour to allow SsPK2 to autophosphorylate. Next, the protein kinase activity of SsPK2 towards BSA (1.0 mg/ml) was determined using a solution assay. Shown are the results of duplicate experiments ± standard error for control (no preheat treatment), preheated SsPK2 preheated, or autophosphorylated SsPK2.

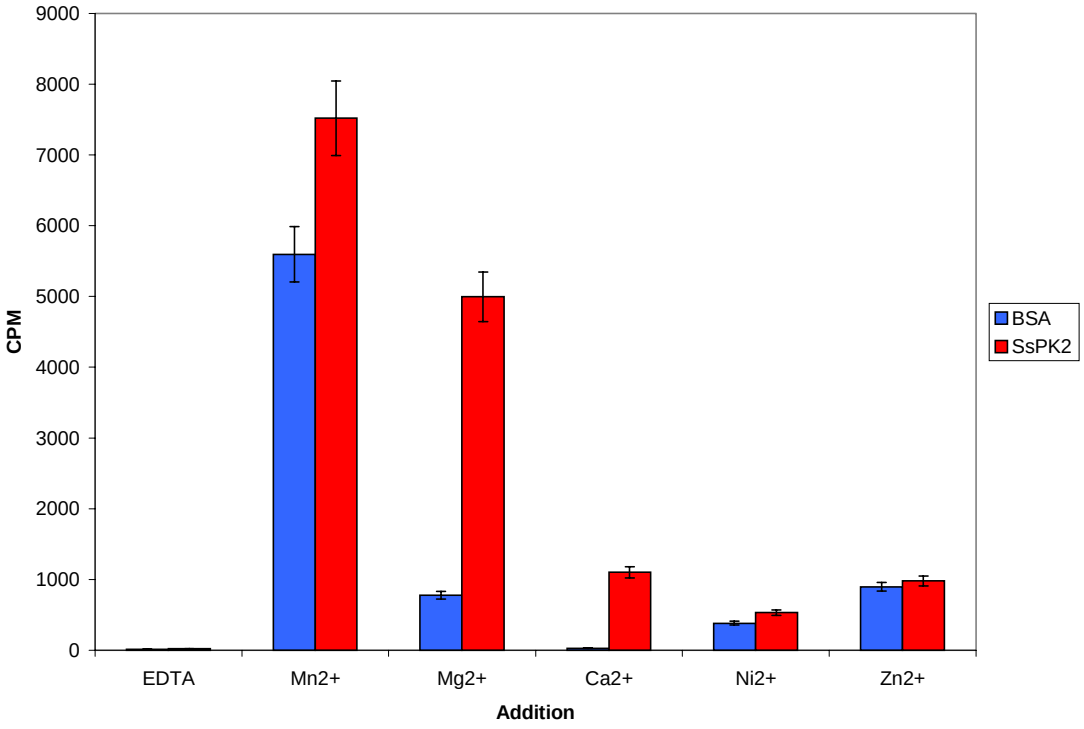


Figure 5-8. Metal Ion Cofactor Preference of SsPK2 for Autophosphorylation and Phosphorylation of Exogenous Substrate.

The activity of the SsPK2 was assayed in the presence of the indicated divalent metal ions at both 37°C and 65°C. Conditions for solution assays were as described in Methods with the exception that, in the place of the standard divalent metal, Mn^{2+}, the compounds listed were added at a final concentration of 5 mM, and assays were performed in 50 mM MOPS, pH 7.0, containing 5 M urea. For the assays conducted at 37°C, Ni fraction, 0.06 µg, was used as the source of the protein kinase activity and BSA was used as the phosphoacceptor substrate at a final concentration of 1.0 mg/ml. Following the assay the amount of [^{32}P] incorporation by BSA was determined by spotting aliquots on Whatman 3MM paper, washing with TCA/NaPPi, and counting the radioactivity using liquid scintillation counter as described in Methods. For the autophosphorylation assays conducted at 65ºC, 35 µg urea soluble fraction was used. Following the assays the amount of [^{32}P] incorporation into SSPK2 was determined by resolving the proteins contained within the urea soluble fraction using SDS-PAGE, then quantifying the amount of [^{32}P] incorporated into the ≈ 72 kDa protein (SsPK2) using electronic autoradiography as described in Methods. Shown is the average of duplicate determinations ± standard error

either to autophosphorylate or to catalyze the transfer of [^{32}P] phosphate to an exogenous substrate from [γ-^{32}P] ATP or [γ-^{32}P] GTP was assayed in solution. Conditions for the assay were as described in Methods with the exception that either [γ-^{32}P] ATP or [γ-^{32}P] GTP was used as the potential nucleotide donor substrate, and assays were performed in 50 mM MOPS, pH 7.0, containing 5 M urea at 37^{o}C (exogenous substrate) or 65^{o}C (autophosphorylation). No protein kinase activity was detected when the SsPK2 contained within the Ni fraction was incubated with a variety of exogenous protein substrates (BSA, casein, mixed histones, MBP, or RCM lysozyme) using [γ-^{32}P] GTP as the potential phosphoryl donor. However, when the urea soluble fraction was incubated with [γ-^{32}P] GTP at 65^{o}C during a solution assay and the proteins within resolved using SDS-PAGE, a [^{32}P] autophosphorylated species corresponding to ≈72 kDa SsPK2 was visualized using autoradiography (Figure 5-9). When compared to ATP, though, GTP was a considerably poorer substrate for SsPK2.

Sensitivity of SsPK2 to Inhibitors of Eukaryotic Protein Kinases

Because the *S. solfataricus* protein kinase displayed an ability to phosphorylate serine residues *in vitro*, a characteristic of eukaryotic protein kinases such as the cAMP-dependent protein kinases, we decided to challenge the enzyme with several known inhibitors of eukaryotic protein kinases. These included PKI peptide, a highly specific inhibitor of the cAMP-dependent protein kinase; genistein, an inhibitor of protein-tyrosine kinases; tamoxifen, an inhibitor or protein kinase C; H7, a specific inhibitor of cAMP-dependent protein kinases; H89, a selective inhibitor of protein kinase A; ML-9, a selective inhibitor of MLC kinase; and staurosporine, a broad range protein kinase inhibitor (reviewed in Hemmings, 1997). A solution assay for determining protein kinase activity was performed at 37^{o}C using casein as the phosphoacceptor substrate as described in Methods, with the exception that protein kinase inhibitors were also added to the reaction mixture and the assay was performed in 50 mM MOPS, pH 7.0, containing 5 M urea. All the compounds were tested in millimolar concentrations, well above the concentrations for which they are known to inhibit their eukaryotic target enzymes.

Only tamoxifen inhibited SsPK2 in a concentration-dependent manner (Figure 5-10), exhibiting an apparent 50% inhibitory concentration (IC_{50}) between 200 – 300 μM, a

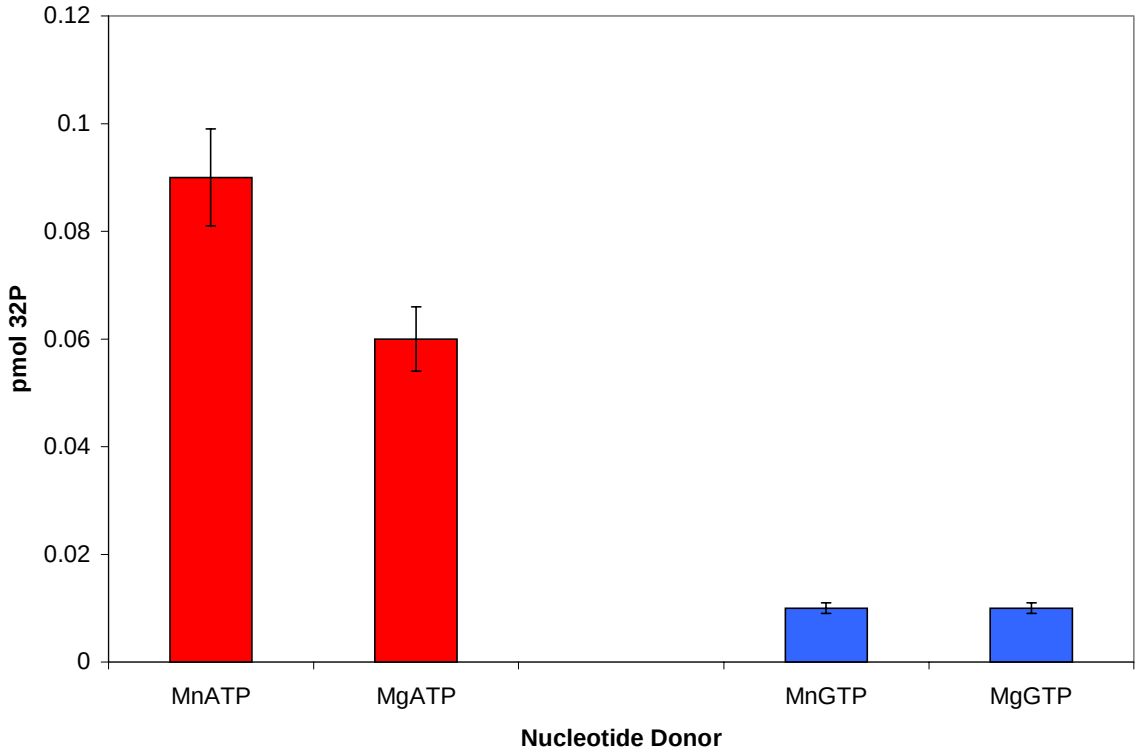


Figure 5-9. Autophosphorylation of SsPK2 Using ATP or GTP as Phosphoryl Donor.

The ability of SsPK2, 35 μg urea soluble fraction, to autophosphorylate using either [γ-^{32}P] ATP or [γ-^{32}P] GTP was examined using a solution assay as described in Methods with the exception that assays were performed at 65^{o}C in a solution of 50 mM MOPS, pH 7.0 containing 5 M urea, and the indicated divalent metal or nucleotide were added to the reaction mixture. Following the assay, proteins were resolved on a 10%(w/v) SDS-polyacrylamide gel and the [^{32}P] phosphorylated proteins visualized by autoradiography. The amount of [32P] incorporated by the ≈ 72 kDa protein, corresponding to SsPK2, was quantified using electronic autoradiography. Shown is the pmol of ^{32}P incorporated by SsPK2 using MnATP, MgATP, MnGTP, or MgGTP as phosphoryl donor.

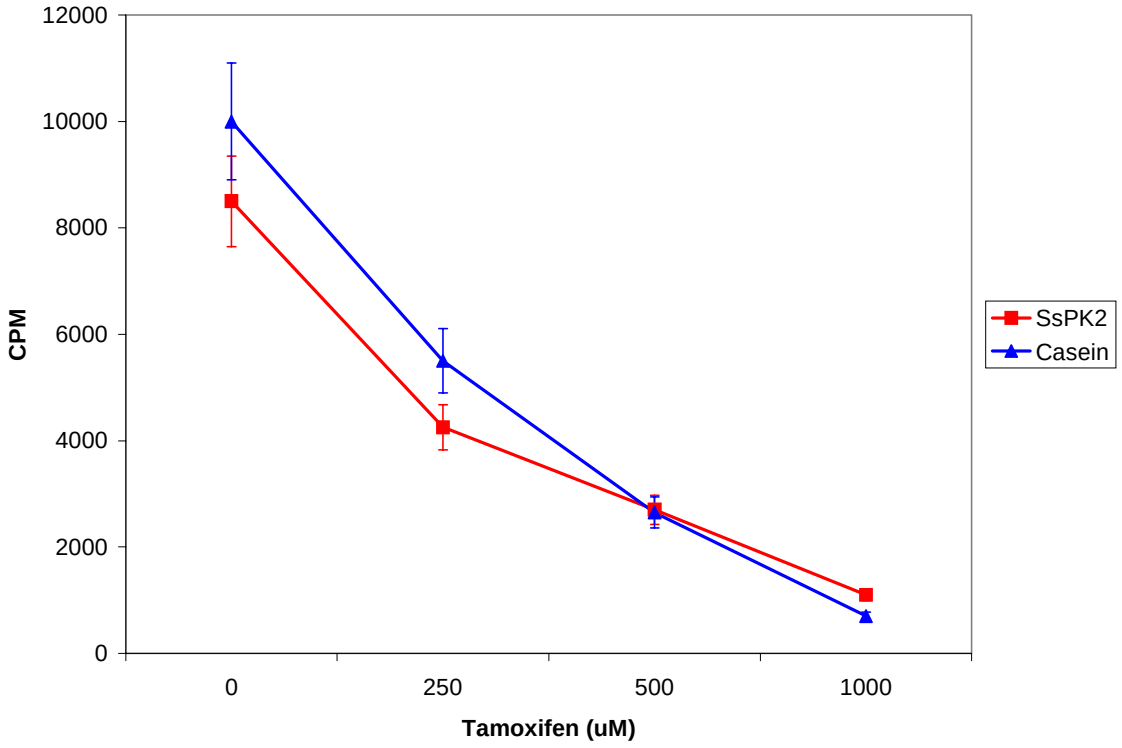


Figure 5-10. Sensitivity of SsPK2 to the Eukaryotic Protein Kinase Inhibitor, Tamoxifen.

The urea soluble fraction, 10 µg of protein, was assayed for protein kinase activity toward casein or autophosphorylation of SsPK2 using the solution assay procedure described in Methods with the exception that, where indicated, tamoxifen was present at the concentrations indicated on the ordinate, and assays were performed in 50 mM MOPS, pH 7.0, containing 5 M urea. Assays using casein, 1.0 mg/ml, were performed at 37°C. Following the assay the amount of [^{32}P] incorporation by casein was determined by spotting aliquots on Whatman 3MM paper, washing with TCA/NaPPi, and counting the radioactivity using a liquid scintillation counter as described in Methods. Assays using SsPK2 alone were performed at 65°C. Following the assays the amount of [^{32}P] incorporation by SSPK2 was determined by resolving the proteins using SDS-PAGE, then quantifying the amount of [^{32}P] incorporated into the ≈ 72 kDa protein (SsPK2) using electronic autoradiography as described in Methods. Shown are the averages of duplicate determinations, ± standard error

range well above the concentrations known to inhibit protein kinase C (IC_{50} = 10 μM). The other protein kinase inhibitors that were tested displayed little/no effect on SsPK2's ability to phosphorylate casein.

We also assayed SsPK2's ability to autophosphorylate at 65°C in the presence of these protein kinase inhibitors. Again, tamoxifen was the only inhibitor observed that inhibited and even prevented autophosphorylation of SsPK2. Like with casein though, an apparent IC_{50} of between 200 – 400 μM was observed, which is well above that for protein kinase C (Figure 5-10).

Construction and Characterization of a Truncated Amino-Terminal Protein Used to Define the Protein Kinase Domain of SsPK2

Upon examination of the primary sequence of SsPK2 (Figure 4-14), one notices that its apparent protein kinase domain is located within the carboxy-terminal half of the protein. Therefore, we decided to construct a truncated form of SsPK2 containing the perceived catalytic domain and determine if this protein also exhibited protein kinase activity. We were also optimistic that a truncated form of SsPK2 would be soluble when expressed in *E. coli*.

The region of SsPK2 that contains the perceived protein kinase subdomains starts at approximately residue 324 (Gly_{50} of subdomain I) and extends through the carboxy terminus of the protein. Therefore an amino terminal deletion mutant of SsPK2 was constructed (corresponding to amino acids 314 – 583) using oligonucleotide primers and the polymerase chain reaction in a procedure similar to that described above for cloning the SsPK2 gene (*ORF sso2387*) into pCR T7/NT TOPO. For the PCR amplification of this deletion, pCR T7/NT TOPO plasmid containing *S. solfataricus* ORF sso2387 was used as template, oligonucleotide Ss67R01 was used as the reverse primer, and oligonucleotide Ss67F02 was used as forward primer (Figure 5-11). Amplification of genomic DNA by PCR using these two oligonucleotides as primers (Figure 5-2), ligation into linearized pCR T7/NT TOPO vector, and transformation into TOP 10 F' cells (Invitrogen) were performed as described above for cloning the gene for SsPK2 into pCR

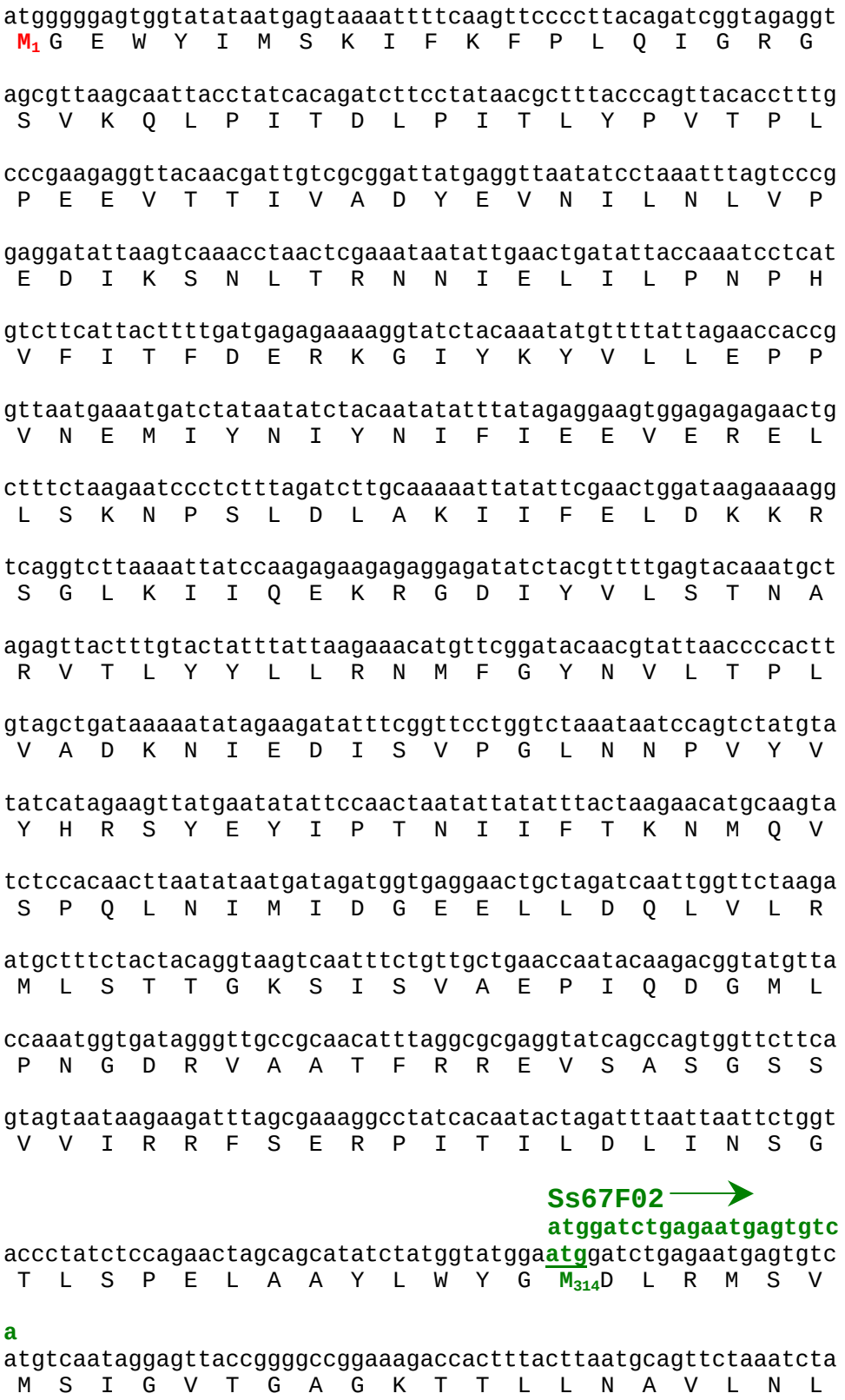

atgggggagtggtatataatgagtaaaattttcaagttccccttacagatcggtagaggt
M1 G E W Y I M S K I F K F P L Q I G R G
agcgttaagcaattacctatcacagatcttcctataacgctttacccagttacacctttg
S V K Q L P I T D L P I T L Y P V T P L
cccgaagaggttacaacgattgtcgcggattatgaggttaatatcctaaatttagtcccg
P E E V T T I V A D Y E V N I L N L V P
gaggatattaagtcaaacctaactcgaaataatattgaactgatattaccaaatcctcat
E D I K S N L T R N N I E L I L P N P H
gtcttcattacttttgatgagagaaaaggtatctacaaatatgttttattagaaccaccg
V F I T F D E R K G I Y K Y V L L E P P
gttaatgaaatgatctataatatctacaatatatttatagaggaagtggagagagaactg
V N E M I Y N I Y N I F I E E V E R E L
ctttctaagaatccctctttagatcttgcaaaaattatattcgaactggataagaaaagg
L S K N P S L D L A K I I F E L D K K R
tcaggtcttaaaattatccaagagaagagaggagatatctacgttttgagtacaaatgct
S G L K I I Q E K R G D I Y V L S T N A
agagttactttgtactatttattaagaaacatgttcggatacaacgtattaaccccactt
R V T L Y Y L L R N M F G Y N V L T P L
gtagctgataaaaatatagaagatatttcggttcctggtctaaataatccagtctatgta
V A D K N I E D I S V P G L N N P V Y V
tatcatagaagttatgaatatattccaactaatattatatttactaagaacatgcaagta
Y H R S Y E Y I P T N I I F T K N M Q V
tctccacaacttaatataatgatagatggtgaggaactgctagatcaattggttctaaga
S P Q L N I M I D G E E L L D Q L V L R
atgctttctactacaggtaagtcaatttctgttgctgaaccaatacaagacggtatgtta
M L S T T G K S I S V A E P I Q D G M L
ccaaatggtgatagggttgccgcaacatttaggcgcgaggtatcagccagtggttcttca
P N G D R V A A T F R R E V S A S G S S
gtagtaataagaagatttagcgaaaggcctatcacaatactagatttaattaattctggt
V V I R R F S E R P I T I L D L I N S G
Ss67F02
atggatctgagaatgagtgtc
accctatctccagaactagcagcatatctatggtatggaatggatctgagaatgagtgtc
T L S P E L A A Y L W Y G M314D L R M S V
a
atgtcaataggagttaccggggccggaaagaccactttacttaatgcagttctaaatcta
M S I G V T G A G K T T L L N A V L N L

Figure 5-11. DNA Derived Amino Acid Sequence of SsPK2trunk.

Shown is the DNA sequence for *S. solfataricus ORF sso2387* and the protein it encodes, SsPK2. The predicted **atg** start codon is underlined and red in color; the predicted **taa** stop codon is underlined and orange in color. The first Met residue encoded by sso2387 is red in color and labeled **M_1**, the first Met contained within SsPK2trunk is green in color and labeled **Met_{314}** while the last amino acid encoded by *sso2387*, **E_{583}**, is blue in color. Also shown are the oligonucleotide primers used to amplify the amino terminal deletion mutant of ORF sso2387 using PCR, as well as the region of the DNA sequence where they anneal. The forward primer (**Ss67F02**) is shown in green and the reverse PCR primer (**Ss67R01**) is shown in blue.

T7/NT TOPO. Recombinant plasmids were purified from TOP 10 F' cells using Wizard Minipreps (Promega) and sequenced using both forward (T7 primer) and reverse (pRSET) primers as described in Methods, to verify that the PCR amplified genomic DNA was indeed the amino terminal deletion mutant of OFR sso2387 and that it had ligated into the vector correctly. Next, the plasmids were transformed into *E. coli* strain BL21(DE3)pLysS as described above. *E. coli* BL21(DE3)pLysS cells transformed with the pCR T7/NT TOPO plasmid containing the deletion mutant of ORF sso2387 were named BL21SsPK2trunk.

The gene encoding the SsPK2trunk was expressed as described above for SsPK2. Unfortunately just like SsPK2, we discovered that the majority (≥ 90%) of SsPK2trunk, when overexpressed in BL21SsPK2trunk, was in the form of inclusion bodies. Therefore we used 50 mM MOPS, pH 7.0 containing 5 M urea to solubilize SsPK2trunk. The urea soluble protein was purified by metal affinity chromatography as described previously (Figure 5-12).

SsPK2trunk was assayed for protein kinase activity using a variety of exogenous protein and polymer substrates. As seen in Table 5-3, SsPK2trunk was able to phosphorylate a variety of exogenous protein substrates using [γ-^{32}P] ATP. The observed activity for SsPK2trunk was similar to that observed for SsPK2. Furthermore, SsPK2trunk was observed to autophosphorylate itself at 65^{o}C as well. Therefore, we concluded that SsPK2trunk did define the catalytic domain of the protein kinase.

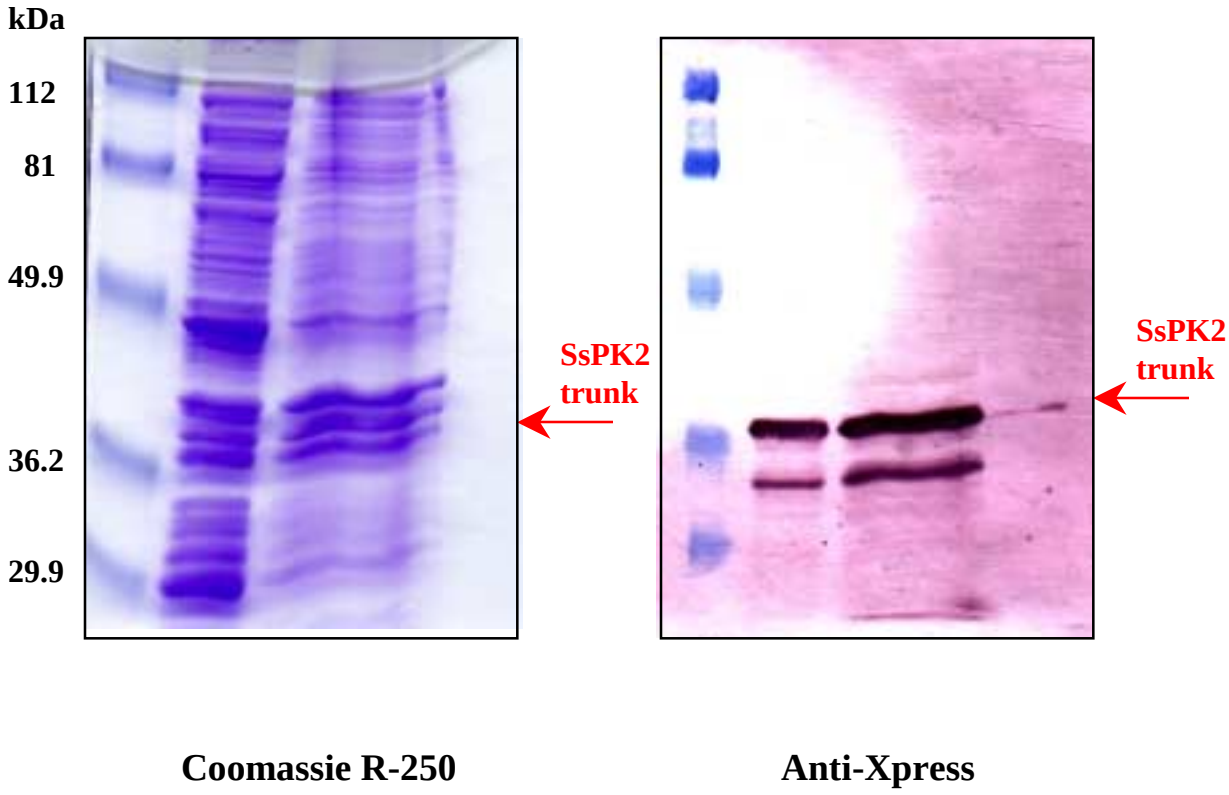


Figure 5-12. SDS-PAGE of SsPK2trunk Stained with Coomassie R-250 or Detected Immunologically Using Anti-Xpress.

Portions of the cell extract, urea soluble fraction, and Nickel-chelating column fraction were resolved on a 10% (w/v) SDS-polyacrylamide gel. Lane 1 contains 8 µl of SDS-PAGE low range molecular weight marker, lane 2 contains 25 µg of the cell extract, lane 3 contains 15 µg of the urea soluble fraction, and lane 4 contains 1.0 µg (70 µl) of the Nickel fraction. The SDS-polyacrylamide gel was stained with Coomassie R-250 as described in Methods (left), or electroblotted to PVDF membrane and then the membrane incubated with Anti-Xpress Antibody as described in Methods to detect the recombinant protein (right).

Protein or Peptide	Activity (pmol ^{32}P/min/mg)
BSA	2000 ± 40
Casein	142 ± 6.7
Mixed Histones	145 ± 16
Histone H4	280 ± 21
RCM lysozyme	132 ± 18
Poly (Glu:Tyr)	n.d.*
Poly (Glu_4:Tyr)	n.d.*

*n.d. = not detectable, < 1.0 pmol ^{32}P / min / mg

Table 5-4. The Activity of SsPK2trunk Towards Exogenous Protein or Amino Acid Copolymer Substrates.

The activity of the recombinant protein SsPK2trunk, 0.05 μg of Ni fraction, toward a range of exogenous proteins and amino acid copolymers was determined using the solution assay technique describes in Methods, with the exception that the assays were performed in a buffer of 50 mM, pH 7.0, containing 5.0 M urea at 37^{o}C. Substrate concentrations were 1.0 mg/ml for each assay. Shown are the averages of duplicate determinations plus or minus standard error.

CHAPTER VI

Identification of Proteins that Undergo Phosphorylation From the Membrane Fraction of *Sulfolobus solfataricus* Using MALDI-MS Analysis.

The ability to identify proteins from minute quantities of material has advanced dramatically over the past several years through new developments in mass spectrometric (MS) techniques and instrumentation. One of the most popular MS techniques is matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) coupled with web-based peptide mass search programs (Patterson, 1995), also known as peptide mass fingerprinting (Link, 1999). Its underlying premise is that the distinct amino acid sequence of a protein will give rise to its own unique set of peptides when it is enzymatically digested with specific proteases. The unique combination of masses characteristic of this set of peptides provides a fingerprint that can be used to identify the protein using computer-searchable databases with a high degree of confidence. Studies have shown that just four to six proteolytic peptides measured with a mass accuracy between 0.1 - 0.01% are required to unambiguously identify proteins in databases containing more than 100,000 entries (Mann et. al., 1993). The mass accuracy for MALDI-MS is generally ± 0.01% or better; therefore, it provides adequately accurate mass estimates for this purpose (Link, 1999).

Peptide mass profiling was employed to identify proteins of the detergent soluble membrane extract of *S. solfataricus* that underwent phosphorylation *in vitro*. *S. solfataricus* DE-52 fraction (see Chapter III) was incubated at 65^{o}C in 20 mM MES, pH 6.5, containing 12.5 mM octyl glucoside, 50 μM ATP, 1 μCi/ul [γ-^{32}P] ATP, 5 mM $MnCl_2$, 2 mM DTT for 1 hour to radiolabel phosphoproteins. Next, proteins were resolved using SDS-PAGE or two-dimensional electrophoresis and stained with Coomassie R-250 as described in Methods. Autoradiography was used to visualize the proteins that underwent [^{32}P] phosphorylation *in vitro*. Five prominent [^{32}P] phosphorylated proteins were observed in the SDS-PAGE gel (Figure 6-1), while sixteen [^{32}P] phosphorylated proteins were observed in the 2-D gel (Figure 6-2).

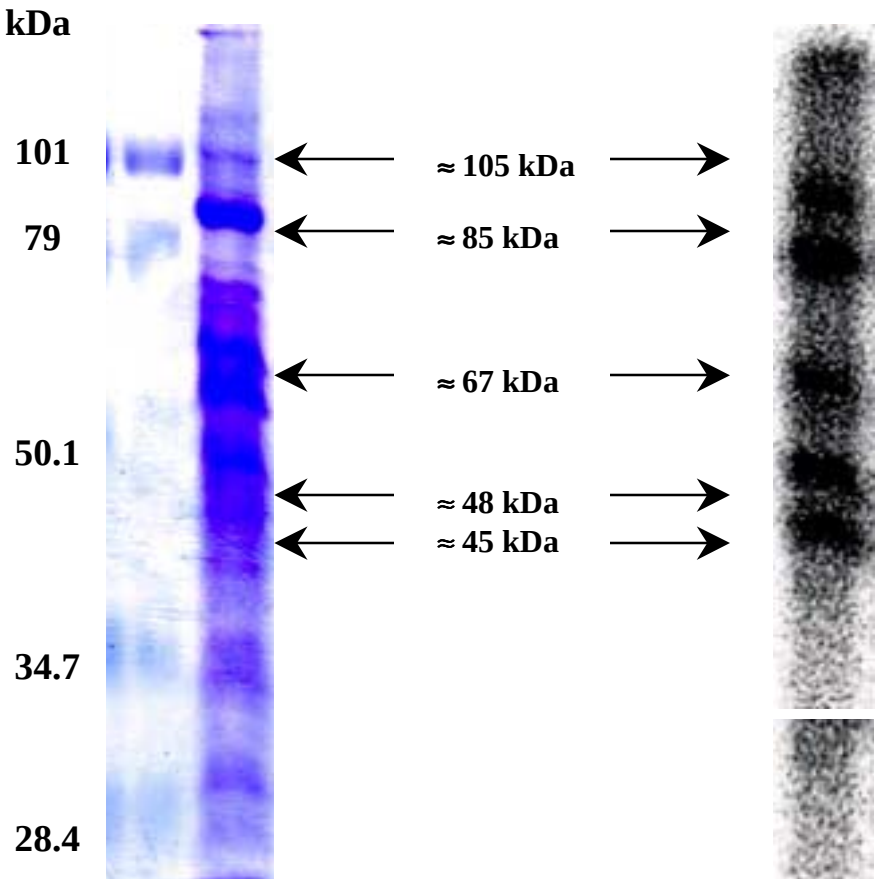


Figure 6-1. Analysis of [^{32}P] Phosphorylated Proteins From *Sulfolobus solfataricus* DE-52 Fraction Using SDS-PAGE.

Reaction mixture (100 µl) contained 20 mM MES, pH 6.5, containing 12.5 mM octyl glucoside, 50 µM ATP, 1 µCi/ul [γ-^{32}P] ATP, 5 mM $MnCl_2$, 2 mM DTT and 45 µg concentrated DE-52 fraction from *S. solfataricus*. In order to phosphorylate proteins in the DE-52 fraction proteins were incubated in the reaction mixture at 65^{o}C for 1 hour. The reaction was stopped and proteins precipitated by the addition of 3 volumes of ice-cold acetone as described in Methods. Precipitated proteins were resuspended in 4x SDS loading buffer by agitation on a vortex mixer, heated to 100^{o}C for 5 minutes, then separated on a 16 cm x 16 cm 10% (w/v) SDS-polyacrylamide gel as described in Methods. The gel was stained with 0.075% (w/v) Coomassie R-250 as described in Methods and analyzed by electronic autoradiography. Shown is the Coomassie stained gel (left) and electronic autoradiogram (right). Lane 1 of the Coomassie stained gel shows the migration of the molecular weight markers to the left. Five prominent [^{32}P] phosphorylated bands were observed at ≈ 105 kDa, ≈ 85 kDa, ≈ 67 kDa, ≈ 48 kDa, and ≈ 45 kDa. These bands were excised using a clean razor blade, digested in gel using sequencing grade trypsin, and the resulting peptide masses determined using MALDI-MS as described in Methods.

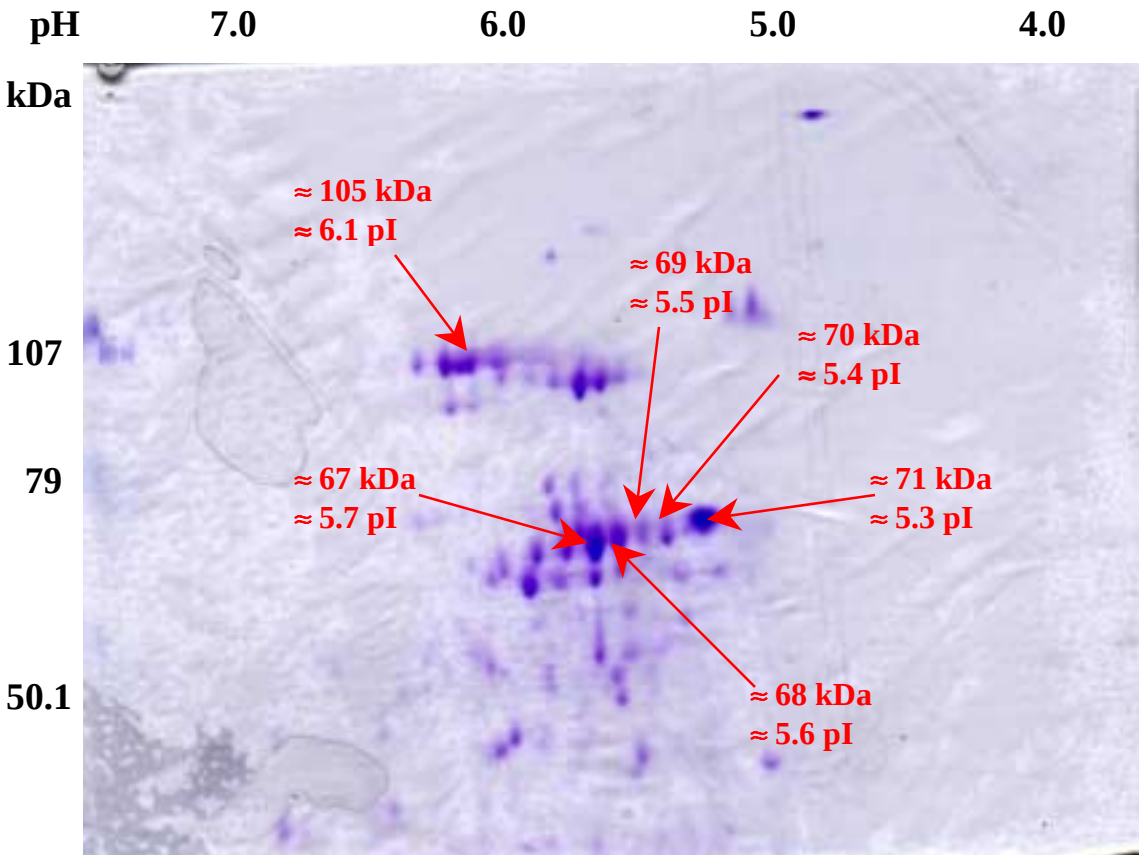


Figure 6-2. Analysis of [^{32}P] Phosphorylated Proteins From *Sulfolobus solfataricus* DE-52 Fraction Using Two-Dimensional Electrophoresis.

Reaction mixture (100 µl) contained 20 mM MES, pH 6.5, containing 12.5 mM octyl glucoside, 50 µM ATP, 1 µCi/ul [γ-^{32}P] ATP, 5 mM $MnCl_2$, 2 mM DTT and ≈ 75 µg concentrated DE-52 fraction from *S. solfataricus*. Proteins were allowed to phosphorylate by incubating the reaction mixture at 65°C for 1 hour. The reaction was stopped and proteins precipitated by the addition of 3 volumes of ice-cold acetone as described in Methods. Precipitated proteins were solubilized in 125 µl rehydration solution and resolved using a 7 cm 4–7 linear IPG strip (1st dimension) and a 8% (w/v) SDS-polyacrylamide gel (2nd dimension) as described in Methods. The gel was stained with 0.075% (w/v) Coomassie R-250 as described in Methods and [^{32}P] containing species visualized by electronic autoradiography. Shown is the Coomassie stained gel as well as the migration of the molecular weight markers to the left and pH gradient across the top. Sixteen [^{32}P] phosphorylated bands were observed. Of these, 6 (marked with red arrows) were selected for MALDI-MS analysis: ≈ 105 kDa, ≈ 6.1 pI; ≈ 67 kDa, ≈ 5.7 pI; ≈ 68 kDa, ≈ 5.6 pI; ≈ 69 kDa, ≈ 5.5 pI; ≈ 70 kDa ≈ 5.4 pI; and ≈ 71 kDa, ≈ 5.3 pI. These bands were excised using a clean razor blade, digested in gel using sequencing grade trypsin, and the resulting peptide masses determined using MALDI-MS as described in Methods.

MALDI-MS analysis was used to identify all five of the [^{32}P] phosphorylated proteins in the SDS-PAGE gel (Figure 6-1) and six of the most prominent [^{32}P] phosphorylated proteins in the 2-D gel (Figure 6-2). Briefly, each Coomassie stained band or spot was excised and incubated with trypsin as described in Methods. The masses of the resulting peptides were determined using MALDI-MS. These masses were used to search a web-based peptide-mass fingerprinting server at Rockefeller University, ProFound (http://www.proteometrics.com/), for *ORFs* in the genome of *S. solfataricus* whose predicted protein products generated theoretical peptide masses matching those observed (Zhang and Chait, 2000). ProFound settings were programmed for: *Archaea* taxonomic category; single protein search; protein mass was estimated using prestained molecular weight markers, ± 5 – 10 kDa; pI (for 2-D gels) was estimated from the linear isoelectric focusing strips, ± 0.5 – 1.0 pH unit; peptide masses were entered as monoisotopic masses with a mass tolerance of ± 1.0 – 2.0 Da and a charge state of MH^+; 1 – 2 missed trypsin cleavage sites were considered; and Cys were treated with iodoacetamide.

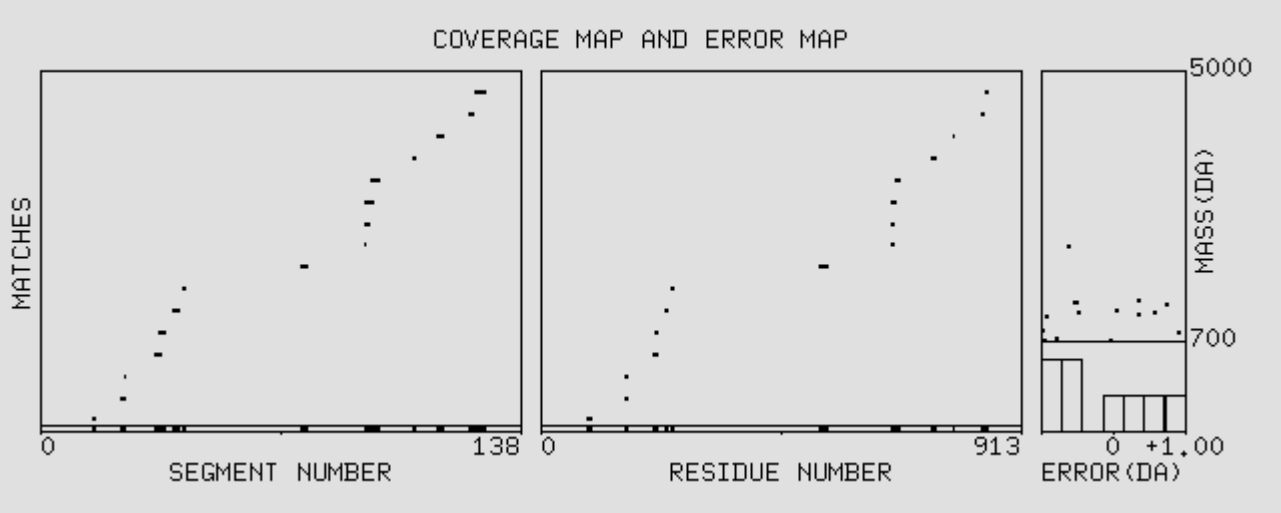
Phosphoamino acid analysis was performed on the [^{32}P] labeled polypeptides to determine the identity of the phosphoamino acids labeled *in vitro* with [γ-^{32}P] ATP. As shown in Table 6-1, eight of the polypeptides contained either phosphoserine or phosphothreonine, while three of the [^{32}P] phosphorylated polypeptides exhibited no detectable amino acid phosphoesters. Possible explanations for these negative results include the relatively low quantities of [^{32}P] phosphate incorporated or incorporation of [^{32}P] phosphate into acid-labile phosphoamino acids, i.e. acyl phosphates or phosphoramides.

ProFound Search Results for the [^{32}P] Phosphorylated Proteins Resolved Using SDS-PAGE

The database search, using the ProFound program, for the identity of the ≈ 105 kDa polypeptide (Figure 6-1) gave the results shown in Figure 6-3. In the search, the following parameters were used: taxonomic category of *Archaebacteria*, molecular weight between 90 – 120 kDa, peptide mass tolerance ± 1 Da, monoisotopic mass, maximum number of missed trypsin cleavages given as 2, and cysteines modified with

gi|12311972|emb|CAC22836.1| (AL512963) ATP-DEPENDENT HELICASE [Sulfolobus solfataricus]

```
Sample ID                   : SDS-PAGE 105 kDa
Digestion chemistry         : Trypsin
Number of measured peptides : 28
Number of matched peptides  : 12
Coverage of protein sequence: 14 %
```



Measured Mass(M)	Avg/ Mono	Computed Mass	Error (Da)	Residues From	To	Missed Cut	Cys	Met
727.48	M	727.49	-0.01	667	672	0	0	0
727.48	M	728.44	-0.96	248	253	0	0	0
760.64	M	761.43	-0.79	838	844	1	0	0
856.55	M	855.58	0.97	667	673	1	0	0
871.51	M	872.50	-0.99	783	789	1	0	0
1109.73	M	1110.65	-0.92	235	244	1	0	0
1127.06	M	1126.67	0.39	160	169	0	0	0
1153.32	M	1152.71	0.61	214	223	1	0	0
1153.32	M	1153.78	-0.46	667	676	2	0	0
1153.32	M	1152.71	0.61	217	226	1	0	0
1185.71	M	1185.65	0.06	845	853	2	0	0
1283.55	M	1282.77	0.78	159	169	1	0	0
1317.20	M	1317.72	-0.52	744	755	0	0	0
1317.20	M	1317.70	-0.50	89	100	0	0	0
1347.12	M	1346.74	0.38	674	684	2	0	0
2240.56	M	2241.18	-0.62	530	549	1	0	0

Figure 6-3. ProFound Identified the ≈ 105 kDa Polypeptide as an ATP-Dependent Helicase From *Sulfolobus solfataricus*.

Shown is a peptide coverage map of the ATP-dependent helicases from *S. solfataricus* (gi|12311972), and the matching peptide masses determined using MALDI-MS.

SDS-PAGE		2-D Electrophoresis	
[^{32}P] Phosphorylated Polypeptide	PAA	[^{32}P] Phosphorylated Polypeptide	PAA
≈ 105 kDa	P-Thr	≈ 105 kDa, ≈ 6.1 pI	P-Thr
≈ 85 kDa	n.d.	≈ 67 kDa, ≈ 5.7 pI	P-Thr
≈ 67 kDa	P-Thr	≈ 68 kDa, ≈ 5.6 pI	P-Thr
≈ 48 kDa	P-Ser	≈ 69 kDa, ≈ 5.5 pI	P-Thr
≈ 45 kDa	*n.d.	≈ 70 kDa ≈ 5.4 pI	P-Thr
		≈ 71 kDa, ≈ 5.3 pI	*n.d.

*n.d. = none detected

Table 6-1. Phosphoamino Acid Analysis of [^{32}P] Phosphorylated Proteins From the *Sulfolobus solfataricus* DE-52 fraction.

Shown are the results of the phosphoamino analysis performed on the [^{32}P] labeled polypeptides from the *S. solfataricus* DE-52 fraction (see Chapter III). The polypeptides were labeled using [γ-^{32}P] ATP and resolved using either SDS-PAGE (Figure 6-1) or two-dimensional electrophoresis (Figure 6-2) as described in the text. Phosphoamino acid analysis was performed as described in Methods.

iodoacetamide. A ProFound search with those parameters selected an ATP-dependent helicase from *S. solfataricus* (gi|12311972), with a probability score of 1.1e-05, as the best candidate. Of the twenty eight peptide masses entered into the ProFound program, twelve matched an ≈ 105 kDa ATP-dependent helicases, yielding 14% protein sequence coverage, ten of the remaining peptide masses were identified as trypsin autoproteolysis products, and six of the peptide masses remained unaccounted for (perhaps due to nonspecific cleavages by trypsin, posttranslational modifications, or a contaminating polypeptide species that comigrated in the SDS-polyacrylamide gel).

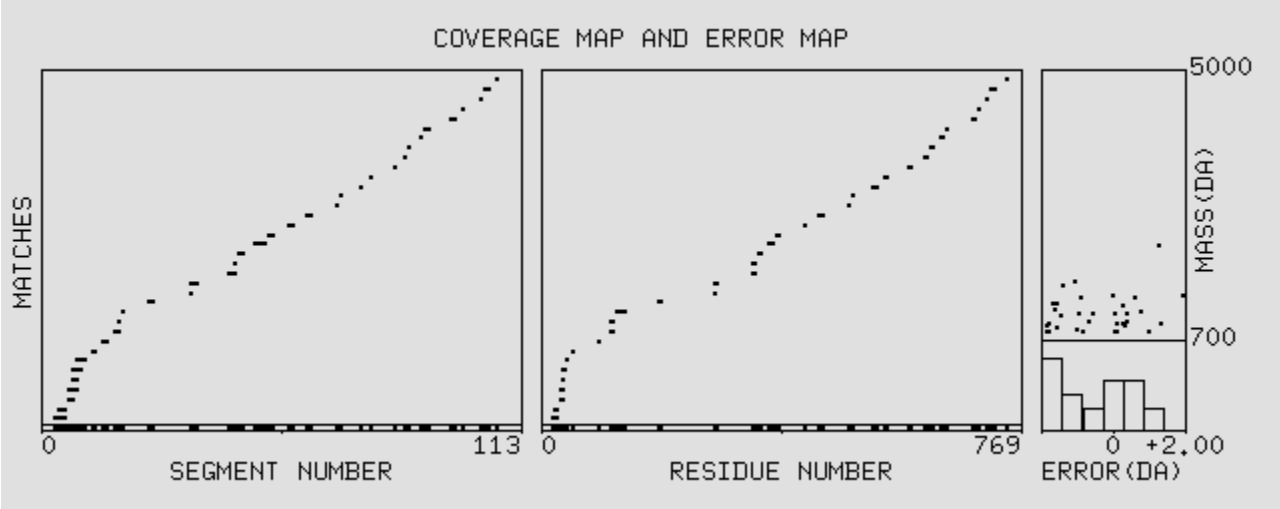
The database search, using the ProFound program, for the identity of the ≈ 85 kDa polypeptide (Figure 6-1) gave the results shown in Figure 6-4. In the search the following parameters were used: taxonomic category of *Archaebacteria*, molecular weight between 75 – 95 kDa, peptide mass tolerance ± 2 Da, monoisotopic mass, maximum number of missed trypsin cleavages given as 2, and cysteines modified with iodoacetamide. A ProFound search with those parameters selected an ≈ 86 kDa protein described as a translational endoplasmic reticulum ATPase from *S. solfataricus* (gi|12311996) with a probability score of 1.1e+00, as the best candidate. Of the thirty-two peptide masses entered into the ProFound program, twenty-four matched the translational endoplasmic reticulum ATPase protein yielding 36% sequence coverage, three of the peptide masses were identified as trypsin autoproteolysis products, and five of the peptide masses remained unaccounted for.

The database search, using the ProFound program, for the identity of the ≈ 67 kDa polypeptide (Figure 6-1) did not yield a match from *S. solfataricus* possessing a high probability score, which may be attributed to the fact that a majority of the peptide masses (22 out of 30) determined using MALDI-MS were assigned as trypsin autolyses products. In the search the following parameters were used: taxonomic category of Archaebacteria, molecular weight between 50 – 80 kDa, peptide mass tolerance ± 2 Da, monoisotopic mass, maximum number of missed trypsin cleavages given as 2, and cysteines modified with iodoacetamide.

The database search, using the ProFound program, for the identity of the ≈ 48 kDa polypeptide (Figure 6-1) gave the results shown in Figure 6-5 and Figure 6-6. In the search the following parameters were used: taxonomic category of *Archaebacteria*,

gi|12311996|emb|CAC22860.1| (AL512963) TRANSITIONAL ENDOPLASMIC RETICULUM ATPASE [Sulfolobus solfataricus]

Sample ID : **SDS-PAGE 85 kDa**
Digestion chemistry : Trypsin
Number of measured peptides : 32
Number of matched peptides : 24
Coverage of protein sequence: 36 %



Measured Mass(M)	Avg/ Mono	Computed Mass	Error (Da)	Residues From	To	Missed Cut	Cys	Met
833.55	M	833.45	0.10	422	427	1	0	0
833.55	M	833.39	0.16	550	556	0	0	0
841.60	M	841.54	0.06	649	655	1	0	0
841.60	M	841.51	0.09	31	37	1	0	0
855.68	M	856.49	-0.81	33	39	1	0	0
855.68	M	857.52	-1.84	743	750	0	0	0
855.68	M	857.48	-1.80	38	44	2	0	0
860.51	M	859.44	1.07	111	118	0	0	0
876.49	M	877.48	-0.99	90	96	1	0	0
892.76	M	894.34	-1.58	713	719	0	1	0
926.78	M	926.42	0.36	498	504	0	0	1
939.71	M	941.57	-1.86	338	346	0	0	0
939.71	M	941.57	-1.86	614	622	0	0	0
967.67	M	967.50	0.17	48	55	0	0	0
971.93	M	970.51	1.42	490	497	0	0	0
974.72	M	974.42	0.30	700	706	0	1	0
974.72	M	976.54	-1.82	690	697	1	0	1
1014.87	M	1015.54	-0.67	110	118	1	0	0
1014.87	M	1014.44	0.43	275	282	0	0	0
1107.15	M	1108.60	-1.45	347	356	1	0	0
1122.10	M	1122.65	-0.55	375	383	1	0	0
1122.10	M	1121.65	0.45	623	632	0	0	0
1125.73	M	1126.68	-0.95	336	346	1	0	0
1125.73	M	1125.67	0.06	31	39	2	0	0

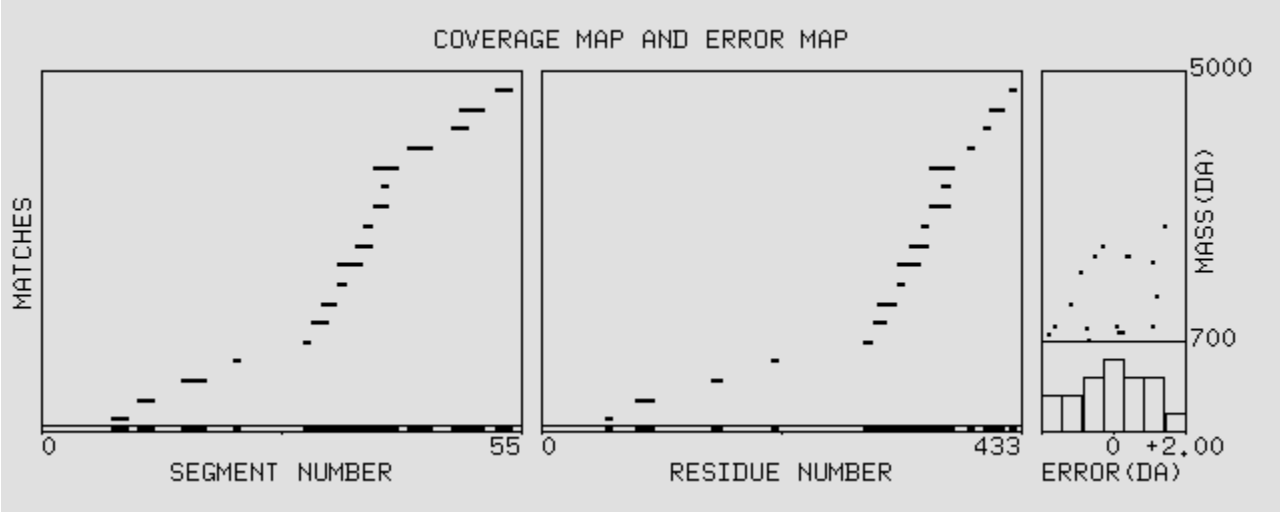
1170.46	M	1169.63	0.83	639	648	0	0	0
1202.94	M	1204.55	-1.61	586	596	0	0	0
1282.03	M	1281.69	0.34	186	197	1	0	0
1282.03	M	1283.71	-1.68	33	43	2	0	0
1282.03	M	1283.63	-1.60	275	284	1	0	0
1392.87	M	1393.77	-0.90	363	374	2	0	0
1405.39	M	1404.74	0.65	529	542	0	0	0
1424.72	M	1424.69	0.03	19	30	1	0	2
1424.72	M	1422.72	2.00	720	731	1	1	1
1579.41	M	1580.79	-1.38	18	30	2	0	2
1666.77	M	1667.83	-1.06	443	456	1	0	2
2239.55	M	2238.19	1.36	119	138	0	0	0

Figure 6-4. ProFound Identified the ≈ 85 kDa Polypeptide as a Translational Endoplasmic Reticulum ATPase From *Sulfolobus solfataricus*.

Shown is a peptide coverage map of the translational endoplasmic reticulum ATPase from *S. solfataricus* (gi|12311996), and the matching peptide masses determined using MALDI-MS.

gi|6015942|emb|CAB57769.1| (Y18930) serine hydroxymethyltransferase [Sulfolobus solfataricus]
gi|12312083|emb|CAC22946.1| (AL512964) serine hydroxymethyltransferase [Sulfolobus solfataricus]

Sample ID : **SDS-PAGE 48 kDa**
Digestion chemistry : Trypsin
Number of measured peptides : 37
Number of matched peptides : 14
Coverage of protein sequence: 37 %



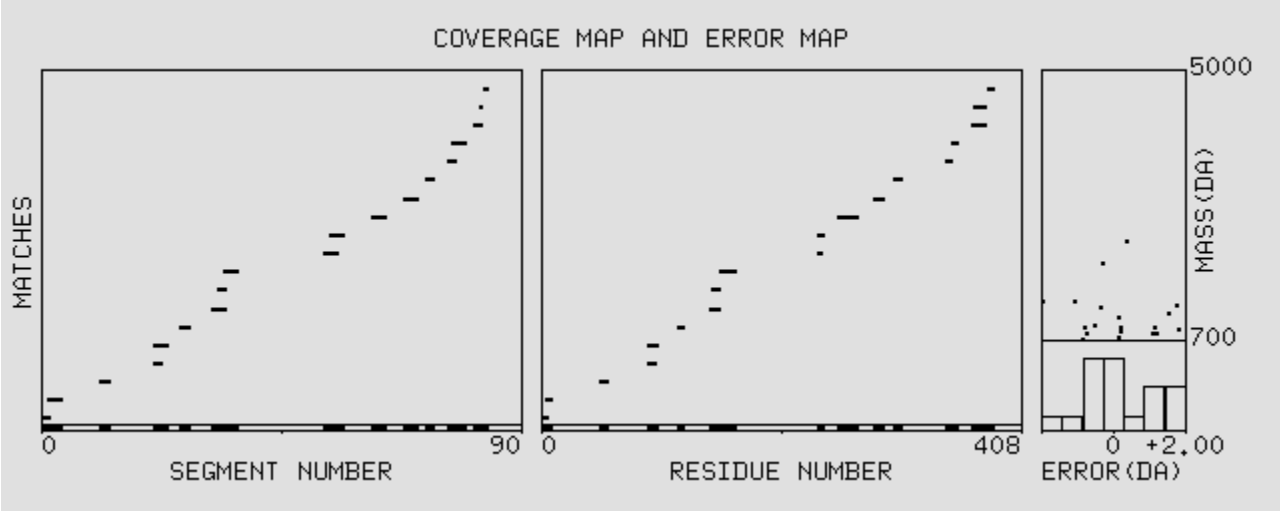
Measured Mass(M)	Avg/ Mono	Computed Mass	Error (Da)	Residues From	To	Missed Cut	Cys	Met
705.66	M	706.34	-0.68	323	329	0	0	0
809.63	M	811.44	-1.81	424	430	1	0	0
842.81	M	842.56	0.25	386	392	2	0	0
842.81	M	842.47	0.34	209	215	0	0	0
861.66	M	861.46	0.20	344	350	0	0	0
913.77	M	914.46	-0.69	59	65	1	0	0
932.64	M	932.48	0.16	362	369	0	0	1
932.64	M	931.49	1.15	400	406	1	0	1
940.87	M	942.51	-1.64	292	300	0	0	0
1283.55	M	1284.75	-1.20	155	165	2	0	1
1426.08	M	1424.77	1.31	301	313	1	0	0
1804.01	M	1804.88	-0.87	406	419	2	0	1
1964.14	M	1962.96	1.18	86	103	1	1	0
2084.56	M	2085.02	-0.46	351	369	1	0	1
2084.56	M	2084.10	0.46	304	322	1	0	0
2084.56	M	2084.15	0.41	333	350	1	0	0
2241.01	M	2241.25	-0.24	323	343	2	0	0
2565.76	M	2564.24	1.52	351	373	2	0	2

Figure 6-5. ProFound Identified the ≈ 48 kDa Polypeptide as a Serine Hydroxymethyltransferase From *Sulfolobus solfataricus*.

Shown is a peptide coverage map of the serine hydroxymethyltransferase (gi|6015942|) from *S. solfataricus* and the matching peptide masses determined using MALDI-MS.

gi|12312957|emb|CAC23674.1| (AL512974) hypothetical protein [Sulfolobus solfataricus]

Sample ID : **SDS-PAGE 48 kDa**
Digestion chemistry : Trypsin
Number of measured peptides : 37
Number of matched peptides : 17
Coverage of protein sequence: 34 %



Measured Mass(M)	Avg/ Mono	Computed Mass	Error (Da)	Residues From	To	Missed Cut	Cys	Met
727.61	M	728.44	-0.83	235	239	2	0	0
762.59	M	762.41	0.18	1	6	1	0	1
804.67	M	803.49	1.18	117	123	1	0	0
809.63	M	808.37	1.26	380	386	0	0	0
813.79	M	814.48	-0.69	236	241	2	0	0
842.81	M	842.57	0.24	4	10	2	0	0
877.36	M	875.45	1.91	344	351	1	1	0
913.77	M	913.54	0.23	300	307	1	0	0
921.74	M	920.51	1.23	146	153	1	0	0
921.74	M	922.49	-0.75	50	57	1	0	0
938.99	M	939.50	-0.51	350	356	2	0	0
1065.74	M	1065.55	0.19	91	99	1	0	0
1127.22	M	1125.60	1.62	368	379	0	0	0
1221.35	M	1221.65	-0.30	91	100	2	0	0
1283.55	M	1281.70	1.85	367	379	1	0	0
1317.71	M	1319.71	-2.00	143	153	2	0	0
1317.71	M	1318.75	-1.04	283	292	2	0	0
1942.89	M	1943.13	-0.24	152	167	2	0	0
2312.74	M	2312.31	0.43	252	271	2	0	0

Figure 6-6. ProFound Identified the ≈ 48 kDa Polypeptide as an ≈ 48 kDa Hypothetical Protein From *Sulfolobus solfataricus*.

Shown is a peptide coverage map of the ≈ 48 kDa hypothetical protein (gi|12312957|) from *S. solfataricus* and the matching peptide masses determined using MALDI-MS.

molecular weight between 35 - 65 kDa, peptide mass tolerance ± 2 Da, monoisotopic mass, maximum number of missed trypsin cleavages given as 2, and cysteines modified with iodoacetamide. Of the sixty-one peptide masses entered into the ProFound program fourteen matched an ≈ 49 kDa serine hydroxymethyltransferase (gi|6015942|), covering 37% of the protein, and having a probability score of 8.8e-04; while seventeen matched a ≈ 48 kDa hypothetical protein (gi|12312957|), covering 34% of the protein, and having a probability score of 5.2e-04. This suggests that both polypeptides were resolved together during the SDS-PAGE. Of the remaining thirty peptide masses, ten were assigned as trypsin autolyses product, while the remaining twenty peptide masses remained unaccounted for.

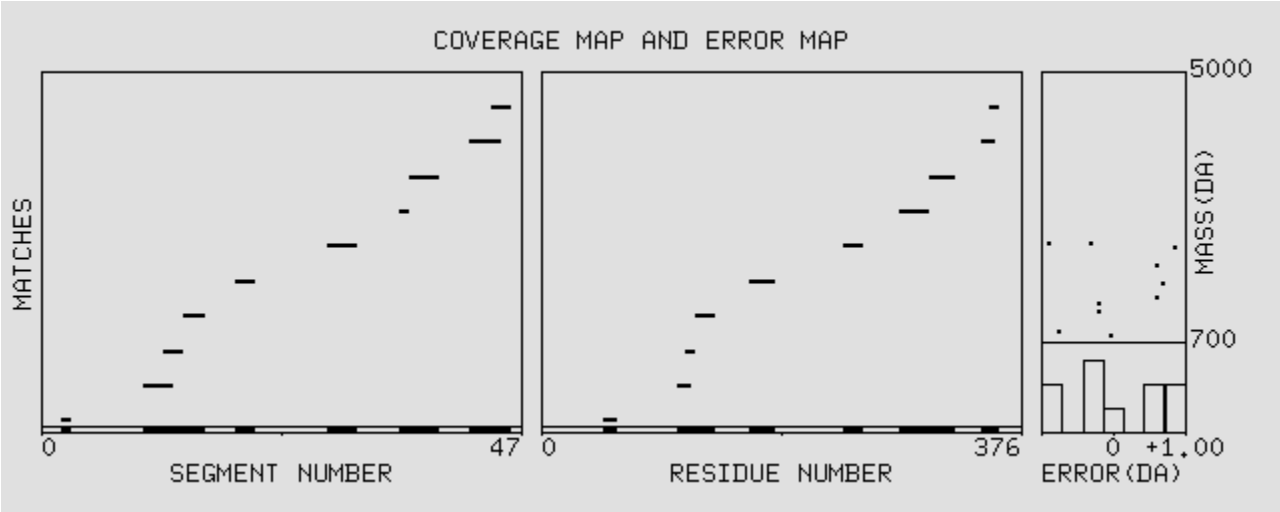
The database search, using the ProFound program, for the identity of the ≈ 45 kDa polypeptide (Figure 6-1) gave the results shown in Figure 6-7. In the search the following parameters were used: taxonomic category of *Archaebacteria*, molecular weight between 35 – 55 kDa, peptide mass tolerance ± 1 Da, monoisotopic mass, maximum number of missed trypsin cleavages given as 2, and cysteines modified with iodoacetamide. A ProFound search with those parameters selected an ≈ 42 kDa hypothetical protein from *S. solfataricus* (gi|7484180) with a probability score of 9.4e-01, as the best candidate. Of the twenty-eight peptide masses entered into the ProFound program, ten matched the hypothetical protein yielding 36% sequence coverage, nine of the peptide masses were identified as trypsin autoproteolysis products, and nine of the peptide masses remained unaccounted for.

ProFound Search Results of the [32P] Phosphorylated Proteins Resolved Using Two-Dimensional Electrophoresis

The database search, using the ProFound program, for the identity of the ≈ 105 kDa, ≈ 6.1 pI polypeptide (Figure 6-2) yielded no significant search results from *S. solfataricus*. However, ProFound did match the peptide masses (7 % total protein coverage) to a large helicase related protein (101 kDa, 6.5 pI) from the archaeon *Thermoplasma acidophilum* (gi|10640115|). Therefore, we presume that this polypeptide was the same as the one observed at ≈ 105 kDa in the SDS-PAGE gel (Figure 6-1) that was identified as an ATP-dependent helicase. One possible explanation as to why no

gi|7484180|pir||S75438 hypothetical protein c04025 - Sulfolobus solfataricus
gi|12311942|emb|CAC22807.1| (AL512962) 5S rRNA [Sulfolobus solfataricus]
gi|1707797|emb|CAA69554.1| (Y08257) orf c04025 [Sulfolobus solfataricus]

```
Sample ID                    : SDS-PAGE 45 kDa
Digestion chemistry          : Trypsin
Number of measured peptides  : 28
Number of matched peptides   : 10
Coverage of protein sequence: 36 %
```



Measured Mass(M)	Avg/ Mono	Computed Mass	Error (Da)	Residues From	To	Missed Cut	Cys	Met
803.40	M	803.41	-0.01	114	120	1	0	0
870.78	M	871.55	-0.77	352	358	1	0	0
1203.43	M	1203.62	-0.19	50	60	0	0	1
1316.53	M	1316.72	-0.19	108	118	2	0	1
1424.54	M	1423.91	0.63	345	356	2	0	0
1668.69	M	1667.97	0.72	238	253	2	0	1
1941.67	M	1941.04	0.63	121	137	1	0	1
2240.22	M	2239.33	0.89	305	324	2	0	0
2298.15	M	2299.06	-0.91	281	304	0	0	1
2313.09	M	2313.38	-0.29	164	183	1	0	0

Figure 6-7. ProFound Identified the ≈ 45 kDa Polypeptide as a ≈ 42 kDa Hypothetical Protein From *S. solfataricus*.

Shown is a peptide coverage map of the ≈ 42 kDa hypothetical protein (gi|7484180|) from *S. solfataricus* and the matching peptide masses determined using MALDI-MS.

significant matches were found in *S. solfataricus* may be attributed to the fact that a majority of the peptide masses (18 out of 24) determined using MALDI-MS were assigned as trypsin autolyses products. In the search, the following parameters were used: taxonomic category of *Archaea*, molecular weight between 95 – 120 kDa, pI between 5.1 – 7.1, peptide mass tolerance ± 1 Da, monoisotopic mass, maximum number of missed trypsin cleavages given as 2, and cysteines modified with iodoacetamide.

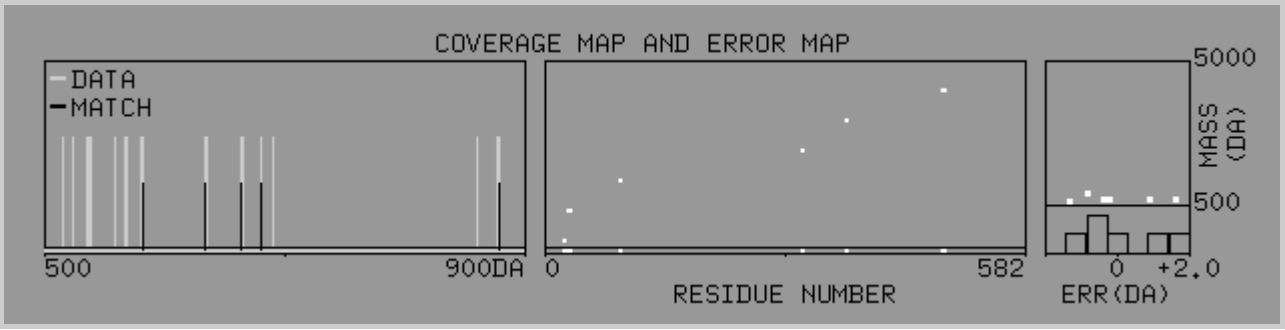
The database search, using the ProFound program, for the identity of the ≈ 67 kDa, ≈ 5.7 pI polypeptide (Figure 6-2) gave the results shown in Figure 6-8. In the search the following parameters were used: taxonomic category of *Archaea*, molecular weight between 57 – 77 kDa, pI between 5.1 – 6.2, peptide mass tolerance ± 2 Da, monoisotopic mass, maximum number of missed trypsin cleavages given as 2, and cysteines modified with iodoacetamide. A ProFound search with those parameters selected an ≈ 67 kDa, ≈ 5.9 pI hypothetical protein from *S. solfataricus* (gi|12312047|) with a probability score of 1.0+e000 as the best candidate. Of the eighteen peptide masses entered into the ProFound program, five matched the hypothetical protein yielding 5% sequence coverage, nine of the peptide masses were identified as trypsin autoproteolysis products, and five of the peptide masses remained unaccounted.

Interestingly, while the probability score and sequence coverage admittedly were low, the primary sequence of this hypothetical protein (encoding by *ORF sso0469*) exhibited noticeable similarities to eukaryotic protein kinases as well as the *S. solfataricus* protein kinase described in Chapter III. The hypothetical protein contains eight of the twelve highly conserved active-site subdomains including I, II, III, V, VI_b, VIII, IX, and XI (Figure 6-9), as well as a leucine zipper pattern spanning residues 537-558 (such patterns are present in many eukaryotic gene regulatory proteins and function to facilitate dimerization between polypeptides), two hypothetical N-glycosylation sites spanning residues 161 – 164 and 371 – 374, and two hypothetical transmembrane domains as identified using DAS (Figure 6-10).

The database search, using the ProFound program, for the identity of the ≈ 68 kDa, ≈ 5.6 pI polypeptide (Figure 6-2) gave the results shown in Figure 6-11. In the search the following parameters were used: taxonomic category of *Archaea*, molecular weight between 60 – 76 kDa, pI between 5.1 – 6.1, peptide mass tolerance ± 1 Da,

gi|12312047|emb|CAC22911.1| (AL512963) hypothetical [Sulfolobus solfataricus]

Sample ID : **67 kDa, 5.7 pI** [Pass:1]
Measured peptides : 19
Matched peptides : 5
Min. sequence coverage: 5%



Measured Mass(M)	Avg/ Mono	Computed Mass	Error (Da)	Residues Start	To	Missed Cut	Peptide sequence
581.032	M	582.312	-1.280	365	369	0	DAHLK
632.962	M	633.315	-0.353	91	95	0	NEIMK
632.962	M	631.317	1.645	312	316	0	QIEDK
662.122	M	662.298	-0.176	22	27	0	QTSDGR
680.262	M	679.365	0.897	28	33	0	GTISFR
877.752	M	878.556	-0.804	481	487	2	LLRHGRK

Unmatched Monoisotopic Masses:

537.780 538.110 559.750 567.800 568.140 581.700 690.340 860.690 877.080

Figure 6-8. ProFound Identified the ≈ 67 kDa, ≈ 5.7 pI Polypeptide as an ≈ 67 kDa, ≈ 5.9 pI Hypothetical Protein From *S. solfataricus*.

Shown is a peptide coverage map of the 67 kDa, ≈ 5.9 pI hypothetical protein (gi|12312047|) from *S. solfataricus* and the matching peptide masses determined using MALDI-MS.

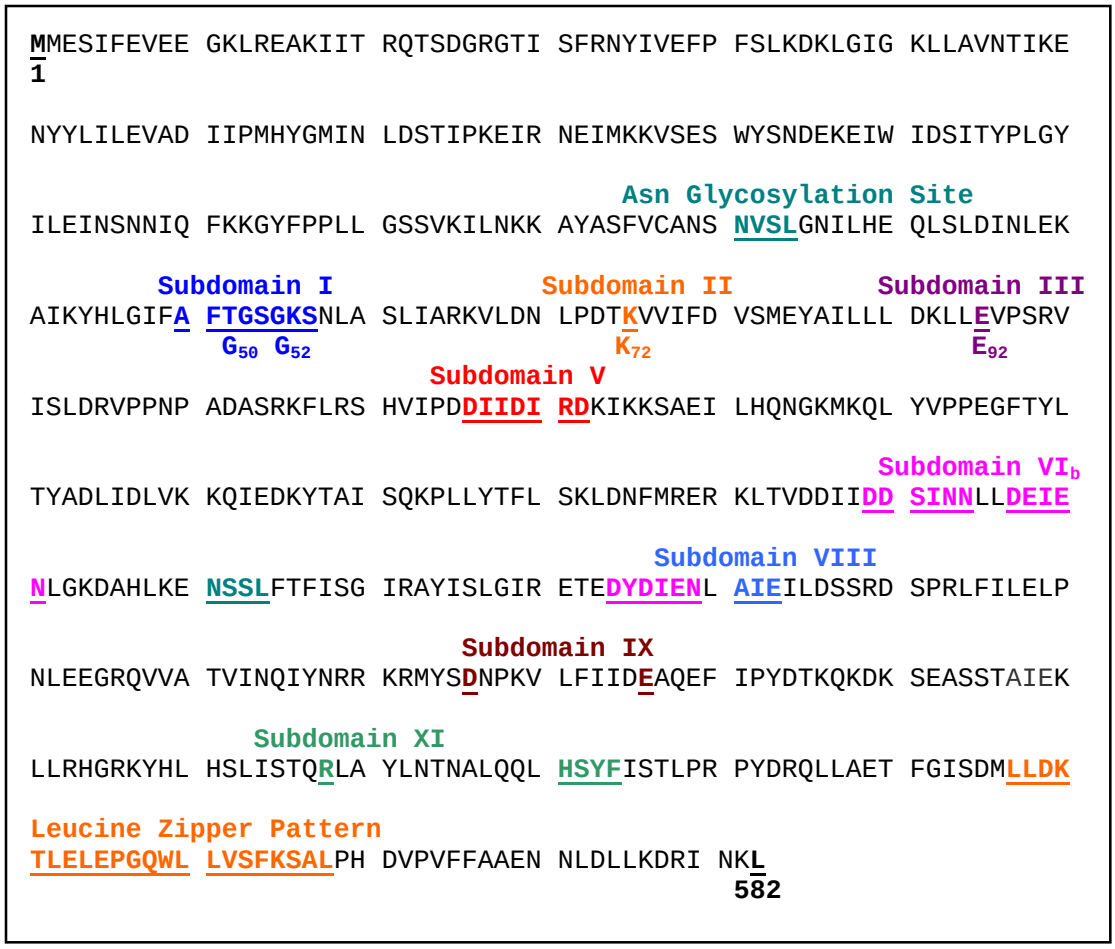
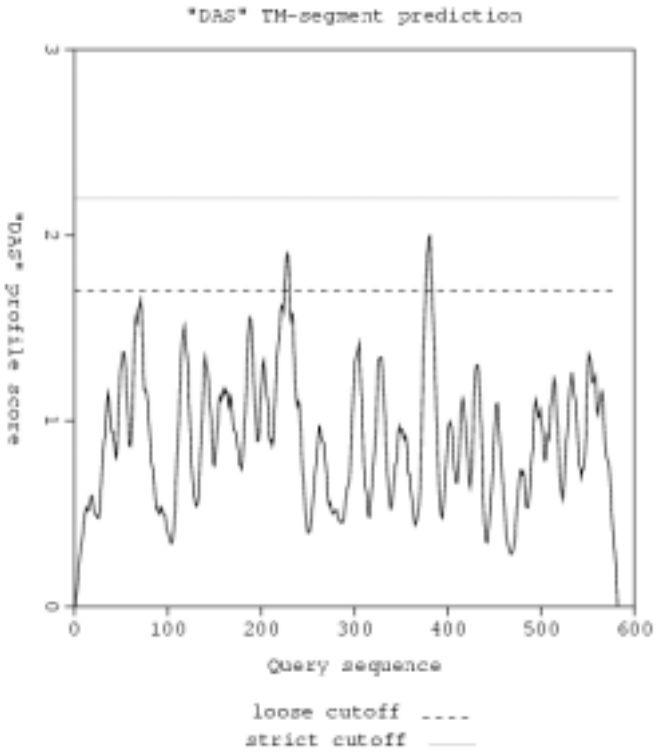


Figure 6-9. The Amino Acid Sequence of the ≈ 67 kDa, pI ≈ 5.7 Hypothetical Protein Encoded by *Sulfolobus solfataricus ORF sso0469* Exhibits Similarity to Eukaryotic Protein Kinases.

Inspection of the partial genome of *S. solfataricus* for homologs of eukaryotic protein kinases, as described in the text, revealed the presence of *ORF sso0469* whose predicted product displays sequence features characteristic of eukaryotic protein-serine/threonine/tyrosine kinases. *S. solfataricus ORF sso0469* encodes a hypothetical protein consisting of 582 amino acids, with a MW of ≈ 67 kDa, and pI of ≈ 5.7. Shown is its deduced amino acid sequence and prospective candidates for conserved subdomains or essential residues characteristic of eukaryotic protein kinases (described in Introduction), as well as hypothetical Asn glycosylation sites and a leucine zipper pattern.



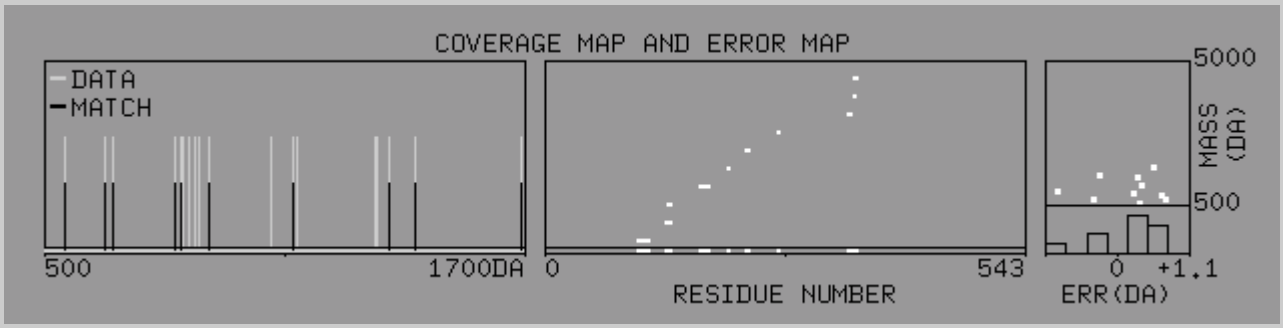
Start	Stop	Length	Cutoff
225	230	6	1.7
377	383	7	1.7

Figure 6-10. Potential Transmembrane Helices of the ≈ 67 kDa, pI ≈ 5.7 Hypothetical Protein Encoded by *Sulfolobus solfataricus ORF sso0469*.

Hypothetical transmembrane domains were identified using the web-based DAS (Dense Alignment Surface Method) Transmembrane Prediction Server (http://www.sbc.su.se/~miklos/DAS/maindas.html) from the Stockholm Bioinformatic Center in Stockholm, Sweden. This algorithm uses low-stringency dot-plots of the query sequence against a collection of non-homologous membrane proteins using a previously derived, special scoring matrix to identify putative transmembrane helices contained within a protein's primary sequence. Shown is the DAS curve for the hypothetical protein encoded by *S.* solfataricus *ORF sso0469* (DAS profile score as a function of the amino acid sequence). A DAS profile score of 1.7 characterizes a loose cutoff and a DAS profile score of 2.2 characterizes a strict cutoff for defining putative transmembrane domains. As is seen in the graph two hypothetical transmembrane domains are defined using DAS.

gi|13816647|gb|AAK43306.1| (AE006909) Acetolactate synthase large subunit homolog (ilvB-6) [Sulfolobus solfataricus]

Sample ID : **2-D, 68 kDa, 5.6 pI** [Pass:0]
Measured peptides : 18
Matched peptides : 10
Min. sequence coverage: 12%



Measured Mass(M)	Avg/ Mono	Computed Mass	Error (Da)	Residues Start	To	Missed Cut	Peptide sequence
549.632	M	549.291	0.341	264	267	0	FDLR
649.082	M	648.315	0.767	207	211	0	EMLEK
672.972	M	673.339	-0.367	349	353	0	QEIER
824.112	M	823.407	0.705	343	348	1	QEEYKK
841.672	M	841.397	0.275	140	145	0	YDFEIR
913.592	M	914.518	-0.926	349	355	1	QEIERLK
1121.942	M	1121.561	0.381	227	234	1	EWFESLRR
1359.992	M	1359.682	0.310	136	145	1	QYVKYDFEIR
1428.462	M	1428.731	-0.269	176	187	1	EVSIQEVNEARR
1693.432	M	1692.854	0.578	106	120	2	SPYTEKGNTASRNLR

Unmatched Monoisotopic Masses:

845.210 861.240 877.220 885.480 1067.050 1131.650 1327.990 1331.950

Figure 6-11. ProFound Identified the ≈ 68 kDa, ≈ 5.6 pI Polypeptide as an Acetolactate Synthase Large Subunit Homolog (ilvB-6) From *S. solfataricus*.

Shown is a peptide coverage map of the acetolactate synthase large subunit homolog (ilvB-6) (gi|13816647|) from *S. solfataricus* and the matching peptide masses determined using MALDI-MS.

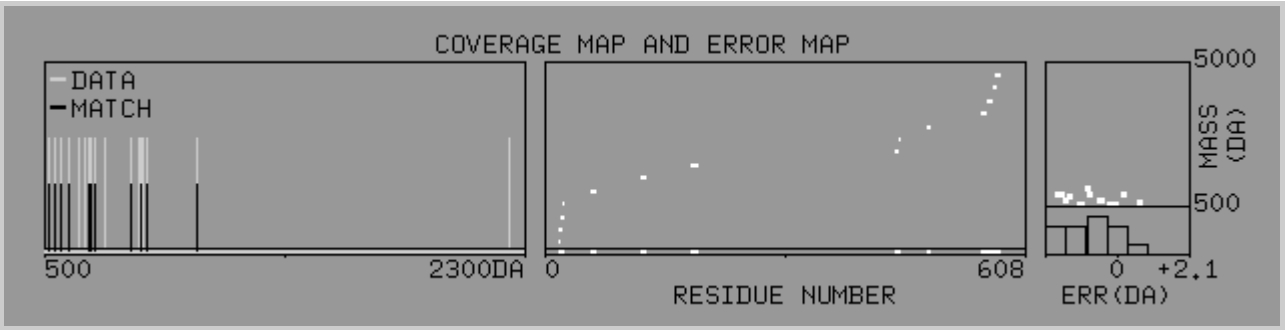
monoisotopic mass, maximum number of missed trypsin cleavages given as 2, and cysteines modified with iodoacetamide. A ProFound search with those parameters selected an acetolactate synthase large subunit homolog (*ilvB-6*, ≈ 62 kDa, ≈ 6.0 pI) (gi|13816647|) from *S. solfataricus* with a probability score of 1.0e+000 as the best candidate. Of the eighteen peptide masses entered into the ProFound program, ten matched the hypothetical protein yielding 12% sequence coverage, five of the peptide masses were identified as trypsin autoproteolysis products, and three of the peptide masses remained unaccounted.

The database search, using the ProFound program, for the identity of the ≈ 69 kDa, ≈ 5.5 pI polypeptide (Figure 6-2) gave the results shown in Figure 6-12. In the search the following parameters were used: taxonomic category of *Archaea*, molecular weight between 60 – 77 kDa, pI between 4.9 – 6.1, peptide mass tolerance ± 2 Da, monoisotopic mass, maximum number of missed trypsin cleavages given as 2, and cysteines modified with iodoacetamide. A ProFound search with those parameters selected an Acetyl-CoA synthetase (acetate-CoA ligase) (*acsA-7,* ≈ 69 kDa, ≈ 5.7 pI) (gi|13815354|) from *S. solfataricus* with a probability of 9.7e-001 as the best candidate. Of the nineteen peptide masses entered into the ProFound program, eleven matched the hypothetical protein yielding 9% sequence coverage, four of the peptide masses were identified as trypsin autoproteolysis products, and four of the peptide masses remained unaccounted.

The database search, using the ProFound program, for the identity of the ≈ 70 kDa, ≈ 5.4 pI polypeptide (Figure 6-2) gave the results shown in Figure 6-13. In the search the following parameters were used: taxonomic category of *Archaea*, molecular weight between 60 – 77 kDa, pI between 4.9 – 6.1, peptide mass tolerance ± 2 Da, monoisotopic mass, maximum number of missed trypsin cleavages given as 2, and cysteines modified with iodoacetamide. A ProFound search with those parameters selected a ≈ 72 kDa, 5.5 pI hypothetical protein (gi|13815259|) from *S. solfataricus* with a probability of 1.0e-000 as the best candidate. Of the fifteen peptide masses entered into the ProFound program, nine matched the hypothetical protein yielding 17% sequence coverage, two of the peptide masses were identified as trypsin autoproteolysis products, and four of the peptide masses remained unaccounted.

gi|13815354|gb|AAK42252.1| (AE006813) Acetyl-CoA synthetase (acetate-CoA ligase) (acsA-7) [Sulfolobus solfataricus]

Sample ID : **2-D, 69 kDa, 5.5 pI** [Pass:0]
Measured peptides : 19
Matched peptides : 11
Min. sequence coverage: 9%



Measured Mass(M)	Avg/ Mono	Computed Mass	Error (Da)	Residues Start	Residues To	Missed Cut	Peptide sequence
514.232	M	514.250	-0.018	449	452	0	NPER
514.232	M	515.281	-1.049	23	26	0	NLNR
537.102	M	537.298	-0.196	18	21	0	FMIK
563.212	M	564.281	-1.069	445	448	0	GFWR
589.022	M	588.334	0.688	569	573	1	TRNAK
670.892	M	671.383	-0.490	22	26	1	RNLNR
675.782	M	677.323	-1.541	485	490	0	SDDTIK
692.932	M	693.399	-0.467	18	22	1	FMIKR
824.112	M	825.532	-1.420	562	568	1	IVKELPK
859.822	M	861.459	-1.637	554	561	0	AFAPSEIK
885.732	M	886.523	-0.791	59	66	1	VLDLSRGK
885.732	M	885.539	0.193	122	129	2	IGLKKGDR
885.732	M	887.512	-1.780	571	577	2	NAKIMRR
1073.752	M	1074.585	-0.833	187	195	2	RGKEVDMLK

Unmatched Monoisotopic Masses:

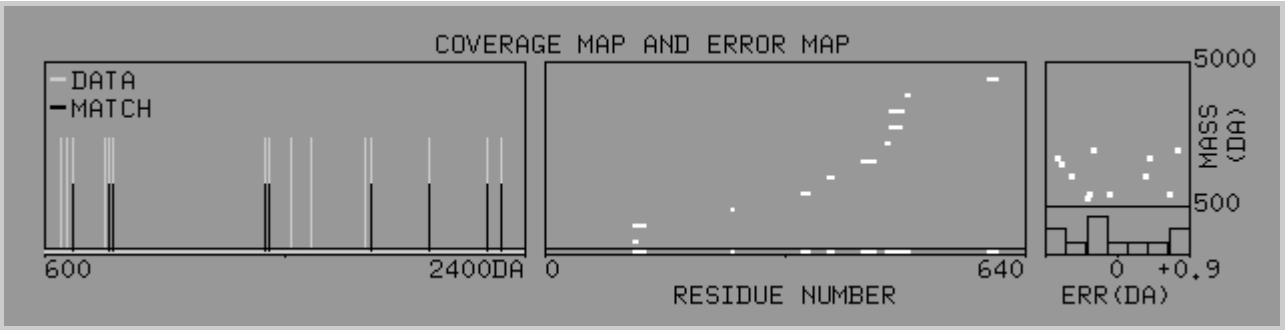
690.960 727.860 854.890 870.810

Figure 6-12. ProFound Identified the ≈ 69 kDa, ≈ 5.5 pI Polypeptide as an Acetyl-CoA synthetase (acetate-CoA ligase) (*acsA-7*) From *S. solfataricus*.

Shown is a peptide coverage map of the Acetyl-CoA synthetase (acetate-CoA ligase) (*acsA-7*) (gi|13815354|) from *S. solfataricus* and the matching peptide masses determined using MALDI-MS.

gi|13815259|gb|AAK42173.1| (AE006805) Hypothetical protein [Sulfolobus solfataricus]

```
Sample ID :  2-D, 70 kDa, 5.4 pI [Pass:0]
Measured peptides     : 15
Matched peptides      : 9
Min. sequence coverage: 17%
```



Measured Mass(M)	Avg/ Mono	Computed Mass	Error (Da)	Residues Start	Residues To	Missed Cut	Peptide sequence
704.022	M	704.389	-0.366	249	254	1	GLMKEK
841.392	M	841.490	-0.098	118	125	0	VASPDLIK
860.172	M	860.496	-0.324	455	462	1	LSSELKGK
860.172	M	859.528	0.644	481	487	1	KYLPALR
1424.182	M	1424.741	-0.559	378	388	2	QLYEEAFNKKR
1442.032	M	1441.672	0.360	343	354	0	EIGDYEDYLVAR
1824.282	M	1824.952	-0.669	591	606	1	ELAKAHALYGAIYYSR
2042.442	M	2042.035	0.407	461	478	1	GKLSVNPEELIYAFGYDK
2042.442	M	2043.172	-0.730	118	137	1	VASPDLIKSVLGALEPLHGK
2261.022	M	2261.300	-0.278	423	442	1	LLEGHLWVLNAVRSVSVLTR
2311.962	M	2311.220	0.742	461	480	2	GKLSVNPEELIYAFGYDKLR

Unmatched Monoisotopic Masses:

825.050 1526.560 1604.660 1803.410

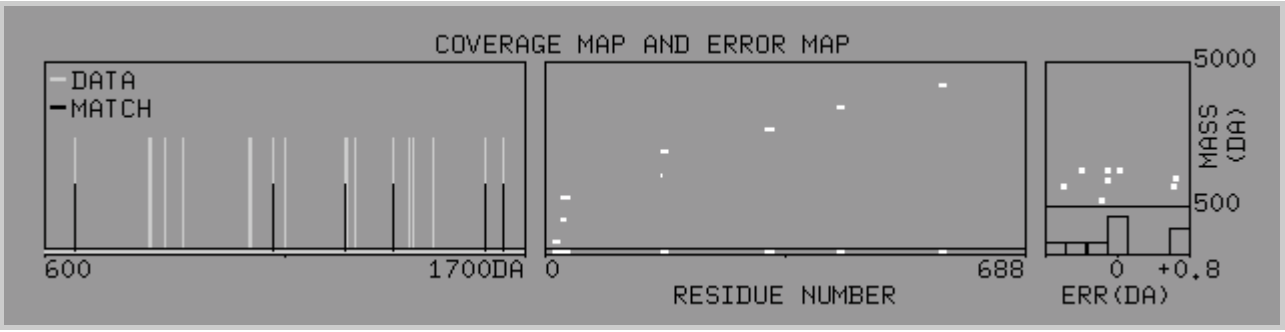
Figure 6-13. ProFound Identified the ≈ 70 kDa, ≈ 5.4 pI Polypeptide as an ≈ 72 kDa, 5.5 pI hypothetical protein From *S. solfataricus*.

Shown is a peptide coverage map of the ≈ 72 kDa, 5.5 pI hypothetical protein (gi|13815259|) from *S. solfataricus* and the matching peptide masses determined using MALDI-MS.

The database search, using the ProFound program, for the identity of the ≈ 71 kDa, ≈ 5.3 pI polypeptide (Figure 6-2) gave the results shown in Figure 6-14. In the search the following parameters were used: taxonomic category of *Archaea*, molecular weight between 61 – 81 kDa, pI between 4.7 – 5.9, peptide mass tolerance ± 1 Da, monoisotopic mass, maximum number of missed trypsin cleavages given as 2, and cysteines modified with iodoacetamide. A ProFound search with those parameters selected an carbon monoxide dehydrogenase, large chain (*cutA-1*, ≈ 75 kDa, ≈ 5.4 pI) (gi|13814405|) from *S. solfataricus* with a probability of 9.9e-001as the best candidate. Of the twenty peptide masses entered into the ProFound program, eight matched the hypothetical protein yielding 11% sequence coverage, six of the peptide masses were identified as trypsin autoproteolysis products, and six of the peptide masses remained unaccounted.

gi|13814405|gb|AAK41457.1| (AE006736) Carbon monoxide dehydrogenase, large chain (cutA-1) [Sulfolobus solfataricus]

Sample ID : **2-D, 71 kDa, 5.3 pI** [Pass:0]
Measured peptides : 20
Matched peptides : 8
Min. sequence coverage: 11%



Measured Mass(M)	Avg/ Mono	Computed Mass	Error (Da)	Residues Start	To	Missed Cut	Peptide sequence
671.252	M	671.407	-0.155	166	170	1	LKQQR
1121.152	M	1120.539	0.613	13	22	0	GSYIDDINPK
1125.102	M	1125.666	-0.563	23	31	0	NVVYLHIIR
1289.642	M	1289.734	-0.092	419	430	0	GVSIVTFAEIVR
1398.332	M	1397.682	0.650	566	577	0	VLEYYAVDDVGR
1610.482	M	1610.860	-0.378	316	329	2	LSNGPYKMKFASIR
1649.872	M	1649.972	-0.100	23	36	1	NVVYLHIIRSPIAR
1652.942	M	1652.902	0.040	166	179	2	LKQQRVVSNPMEPK

Unmatched Monoisotopic Masses:

919.590 1152.040 1294.720 1310.700 1443.750 1492.140

Figure 6-14. ProFound Identified the ≈ 71 kDa, ≈ 5.3 pI Polypeptide as a Carbon Monoxide Dehydrogenase, Large Chain (cutA-1) From *S. solfataricus*.

Shown is a peptide coverage map of the carbon monoxide dehydrogenase, large chain (cutA-1) (gi|13814405|) from *S. solfataricus* and the matching peptide masses determined using MALDI-MS.

CHAPTER VII

Discussion

Fossils of *Archaea* have been discovered dating as far back as 3.5 billion years (Rasmussen, 2000; Gogarten-Boekels et. al., 1999). Thus, the members of this domain inhabited the earth a mere one billion years after its formation, implicating them as the forerunners of all contemporary life on this planet including humans. The overall goal of the research described herein was to examine the nature of protein O-phosphorylation in these unique organisms. Through the study of protein phosphorylation-dephosphorylation in these "protoeukaryotes," its is our intent to gain insight into the form and function of the first phosphorylation networks and to trace their evolution into the many-fold more complex phosphorylation labyrinths observed in higher eukaryotes. To this end, a threonine-specific protein kinase activity was purified and characterized from the membrane fraction of the archaeon *Sulfolobus solfataricus*, and its genome searched for *ORFs* encoding eukaryotic-like protein kinases. One of the latter, *ORF sso2387*, was cloned and expressed in *E. coli* and its encoded protein, SsPK2, was demonstrated to possess protein kinase activity that preferentially phosphorylated serine residues *in vitro*. SsPK2 represents the first bona fide "eukaryotic-like" protein kinase from the *Archaea*. Finally, proteins from the membrane fraction of *S. solfataricus* subject to phosphorylation *in vitro* on serine or threonine residues were identified using MALDI-MS / peptide fingerprinting techniques.

Identification, Purification, and Characterization of a Membrane-Associated Protein Kinase From *Sulfolobus solfataricus*.

Identification and Solubilization of a Protein Kinase Activity in Sulfolobus solfataricus

The presence of a membrane-associated protein kinase activity in the archaeon *S. solfataricus* was first detected when a detergent extract of the membrane fraction was subjected to SDS-PAGE in gels containing copolymerized casein. Following a denaturation-renaturation cycle with guanidine hydrochloride and renaturation buffer,

protein kinase activity was assayed *in situ* by incubating the gel with [γ-^{32}P] ATP and divalent metal cofactors Mn^{2+} and Mg^{2+}. After the radiolabeled ATP was washed away, autoradiography revealed the presence of a discrete band at ≈ 67 kDa where [^{32}P] phosphate had been retained within the gel. Subsequent phosphoamino acid analysis revealed that the phosphate was bound as phosphothreonine, establishing that a covalent protein-phosphate bond had been formed and that the activity detected displayed the essential characteristics of a protein kinase. Efforts to solubilize the protein kinase by alternative means such as extraction with mildly acidic buffers and/or high salt concentrations proved unsuccessful. The enzyme activity could only be extracted using nonionic detergents such as Triton X-100 and octyl glucoside. Following solubilization the presence of detergent was required to preserve catalytic activity. Therefore, we concluded that the protein kinase was an integral membrane protein.

Gel filtration chromatography of the DE-52 fraction showed the protein kinase activity eluting with an apparent molecular weight of ≈ 125 kDa, suggesting that the enzyme exists as a homodimer or some other heteroligomer. The potential physiological significance of the ≈ 125 kDa oligomer must be viewed with caution, as the protein kinase activity present in crude detergent soluble membrane extracts eluted in the void volume (> 250 kDa) of the S-200 gel filtration column. Nonetheless, the ability to recover protein kinase activity following SDS-PAGE or two-dimensional electrophoresis indicates that if the enzyme does exist as an oligomer, the essential determinants for its catalytic activity reside within a single polypeptide subunit of ≈ 67 kDa.

Phosphorylation of Exogenous Proteins and Peptides by the Sulfolobus solfataricus Protein Kinase

The *S. solfataricus* protein kinase proved to be relatively selective when challenged with a number of exogenous protein and peptide substrates. Only casein, RCM lysozyme, histone H4, and MLC peptide (KKRAARA**TS**NVFA) were phosphorylated at appreciable levels *in vitro*. Phosphoamino acid analysis revealed that the protein kinase had a preference for phosphorylating threonine residues *in vitro* as casein, RCM lysozyme, and MLC peptide were phosphorylated exclusively on threonine. Histone H4 was phosphorylated on serine in addition to threonine. While it should be

noted that all these substrates are nonphysiological, it was also observed that autophosphorylation of the *S. solfataricus* protein kinase occurred on threonine residues as well.

This apparent preference for phosphorylating threonine residues *in vitro* was examined further using a variant of the MLC peptide, named T8S peptide (KKRAARA**SS**NVFA), in which the threonine at position eight was changed to a serine. Kinetic experiments revealed a preference on the part of the *S. solfataricus* protein kinase to phosphorylate the MLC peptide over the T8S peptide. While the K_m for the two peptides were similar, the V_{max} for phosphorylation of the threonine residue in the MLC peptide was over 20-fold higher.

The *S. solfataricus* protein kinase appears to be rather unique in its predilection for phosphorylating threonine residues, as few threonine-specific or threonine-preferring protein kinases have been reported in the literature. These remnants include phosphotyrosine picked threonine kinase, PYT (Lindberg et. al., 1993), and a plasma cell membrane glycoprotein, PC-1 (Oda et. al., 1991). The primary sequence of PYT conforms to the cAMP dependent protein kinase superfamily. However, PYT unlike the *S. solfataricus* protein kinase, was shown to be a dual-specific protein kinase, capable of phosphorylating both threonine and tyrosine residues. PC-1, while its primary sequence does not conform to the eukaryotic protein kinase paradigm, was shown to undergo autophosphorylation exclusively on threonine.

Metal Ion Preference of the Sulfolobus solfataricus Protein Kinase

The *S. solfataricus* protein kinase exhibited a preference for Mn^{2+} *in vitro* over the other divalent metal ions tested (Ca^{2+}, Mg^{2+}, Ni^{2+}, and Zn^{2+}). This is fairly infrequent occurrence among protein kinases as most prefer Mg^{2+} as cofactor. This phenomenon was observed previously with serine/threonine phosphatases from the *Archaea* (Kennelly et. al., 1993; Mai et. al., 1998; Solow et. al., 1997), as well as in experiments investigating protein kinase activity in *Sulfolobus acidocaldarius* (Skorko, 1989). However, Mn^{2+} may not serve as the principal cofactor *in vivo* as assays conducted at temperatures well within the physiological range of the organism demonstrated that Mg^{2+} surpassed the effectiveness of Mn^{2+}.

Nucleotide Preference of the Sulfolobus solfataricus Protein Kinase

Another unusual property of the *S. solfataricus* protein kinase was its utilization of a variety of purine nucleotide di- and triphosphoryl donor substrates *in vitro*, including ATP, ADP, GTP, and GDP. The identification of cellular substrates of individual protein kinases, including phosphoryl donor substrate, remains a key challenge in the field protein phosphorylation. While most protein kinases exhibit dependence on ATP as phosphoryl donor substrate, several exceptions to this pattern have been reported. The best-documented example of these is casein kinase II, which can use either ATP or GTP *in vitro* (Jakobi and Traugh, 1992). Other examples include a novel human serine/threonine protein kinase, Mst-3, which interestingly, prefers Mn^{2+} to Mg^{2+} as a divalent cofactor, and can use both GTP and ATP as phosphate donors (Schinkmann and Blenis, 1997) and the HPr kinase from *Bacillus subtilis* which exhibits a similar nucleotide preference (Reizer et. al., 1998). Protein kinases that catalyze the transfer of β-phosphate from ADP to a serine and/or threonine residues have been discovered in *Zea mays* (Ashton et. al., 1984; Budde et. al., 1986), *Hordeum vulgare* (Kim et. al., 1999), and *Chlamydomonas reinhardtii* (Danon and Mayfield, 1994). It has also been reported that pyrophosphate can serve as a source of phosphoryl groups for the phosphorylation of proteins in *E. coli* (Duclos et. al., 1996). When an "in gel" assay was performed, however, using [^{32}P] $NaPP_i$ as phosphoryl donor, no protein kinase activity was detected in either the membrane fraction nor the DE-52 fraction, demonstrating that pyrophosphate did not serve as a phosphoryl donor substrate for the *S. solfataricus* protein kinase. The ability of archaeal enzymes to utilize "unusual" phosphoryl donors has been demonstrated previously. Examples include a phosphofructokinase from *Thermoproteus tenax*, which uses PP_i instead of ATP as the phosphoryl donor (Siebers et. al., 1998), and experiments probing protein kinase activity in *Sulfolobus acidocaldarius*, which demonstrated that polyphosphate could serve as phosphoryl donor *in vitro* (Skorko, 1989).

Kinetic experiments performed at elevated temperatures under which the organism lives, suggest the protein kinase prefers ATP to GTP, as the K_m for ATP was shown to be approximately 5-fold lower, and the V_{max} for ATP was shown to be approximately 2-fold higher than that of GTP. These results were verified when the

ability of the DE-52 fraction to catalyze the transfer of [^{32}P] phosphate from [γ-^{32}P] ATP, [γ-^{32}P] GTP, or [β-^{32}P] GDP was assayed in solution or "in gel," with the degree of phosphoryl transfer exhibiting the order ATP > GTP >> GDP. Since the intracellular concentrations of these nucleotides within *S. solfataricus* are unknown, it is difficult to determine the physiological relevance of these results. Once the identity of the protein kinase is determined and its primary sequence established, it will be interesting to see if those subdomains responsible for binding the nucleotide phosphoryl donor (subdomain I - VI) display any unique features that would distinguish them form previously characterized sequences of eukaryotic protein kinases, which have been shown to only utilize ATP as phosphoryl donor substrate.

Sensitivity of the Sulfolobus solfataricus Protein Kinase to Inhibitors of Eukaryotic Protein Kinases

When the *Sulfolobus solfataricus* protein kinase was challenged with several known inhibitors of eukaryotic protein kinases, only three of the compounds tested, H7 (a specific inhibitor of cAMP-dependent protein kinases), Genistein (an inhibitor of protein-tyrosine kinases), and ML-9 (a selective inhibitor of MLC kinase), were shown to inhibit protein kinase activity. Even so, all three compounds inhibited catalytic activity at millimolar concentration, which is 10^3 - 10^6 times more concentrated than those at which they are known to inhibit their eukaryotic target enzymes, making it difficult to characterize the *S. solfataricus* protein kinase as a member of the extended family of eukaryotic protein kinases based on these inhibition studies. Given the variety of both specific and nonspecific serine/threonine/tyrosine protein kinase inhibitors tested and the weak responses obtained, it seems unlikely that testing other eukaryotic protein kinase inhibitors would offer further insight into the enzyme's nature.

Autophosphorylation of the Sulfolobus solfataricus Protein Kinase

Most protein kinases catalyze autophosphorylation, a process that often alters the functional properties of the protein kinase and is sometimes controlled by regulatory ligands (Smith et. al., 1993). Serine, threonine, or tyrosine residues can act as the phosphoacceptor, and typically, multiple sites on the protein kinase are phosphorylated.

When the *S. solfataricus* protein kinase was completely resolved from other proteins using two-dimensional electrophoresis, a subsequent "in gel" assay conducted at 65^{o}C revealed that the protein kinase had undergone autophosphorylation. Phosphoamino acid analysis demonstrated that the protein kinase could autophosphorylate on threonine residues using both [γ-^{32}P] ATP and [γ-^{32}P] GTP as phosphoryl donors. Incubation of the autophosphorylated protein kinase that had been renatured in gel with diffusible phosphoacceptor substrate had no effect on the level of [^{32}P] phosphate associated with the protein kinase. These results indicating that the phosphothreonine bond was indeed formed as the result of an autophosphorylation event and did not function as a catalytic intermediate, as under the latter circumstances transfer of [^{32}P] phosphate to the exogenous substrate should have been observed.

We were unable to determine if autophosphorylation had any effect on the catalytic activity of the protein kinase, as its stability could not be maintained during the prolonged exposure at the elevated temperatures that were required to autophosphorylate the enzyme. Another approach that could be used to determine the consequences of autophosphorylation would be to mutationally alter the threonine residue(s) that undergo autophosphorylation such that the enzyme can no longer autophosphorylate. The protein kinase activity of this mutationally derived protein could then be compared to wild type protein kinase to determine if autophosphorylation had any effect on the catalytic ability of the enzyme. However, until the identity of the polypeptide responsible for the protein kinase activity is determined and its gene subsequently cloned, such experiments can not be accomplished.

Glycosylation of the Sulfolobus solfataricus Protein Kinase

While glycosylation of proteins in eukaryotes has been extensively studied, this phenomenon has only recently received attention in prokaryotes. As was the case with another covalent modification, protein phosphorylation, for a long time prokaryotes were considered to be unable to glycosylate proteins. However, it is now clear that protein glycosylation (both O- and N-glycosylation) is widespread among members of the *Bacteria* and *Archaea* (reviewed in Moens and Vanderleyden, 1997). As in eukaryotes, the glycosylation of proteins in prokaryotes has been shown or suggested to serve several

functions, including maintenance of protein conformation and stability, protection against proteolysis, enhancement of heat stability, surface and intracellular recognition, and cell adhesion (Moens and Vanderleyden, 1997).

When autophosphorylated *S. solfataricus* protein kinase was resolved using two-dimensional electrophoresis, multiple [^{32}P] phosphorylated spots were observed. Such behavior indicated the presence of a post-translational modification such as glycosylation. To determine if the protein kinase was a glycoprotein, SDS-PAGE gels in which the *S. solfataricus* protein kinase was electrophoresed were stained with a GelCode Glycoprotein Staining Reagent (Pierce). A polypeptide representing the ≈ 67 kDa autophosphorylated protein kinase was observed, indicating that the protein kinase was glycosylated. These results were verified by showing that protein kinase treated with trifluoromethanesulfonic acid i) exhibited a shift in M_r on SDS-PAGE and ii) did not react with the GelCode Glycoprotein Staining Reagent. Subsequent attempts to identify and characterize the carbohydrate structure of the *S. solfataricus* protein kinase using conjugated lectins indicated the presence "high mannose" N-glycan chains or O-glycosidically linked mannoses. However, this must be viewed with caution as a good deal of variation has been observed in the carbohydrate constituents and linkage units of prokaryotic glycoproteins (Moens and Vanderleyden, 1997).

The exact function of glycosylation of the protein kinase remains a mystery, however, given that most prokaryotic and eukaryotic glycoproteins are especially abundant in the plasma membrane, or synthesized as precursors with a signal peptide, suggests that glycosylation may function to associate the protein kinase with the membrane (Hemmings, 1997). Since the harsh chemical treatment of the protein kinase with trifluoromethanesulfonic acid resulted in a total loss of catalytic activity, while attempts to enzymatically deglycosylate the protein kinase failed, we were unable to compare the characteristics of the glycosylated enzyme with that of the deglycosylated enzyme. We also attempted to grow *S. solfataricus* in the presence of bacitracin, an antibiotic that was previous demonstrated by Lechner and Wieland to inhibit the glycosylation of glycoproteins in prokaryotes, however the antibiotic proved fatal to the cells (Lechner and Wieland, 1989).

Identification of Open Reading Frames in *Sulfolobus solfataricus* Whose Predicted Products Exhibit Homology to Eukaryotic Protein Kinases.

Comparison of the primary sequence of over 400 known and potential protein kinases from eukaryotes indicates that the catalytic domain is approximately 280 amino acids in length and contains twelve conserved subdomains, numbered I - V, VI_a, VI_b, VII – XI (Hanks and Hunter, 1995). While there may be some insertions or deletions within the catalytic domain, all members function in the same manner. The domain folds into a bilobal unit with an N-terminal lobe, comprised of subdomains I – IV, whose primary function is binding and orientating the MgATP complex that serves as the phosphoryl donor and the larger C-terminal lobe, comprised of subdomains VI_a – XI, whose primary function is binding the peptide or protein substrate and initiating the phosphoryl transfer (Smith et. al., 1997).

We employed two methods to search the partial genome sequence of *S. solfataricus* for homologs of eukaryotic protein kinases. The first method used two eukaryotic protein kinase "signature motifs," subdomain I (Gly-X-Gly-X-X-Gly-X) and subdomain VI_b (Asp-X_4-N), as templates for the initial identification. These two motifs were selected because of their absolute conservation among eukaryotic protein kinases. Subdomain I represents conserved residues found in the glycine-rich nucleotide binding domain, while subdomain VI_b includes the residues thought to participate directly in catalysis. The second method used the amino acid sequence of the catalytic subunit of cAMP-dependent protein kinase (gi|11596395|) from *Candida albicans* to identify *ORFs* encoding hypothetical protein kinases. This protein kinase was selected as template because it contained all twelve characteristic subdomains of eukaryotic protein kinases.

Figure 7-1 summarized the results from the search of *S. solfataricus* for homologs of eukaryotic protein kinases. A total of six *ORFs* were identified whose predicted products contained sequence features considered essential for phosphotransferase activity. While none of the hypothetical proteins conformed exactly to the eukaryotic prototype, all of the hypothetical proteins contained plausible candidates for at least seven of the twelve conserved subdomains. Considering that no invariant or nearly invariant residues are present in subdomains IV and VI_a, these hypothetical proteins

SD		I	II	III	IV	V	VI_a	VI_b	VII	VIII	IX	X	XI
EukPK		$G_{50}G_{52}$	K_{72}	E_{91}		$M_{120}EYVE_{127}$		$D_{166}N_{171}$	D_{184} FG	APE_{208}	D_{220}		R_{280}
ORF sso-	**#AA**												
0197	**287**	$G_{50}G_{52}$	K_{71}	E_{90}		$M_{132}EYIE_{139}$		$D_{176}N_{181}$	$D_{194}W$		D_{222}		
0433	**223**	G_{49}	K_{72}	E_{91}		$E_{121}I_{123}E_{126}$		$D_{164}N_{169}$	$D_{183}FG_{185}$	E_{208}	D_{220}		
2291	**554**	G_{50}	K_{68}	E_{94}		$M_{142}E_{143}$		$D_{189}N_{194}$	$D_{223}LG_{225}$		D_{262}		R_{310}
2387	**583**	$G_{50}G_{53}$	K_{72}	E_{81}		$D_{120}V_{123}E_{126}$		$D_{222}N_{227}$	$E_{240}AG_{242}$	E_{267}			R_{301}
3182	**537**	$G_{50}G_{52}$	K_{70}	E_{86}		E_{134}		$D_{183}N_{188}$	$D_{217}LG_{219}$	D_{239}	D_{254}		R_{328}
3207	**669**	$G_{50}G_{52}$	K_{70}	E_{84}		$M_{122}EYD_{129}$		$D_{175}N_{180}$	$D_{209}LG_{211}$		D_{243}		R_{320}

Figure 7-1. Schematic Representation of the Essential Sequence Features of Eukaryotic Protein Kinases Found in Potential Homologs From *Sulfolobus solfataricus.*

A linear representation of the twelve conserved subdomains (**SD I – XI**) common to the 280 amino acid catalytic core of eukaryotic protein kinases is listed in bold blue color across the top (**EukPK**). Amino acids considered to be invariant or nearly invariant in eukaryotic protein kinases are listed using the numbering system adapted by Taylor et al. (Taylor et. al., 1993). Shown down the left hand side are the ***ORFs*** (***bold italics***) from *S. solfataricus* and their predicted amino acid sequence features, as well as total number of amino acids (**#AA**).

appear to contain nearly the entire conserved catalytic domain of eukaryotic protein kinases.

All of the hypothetical proteins, with the exception of perhaps *sso0433* which contains an "atypical" subdomain I, appeared to contain subdomain I – VII. Most deviations from the eukaryotic protein kinase paradigm were found in the C-terminal regions of the catalytic core, i. e. subdomains VIII – XI. This trend was observed previously in genome searches of eukaryotic protein kinases among members of the *Archaea* (Shi et. al., 1998; King and Smith, 1995). Since these regions comprise portions of the protein kinase that are removed from the immediate sites of nucleotide binding and catalysis, variations in these regions may not be inconsistent with protein kinase activity.

The hypothetical protein product of *ORF sso0433* must be viewed with caution as its sequence displays two key deviations from the eukaryotic prototype. First, while a glycine residue is present at position 49, it lacks the characteristic sequence of subdomain I, Gly_{50}-X-Gly_{52}. Studies on other phosphotransferases indicate that the central glycine, G_{52}, may not be as essential for nucleotide binding as its prevalence in eukaryotic protein kinases suggests. The bacterial antibiotic phosphotransferases bind and utilize ATP in a manner similar to protein kinases, but do so without using the subdomain I motif of eukaryotic protein kinases (Brenner, 1988; Martin et. al., 1988). The antibiotic phosphotransferases do, however, contain the conserved lysine of subdomain II and glutamate of subdomain III, suggesting that these comprise essential features of nucleotide binding. Second, its length of 223 amino acids falls well short of that characteristic of eukaryotic protein kinases, which span ≈ 280 amino acids. The hypothetical protein does, however, contain the conserved glutamate reside (E_{208}) of subdomain VII, as well as the conserved aspartate residue (D_{220}) of subdomain IX. However, as discussed previously, because region XI is located outside of the enzyme's active site it may be dispensable. One possible explanation why the hypothetical protein encoded by *ORF sso0433* lacks an apparent subdomain I and XI is that this enzyme is not a protein kinase per se, but rather possesses a phosphotransferase activity, such as the bacterial antibiotic phosphotransferases that acts on nonprotein substrates as is the case for bacterial antibiotic phosphotransferases.

Interestingly, the regions of the hypothetical proteins encoded by *ORFs sso2291, sso3182, sso3207,* and *sso2387* that displayed sequence features characteristic of eukaryotic protein kinases (subdomain I – XI) were located in the C-terminal half of the protein. Such an arrangement suggests that the N-terminal half of these hypothetical proteins serves a structural and/or regulatory function. Circumstantial evidence for such an arrangement was provided by the identification of potential transmembrane domains, N-glycosylation sites, myristylation sites, and phosphorylation sites in the N-terminal half of these hypothetical proteins by computer homology searches.

Cloning and Expression of a Gene From *Sulfolobus solfataricus, sso2387,* Purification and Characterization of its Encoded Protein Kinase, SsPK2.

Analysis of the Putative Catalytic Domain of SsPK2

Since biochemical efforts to reveal the identity of the *S. solfataricus* protein kinase (described in Chapter III) failed, we attempted to identify the gene encoding this enzyme within the genome of *S. solfataricus*. Our working hypothesis was that the source of the protein kinase activity was a member of the "eukaryotic" family of protein kinases, hence its catalytic domain should display features characteristic of eukaryotic protein kinases. Six *ORFs* were identified whose predicted products contained sequence features considered essential for phosphotransferase activity. Of these six hypothetical proteins, three (encoded by *ORFs: sso2291, sso2387*, and *sso3182*) exhibit noticeable similarities to the previously characterized *S. solfataricus* protein kinase, which included molecular weight, consensus sequence motifs for potential N-glycosylation sites, and numerous potential threonine phosphorylation sites as identified using computer homology searches. *ORF sso2387* was selected for further investigation because this gene was discovered well in advance of the other *ORFs* from *S. solfataricus*.

The protein product of *ORF sso2387*, designated SsPK2, displays nine of twelve essential subdomains characteristic of eukaryotic protein kinases, including subdomains I – VII, VIII, XI (Figure 5-14 and 7-1). The apparent catalytic domain of SsPK2 deviates somewhat from the eukaryotic prototype however. Its sequence contains the key features of subdomain I including the conserved aliphatic hydrophobic (Val_{51}, Ala_{54}) and glycine

(Gly_{50}, Gly_{53}, and Gly_{55}) residues, however, the invariant glycine residue, Gly_{52}, characteristic of the eukaryotic subdomain I is conspicuously absent. The fact that SsPK2 was shown to be catalytically active suggests that Gly_{52} may not be as indispensable for nucleotide binding as its near absolute conservation among eukaryotic protein kinases suggests. Also, while most eukaryotic protein kinases contain a highly conserved aliphatic hydrophobic residue at the end of subdomain I, SsPK2 contains a hydroxyl residue (Thr_{57}), a residue, which is only occasionally observed in the eukaryotic archetype (Smith et. al., 1997).

While SsPK2 possesses the invariant Lys_{72} residue of subdomain II, the alleged glutamate residue of subdomain III, Glu_{81} if present, is slightly removed from its characteristic position of Glu_{91} observed in eukaryotic protein kinases. Curiously, SsPK2 contains a glutamine residue this position, Gln_{91}. While the function of subdomain III is to stabilize the interactions between Lys_{72} and the alpha and beta phosphates of MgATP, a conservative substitution such as an glutamine at this position, which also possesses hydrogen bonding potential, could conceivably produce a catalytically active enzyme. Whether or not this is the case for SsPK2 remains to be determined. If a mutational alteration of Gln_{91} to an amino acid that lacks hydrogen bonding potential, such as alanine, resulted in a catalytically inactive enzyme this would suggest that subdomain III of SsPK2 is characterized by a Gln_{91} rather than a Glu_{81}. If on the other hand such an enzyme were catalytically active, than this would suggest that Glu_{81} defines subdomain III of SsPK2.

The putative subdomain V of SsPK2 (Asp_{120}-Ile-Ile-Val-Val-Gly-E_{126}) is noticeably similar to that of eukaryotic protein kinases with the exception that the aspartate residue (Asp_{120}) is substituted for the typical glutamate residue. Given the highly conservative nature of this substitution, it is reasonable to assume that the function of this subdomain, which is to help anchor the MgATP through hydrogen bonding and hydrophobic interactions, is maintained in SsPK2.

There appears to be an insertion of ≈ 60 amino acids between the apparent subdomains V and VI_b of SsPK2 that is not typically observed in eukaryotic protein kinases. For most eukaryotic protein kinases subdomain V spans residues 120 - 127, while subdomain VI_b spans residues 166 – 171. In SsPK2, the perceived subdomain V

spans residues 120 – 126, consistent with eukaryotic protein kinases, however subdomain VI_b (Asp_{222}-X_4-Asn_{227}) spans residues 222 – 227. In eukaryotic protein kinases subdomain VI_a is defined as the region between these two subdomains. Since there are no invariant residues, or any other residues present in subdomain VI_a that seem to participate in nucleotide or substrate interactions, its function appears to be only structural. Therefore, the insertion in SsPK2 observed within its presumed subdomain VI_a most likely serves a structural role as well.

The sequence of SsPK2 also contains plausible candidates for all three of the highly conserved active-site residues, which define subdomains VI_b and VII and are thought to participate directly in catalysis. The presumed catalytic loop (subdomain VI_b) of SsPK2 possesses the characteristic Asp-X_4-Asn sequence spanning Asp_{222}-X_4-Asn_{227}. In eukaryotic protein kinases, subdomain VII contains an invariant aspartate residue at position 184 that functions to chelate the Mg^{2+} ions that bridge the alpha and beta phosphates thereby orienting the gamma phosphate of MgATP for transfer to the substrate (Smith et. al., 1997). In SsPK2, a glutamate also is present at this position. However, since this amino acid is also capable of chelating Mg^{2+}, it is reasonable to assume that it performs a similar function.

In eukaryotic protein kinases, the glutamate (Glu_{208}) of subdomain VIII is a nearly invariant residue that forms a salt bridge with the arginine (Arg_{280}) of subdomain XI, stabilizing the protein/peptide-binding lobe of the protein kinase (Smith et. al., 1997). SsPK2 contains plausible candidates for both subdomain VIII (Glu_{267}) and subdomain XI (Arg_{301}).

There appears to be a deletion of ≈ 40 amino acids in the sequence of SsPK2 between subdomains VIII – XI that has resulted in a loss of either subdomain IX or subdomain X (although this is difficult to determine as subdomain X lacks any invariant residues). For most eukaryotic protein kinases, a conserved aspartate residue of subdomain IX is positioned at residue 220, which is located just downstream of the conserved Glu_{208} of subdomain VIII. The aspartate of subdomain IX forms hydrogen bonds with backbone amides of subdomain VI_b to stabilize the catalytic loop (Smith et. al., 1997). In SsPK2, a glutamate (putative subdomain VIII) is positioned at residue 267 and an arginine (putative subdomain XI) is positioned at residue 301, however a lysine

residue (Lys_{279}) is observed where the conserved aspartate residue of subdomain IX is predicted to be located. The only plausible candidates for subdomain IX within SsPK2 are glutamate residues positioned at E_{272}, E_{273}, E_{290}, and E_{291}. Given the important role that the aspartate of subdomain VIII plays in stabilizing the catalytic loop, it seems unlikely that SsPK2 would lack this subdomain, suggesting that one of these glutamate residues (E_{272}, E_{273}, E_{290}, and E_{291}) plays an analogous role. Mutationally altering each of these four glutamate residues separately to an amino acids lacking the ability to form hydrogen bonds, i.e. alanine, and determining which, if any, resulted in a catalytically inactive enzyme this would offer insight into plausible candidates for subdomain IX. If none of these mutationally altered enzymes result in loss of activity this would suggest that the Asp_{220} of subdomain IX is not as essential as its absolute conservation among eukaryotic protein kinases implies.

Cloning ORF sso2387 and its Amino Terminal Deletion Mutant and Expression and Purification of Their Protein Products SsPK2 and SsPK2trunk

Upon examination of the primary sequence of SsPK2 (Figure 4-14), one notices that its apparent catalytic domain is located within the carboxy-terminal half of the protein. Therefore, we decided to construct a truncated form of SsPK2, which we called SsPK2trunk, containing the putative catalytic domain to determine if this protein also exhibited protein kinase activity.

The amplification of *ORF sso2387* from genomic DNA was performed using the polymerase chain reaction and two oligonucleotides primers designed from the DNA sequence of this *ORF*. The PCR reaction produced a single product of ≈ 1.8 kbp as observed using agarose gel electrophoresis. The amplification of the amino terminal deletion mutant of *ORF sso2387* performed using polymerase chain reaction, plasmid DNA containing *ORF sso2387*, and two oligonucleotide primers designed from the DNA sequence of *ORF sso2387* resulted in a single product of ≈ 0.8 kbp as observed using agarose gel electrophoresis. Each PCR products was ligated into a pCR T7/NT TOPO vector and sequenced using both forward and reverse primers to verify that the PCR amplified DNA was indeed *ORF sso2387* or the amino terminal deletion mutant of *OFR sso2387* and that each had ligated into the vector properly. The pCR T7/NT TOPO

vector added a sequence of DNA encoding six successive histidine residues 5' to the cloned gene to facilitate the purification of its recombinant protein product by metal affinity chromatography.

When expressed in *E. coli*, the majority of SsPK2 and SsPK2trunk were present in the form of inclusion bodies. Attempts to produce soluble recombinant proteins by growing *E. coli* cells at a lower temperature also resulted in the formation of inclusion bodies. A number of alternative methods which might be used in future large-scale expressions of SsPK2 and SsPK2trunk to increase the soluble fraction of these recombinant proteins include growing cells in different media (Doonan, 1996), using a different host strain (Gribskov and Burgess, 1983), or cloning the genes into an expression vector that creates a fusion protein with a highly soluble protein, such as glutathione-S-transferase or thioredoxin (Smith and Johnson, 1988; LaVallie et. al., 1993).

Since the majority of the recombinant proteins were expressed in the form of inclusion bodies, they were easily separated from a large proportion of the soluble proteins from *E. coli* using centrifugation. Western Blotting demonstrated that the majority of the protein contained within the inclusion bodies was either SsPK2 or SsPK2trunk, respectively. Both of the recombinant proteins were solubilized from inclusion bodies using a buffer solution containing 5M urea, as attempts to solubilize the recombinant proteins using nonionic detergents were only marginally effective. Since 5M urea was required for solubilization, all subsequent protein kinase assays were performed in buffer containing 5M urea.

Another problem that arose during the expression of both SsPK2 and SsPK2trunk was that these proteins were being proteolytically degradated inside *E. coli*. This resulted in a substantial loss of full-length recombinant proteins. One way to overcome this problem in the future would be to explore the use of other strains of *E. coli* that have been genetically manipulated to be deficient in proteases.

When the urea soluble fractions of SsPK2 and SsPK2truck, respectively were incubated with [γ-^{32}P] ATP, at 65^{o}C both were observed to undergo autophosphorylation. In order to verify that the source of this protein kinase activity was from SsPK2 or SsPK2trunk respectively, the recombinant proteins were purified using metal affinity

chromatography. Each of the fractions collected from the nickel column were analyzed using SDS-PAGE, and those which contained SsPK2 or SsPK2trunk alone, as identified by Coomassie R-250 staining and Western Blot Analysis, were pooled and retained as the Ni fraction. When Ni fractions for SsPK2 or SsPK2trunk were assayed for protein kinase activity, each displayed an ability to phosphorylate several exogenous protein substrates *in vitro*.

It also should be noted that a control experiment was performed in which *E. coli* BL21(DE3)pLysS, pCR T7/NT TOPO cells were grown, harvested, and proteins separated on a metal chelating column by a procedure identical to that described for SsPK2. Using a solution assay, as described in Methods, and BSA as phosphoacceptor substrate, very little kinase activity (≈ 5.0 pmol / min / mg, which corresponded to < 10% of that observed with SsPK2) was detected from the urea soluble fraction and no activity was detected in the nickel chelating column fraction. These results confirmed that the protein kinase activity observed in the various fractions was from SsPK2 or SsPK2trunk and not from an endogenous *E. coli* enzyme.

Autophosphorylation of SsPK2 on Serine

One of the questions we had when we expressed SsPK2 was whether this putative protein kinase was the source of the threonine-specific protein kinase activity from *S. solfataricus* that was described in Chapter III. When SsPK2 was incubated with [γ-^{32}P] ATP, at 65^{o}C it was observed to undergo autophosphorylation. Subsequent phosphoamino acid analysis, however, revealed a phosphoserine bond had been formed, indicating that this *ORF* was not the gene encoding the *S. solfataricus* protein kinase described in Chapter III.

MALDI-MS analysis was employed to identify the site(s) of autophosphorylation in SsPK2. Three peaks in the mass spectrum were observed whose masses corresponded to those of hypothetical peptides containing serine residues as well as one or more PO_3 groups. However, these assignments should be viewed with caution as each contains two missed trypsin cleavage sites, and no masses corresponding to the completely digested peptides (zero missed trypsin cleavages) were observed in the MALDI-MS spectrum. Possible explanations for this discrepancy include a failure of SsPK2 to digest well with

trypsin, or poor ionization of the completely digested peptides during the MS analysis. Two of the three peptides identified by MALDI-MS analysis (2300.40 MH^+ and 2236.27 MH^+) implicate the same serine residues (Ser_{538} and Ser_{539}) as the sites where autophosphorylation occurs. Curiously, these two serine residues appear to be within the putative subdomain VIII of SsPK2, the region known to contain the major autophosphorylation sites of eukaryotic protein kinases (Smith et. al., 1997). Future experiments that could be designed to verify these result would include i) synthesizing peptides modeled after these autophosphorylation sites that could be tested as potential substrates for SsPK2 and ii) mutationally altering these serine residues to another amino acid that cannot undergo phosphorylation, such as alanine, and asking if autophosphorylation still occurs.

Metal Ion Cofactor Preference of SsPK2

When SsPK2 was assayed at 37°C using BSA as the phosphoacceptor substrate, it exhibited a preference for Mn^{2+} over the other divalent metal ions tested (Ca^{2+}, Mg^{2+}, Ni^{2+}, and Zn^{2+}). While this is a fairly infrequent occurrence among protein kinases, this phenomenon had been observed previously with the *S. solfataricus* protein kinase described in Chapter III, as well as serine/threonine phosphatases from the *Archaea* (Kennelly et. al., 1993; Mai et. al., 1998; Solow et. al., 1997), and in experiments investigating protein kinase activity in *Sulfolobus acidocaldarius* (Skorko, 1989). However, Mn^{2+} may not serve as the principal cofactor *in vivo*, as assays conducted at temperatures within the physiological range of the organism demonstrated that SsPK2 could autophosphorylate using Mg^{2+} as well as Mn^{2+} as the metal ion cofactor. Interestingly, SsPK2 exhibited the ability to autophosphorylate at 65°C utilizing a variety of other, unusual divalent metals cofactors exhibiting a preference for $Ca^{2+} \approx Zn^{2+} > Ni^{2+}$.

Phosphorylation of Exogenous Substrates by SsPK2

SsPK2 proved to be relatively selective when challenged with a number of exogenous protein and peptide substrates. Only BSA, casein, RCM lysozyme, and mixed histones were phosphorylated at appreciable levels *in vitro*. Phosphoamino acid analysis revealed that the protein kinase had a preference for phosphorylating serine residues *in*

vitro as BSA, casein, and mixed histones were phosphorylated exclusively on serine. While it should be noted that these substrates were nonphysiological, it was also observed that autophosphorylation of SsPK2 occurred on serine residues as well.

The polycation, polyLys, was observed to stimulate SsPK2 activity ≈ 2-fold towards macromolecular substrates such as BSA and casein. We believe this phenomenon could be attributed to the ability of polyLys to recruit SsPK2 and its substrates (ATP and BSA or casein) to the same microenvironment. Interestingly, in assays in which poly(Glu:Tyr) was substituted for polyLys, little/no change in SsPK2 activity towards macromolecular substrates was observed. These results support the argument that polycations stimulate SsPK2 activity by localizing its substrates. Current models of signal transduction in eukaryotes demonstrate that many protein kinases, protein phosphatases, and their substrates are often specifically localized within the cell, and that such a spatiotemporal arrangement is critical for the regulation of these pathways (Hunter, 2000). These data suggest that SsPK2 may be regulated *in vivo* in such a manner, being concentrated about a macromolecular polycation or attracted to proteins bearing a net positive charge.

Nucleotide Preference of SsPK2

Unlike the *S. solfataricus* protein kinase described in Chapter III, SsPK2 was unable to phosphorylate exogenous protein substrates using [γ-^{32}P] GTP as the potential phosphoryl donor. However, when SsPK2 was incubated with [γ-^{32}P] GTP at 65^{o}C, it was observed to undergo autophosphorylation. When compared to ATP, however, GTP was a considerably poorer substrate for SsPK2 autophosphorylation, suggesting the enzyme prefers ATP as phosphoryl donor substrate *in vivo*.

Sensitivity of SsPK2 to Inhibitors of Eukaryotic Protein Kinases

When SsPK2 was challenged with several known inhibitors of eukaryotic protein kinases, only tamoxifen (a known inhibitor of protein kinase C) was shown to inhibit protein kinase activity. Even so, with an apparent $IC_{50} \approx 200 – 300$ μM, that was 20 - 30 times more concentrated than those at which tamoxifen is known to inhibit its eukaryotic

target enzyme, protein kinase C (IC_{50} = 10μM), it is difficult to determine the significance of such results.

SsPK2trunk Also Possesses Protein Kinase Activity

SsPK2trunk was able to phosphorylate a variety of exogenous protein substrates using [γ-^{32}P] ATP, exhibiting activity similar to that observed for SsPK2. SsPK2trunk was also observed to undergo autophosphorylation at 65^{o}C, suggesting that at least one of the autophosphorylation sites was contained within the C-terminal half of SsPK2. These results are consistent with the MALDI-MS analysis implicating Ser_{538} and Ser_{539} as the residues that underwent autophosphorylation. From these results, we concluded that SsPK2trunk did define the catalytic domain of the SsPK2.

Identification of Proteins that Undergo Phosphorylation From the Membrane Fraction of *Sulfolobus solfataricus* Using MALDI-MS Analysis.

In eukaryotes, protein phosphorylation-dephosphorylation participates in the regulation of virtually all aspects of cell physiology and development. The importance of this regulatory mechanism is underscored by the large number of protein kinases and protein phosphatases that are present in various eukaryotic genomes (Hunter, 1995; Hunter and Plowman, 1997) as well as its pervasiveness throughout the three domains (Cozzone, 1998; Kennelly and Potts, 1996; Shi et. al., 1998; Kennelly and Potts, 1999; Zhang, 1996). Despite this fact, our understanding of the depth and breadth of this regulatory mechanism remains relatively limited. This appears to be changing, however, with the recent advances in the use of mass spectrometry in conjunction with protein/DNA-sequence database search-algorithms that permit the identification of proteins with unprecedented speed (Oda et. al., 1999; Link, 1999; Vener et. al., 2001). One of the most popular MS techniques is matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) coupled with web-based peptide mass search programs (Patterson, 1995), also known as peptide mass fingerprinting (Link, 1999).

To determine the potential of this technique for our laboratory's future use we analyzed the major phosphoproteins in the *S. solfataricus* DE-52 fraction that underwent

phosphorylation *in vitro* using [γ-^{32}P] ATP. Five of the most prominent [^{32}P] phosphorylated proteins observed in the SDS-PAGE gel (Figure 6-1) and six of the most prominent [^{32}P] phosphorylated proteins observed in the 2-D gel (Figure 6-2) were identified by MALDI-MS on the basis of peptide mass fingerprinting, following in-gel digestion with trypsin and matching with the theoretical peptide masses of all archaeal proteins in the ProFound database.

Table 7-1 summarized the results from the MALDI-MS analysis. As seen in the table, nine of the eleven [^{32}P] phosphorylated species were tentatively identified using the ProFound search engine. A probability score of 1.0e+000 represents the highest value possible and denotes that the candidate in a database search most likely is the protein being analyzed. In some instances, ProFound also will provide an indicator of the quality of the search result, called a Z score, which can be used to indicate the likelihood that a candidate belongs to a random match population in the sense of traditional statistics. The Z score is the distance to the population mean in units of standard deviation and is estimated by comparing the search result against an estimated random match population. For instance, a Z score of 1.65 for a search means that the search is in the 95th percentile. In other words, there are about 5% of random matches that could yield higher Z scores than this search (Z score of 1.28 corresponds to the 90th percentile; Z score of 1.65 corresponds to the 95th percentile; Z score of 2.33 corresponds to the 99th percentile; and Z score of 3.09 corresponds to the 99.9th percentile) (Tang et. al., 2000).

The candidate proteins having high probability values (approaching 1.0e+000), and uniform/extensive peptide coverage (> 20%) could be assigned the identity determined using ProFound with a relative high degree of confidence. Those meeting these criteria include ≈ 85 kDa polypeptide (SDS-PAGE) identified as a translational endoplasmic reticulum ATPase and the ≈ 45 kDa polypeptide (SDS-PAGE) identified as a ≈ 42 kDa hypothetical protein from *S. solfataricus*. The Z-values for the ≈ 85 kDa and the ≈ 45 kDa polypeptides were also relatively high, 1.41 and 1.14 respectively (corresponding to approximately the 90th percentile), suggesting a positive identification as well. While the ProFound results obtained for several of the other polypeptides are highly indicative as to their respective identities, an unambiguous identification cannot be made. In the case of two of the [^{32}P] phosphorylated polypeptides from *S. solfataricus*

Excised Band / Spot	# Matching Peptides	% Coverage	P Score	Identified Protein From *S. solfataricus*
SDS-PAGE				
≈ 105 kDa	12	14	1.1e-05	ATP-Dependent Helicase
≈ 85 kDa	24	36	1.00e+00	Translational ER ATPase
≈ 67 kDa	-	-	-	-
≈ 48 kDa	14	37	8.8e-04	Ser Hydroxymethyltransferase
≈ 45 kDa	17	34	9.4e-01	Hypothetical (gi|12312957|)
2-D Electrophoresis				
≈ 105 kDa ≈ 6.1 pI	-	-	-	-
≈ 67 kDa ≈ 5.7 pI	5	5	8.7e-001	Hypothetical (gi|12312047|)
≈ 68 kDa ≈ 5.6 pI	10	12	1.0e+000	Acetolactate Synthase Homolog
≈ 69 kDa ≈ 5.5 pI	11	9	9.7e-001	Acetyl-CoA Synthetase
≈ 70 kDa ≈ 5.4 pI	9	17	1.0e+000	Hypothetical (gi|13915259|)
≈ 71 kDa ≈ 5.3 pI	8	11	9.9e-001	CO Dehydrogenase

Table 7-1. Summary of the *Sulfolobus solfataricus* Proteins Identified Using Peptide Mass Fingerprinting.

These proteins were identified by Matrix-assisted laser desorption/ionization-mass spectrometry (MALDI-MS) on the basis of peptide mass fingerprinting, following in-gel digestion with trypsin and matching with the theoretical peptide masses of all known proteins from *S. solfataricus* using the ProFound search-algorithm at Rockefeller University (http://www.proteometrics.com/). Shown, are the [^{32}P] phosphorylated bands (SDS-PAGE) or spots (2-D Electrophoresis), the number of matching peptides, the percent coverage (defined as the ratio of the portion of protein sequence covered by matched peptides to the whole length of protein sequence), the probability score (P score), and the identity of the protein from *S. solfataricus*. Note that a P score of 1.0e+000 denotes that the candidate in a database search most likely is the protein being analyzed.

proteins (SDS-PAGE ≈ 105 kDa and ≈ 48 kDa), relatively low probability scores were obtained. This indicated that the ProFound identification of these proteins was most likely incorrect. Interestingly though, while the probability score of the ≈ 48 kDa was relatively low (8.8e-04), 37% of the candidate protein was covered by the observed peptide masses. Such a result is probably a caveat of the ProFound software as the probability score is based on input data from the user. The input data for this polypeptide included an additional twenty unidentified peptide masses, which no doubt contributed to the low probability value obtained from ProFound.

In the future when MALDI-MS / peptide mass fingerprinting is employed to identify phosphoproteins in *S. solfataricus*, it would be beneficial to digest the [^{32}P] labeled polypeptide with a second protease, such as V8 protease, in addition to trypsin. The peptide masses derived from a digestion with V8 protease could then be used in a separate peptide mass search to confirm the identity obtained using peptide masses derived from trypsin. An unambiguous identification could then be assigned to those proteins identified by both the V8 protease peptide masses and the trypsin derived peptide masses. Such a procedure could be used to verify the results obtained here as well. Still another method that could be used to verify the identity of these phosphoproteins would be to produce an antibody directed against a specific phosphoprotein that could be used to immunoprecipitate [^{32}P] labeled polypeptides from *S. solfataricus* cell extracts.

It should be mentioned that in peptide mass fingerprinting, there are frequently peptides whose masses cannot be explained, as was the case here as well. These are usually attributed to either a missed cleavage during the enzymatic digestion process, incomplete reduction and alkylation of a protein, protein modifications such as the oxidation of methionine, or the presence of protein post-translational modifications, as well as to instrument, database and human errors. Many of the unidentified peptide masses observed here no doubt could be credited to these aforementioned reasons.

Collectively these results suggest that these techniques coupled with [^{32}P] labeling of intracellular phosphoproteins will have a tremendous potential for mapping out a major portion of the phosphoproteome (i.e. the subset of proteins in the proteome that becomes modified *in vivo* by phosphorylation) from *S. solfataricus*. While the dynamics

and transient nature of many protein phosphorylation events underscores the difficulties of resolving the entire phosphoproteome, nevertheless, the identification of the principal cellular phosphoproteins in *S. solfataricus* should bring significant biological insight.

Conclusion

Protein phosphorylation was, until recently, largely regarded as a refined regulatory mechanism which emerged late in evolutionary time to meet the specific needs of the multicellular organisms. The discovery of what were once considered to be "eukaryotic" protein O-kinases and O-phosphatases in prokaryotes, coupled with discoveries of what were once considered "prokaryotic" two-component signalling modules in eukaryotes, has raised some interesting questions concerning the origin and evolution of these regulatory mechanisms. Commonality suggests that reversible phosphorylation arose early in evolution as a dynamic and versatile regulatory mechanism. The detection of such components in the archaeon *S. solfataricus* further emphasizes the antiquity of protein phosphorylation, as these organisms are believed to be among the earliest inhabitants of Earth. Future studies on protein phosphorylation networks in *S. solfataricus* offers great promise for tracing its origins and evolution, as well as presenting a unique, and largely untapped library of phylogenetically diverse protein kinases, protein phosphatases, and phosphoproteins with which to advance our understanding of a regulatory mechanism that is crucial to all organisms.

REFERENCES

Albers S. V., J. L. van de Vossenberg, A. J. Driessen, and W. N. Konings WN (2000) "Adaptations of the archaeal cell membrane to heat stress," Front. Biosci. 1; 5, D813-820.

Alex, L. A. and M. I. Simon (1994) "Protein kinase and signal transduction in prokaryotes and eukaryotes," Trends in Genet. 10, 133-138.

Alex, L. A., K. A. Borkovich and M. I. Simon (1996) "Hyphal development in *Neurospora crassa*: involvement of a two-component histidine kinase," Proc. Natl. Acad. Sci. USA 93, 3416-3421.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990) "Basic local alignment search tool," J. Mol. Biol. 215, 403-410.

Armbruster D.W., C. J. Daniels (1997) "Splicing of intron-containing tRNATrp by the archaeon *Haloferax volcanii* occurs independent of mature tRNA structure," J. Biol. Chem. 272(32), 19758-19762.

Arvalli, R. N. and R. A. Garret (1997) "Shuttle vectors for hyperthermophilic archaea," Extremophiles 1, 183-191.

Ashton A. R., J. N. Burnell, M. D. Hatch (1984) "Regulation of C4 photosynthesis: inactivation of pyruvate, Pi dikinase by ADP-dependent phosphorylation and activation by phosphorolysis," Arch. Biochem. Biophys. 230(2), 492-503.

Av-Gay J. and M. Everett (2000) "The eukaryotic-like Ser/Thr protein kinases of *Mycobacterium tuberculosis*," Trends Microbiol. 8(5), 238-244.

Av-Gay Y., S. Jamil, and S. J. Drews (1999) "Expression and Characterization of the *Mycobacterium tuberculosis* Serine/Threonine Protein Kinase PknB," Infect. Immun. 67(11), 5676-5682.

Barford, D. (1996) "Molecular mechanisms of the protein serine/threonine phosphatases," Trends Biochem. Sci. 21(11), 407-412.

Barton, G. J., P. T. Cohen and D. Bradford (1994) "Conservation analysis and structure prediction of the protein serine/threonine phosphatases. Sequence similarity with diadenosine tetraphosphatase from *Escherichia coli* suggests homology to the protein phosphatases," Eur. J. Biochem. 220, 225-237.

Becker W, H.G. Joost (1999) "Structural and functional characteristics of Dyrk, a novel subfamily of protein kinases with dual specificity," Prog. Nucleic Acid Res. Mol. Biol. 62, 1-17.

Bell S.D., P. L. Kosa, P. B. Sigler, S. P. Jackson (1999) "Orientation of the transcription preinitiation complex in Archaea," Proc. Natl. Acad. Sci. USA 96 (24), 13662-13667.

Biemann K. and H. A. Characterization by Tandom Mass Spectrometry of Structural Modifications in Proteins," Science, 239, 992-995.

Bischoff, K. M. and P. J. Kennelly (1999) "In-Gel Assay for Identifying Alternative Nucleotide Substrates for Protein Kinases," Anal. Biochem. 271, 199-202.

Bradford, M. M. (1976) " A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding," Analytical Biochemistry 72, 248-254.

Bray, D. (1995) "Protein molecules as computational elements in living cells," Nature 376, 307-312.

Brenner, S. (1987) "Phosphotransferase sequence homology," Nature 329, 21.

Brochu D. and C. Vadeboncoeur (1999) "The HPr (Ser) kinase of *Streptococcus salivarius*: purification, properties, and cloning of the hprK gene," J. Bacteriol. 181, 709-717.

Brock, T. D., K. M. Brock, R. T. Belly and R. L. Weiss (1972) "Sulfolobus: A New Genus of Sulfur-Oxidizing Bacteria Living at Low pH and High Temperature," Arch. Microbiol. 84, 54-68.

Bröckl G., M. Behr, S. Fabry, R. Hensel, H. Kaudewitz, and E. Biendl (1991) "Analysis and nucleotide sequence of the genes encoding the surface-layer glycoproteins of the hyperthermophilic methanogens *Methanothermus fervidus* and *Methanothermus sociabili*s, Eur. J. Biochem. 199, 147–152.

Brown, J. M. and R. A. Firtel (1998) "Phosphorelay signalling: new tricks for an ancient pathway," Curr. Biol. 8, R662-665.

Buchanan C. L., H. Connaris, M. J. Danson, C. D. Reeve, D. W. Hough (1999) "An extremely thermostable aldolase from Sulfolobus solfataricus with specificity for non-phosphorylated substrates," Biochem. J. 343 Pt 3, 563-570.

Budde R. J., S. M. Ernst, and R. Chollet (1986) "Substrate specificity and regulation of the maize (Zea mays) leaf ADP: protein phosphotransferase catalysing phosphorylation / inactivation of pyruvate, orthophosphate dikinase," Biochem. J. 236(2), 579-584.

Bult C. J., O. White, G. J. Olsen, L. Zhou, R. D. Fleischmann, G. G. Sutton, J. A. Blake, L. M. FitzGerald, R. A. Clayton, J. D. Gocayne, A. R. Kerlavage, B. A. Dougherty, J. F. Tomb, M. D. Adams, C. I. Reich, R. Overbeek, E. F. Kirkness, K. G. Weinstock, J. M.

Merrick, A. Glodek, J. L. Scott, N. S. Geoghagen, and J. C. Venter (1996) "Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*," Science 1996 273(5278), 1058-1073.

Calera, J. A., G. H. Choi and R. A. Calderone (1998) "Identification of a putative histidine kinase two-component phosphorelay gene (CaHK1) in Candida albicans," Yeast 14, 665-674.

Cann I. K. and Y. Ishino (1999) "Archaeal DNA replication: identifying the pieces to solve a puzzle," Genetics 152(4), 1249-1267.

Cartier, C., P. Sivard, C. Tranchat, D. Decimo, C. Desgranges and V. Boyer (1999) "Identification of three major phosphorylation sites within HIV-1 capsid. Role of phosphorylation during the early steps of infection," J. Biol. Chem. 274, 19434-19440.

Chang, C., S. F. Kwok, A. B. Bleecker, A. B. and E. M. Meyerowitz (1993) "*Arabidopsis* ethylene-response gene ETR-1: similarity of product to two-component regulators," Science 262, 539-544.

Cohen, P. (1982) "The role of protein phosphorylation in neural and hormonal control of cellular activity," Nature 296, 613-620.

Cohen, P. (1988) "Protein phosphorylation and hormone action," Proc. Royal Soc. London, Series B 234, 115-144.

Cohen, P. (1991) "Classification of Protein-Serine/Threonine Phosphatases: Identification and Quantitation in Cell Extracts," Methods Enzymol. 201, 389-398.

Corbin, J. D. and E. M. Reimann (1974) "Assay of cyclic AMP-dependent protein kinases," Methods Enzymol. 38, 287-290.

Cortay, J. C., F. Bleicher, C. Rieul, H. C. Reeves, and A. J. Cortay (1988) "Nucleotide sequence and expression of the *aceK* gene coding for isocitrate dehydrogenase kinase/phosphatase in *Escherichia coli*," J. Bacteriol. 170, 89-97.

Cozzone, A. J. (1998) "Post-translational modification of proteins by reversible phosphorylation in prokaryotes," Biochimie. 80, 43-48.

Cserzo M., E. Wallin, I. Simon, G. von Heijne, and A. Elofsson (1997) "Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the Dense Alignment Surface method," Prot. Eng. 10 (6), 673-676.

Cyert, M. S. and J. Thorner (1989) "Putting it on and taking it off: phosphoprotein phosphatase involvement in cell cycle regulation," Cell 57, 891-893.

Dalgaard J.Z. and R. A. Garrett (1992) “Protein-coding introns from the 23S rRNA-encoding gene form stable circles in the hyperthermophilic archaeon Pyrobaculum organotrophum,” Gene 121(1), 103-110.

Danon A., S. P. Mayfield (1994) “ADP-dependent phosphorylation regulates RNA-binding in vitro: implications in light-modulated translation,” EMBO J. 13(9), 2227-2235.

De Lanerolle, P. and R. J. Paul (1991) “Myosin phosphorylation/dephosphorylation and regulation of airway smooth muscle contraction,” Am. J. Physiol. 261, 1-14.

DeLong, E. (1998) “Archaeal Means and Extremes,” Science 280, 542-543.

De Rosa, M, A. Gambacorta, and J. D. Bullock (1975) “Extremely thermophilic acidophilic bacteria convergent with *Sulfolobus acidocaldarius*,” J. Gen. Microbiol. 86, 156-164.

Doolittle, R. F. (1995) “Of Archae and Eo: What’s in a name?,” Proc. Natl. Acad. Sci. USA 92, 2421-2423.

Doolittle, W. F. (1999) “Phylogenetic classification and the universal tree,” Science 284(5423), 2124-2129.

Doonan, S. (1996) “Protein Purification Protocols,” Methods I Molecular Biology, Vol 59, Humana Press Inc., Totowa, New Jersey.

Duclos, B., S. Marcandier and A. J. Cozzone (1991) "Chemical properties and separation of phosphoamino acids by thin-layer chromatography and/or electrophoresis," Meth. Enzym. 201, 10-21.

Duclos B., E. Vaganay, M. Dadssi, and A. J. Cozzone (1996) “Pyrophosphate is a source of phosphoryl groups for Escherichia coli protein phosphorylation,” FEMS Microbiol. Lett. 145(1), 49-54.

Eckstein, J. W., P. Beer-Romero, and I. Berdo (1996) “Identification of an essential acidic residue in Cdc25 protein phosphatase and a general three dimensional model for a core region in protein phosphatases,” Protein Sci. 5, 5-12.

Eichinger L., L. Bomblies, J. Vandekerckhove, M. Schleicher and J. Gettemans (1996) “A novel type of protein kinase phosphorylates actin in the actin-fragment complex,” EMBO J.15, 5547-5556

Erikson, R. L. (1991) “Structure, expression, and regulation of protein kinases involved the phosphorylation of ribosomal protein S6,” J. Biol. Chem. 266, 6007-6010.

Fairbanks, G., T. L. Steck, and D. F. H. Wallach (1971) " Electrophoretic transfer of the major polypeptides of the human erythrocyte membrane," Biochemistry 10, 2606-2617.

Feng D., G. Cho, and R. F. Doolittle (1997) "Determining divergence times with a protein clock: Update and reevaluation," Proc. Natl. Acad. Sci. USA 94, 13028-13033.

Forchhammer, K. and Marsac N. Tandeau de (1995) "Phosphorylation of the PII protein (glnB gene product) in the cyanobacterium *Synechococcus* sp. Strain PCC 7942: analysis of *in vitro* kinase activity," J. Bacteriol. 177, 5812-5817.

Forterre, P. (1997) "Archaea: what can we learn from their sequences?," Curr. Opin. Genet. Dev. 7, 764-770.

Frasch, S. C. and M. Dworkin (1996) "Tyrosine phosphorylation in *Myxococcus xanthus*, a multicellular prokaryote," J. Bacteriol. 178, 4084-4088.

Futey L. M., Q. G. Medley, and P. C. Graham (1999) "Structural Analysis of Myosin Heavy Chain Kinase A from *Dictyostelium*," J. Biol. Chem. 270 (2), 523-529.

Gaidenko, T. A., X. Yang, Y. M. Lee and C. W. Price (1999) "Threonine phosphorylation of modulator protein RsbR governs its ability to regulate a serine kinase in the environmental stress signaling pathway of *Bacillus subtilis*," J. Mol. Biol. 288, 29-39.

Galinier A, M. Kravanja, R. Engelmann, W. Hengstenberg, M. C. Kilhoffer, J. Deutscher, and J. Haiech (1998) "New protein kinase and protein phosphatase families mediate signal transduction in bacterial catabolite repression," Proc. Natl. Acad. Sci. U S A 95(4), 1823-1828.

Glover, D. M. and B. D. Hames (1995) DNA Cloning 2, Expression Systems, 2nd ed. (Glover and Hames eds.) Oxford University Press, New York.

Gogarten-Boekels M., E. Hilario, J. P. Gogarten (1995) "The effects of heavy meteorite bombardment on the early evolution--the emergence of the three domains of life," Orig. Life Evol. Biosph. 25(1-3): 251-264.

Gonzalez L., V. Phalip, and C. C. Zhang (2001) "Characterization of PknC, a Ser/Thr kinase with broad substrate specificity from the cyanobacterium Anabaena sp. strain PCC 7120," Eur. J. Biochem. 268(6), 1869-1875.

Gorg, A, W. Postel, S. Gunther, J. Weser (1985) "Improved horizontal two-dimensional electrophoresis with hybrid isoelectric focusing in immobilized pH gradients in the first dimension and laying-on transfer to the second dimension," Electrophoresis 6, 599-604.

Gorg, A., W. Postel, S. Gunther (1988) "The current state of two-dimensional electrophoresis with immobilized pH gradients," Electrophoresis 9, 531-546.

Grangeasse, C., P. Doublet, C. Vincent, E. Vaganay, M. Riberty, B. Duclos and A. J. Cozzone (1998) "Functional characterization of the low-molecular-mass phosphotyrosine-protein phosphatase of *Acinetobacter johnsonii*," J. Mol. Biol. 278, 339-347.

Greengard, P. (1978) "Phosphorylated proteins as physiological effectors," Science 199, 146-152.

Gribskov, M. and R. R. Burgess (1983) "Overexpression and purification of the sigma subunit of *Escherichia coli* RNA polymerase," Gene 26, 109-118.

Grose, C. (1990) "Glycoproteins encoded by varicella-zoster virus: biosynthesis, phosphorylation, and intracellular trafficking," Annu. Rev. Microbiol. 44, 59-80.

Guan, K. and J. E. Dixon (1991) "Evidence for protein-tyrosine phosphatase catalysis proceeding via a cysteine phosphate intermediate," J. Biol. Chem. 266, 17026-17030.

Gustin M. C., A. Jacobus, M. Alexander, and K. Davenport (1998) "MAP Kinase Pathways in the Yeast *Saccharomyces cerevisiae*," Microbiol. Mol. Biol. Rev. 62(4), 1264-1300.

Hakenbeck R., T. Grebe, D. Zahner, J. B. Stock (1999) "beta-lactam resistance in Streptococcus pneumoniae: penicillin-binding proteins and non-penicillin-binding proteins," Mol. Microbiol. 33(4), 673-678.

Hallenbeck, P. C. and D. A. Walsh (1983) "Autophosphorylation of Phosphorylase Kinase," J. Biol. Chem. 258, 13493-13501.

Hanks, S. K. and T. Hunter (1995) "The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification," FASEB J. 9, 576-596.

Hanzelka B. L., T. J. Darcy, and J. N. Reeve (2001) "TFE, an archaeal transcription factor in Methanobacterium thermoautotrophicum related to eucaryal transcription factor TFIIE alpha," J. Bacteriol. 183(5), 1813-1818.

Harris R. A., K. M. Popov, Y. Shimomura, Y. Zhao, J. Jaskiewicz, N. Nanaumi, and M. Suzuki (1992) "Purification, characterization, regulation and molecular cloning of mitochondrial protein kinases," Adv. Enzyme Regul. 32, 267-284.

Hawdon, J. M., B. F. Jones, D. R. Hoffman and P. J. Hotez (1996) "Cloning and Characterization of *Ancylostoma*-secreted Protein," J. Biol. Chem. 271, 6672-6678.

Hemmings H. C. Jr. (1997) "Regulatory Protein Modification," Neuromethods 30, Humana Press, New York, New York.

Herskowitz, I. (1995) “MAP kinase pathways in yeast: For mating and more,” Cell 80, 187-197.

Hidaka, H. and R. Kobayashi (1992) “Pharmacology of protein kinase inhibitors,” Annu. Rev. Pharmacol. Toxicol. 32, 377-397.

Hirano, H., S. Komatsu, H. Takakura, F. Sakiyama and S. Tsunasawa (1992) “Deblocking and Subsequent Microsequence Analysis of N-Blocked Proteins Electroblotted onto PVDF Membrane,” J. Biochem. 111, 754-757.

Ho, S. N., H. D. Hunt, R. M. Horton, J. K. Pullen and L. R. Pease (1989) "Site-directed mutagenesis by overlap extension using PCR," Gene 77, 51-59.

Hoekstra, M. F., A. J. De Maggio and N. Dhillon (1991) “Genetically identified protein kinases in yeast. II:DNA metabolism and meiosis,” Trends Genet 9, 293-297.

Hunt, D. F., J. R. Yates III, J. Shabanowitz, S. Winston and C. R. Hauer (1986) “Protein Sequencing by Tandem Mass Spectrometry,” Proc. Natl. Acad. Sci. USA 83, 6233-6238.

Hunter, T. (1995) “Protein Kinases and Phosphatases: The Yin and Yang of Protein Phosphorylation and Signaling,” Cell 80, 225-236.

Hunter, T and G. D. Plowman (1997) “The protein kinases of budding yeast: six score and more,” Trends Biochem. Sci. 22(1), 18-22.

Hunter, T. (2000) “Signaling-2000 and Beyond,” Cell 100, 113-127.

Ichihara Y. and Y. Kurosawa (1993) “Construction of new T vectors for direct cloning of PCR products,” Gene 130, 153-154.

Ingebritsen, T. S. and P. Cohen (1983) “The protein phosphatases involved in cellular regulation 1. Classification and substrate specificities,” Eur. J. Biochem 132, 25-261.

Inouye S, R. Jain, T. Ueki, H. Nariya, C. Y. Xu, M. Y. Hsu, B. A. Fernandez-Luque, J. Munoz-Dorado, E. Farez-Vidal, M. Inouye (2000) “A large family of eukaryotic-like protein Ser/Thr kinases of Myxococcus xanthus, a developmental bacterium,” Microb. Comp. Genomics 5(2), 103-120.

Itoh T., K. Suzuki, T. Nakase (1998) “Occurrence of introns in the 16S rRNA genes of members of the genus *Thermoproteus*, Arch. Microbiol. 170(3), 155-161.

Jakobi R. and J. A. Traugh (1992) “Characterization of the phosphotransferase domain of casein kinase II by site-directed mutagenesis and expression in Escherichia coli,” J. Biol. Chem. 267(33), 23894-23902.

Janiak-Spens, F., J. M. Sparling, M. Gurfinkel and A. H. West (1999) "Differential stabilities of phosphorylated response regulator domains reflect functional roles of the yeast osmoregulatory SLN1 and SSK1 proteins," J. Bacteriol. 181, 411-417.

Johnson, L. N. and D. Barford (1993) "The Effects of Phosphorylation on the Structure and Function of Proteins," Annu. Rev. Biophys. Biomol. Struct. 22, 199-232.

Jones, J. I., W. H. Busby, Jr., G. Wright, C. E. Smith, N. M. Kimack and D. R. Clemmons (1993) "Identification of the Sites of Phosphorylation in Insulin-like Growth Factor Binding Protein-1," J. Biol. Chem. 268, 1125-1131.

Julien, J. P. and W. E. Mushynski (1998) "Neurofilaments in health and disease," Prog. Nucleic Acid Res. Mol. Biol. 61, 1-23.

Kakimoto, T. (1996) "CKl1, a Histidine Kinase Homolog Implicated in Cytokinin Signal Transduction," Science 274, 982-985.

Kamei A., T. Yuasa, K. Orikawa, X. X. Geng, and M. Ikeuchi (2001) "A eukaryotic-type protein kinase, SpkA, is required for normal motility of the unicellular Cyanobacterium *Synechocystis* sp. strain PCC 6803," J. Bacteriol. 183(5), 1505-1510.

Kamps, M. P. and B. M. Sefton (1989) "Acid and base hydrolysis of phosphoproteins bound to Immobilon facilitates analysis of phosphoamino acids in gel-fractionated proteins," Anal. Biochem. 176, 22-27.

Kaufman, R. J. (1997) "DNA transfection to study translational control in mammalian cells," Methods 4, 361-370.

Kawamoto S, Hidaka (1984) "H 1-(5-Isoquinolinesulfonyl)-2-methylpiperazine (H-7) is a selective inhibitor of protein kinase C in rabbit platelets,".Biochem. Biophys. Res. Commun. 125, 258-264.

Keeling, P. J. and W. F. Doolittle (1995) "Archaea: Narrowing the gap between prokaryotes and eukaryotes," Proc. Natl. Acad. Sci. USA 92, 5761-5764.

Kennelly, P. J. (1991) "New perspectives on the role of protein phosphorylation in the regulation of 3-hydroxy-3-methylglutaryl coenzyme A reductase," Adv. Lipids Res. 1, 19-26.

Kennelly, P. J. (1994) "Identification of Sites of Serine and Threonine Phosphorylation via Site-Directed Mutagenesis-Site Transformation versus Site Elimination," Anal. Biochem. 219, 384-386.

Kennelly, P. J. (1999) "Bits for an organic microprocessor: Protein phosphorylation/dephosphorylation," In: Introduction to Cellular Signal Transduction (Sitaramayya, A., ed.), pp. 235-263, Birkhauser, Boston.

Kennelly, P. J., A. M. Edelman, D. K. Blumenthal, and E. G. Krebs (1987) “Rabbit skeletal muscle myosin light chain kinase: the calmodulin-binding domain as a potential active site-directed inhibitory domain,” J. Biol. Chem. 262, 11958-11963.

Kennelly, P. J., K. A. Oxenrider, J. Leng, J. S. Cantwell and N. Zhao (1993) “Identification of a Serine/Threonine-specific Protein Phosphatase from the Archaebacterium *Sulfolobus solfataricus*," J. Biol. Chem. 268, 6505-6510.

Kennelly, P. J. and K. M. Bischoff (1999) “An “in gel” assay for alternative nucleotide donors for protein kinases,” Anal. Biochem. 271, 199-202.

Kennelly, P. J. and M. Potts (1996) “Fancy Meeting You Here! A Fresh Look at “Prokaryotic Protein Phosphorylation,” J. Bacteriol. 178, 4759-4764.

Kennelly, P. J. and M. Potts (1999) “Life Among the Primitives: Protein O-Phosphatases in Prokaryotes,” Frontiers Biosci. 4, d372-d385.

Kim D. J. and S. Forst (2001) “Genomic analysis of the histidine kinase family in bacteria and archaea,” Microbiology 147(Pt 5), 1197-1212.

Kim M., D. A. Christopher, and J. E. Mullet (1999) “ADP-Dependent phosphorylation regulates association of a DNA-binding complex with the barley chloroplast psbD blue-light-responsive promoter,” Plant Physiol. 119(2), 663-670.

Kimura Y., H. Nakano, H. Terasaka, and K. Takegawa (2001) “*Myxococcus xanthus* mokA encodes a histidine kinase-response regulator hybrid sensor required for development and osmotic tolerance,” J. Bacteriol. 183(4), 1140-1146.

Klenk, H. P., R. A. Clayton, J. F. Tomb, O. White, K. E. Nelson, K. A. Ketchum, R. J. Dodson, M. Gwinn, E. K. Hickey, J. D. Peterson, D. L. Richardson, A. R. Kerlavage, D. E. Graham, N. C. Kyrpides, R. D. Fleischmann, J. Quackenbush, N. H. Lee, G. G. Sutton, S. Gill, E. F. Kirkness, B. A. Dougherty, K. McKenney, M. D. Adams, B. Loftus, S. Peterson, C. I. Reich, L. K. McNeil, J. H. Badger, A. Glodek, L. Zhou, R. Overbeek, J. D. Gocayne, J. F. Weidman, L. McDonald, T. Utterback, M. D. Cotton, T. Spriggs, P. Artiach, B. P. Kaine, S. M. Sykes, P. W. Sadow, K. P. D'Andrea, C. Bowman, C. Fujii, S. A. Garland, T. M. Mason, G. J. Olsen, C. M. Fraser, H. O. Smith, C. R. Woese, and J. C. Venter (1997) "The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon *Archaeoglobus fulgidus*," Nature 390, 364-370.

Klumpp, D. J., D. W. Plank, L. J. Bowdin, C. S. Stueland, T. Chung, and D. C. La Porte (1988) “Nucleotide sequence of *aceK*, the gene encoding isocitrate dehydrogenase kinase/phosphatase,” J. Bacteriol. 170, 2763-2769.

Kravanja M., R. Engelmann, V. Dossonnet, M. Bluggel, H. E. Meyer, R. Frank, A. Galinier, J. Deutscher, N. Schnell and W. Hengstenberg (1999) “The hprK gene of

Enterococcus faecalis encodes a novel bifunctional enzyme: the HPr kinase/phosphatase," Mol. Microbiol. 31, 59-66.

Krebs, E. G. and E. H. Fisher (1956) "The phosphorylase b to a converting enzyme of rabbit skeletal muscle," Biochim. Biophys. Acta 20, 150-157.

Kwak, J. H. and M. Y. Kim (1995) "Construction of T-Vector for Direct Cloning and Expression of Cloned Genes in *Escherichia coli*," Anal. Biochem. 228, 178-180.

Laemmli, U. K. (1970) "Cleavage of structural proteins during the assembly of bacteriophage t4," Nature 227, 680-685.

LaPorte, D. C. and D. E. Koshland Jr. (1983) "A protein with kinase and phosphatase activities involved in regulation of the tricarboxylic acid cycle," Nature 300, 458-460.

LaPorte, D. C. and T. Chung (1985) "A single gene encodes for the kinase and phosphatase which regulate isocitrate dehydrogenase," J. Biol. Chem. 260, 15291-15297.

Laurent E., M. Talpaz, H. Kantarjian, R. Kurzrock (2001) "The BCR gene and philadelphia chromosome-positive leukemogenesis," Cancer Res. 61(6), 2343-55.

LaVallie, E. R., E. A. DiBlasio, S. Kovacic, K. L. Grant, P. F. Schendel, and J. M. McCoy (1993) "A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm," BioTechnology 11, 187-193.

Lechner J. and F. Wieland (1989) "Structure and biosynthesis of prokaryotic glycoproteins," Annu. Rev. Biochem. 58, 173–194.

Lee J. H., S. K. Choi, A. Roll-Mecak, S. K. Burley, and T. E. Dever (1997) "Universal conservation in translation initiation revealed by human and archaeal homologs of bacterial translation initiation factor IF2," Proc. Natl. Acad. Sci. USA 96(8), 4342-4347.

Leng, J, A. J. Cameron, S. Buckel and P. J. Kennelly (1995) "Isolation and Cloning of a Protein-Serine/Threonine Phosphatase from an Archaeon," J. Bacteriol. 177, 6510-6517.

Leonard, C. J., L Aravind and E. V. Koonin (1998) "Novel families of putative protein kinases in bacteria and archaea: evolution of the "eukaryotic" protein kinase superfamily," Genome Res. 8, 1038-1047.

Levin D. E. and B. Errede (1995) "The proliferation of MAP kinase signaling pathways in yeast," Curr. Opin. Cell Biol. 7(2), 197-202.

Lin, B., S. K. Hollingshead, J. E. Coligan, M. L. Egan, J. R. Baker and D. G. Pritchard (1994) "Cloning and Expression of the Gene for Group B Streptococcal Hyaluronate Lyase," J. Biol. Chem. 269, 30113-30116.

Lindberg R. A., W. H. Fischer, and T. Hunter (1993) "Characterization of a human protein threonine kinase isolated by screening an expression library with antibodies to phosphotyrosine," Oncogene 8(2), 351-359.

Link, A. (1999) "2-D Proteome Analysis Protocols," Humana Press Inc., Totowa, New Jersey.

Locht, C. (1999) "Molecular aspects of Bordetella pertussis pathogenesis," Int. Microbiol. 2(3), 137-144.

Lower, B. H., P. J. Kennelly, and K. M. Bischoff (2000) "The Archaeon *Sulfolobus solfataricus* Contains a Membrane-Associated Protein Kinase Activity That Preferentially Phosphorylates Threonine Residues In Vitro," J. Bacteriol 182 (12), 3452-3459.

Madigan, M. T., J. M. Martinko, and J. Parker (1997) "Brock Biology of Microorganisms," Prentice Hall, Inc., Upper Saddle River, New Jersey, USA.

Maeda, T., Wurgler-Murphy, S. M. and H. Saito (1994) "A two-component system that regulates an osmosensing MAP kinase cascade in yeast," Nature 369, 242-245.

Mai, B., G. Frey, R. V. Swanson, E. J. Mathur and K. O. Stetter (1998) "Molecular cloning and functional expression of a protein-serine/threonine phosphatase from the hyperthermophilic archaeon *Pyrodictium abyssi* TAG1," J. Bacteriol. 180, 4030-4035.

Mann M., P. Hojrup, and P. Roepstorff (1993) "Use of mass spectrometric molecular weight information to identify proteins in sequence databases," Biol. Mass Spectrom. 22(6), 338-345.

Martin P., E. Jullien, and P. Courvalin (1988) "Nucleotide sequence of Acinetobacter baumannii aphA-6 gene: evolutionary and functional implications of sequence homologies with nucleotide-binding proteins, kinases and other aminoglycoside-modifying enzymes," Mol. Microbiol. 2(5), 615-625.

Maru Y. and O.N. Witte (1991) "The BCR gene encodes a novel serine/threonine kinase activity within a single exon," Cell 67(3), 459-468.

Mathews, M. B. (1990) "Control of translation in adenovirus-infected cells," Enzyme 44, 250-264.

Matsudaira, P. T. (1993) "A practical guide to protein and peptide purification for microsequencing (2nd ed.)," In: Academic Press, San Diego, CA.

Medley, Q. G., W. L. Bagshaw, T. Truong and G. P. Cote (1992) "Dictyostelium myosin II heavy-chain kinase A is activated by heparin, DNA and acidic phospholipids and inhibited by polylysine, polyarginine and histones," Biochimica et Biophysica Acta. 1175, 7-12.

Mendenhall, M. D. and A. E. Hodge (1998) “Regulation of Cdc28 cyclin-dependent protein kinase activity during the cell cycle of the yeast *Saccharomyces cerevisiae*,” Microbiol. Mol. Biol. Rev. 62, 1191-243.

Miller S. P., E. J. Karschnia, T. P. Ikeda, and D. C. LaPorte (1996) “Isocitrate dehydrogenase kinase/phosphatase. Kinetic characteristics of the wild type and two mutant proteins,” J. Biol. Chem. 271(32), 19124-19128.

Migita K, Eguchi K, Kawabe Y, Mizokami A, Tsukada T, Nagataki S (1994) "Prevention of anti-CD3 monoclonal antibody-induced thymic apoptosis by protein tyrosine kinase inhibitors," J. Immunol. 153, 3457-3465.

Moens S. and J. Vanderleyden (1997) “Glycoproteins in prokaryotes,” Arch. Microbiol. 168, 169–175.

Morell V. (1997) “Microbial Biology: Microbiology's Scarred Revolutionary,” *Science* 276, 699-702.

Motley, S. T. and S. Lory (1999) “Functional Characterization of a Serine/Threonine Protein Kinase of *Pseudomonas aeruginosa*,” Infect. Immun. 67, 5386-5394.

Novak R., E. Charpentier, J. S. Braun, E. Tuomanen (2000) “Signal transduction by a death signal peptide: uncovering the mechanism of bacterial killing by penicillin,” Mol. Cell 5(1), 49-57.

Oda Y., M. D. Kuo, S. S. Huang, and J. S. Huang (1991) “The plasma cell membrane glycoprotein, PC-1, is a threonine-specific protein kinase stimulated by acidic fibroblast growth factor,” J. Biol. Chem. 266(25), 16791-16795.

Oda Y., K. Huang, F. R. Cross, D. Cowburn, and B. T. Chait (1999) “Accurate quantitation of protein expression and site-specific phosphorylation,” Proc. Natl. Acad. Sci. USA 96(12), 6591-6596.

Ostrovsky, P. C. and S. Maloy (1995) “Protein phosphorylation on serine, threonine, and tyrosine residues modulates membrane-protein interactions and transcriptional regulation in Salmonella typhimurium,” Genes Dev. 9, 2034-041.

Ota, I. M. and A. Varshavsky (1993) “A yeast protein similar to bacterial two-component regulators,” Science 262, 566-569.

Pain, V. M. (1994) “Translational control during amino acid starvation,” Biochimie. 8, 718-728.

Parkinson, J. S. (1993) “Signal transduction scheme of bacteria,” Cell 73, 857-871.

Pavur K.S., A. N. Petrov, and A. G. Ryazanov (2000) "Mapping the functional domains of elongation factor-2 kinase," Biochemistry 39(40), 12216-12224.

Perego M., C. Hanstein, K. M. Welsh, T. Djavakhishvili, P. Glaser, and J.A. Hoch (1994) "Multiple protein-aspartate phosphatases provide a mechanism for the integration of diverse signals in the control of development in B. subtilis," Cell 79(6), 1047-1055.

Perego M. and J. A. Hoch (1996) "Protein aspartate phosphatases control the output of two-component signal transduction systems," Trends Genet. 12(3), 97-101.

Patterson S.D. (1995) "Matrix-assisted laser-desorption/ionization mass spectrometric approaches for the identification of gel-separated proteins in the 5-50 pmol range," Electrophoresis 16(7), 1104-1114.

Ponting C. P., L. Aravind, J. Schultz, P. Bork, E. V. Koonin (1999) "Eukaryotic signalling domain homologues in archaea and bacteria. Ancient ancestry and horizontal gene transfer," J. Mol. Biol. 289(4), 729-745.

Potts, M., H. Sun, K. Mockaitis, P. J. Kennelly, D. Reed and N. K. Tonks (1993) "A Protein-tyrosine/serine Phosphatase Encoded by the Genome of the Cyanobacterium Nostoc commune UTEX 584," J. Biol. Chem. 268, 7632-7635.

Ramwani, J. and R. K. Mishra (19860 "Purification of bovine striatal dopamine D2 receptor by affinity chromatography," J. Biol. Chem. 261, 8894-8898.

Rasmussen B. (2000) "Filamentous microfossils in a 3,235-million-year-old volcanogenic massive sulphide deposit," Nature 405, 676 – 679.

Raya A., F. Revert, S. Navarro, and J. Saus (1999) "Characterization of a novel type of serine/threonine kinase that specifically phosphorylates the human goodpasture antigen," J. Biol. Chem. 274(18), 12642-12649.

Reddy T.R. and T. Suryanarayana (1989) "Archaebacterial histone-like proteins. Purification and characterization of helix stabilizing DNA binding proteins from the acidothermophile Sulfolobus acidocaldarius," J. Biol. Chem. 264(29), 17298-17308.

Reizer J., C. Hoischen, F. Titgemeyer, C. Rivolta, R. Rabus, J. Stulke, D. Karamata, M. H. Saier Jr, and W. Hillen (1998) "A novel protein kinase that controls carbon catabolite repression in bacteria" Mol. Microbiol. 27(6), 1157-1169.

Rivera M. C., R. Jain, J. E. Moore, and J. A. Lake (1998) "Genomic evidence for two functionally distinct gene classes," Genetics 95 (11), 6239-6244.

Robb, R. T. and A. R. Place (1995) "Archaea a Laboratory Manual Thermophiles," Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Roux, K. H. (1994) "Using mismatched primer-template pairs in touchdown PCR," BioTechniques 16, 812-814.

Rudolph, J. and D. Oesterhelt (1995) “Chemotaxis and phototaxis require a CheA histidine kinase in the archaeon *Halobacterium salinarium*,” EMBO J. 14, 667-673.

Rudolph, J., N. Tolliday, C. Schmitt, S. C. Schuster and D. Oesterhelt (1995) “Phosphorylation in halobacterial signal transduction,” EMBO J. 14, 4249-4257.

Ryazanov AG, M. D. Ward, C. E. Mendola, K. S. Pavur, M. V. Dorovkov, M. Wiedmann, H. Erdjument-Bromage, P. Tempst, T. G. Parmer, C. R. Prostko, F. J. Germino, and W. N. Hait (1997) “Identification of a new class of protein kinases represented by eukaryotic elongation factor-2 kinase,” Proc. Natl. Acad. Sci. U S A 94(10), 4884-4889.

Saitoh M, Ishikawa T, Matsushima S, Naka M, Hidaka H, (1987) "Selective inhibition of catalytic activity of smooth muscle myosin light chain kinase," J. Biol. Chem. 262, 7796-7801.

Sambrook, J., E. F. Fritsch and T. Maniatis (1989) "Molecular Cloning, A Laboratory Manual (2nd ed.), Cold Spring Harbor Laboratories, Cold Spring Harbor, NY.

Sanger, F., S. Nicklen and A. R. Russell (1977) “DNA sequencing with chain-terminating inhibitors,” Proc. Natl. Acad. Sci. USA 74, 5463-5467.

Schafer G. (1999) “How can archaea cope with extreme acidity?,” Novartis Found. Symp. 221:131-44, discussion 145-151.

Schinkmann K. and J. Blenis (1997) “Cloning and characterization of a human STE20-like protein kinase with unusual cofactor requirements,” J. Biol. Chem. 272(45), 28695-28703.

Schleper, C., K. Kubo and W. Zillig (1992) “The particle SSV1 from the extremely thermophilic archaeon *Sulfolobus* is a virus: demonstration of infectivity and of transfection with viral DNA,” Proc. Natl. Acad. Sci. USA 89, 7645-7649.

Schuster M., R. E. Silversmith, and R. B. Bourret (2001) “Conformational coupling in the chemotaxis response regulator CheY,” Proc. Natl. Acad. Sci. U.S.A. 98 (11), 6003-6008.

Selkirk, J. K., C. He, R. M. Patterson and B. A. Merrick (1996) “Tumor suppressor p53 gene forms multiple isoforms: evidence for single locus origin and cytoplasmic complex formation with heat shock proteins,” Electrophoresis 11, 1764-1771.

Sellers, J. R. and R. S. Adelstein (1987) “Regulation of contractile activity,” Enzymes 18, 381-418.

She Q., R. K. Singh, F. Confalonieri, Y. Zivanovic, G. Allard, M. J. Awayez, C. .C Chan-Weiher, I. G. Clausen, B. A. Curtis, A. De Moors, G. Erauso, C. Fletcher, P. M. Gordon, I. Heikamp-De Jong , A. C. Jeffries, C. J. Kozera, N. Medina, X. Peng, H. P. Thi-Ngoc, P. Redder, M. E. Schenk, C. Theriault, N. Tolstrup, R. L. Charlebois, W. F. Doolittle, M. Duguet, T. Gaasterland, R. A. Garrett, M. A. Ragan, C. W. Sensen, J. Van Der Oost (2000) "The complete genome of the crenarchaeon Sulfolobus solfataricus P2," Proc. Natl. Acad. Sci. USA 98 (14).

Sheppard, D. N. and M. J. Welsh (1999) "Structure and function of the CFTR chloride channel,"Physiol. Rev. 79, 23-45.

Shi, L., P. J. Kennelly and M. Potts (1998) "The serine, threonine, and/or tyrosine-specific protein kinases and protein phosphatases of prokaryotic organisms: a family portrait," FEMS Microbiol. Rev. 22, 229-253.

Shi, L., K. M. Bischoff and P. J. Kennelly (1999) "The icfG Gene Cluster of Synechocystis sp. Strain PCC 6803 Encodes an Rsb/Spo-Like Protein Kinase, Protein Phosphatase, and Two Phosphoproteins," J. Bacteriol. 181, 4761-4767.

Siebers B., H. P. Klenk, and R. Hensel (1998) "PPi-dependent phosphofructokinase from Thermoproteus tenax, an archaeal descendant of an ancient line in phosphofructokinase evolution," J. Bacteriol. 180(8), 2137-2143.

Sjoholm A. (1998) "Aspects of novel sites of regulation of the insulin stimulus-secretion coupling in normal and diabetic pancreatic islets," Endocrine 1, 1-13.

Skorko, R. (1984) "Protein phosphorylation in the archaebacterium *Sulfolobus acidocaldarius*," Eur. J. Biochem. 145, 617-622.

Skorko, R. (1989) "Polyphosphate as a source of phosphoryl group in protein modification in the archaebacterium *Sulfolobus acidocaldarius*," Biochimie 71, 1089-1093.

Smith C.M., I. N. Shindyalov, S. Veretnik, M. Gribskov, S. S. Taylor, L. F. Ten Eyck, and P. E. Bourne (1997) "The protein kinase resource," Trends Biochem. Sci. 22(11), 444-446.

Smith J. A., S. H. Francis SH and J. D. Corbin (1993) "Autophosphorylation: a salient feature of protein kinases," Mol. Cell Biochem. 127-128, 51-70.

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

Smith, R. F. and K. Y. King (1995) "Identification of a eukaryotic-like protein kinase gene in Archaebacteria," Protein Science 4, 126-129.

Smith, R. S. et. al. (1997) "Complete Genome Sequence of *Methanobacterium thermoautotrophicum* ΔH: Functional Analysis and Comparative Genomics," J. Bacteriol. 179, 7135-7155.

Sobel, A and A. H. Tashjian Jr. (1983) "Distinct patterns of cytoplasmic protein phosphorylation related to regulation of synthesis and release of prolactin by GH cells," J. Biol. Chem. 258, 10312-10324.

Soderling, T. R. (1993) "Calcium/calmodulin-dependent protein kinase II: role in learning and memory," Mo. Cell Biochem. 127, 93-101.

Solow, B., J. C. Young and P. J. Kennelly (1997) "Gene Cloning and Expression and Characterization of a Toxin-Sensitive Protein Phosphatase from the Methanogenic Archaeon *Methanosarcina thermophila* TM-1," J. Bacteriol. 179, 5072-5075.

Solow, B., K. M. Bischoff, M. J. Zylka and P. J. Kennelly (1998) "Archaeal phosphoproteins. Identification of a hexosephosphate mutase and the α-subunit of succinyl-CoA synthetase in the extreme acidothermophile *Sulfolobus solfataricus*," Protein Sci. 7, 105-111.

Sprott G. D.(1992) "Structures of archaebacterial membrane lipids," J. Bioenerg. Biomembr. 24(6), 555-566.

Spudich, J. L. and W. Stoeckenius (1980) "Light-regulated retinal-dependent reversible phosphorylation of *Halobacterium* proteins," J. Biol. Chem. 255, 5501-5503.

Stedman, K. M., C. Schleper, E. Rumpf and W. Zillig (1999) "Genetic Requirements for the Function of the Archaeal Virus SSV1 in Sulfolobus sulfataricus: Construction and Testing of Viral Shuttle Vectors," Genetics 152, 1397-1405.

Stetter, K. O., R. Huber, E. Blochl, M. Kurr, R. D. Eden, M. Fielder, H. Cash and I. Vance (1993) "Hyperthermophilic archaea are thriving in deep North Sea and Alaskan oil reservoirs," Nature 365, 743-745.

Stock, J. B., A. M. Stock and J. M. Mottonen (1990) "Signal transduction is bacteria," Nature 344, 395-400.

Stock, J. B., M. G. Surette, W. R. McCleary and A. M. Stock (1992) "Signal transduction in bacterial chemotaxis," J. Biol. Chem. 267, 19753-19756.

Stravopodis, D. J. and N. C. Kyrpides (1999) "Identification of protein-tyrosine phosphatases in Archaea," J. Mol. Evol. 48, 625-627.

Stetter, K. O. and W. Zillig (1985) "*Thermoplasma* and the Thermophilic Sulfur-Dependent Archaebacteria," In: The Bacteria, Vol. VIII (C. R. Woese and R. S. Wolfe eds.), p. 85-170, Academic Press, Inc., Orlando, Florida.

C. Tang, W. Zhang, B. T. Chait, and D. Fenyö (2000) "A method to evaluate the quality of database search results," 48th ASMS annual meeting June, Long Beach, CA.

Tao, W., R. J. Deschenes and J. S. Fassler (1999) "Intracellular glycerol levels modulate the activity of Sln1p, a Saccharomyces cerevisiae two-component regulator," J. Biol. Chem. 274, 360-367.

Taylor, B. L. and I. B. Zhulin (1998) "In search of higher energy: metabolism-dependent behavior in bacteria," Mol. Microbiol. 28, 683-690.

Taylor, S. S., D. R. Knighton, J. Zheng, L. F. ten Eyck and J. M. Sowadski (1992) "Structural framework for the protein kinase family," Annu. Rev. Cell Biol. 8, 429-62.

Thomason P. and R. Kay (2000) "Eukaryotic signal transduction via histidine-aspartate phosphorelay," J. Cell Sci. 113 (Pt 18), 3141-3150.

Thomason P., D. Traynor and R. Kay (1999) "Taking the plunge. Terminal differentiation in *Dictyostelium*," Trends Genet. 15, 15-19.

Thomm, M. (1996) "Archaeal transcription factors and their role in transcription initiation," FEMS Microbiol. Rev. 18(2-3), 159-171.

Tonks, N. K., C. D. Diltz, and E. H. Fischer (1988) "Purification of the major protein-tyrosine phosphatases of human placenta," J. Biol. Chem. 263, 6722-6730.

Udo H., M. Inouye and S. Inouye (1997) "Biochemical characterization of Pkn2, a protein Ser/Thr kinase from *Myxococcus xanthus*, a Gram-negative developmental bacterium," FEBS Lett. 400, 188-192.

Uhl M.A., and J. F. Miller (1996) "Integration of multiple domains in a two-component sensor protein: the *Bordetella pertussis* BvgAS phosphorelay," EMBO J. 15(5):1028-1036.

Vener, A. V., A. Harms, M R. Sussman, and R. D. Vierstra (2001) "Mass Spectrometric Resolution of Reversible Protein Phosphorylation in Photosynthetic Membranes of *Arabidopsis thaliana*," J. Biol. Chem. 276(10), 6959-6966.

Vincent C., P. Doublet, C. Grangeasse, E. Vaganay, A. J. Cozzone and B. Duclos (1999) "Cells of *Escherichia coli* contain a protein-tyrosine kinase, Wzc, and a phosphotyrosine-protein phosphatase, Wzb," J. Bacteriol. 181, 3472-3477.

Vlamis-Gardikas A., F. Aslund, G. Spyrou, T. Bergman and A. Holmgren (1997) "Cloning, Overexpression, and Characterization of Glutaredoxin 2, An Atypical Glutaredoxin from *Escherichia coli*, J. Biol. Chem. 272, 11236-11243.

Walaas, D. A. and D. B. Glass (1991) "Protein phosphorylation and neuronal function," Pharma. Rev. 43, 299-349.

Walsh D. A., C. O. Brostrom, M. A. Brostrom, L. Chen, J. D. Corbin, E. Reimann, T. R. Soderling, E. G. Krebs (1972) "Cyclic AMP-dependent protein kinases from skeletal muscle and liver," Adv. Cyclic Nucleotide Res. 1, 33-45.

Wang, N., F. Soderbom, C. Anjard, G. Shaulsky and W. R. Loomis (1999) "SDF-2 induction of terminal differentiation in *Dictyostelium discoideum* is mediated by the membrane-spanning sensor kinase DhkA," Mol. Cell Biol. 19, 4750-4756.

Wang, Y., C. J. Fiol, A. A. DePaoli-Roach, A. W. Bell, M. A. Hermodson and P. J. Roach (1988) "Identification of Phosphorylation Sites in Peptides Using a Gas-Phase Sequencer," Anal. Biochem. 174, 537-547.

Wellner, D., C. Panneerselvam and B. L. Horecker (1990) "Sequencing of Peptides and Proteins with Blocked N-terminal Amino Acids: N-Acetylserine or N-Acetylthreonine," Proc. Natl. Acad. Sci. USA 87, 1947-1949.

Weng, L. P., J. Yuan and Q. Yu (1998) "Overexpression of the transmembrane tyrosine phosphatase LAR activates the caspase pathway and induces apoptosis," Curr. Biol. 26, 247-256.

Werner F., J. J. Eloranta, R. O. Weinzierl (2000) "Archaeal RNA polymerase subunits F and P are bona fide homologs of eukaryotic RPB4 and RPB12," Nucleic Acids Res. 28(21), 4299-305.

Westheimer, F. H. (1987) "Why nature chose phosphates," Science 235, 1173-1178.

Whitman, W. B., F. Pfeifer, P. Blum and A. Klein (1999) "What Archaea Have to Tell Biologists," Genetics 152, 1245-1248.

Woese, C. R. and G. E. Fox (1977) "Phylogenetic structure of the prokaryotic domain: The primary kingdoms," Proc. Natl. Acad. Sci. USA 74, 5088-5090.
Woese, C. R., and R. S. Wolfe (1985) Archaebacteria: The Urikingdom. The Bacteria VIII: 561-564.

Woese, C. R., O. Kandler and M. L. Wheelis (1990) "Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya," Proc. Natl. Acad. Sci. USA 87, 4576-4579.

Zhang, C. C. (1996) "Bacterial signalling involving eukaryotic-type protein kinases," Mol. Microbiol. 20, 9-15.

Zhang, C. C., A. Friry and L. Peng (1998) "Molecular and genetic analysis of two closely linked genes that encode, respectively, a protein phosphatase 1/2A/2B homolog and a

protein kinase homolog in the cyanobacterium *Anabaena* sp. Strain PCC 7120," J. Bacteriol. 180, 2616-2622.

Zhang, C. C. and L. Libs (1998) "Cloning and characterization of the pknD gene encoding an eukaryotic-type protein kinase in the cyanobacterium *Anabaena* sp. PCC 7120," Mol. Gen. Genet. 258, 26-33.

Zhang, Z., E. Harms, and R. L. Van Etten (1994) "Asp 129 of low molecular weight protein tyrosine phosphatase is involved in leaving group protonation," J. Biol. Chem. 269, 25947-25950.

Zhang W. and B. T. Chait (2000) "ProFound: an expert system for protein identification using mass spectrometric peptide mapping information," Anal. Chem. 72(11), 2482-2489.

Zheng, J., D. R. Knighton, J. Parello, S. S. Taylor and J. M. Sowadski (1991) "Crystallization of Catalytic Subunit of Adenosine Cyclic Monophosphate-Dependent Protein Kinase," Meth. Enzym. 200, 508-521.

Zheng, J., D. R. Knighton, L. F. ten Eyck, R. Karlsson, N. Xuong, S. S. Taylor and J. M. Sowadski (1993) "Crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with MgATP and peptide inhibitor," Biochemistry 32, 2154-2161.

Zillig, W., H. P. Arnold, I. Holz, D. Prangishvili, A. Schweier, et al. (1998) "Genetic elements in the extremely thermophilic archaeon *Sulfolobus*," Extremophiles 2, 131-140.

CURRICULUM VITAE

Brian Howard Lower
Department of Biochemistry
Virginia Polytechnic Institute and State University (Virginia Tech)
105 Engel Hall, Blacksburg, Virginia 24061
Tel: (540) 231-3062 Fax: (540) 231-9070
email: blower@vt.edu

Education

Ph.D. BIOCHEMISTRY (8/97-7/01) *Department of Biochemistry, Virginia Tech, Blacksburg, VA*.
Advisor: Prof. Peter J. Kennelly.
Thesis: "Protein Phosphorylation in the *Archaea*: Purification and Characterization of a Membrane-Associated Threonine Specific Protein Kinase and Identification, Cloning, and Characterization of a Serine Specific Protein Kinase from the Archaeon *Sulfolobus solfataricus*.

B.S. ZOOLOGY (9/90-12/95) *Department of Biological Sciences, Kent State University, Kent, OH.*

Research Interests

Functional Genomics: Proteomics with a focus on identifying those proteins involved in protein phosphorylation-dephosphorylation in prokaryotic organisms using a combination of 2-D electrophoresis, MALDI-MS, and peptide-mass search programs.

Biochemistry/Molecular Biology: Cloning and expression of the genes encoding these phosphoproteins and evaluation of their functional properties (substrate specificity, phosphorylation site mapping, functional consequences of phosphorylation, etc.) *in vitro*. Evaluating the physiological role of individual phosphorylation events by analyzing how environmental factors influence the expression, activity, and pattern of expression of protein kinases, protein phosphatases, and phosphoproteins in both normal and genetically altered cells.

Microbiology: Extremophiles particularly their evolution and molecular mechanisms that have allowed them to thrive in extreme environments.

Professional Experiences

GRADUATE RESEARCH ASSISTANT (5/98-present), Prof. Peter J. Kennelly, Department of Biochemistry, Virginia Tech, Blacksburg, VA.

Responsibilities include primary doctoral research of investigating protein O-phosphorylation in prokaryotic organisms, development of assay and purification protocols for protein kinases and protein phosphatases, culturing hyperthermophilic microorganisms, mentoring undergraduate and graduate student researchers.

GRADUATE RESEARCH ASSISTANT (12/97-4/98), Prof. Timothy J. Larson, Department of Biochemistry, Virginia Tech, Blacksburg, VA.

Cloned a potential phosphatase (Ygg F) from *E. Coli*, involved in the purification of a second phosphatase (Glp X) from *E. coli.*

GRADUATE TEACHING ASSISTANT (12/97-4/98), Prof. Thomas O. Sitz, Department of Biochemistry, Virginia Tech, Blacksburg, VA.

Assisted in teaching the undergraduate Organic Biochemistry course offered to science majors at Virginia Tech.

LABORATORY TECHNICIAN (8/94-5/95), Prof. Ratherida S. Bose, Department of Chemistry, Kent State University, Kent, OH.

Responsibilities included the preparation of undergraduate organic laboratory experiments and managing the chemistry stock room.

LABORATORY TECHNICIAN (8/90-5/94), Prof. Benjamin A. Foote, Department of Biological Sciences, Kent State University, Kent, OH.

Responsibilities included managing the departmental greenhouse, examining the geographic range and distribution of disease vectoring mosquitoes in Northeastern Ohio.

Professional Skills

MOLECULAR BIOLOGY: PCR, cloning with plasmid and phage vectors, library construction, sequencing, Southern blot analysis, genome analysis.

BIOCHEMISTRY: recombinant protein expression purification and characterization, liquid chromatography, HPLC, FPLC, analytical and preparative SDS-PAGE, two-dimensional electrophoresis, phosphoamino acid analysis, purification of native proteins, spectrophotometric and radioisotopic assay of enzymes, Western blot analysis, N-terminal and internal amino acid analysis, autoradiography.

INSTRUMENTATION: electronic autoradiography, atomic force microscopy, confocal laser scanning microscopy, MALDI-MS.

Curricular-Related Activities

MEMBER, SIGMA XI NATIONAL RESEARCH SOCIETY (2001-present).

MEMBER, PHI LAMBDA UPSILON CHEMISTRY HONOR SOCIETY (1997-2001).

COMMITTEE MEMBER, COMMISSION OF UNDERGRADUATE STUDIES AND POLICIES AT VIRGINIA TECH (1999-2000).

MEMBER, VIRGINIA ACADEMY OF SCIENCE (1998-2001).

Honors and Awards

JOHN JOHNSON MEMORIAL SCHOLARSHIP FOR OUTSTANDING GRADUATE STUDENT IN THE FIELD OF MOLECULAR BIOLOGY (2001), Virginia Tech.

JAMES H. EHEART SCHOLARSHIP (2000), Department of Biochemistry, Virginia Tech.

KENDALL W. KING MEMORIAL SCHOLARSHIP FOR OUTSTANDING SENIOR GRADUATE STUDENT (2000), Department of Biochemistry, Virginia Tech.

GEOGRAPHY DEPARTMENT STUDENT OF THE YEAR (1993, 1994), Department of Geography, Kent State University.

PRESIDENTIAL ACADEMIC AWARD (1993, 1994), Kent State University.

PAUL SWAGART UNDERGRADUATE ACADEMIC SCHOLARSHIP (1990), Kent State University.

Publications

Brian H. Lower, Kenneth M. Bischoff, and Peter J. Kennelly (2000). "The Archaeon *Sulfolobus solfataricus* Contains a Membrane-Associated Protein Kinase Activity That Preferentially Phosphorylates Threonine Residues in Vitro," *J. Bacteriol*. 182: 3452-3459.

Brian H. Lower and Peter J. Kennelly. "Identification, Cloning, and Characterization of a Protein Serine Kinase (SsPK2) from the Archaeon *Sulfolobus solfataricus*, Which Represents the First Bona Fide Protein O-Kinase From the *Archaea*" (manuscript in preparation).

Brian H. Lower and Peter J. Kennelly. "Identification and Characterization of a Unique Protein Serine/Threonine Kinase from the Archaeon *Sulfolobus solfataricus* Which undergoes Glycosylation *In Vivo*, as Well as Demonstrates an ability to use GTP as a Phosphoryl Donor *In Vitro*," (manuscript in preparation).

Brian H. Lower and Peter J. Kennelly. "Identification of Proteins that Undergo Phosphorylation on Serine and Threonine, From the Membrane Fraction of the Archaeon *Sulfolobus solfataricus* using Peptide Masses derived from MALDI-MS," (manuscript in preparation).

Presentations

Gordon Summer Research Conference on Second Messengers and Protein Phosphorylation, NH. June 11-16, 2000. Topic: "Protein Phosphorylation in the Archaea: The Archaeon *Sulfolobus solfataricus* Contains a Membrane-Associated Protein Kinase Activity That Preferentially Phosphorylates Threonine Residues In Vitro."

16th Annual Research Symposium, Virginia Tech, Blacksburg, VA. March 27, 2000. Topic: "Protein Phosphorylation in the *Archaea*: Protein-Serine/Threonine Kinase Activity of a Membrane Protein From The Archaeon *Sulfolobus solfataricus*."

Department of Biochemistry, Virginia Tech, Blacksburg, VA. August 16, 1999. Topic: "Characterization of a Membrane-Associated Protein Kinase from the Archaeon *Sulfolobus solfataricus*."