

BIOCHEMICAL AND STRUCTURAL STUDIES OF THE RNAPII CTD PHOSPHATASE

FCP1

by

KAREN LYNN ABBOTT

(Under the Direction of Pascale Legault)

ABSTRACT

FCP1 (TFIIF-associated C-terminal domain phosphatase) is the first identified phosphatase specific for the carboxyl-terminal domain (CTD) of RNA polymerase II (RNAPII). FCP1 is an essential component of the RNAPII holoenzyme that plays important roles in transcription elongation as well as polymerase recycling. Human FCP1 can be divided into three main domains based on sequence homology with yeast: the amino-terminal FCP1 homology domain (FCPH), the central domain containing a single BRCT (BRCA1 C-terminus) repeat, and the highly acidic carboxyl-terminal domain. Regulation of FCP1 phosphatase activity is complex involving multiple protein-protein interactions. The large subunit of TFIIF (RAP74) interacts directly with two conserved hydrophobic motifs located in the central and carboxyl-terminal domains of FCP1 to stimulate the phosphatase activity which is vital for polymerase recycling. Stimulation by RAP74 can be inhibited by the general transcription factor TFIIB. The HIV-1 transcriptional activator Tat has also been shown to interact with the central domain of FCP1, and this interaction inhibits phosphatase activity *in vitro*. FCP1 is a phosphoprotein, and phosphorylation by protein casein kinase II (CK2) stimulates phosphatase activity *in vitro*. We have used biochemical and NMR-based structural studies to analyze the interactions of the FCP1 central domain with HIV-1 Tat, RAP74, CK2, and nucleic acids. This work has identified and characterized a conserved region rich in

acidic and hydrophobic amino acids adjacent to the BRCT repeat within the central domain of FCP1 important for interactions with RAP74 and HIV-1 Tat. Tat binds FCP1 at two distinct non-overlapping sites located within the acidic/hydrophobic region and the BRCT domain. A novel threonine CK2 phosphorylation site is identified between these two Tat-binding sites, and binding of HIV-1 Tat inhibits phosphorylation by CK2. Nucleic acid-binding activity was identified for the central domain of FCP1, and this led to the discovery of a novel endonuclease activity. This work has illuminated novel activities for FCP1 and suggests that FCP1 may have additional roles in repair, termination, and/or mRNA processing events.

INDEX WORDS: BRCT, CK2, CKII, CTD, Endonuclease, FCP1, HIV-1 Tat, NMR, Nucleic Acid Binding, Phosphatase, RAP74, RNAPII, Transcription

BIOCHEMICAL AND STRUCTURAL STUDIES OF THE RNAPII CTD PHOSPHATASE

FCP1

by

KAREN LYNN ABBOTT

B.S., Georgia College, 1988

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial
Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2004

© 2004

Karen Lynn Abbott

All Rights Reserved

BIOCHEMICAL AND STRUCTURAL STUDIES OF THE RNAPII CTD PHOSPHATASE

FCP1

by

KAREN LYNN ABBOTT

Major Professor: Pascale Legault

Committee: Claiborne Glover
Mike McEachern
Michael Terns
Robert Scott

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2004

DEDICATION

I dedicate this thesis to my parents who have inspired me with a love for scientific discovery.

ACKNOWLEDGEMENTS

I would like to thank Jim Omichinski and Pascale Legault for their support and guidance over the years. Jim and Pascale allowed me to focus solely on my project, and they have been instrumental in my scientific training. They have taught me to focus on scientific quality instead of quantity. Pascale is an excellent writer, and I would like to thank her for her patient critique of my thesis. I would like to thank all past and present members of the Legault and Omichinski lab for their friendship. I would especially like to thank Lisa Miller Jenkins for always being there with sound advice. I would like to acknowledge the sacrifices made by my husband and my children. Without their team effort my completion of this degree would not have been possible. The love and support of my family has been unwavering. I would like to express my gratitude to my committee for their mentorship and guidance during my graduate training. I would like to acknowledge Dr. Claiborne Glover in particular for his abundant wisdom and leadership.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	v
CHAPTER	
1 INTRODUCTION AND LITERATURE REVIEW	1
Eukaryotic Transcription by RNA Polymerase II: Emerging Importance of CTD	
Phosphorylation.....	1
TFIIF-Associated CTD phosphatase (FCP1): Initial Discovery and Review of	
Published Literature	4
Domain Structure of FCP1	6
Factors That Regulate FCP1 Phosphatase Activity: HIV-1 Tat, RAP74, and CK2	
Phosphorylation.....	8
FCP1 Nuclease Activity	12
Goals.....	13
References	15
2 INTERACTIONS OF THE HIV-1 TAT AND RAP74 PROTEINS WITH THE	
RNAPII CTD PHOSPHATASE FCP1	32
Abstract	33
Introduction	34
Materials and Methods	36
Results	45

	Discussion	55
	Acknowledgements	60
	References	61
3	INVESTIGATIONS INTO THE ENHANCED BINDING OF THE RNAPII CTD PHOSPHATASE FCP1 TO RAP74 FOLLOWING CK2 PHOSPHORYLATION	78
	Abstract	79
	Introduction	80
	Materials and Methods	84
	Results	94
	Discussion	105
	Acknowledgements	111
	References	112
4	NOVEL ENDONUCLEASE ACTIVITY ASSOCIATED WITH THE CTD PHOSPHATASE FCP1	126
	Abstract	127
	Results	127
	Methods	131
	References	134
	Supplementary Material	141

5	^1H , ^{15}N , and ^{13}C ASSIGNMENT AND CSI ANALYSIS OF A COMPLEX FORMED BY THE CARBOXYL-TERMINAL DOMAIN OF HUMAN RAP74 AND A PHOSPHORYLATED PEPTIDE FROM THE CENTRAL DOMAIN OF FCP1	146
	Biological Significance	147
	Methods	149
	Extent of Assignments.....	150
	Acknowledgements	151
	References	152
6	CONCLUSIONS.....	161
	References	166

LIST OF ABBREVIATIONS

AI-ECD	Activated Ion Electron Capture Dissociation
ARM	Arginine-rich Motif
AT	3-Amino-Triazole
ATP	Adenosine Triphosphate
BLAST	Basic Local Alignment Search Tool
BRCA1	Breast Cancer 1
BRCT	BRCA1-C-Terminus
CD	Central Domain
CK2	Casein Kinase 2
CSI	Chemical Shift Index
CTD	C-Terminal Domain
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic Acid
DSIF	DRB Sensitivity Inducing Factor
DTT	Dithiothreitol
ECD	Electron Capture Detector
ECL	Enhanced Chemiluminescent
EDT	Ethanedithiol
EDTA	(Ethylenedinitrilo) Tetraacetic Acid Disodium Salt

ESI	Electrospray Ionization
FCP1	TFIIF-Associated CTD Phosphatase 1
FCPH	FCP1 Homology Domain
FT	Fourier Transformation
GSH	Glutathione (γ -Glutamyl-Cysteinyl-Glycine)
GST	Glutathione S-Transferase
HA	Hemagglutinin
HEK	Human Embryonic Kidney
hFCP1	Human FCP1
HPLC	High Performance Liquid Chromatography
HSQC	Heteronuclear Single Quantum Coherence
ICR	Ion-Cyclotron Resonance
IgG	Immunoglobulin G
IRMPD	InfraRed Multi-Photon Dissociation
I κ B β	Inhibitor Kappa B-Beta
KCl	Potassium Chloride
mRNA	Messenger RNA
MS	Mass Spectrometry
NaCl	Sodium Chloride
NELF	Negative Elongation Factor
NH ₄ Cl	Ammonium Chloride
NIH	National Institutes of Health

NMR	Nuclear Magnetic Resonance
NP-40	Nonidet P-40
PFAM	Protein Families Database of Alignments
P-TEFb	Positive Transcription Elongation Factor b
PTIP	Pax Transactivation Domain-Interacting Protein
RAP30	RNA Polymerase II-Associated Protein 30
RAP74	RNA Polymerase II-Associated Protein 74
RNA	Ribonucleic Acid
RNAPII	RNA Polymerase II
RNase T1	Ribonuclease T1
SCP1	Small Molecule CTD Phosphatase 1
SD	Synthetic Dropout
SDS	Sodium Dodecyl Sulfate
Ser	Serine
SWIFT	Stored Waveform Inverse Fourier Transform
TAR	Trans-Activation Responsive Region
TFA	Trifluoroacetic Acid
TOCSY	Total Correlation Spectroscopy
Tris-HCl	Tris(hydroxymethyl)Aminomethane Hydrochloride
UV	Ultraviolet
XRCC1	X-Ray Repair Cross Complementing 1
yFCP1	Yeast FCP1

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Eukaryotic transcription by RNA polymerase II: Emerging importance of CTD phosphorylation

Transcription of eukaryotic mRNA is a complex biochemical process involving RNA polymerase II (RNAP II) and an assortment of transcription factors. RNAP II is a large multi-subunit polymerase conserved from yeast to mammals. RNAP II is a remarkably processive enzyme with the capability of transcribing transcripts as long as 2 million nucleotides without dissociating from the DNA template.

Transcription by RNAP II is a cyclic process broadly divided into three main phases: initiation, elongation, and termination. Each phase of transcription is highly regulated. The initiation phase of transcription is regulated by the recruitment of a class of general transcription factors (TFIIB, TFIID, TFIIIE, TFIIF, TFIIH) required for formation of an active pre-initiation complex (Roeder 1996). Transition to the elongation phase of transcription is regulated partly by phosphorylation of the carboxyl-terminal domain (CTD) of the largest subunit of RNAP II. The largest subunit of RNAP II has a CTD that contains tandem repeats of a heptapeptide sequence (Y-S-P-T-S-P-S) (Corden et al. 1985). The presence of the heptapeptide repeats within the CTD is a conserved feature among eukaryotic RNAP II with the number of repeats varying according to the complexity of the organism (Allison et al. 1985). For example yeast RNAP II CTD has 26-

27 repeats, *C.elegans* has 34, *Drosophila* has 43, and human has 52. Different forms of RNAP II exist within cells and are identified based on the phosphorylation state of the CTD. The hypophosphorylated (RNAP IIA) form of the polymerase is found in initiating transcription complexes, whereas the hyperphosphorylated (RNAP IIO) form of the enzyme is isolated from elongating transcription complexes (Lu et al. 1991; O'Brien et al. 1994). These different isolated forms of RNAP II firmly link the transition from transcription initiation to elongation with changes in CTD phosphorylation. Indeed, enzymes capable of phosphorylating and dephosphorylating the CTD of RNAP II have been identified.

Many kinases are capable of phosphorylating the CTD of RNAP II, including TFIIF (Cdk7) (Feaver et al. 1991; Lu et al. 1992; Serizawa et al. 1992), Srb10/Cdk8 kinase (Liao et al. 1995), and P-TEFb (Cdk9) (Marshall et al. 1996). The importance of CTD phosphorylation for the transition to an elongating transcription complex is highlighted by the fact that the HIV-1 transcription activator Tat interacts with two of the kinase complexes capable of phosphorylating the CTD: TFIIF and P-TEFb (Mancebo et al. 1997; Zhu et al. 1997). In fact, Tat has been shown to alter the substrate specificity of Cdk9 to allow phosphorylation at serine 5 as well as at the preferred site serine 2 (Zhou et al. 2000).

Phosphatases capable of dephosphorylating the CTD have also been identified. The first CTD-specific phosphatase identified is FCP1 (TFIIF-associated CTD phosphatase) (Chambers and Dahmus 1994). Recently, other phosphatases capable of dephosphorylating the CTD have been identified. One is called SCP1 (small CTD phosphatase) (Yeo et al. 2003), and the other is the well known protein phosphatase-1 (PP1) (Washington et al. 2002). The HIV-1 transcriptional activator Tat makes critical interactions with PP1 (Ammosova et al. 2003) and FCP1 (Abbott et al. 2004a). HIV-1 Tat has been shown to interact with the CTD phosphatase FCP1 to

suppress phosphatase activity *in vitro* (Marshall et al. 1998). Mutations in Tat that reduce transactivation also reduce the ability of Tat to suppress phosphatase activity (Marshall et al. 1998). This further illustrates the importance of maintaining a hyperphosphorylated CTD to enhance transcription elongation.

The phosphorylation status of the CTD is also important in regulating the association of mRNA processing factors. Capping enzymes bind to the phosphorylated CTD *in vitro* (Cho et al. 1997; McCracken et al. 1997; Yue et al. 1997). Although the link between splicing and the CTD is not as definitive, phosphorylated CTD stimulates splicing indirectly, and truncation of the CTD inhibits splicing (Corden and Patturajan 1997). Transcript cleavage and polyadenylation are also linked to the phosphorylation state of the CTD. In yeast the cleavage/Poly-(A) factor pCF11 directly associates with the phosphorylated CTD (Rodriguez et al. 2000). Cleavage and polyadenylation factors are more active in the presence of the CTD (Hirose and Manley 1998).

Research over the past few years has illustrated that CTD phosphorylation is not uniform. Studies in yeast have shown that differential phosphorylation of the CTD occurs during the various phases of transcription (Kormarnitsky et al. 2000). Phosphorylation at serine 5 is found at promoter regions, whereas RNAP II CTD immunoprecipitated from coding regions has a predominance of serine 2 phosphorylation. Interestingly, yeast FCP1 has been isolated by chromatin immunoprecipitation from elongating complexes, and mutation of the FCP1 gene leads to increases in serine 2 phosphorylation (Cho et al. 2001).

Clearly complex regulation occurs to maintain a specified phosphorylated configuration of the CTD. Understanding the regulation of kinases and phosphatases coordinating CTD phosphorylation will enhance our knowledge of the transcription process and could lead to new therapeutic treatments for cancer and viral infections.

TFIIF-associated CTD phosphatase (FCP1): Initial discovery and review of published literature

FCP1 was originally partially purified from HeLa cell extracts (Chambers and Dahmus 1994). Early experiments using partially purified FCP1 identified TFIIF and TFIIB as having roles in regulating FCP1 activity (Chambers et al. 1995). TFIIF was found to stimulate FCP1 phosphatase activity, whereas TFIIB inhibits the stimulation by TFIIF. Partial human FCP1 cDNA clones were isolated from a yeast two-hybrid screen using the conserved carboxyl-terminal domain of the large subunit of TFIIF (RAP74) as bait (Archambault et al. 1998). Two fragments of FCP1 were isolated: an 842 amino acid form that lacks the first 119 amino acids, referred to as FCP1a, and a form that is missing the last 128 amino acids from the carboxyl-terminus, known as FCP1b. Subsequently, Reinberg and coworkers were able to isolate the full-length 961 amino acid clone known as FCP1 (Cho et al. 1999) (Figure 1.1). Initially, it was unclear whether FCP1 was the phosphatase or simply a regulatory subunit of the phosphatase since the amino acid sequence of FCP1 did not show homology with any known proteins of the serine/threonine phosphatase family. Studies conducted using *in vitro* translated His-tagged FCP1 immunoprecipitated from reticulocyte lysates demonstrated that FCP1 (1-961) or FCP1b (1-833) were capable of dephosphorylating isolated RNAP II O *in vitro* (Cho et al. 1999), but because FCP1 was not active on synthetic CTD substrates or isolated Rpb1, it was not definitively accepted as the CTD phosphatase. Database alignments found that the amino terminus of FCP1 contained a conserved hydrophobic motif ($\Psi\Psi\Psi\text{DXDX}(\text{T/V})\Psi\Psi$) (where Ψ denotes hydrophobic residues) (Fig. 1.1) found in a class of small molecule phosphotransferases and phosphohydrolases (Collet et al. 1998). Mutation studies in yeast have

demonstrated that the conserved aspartate residues in this motif are essential for CTD phosphatase activity (Kobor et al. 1999). FCP1 is currently accepted as a CTD phosphatase, essential for yeast cell viability and required for RNAP II transcription *in vivo* (Kobor et al. 1999). Human FCP1 has been shown to be a stoichiometric component of the RNAP II holoenzyme initiating complex (Archambault et al. 1998).

New roles for FCP1 in transcription are emerging. FCP1 has been found to function as a positive regulator of transcription elongation. This activity is independent of its CTD phosphatase activity (Cho et al. 1999; Mandal et al. 2002). Yeast studies have established a genetic link between FCP1 and various transcription elongation factors such as TFIIS and DSIF (spt4, spt5) (Costa and Arndt 2000; Lindstrom and Hartzog 2001; Mandal et al. 2002). FCP1 has also been isolated by chromatin immunoprecipitation experiments from active elongating complexes, suggesting that FCP1 remains stably associated with the polymerase during elongation (Cho et al. 2001). These experiments firmly establish that FCP1 plays an essential role in transcription elongation.

A role for FCP1 in mRNA processing is also beginning to emerge. FCP1 has been shown to co-immunoprecipitate with MEP50, a component of the methylosome complex that binds to Sm proteins (Licciardo et al. 2003). This same study was able to show an interaction between FCP1 and the U1 snRNA. Yeast capping enzymes Ceg1 and Abd1 are recruited to RNAP II via the phosphorylated CTD, and FCP1 phosphatase activity is essential for the subsequent release of these capping enzymes (Schroeder et al. 2000). Identification of FCP1 associating with components of the pre-mRNA processing machinery further highlights the global importance of FCP1 in RNAP II transcription.

Domain structure of FCP1

FCP1 is a ubiquitous serine phosphatase with orthologs present in all eukaryotic species. Sequence alignments of human and yeast FCP1 reveal three main areas of conservation, with non-conserved sequence located between these regions. This type of organization suggests that these regions may represent functional domains (Archambault et al. 1997). The amino-terminal region (hFCP1 122-330), named the FCP1 homology domain, contains the conserved phosphotransferase motif. The central region containing the BRCT domain (hFCP1 562-738) and the carboxyl-terminal region (hFCP1 812-961) of FCP1 are thought to serve as regulatory domains (Fig. 1.1).

Studies in fission and budding yeast have defined a portion of FCP1 that includes the FCP1 homology domain through the BRCT domain that is required for phosphatase activity (Kobor et al. 2000; Hausmann and Shuman 2003). FCP1 is the only protein within the general transcription machinery known to have a BRCT domain. The BRCT (BRCA1 C-terminus) domain, named for its initial isolation from the breast tumor suppressor gene BRCA1, is a small ~95 amino acid autonomously folding domain found in proteins of diverse function (Callebaut and Mornon 1997). There are over 570 proteins in the database with identified BRCT domains with functions in cell cycle regulation, DNA repair, chromatin remodeling, and transcription activation (Bork et al. 1997; Callebaut and Mornon 1997; Miyake et al. 2000; Ye et al. 2001; Taylor et al. 2002; Morales et al. 2003). BRCT domains can interact in various ways with other molecules, as homo- or hetero-BRCT dimers, in complexes with non-BRCT containing proteins, or complexed with DNA (Taylor et al. 1998; Soulier and Lowndes 1999; Yamane and Tsuruo 1999). Several structures of BRCT domains have been determined by X-ray crystallography

(Zhang et al. 1998; Williams et al. 2001; Derbyshire et al. 2002; Joo et al. 2002). As shown in Figure 1.2 these structures reveal a core motif composed of four-stranded parallel beta sheets surrounded by three alpha helices. The central regulatory domain of FCP1 contains a single BRCT repeat. Yeast two-hybrid studies reveal that this region of FCP1 is important for mediating interactions with the HIV-1 transcriptional activator Tat and the large subunit of TFIIF (RAP74) (Archambault et al. 1997; Archambault et al. 1998). Currently, there are no structural studies available for the FCP1 BRCT domain; characterizing interactions of the FCP1 BRCT domain with other molecules will provide valuable information for future structural and biochemical studies.

The carboxyl-terminal domain of FCP1 is highly acidic, reminiscent of acidic activation domains (Nguyen et al. 2003a). In fact, the yeast or human FCP1 carboxyl terminus is capable of activating transcription when artificially tethered to a DNA-binding domain (Kobor et al. 2000; Licciardo et al. 2001). The carboxyl terminus of human FCP1 also contains a nuclear localization signal. Constructs containing a deletion of the carboxyl-terminal domain localize primarily in the cytoplasm of the cell (Licciardo et al. 2001). Binding studies *in vitro* and *in vivo* have demonstrated that the carboxyl-terminal domain of FCP1 interacts with the carboxyl-terminal region of the largest subunit of TFIIF (RAP74) and the first cyclin repeat of TFIIB (Archambault et al. 1997; Archambault et al. 1998; Kobor et al. 2000). TFIIB and TFIIF share a short conserved amino acid sequence (KEFGK) responsible for binding FCP1 (Kobor et al. 2000). The regulation of phosphatase activity by RAP74 and TFIIB, which share a similar binding motif for FCP1, suggests mutually exclusive interaction with FCP1 although this has not been directly demonstrated.

Two high-resolution structures are available for the carboxyl-terminus of FCP1 complexed to RAP74 (Kamada et al. 2003; Nguyen et al. 2003a). The first identified RAP74-binding site of FCP1 is located in the carboxyl-terminal domain (881-961). This region of FCP1 contains a conserved hydrophobic motif similar to the LXXLL motif previously described for steroid hormone receptor coactivators and corepressors (Heery et al. 1997). The last 17 amino acids of FCP1 forms an amphipathic helix and binds to a hydrophobic groove formed at the surface of the winged-helix carboxyl-terminal RAP74 structure (Nguyen et al. 2003a) (Figure 1.3). This type of interaction is similar to the interaction of nuclear receptor (NR) coactivators or corepressors with the ligand binding domain (LBD) of the nuclear receptor. In the NR coactivators and corepressors as well as FCP1 the LXXLL motif forms an amphipathic helix that binds in a hydrophobic groove. Furthermore, intermolecular interactions in both cases involve the use of nonaromatic side chains predominately. Finally, a charged clamp is present at the carboxyl-terminal and amino-terminal ends of the helix orienting the helix into a hydrophobic groove (Nguyen et al 2003a).

Factors that regulate FCP1 phosphatase activity: HIV-1 Tat, RAP74, and CK2 phosphorylation.

The HIV-1 transcriptional activator Tat is essential for viral replication and is required for viral gene expression (Cullen 1991; Jones and Peterlin 1994). Despite the presence of numerous enhancer elements within the HIV-1 LTR, aborted transcripts are predominately found in the absence of Tat. Tat primarily stimulates the efficiency of transcription elongation (Marciniak and Sharp 1991). Tat transactivation requires the CTD of RNAP II (Chun and Jeang 1996). The ability of Tat to activate transcription correlates with hyperphosphorylation of the

CTD (Parada and Roeder 1996). Over the years, Tat has been shown to interact directly with several cellular cofactors that are essential for Tat transactivation (Wu-Baer et al. 1996; Zhou and Sharp 1996; Mancebo et al. 1997; Zhu et al. 1997; Emerman and Malim 1998; Ammosova et al. 2003; Abbott et al. 2004a). The unifying theme of these cofactors is that they all influence phosphorylation of the RNAP II CTD either directly or indirectly.

Tat is a unique transcription factor because it acts via an RNA element rather than a DNA element. Tat and other cellular cofactors are recruited to the 5' end of the viral mRNA via interaction with the transactivation responsive region (TAR) RNA element (Jones and Peterlin 1994; Garber et al. 1998; Bieniasz et al. 1999; Richter et al. 2002). TAR RNA adopts a stable stem-loop structure that is required for efficient transactivation by Tat. Tat proteins from all lentiviruses are highly conserved, providing an ideal target for drug development. Structural studies of HIV-1 Tat have been hampered by the dynamic nature of the protein. Tat does not adopt stable elements of regular secondary structure (Bayer et al. 1995). However, three flexible sequence regions can be defined: the cysteine-rich amino-terminal activation domain, the centrally located highly basic arginine-rich RNA binding motif (ARM), and the carboxyl-terminal glutamine-rich region (Figure 1.4). The ARM of Tat is responsible for binding to a U-rich bulge in the stem of the TAR element. The ARM cannot function independently and requires the activation domain for high-affinity binding to the TAR element *in vivo* (Jones and Peterlin 1994). The recruitment of Tat to TAR is significantly enhanced by the CTD kinase complex P-TEFb (cyclin T1) (Zhang et al. 2000).

Yeast-two hybrid studies conducted by our collaborator Dr. Jack Greenblatt have determined that Tat interacts with the central domain of FCP1. Using purified proteins *in vitro*, Tat was shown to inhibit phosphatase activity (Marshall et al. 1998). Mutations in Tat that impair

transactivation activity also impair the ability of Tat to inhibit CTD phosphatase activity (Marshall et al. 1998). These early studies indicated that the interaction of Tat with FCP1 was an important component of HIV-1 Tat transactivation. Detailed analysis of the interaction between FCP1 and Tat will help us better understand how HIV-1 commands cellular transcription controls and may lead to novel anti-viral therapies.

The general transcription factor TFIIF influences multiple facets of RNAP II transcription. TFIIF is composed of a large subunit (RAP74) and a small subunit (RAP30) and is believed to exist as a heterodimer *in vivo* (Conaway and Conaway 1993). During the initiation phase of transcription, TFIIF is required for recruiting RNAP II to the promoter through direct interaction of RNAP II with RAP30 (McCracken and Greenblatt 1991). TFIIF has some interesting characteristics similar to bacterial sigma factors. These include the ability to bind tightly to the polymerase, suppression of nonspecific binding by the polymerase, and stabilization of a pre-initiation complex (Greenblatt 1991; Conaway and Conaway 1993). Photo-cross-linking studies have revealed that both RAP74 and RAP30 bind promoter DNA between the TATA box and the transcription start site (Robert et al. 1996). The amino-terminal portion of RAP74 through its contacts with the promoter induces conformational changes in the polymerase complex that promotes wrapping of the DNA around the polymerase (Robert et al. 1998). TFIIF also stimulates transcription elongation by suppressing pausing of the polymerase (Tan et al. 1995). On non-chromatin templates TFIIF can stimulate RNAP II transcription elongation to near physiological rates (Izban and Luse 1992). The RAP74 subunit of TFIIF makes direct contact with DNA, RNAP II, TFIIB, and FCP1 (Tan et al. 1995; Wang and Burton 1995; Fang and Burton 1996; Archambault et al. 1998; Kobor et al. 2000). Through these interactions RAP74 is capable of influencing transcription initiation, elongation, and polymerase recycling.

Human RAP74 is a 517 amino acid protein that can be divided into three main domains based on sequence: the amino-terminal (1-217) highly basic globular domain that binds to RAP30 as well as DNA, the highly acidic central domain (218-398), and the globular basic carboxyl-terminal domain (399-517) important for interactions with FCP1 (Archambault et al. 1998; Lei et al. 1998; Nguyen et al. 2003a; Nguyen et al. 2003b).

FCP1 was originally identified based on its ability to bind the evolutionarily conserved carboxyl-terminus of RAP74 (Archambault et al. 1998). Yeast-two hybrid studies have revealed that the carboxyl-terminus of RAP74 interacts with the central and carboxyl-terminal regions of FCP1 (Archambault et al. 1998). The carboxyl-terminal domain of RAP74 stimulates the CTD phosphatase activity of FCP1 and is important for stimulating multiple rounds of transcription *in vitro* (Chambers et al. 1995; Lei et al. 1998). The structure of the carboxyl-terminal domain of RAP74 (cterRAP74) has been solved by x-ray crystallography and NMR spectroscopy (Kamada et al. 2001; Nguyen et al. 2003b). Both structures reveal that the carboxyl terminus of RAP74 has an overall fold similar to the winged-helix domain. Initial work from our laboratory determining the structure of the complex of cterRAP74 with the carboxyl domain of FCP1 (cterFCP1) revealed a small conserved hydrophobic motif within FCP1 required for binding of RAP74 (Nguyen et al. 2003a). This led to the identification of a similar small hydrophobic RAP74 binding site located in the central domain of FCP1 (Abbott et al. 2004a). Biochemical and structural analysis of the cterRAP74/FCP1 interactions described previously and within this dissertation highlight a conserved mechanism of interaction with similarities to the steroid hormone receptor superfamily (Nguyen et al. 2003a).

Protein kinase CK2, formerly known as casein kinase II, is an abundant, constitutively active, pleiotropic kinase. CK2 is essential for cell survival and is influential in class I, II, and III

gene transcription (Dahmus et al. 1984; Hannan et al. 1998; Egyhazi et al. 1999; Ghavidel and Schultz 2001). Ironically, twenty years ago the challenge was to find physiological substrates for CK2. Today there are over 300 cellular proteins identified to be substrates of CK2 (Meggio and Pinna 2003). FCP1 contains numerous potential CK2 phosphorylation sites, and CK2 has been identified as a kinase capable of phosphorylating FCP1 *in vitro* (Palancade et al. 2002; Friedl et al. 2003) (Abbott et al., 2003a submitted) (Abbott et al., 2003b submitted). Human FCP1 isolated from baculovirus exists in a phosphorylated state and is more active in stimulating transcription elongation than dephosphorylated FCP1 (Friedl et al. 2003). Interestingly, elongation reactions conducted in the presence of CK2 are inhibited (Friedl et al. 2003). This effect appears to be FCP1-specific, as CK2 added to elongation reactions in the absence of FCP1 has no effect (Friedl et al. 2003). *Xenopus* CK2 co-purifies with FCP1; therefore CK2 appears to be a principal *in vivo* kinase for FCP1 (Palancade et al. 2002). Phosphorylation of FCP1 by CK2 enhances RAP74 binding and stimulates CTD phosphatase activity (Friedl et al. 2003) (Palancade et al. 2002). The minimal RAP74 binding sites located within the central and carboxyl-terminal regulatory domains of FCP1 contain consensus CK2 phosphorylation sites (Abbott et al. 2004b). Analysis of these CK2 phosphorylation sites *in vitro* and *in vivo* will shed light on the possible biological consequences of FCP1 phosphorylation.

FCP1 Nuclease Activity

Nucleases play vital roles in several cellular processes such as replication, recombination, repair, RNA maturation, telomere maintenance, and transcription. The enzymatic hydrolysis of nucleic acid was first discovered over 100 years ago (Araki 1903). The phosphodiester bond is remarkably resistant to cleavage in part due to electrostatic repulsion between the phosphodiester

anion and the approaching potential nucleophile required for hydrolysis. Therefore, diverse mechanisms have evolved to cleave the phosphodiester bond. Some enzymes utilize water or a hydroxide anion to catalyze hydrolysis of the (P-O) bond while others utilize phosphate and cleave the (P-O) bond using a phosphorolytic mechanism. Enzymes involved in repair of damaged DNA are known to cleave the (C-O) bond at abasic sites by β elimination (Gerlt 1993). Nucleases are classified based on several considerations: nature of the nucleic acid hydrolyzed (DNA vs RNA and single-strand vs double-strand), type of attack (endonucleolytic or exonucleolytic), and the nature of the termini created. Single strand-specific nucleases, first discovered in 1960 (Lehman 1960), recognize single-stranded nucleic acids or single-stranded regions of double stranded nucleic acid structures. Many single strand-specific nucleases have diverse substrate specificities. For example, the nuclease Rsn from *Rhizopus stolonifer* hydrolyzes ssDNA>dsDNA>RNA making it a single stranded DNA-preferring nuclease (Rangarajan and Shankar 1999). Several single strand-specific nucleases have multiple activities such as the endo-exonuclease Mre11 important for repair and recombination (Connelly and Leach 1996; Connelly et al. 1997). Single strand-specific nucleases can also create different termini on different substrates. For example, the wheat chloroplast nuclease hydrolyzes ssDNA endonucleolytically to produce oligonucleotides that have 3'-OH and 5'-PO₄ termini, whereas endonucleolytic cleavage of ssRNA by the same enzyme yields products having 3'-PO₄ and 5'-OH (Kuligowska et al. 1988). During the course of this dissertation project, a nuclease activity was discovered associated with the central domain of FCP1. The data presented herein are the first attempt to classify the possible activities of this enzyme.

Goals

FCP1 is a ubiquitous nuclear phosphatase that has not been extensively studied. The initial goals of my thesis work were to characterize, using purified proteins, the molecular interactions of FCP1 with the proteins known to regulate its phosphatase activity, HIV-1 Tat and RAP74. We also wanted to determine the involvement of the HIV-1 TAR element in the formation of the FCP1/Tat complex. These biochemical experiments led into two new areas of study: (i) NMR-based structural and biochemical analysis of how CK2 directed FCP1 phosphorylation results in enhanced binding of RAP74 (ii) the discovery and characterization of an endonuclease activity associated with the central domain of FCP1.

REFERENCES

- Abbott, K.L., J. Archambault, H. Xiao, B.D. Nguyen, R.G. Roeder, J. Greenblatt, J.G. Omichinski, and P.L. Legault. 2004a. Interactions of the HIV-1 Tat and RAP74 Proteins with the RNAPII CTD Phosphatase FCP1. *(To be submitted to J. Biol. Chem.)*
- Abbott, K.L., M.B. Renfrow, M.J. Chalmers, B.D. Nguyen, J. Greenblatt, A.G. Marshall, P. Legault, and J.G. Omichinski. 2004b. Investigations into the enhanced binding of the RNAPII CTD phosphatase FCP1 to RAP74 following CK2 phosphorylation. *(To be submitted to J. Biol. Chem.)*
- Allison, L.A., M. Moyle, M. Shales, and C.J. Ingles. 1985. Extensive homology among the large subunits of eukaryotic and prokaryotic RNA polymerases. *Cell* **42**: 599-610.
- Ammosova, T., M. Jerebtsova, M. Beullens, Y. Voloshin, P.E. Ray, A. Kumar, M. Bollen, and S. Nekhai. 2003. Nuclear protein phosphatase-1 regulates HIV-1 transcription. *J Biol Chem* **278**: 32189-94.
- Araki, T. 1903. Enzymatic decomposition of nucleic acids. *Chem.* **38**: 84-92.
- Archambault, J., R.S. Chambers, M.S. Kobor, Y. Ho, M. Cartier, D. Bolotin, B. Andrews, C.M. Kane, and J. Greenblatt. 1997. An essential component of a C-terminal domain phosphatase that interacts with transcription factor IIF in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **94**: 14300-14305.
- Archambault, J., G. Pan, G.K. Dahmus, M. Cartier, N. Marshall, S. Zhang, M.E. Dahmus, and J. Greenblatt. 1998. FCP1, the RAP74-interacting subunit of a human protein phosphatase that dephosphorylates the carboxy-terminal domain of RNA polymerase II. *J. Biol. Chem.* **273**: 27593-27601.

- Bayer, P., M. Kraft, A. Ejchart, M. Westendorp, R. Frank, and P. Rosch. 1995. Structural studies of HIV-1 Tat protein. *J. Mol. Biol.* **247**: 529-535.
- Bieniasz, P.D., T.A. Gridina, H.P. Bogerd, and B.R. Cullen. 1999. Analysis of the effect of natural sequence variation in Tat and in cyclin T on the formation and RNA binding properties of Tat/cyclin T complexes. *J. Virol.* **73**: 5777-5786.
- Bork, P., K. Hofmann, P. Bucher, A.F. Neuwald, S. Altschul, and E.V. Koonin. 1997. A superfamily of conserved domains in DNA damage-responsive cell cycle checkpoint proteins. *FASEB J* **11**: 68-76.
- Callebaut, I. and J.P. Mornon. 1997. From BRCA1 to RAP1: a widespread BRCT module closely associated with DNA repair. *FEBS Lett.* **400**: 25-30.
- Chambers, R.S. and M.E. Dahmus. 1994. Purification and characterization of a phosphatase from HeLa cells which dephosphorylates the C-terminal domain of RNA polymerase II. *J. Biol. Chem.* **269**: 26243-26248.
- Chambers, R.S., B.Q. Wang, Z.F. Burton, and M.E. Dahmus. 1995. The activity of COOH-terminal domain phosphatase is regulated by a docking site on RNA polymerase II and by the general transcription factors IIF and IIB. *J. Biol. Chem.* **270**: 14962-14969.
- Cho, E.J., M.S. Kobor, M. Kim, J. Greenblatt, and S. Buratowski. 2001. Opposing effects of Ctk1 kinase and FCP1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. *Genes & Dev.* **15**: 3319-3329.
- Cho, E.J., T. Takagi, C.R. Moore, and S. Buratowski. 1997. mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxyl-terminal domain. *Genes & Dev.* **11**: 3319-3326.

- Cho, H., T.K. Kim, H. Mancebo, W.S. Lane, O. Flores, and D. Reinberg. 1999. A protein phosphatase functions to recycle RNA polymerase II. *Genes & Dev.* **13**: 1540-1552.
- Chun, R. and K. Jeang. 1996. Requirements for RNA polymerase II carboxyl-terminal domain for activated transcription of human retroviruses human T-cell lymphotropic virus I and HIV-1. *J. Biol. Chem.* **271**: 27888-27894.
- Collet, J.F., V. Stroobant, M. Pirard, G. Delpierre, and E.V. Schaftingen. 1998. A new class of phosphotransferases phosphorylated on an aspartate residue in an Amino-terminal DXDX(T/V) motif. *J. Biol. Chem.* **273**: 14107-14112.
- Conaway, R.C. and J.W. Conaway. 1993. General initiation factors for RNA polymerase II. *Annu Rev. Biochem.* **62**: 161-190.
- Connelly, J.C., E.S. de Leau, E.A. Okely, and D.R. Leach. 1997. Overexpression, purification, and characterization of the SbcCD protein from Escherichia coli. *J Biol Chem* **272**: 19819-26.
- Connelly, J.C. and D.R. Leach. 1996. The sbcC and sbcD genes of Escherichia coli encode a nuclease involved in palindrome inviability and genetic recombination. *Genes Cells* **1**: 285-91.
- Corden, J.L., D.L. Cadena, J.M. Ahearn, and M.E. Dahmus. 1985. A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II. *Proc. Natl. Acad. Sci. USA* **82**: 7934-7938.
- Corden, J.L. and M. Patturajan. 1997. A CTD function linking transcription to splicing. *Trends Biochem. Sci.* **22**: 413-416.

- Costa, P.J. and K.M. Arndt. 2000. Synthetic lethal interactions suggest a role for the *Saccharomyces cerevisiae* Rtf1 protein in transcription elongation. *Genetics* **156**: 535-547.
- Cullen, B.R. 1991. Regulation of HIV-1 gene regulation. *FASEB J.* **5**: 2361-2368.
- Dahmus, G.K., C.V. Glover, D.L. Brutlag, and M.E. Dahmus. 1984. Similarities in structure and function of calf thymus and Drosophila casein kinase II. *J. Biol. Chem.* **259**: 9001-6.
- Derbyshire, D.J., B.P. Basu, L.C. Serpell, W.S. Joo, T. Date, K. Iwabuchi, and A.J. Doherty. 2002. Crystal structure of human 53BP1 BRCT domains bound to p53 tumour suppressor. *EMBO J.* **21**: 3863-3872.
- Egyhazi, E., A. Ossoinak, O. Filhol-Cochet, C. Cochet, and A. Pigon. 1999. The binding of the a subunit of protein kinase CK2 and RAP74 subunit of TFIIF to protein-coding genes in living cells is DRB sensitive. *Mol. and Cell. Biochem.* **191**: 149-159.
- Emerman, M. and M.H. Malim. 1998. HIV-1 Regulatory/Accessory genes: keys to unraveling viral and host cell biology. *Science* **280**: 1880-1884.
- Fang, S.M. and Z.F. Burton. 1996. RNA polymease II-associated protein (RAP) 74 binds transcription factor (TF) IIB and blocks TFIIB-RAP30 binding. *J. Biol. Chem.* **271**: 11700-11709.
- Feaver, W.J., O. Gileadi, Y. Li, and R.D. Kornberg. 1991. CTD Kinase associated with yeast RNA polymerase II initiation factor b. *Cell* **67**: 1223-1230.
- Friedl, E.M., W.S. Lane, H. Erdjument-Bromage, P. Tempst, and D. Reinberg. 2003. The C-terminal domain phosphatase and transcription activities of FCP1 are regulated by phosphorylation. *Proc. Natl. Acad. Sci.* **100**: 2328-2333.

- Garber, M.E., P. Wei, V.N. KewalRamani, T.P. Mayall, C.H. Herrmann, A.P. Rice, D.R. Littman, and K.A. Jones. 1998. The interaction between the HIV-1 Tat and human cyclin T1 requires zinc and a critical cysteine residue that is not conserved in the murine CycT1 protein. *Genes & Dev.* **12**: 3512-3527.
- Gerlt, J. 1993. Mechanistic Principles of Enzyme-catalyzed Cleavage of Phosphodiester Bonds. In *Nucleases* (ed. S.M. Linn). Cold Spring Harbor Laboratory Press, New York.
- Ghavidel, A. and M.C. Schultz. 2001. TATA binding protein-associated CK2 transduces DNA damage signals to the RNA polymerase III transcriptional machinery. *Cell* **106**: 575-84.
- Greenblatt, J. 1991. RNA polymerase associated transcription factors. *Trends Biochem. Sci.* **16**: 408-411.
- Hannan, R.D., W.M. Hempel, A. Cavanaugh, T. Arino, S.I. Dimitrov, T. Moss, and L. Rothblum. 1998. Affinity purification of mammalian RNA polymerase I. Identification of an associated kinase. *J. Biol. Chem.* **273**: 1257-67.
- Hausmann, S. and S. Shuman. 2003. Defining the active site of *Schizosaccharomyces pombe* C-terminal domain phosphatase Fcp1. *J. Biol. Chem.* **278**: 13627-13632.
- Heery, D.M., E. Kalkhoven, S. Hoare, and M.G. Parker. 1997. A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature* **387**: 733-736.
- Hirose, Y. and J.L. Manley. 1998. RNA polymerase II is an essential mRNA polyadenylation factor. *Nature* **395**: 93-96.
- Izban, M.G. and D.S. Luse. 1992. Factor-stimulated RNA polymerase II transcribes at physiological rates on naked DNA but very poorly on chromatin templates. *J. Biol. Chem.* **267**: 13647-13655.

- Jones, K.A. and B.M. Peterlin. 1994. Control of RNA initiation and elongation at the HIV-1 promoter. *Annu. Rev. Biochem.* **63**: 717-743.
- Joo, W.S., P.D. Jeffrey, S.B. Cantor, M.S. Finnin, D.M. Livingston, and N.P. Pavletich. 2002. Structure of the 53BP1 BRCT region bound to p53 and its comparison to the Brca1 BRCT structure. *Genes & Dev.* **16**: 583-93.
- Kamada, K., J.D. Angelis, R. Roeder, and S. Burley. 2001. Crystal structure of the C-terminal domain of the RAP74 subunit of human transcription factor IIF. *Proc. Natl. Acad. Sci.* **98**: 3115-20.
- Kamada, K., R.G. Roeder, and S.K. Burley. 2003. Molecular mechanism of recruitment of TFIIF-associating RNA polymerase C-terminal domain phosphatase (FCP1) by transcription factor IIF. *Proc. Natl. Acad. Sci.* **100**: 2296-2299.
- Kobor, M.S., J. Archambault, W. Lester, F.C.P. Holstege, O. Gileadi, and J. Greenblatt. 1999. An Unusual Eukaryotic Protein Phosphatase Required for Transcription by RNA Polymaerase II and CTD Dephosphorylation in *S. cerevisiae*. *Molecular Cell* **4**: 55-62.
- Kobor, M.S., L.D. Simon, J. Omichinski, G. Zhong, J. Archambault, and J. Greenblatt. 2000. A motif shared by TFIIF and TFIIB mediates their interaction with the RNA polymerase II carboxyl-terminal domain phosphatase Fcp1p in *Saccharomyces cerevisiae*. *Mol. Cell Biol.* **20**: 7438-7449.
- Kormarnitsky, P., E.J. Cho, and S. Buratowski. 2000. Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes & Dev.* **14**: 2452-2460.

- Kuligowska, E., D. Klarkowska, and J.W. Szarkowski. 1988. Purification and properties of endonuclease from wheat chloroplasts, specific for single-strand DNA. *Phytochemistry* **27**: 1275-1279.
- Lehman, I.R. 1960. The deoxyribonucleases of Escherichia Coli. I Purification and properties of a phosphodiesterase. *J. Biol. Chem.* **235**: 1479-1487.
- Lei, L., D. Ren, A. Finkelstein, and Z.F. Burton. 1998. Functions of the N- and C-terminal domains of human RAP74 in transcriptional initiation, elongation and recycling of RNA polymerase II. *Mol. Cell. Biol.* **18**: 2130-2142.
- Liao, S.M., J. Zhang, D.A. Jeffery, A.J. Koleske, C.M. Thompson, D.M. Chao, M. Viljoen, H.J.V. Vuuren, and R.A. Young. 1995. A kinase-cyclin pair in the RNA polymerase II holoenzyme. *Nature* **374**: 193-6.
- Licciardo, P., S. Amente, L. Ruggiero, M. Monti, P. Pucci, L. Lania, and B. Majello. 2003. The FCP1 phosphatase interacts with RNA polymerase II and with MEP50 a component of the methylosome complex involved in the assembly of snRNP. *Nucleic Acids Res.* **31**: 999-1005.
- Licciardo, P., L. Ruggiero, L. Lania, and B. Majello. 2001. Transcription activation by targeted recruitment of RNA polymerase II CTD phosphatase FCP1. *Nucleic Acids Research* **29**: 3539-3545.
- Lindstrom, D.L. and G.A. Hartzog. 2001. Genetic interactions of Spt4-Spt5 and TFIIIS with the RNA polymerase II CTD and CTD modifying enzymes in *Saccharomyces cerevisiae*. *Genetics* **159**: 487-97.

- Lu, H., O. Flores, R. Weinmann, and D. Reinberg. 1991. The nonphosphorylated form of the RNA polymerase II preferentially associates with the preinitiation complex. *Proc. Natl. Acad. Sci. USA* **88**: 10004-10008.
- Lu, H., L. Zawel, L. Fisher, J.M. Egly, and D. Reinberg. 1992. Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. *Nature* **358**: 641-645.
- Mancebo, H.S.Y., G. Lee, J. Flygare, J. Tomassini, P. Luu, Y. Zhu, J. Peng, C. Blau, D. Hazuda, D. Price, and O. Flores. 1997. P-TEFb kinase is required for HIV Tat transcriptional activation *in vivo* and *in vitro*. *Genes & Dev.* **11**: 2633-2644.
- Mandal, S.S., H. Cho, S. Kim, K. Cabane, and D. Reinberg. 2002. FCP1, a phosphatase specific for the heptapeptide repeat of the largest subunit of RNA polymerase II, stimulates transcription elongation. *Mol. Cell Biol.* **22**: 7543-7552.
- Marciniak, R. and P. Sharp. 1991. HIV-1 Tat promotes formation of more processive elongation complexes. *EMBO J.* **10**: 4189-4196.
- Marshall, N.F., G.K. Dahmus, and M.E. Dahmus. 1998. Regulation of carboxyl-terminal domain phosphatase by HIV-1 Tat protein. *J. Biol. Chem.* **273**: 31726-31730.
- Marshall, N.F., J. Peng, Z. Xie, and D.H. Price. 1996. Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. *J. Biol. Chem.* **271**: 27176-27183.
- McCracken, S., N. Fong, K. Yankulov, G. Bothers, D. Siderovski, A. Hessel, S. Forster, A.E. Program, S. Shuman, and D.L. Bentley. 1997. 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated caroxy-terminal domain of RNA polymerase II. *Genes & Dev.* **11**: 3306-3318.

- McCracken, S. and J. Greenblatt. 1991. Related RNA polymerase-binding regions in human RAP30/74 and *Eschericia coli* σ 70. *Science* **253**: 900-902.
- Meggio, F. and L.A. Pinna. 2003. One-thousand-and-one substrates of protein kinase CK2. *FASEB J.* **17**: 349-368.
- Miyake, T., Y.F. Hu, D.S. Yu, and R. Li. 2000. A functional comparison of BRCA1 C-terminal domains in transcription activation and chromatin remodeling. *J. Biol. Chem.* **275**: 40169-73.
- Morales, J.C., Z. Xia, T. Lu, M.B. Aldrich, B. Wang, C. Rosales, R.E. Kellems, W.N. Hittelman, S.J. Elledge, and P.B. Carpenter. 2003. Role for the BRCT protein 53BP1 in maintaining genomic stability. *J. Biol. Chem.* **278**: 14971-77.
- Nguyen, B.D., K.L. Abbott, K. Potempa, M.S. Kobor, J. Archambault, J. Greenblatt, P. Legault, and J.G. Omichinski. 2003a. NMR structure of a complex containing the TFIIF subunit RAP74 and the RNA polymerase II carboxyl-terminal domain phosphatase FCP1. *Proc. Natl. Acad. Sci. USA* **100**: 5688-5693.
- Nguyen, B.D., H. Chen, A.S. Kobor, J. Greenblatt, P. Legault, and J.G. Omichinski. 2003b. Solution structure of the carboxyl-terminal domain of RAP74 and NMR characterization of the FCP1-binding sites of RAP74 and Human TFIIB. *Biochemistry*.
- O'Brien, T.S., S. Hardin, A. Greenleaf, and J.T. Lis. 1994. Phosphorylation of RNA polymerase II C-terminal domain and transcriptional elongation. *Nature* **370**: 75-77.
- Palancade, B., M.F. Dubois, and O. Bensaudé. 2002. FCP1 phosphorylation by casein kinase 2 enhances binding to TFIIF and RNA polymerase II carboxyl-terminal domain phosphatase activity. *J. Biol. Chem.* **277**: 36061-36067.

- Parada, C. and R. Roeder. 1996. Enhanced processivity of RNA polymerase II triggered by Tat-induced phosphorylation of its carboxyl-terminal domain. *Nature* **384**: 375-378.
- Rangarajan, S. and V. Shankar. 1999. Extracellular nuclease from *Rhizopus stolonifer*: purification and characteristics as a single-strand preferential-deoxyribonuclease activity. *Biochim. Biophys. Acta*. **1473**: 293-304.
- Richter, S., Y. Ping, and T.M. Rana. 2002. TAR RNA loop: A scaffold for the assembly of a regulatory switch in HIV replication. *Proc. Natl. Acad. Sci.* **99**: 7928-7933.
- Robert, F., M. Douziech, D. Forget, J.M. Egly, J. Greenblatt, Z.F. Burton, and B. Coulombe. 1998. Wrapping of promoter DNA around the RNA polymerase II initiation complex induced by TFIIF. *Mol. Cell*. **2**: 341-51.
- Robert, F., D. Forget, J. Li, J. Greenblatt, and B. Coulombe. 1996. Localization of subunits of transcription factors IIE and IIF immediately upstream of the transcriptional initiation site of the adenovirus major late promoter. *J. Biol. Chem.* **271**: 8517-8520.
- Rodriguez, C.R., E.J. Cho, M.C. Keogh, C.L. Moore, A.L. Greenleaf, and S. Buratowski. 2000. Kin 28, the TFIIH-associated carboxyl-terminal domain kinase, facilitates the recruitment of mRNA processing machinery to RNA polymerase II. *Mol. Cell. Biol.* **20**: 104-112.
- Roeder, R.G. 1996. The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem. Sci.* **21**: 327-335.
- Schroeder, S.C., B. Schwer, S. Shuman, and D. Bentley. 2000. Dynamic association of capping enzymes with transcribing RNA polymerase II. *Genes Dev.* **14**: 2435-40.
- Serizawa, H.R., R.C. Conaway, and J.W. Conaway. 1992. A carboxyl-terminal-domain kinase associated with RNA polymerase II transcription factor delta from rat. *Proc. Natl. Acad. Sci. USA* **89**: 7476-7480.

- Soulier, J. and N.F. Lowndes. 1999. The BRCT domain of the *S. cerevisiae* checkpoint protein Rad9 mediates a Rad9-Rad9 interaction after DNA damage. *Curr. Biol.* **9**: 551-554.
- Tan, S., R.C. Conaway, and J.W. Conaway. 1995. Dissection of transcription factor TFIIF functional domains required for initiation and elongation. *Proc. Natl. Acad. Sci. USA* **92**: 6042-6046.
- Taylor, R.M., A. Thistlewhite, and K.W. Caldecott. 2002. Central role for the XRCC1 BRCT 1 domain in mammalian DNA single-strand break repair. *Mol. Cell Biol.* **22**: 2556-63.
- Taylor, R.M., B. Wickstead, S. Cronin, and K.W. Caldecott. 1998. Role of a BRCT domain in the interaction of DNA ligase III- α with the DNA repair protein XRCC1. *Curr. Biol.* **8**: 877-880.
- Wang, B.Q. and Z.F. Burton. 1995. Functional domains of human RAP74 including a masked polymerase binding domain. *J. Biol. Chem.* **270**: 27035-27044.
- Washington, K., T. Ammosova, M. Beullens, M. Jerebtsova, A. Kumar, M. Bollen, and S. Nekhai. 2002. Protein phosphatase-1 dephosphorylates the C-terminal domain of RNA polymerase-II. *J Biol Chem* **277**: 40442-8.
- Williams, R.S., R. Green, and J.N. Glover. 2001. Crystal structure of the BRCT repeat region from the breast cancer-associated protein BRCA1. *Nat. Struct. Biol.* **8**: 838-842.
- Wu-Baer, F., W.S. Lane, and R.B. Gaynor. 1996. Identification of a group of cellular cofactors that stimulate binding of RNA polymerase II and TRP-185 to Human Immunodeficiency Virus I TAR RNA. *J. Biol. Chem.* **271**: 4201-4208.
- Yamane, K. and T. Tsuruo. 1999. Conserved BRCT regions of TopBP1 and of the tumor suppressor BRCA1 bind strand breaks and termini of DNA. *Oncogene* **18**: 5194-5203.

- Ye, Q., Y.F. Hu, H. Zhong, A.C. Nye, A.S. Belmont, and R. Li. 2001. BRCA1-induced large-scale chromatin unfolding and allele-specific effects of cancer-predisposing mutations. *J. Cell Biol.* **155**: 911-21.
- Yeo, M., P.S. Lin, M.E. Dahmus, and G.N. Gill. 2003. A novel RNA polymerase II C-terminal domain phosphatase that preferentially dephosphorylates serine 5. *J. Biol. Chem.* **278**: 26078-85.
- Yue, Z., E. Maldonado, R. Pilluta, H. Cho, D. Reinberg, and A. Shatkin. 1997. Mammalian capping enzyme complements mutant *Saccharomyces cerevisiae* lacking mRNA guanytransferase and selectively binds the elongating form of RNA polymerase II. *Proc. Natl. Acad. Sci.* **94**: 12898-12903.
- Zhang, J., N. Tamilarasu, S. Hwang, M.E. Garber, and I. Huq. 2000. HIV-1 TAR RNA Enhances the interaction between Tat and cyclin T1. *JBC* **275**: 34314-34319.
- Zhang, Z., S. Morera, P.A. Bates, P.C. Whitehead, A.I. Coffey, K. Hainbucher, R.A. Nash, M.J.E. Sternberg, T. Lindahl, and P.S. Freemont. 1998. Structure of an XRCC1 BRCT domain: a new protein:protein interaction module. *EMBO J.* **17**: 6404-6411.
- Zhou, M., M.A. Halanski, M.F. Radonovich, F. Kashanchi, J. Peng, D.H. Price, and J.N. Brady. 2000. Tat modifies the activity of CDK9 to phosphorylate serine 5 of the RNA polymerase II carboxyl-terminal domain during human immunodeficiency virus type I transcription. *Mol. Cell Biol.* **20**: 5077-86.
- Zhou, Q. and P.A. Sharp. 1996. Tat-SF1: Cofactor for stimulation of transcriptional elongation by HIV-1 Tat. *Science* **274**: 605-610.

Zhu, Y., T. Pe'ery, J. Peng, Y. Ramanathan, N. Marshall, T. Marshall, B. Amendt, M.B.

Mathews, and D.H. Price. 1997. Transcriptional elongation factor P-TEFb is required for HIV-1 Tat transactivation in vitro. *Genes & Dev.* **11**: 2622-2632.

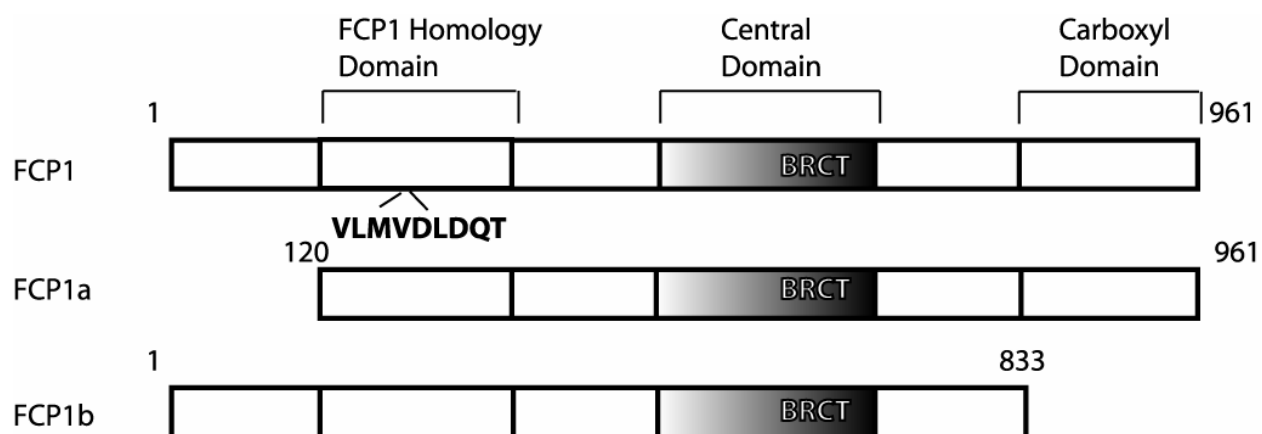


Figure 1.1. Diagram of isolated FCP1 fragments. Domains defined by sequence homology are labeled.

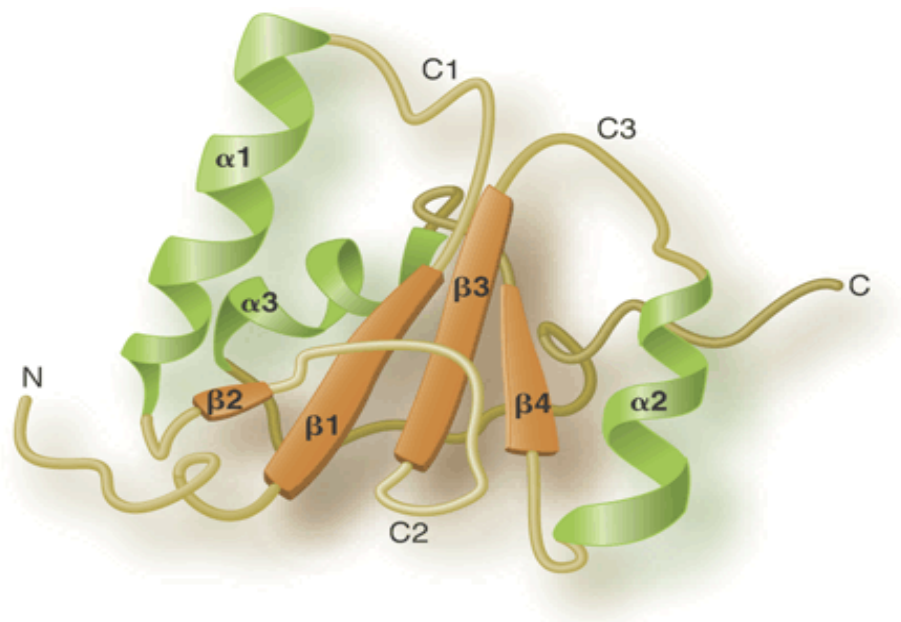


Figure 1.2. Conserved BRCT structure. The BRCT domain is composed of three alpha helices surrounding a central β sheet (Caldecott 2003).

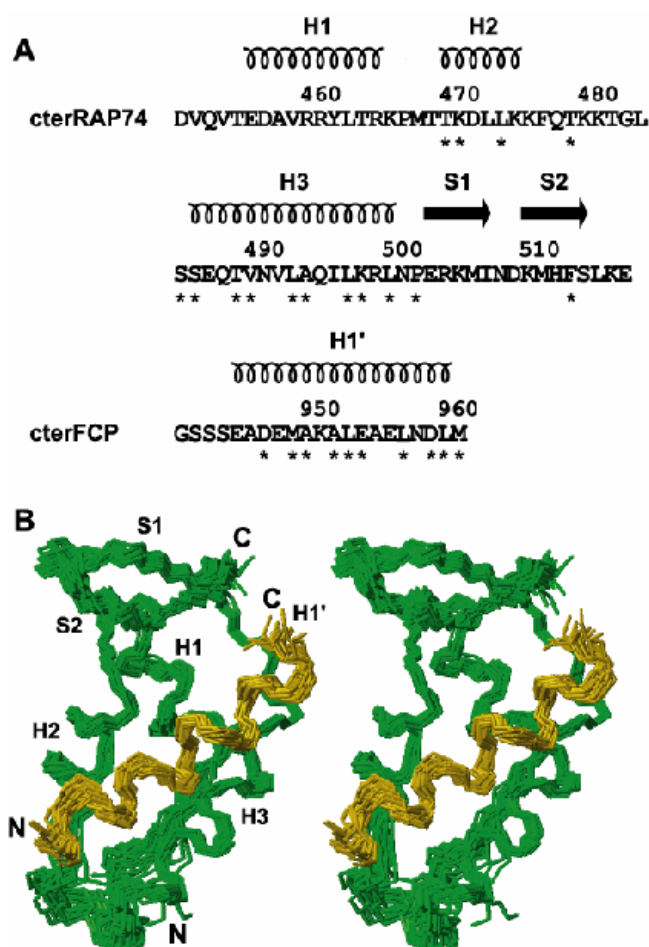


Figure 1.3. Sequence and structure of RAP74 and the carboxyl-terminus of FCP1. (A) Amino acid sequence and secondary structure of RAP74 (436-517) and FCP1 (879-961). Amino acids showing intermolecular NOEs are indicated by an asterisk. (B) A stereoview of the 20 lowest-energy NMR structures showing the backbone trace (N, C α , and C') of residues 436-517 of RAP74 (green) and residues 945-961 of FCP1 (gold) (Nguyen et al. 2003a)

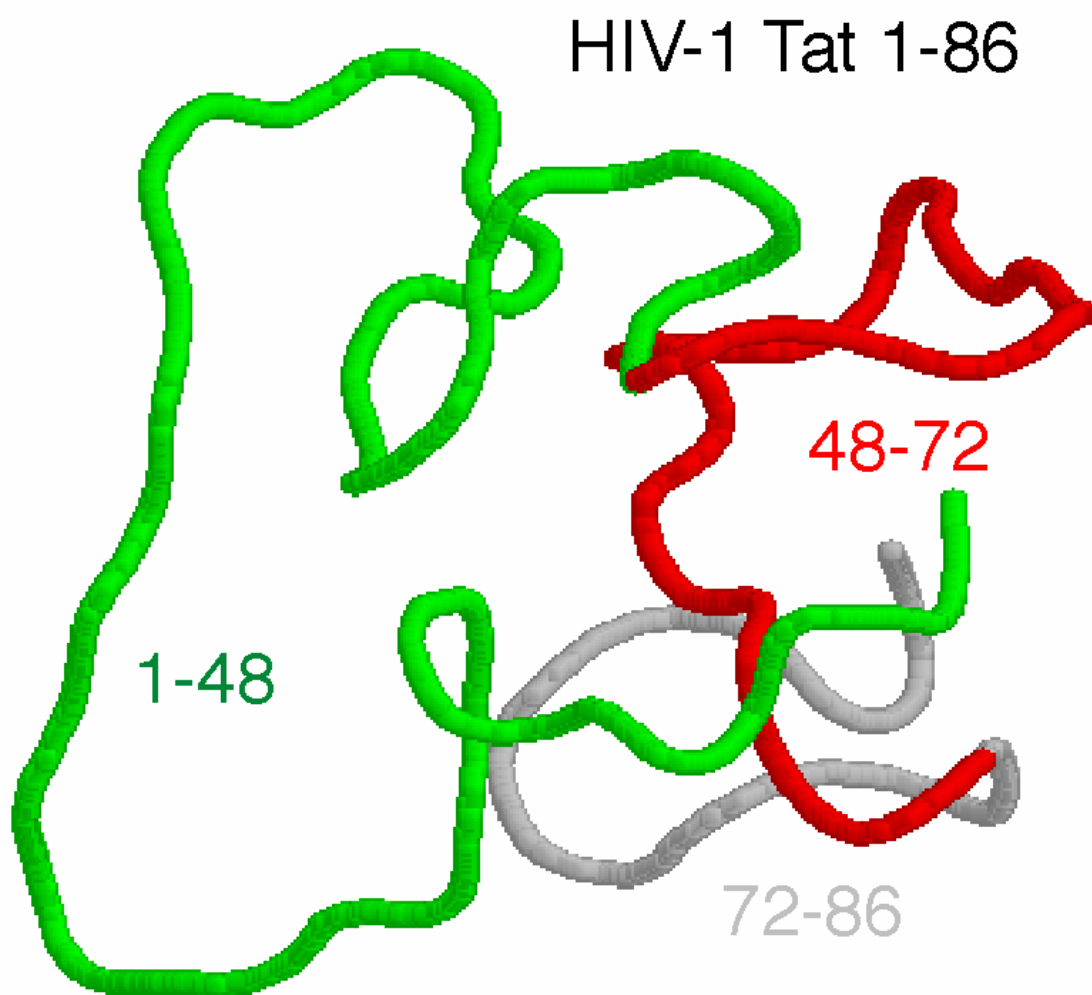


Figure 1.4. Representation of the HIV-1 Tat structure. Three distinct domains of varying flexibility are shown: the cysteine-rich activation domain (green), the basic rich region (red), and the glutamine rich region (grey). This figure is a C α trace adapted from data previously published (Bayer et al. 1995).

CHAPTER 2

INTERACTIONS OF THE HIV-1 TAT AND RAP74 PROTEINS WITH THE RNAPII CTD
PHOSPHATASE FCP1₁

¹Karen L. Abbott, Jacques Archambault, Hua Xiao, Bao D. Nguyen, Robert G. Roeder, Jack Greenblatt, James G. Omichinski, and Pascale Legault. To be submitted to J. Biol. Chem.

ABSTRACT

FCP1, a phosphatase specific for the carboxyl-terminal domain of the largest subunit of RNA polymerase II, is regulated *in vitro* by the HIV-1 Tat protein, CK2 and the large subunit of TFIIF (RAP74). Whereas Tat inhibits FCP1 phosphatase activity, CK2, and RAP74 increase phosphatase activity. We have characterized the interactions of Tat and RAP74 with the BRCT-containing central domain of FCP1 (FCP1₅₆₂₋₇₃₈). We demonstrate that FCP1 is required for Tat-mediated transactivation *in vitro* and that amino acids 562-738 of FCP1 are necessary for Tat interaction *in vivo*. From sequence alignments, we identified a conserved acidic/hydrophobic region in FCP1 adjacent to its highly conserved BRCT domain. *In vitro* binding studies with purified proteins indicate that HIV-1 Tat interacts with both the acidic/hydrophobic region and the BRCT domain of FCP1, whereas RAP74₄₃₆₋₅₁₇ interacts solely with a portion of the acidic/hydrophobic region containing a conserved LXXLL-like motif. HIV-1 Tat inhibits the binding of RAP74₄₃₆₋₅₁₇ to FCP1. In a parallel study, we identified a novel CK2 site adjacent to this conserved LXXLL-like motif. Phosphorylation of FCP1₅₆₂₋₆₁₉ by CK2 at this site increases binding to RAP74₄₃₆₋₅₁₇, but this phosphorylation is inhibited by Tat. These results suggest that Tat can regulate the FCP1 phosphatase activity both by preventing RAP74 binding and inhibiting CK2 phosphorylation. In addition to providing new insights into the role of HIV-1 Tat in transcriptional regulation, this study defines a clear role for regions adjacent to the BRCT domain in promoting important protein-protein interactions.

INTRODUCTION

Transcription by RNA polymerase II (RNAPII)¹ is a complex process regulated in part by the phosphorylation of the carboxyl-terminal domain (CTD) of the largest subunit of RNAPII (1,2). The RNAPII CTD contains several heptapeptide repeats that are phosphorylated during the transcription cycle (3). Transcription begins with the recruitment of the hypophosphorylated form of the polymerase. Following the assembly of the initiation complex, several kinases, such as cdk7 and cdk9, are recruited, allowing hyperphosphorylation of the CTD (4,5). The hyperphosphorylated form of the polymerase is the form primarily found in elongating complexes (1,6). Upon completion of transcription, the CTD must be dephosphorylated before recycling of the polymerase can occur. FCP1 is capable of processively dephosphorylating the CTD of RNAPII at Ser² and Ser⁵ (7,8) and is an essential part of the preinitiation complex (9-11). FCP1 has also been shown to enhance elongation, and this function is independent of the phosphatase activity (9,12). More recently, it has been demonstrated that FCP1 is present in active elongating complexes and may help modulate phosphorylation levels of the elongating polymerase (13).

Three domains have been identified in FCP1 based on sequence similarity between human FCP1 (hFCP1) and yeast FCP1 (yFCP1) (11). The amino terminus contains the FCP homology region with the conserved $\Psi\Psi\Psi\text{D}\text{X}\text{D}\text{X}(\text{T}/\text{V})\Psi\Psi$ motif (Ψ is a hydrophobic residue) characteristic of a family of small molecule phosphotransferases and phosphohydrolases (14,15). Yeast studies revealed that mutation of the first aspartic acid of this motif abolishes catalytic activity (7) and strains carrying this mutation in FCP1 are not viable, indicating that the phosphatase activity of FCP1 is essential in yeast (14). Yeast studies have also revealed that the central domain of FCP1, containing the Brca-1 carboxyl-terminal (BRCT) domain, is essential

for cell viability and phosphatase activity (7,16). The BRCT domain is an independently folded domain of ~95 amino acid residues that is conserved as a single domain or as multiple repeats in proteins and appears important for mediating protein-protein interactions. BRCT domains have been identified in over 450 proteins, which function in cell cycle regulation, DNA repair, recombination, chromatin remodeling, and transcription activation (17-23). Recently, it was found that the BRCT domains of Brca1 and PTIP specifically recognize phosphorylated protein targets, defining a new role for the BRCT domain in cellular signaling (24,25). Several high-resolution structures of various BRCT domains have been determined either as single domains (26-29) or as tandem repeats (28,30-32). The BRCT fold is characterized by a four-stranded parallel β -sheet surrounded by two or three α -helices. In tandem BRCT domains, the two domains pack together in a head-to-tail arrangement to form a stable compact structure (30) (28,32). The carboxyl-terminal domain of FCP1 is highly acidic and is capable of activating transcription when tethered to a promoter by fusion to a heterologous DNA-binding domain (16). The carboxyl-terminus of FCP1 (FCP1₈₇₉₋₉₆₁) directly interacts with the carboxyl-terminal domain of RAP74 (RAP74₄₃₆₋₅₁₇) (11). High-resolution structures of the complex formed by these two domains have been determined by X-ray crystallography (33) and NMR spectroscopy (34).

Several factors have been shown to influence FCP1 phosphatase activity. The largest subunit of the general transcription factor TFIIF (RAP74) has been shown to stimulate FCP1 phosphatase activity (35), whereas the general transcription factor TFIIB prevents the stimulation by TFIIF (35). The HIV-1 transcription activator protein Tat inhibits the dephosphorylation of the CTD by FCP1 (36). Interestingly, Tat also interacts with the cellular kinase complex P-TEFb (cyclin T1 component) and has been shown to stimulate both Ser² and Ser⁵ phosphorylation of

the CTD (37). FCP1 also contains multiple CK2 (formerly known as casein kinase II) phosphorylation sites, and a recent study demonstrates that phosphorylation of *Xenopus laevis* FCP1 by CK2 *in vitro* enhances phosphatase activity (38). In this study, we characterized the interactions of the central domain of FCP1 (FCP1₅₆₂₋₇₃₈) with the HIV-1 Tat protein and the carboxyl-terminal domain of the RAP74 subunit of TFIIF. We also analyzed the effect of HIV-1 Tat on CK2 phosphorylation of purified FCP1₅₆₂₋₆₁₉ *in vitro*. Our results suggest a mechanism whereby Tat can inhibit FCP1 phosphatase activity by both preventing RAP74 interaction and inhibiting CK2 phosphorylation with the central domain of FCP1. This study brings a new understanding for the role of HIV-1 Tat in transcriptional regulation and defines a role for regions adjacent to the BRCT domain in promoting important protein-protein interactions.

MATERIALS and METHODS

Antibodies.

The HIV-1 Tat monoclonal antibody (15.1) and the Tat polyclonal antibody (705) used in this study were obtained through the NIH AIDS Research and Reference Reagent Program (39,40). The RAP74 polyclonal antibody (C-18) and all secondary antibodies were purchased from Santa Cruz Biotechnologies (Santa Cruz, CA). The polyhistidine monoclonal antibody was purchased from Novagen (Madison, WI).

Plasmids.

Gal4 yeast expression constructs were prepared for the yeast two-hybrid studies and details will be provided upon request. All plasmids expressing the glutathione-S-

transferase(GST)-human FCP1 fragments used for *in vitro* binding studies were constructed by PCR amplification using pJA533 as the template (11). The amplified DNA fragments were digested with *Bam*H1 and *Eco*R1, and then inserted in frame into the *Bam*H1 and *Eco*R1 sites of the pGEX5X-1 GST-fusion expression vector (Amersham Biosciences, NJ). The pGEX2T expression vectors for HIV-1 Tat proteins (Tat₁₋₈₆, Tat₁₋₇₂, Tat₁₋₄₈, and Tat_{1-48C22G}) (41) were obtained from the AIDS Research and Reference Reagent Program. The pGEX2T expression vector for HIV-1 Tat₄₈₋₈₆ was constructed by PCR amplification of Tat cDNA using pGEX2T HIV-1 Tat₁₋₈₆ as the template. The PCR fragments were digested with *Bam*H1 and *Eco*R1 and inserted in frame into the *Bam*H1 and *Eco*R1 sites of the GST expression vector pGEX2T (Amersham Biosciences, NJ). An expression vector for the hexahistidine (His₆)-Tat₁₋₇₂ fusion protein was prepared by inserting a PCR-amplified Tat cDNA cut with *Bgl*II into the *Bam*H1 site of pRSET-C (Invitrogen, CA). The expression vector for RAP74₄₃₆₋₅₁₇ has been described previously (42). The alanine substituted mutants GST-FCP1₈₇₉₋₉₆₁ (E956A and L957A) (34), GST-FCP1₅₆₂₋₇₃₈ and GST-FCP1₅₆₂₋₅₉₉ (Y592A and L593A), as well as the CK2 site mutant GST-FCP1₅₆₂₋₆₁₉ (T584E) were prepared by site-directed mutagenesis (Stratagene, CA). The expression vectors for preparing purified FCP1₅₆₂₋₆₁₉ and thioredoxin-His₆-FCP1₅₆₂₋₇₃₈ were created by inserting the *Bam*H1/*Eco*R1 fragments from pGEX5X-1 FCP1₅₆₂₋₆₁₉ and pGEX5X-1 FCP1₅₆₂₋₇₃₈, respectively, into the *Bam*H1/*Eco*R1 sites of the pET32C (Novagen, WI) expression vector. The pHIV+TARG400 template used for *in vitro* transcription experiments has been described previously (43). The expression vector for the carboxyl-terminal histidine-tagged full-length hFCP1 was generously provided by Danny Reinberg. All plasmids created for this study were validated by DNA sequencing.

Immunodepletion of FCP1 from HeLa Nuclear Extract.

Five-hundred μ l of HeLa nuclear extract prepared as previously described (44) were incubated with 1 μ l of affinity-purified polyclonal FCP1 antibody or control IgG (Bio-Rad, CA) for 1 h at 4 °C. Extracts were loaded onto a 20 μ l protein A-Sepharose columns that had been equilibrated with 400 μ l BC100 buffer (20 mM Tris-HCl pH 7.9, 0.2 mM EDTA, 0.2 mM DTT, 20% glycerol, and 100 mM KCl) (45) containing 250 mg/ml BSA prior to loading. The flow-through fractions were aliquoted and stored at -80 °C until needed for *in vitro* transcription assays with pHIV+TARG400.

***In Vitro* Transcription Assay.**

In vitro transcription reactions with Tat₁₋₇₂ were performed as described previously (43,46) with some modifications. For each 25 μ l reaction, we used 100 ng of supercoiled pHIV+TARG400 template, 30 μ g of nuclear extract that was FCP1 or IgG depleted, and various amounts of Tat₁₋₇₂. The His₆-Tat₁₋₇₂ protein used for these studies was purified under denaturing conditions. The reactions were incubated at 30 °C for 30 min prior to the addition of 1 μ l of [α ³²P]-CTP 3000 Ci/mM (Amersham Biosciences, NJ). The reactions were further incubated for 30 min and then stopped with the addition of 1 μ l of 250 mM EDTA. Each reaction was then supplemented with 2 μ l of RNase T1 (5 U/ μ l) and incubated at 37 °C for 10 min. The G-less RNA transcripts were resolved on a denaturing polyacrylamide gel. Gels were vacuum-dried and ³²P-labeled RNA was detected by PhosphorImager analysis (Molecular Dynamics-Amersham Biosciences, NJ).

Yeast Two-Hybrid Assay.

Yeast manipulations and growth media were described previously (47). To test the ability of yeast cells to grow on medium containing 3-aminotriazole (AT), 3 μ l of a cell suspension (approximately 2,000 cells) were applied onto SD medium (BD Biosciences, NJ) containing 100 μ g/ml adenine and 40 mM AT. As a control, a similar number of cells were applied onto medium lacking AT, but containing 100 μ g/ml of adenine and histidine.

Protein Purification.

The full-length FCP1 used for *in vitro* transcription of pHIV+TARG400 was purified from HeLa cells as described previously (11). Recombinant full-length His₆-tagged human FCP1 used for binding studies was purified from baculovirus-infected insect cells as described previously for yeast FCP1 (14). GST-FCP1 fragments were expressed in DH5 α (Stratagene, CA). These cells were grown at 37 °C, and protein expression was induced with 0.7 mM isopropyl-beta-D-thiogalactoside (IPTG) for 3 h at 30 °C. The cells were harvested and resuspended in EBC buffer (50 mM Tris-HCl pH 8.0, 120 mM NaCl, 0.5% NP-40, and 2 mM DTT) supplemented with protease inhibitors (Sigma, MO): 1 μ g/ml leupeptin, 2 μ g/ml aprotinin, and 0.9 mg/ml PMSF. Cells were lysed by passage through a French press cell and centrifuged at 100,000 g for 40 min. The supernatant obtained from the high-speed centrifugation was added to glutathione(GSH)-Sepharose resin (Amersham Biosciences, NJ) and incubated with mixing at 4 °C for 30 min. The resin was collected by centrifugation and washed four times with wash buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5% NP-40, 0.05% SDS, and 1 mM DTT). GST-FCP1 fragments were eluted off the resin with 15 mM reduced glutathione (Sigma, MO) for 10 min at 25 °C, dialyzed into protein storage buffer (50 mM Tris-HCl pH 8.0,

100 mM NaCl, 2 mM DTT, and 20% glycerol), aliquoted and stored at -80 °C. GST-FCP1 fragments were rebound to fresh GSH-Sepharose resin prior to protein-protein binding studies.

All HIV-1 Tat proteins used for binding studies were expressed as GST-Tat constructs and purified as stated for the GST-FCP1 fragments up to the elution step. At this step, Tat proteins were cleaved from the GST using thrombin (Calbiochem, CA) and subsequently dialyzed into 5% acetic acid and loaded on a C4-reverse phase HPLC column (Vydac, CA). Tat proteins were eluted from the column using an acetonitrile gradient (25% to 45% over 20 min) in 0.05% TFA at a flow rate of 8 ml/min. Purified Tat proteins were stored lyophilized in aliquots at -20 °C. The GST-Tat₁₋₈₆ used in binding studies with His₆-tagged FCP1 was purified as described for GST-FCP1 fragments up to the end of the high-speed centrifugation step. The supernatant from the high-speed centrifugation was snap frozen in liquid nitrogen and stored at -80 °C until needed. Prior to each experiment, this partially-purified GST-Tat₁₋₈₆ was freshly purified by capture and washes on GSH-Sepharose resin as described for the GST-FCP1 fragments.

The RAP74₄₃₆₋₅₁₇ was purified as described previously (42). The FCP1 peptides (FCP1₅₇₉₋₆₀₀ and FCP1₅₈₃₋₆₀₇) used for the protein-protein binding studies and for the 2D ¹H-¹⁵N heteronuclear single quantum coherence (HSQC) spectrum were chemically synthesized at the Medical College of Georgia with an extra glycine residue at the amino terminus. The peptides were purified to homogeneity on a C-4 reverse phase HPLC column using an acetonitrile gradient (30% to 50% over 20 min) in 0.05% TFA at a flow rate of 8 ml/min. The purified FCP1₅₇₉₋₆₀₀ and FCP1₅₈₃₋₆₀₇ peptides were stored lyophilized in aliquots at -20 °C.

The thioredoxin-His₆-FCP1₅₆₂₋₇₃₈ fusion protein used for *in vitro* binding studies was obtained by expression in BL21(DE3). The cells were induced with 1 mM IPTG for 4 h at 37 °C.

The cells were harvested, lysed by passage through a French press in PE buffer (25 mM phosphate pH 7.2, 1 mM EDTA), and centrifuged at 15,000 g for 30 min. The pellet, which contained the protein in an inclusion body, was washed twice at 25 °C in PE buffer supplemented with 1 M urea and 0.5% triton X-100 to remove contaminants. The washed pellet was then solubilized in 6 M guanidinium HCl, 10 mM Tris pH 8.0 buffer for 15 min at 50 °C and centrifuged at 100,000 g for 40 min at 4 °C. The fusion protein from the high-speed supernatant was first bound to Ni²⁺-charged Chelating Sepharose Fast Flow resin (Amersham Biosciences, NJ). The captured fusion protein was gradually refolded on the resin at 25 °C by stepwise incubation in refolding buffer A (10 mM Tris pH 8.0, 500 mM NaCl, 20% glycerol, 1 mM β -mercaptoethanol) supplemented with 4 M, 2 M, and 1 M urea. The fusion protein was eluted from the resin using 250 mM imidazole in phosphate buffered saline, dialyzed overnight into protein storage buffer, and aliquots were stored at -80 °C. The FCP1₅₆₂₋₆₁₉ protein used in the CK2 assay was obtained by expression of a thioredoxin-His₆-FCP1₅₆₂₋₆₁₉ fusion protein in BL21(DE3). The cells were induced with 1 mM IPTG for 4 h at 37 °C. The thioredoxin-His₆-FCP1₅₆₂₋₆₁₉ was purified from the inclusion body under denaturing conditions, as described for the thioredoxin-His₆-FCP1₅₆₂₋₇₃₈. After the high-speed centrifugation, the fusion protein from the supernatant was refolded by stepwise dialysis at 4 °C in refolding buffer A supplemented with 4 M, 2 M, and 1 M urea and finally in phosphate buffered saline. FCP1₅₆₂₋₆₁₉ was then cleaved from thioredoxin-His₆ with thrombin (Calbiochem, CA). The FCP1₅₆₂₋₆₁₉ protein was applied to a Q-Sepharose High Performance (Amersham Biosciences, NJ) column (100 ml bed volume) equilibrated with buffer A (20 mM phosphate buffer pH 7.2, 1 mM DTT, and 1 mM EDTA). FCP1₅₆₂₋₆₁₉ was eluted from the column using a gradient (from 0% to 100% over 700 ml) of buffer B (20 mM phosphate pH 7.2, 1 mM DTT, 1 mM EDTA, and 1 M NaCl) and dialyzed

overnight at 4 °C into 3% acetic acid. The eluent volume was reduced using a rotovap, and FCP1₅₆₂₋₆₁₉ was applied to a C-4 reverse-phase HPLC column (Vydac, CA). FCP1₅₆₂₋₆₁₉ was eluted from this column using an acetonitrile gradient (30% to 50% over 20 min) in 0.05% TFA at a flow rate of 8 ml/min. Eluted FCP1₅₆₂₋₆₁₉ was dialyzed overnight into protein storage buffer, and aliquots were stored at -80 °C. Protein concentrations were estimated using a combination of UV spectroscopy at 280 nm (48) and comparison with known amounts of bovine serum albumin (Promega) on Coomassie blue-stained gels.

Protein-Protein Binding Assay.

For all *in vitro* binding experiments with immobilized GST-FCP1 we used the following procedure. GST-FCP1 fragments were first coupled to 10 µl of GSH-Sepharose resin (Amersham Biosciences, NJ) in 500 µl of EBC buffer for 1 h. The bound GST-FCP1 was then collected by centrifugation, washed with EBC buffer, and equilibrated with binding buffer PP (40 mM HEPES pH 7.9, 120 mM NaCl, 0.5% NP-40, 10 mM DTT). The binding reactions contained equimolar amounts of various GST-FCP1 fragments and purified HIV-1 Tat proteins (0.5 µM Tat₁₋₈₆, Tat₁₋₇₂, Tat₁₋₄₈, Tat_{1-48C22G}, or Tat₄₈₋₈₆) or RAP74₄₃₆₋₅₁₇ (10 µM). Binding reactions were performed in 500 µl of binding buffer PP at 4 °C for 1 h. For binding with Tat proteins, this buffer was supplemented with 80 µM zinc sulfate. The protein-protein complexes were collected by centrifugation and then washed (2 to 4 times) with binding buffer supplemented with 350 mM NaCl (for binding assays including RAP74 no additional salt was added). The washed pellet was resuspended in NuPage LDS 4X sample buffer (Invitrogen, CA), and proteins were resolved on a 12% Bis-Tris NuPage gel using 1X MES electrophoresis buffer (Invitrogen, CA). Proteins were transferred to a PVDF membrane (Millipore, MA) using the

Invitrogen transfer system. The Tat₁₋₇₂, Tat₁₋₄₈, and Tat_{1-48C22G} proteins were detected using a 1:1,000 dilution of the Tat monoclonal antibody (15.1) and a 1:10,000 dilution of anti-mouse horseradish peroxidase (HRP) conjugated secondary antibody. Tat₄₈₋₈₆ was detected using a 1:5,000 dilution of the Tat polyclonal antibody (705) and a 1:20,000 dilution of the anti-rabbit HRP conjugated secondary antibody. The RAP74₄₃₆₋₅₁₇ protein was detected using a 1:400 dilution of the RAP74 polyclonal antibody C-18 and a 1:7,000 dilution of anti-rabbit HRP conjugated secondary antibody.

For binding reactions with immobilized GST-Tat (Fig. 4A), control GST or GST-Tat₁₋₈₆ were purified on GSH-Sepharose as described above. Binding experiments were performed in 500 µl of binding buffer PP containing 0.5 µM purified His₆-FCP1₁₋₉₆₁ or thioredoxin-His₆-FCP1₅₆₂₋₇₃₈ and 0.5 µM GST-Tat₁₋₈₆. Bound His₆-FCP1 proteins were collected and washed (2 to 4 times) in binding buffer PP supplemented with 350 mM NaCl. The washed pellet was resuspended in NuPage LDS 4X sample buffer, and proteins were resolved on a NuPage 4-12% Bis-Tris gradient gel in 1X MES electrophoresis buffer. Proteins were transferred to an Immobilon-P membrane (Millipore, MA) and detected using a 1:1,000 dilution of a His₆ monoclonal antibody (Novagen, WI) and a 1:10,000 dilution of an HRP-conjugated secondary antibody. The bound secondary antibodies were detected by chemiluminescence using the ECL-Plus kit (Amersham Biosciences, NJ), and all membranes were subsequently stained with Coomassie blue to verify equivalent GST-fusion protein input.

NMR Spectroscopy.

The ¹⁵N-labeled RAP74₄₃₆₋₅₁₇ was expressed from a pGEX2T vector (Amersham Biosciences, NJ) in BL21(DE3). Labeled protein was obtained by growth in a modified minimal

medium containing ^{15}N -labeled NH_4Cl as the sole source of nitrogen (42). The NMR sample consisted of 1 mM ^{15}N -labeled RAP74₄₃₆₋₅₁₇ / 1 mM unlabeled FCP1₅₇₉₋₆₀₀ in 500 μl of 20 mM sodium phosphate pH 6.5 and 1 mM EDTA (90% H_2O , 10% D_2O). The 2D ^1H - ^{15}N HSQC (49) was collected at 27 °C using Varian Unity^{Inova} 600 MHz NMR spectrometer equipped with a z pulsed-field gradient unit and a HCN triple resonance probe. The NMR data was processed with NMRPipe / NMRDraw (50).

Phosphorylation of FCP1 Fragments with CK2.

For phosphorylation of FCP1, the FCP1₅₆₂₋₆₁₉ protein (or control protein) was first diluted at 1.25 mg/ml in 20 μl of 1.25X phosphorylation buffer (20 mM Tris-HCl pH 7.5, 25 mM KCl, 10 mM MgCl_2 , 350 μM ATP, 2.5 μCi [$\gamma^{32}\text{P}$]-ATP (4500 Ci/mmol)). The reaction was initiated by the addition of 10 units (5 μl) of CK2 enzyme (New England Biolabs, MA) diluted in 20 mM Tris-HCl pH 7.5, 200 mM NaCl, 0.5 mM DTT, 10% glycerol, 0.5% triton X-100. The phosphorylation reactions were incubated at 25 °C for 30 min and terminated by the addition of NuPage LDS 4X sample buffer. For mock phosphorylation, the reaction was performed similarly, but 1X CK2 buffer replaced the CK2 enzyme. Proteins were resolved on 12% Bis-Tris NuPage gels using 1X MES electrophoresis buffer (Invitrogen, CA). The phosphorylated proteins were detected by PhosphorImager analysis (Molecular Dynamics-Amersham Pharmacia, NJ). Phosphorylation of GST-FCP1 fusion protein used in protein-protein binding assays was obtained by capturing GST-FCP1₅₆₂₋₆₁₉ (or control GST) on GSH-Sepharose resin and exchanging into CK2 phosphorylation buffer. Following the CK2 phosphorylation reaction, the resin was washed (2 to 4 times) in binding buffer PP.

RESULTS

The FCP1 / Tat interaction is essential for Tat-mediated transcription *in vitro*.

Although previous studies have shown that Tat is capable of inhibiting the phosphatase activity of FCP1 *in vitro* (36), the influence of FCP1 on Tat-mediated transcription activation has remained elusive. *In vitro* transcription assays using the pHIV+TARG400 template and either IgG-depleted or FCP1-depleted nuclear extracts were used to assess the role of FCP1 in Tat-mediated transcription (Fig. 1). HeLa cell nuclear extracts could be completely depleted of FCP1 using FCP1 antibodies prior to the transcription assays (Fig. 1A). In the presence of HeLa cell nuclear extract containing FCP1 (IgG-depleted extract) there was an increase in pHIV+TARG400 transcription with increasing concentration of Tat (Fig. 1B lanes 1-3). However, in the presence of FCP1-depleted extracts there was no increase in pHIV+TARG400 transcription with increasing concentration of Tat (Fig. 1B lanes 4-6). When purified FCP1 was added back to the transcription reaction, we observed an increase in transcription from the pHIV+TARG400 template in the presence of Tat (Fig. 1C). As expected, the immunodepletion of FCP1 from nuclear extracts did not remove RAP74 (Fig. 1A), because the antibody recognition site overlaps with the RAP74-binding site of FCP1. Furthermore, it is unlikely that immunodepletion of FCP1 removed the RNAPII holoenzyme from the HeLa nuclear extract since the levels of human SRB7, an intrinsic component of the holoenzyme, are similar in both IgG-depleted and FCP1-depleted extracts (data not shown). Therefore, the reduction in Tat-mediated transcription in the FCP1-immunodepleted extracts is not due to removal of either RAP74 or RNAPII. Since the HIV-1 promoter of the pHIV+TARG400 template contains binding sites for Sp1, which are important for activation by Tat (51,52), we tested if the depletion of FCP1 had any effect on SP1-activated transcription. Our studies revealed that

activation of a core promoter containing upstream GAL4-binding sites by the chimeric activator protein GAL4-Sp1 was unaffected by removal of FCP1 from the nuclear extract (data not shown). The results presented in Fig. 1 indicate that FCP1 is an essential cellular cofactor for Tat-mediated transcription *in vitro*.

HIV-1 Tat interacts with FCP1 *in vivo*.

Yeast-two hybrid studies were performed using various fragments of human FCP1 fused to the GAL4 activation domain and HIV-1 Tat₁₋₇₂ fused to the GAL4 DNA-binding domain (Fig. 2). These studies demonstrate that Tat₁₋₇₂ interacts with a carboxyl-terminal fragment of human FCP1 (FCP1₅₆₂₋₉₆₁) *in vivo*. In addition, carboxyl-terminal deleted fragments of FCP1₅₆₂₋₉₆₁ (FCP1₅₆₂₋₈₇₈, FCP1₅₆₂₋₇₈₈, and FCP1₅₆₂₋₆₈₅) retain *in vivo* binding to HIV-1 Tat. However, further carboxyl-terminal and amino-terminal deletions that yield the FCP1₅₆₂₋₆₃₇ and FCP1₆₉₈₋₉₆₁ fragments lead to impaired Tat binding in yeast two-hybrid studies. It therefore appears that the central domain of FCP1 (FCP1₅₆₂₋₆₈₅), which is essential for cell viability and phosphatase activity (16), is also sufficient for Tat binding *in vivo* (Fig. 2).

Sequence conservation in the central domain of FCP1.

Prior to initiating *in vitro* binding studies with purified human FCP1 fragments, conserved amino acid regions within the central domain of FCP1 were identified using Pfam (53) and BLAST (54) searches (Fig. 3). The boundaries of the BRCT domain (FCP1₆₃₆₋₇₂₈) were initially defined with the Pfam multiple alignment program (53) (data not shown). We found that the BRCT domain of human FCP1 is defined by amino acids 636 to 728 (Fig. 3A). The BRCT domains of published FCP1 sequences were subsequently aligned with BLAST (Fig. 3A). The

percentage of sequence identity to human FCP1 within the BRCT domain is very high for *M. musculus* (95%) and *X. laevis* (80%) and lower (27% to 42%) for other eukaryotic sequences (Fig. 3A). Interestingly, the region of FCP1 located immediately at the amino-terminus of the BRCT domain (FCP1₅₆₂₋₆₃₅) also showed considerable sequence identity among eukaryotic FCP1 proteins. We defined this region as the acidic/hydrophobic region of FCP1 (FCP1₅₆₂₋₆₃₅) based on the high occurrence of acidic and hydrophobic amino acids. The percentage of sequence identity to human FCP1 within the acidic/hydrophobic region is also very high for *M. musculus* (91%) and *X. laevis* (84%) and lower (23% to 38%) for other eukaryotic sequences (Fig. 3B). The exceptionally high level of sequence identity within the BRCT domain and the acidic/hydrophobic region of FCP1 from human, mouse, and frog suggest that the central domain of FCP1 may have a distinct function(s) in vertebrates.

HIV-1 Tat interacts directly with two non-overlapping regions of the central domain of FCP1 *in vitro*.

The yeast two-hybrid results revealed that HIV-1 Tat₁₋₇₂ interacts *in vivo* with a minimal FCP1 fragment including residues 562 to 685 and that further amino-terminal and carboxyl-terminal deletions of the FCP1₅₆₂₋₉₆₁ fragment abrogates binding. We attempted to define the minimal Tat recognition motif of FCP1 by using an *in vitro* batch-affinity assay with purified proteins and protein fragments (Figs. 4 and 5). Initial binding experiments were performed using His₆-tagged FCP1₁₋₉₆₁ and thioredoxin-His₆-FCP1₅₆₂₋₇₃₈ to compare the relative affinity of these proteins for GST-Tat₁₋₈₆ immobilized on GSH-Sepharose beads. In this assay, His₆-FCP1₁₋₉₆₁ (Fig. 4A, lanes 1-3) and thioredoxin-His₆-FCP1₅₆₂₋₇₃₈ (Fig. 4A, lanes 4-6) appear to have similar

affinity for Tat₁₋₈₆. These results indicate that the central domain of FCP1 is sufficient for interaction with the Tat protein.

Several carboxyl- and amino-terminal truncated versions of FCP1₅₆₂₋₇₃₈ were then purified as GST-fusion proteins in *E. coli* and assayed for Tat₁₋₇₂ binding by using our *in vitro* batch-affinity binding assay. Carboxyl-terminal deletion fragments FCP1₅₆₂₋₆₃₇ and FCP1₅₆₂₋₆₁₉ retained the same level of binding to Tat₁₋₇₂ as FCP1₅₆₂₋₇₃₈. (Figure 4B, Lanes 1-3) However, FCP1₅₆₂₋₅₉₉ bound Tat₁₋₇₂ with lower affinity (Fig. 4B lane 4), indicating that FCP1₅₆₂₋₆₁₉, which comprises most of the acidic/hydrophobic region, is a minimal Tat binding domain. Since deletions of FCP1 residues 599 to 619 caused a drastic change in binding, we tested the binding of Tat₁₋₇₂ to GST-FCP1₅₉₉₋₆₃₇ (Fig. 4B, lane 5). Since no binding was detected with this fragment, we concluded that although residues 599-619 are important, they are not sufficient for binding. Amino-terminal deletion fragments of FCP1₅₆₂₋₇₃₈, FCP1₆₂₆₋₇₃₈ and FCP1₅₉₉₋₇₃₈ also bound Tat, although poorly (Fig. 4B lanes 6-7), indicating that FCP1₆₂₆₋₇₃₈, which contains the BRCT domain, binds independently although with lower affinity to Tat₁₋₇₂. Similar results were obtained using Tat₁₋₈₆ instead of Tat₁₋₇₂ (data not shown). These binding studies reveal that there are two non-overlapping Tat binding sites in the central domain of FCP1, one located within the acidic/hydrophobic region (562-619), and the other within the 626-738 fragment that contains the BRCT domain (636-728).

Additional binding studies were performed using Tat₁₋₄₈ or a transactivation-negative C22G mutant of Tat₁₋₄₈ (36,55) and the same GST-FCP1 deletion fragments used with Tat₁₋₇₂ (Fig. 4C). The Tat₁₋₄₈ fragment contains only the activation domain, but not the arginine-rich and glutamine-rich regions present in residues 48-72 of Tat (56). The binding results with Tat₁₋₄₈ are similar to those obtained with Tat₁₋₇₂, with the exception that Tat₁₋₄₈ binds better than Tat₁₋₇₂ to

FCP1₆₂₆₋₇₃₈ and FCP1₅₉₉₋₇₃₈ (Fig. 4C, lanes 1-8), indicating that Tat residues 48-86 prevent an optimal interaction of Tat₁₋₄₈ with the BRCT domain (FCP1₆₃₆₋₇₂₈). Binding was not observed between the transactivation-negative C22G mutant of Tat₁₋₄₈ (Tat₁₋₄₈C22G) and any of the GST-FCP1 fragments (Fig. 4C, lanes 9-16), suggesting that the interactions with wild-type Tat proteins are specific.

To further explore the interaction of Tat₁₋₄₈ with the BRCT domain of FCP1, several deletion fragments of GST-FCP1₆₂₆₋₇₃₈ were purified and used for *in vitro* binding studies with purified Tat₁₋₄₈ (Fig. 4D). An amino-terminal deletion fragment of the FCP1 BRCT domain (GST-FCP1₆₃₉₋₇₃₈) was tested for Tat₁₋₄₈ binding. This more strictly defined BRCT domain binds Tat₁₋₄₈ as well as GST-FCP1₆₂₆₋₇₃₈ (data not shown). However, we detected little binding with FCP1₆₂₆₋₇₂₃ and FCP1₆₂₆₋₇₁₁, and no binding with FCP1₆₂₆₋₆₉₈, indicating that the carboxyl-terminal portion of the BRCT domain (698-738) is essential for optimal Tat₁₋₄₈ binding. Although truncation of the BRCT domain may destabilize the overall fold of the protein fragment, the carboxyl-terminal deletions used in this study were planned using a predicted secondary structure alignment to minimize truncations within a secondary structure element (data not shown) (57).

We noted another difference in the binding results between full-length Tat and Tat₁₋₄₈, which is that Tat₁₋₇₂ (Fig. 4B lane 4) and Tat₁₋₈₆ (not shown) interact weakly with GST-FCP1₅₆₂₋₅₉₉, whereas no such interaction was observed for Tat₁₋₄₈ (Fig. 4C lane 5). These results suggest that a binding site for the carboxyl-terminal region of Tat (Tat₄₈₋₈₆) is located within FCP1₅₆₂₋₅₉₉. To further investigate this possibility, the Tat₄₈₋₈₆ fragment was purified and used in our binding studies with various GST-FCP1 fragments (Fig. 4E). Binding was not detected between Tat₄₈₋₈₆ and the BRCT-containing fragment FCP1₅₉₉₋₇₃₈ (Fig. 4E, lane 7). Tat₄₈₋₈₆ only bound to FCP1

fragments FCP1₅₆₂₋₆₁₉, FCP1₅₈₄₋₅₉₉, and FCP1₅₆₂₋₅₉₉ that contained the acid-rich portion of the acidic/hydrophobic region (Fig. 4E, lanes 3-6). This is not surprising given that Tat₄₈₋₈₆ is positively charged. Unexpectedly, the binding of Tat₄₈₋₈₆ to GST-FCP1₅₆₂₋₇₃₈ was weaker than the binding to GST-FCP1₅₆₂₋₅₉₉ (Fig. 4E lanes 3 and 6), suggesting that intra-molecular interactions between the acidic/hydrophobic region and the BRCT domain prevent optimal binding of Tat₄₈₋₈₆ (see Discussion). It is also noteworthy that Tat₁₋₇₂, which binds well to GST-FCP1₅₆₂₋₇₃₈ (Fig. 4B, lane 1), binds poorly to GST-FCP1₅₆₂₋₅₉₉ (Fig. 4B, lane 4), indicating that the positively-charged 48-86 region of Tat does not interact optimally with the acid-rich FCP1₅₆₂₋₅₉₉ in the context of the full-length Tat protein. A summary of the *in vitro* binding results between Tat and GST-FCP1 fragments is presented in Fig. 5.

The carboxyl-terminal domain of RAP74 binds to the acidic/hydrophobic region within the central domain of FCP1.

A small conserved acidic/hydrophobic sequence motif located within the carboxyl-terminal region of FCP1 (FCP1₈₇₉₋₉₆₁), has been shown to be important for mediating interactions between FCP1 and the carboxyl-terminal domain of RAP74 (RAP74₄₃₆₋₅₁₇) (34). The consensus sequence of this motif in FCP1₈₇₉₋₉₆₁ was derived from vertebrates and yeast sequences as D/EXD/EXD\EXXXL\IEXELXDhh (34). Comparison of this consensus sequence with the acidic/hydrophobic region from the central domain of human FCP1 (FCP1₅₆₂₋₆₃₅) revealed a similar small conserved acidic/hydrophobic motif from approximately residues 578 to 605 (Figs. 3B and 6A). As for the central domain of FCP1, the sequence conservation of this motif is very high for the vertebrates *M. musculus* and *X. laevis* when compared to human FCP1 (Fig. 3), but lower for other eukaryotic sequences. Based on the sequence conservation (Fig. 3B), we defined

a DXLXXL/IXXXLXXV/IHXXF/YY consensus among eukaryotes and a E/DEE/DXEDTDE/DDDHLLIXLEEILVRVHTDYY consensus among vertebrates (Fig. 6A).

In vitro binding studies were used to define the minimal RAP74-binding site within the central domain of FCP1 (Fig. 6B). Purified RAP74₄₃₆₋₅₁₇ was used in batch-affinity binding assays that contained various GST-FCP1 fragments coupled to GSH-Sepharose resin. FCP1 fragments that contain the small acidic/hydrophobic motif (FCP1₅₆₂₋₇₃₈, FCP1₅₆₂₋₆₃₇, and FCP1₅₆₂₋₆₁₉) interacted with RAP74₄₃₆₋₅₁₇ (Fig. 6B lanes 3-6), whereas GST-FCP1₅₉₉₋₇₃₈ that does not contain this motif did not interact with RAP74₄₃₆₋₅₁₇ (Fig. 6B lane 7).

A small peptide (FCP1₅₇₉₋₆₀₀), which contains the small acidic/hydrophobic motif sequence (Figs. 6A and 6C) was chemically synthesized and used as a competitor in our binding assay. FCP1₅₇₉₋₆₀₀ substantially inhibited the interaction between GST-FCP1₅₆₂₋₆₁₉ and RAP74₄₃₆₋₅₁₇ at concentrations (1-10 μ M) similar to those of the other proteins in this assay (Fig. 6C). Another small peptide (FCP1₅₈₃₋₆₀₇), which contains the small acidic/hydrophobic motif and other conserved residues, was chemically synthesized. However, because of its poor solubility in aqueous solution, this peptide could not be used reliably in binding studies. Therefore, FCP1₅₇₉₋₆₀₀ represents a minimal conserved RAP74-binding site within the central domain of FCP1.

The small hydrophobic motifs located within the central and carboxyl-terminal domains of FCP1 resemble the LXXLL motifs found in of the steroid hormone receptor superfamily (34). We previously examined the requirement of the first conserved leucine in the LXXLL-like motif of the carboxyl-terminus of FCP1 and found that replacements of the first conserved leucine and its preceding residue by alanines significantly weakened the binding of RAP74₄₃₆₋₅₁₇ to FCP1₈₇₉₋₉₆₁ (34). To test if the first leucine of the LXXLL-like motif within the central domain of FCP1 is required for RAP74₄₃₆₋₅₁₇ binding, we made a double alanine mutant (Y592A and L593A) within

the minimal fragment of the central domain of FCP1 that binds RAP74 (GST-FCP1₅₆₂₋₅₉₉) (Fig 6A). We observed significantly reduced binding of RAP74₄₃₆₋₅₁₇ to this double alanine mutant GST-FCP1₅₆₂₋₅₉₉ compared to the wild-type GST-FCP1₅₆₂₋₅₉₉ (Fig. 6D). These results confirm that a small conserved acidic/hydrophobic motif is required for RAP74 interactions with both the central and carboxyl-terminal domains of FCP1. A summary of the *in vitro* binding results between Tat and GST-FCP1 fragments is presented in Fig. 6E.

FCP1₅₇₉₋₆₀₀ binds to a shallow groove between α helices H2 and H3 on the surface of the carboxyl-terminal domain of RAP74.

We have previously determined the NMR structure of the free carboxyl-terminal domain of RAP74 (RAP74₄₃₆₋₅₁₇) (42) and mapped the FCP1₈₇₉₋₉₆₁-binding domain on this structure by computing chemical shift changes in RAP74₄₃₆₋₅₁₇ before and after formation of the FCP1₈₇₉₋₉₆₁/RAP74₄₃₆₋₅₁₇ complex (42). We found that FCP1₈₇₉₋₉₆₁ binds in a shallow groove formed by α helices H2 and H3 on the surface of RAP74 (42). Detailed views of the interaction were subsequently obtained from high-resolution structure determination of the FCP1₈₇₉₋₉₆₁/RAP74₄₃₆₋₅₁₇ complex by NMR spectroscopy (34) and of a similar complex by X-ray crystallography (33). Here we have also used chemical shift mapping for studying the FCP1₅₇₉₋₆₀₀-binding site on this RAP74₄₃₆₋₅₁₇ domain (Fig. 7A) (58,59). The changes in amide resonance chemical shifts (Δ_{HN}) between the free RAP74₄₃₆₋₅₁₇ and the FCP1₅₇₉₋₆₀₀/RAP74₄₃₆₋₅₁₇ complex were calculated from 2D ¹H-¹⁵N HSQC spectra (49) using the equation: $\Delta_{\text{HN}} = [(\Delta_{\text{H}})^2 + (0.17 \times \Delta_{\text{N}})^2]^{1/2}$. As observed previously with FCP1₈₇₉₋₉₆₁ (42), many of the amide resonances of RAP74 have similar chemical shifts between the free and bound forms of RAP74₄₃₆₋₅₁₇, indicating that the overall structure of RAP74 remains essentially the same in this complex. However, 33 out of the 82 amide

resonances underwent significant chemical shift changes ($\Delta \geq 0.13$ ppm), and these changes were mapped on the structure of free RAP74₄₃₆₋₅₁₇ (Fig. 7B). Interestingly, the RAP74 residues affected by FCP1₅₇₉₋₆₀₀ binding are for the most part, the same as those affected by FCP1₈₇₉₋₉₆₁ binding (Fig. 7B). It appears that, similarly to FCP1₈₇₉₋₉₆₁, (34,42), FCP1₅₇₉₋₆₀₀ interacts with the shallow groove formed by α helices 2 and 3 on the carboxyl-terminal domain of RAP74.

Tat blocks RAP74 binding to FCP1 *in vitro*.

Our *in vitro* binding studies revealed that two regions of FCP1, the acidic/hydrophobic region and the BRCT domain, interact with Tat. The small acidic/hydrophobic motif that is essential for the interaction of RAP74₄₃₆₋₅₁₇ with FCP1₅₆₂₋₇₃₈ lies within one of these two Tat-binding sites. Therefore, we were interested to see if Tat could recognize this small motif. GST-FCP1 fragments coupled to GSH-Sepharose resin were incubated with purified Tat₁₋₈₆ in the absence or presence of FCP1₅₇₉₋₆₀₀ (Fig. 8A). The FCP1₅₇₉₋₆₀₀ peptide, which inhibits RAP74₄₃₆₋₅₁₇ binding to FCP1₅₆₂₋₆₁₉ (Fig. 6C), did not inhibit Tat₁₋₈₆ binding to either GST-FCP1₅₆₂₋₇₃₈ or GST-FCP1₅₆₂₋₆₁₉ even when the concentration of FCP1₅₇₉₋₆₀₀ was 200 times higher than that of Tat₁₋₈₆ (Fig. 8A). Therefore, it appears that, unlike RAP74₄₃₆₋₅₁₇, Tat does not specifically recognize the small acidic/hydrophobic motif of FCP1₅₆₂₋₇₃₈.

To determine if this acidic/hydrophobic motif made any contribution to the Tat/FCP1 interaction, we tested the binding of Tat₁₋₈₆ with the double alanine mutants (Y592A and L593A) of both GST-FCP1₅₆₂₋₇₃₈ and GST-FCP1₅₆₂₋₅₉₉ (Fig. 8B). *In vitro* binding studies revealed that these mutants bound Tat₁₋₈₆ as well as their wild-type counterparts, indicating that the small hydrophobic motif is not essential for the Tat/FCP1 interaction (Fig. 8B).

Because Tat recognizes sequences in FCP1 located on both sides of the RAP74-binding site, we were interested in determining if purified Tat could inhibit RAP74₄₃₆₋₅₁₇ binding to the central domain of FCP1 (Fig. 8C). GST-FCP1₅₆₂₋₇₃₈ coupled to GSH-Sepharose resin was incubated with purified RAP74₄₃₆₋₅₁₇ in the absence or presence of purified HIV-1 Tat₁₋₄₈ or Tat₁₋₇₂. At equimolar concentration, Tat₁₋₇₂ completely inhibited RAP74₄₃₆₋₅₁₇ binding to GST-FCP1₅₆₂₋₇₃₈ (Fig. 8C, lane 8), indicating that Tat₁₋₇₂ has a higher affinity than RAP74₄₃₆₋₅₁₇ for GST-FCP1₅₆₂₋₇₃₈. However, under similar conditions Tat₁₋₄₈ did not inhibit the interaction of RAP74₄₃₆₋₅₁₇ with GST-FCP1₅₆₂₋₇₃₈ (Fig. 8C lane 9).

To evaluate if both regions of the central domain of FCP1 are required for the inhibition of RAP74 binding to FCP1 by Tat₁₋₇₂, purified RAP74₄₃₆₋₅₁₇, Tat₁₋₇₂, or Tat₁₋₄₈ were added to GST-FCP1₅₆₂₋₆₁₉ protein. This fragment contains only a portion of the acidic/hydrophobic region of GST-FCP1₅₆₂₋₇₃₈ (Fig. 8C). When equimolar concentrations of Tat₁₋₇₂ and RAP74₄₃₆₋₅₁₇ were allowed to bind GST-FCP1₅₆₂₋₆₁₉, both Tat₁₋₇₂ and RAP74₄₃₆₋₅₁₇ interacted with FCP1 (Fig. 8C lane 12). Similar concentrations of Tat₁₋₄₈ did not prevent RAP74₄₃₆₋₅₁₇ from binding to both GST-FCP1₅₆₂₋₆₁₉ and GST-FCP1₅₆₂₋₇₃₈ (Fig. 8C lanes 9 and 13). These results indicate that the binding of Tat₁₋₇₂ to two distinct non-overlapping sites within the central domain of FCP1 prevents RAP74₄₃₆₋₅₁₇ from interacting with the central domain of FCP1.

Tat inhibits CK2 phosphorylation at a novel site within the central domain of FCP1.

In a parallel study, sequence analysis of the acidic/hydrophobic region has revealed a novel threonine CK2 phosphorylation site (TD_{ED}) conserved in FCP1 among vertebrates (Fig. 3 and Fig. 9A) (60). We have also determined that T584 is the primary site of CK2 phosphorylation within FCP1₅₆₂₋₆₁₉ (60). Interestingly, we found that RAP74₄₃₆₋₅₁₇ has a higher

affinity for CK2-phosphorylated GST-FCP1₅₆₂₋₆₁₉ than for the non-phosphorylated protein (60). We were therefore interested to test the effect of T584 phosphorylation on Tat binding (Fig. 9B). Binding reactions were set up using CK2-phosphorylated and non-phosphorylated GST-FCP1₅₆₂₋₆₁₉ with purified Tat₁₋₈₆. Although CK2 phosphorylation of GST-FCP1₅₆₂₋₆₁₉ enhances RAP74₄₃₆₋₅₁₇ binding (60), it did not affect Tat₁₋₇₂ binding (Fig. 9B lanes 4 and 5). As a control, the T584E mutant subjected or not to the CK2 phosphorylation reaction interacted with Tat₁₋₇₂ with similar affinity (Fig. 9B, lanes 4 and 5). These results indicate that T584 does not contribute significantly to Tat binding.

Although the phosphorylation of T584 did not affect Tat binding, we wanted to know if Tat binding could affect the phosphorylation level of T584. To test this, we formed a Tat₁₋₈₆/FCP1₅₆₂₋₆₁₉ complex and assayed the ability of CK2 to phosphorylate FCP1 (Fig. 9C). Tat inhibited CK2 phosphorylation of FCP1 in a concentration dependent manner (Fig. 9C lanes 7-9). By contrast, formation of a RAP74₄₃₆₋₅₁₇/FCP1₅₆₂₋₆₁₉ complex did not inhibit CK2 phosphorylation of FCP1₅₆₂₋₆₁₉ (Fig. 9C lanes 10-12). Our results indicate that Tat not only prevents RAP74₄₃₆₋₅₁₇ binding to the central domain of FCP1, but also inhibits phosphorylation at a novel conserved CK2 site (60).

DISCUSSION

With this manuscript, we characterized the interaction of the central domain of FCP1 with the carboxyl-domain of RAP74 and the HIV-1 Tat protein, and investigated the role of CK2 phosphorylation on these interactions. It had been previously found that the HIV-1 Tat protein inhibits the *in vitro* CTD phosphatase activity of FCP1 (36). Here we showed that FCP1 is essential for the role of Tat in transcriptional transactivation and that HIV-1 Tat and FCP1 directly interact both in the yeast two-hybrid system and *in vitro*. The interaction between HIV-1

Tat and FCP1 is likely directly related to the Tat-inhibition of the CTD phosphatase activity (36) and the FCP1-dependent transcriptional transactivation by Tat.

We characterized intermolecular interactions with FCP1 using an *in vitro* protein-protein batch-affinity assay with purified proteins. We found that the carboxyl-terminal domain of RAP74 interacts with a small fragment of the acidic/hydrophobic region of FCP1, which contains a conserved LXXLL-like motif. This motif is similar to the one found in the carboxyl-terminal domain of FCP1 (FCP1₈₇₉₋₉₆₁) and involved in the FCP1₈₇₉₋₉₆₁ / RAP74₄₃₆₋₅₁₇ interaction (34). We found that, like the H1' α -helix of FCP1₈₇₉₋₉₆₁, FCP1₅₇₉₋₆₀₀ interacts with the shallow groove formed by α helices H2 and H3 of the carboxyl-terminal domain of RAP74. Furthermore, secondary structure prediction of the FCP1₅₇₉₋₆₀₀ indicates that residues 593-600 within this fragment have a propensity to form a α -helix (data not shown). FCP1₅₇₉₋₆₀₀ likely adopts an amphiphilic helix that forms hydrophobic interactions and key polar interactions with RAP74, as previously described for the FCP1₈₇₉₋₉₆₁/RAP74₄₃₆₋₅₁₇ complex (34). It is interesting to note that the same domain of RAP74 interacts with both the central and the carboxyl-terminal domains of FCP1. It has been previously shown that TFIIF forms a $\alpha_2\beta_2$ heterotetramer, where two RAP74-RAP30 dimers assemble via homomeric RAP74 interactions (61). Formation of such a tetramer would be consistent with simultaneous binding of two RAP74 subunits to the two interacting domains of FCP1.

We found that HIV-1 Tat interacts with two independent domains of FCP1, the amino-terminal portion of the acidic/hydrophobic region and the BRCT domain. Our results suggest that HIV-1 Tat protein does not directly interact with the peptide fragment that contains the conserved acidic/hydrophobic motif and which is essential for the interaction with RAP74₄₃₆₋₅₁₇. Nevertheless, we have clearly shown that binding of the HIV-1 Tat protein to the central domain

of FCP1 inhibits the binding of RAP74₄₃₆₋₅₁₇. The full-length HIV-1 Tat and the full-length central domain of FCP1 are required for this inhibition. The inhibition of the RAP74/FCP1 interaction by Tat is intriguing in light of the fact that Tat and RAP74 do not appear to share a common binding site of FCP1. It is possible that binding of Tat to FCP1 reduces the accessibility of the RAP74 binding site on FCP1, either by steric hindrance or as a result of a conformation change in the RAP74-binding site.

We have defined a conserved acidic/hydrophobic region adjacent to the BRCT domain, which is important for the Tat and RAP74 interactions with FCP1. In proteins with tandem BRCT repeats, the BRCT domains pack to form a compact structure, which involves an α -helix from the linker region at the interface of the two repeats (28,30,32). The linker between the BRCT domains often serves to stabilize the BRCT fold and may participate in ligand interactions (19). Interestingly, the conserved acidic/hydrophobic region of FCP1 has similarity to the acid-rich linker region between BRCT I and BRCT II of XRCC1, and there are two CK2 sites located in the conserved region II prior to the BRCT domain II of XRCC1 (62). We propose that the conserved acidic/hydrophobic region in FCP1 similarly interacts with the BRCT domain and helps to strengthen its interaction with the Tat protein. Given the importance of BRCT domains in protein-protein interactions, it is likely that BRCT domains form intramolecular interactions not only with adjacent BRCT domains, but with other protein domains as well. BRCT domains have been shown to be phosphopeptide recognition modules (24,25,63), therefore it is also likely that the strength of these intramolecular interactions could be controlled by phosphorylation. Further studies are needed to address the role of CK2 phosphorylation in promoting intramolecular interactions between the acidic/hydrophobic region and the BRCT domain of FCP1.

Although Tat inhibits the phosphatase activity of FCP1 and FCP1 is essential for transcriptional transactivation by Tat, the molecular details of these functional interactions are unclear. We have shown that Tat prevents the interaction of RAP74 with FCP1, suggesting that Tat inhibits the phosphatase activity by preventing RAP74 from interacting with FCP1. This may be only one mode of phosphatase inhibition by Tat, since it has been shown that Tat inhibits the CTD phosphatase activity of FCP1 in the absence of RAP74 (36). Tat may also block amino acids that are essential for the phosphatase activity of FCP1. In fact, a recent study characterizing the active site of *Schizosaccharomyces pombe* FCP1 reveals that the BRCT domain is required for catalytic activity (64). It is also possible that the interaction of Tat with FCP1 prevents negative transcription elongation factors from interacting with FCP1. It has been previously demonstrated that the positive transcription elongation effect of TFIIF is partly due to competition with the negative elongation factors DSIF and NELF (65). Another interesting possibility is that FCP1 serves as a molecular switch, and when Tat interacts with FCP1 a conformational change occurs within FCP1 and the transcription complex that enhances transcription elongation.

In an accompanying manuscript (60), we identified a novel CK2 site located within the Tat-binding domain of FCP1. The ability of Tat to inhibit phosphorylation by CK2 at this site suggests that this phosphorylation event may be an important regulator of the phosphatase activity or the positive transcription elongation effect of FCP1. It is not clear whether this particular site is phosphorylated *in vivo* or not; future studies will be necessary to understand the importance of this discovery. It was recently established that phosphorylation of *X. Laevis* FCP1 by CK2 enhances the phosphatase activity of FCP1(38). We propose that Tat inhibits FCP1 phosphatase activity by blocking CK2 phosphorylation of FCP1, which could affect the

phosphatase activity directly or indirectly by preventing RAP74 binding to the central domain of FCP1. Alternatively, it was recently found that CK2 inhibits the positive transcription elongation effect of FCP1(66). In this view, Tat could also function by preventing the CK2 inhibition and thereby promoting the positive transcription elongation effect of FCP1. Finally, phosphorylation of FCP1 by CK2 may serve to regulate a function not yet defined for FCP1 that is regulated by Tat.

In addition to inhibiting the phosphatase activity of FCP1, Tat also stimulates the enzymatic activity of two cellular kinases, which contribute to producing the hyperphosphorylated form of RNAPII: cdk7, the kinase component of TFIIF, and cdk9, the kinase component of pTEFb. Tat interacts with the cyclin T1 component of pTEFb to form an interaction that is stabilized by the presence of the HIV-1 transactivation response element (TAR) RNA (67). The TAR RNA is important for recruiting the kinase to the early elongation complex. There are important differences between the cyclin T1/Tat and the FCP1/Tat interaction. First, the interaction of Tat with cyclin T1 only involves the activation domain of Tat, whereas the interaction of Tat with FCP1 involves both the activation domain and the arginine-rich carboxyl terminus (67). Initial binding studies with the TAR RNA also indicate that, unlike the cyclinT1/Tat complex (67), the FCP1/Tat complex does not bind TAR and is not stabilized in the presence of TAR (unpublished results). Given its role in phosphorylating the initiating RNAPII, there may be an advantage for pTEFb to be anchored to the TAR RNA element located at the 5'-end of the mRNA; however, such an advantage may be lost for FCP1, which dephosphorylates non-initiating forms of the polymerase.

In this study, we have shown that FCP1 is essential for the transactivation activity of the Tat protein and that FCP1 and Tat interact both *in vitro* and *in vivo*. Using purified protein

fragments and an *in vitro* batch affinity-binding assay, we have defined the interaction of FCP1 with HIV-1 Tat and the carboxyl-terminal domain of RAP74. Tat interacts with both the acidic/hydrophobic domain and the BRCT domain, whereas RAP74 interacts only with a portion of the acidic/hydrophobic domain. Our studies highlight the importance of regions adjacent to the BRCT domain in promoting protein-protein interactions. Although Tat and RAP74 do not appear to share a common binding site on FCP1, Tat inhibits RAP74 from binding FCP1. Tat also inhibits CK2 phosphorylation at T584, and this phosphorylation has been previously shown to enhance RAP74 binding. Our results suggest that HIV-1 Tat protein regulates FCP1 phosphatase activity by both preventing RAP74 binding and inhibiting CK2 phosphorylation, providing a better understanding of the role of HIV-1 Tat in transcriptional regulation.

ACKNOWLEDGEMENTS.

We thank the NIH AIDS Research and Reference Reagent Program for HIV-1 Tat expression vectors and antibodies, Danny Reinberg for providing the full-length hFCP1 clone, Barbara Potempa for assistance in the preparation of protein samples, and Claiborne Glover for critical reading of the manuscript. This work was supported by the National Institutes of Health Grant RO1 GM60298-01 (to J.G.O and P.L.), by a grant from the Canadian Institutes of Health Research (to J.G.), and by (to R.G.R.)

REFERENCES

1. Payne, J. M., Laybourn, P. J., and Dahmus, M. E. (1989) *J. Biol. Chem.* **264**, 19621-19629
2. Weeks, J. R., Hardin, S. E., Shen, J., Lee, J. M., and Greenleaf, A. L. (1993) *Genes & Dev.* **7**, 2329-2344
3. Corden, J. L. (1990) *Trends Biochem. Sci.* **15**, 383-387
4. Marshall, N. F., Peng, J., Xie, Z., and Price, D. H. (1996) *J. Biol. Chem.* **271**, 27176-27183
5. Lu, H., Zawel, L., Fisher, L., Egly, J. M., and Reinberg, D. (1992) *Nature* **358**, 641-645
6. Kobor, M. S., and Greenblatt, J. (2002) *Biochim. Biophys. Acta.* **1577**, 261-275
7. Hausmann, S., and Shuman, S. (2002) *J. Biol. Chem.* **277**, 21213-21220
8. Lin, P. S., Dubois, M.-F., and Dahmus, M. E. (2002) *J. Biol. Chem.* **277**, 45949-45956
9. Cho, H., Kim, T. K., Mancebo, H., Lane, W. S., Flores, O., and Reinberg, D. (1999) *Genes & Dev.* **13**, 1540-1552
10. Chambers, R. S., and Dahmus, M. E. (1994) *J. Biol. Chem.* **269**, 26243-26248
11. Archambault, J., Pan, G., Dahmus, G. K., Cartier, M., Marshall, N., Zhang, S., Dahmus, M. E., and Greenblatt, J. (1998) *J. Biol. Chem.* **273**, 27593-27601
12. Mandal, S. S., Cho, H., Kim, S., Cabane, K., and Reinberg, D. (2002) *Mol. Cell. Biol.* **22**, 7543-7552
13. Cho, E.-J., Kobor, M. S., Kim, M., Greenblatt, J., and Buratowski, S. (2001) *Genes & Dev.* **15**, 3319-3329
14. Kobor, M. S., Archambault, J., Lester, W., Holstege, F. C. P., Gileadi, O., Jansma, D. B., Jennings, E. G., Kouyoumdjian, F., Davidson, A. R., Young, R. A., and Greenblatt, J. (1999) *Mol. Cell* **4**, 55-62

15. Collet, J. F., Stroobant, V., Pirard, M., Delpierre, G., and Schaftingen, E. V. (1998) *J. Biol. Chem.* **273**, 14107-14112
16. Kobor, M. S., Simon, L. D., Omichinski, J., Zhong, G., Archambault, J., and Greenblatt, J. (2000) *Mol. Cell. Biol.* **20**, 7438-7449
17. Gaboriaud, C., Bissery, V., Benchetrit, T., and Mornon, J. P. (1987) *FEBS Lett.* **224**, 149-155
18. Bork, P., Hofmann, K., Bucher, P., Neuwald, A. F., Altschul, S., and Koonin, E. V. (1997) *FASEB J* **11**, 68-76
19. Callebaut, I., and Mornon, J.-P. (1997) *FEBS Letters* **400**, 25-30
20. Taylor, R. M., Thistlethwaite, A., and Caldecott, K. W. (2002) *Mol. Cell. Biol.* **22**, 2556-2563
21. Morales, J. C., Xia, Z., Lu, T., Aldrich, M. B., Wang, B., Rosales, C., Kellems, R. E., Hittelman, W. N., Elledge, S. J., and Carpenter, P. B. (2003) *J. Biol. Chem.* **278**, 14971-14977
22. Miyake, T., Hu, Y.-F., Yu, D. S., and Li, R. (2000) *J. Biol. Chem.* **275**, 40169-40173
23. Ye, Q., Hu, Y.-F., Zhong, H., Nye, A. C., Belmont, A. S., and Li, R. (2001) *J. Cell Biol.* **155**, 911-921
24. Yu, X., Chini, C. C. S., He, M., Mer, G., and Chen, J. (2003) *Science* 639-642
25. Manke, I. A., Lowery, D. M., Nguyen, A., and Yaffe, M. B. (2003) *Science* **302**, 636-639
26. Zhang, Z., Morera, S., Bates, P. A., Whitehead, P. C., Coffey, A. I., Hainbucher, K., Nash, R. A., Sternberg, M. J. E., Lindahl, T., and Freemont, P. S. (1998) *EMBO J.* **17**, 6404-6411
27. Lee, J. Y., Chang, C., Song, H. K., Moon, J., Yang, J. K., Kim, H.-K., Kwon, S.-T., and Suh, S. W. (2000) *EMBO J.* **19**, 1119-1129

28. Joo, W. S., Jeffrey, P. D., Cantor, S. B., Finnin, M. S., Livingston, D. M., and Pavletich, N. P. (2002) *Genes & Dev.* **16**, 583-593
29. Krishnan, V. V., Thornton, K. H., Thelen, M. P., and Cosman, M. (2001) *Biochemistry* **40**, 13158-13166
30. Williams, R. S., Green, R., and Glover, J. N. M. (2001) *Nature Struct. Biol.* **8**, 838-842
31. Williams, R. S., and Glover, J. N. M. (2003) *J. Biol. Chem.* **278**, 2630-2635
32. Derbyshire, D. J., Basu, B. P., Serpell, L. C., Joo, W. S., Date, T., Iwabuchi, K., and Doherty, A. J. (2002) *EMBO J.* **21**, 3863-3872
33. Kamada, K., Roeder, R. G., and Burley, S. K. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100**, 2296-2299
34. Nguyen, B. D., Abbott, K. L., Potempa, K., Kobor, M. S., Archambault, J., Greenblatt, J., Legault, P., and Omichinski, J. G. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100**, 5688-5693
35. Chambers, R. S., Wang, B. Q., Burton, Z. F., and Dahmus, M. E. (1995) *J. Biol. Chem.* **270**, 14962-14969
36. Marshall, N. F., Dahmus, G. K., and Dahmus, M. E. (1998) *J. Biol. Chem.* **273**, 31726-31730
37. Zhou, M., Halanski, M. A., Radonovich, M. F., Kashanchi, F., Peng, J., Price, D. H., and Brady, J. N. (2000) *Mol. Cell. Biol.* **20**, 5077-5086
38. Palancade, B., Dubois, M.-F., and Bensaude, O. (2002) *J. Biol. Chem.* **277**, 36061-36067
39. Campioni, D., Corallini, A., Zauli, G., Possati, L., Altavilla, G., and Barbanti-Brodano, G. (1995) *Retroviruses* **11**, 1039-1048
40. Hauber, J., Perkins, A., Heiner, E., and Cullen, B. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 6364-6368

41. Rhim, H., Echeteu, C. O., and Rice, A. P. (1994) *J. Acquired Immune Defic. Syndr.* **7**, 1116-1121
42. Nguyen, B. D., Chen, H.-T., Kobor, M. S., Greenblatt, J., Legault, P., and Omichinski, J. G. (2003) *Biochemistry* **42**, 1460-1469
43. Zhou, Q., and Sharp, P. A. (1995) *EMBO J.* **14**, 321-328
44. Dignam, J. D., Lebovitz, R. M., and Roeder, R. G. (1983) *Nucleic Acids Res.* **11**, 1475-1489
45. Ge, H., and Roeder, R. G. (1994) *J. Biol. Chem.* **269**, 17136-17140
46. Lapsia, M. F., Rice, A. P., and Mathews, M. B. (1989) *Cell* **59**, 283-292
47. Sherman, F., Fink, G. R., and Hicks, J. B. (1986) *Laboratory course manual for methods in yeast genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y.
48. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) *Protein Science* **4**, 2411-2423
49. Kay, L. E., Keifer, P., and Saarinen, T. (1992) *J. Am. Chem. Soc.* **114**, 10663-10665
50. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR* **6**, 277-293
51. Southgate, C. D., and Green, M. R. (1991) *Genes & Dev.* **5**, 2496-2507
52. Blau, J., Xiao, H., MacCracken, S., O'Hare, P., Greenblatt, J., and Bently, D. (1996) *Mol. Cell Biol.* **16**, 2044-2055
53. Sonnhammer, E. L. L., Eddy, S. R., and Durbin, R. (1997) *Proteins* **28**, 405-420
54. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res.* **25**, 3389-3402
55. Kato, H., Sumimoto, H., Pognonec, P., Chen, C.-H., Rosen, C. A., and Roeder, R. G. (1992) *Genes & Dev.* **6**, 655-666

56. Bayer, P., Kraft, M., Ejchart, A., Westendorp, M., Frank, R., and Rosch, P. (1995) *J. Mol. Biol.* **247**, 529-535
57. Baldi, P., Brunak, S., Frasconi, P., Soda, G., and Pollastri, G. (1999) *Bioinformatics* **15**, 937-946
58. Farmer, B. T., Constantine, K. L., Goldfarb, V., Friedrichs, M. S., Wittekind, M., Yanchunas, J., Robertson, J. G., and Mueller, L. (1996) *Nature Struct. Biol.* **3**, 995-997
59. Shuker, S. B., Hajduk, P. J., Meadows, R. P., and Fesik, S. W. (1996) *Science* **274**, 1531-1534
60. Abbott, K. L., Renfrow, M. B., Chalmers, M. J., Nguyen, B. D., Appella, E., Marshall, A. G., Legault, P., and Omichinski, J. G. (2004) *J. Biol. Chem. In Preparation*
61. Robert, F., Douziech, M., Forget, D., Egly, J.-M., Greenblatt, J., Burton, Z. F., and Coulombe, B. (1998) *Molecular Cell* **2**, 341-351
62. Marintchev, A., Robertson, A., Dimitriadis, E. K., Prasad, R., Wilson, S. H., and Mullen, G. P. (2000) *Nucleic Acids Res.* **28**, 2049-2059
63. Caldecott, K. W. (2003) *Science* **302**, 579-580
64. Hausmann, S., and Shuman, S. (2003) *J. Biol. Chem.* **278**, 13627-13632
65. Renner, D. B., Yamaguchi, Y., Wada, T., and Price, D. H. (2001) *J. Biol. Chem.* **276**, 42601-42609
66. Friedl, E. M., Lane, W. S., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100**, 2328-2333
67. Garber, M. E., Wei, P., KewalRamani, V. N., Mayall, T. P., Herrmann, C. H., Rice, A. P., Littman, D. R., and Jones, K. A. (1998) *Genes & Dev.* **12**, 3512-3527

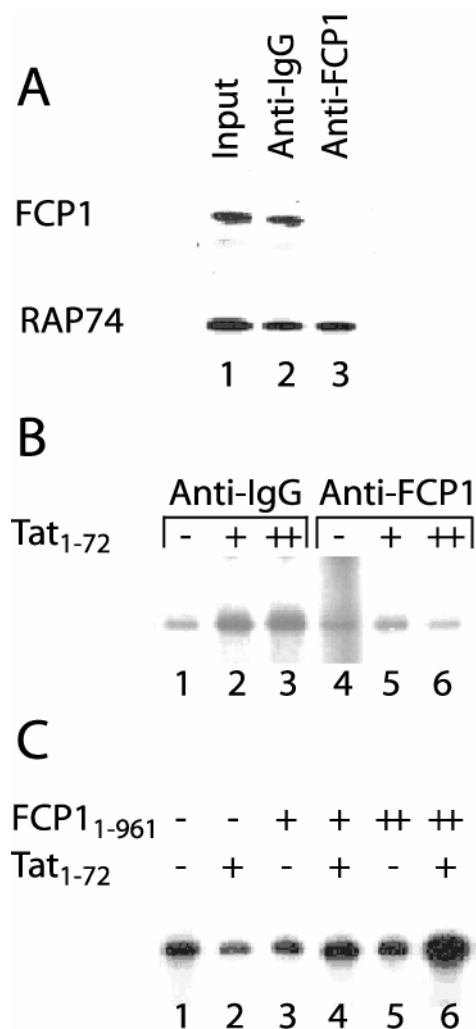


Figure 2.1. FCP1 mediates transcriptional activation by HIV-1 Tat *in vitro*. (A) Immuno-depletion of endogenous FCP1. HeLa cell nuclear extracts (30 μ g) depleted with control IgG (lane 1) or affinity-purified FCP1 polyclonal (lane 2) antibodies were analyzed by western blot using FCP1 and RAP74 antibodies. (B) Effect of FCP1 depletion on transcription activated by Tat. Nuclear extracts depleted with control IgG (lanes 1-3) or FCP1 affinity-purified antibodies (lanes 4-6) were tested for their ability to support transcription from the pHIV+TARG400 template in the absence (-) or presence of 200 ng (+) or 400 ng (++) Tat₁₋₇₂. (C) Restoration of Tat activation in nuclear extracts treated with FCP1 affinity-purified antibody after the addition of full-length purified FCP1. *In vitro* transcription reactions were performed with nuclear extracts treated with FCP1 affinity-purified antibody in the absence (-) or presence (+) of 200 ng Tat₁₋₇₂ and either 0.25 μ l (+) or 0.5 μ l (++) of full-length FCP1 (Mono Q fraction 34) (11).

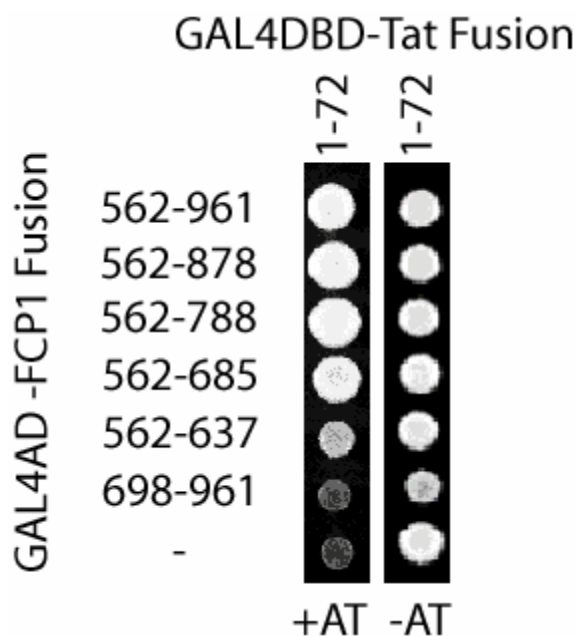


Figure 2.2 FCP1 and HIV-1 Tat interact in the yeast two-hybrid system. *S. cerevisiae* cells were co-transformed with plasmids expressing FCP1 fragments fused to the GAL4 activation domain (GAL4AD) and a plasmid expressing Tat₁₋₇₂ fused to the GAL4 DNA-binding domain (GAL4DBD). For each of the resulting strains, a drop of a cell suspension was applied onto SD medium containing (+AT) or lacking (-AT) 3-aminotriazole and allowed to grow at 30 °C for 4 days.

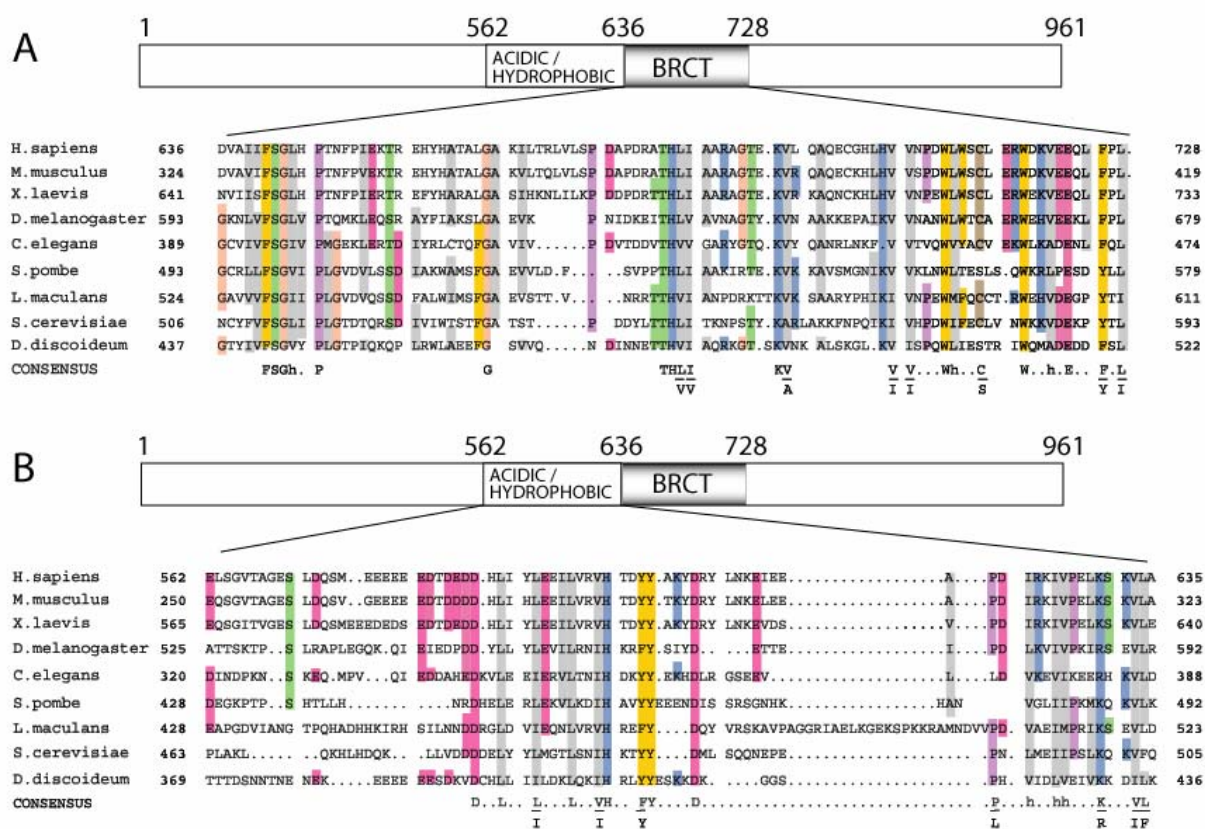
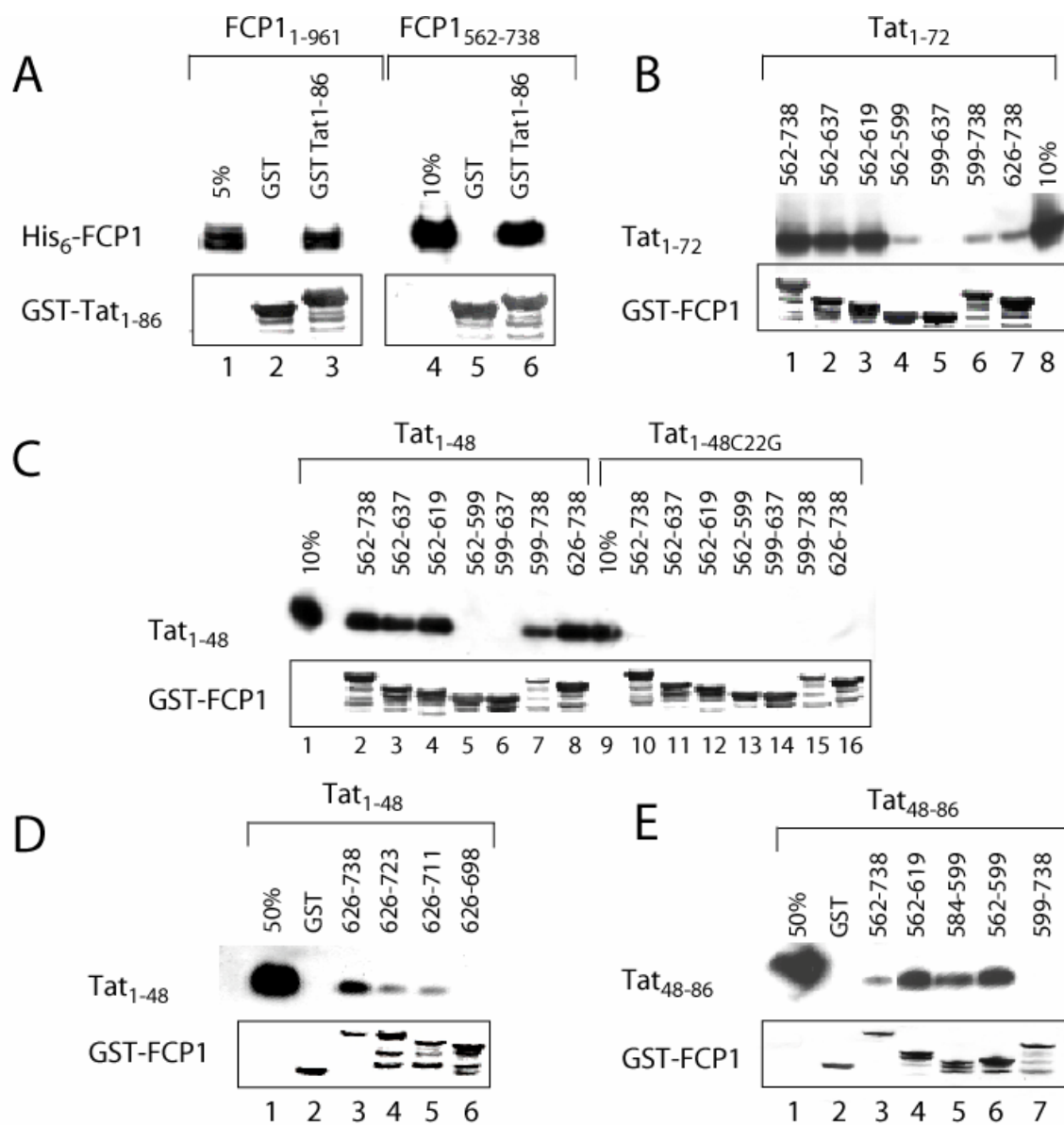


Figure 2.3. The FCP1 central domain is conserved in eukaryotes. The sequences of (A) the BRCT and (B) the acidic/hydrophobic regions of FCP1 from *Homo sapiens* (NP_004706.2), *Mus musculus* (AAH53435), *Xenopus laevis* (Q98SN2), *Drosophila melanogaster* (Q9W147), *Caenorhabditis elegans* (NP_492423), *Schizosaccharomyces pombe* (Q9P376), *Leptosphaeria maculans* (AAM81360), *Saccharomyces cerevisiae* (NP_014004), and *Dictyostelium discoideum* (AAM33170) were aligned using Pfam (53) and BLAST (54). Residues conserved in at least four species are highlighted using the following color code; pink, acidic residues (D,E); sky blue, basic residues (H,K,R); grey, hydrophobic residues (A,I,L,M,V); gold, aromatic residues (F,Y,W); green, threonine and serine residues (T, S); peach, glycine residues (G); purple, proline residues (P); and beige, cysteine residues (C). A consensus is provided for residues that are conserved in all species.

Figure 2.4. Identification and characterization of the Tat-binding regions of FCP1 using *in vitro* binding studies. (A) Recombinant full-length His₆-FCP1 and thioredoxin-His₆-FCP1₅₆₂₋₇₃₈ interact with GST-Tat₁₋₈₆ with similar affinity. GST-Tat₁₋₈₆ was coupled to GSH-Sepharose and incubated with recombinant His₆-FCP1₁₋₉₆₁ or thioredoxin-His₆-FCP1₅₆₂₋₇₃₈ (5 % His₆-FCP1 and 10% FCP1₅₆₂₋₇₃₈ inputs are shown). Bound FCP1 was detected using a polyhistidine monoclonal antibody. (B) Interaction of purified HIV-1 Tat₁₋₇₂ with various GST-FCP1 fragments reveals two Tat-binding regions in FCP1. GST-FCP1 fusion proteins were coupled to GSH-Sepharose resin and incubated with purified HIV-1 Tat₁₋₇₂ (10 % of the Tat₁₋₇₂ input is shown). Bound Tat protein was detected using the Tat monoclonal antibody (15.1). (C) The HIV-1 Tat activation domain (Tat₁₋₄₈) interacts specifically with two regions of FCP1. Purified HIV-1 Tat₁₋₄₈ or purified HIV-1 Tat_{1-48C22G} were incubated with various GST-FCP1 fragments (10% inputs for each Tat fragment are shown). Bound wild-type and C22G mutant Tat₁₋₄₈ proteins were detected using the Tat monoclonal primary antibody (15.1) which recognizes amino acids 1-16 of HIV-1 Tat. (D) The carboxyl-terminal region of the FCP1 BRCT domain is essential for interaction with Tat₁₋₄₈. Purified HIV-1 Tat₁₋₄₈ protein was added to various carboxyl-terminal BRCT deletion fragments of GST-FCP1₆₂₆₋₇₃₈ (50 % Tat₁₋₄₈ input is shown). Bound Tat₁₋₄₈ protein was detected using the Tat monoclonal (15.1) primary antibody. (E) The arginine-rich carboxyl terminus of HIV-1 Tat₄₈₋₈₆ interacts with the acidic/hydrophobic region of the central domain of FCP1. Purified Tat₄₈₋₈₆ was incubated with various GST-FCP1 fragments (50% Tat₄₈₋₈₆ input is shown). Tat₄₈₋₈₆ protein was detected using the Tat polyclonal primary antibody (705). For (A)-(E) equivalent inputs of GST-Tat or GST-FCP1 fusion proteins were determined by performing a Coomassie blue stain of the membrane (boxed).



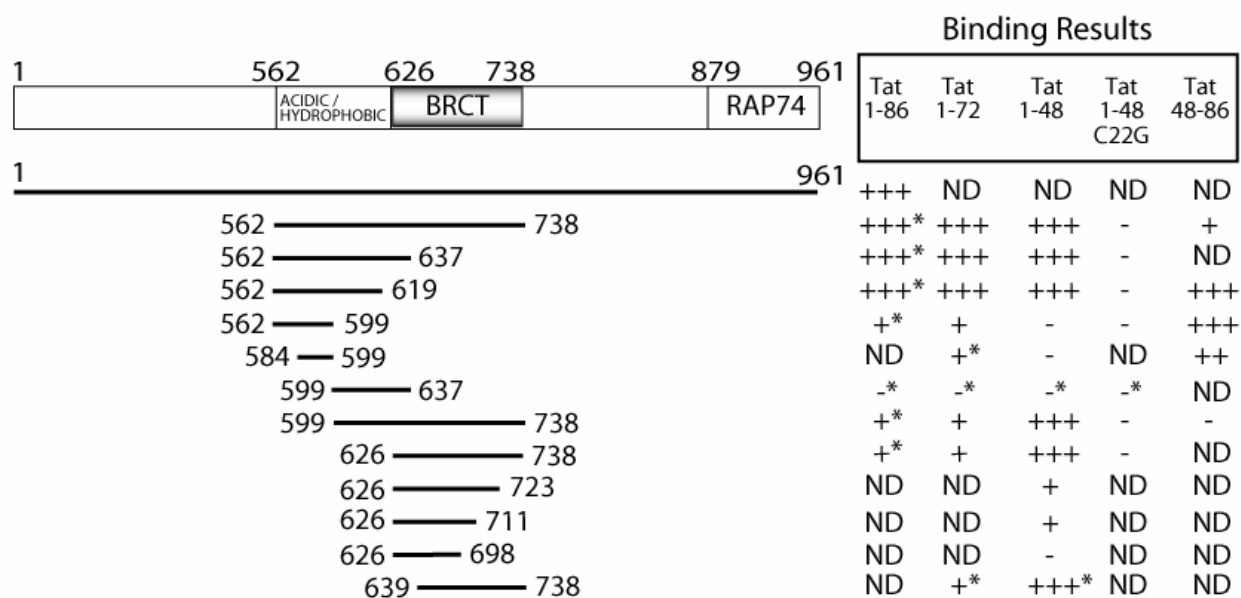
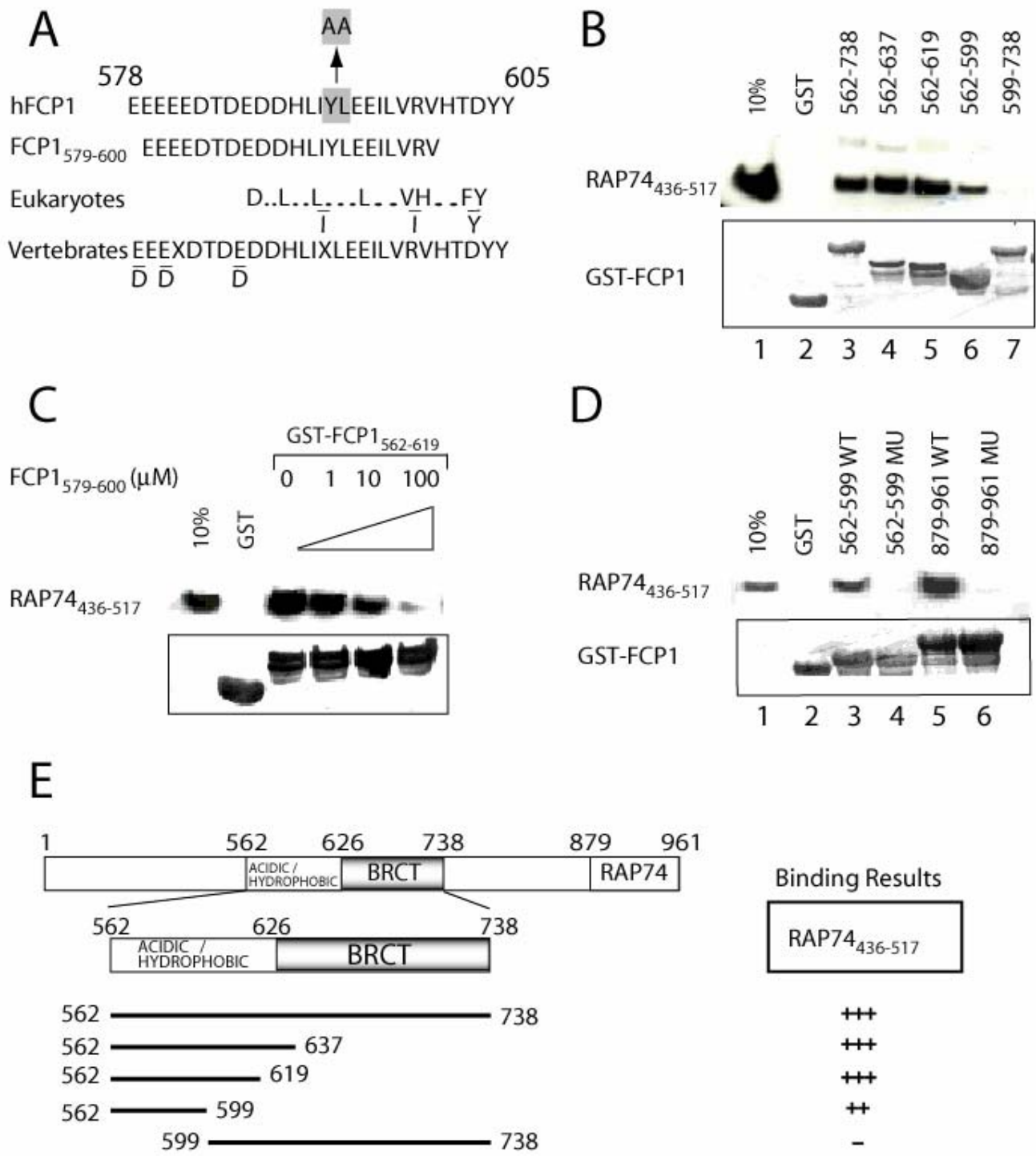


Figure 2.5. Summary of the interactions of HIV-1 Tat proteins with various GST-FCP1 fragments. Strong (+++), medium (++), weak (+), and no (-) binding was detected between Tat proteins and GST-FCP1 fragments. ND means not determined, and * indicates that the results have been obtained but are not explicitly shown.

Figure 2.6. RAP74 interacts with a small conserved acidic/hydrophobic motif in the central domain of FCP1. (A) Region of human FCP1 (amino acids 578-605), which contains a conserved acidic/hydrophobic motif, and consensus sequences from eukaryotes and vertebrates (See Fig. 3B). The location of the Y592A/L593A mutation and the sequence FCP1₅₇₉₋₆₀₀ are also shown. **(B)** Interaction of RAP74₄₃₆₋₅₁₇ with various GST-FCP1 fragments. GST-FCP1 fragments were coupled to GSH-Sepharose resin and incubated with purified RAP74₄₃₆₋₅₁₇ (10 % RAP74₄₃₆₋₅₁₇ input is shown). Bound RAP74₄₃₆₋₅₁₇ was detected using the RAP74 polyclonal antibody (C-18). **(C)** A small peptide derived from FCP1 (FCP1₅₇₉₋₆₀₀) interacts with RAP74₄₃₆₋₅₁₇. GST-FCP1₅₆₂₋₆₁₉ (10 μ M) and purified RAP74₄₃₆₋₅₁₇ (10 μ M) proteins were incubated in the presence of increasing amounts of FCP1₅₇₉₋₆₀₀ (0, 1, 10, and 100 μ M). The bound RAP74₄₃₆₋₅₁₇ was detected using the RAP74 polyclonal antibody (C-18). **(D)** Alanine substitutions within the conserved hydrophobic motif inhibit RAP74 binding to the central domain of FCP1. We examined the interaction of RAP74₄₃₆₋₅₁₇ with the GST-FCP1₅₆₂₋₅₉₉ Y592A/L593A mutant (562-599 MU) and compared with the wild-type FCP1₅₆₂₋₅₉₉ (562-599 WT). As a control we repeated the binding experiment of RAP74₄₃₆₋₅₁₇ with GST-FCP1₈₇₉₋₉₆₁ E956A/L957A mutant (879-961 MU) and the wild-type GST-FCP1₈₇₉₋₉₆₁ (879-961 WT) (34). For **(B)**, **(C)**, and **(D)** equivalent input of GST-FCP1 fusion protein was verified by performing a Coomassie blue stain of the membrane (boxed). **(E)** Summary of the RAP74₄₃₆₋₅₁₇ interactions with various GST-FCP1 fragments. Strong (+++), medium (++), and no (-) binding was detected between RAP74₄₃₆₋₅₁₇ and GST-FCP1 fragments.



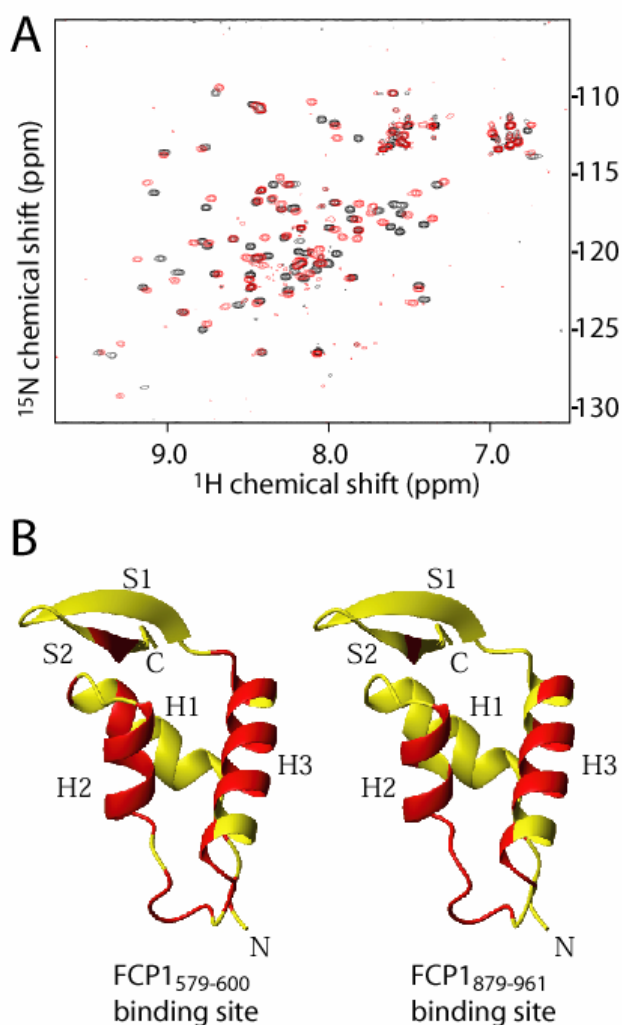


Figure 2.7. FCP1₅₇₉₋₆₀₀ binds to a shallow groove between helices H2 and H3 on the surface of RAP74₄₃₆₋₅₁₇. (A) Overlay of 2D ¹H-¹⁵N HSQC spectra of the free ¹⁵N-labeled RAP74₄₃₆₋₅₁₇ (black) and of the ¹⁵N-labeled RAP74₄₃₆₋₅₁₇/FCP1₅₇₉₋₆₀₀ complex (red). (B) Chemical shift mapping of the FCP1₅₇₉₋₆₀₀-binding site (left) and FCP1₈₇₉₋₉₆₁-binding site (right) (42) on the ribbon diagram of RAP74₄₃₆₋₅₁₇. Residues for which we observed a significant change in amide chemical shifts ($\Delta_{\text{HN}} \geq 0.13$ ppm and $\Delta_{\text{HN}} = [(\Delta_{\text{H}})^2 + (0.17 \times \Delta_{\text{N}})^2]^{1/2}$) upon addition of FCP1₅₇₉₋₆₀₀ are indicated in red.

Figure 2.8. HIV-1 Tat inhibits RAP74₄₃₆₋₅₁₇ binding to the central domain of FCP1. (A) A small peptide derived from FCP1 (FCP1₅₇₉₋₆₀₀), which interacts with RAP74₄₃₆₋₅₁₇, does not inhibit the binding of Tat₁₋₈₆ to the central domain of FCP1. GST-FCP1₅₆₂₋₇₃₈ (0.5 μ M) or GST-FCP1₅₆₂₋₆₁₉ (2 μ M) were incubated with purified Tat₁₋₈₆ (0.5 μ M) protein in the presence of increasing amounts of FCP1₅₇₉₋₆₀₀ (0, 1, 10, and 100 μ M). **(B)** Alanine substitutions within the conserved hydrophobic motif do not affect Tat₁₋₈₆ binding to the central domain of FCP1. We examined the interaction of Tat₁₋₈₆ with the GST-FCP1₅₆₂₋₇₃₈ Y592A/L593A mutant (562-738 MU) and with the GST-FCP1₅₆₂₋₅₉₉ Y592A/L593A mutant (562-599 MU) as well as with the corresponding wild-type proteins (562-738 WT and 562-599 WT). Lane 1 shows 10 % Tat₁₋₈₆ input and lane 2 the control reaction with the GST protein. In lanes 2-4, each binding reaction contained 0.5 μ M of Tat₁₋₈₆ and 0.5 μ M of GST or GST-fusion protein. In lanes 5 and 6, each binding reaction contained 0.5 μ M of Tat₁₋₈₆ and 2 μ M of GST-fusion protein. **(C)** Tat₁₋₇₂ inhibits RAP74₄₃₆₋₅₁₇ binding to FCP1, and the full BRCT domain is required for this inhibition. Purified RAP74₄₃₆₋₅₁₇, Tat₁₋₇₂, or Tat₁₋₄₈ proteins were added to binding reactions with 10 μ M GST-FCP1₅₆₂₋₇₃₈ or GST-FCP1₅₆₂₋₆₁₉ as indicated (concentrations in μ M). Lanes 1, 2, and 3 show 10 % input of purified proteins, lanes 4 and 5 represent GST control reactions. For **(A)** and **(B)**, bound Tat protein was detected using the Tat monoclonal antibody (15.1). For **(C)**, bound RAP74₄₃₆₋₅₁₇ protein was detected using the RAP74 polyclonal antibody (C-18), and subsequently bound Tat₁₋₄₈ and Tat₁₋₇₂ were detected using the Tat monoclonal antibody (15.1). Equivalent GST-FCP1 fragment input was determined by performing a Coomassie blue stain of the membrane (shown in **(A)** and **(B)** only).

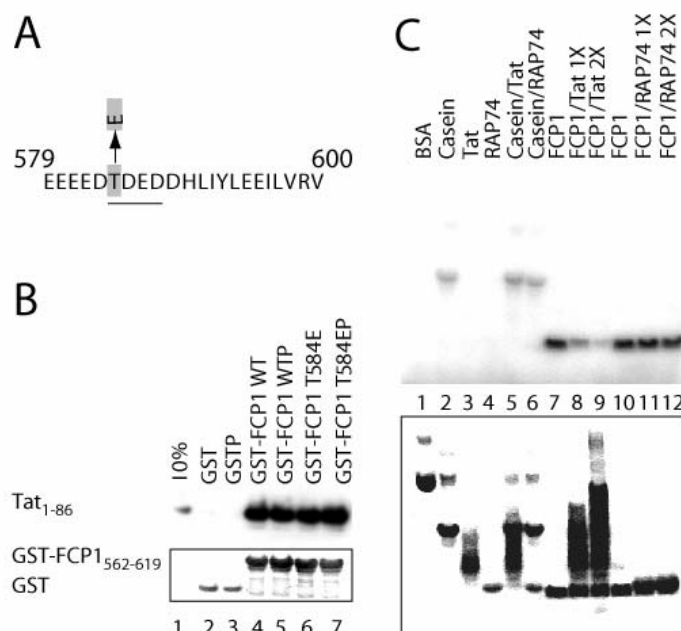


Figure 2.9. Role of a novel CK2 site within the acidic/hydrophobic region of the central domain of FCP1. (A) Sequence of FCP1₅₇₉₋₆₀₀ showing the location of the conserved CK2 site (TDED; underlined), the phosphorylated amino acid (T584; shaded), and the T584E mutation studied here. (B) Phosphorylation of T584 does not affect the binding of Tat₁₋₈₆ to the central domain of FCP1. Lane 1 shows 10% Tat₁₋₈₆ input, and lanes 2 and 3 represent binding of Tat₁₋₈₆ to the GST control resin after mock phosphorylation and phosphorylation reactions, respectively. Each binding reaction contained 0.5 μ M of GST-FCP1₅₆₂₋₆₁₉ wild-type (WT) or GST-FCP1₅₆₂₋₆₁₉ mutant (T584E) and 0.5 μ M Tat₁₋₈₆. Mock phosphorylated (GST, GST-FCP1 WT, or GST-FCP1 T584E) and CK2 phosphorylated (GSTP, GST-FCP1 WTP, or GST-FCP1 T584EP) proteins are indicated. Bound Tat₁₋₈₆ was detected using the Tat monoclonal antibody (15.1). (C) The binding of Tat inhibits phosphorylation at T584 by CK2. Phosphorylation reactions were performed in the presence of [γ ³²P]-ATP. Control phosphorylation reactions with BSA (lane 1), casein (lane 2), Tat₁₋₈₆ (lane 3), RAP74 (lane 4), casein and Tat₁₋₈₆ (lane 5), and casein and RAP74₄₃₆₋₅₁₇ (lane 6) are shown. CK2 phosphorylation of FCP1₅₆₂₋₆₁₉ (25 μ g or 161 μ M) in the absence (lane 7) or the presence of 1X (25 μ g or 100 μ M; lane 8) or 2X (50 μ g or 200 μ M; lane 9) Tat₁₋₈₆. CK2 phosphorylation of FCP1₅₆₂₋₆₁₉ (25 μ g or 161 μ M) in the absence (lane 10) or the presence of 1X (25 μ g or 114 μ M; lane 11) or 2X (50 μ g or 228 μ M; lane 12) RAP74₄₃₆₋₅₁₇. The ³²P-labeled proteins were detected by PhosphorImager analysis. A Coomassie blue stain of the membrane (B) or the gel (C) was performed to verify equivalent protein input (boxed).

CHAPTER 3

INVESTIGATIONS INTO THE ENHANCED BINDING OF THE RNAPII CTD PHOSPHATASE FCP1 TO RAP74 FOLLOWING CK2 PHOSPHORYLATION²

²Karen L. Abbott, Mathew B. Renfrow, Michael J. Chalmers, Bao D. Nguyen, Alan G. Marshall, Pascale Legault, and James G. Omichinski. To be submitted to J. Biol. Chem.

ABSTRACT

FCP1 (TFIIF associated CTD phosphatase) is the first identified CTD-specific phosphatase required to recycle RNA polymerase II (RNAPII). FCP1 activity has been shown to be regulated by the general transcription factors TFIIF (RAP74) and TFIIB, protein kinase CK2 (CK2), and the HIV-1 transcriptional activator Tat. Phosphorylation of FCP1 by CK2 stimulates FCP1 phosphatase activity and enhances binding of RAP74 to FCP1. We have examined consensus CK2 phosphorylation sites located immediately adjacent to both RAP74 binding sites of FCP1. We demonstrate that both of these consensus CK2 sites can be phosphorylated *in vitro*, and that phosphorylation at either CK2 site results in enhanced binding of RAP74 to FCP1. The CK2 site adjacent to the RAP74-binding site in the central domain is phosphorylated at a single threonine site (T584), whereas the CK2 site adjacent to the RAP74-binding site in the carboxyl-terminal domain is phosphorylated at three successive serine residues (S942-S944). Using mass spectral analysis, we demonstrate that the phosphorylation of S942-S944 occurs in a highly ordered fashion with the first phosphorylation occurring at S942 followed by subsequent phosphorylation at S944, and finally S943. Using Nuclear Magnetic Resonance (NMR) spectroscopy, we identify and map chemical shift changes onto the solution structure of the carboxyl-terminal domain of RAP74 (RAP74₄₃₆₋₅₁₇) that occur when RAP74₄₃₆₋₅₁₇ is complexed with phosphorylated FCP1 peptides. These results provide new functional and structural information on the role of phosphorylation in the recognition of acidic-rich activation domains involved in transcriptional regulation, and bring insights into how CK2 and TFIIF regulate FCP1 function.

INTRODUCTION

RNA polymerase II (RNAP II) is a multi-subunit enzyme responsible for transcription of eukaryotic messenger RNA. The carboxyl-terminal domain (CTD) of the largest subunit of RNAP II contains a conserved repeat of the heptapeptide sequence (Y-S-P-T-S-P-S) (1,2). Phosphorylation and dephosphorylation at serine-2 and serine-5 within this repeat play an important role in regulating progression through the RNAP II transcription cycle (3,4). The hypophosphorylated form of RNAP II (RNAP IIA) is associated with the pre-initiation complex (PIC), whereas the hyperphosphorylated form of RNAP II (RNAP IIO) is associated with the elongating RNAP II complexes (5,6). In addition to its role in passing from the initiation phase to the elongation phase of transcription, CTD phosphorylation has also been implicated in regulating post-transcriptional mRNA processing (7). Numerous kinases and phosphatases have been shown to specifically associate with the CTD (8-18), and it is becoming increasingly clear that a complex regulatory network is required to maintain the precise phosphorylation state of the CTD.

FCP1 (TFIIF-associated CTD phosphatase 1) is a ubiquitous serine phosphatase with orthologs present in all eukaryotic species. FCP1 was the first RNAP II CTD-specific phosphatase identified (14-18), and it has been demonstrated to be a component of the mammalian RNAP II PIC (16,19). It has been proposed that one key function of FCP1 is to dephosphorylate the CTD of RNAP II, thus enabling the polymerase to undergo another transcription cycle (18,19). In addition, yeast capping enzymes Ceg1 and Abd1 are recruited to RNAP II via the phosphorylated CTD, and FCP1 phosphatase activity is essential for release of these capping enzymes (20). Besides its role in the recycling of RNAP II through its CTD phosphatase activity, FCP1 appears to serve other functions in regulating transcription. FCP1

functions as a positive regulator of transcription elongation, and this activity has been shown to be independent of its CTD phosphatase activity (18,21). Studies in yeast have established a genetic link between FCP1 and several transcription elongation factors (21-23). Furthermore, chromatin immunoprecipitation experiments demonstrate that FCP1 remains associated with the polymerase during elongation (24). Recently, a role for FCP1 in regulating mRNA processing has also been suggested although it is not clear if this is associated with its phosphatase activity (25,26). FCP1 has been shown to co-immunoprecipitate with the methylosome complex (27), and FCP1 also associates with spliceosomal proteins (27).

FCP1 function is controlled by a number of nuclear factors, and these factors have been shown to interact with specific domains in FCP1. Human FCP1 (hFCP1) can be divided into three main domains based on sequence homology with yeast FCP1 (yFCP1) (15,16). The N-terminal FCP1 homology domain (hFCP1 122-330) contains a conserved hydrophobic motif $\Psi\Psi\Psi\text{D}\text{X}\text{D}\text{X}(\text{T}/\text{V})\Psi\Psi$ (where Ψ denotes hydrophobic residues) found in a family of small molecule phosphotransferases and phosphohydrolases (18,19,28). The central domain (hFCP1 544-728) contains a single BRCT (BRCA1 C-terminus) repeat, and yeast two hybrid studies have revealed that this region of FCP1 is important for mediating interactions with the HIV-1 transcriptional activator Tat (Abbott et al. unpublished) and the large subunit of TFIIF (RAP74) (16). Binding of Tat and RAP74 to FCP1 may be mutually exclusive, since Tat can inhibit binding of RAP74 to this central domain of FCP1 (Abbott et al., unpublished). These interactions of Tat and TFIIF with the central domain of FCP1 may in part explain why Tat can inhibit FCP1 phosphatase activity *in vitro* (29), because RAP74 can stimulate FCP1 phosphatase activity (30). The CTD of FCP1 (hFCP1 812-961) is highly acidic. The amino acid composition of this domain is very similar to that found in acidic activation domains. In fact, FCP1 is capable

of activating transcription both *in vitro* and *in vivo* when artificially recruited to a promoter (19,31). In addition, *in vitro* and *in vivo* binding studies have demonstrated that the carboxyl-terminal domain of FCP1 interacts with the carboxyl-terminal region of RAP74 and the first cyclin repeat of TFIIB (16,30,32). The FCP1-binding sites for both RAP74 and TFIIB consist of a shallow groove on the surface of the protein that is rich in both hydrophobic and basic amino acids (33). These interactions are believed to be responsible in part for the observation that TFIIF stimulates FCP1 phosphatase activity whereas TFIIB inhibits the stimulation by TFIIF (30).

In addition to being a CTD phosphatase, it has been demonstrated that FCP1 itself exists as a phosphoprotein (34,35) that co-purifies with protein kinase CK2 (formerly known as casein kinase II) (34). These results suggest that CK2 could be an *in vivo* kinase for FCP1. *In vitro* studies examining the role of CK2 in FCP1 function demonstrated that phosphorylation of FCP1 by CK2 not only enhances RAP74 binding, but also stimulates CTD phosphatase activity (34,35). In contrast, elongation reactions conducted in the presence of CK2 are inhibited and this effect appears to be FCP1-specific, as CK2 added to elongation reactions in the absence of FCP1 had no effect (34). FCP1 is a highly acidic protein, and it contains several potential CK2 phosphorylation sites (35). The two RAP74-binding sites within FCP1 have been precisely mapped to highly conserved LXXLL-like hydrophobic motifs (36). The first RAP74-binding site is located within the central domain of FCP1 (16) (Abbott et al. unpublished), and the second one is located in the carboxyl-terminal domain of FCP1 (16,36). Interestingly, both of these conserved RAP74-binding sites contain a consensus CK2 phosphorylation site located immediately adjacent to it on its amino-terminal side. Given that CK2 phosphorylation enhances RAP74 binding to FCP1 (35), it is highly possible that phosphorylation of one or both of these

two CK2 sites located adjacent to the RAP74-binding sites may enhance RAP74 binding to FCP1.

In this manuscript, we further examined CK2 phosphorylation of FCP1. We verified that FCP1 isolated from human cells is phosphorylated *in vivo*. We identify consensus CK2 phosphorylation sites located immediately adjacent to both RAP74 binding sites of FCP1 that are highly conserved in all vertebrate FCP1 sequences. We demonstrate that both of these consensus CK2 sites can be phosphorylated *in vitro*, and that phosphorylation at either the central or carboxyl-terminal CK2 site results in enhanced binding of RAP74 to FCP1. The CK2 site adjacent to the RAP74-binding site in the central domain is phosphorylated at a single threonine site (T584), whereas the CK2 site adjacent to the RAP74-binding site in the carboxyl-terminal domain is phosphorylated at three serine residues (S942, S943 and S944). Using mass spectral analysis with synthetic peptide models of the carboxyl-terminal CK2 phosphorylation site, we demonstrate that the phosphorylation of the S942, S943 and S944 may represent a mechanism of phosphorylation by CK2 that has not been previously described. The three phosphorylations occur in a highly ordered fashion with the first phosphorylation occurring at S942 followed by subsequent phosphorylation at S944, and finally at S943. Using Nuclear Magnetic Resonance (NMR) spectroscopy, we identify chemical shift changes that occur within residues of RAP74 following interaction with FCP1 peptides phosphorylated at T584, S942, and S944. In addition, we map these chemical shift changes on the solution structure of the carboxyl-terminal domain of RAP74 (RAP74₄₃₆₋₅₁₇). These results provide the first detailed analysis of the structural changes in RAP74 following interaction with phosphorylated FCP1. In addition, these results provide new functional and structural information on the role of phosphorylation in the

recognition of acidic-rich domains involved in transcriptional regulation, and bring insights into how CK2 and TFIIF regulate FCP1 function.

MATERIALS and METHODS

Plasmids

The expression vectors for GST-RAP74₄₃₆₋₅₁₇ and GST-FCP1₈₇₉₋₉₆₁ have been described previously (16). The plasmid expressing the GST-FCP1₅₆₂₋₆₁₉ fragment used for *in vitro* binding studies was constructed by PCR amplification using pJA533 as a template (16). The amplified FCP1 DNA fragment was digested with *Bam*H1 and *Eco*R1 and then inserted in frame into the *Bam*H1 and *Eco*R1 sites of the pGEX5X-1 GST-fusion expression vector (Amersham Biosciences, NJ). Clones expressing GST-FCP1₅₆₂₋₆₁₉ (T584E), GST-FCP1₈₇₉₋₉₆₁(S942A) and GST-FCP1₈₇₉₋₉₆₁ (S944A) were prepared by site-directed mutagenesis of the wild-type clones using the Quick Change kit (Stratagene, CA). The mammalian expression construct HA-FCP1₁₋₉₆₁ was constructed in two steps. In the first step, oligonucleotides containing *Bam*H1 and *Eco*R1 overhangs encoding the hemagglutinin (HA) epitope (HA sense 5' GATCCCCATGGGATCC TACCCTTACGA CGTTCCTGATTACGCTAGCCTCG 3'; HA antisense 5' AATTCGAGGCT AGCGTAATCA GGAACGTCGTAAGGGTAGGATCCCATGGG 3') were annealed and cloned into the pCDNA3.1 vector (Invitrogen, CA). In the second step, full-length FCP1 cDNA was isolated by *Eco*R1 digestion of BSK4.1 (gift from D. Reinberg) and cloned into the *Eco*R1 digested pCDNA3.1-HA plasmid.

Antibodies

The FCP1 polyclonal antibody used was prepared as previously described (Archambault et al. 1998). The RAP74 polyclonal antibody (C-18) and all secondary antibodies were purchased from Santa Cruz Biotechnologies (Santa Cruz, CA). The HA monoclonal antibody (clone 12CA5) was purchased from Roche Molecular Biochemicals (Indianapolis, IN)

Purification of GST-FCP1 Fusion Proteins

GST-FCP1 fusion proteins were expressed in DH5 α cells (Stratagene, CA). Cells were grown at 37 °C and protein expression was induced with 0.7 mM isopropyl-thiogalactoside (IPTG) for 3 h at 30 °C. The cells were harvested and resuspended in EBC buffer (50 mM Tris-HCl pH 8.0, 120 mM NaCl, 0.5 % NP-40, and 2 mM DTT) supplemented with protease inhibitors (Sigma, MO): 1 μ g/ml leupeptin, 2 μ g/ml aprotinin, and 0.9 mg/ml PMSF. The cells were lysed by passage through a French press cell and centrifuged at 100,000 x g for 40 min. The supernatant obtained from the high-speed centrifugation was added to glutathione (GSH)-sepharose resin (Amersham Biosciences, NJ) and incubated with mixing at 4 °C for 30 min. The resin was collected by centrifugation and washed four times with wash buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5 % NP-40, 0.05 % SDS, and 1 mM DTT). GST-FCP1 fragments were eluted off the resin by incubation with 15 mM reduced GSH (Sigma, MO) for 10 min at 25 °C. The proteins were then dialyzed into storage buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 2 mM DTT, and 20 % glycerol), and stored at -80 °C. The GST-FCP1 fragments were rebound to fresh GSH-sepharose resin prior to protein-protein binding studies.

Peptide Synthesis and purification

Peptides FCP1937-961, FCP1937-961(S942P0₄), and FCP1937-961(S942P0₄/S944P0₄) were purchased from the peptide synthesis Facility at the Medical College of Georgia (Augusta, GA). The FCP1579-600, FCP1579-600 (T584P0₄), FCP1941-961, and FCP1584-607 peptides were synthesized by the solid phase method with Fmoc chemistry using an Applied Biosystems (Foster City, CA) 430A peptide synthesizer. The FCP1579-600, FCP1579-600 (T584P0₄) peptides were prepared with a glycine cap at the amino-terminus. The phosphothreonine was incorporated as FmocThr(PO(OBzl)OH)-OH (Novabiochem, San Diego, CA). The FCP1579-600 and FCP1579-600 (T584P0₄) peptides were cleaved from the resin, and side chain protecting groups were removed by incubating in reagent K (trifluoroacetic acid (TFA) : phenol : thioanisole : H₂O : ethanedithiol, 82.5:5:5:5:2.5) for 3 h at room temperature. All seven peptide were purified to homogeneity on a Vydac C₄ (Hesperia, CA) reverse phase HPLC column (22 X 250 mm) using an acetonitrile gradient (30% to 50% over 20 min) in 0.05% TFA at a flow rate of 8 ml/min. The masses of peptides were confirmed by electrospray ionization mass spectrometry (MS).

Transfection and Immunoprecipitation of HA-FCP1₁₋₉₆₁

Human embryonic kidney (HEK) 293 cells were grown in a humidified 5% CO₂ atmosphere at 37 °C in growth media consisting of DMEM (4.5 µg/ml glucose) and 10% fetal bovine serum. Cells were plated at 70% confluence in 10 cm diameter dishes and transfected with 40 µg HA-FCP1₁₋₉₆₁ plasmid DNA using calcium phosphate and 25 µM chloroquine for 6-12 h before refeeding with fresh growth media. Forty-eight hours after the transfection, the cells were washed twice with phosphate buffered saline and harvested in RIPA-2 buffer (50 mM Tris-

HCl pH 8.0, 150 mM NaCl, 1% Nonidet P-40, 0.5% deoxycholate, and 0.1% SDS,) containing 90 µg/ml phenylmethylsulfonyl fluoride, 0.2 TIU/ml aprotinin, 0.1 mM leupeptin, 0.2 mM sodium orthovanadate, 5 µg/ml pepstatin A, and 5 µg/ml soybean trypsin inhibitor. Insoluble material was pelleted by centrifugation and the supernatant (lysate) was frozen in liquid nitrogen and stored at -80 °C until needed.

To determine the phosphorylation state of FCP1 from human cells, ~ 2 µg of the anti-HA monoclonal antibody (Roche Diagnostics) was added to 200 µl lysate, which corresponded to ~ 1×10^7 cells, and the reaction was rotated at 4 °C for 2 h. Next, 20 µl of agarose A/G conjugate (Santa Cruz Biotechnologies, CA) was added and incubated with rotation overnight at 4 °C. The agarose A/G resin bound proteins were centrifuged to a pellet and then washed 4 times with ice-cold RIPA-2 buffer with repelleting after each wash step. Following the final wash, the resin-bound proteins were resuspended in 1X SDS electrophoresis buffer (62.5 mM Tris-Cl pH 6.8, 10% glycerol, 1.5% DTT, 0.001% bromophenol blue, and 2% SDS) and resolved on 7% SDS polyacrylamide gel. The resolved FCP1 proteins were then detected by Western blot analysis. Briefly, the proteins were transferred to Immobilon P membrane (Millipore, MA) and identified using the rabbit FCP1 polyclonal antibody and a goat HRP-conjugated secondary antibody visualized using ECL Plus Western Blotting Detection Kit (Amersham Pharmacia, NJ).

FCP1 protein for MS analysis was obtained using the immunoprecipitation procedure described above except that FCP1 polyclonal antibody was used instead of the anti-HA monoclonal antibody. The immunoprecipitated proteins were resuspended in 1X SDS electrophoresis buffer, resolved on a 7% SDS polyacrylamide gel, stained with Coomassie blue, and excised from the stained gel prior to MS analysis.

Phosphorylation Studies

Phosphorylation of HA-Tagged Immunoprecipitated FCP1: HA-tagged FCP1 protein was immunoprecipitated as described above using 2 µg of HA monoclonal antibody and 10 µl of protein A/G agarose resin. The resin-bound protein was then diluted to 25 µl in the same reaction buffer described above. Phosphorylation reactions were initiated by the addition of 5 µl (10 units) protein kinase CK2 enzyme (New England Biolabs, MA) diluted in 20 mM Tris-HCl pH 7.5, 200 mM NaCl, 0.5 mM DTT, 10% glycerol, and 0.5% Triton X-100. The reactions were incubated at 25 °C for 1 h, and then terminated by the addition of 1X SDS electrophoresis buffer. For immunoprecipitation reactions treated with alkaline phosphatase, 10 units of calf intestine alkaline phosphatase (NEB, MA) was added to 10 µl of protein A/G agarose containing immunoprecipitated HA-FCP1₁₋₉₆₁. The reactions were incubated at 37 °C for 1 h, and then terminated by the addition of 1X SDS electrophoresis buffer. Immunoprecipitated HA-FCP1₁₋₉₆₁ phosphorylation, and alkaline phosphatase reactions were resolved on 7% SDS polyacrylamide gels in 1X Tris-glycine buffer. Immunoprecipitated FCP1 protein was detected by Western blot analysis as described above.

Phosphorylation of Purified GST-FCP1 Fusion Proteins: Purified GST-FCP1 fusion proteins were first captured on 10 µl of GSH-sepharose resin. The resin-bound proteins were then diluted to a final protein concentration of 1 µM in 25 µl of CK reaction buffer (20 mM Tris-HCl pH 7.5, 25 mM KCl, 10 mM MgCl₂, 350 µM ATP, 2.5 µCi [γ ³²P]-ATP (for phosphorylation reactions without ³²P-ATP cold ATP was added at 0.5 mM)). Phosphorylation reactions were initiated by the addition of 5 µl (10 units) protein kinase CK2 enzyme (NEB, MA) diluted in 20 mM Tris-HCl pH 7.5, 200 mM NaCl, 0.5 mM DTT, 10% glycerol, and 0.5% Triton X-100. The reactions were terminated by addition of SDS electrophoresis buffer, and the proteins were

resolved on NuPage 12% gels in 1X MES electrophoresis buffer (Invitrogen, CA). ^{32}P -labeled GST-FCP1 fusion proteins were detected by PhosphorImager analysis (Molecular Dynamics, Amersham Pharmacia, NJ).

Phosphorylation of Purified Synthetic Peptides: For phosphorylation reactions, peptides (1 μM) and the appropriate controls (1 μM) were dissolved in the CK2 reaction buffer described above. Phosphorylation reactions were initiated by the addition of 5 μl (10 units) protein kinase CK2 enzyme (NEB, MA) diluted in 20 mM Tris-HCl pH 7.5, 200 mM NaCl, 0.5 mM DTT, 10% glycerol, and 0.5% Triton X-100. The reactions were terminated by the addition of 1X SDS electrophoresis buffer, and phosphorylated peptides were resolved on 15% SDS polyacrylamide gels in 1X Tris-glycine buffer. ^{32}P -labeled peptides were detected by PhosphorImager analysis (Molecular Dynamics, Amersham Pharmacia, NJ).

Protein-Protein Binding Assay

For binding reactions quantified and graphed in Fig. 3, GST-FCP1₈₇₉₋₉₆₁, GST-FCP1₈₇₉₋₉₆₁(S942A), and GST-FCP1₈₇₉₋₉₆₁(S944A) were first purified on GSH-sepharose. The purified, resin-bound GST-FCP1 fusion proteins were then phosphorylated with CK2 prior to the binding studies. The CK2 phosphorylation reactions were performed at 25 °C from 0 to 60 min as described above, except the reactions were terminated by washing twice with 20 mM Tris-HCl pH 8.0, 120 mM NaCl, 25 mM EDTA. Binding experiments were then performed at 4 °C for 1 h in 250 μl binding buffer (40 mM HEPES pH 8.0, 120 mM NaCl, 0.5% NP-40, 10 mM DTT). The binding experiment consisted of 1 μM of the CK2-treated GST-FCP1 fusion proteins (GST-FCP1₈₇₉₋₉₆₁, GST-FCP1₈₇₉₋₉₆₁ (S942A), or GST-FCP1₈₇₉₋₉₆₁ (S944A)) and 1 μM of purified RAP74₄₃₆₋₅₁₇. The resin-bound protein-protein complexes were collected and washed three times in binding buffer. Proteins were resuspended in 1X SDS electrophoresis buffer and resolved on

NuPage Bis-Tris 12% gels in 1X MES electrophoresis buffer (Invitrogen, CA) and transferred to Immobilon-P membrane (Millipore, MA). The bound RAP74 was detected using a 1:400 dilution of RAP74 (C-18) polyclonal antibody and a 1:10,000 dilution of an HRP-conjugated secondary antibody. Bound secondary antibodies were detected by chemiluminescence using the ECL-Plus kit (Amersham Biosciences, NJ). Bands were quantified using the Bio-Rad Versa Doc Imaging System. The membrane was stained subsequently with Coomassie blue to verify equivalent GST-fusion protein input. Normalized values from three separate experiments were averaged and graphed in Figure 3.

Mass Spectrometry

Phosphorylation of Peptides for Mass Spectrometry: Peptides were dissolved in 450 μ l of CK2 reaction buffer described above. Phosphorylation reactions were initiated by the addition of 50 μ l (100 units) protein kinase CK2 enzyme (NEB, MA) diluted in 20 mM Tris-HCl pH 7.5, 200 mM NaCl, 0.5 mM DTT, 10% glycerol, and 0.5% Triton X-100 at 25 °C. After addition of the enzyme, the final peptide concentration was 0.3 mM. 25 μ l samples were removed 1 h, 2 h, 4 h, 8 h and 24 h after addition of the CK2 enzyme. The reactions were terminated by the addition of 25 μ l of glacial acetic acid and lyophilized. In some cases, the phosphorylated peptides were resuspended in 25% aqueous acetic acid and resolved on a Vydac C₄ (Hesperia, CA) reverse phase HPLC column (4.6 X 250 mm) using an acetonitrile gradient (30% to 50% over 20 min) in 0.05% TFA at a flow rate of 1 ml/min.

Preparation of Samples for ESI: Samples from the peptide phosphorylation reactions were resuspended in 50 μ L of water, and 10 μ l of the stock solution was diluted to 50 μ L. Prior to electrospray ionization (ESI), samples were desalted with a C₁₈ ZipTip (Millipore, Billerica, MA) into 30 μ l of a 4:1 CH₃CN:H₂O solution containing 0.1% - 1% formic acid. Desalted samples were then microelectrosprayed from an emitter consisting of a 50 μ m i.d. fused silica

capillary that had been mechanically ground to a uniform thin-walled tip (37) at a flow rate of 400 nl/min.

β -elimination/Michael Addition: Reactions were carried out based upon a slightly modified version of the procedure described by Molloy and Andrews (38). 3.7 μ l (5 μ g) of the FCP1₉₃₇₋₉₆₁ peptide stock solution following phosphorylation was combined with 13 μ l H₂O, 9 μ l 1-propanol, 3 μ l of 200 mM Ba(OH)₂, and 1.3 μ l ethanethiol to yield final concentrations of 30% (v/v) 1-propanol, 20 mM Ba(OH)₂, and 0.5 M ethanethiol. The solution was then incubated at 45 °C for between 0.5 h and 5 h and the reaction was terminated by the addition of 15 μ l of 1 M ammonium sulfate to form a barium salt. The barium sulfate was then removed by centrifugation at 10,000 $\times$ g for 5 min, and the supernatant was desalted and electrosprayed as described above.

9.4 Tesla Fourier Transform-Ion Cyclotron Resonance (FT-ICR) MS: Experiments were performed with a homebuilt 9.4 Tesla ESI Q FT-ICR mass spectrometer (39) under the control of a modular ICR data acquisition system (MIDAS) (40). Ions were transported through a Chait-style atmosphere-to-vacuum interface (41) and accumulated within a linear octopole ion trap, modified to allow improved ion ejection along the z-axis (42). Analyte ions were then transferred (1.0 - 1.4 ms) through an octopole ion guide to an open cylindrical ICR cell in which ions were captured by gated trapping. Ions were subjected to chirp (72-480 kHz at 150 Hz / μ s) excitation (43,44) and direct-mode broadband detection (512 k time-domain data). Hanning apodization and one zero-fill were applied to all data prior to fast Fourier transformation and magnitude calculation (45). Frequency-domain spectra were calibrated (46) externally from the measured ICR frequencies of Agilent calibration mixture ions (m/z 622.02895, 922.00979, 1521.97146). Each displayed spectrum represents a sum of 1-50 time-domain transients.

Masses and m/z values were calculated with Isopro 3.1 (MS/MS software, <http://members.aol.com/msmssoft/>).

FT-ICR MS/MS: Precursor ions were mass-selectively accumulated externally for 20-40 s (47). Following transfer to the ICR cell (1.0 ms), stored-waveform inverse Fourier transform (SWIFT) (48,49) ejection was applied to further isolate the peptide ion under investigation. (Note: the instrumental configuration and experimental parameters for Electron Capture Detection (ECD), InfraRed Multi Photon Dissociation (IRMPD), and Activated Ion Electron Capture Dissociation (AI-ECD) are described in depth elsewhere) (39).

For IRMPD, precursor ions were irradiated with a 40 W, 10.6 μm , CO_2 laser (no beam expander, Synrad, Mukilteo, WA), Photon irradiation was for 100 ms at 35 % laser power (14 W). For ECD, precursor ion populations were irradiated with low energy electrons for 10 ms. For activated ion AI- ECD, precursor ion isolation and electron irradiation was performed as described above. Following electron irradiation, ions were irradiated with the IR laser for 100 ms at 10-20% laser power (4-8 W). Following MS/MS (ECD, IRMPD, or AI-ECD), product ions were subjected to chirp excitation (58 to 480 kHz at 150 Hz/ μs) and direct-mode broadband detection (512 K time-domain data). Displayed spectra represent the sum of 50-60 time domain transients.

NMR Spectroscopy.

The ^{15}N -labeled RAP74₄₃₆₋₅₁₇ was expressed from a pGEX-2T vector (Amersham Biosciences, NJ) in BL21(DE3). Labeled protein was obtained by growth in a modified minimal medium containing ^{15}N -labeled NH_4Cl as the sole source of nitrogen as previously described (33). NMR samples of the complexes were formed by addition in four equal increments of the

appropriate peptide to ^{15}N -labeled RAP74₄₃₆₋₅₁₇ to obtain the final 1:1 complex. The final NMR sample consisted of 0.20 mM ^{15}N -labeled RAP74₄₃₆₋₅₁₇ and 0.20 mM FCP1 peptides (FCP1₅₇₉₋₆₀₀, FCP1₅₇₉₋₆₀₀ (T584PO₄), FCP1₉₃₇₋₉₆₁, FCP1₉₃₇₋₉₆₁(S942PO₄) and FCP1₉₃₇₋₉₆₁(S942PO₄/S944PO₄) in 500 μl of 20 mM sodium phosphate pH 6.5 and 1 mM EDTA (90% H₂O/ 10% D₂O). The 2D ^1H - ^{15}N HSQC's (50) were collected at 27 °C using a Varian Inova Unity 600 MHz NMR spectrometer equipped with a z pulsed-field gradient unit and a HCN triple resonance probe. The NMR data were processed with NMRPipe / NMRDraw (51). and ribbon diagrams were made using the program MOLMOL (52).

Model of FCP1₈₇₉₋₉₆₁(S942PO₄/S944PO₄)/ RAP74₄₃₆₋₅₁₇ Complex.

Three dimensional structures of the FCP1₈₇₉₋₉₆₁ (S942PO₄/S944PO₄)/ RAP74₄₃₆₋₅₁₇ were calculated using Torsion Angle Molecular Dynamics protocol of CNS, starting from two extended structures (one for RAP74 (residue 451-517) and one for FCP1 (residue 879-961) with standard geometry and modified phosphoserine (residues 942 and 944). To model the potential ionic bridges between K475 / S942PO₄ and K480 / S944PO₄ two constraints were added between K475 HZ and S942PO₄ OP (oxygen of the phosphate group) and between K480 HZ and S944PO₄ OP. These constraints were added on the basis of chemical shift changes in the NMR studies. These restraints are set from 1.8 to 2.5 Å using the force constant of 75 kcal/mole, similar to those used for the NOE restraints. 50 structures were calculated and the 20 structures with lowest energies were chosen for further evaluation and for determination of the average structure. These structures had no NOE violation greater than 0.2 Å and no dihedral angle (phi and psi) violation greater than 5 degree.

RESULTS

FCP1 isolated from human cells is phosphorylated.

Previous experiments have demonstrated that both *Xenopus* FCP1 (xFCP1) isolated from *Xenopus* A6 cells (35) and human FCP1 isolated from baculovirus-infected SF9 cells exist as phosphoproteins (34). To determine if human FCP1 isolated from human cells is also a phosphoprotein, HEK 293 cells were transfected with a HA-FCP11-961 expression construct, and the HA-tagged full-length FCP1 protein produced in the HEK 293 cells was immunoprecipitated with a monoclonal antibody specific for the HA-tag. The immunoprecipitated FCP1 protein was then subjected to treatment with either CK2 (Fig. 1 lane 3) or alkaline phosphatase (Fig. 1 lane 4). Following the enzyme treatments, the mobility of treated FCP1 was compared with the mobility of untreated immunoprecipitated FCP1 (Fig. 1 lane 2). The mobilities were determined using denaturing gel electrophoresis followed by Western blot detection with a polyclonal antibody against FCP1. Immunoprecipitation of FCP1 followed by treatment with CK2 did not alter the mobility of the FCP1 relative to the untreated immunoprecipitated protein, but immunoprecipitation of FCP1 followed by alkaline phosphatase treatment resulted in an FCP1 protein with increased mobility relative to the untreated FCP1 protein. Based on these results, we conclude that FCP1 does exist as a phosphoprotein when expressed in human embryonic kidney cells.

CK2 phosphorylation sites in FCP1 located adjacent to RAP74 binding sites.

Earlier studies have demonstrated that phosphorylation of xFCP1 by CK2 leads to enhanced binding of xFCP1 to xRAP74 *in vitro*, and it is believed that this enhanced binding to

RAP74 may be crucial for FCP1 function (35). However, the exact mechanism by which CK2 phosphorylation of FCP1 leads to enhanced binding to RAP74 is unknown. There are two known RAP74-binding sites in FCP1 and either of these two sites or both could be responsible for the enhanced binding observed after phosphorylation of FCP1 by CK2 (15,16,32,35). The first RAP74-binding site is located in the central domain of FCP1 on the amino-terminal side of the BRCT domain (Figure 2A, residues 579-600) (Abbott et al., unpublished), and the second RAP74-binding site is located at the extreme carboxy-terminus of FCP1 (Figure 2B, residues 937-961). Interestingly, both RAP74-binding sites in FCP1 are located immediately adjacent to consensus CK2 phosphorylation sites (53) and these predicted CK2 phosphorylation sites are highly conserved in many species with higher conservation in vertebrates (highlighted in Figure 2A and Figure 2B). The central domain contains a threonine CK2 phosphorylation site (T584), whereas the carboxyl-terminal domain contains two serine CK2 phosphorylation sites (S942 and S944). Our initial goal was to determine if the predicted CK2 sites at either the central domain or the carboxyl-terminal domain could be phosphorylated by CK2 *in vitro* and if phosphorylation at these sites by CK2 resulted in enhanced binding of FCP1 to RAP74.

CK2 phosphorylation of T584 enhances RAP74 binding to the FCP1 *in vitro*.

As mentioned above, the central domain of FCP1 contains a highly conserved consensus CK2 phosphorylation site, and in hFCP1 the predicted site of phosphorylation is at amino acid T584 (Figure 3A). In order to determine if the central domain of FCP1 could be phosphorylated *in vitro* by CK2 at T584, equal molar concentrations (1 μ M) of casein, GST, GST-FCP1₅₆₂₋₆₁₉, and GST-FCP1₅₆₂₋₆₁₉ (T584E) were incubated for 30 minutes with recombinant CK2 in the presence of [γ -³²P]-ATP. Following the phosphorylation reaction with CK2, the various proteins

were checked for incorporation of the ^{32}P -label (Figure 3B). It is clear that both the control casein and GST-FCP1₅₆₂₋₆₁₉ were extensively phosphorylated by CK2 *in vitro*, but that GST and GST-FCP1₅₆₂₋₆₁₉ (T584E) were not phosphorylated to any significant degree. Based on these experimental results (Figure 3B), we conclude that CK2 is capable of phosphorylating GST-FCP1₅₆₂₋₆₁₉ specifically at residue T584 *in vitro*, and that T584 is the primary CK2 phosphorylation site within the central domain *in vitro*.

Since CK2 is capable of phosphorylating GST-FCP1₅₆₂₋₆₁₉ at T584 *in vitro*, we attempted to determine if phosphorylation at T584 led to enhanced binding of RAP74 to FCP1₅₆₂₋₆₁₉. Resin-bound GST and GST-FCP1₅₆₂₋₆₁₉ were first treated for 30 min with either a mock CK2 phosphorylation reaction (all components except enzyme) or a CK2 phosphorylation reaction. Following these CK2 treatments, the resin-bound GST and GST-FCP1₅₆₂₋₆₁₉ were used in binding reactions with purified RAP74₄₃₆₋₅₁₇ (Figure 3C). RAP74 binding to FCP1 was determined by Western blot detection using a polyclonal antibody against RAP74. In these experiments, RAP74₄₃₆₋₅₁₇ bound to both GST-FCP1₅₆₂₋₆₁₉ treated with the mock reaction (Figure 3C lane 4) and to GST-FCP1₅₆₂₋₆₁₉ treated with the CK2 reaction (Figure 3C lane 5). However, the CK2-treated GST-FCP1₅₆₂₋₆₁₉ displays enhanced binding to RAP74₄₃₆₋₅₁₇ relative to the mock-treated GST-FCP1₅₆₂₋₆₁₉. As expected, RAP74₄₃₆₋₅₁₇ did not bind to either the GST mock or GST CK2 reaction (Figure 3C lanes 2 and 3, respectively).

CK2 phosphorylation of S942 and S944 enhances RAP74 binding to the FCP1 *in vitro*.

Like the central domain, the RAP74-binding domain at the extreme carboxyl-terminus of FCP1 is also located adjacent to a CK2 phosphorylation site. The consensus CK2 site at the carboxyl-terminus contains three consecutive serine residues (S942-S944) sandwiched between

numerous acidic amino acids (Figure 3A). It is predicted that both S942 and S944 should be ideal targets for CK2 phosphorylation (acidic residue n+3 to serine) (53). In order to determine if the carboxyl-terminal domain of FCP1 could be phosphorylated *in vitro* by CK2, equal molar concentrations (1 μ M) of casein, GST, GST-FCP1₈₇₉₋₉₆₁, GST-FCP1₈₇₉₋₉₆₁ (S942A), and GST-FCP1₈₇₉₋₉₆₁ (S944A) (Figure 3D) were incubated for 10 and 30 min with recombinant CK2 in the presence of [γ -³²P]-ATP. Following the phosphorylation reaction with CK2, the various proteins were checked for incorporation of the ³²P-label (Figure 3D). It is clear that CK2 is capable of phosphorylating GST-FCP1₈₇₉₋₉₆₁, GST-FCP1₈₇₉₋₉₆₁ (S942A), and GST-FCP1₈₇₉₋₉₆₁ (S944A), and they all appear to be phosphorylated to similar extents. There are two plausible explanations for these results. The first explanation is that both S942 and S944 are phosphorylated by CK2 *in vitro* as predicted. The second explanation is that CK2 is phosphorylating the carboxyl-terminal region of FCP1 at residues other than either S942 or S944.

Since CK2 is capable of phosphorylating FCP1₈₇₉₋₉₆₁ *in vitro*, we attempted to determine if this phosphorylation led to enhanced binding of RAP74 to FCP1₈₇₉₋₉₆₁. Resin-bound GST and GST-FCP1₈₇₉₋₉₆₁ were first treated for 30 min with either a mock CK2 phosphorylation reaction or a CK2 phosphorylation reaction. Following the treatments, the resin-bound GST and GST-FCP1₈₇₉₋₉₆₁ were used in binding reactions with purified RAP74₄₃₆₋₅₁₇ (Figure 4A). As before, RAP74 binding was determined by Western blot detection using a polyclonal antibody against RAP74. RAP74₄₃₆₋₅₁₇ bound to both FCP1₈₇₉₋₉₆₁ treated with the mock reaction (Figure 4A lane 4) and to FCP1₈₇₉₋₉₆₁ treated with the CK2 reaction (Figure 4A lane 5), but the CK2-treated FCP1₈₇₉₋₉₆₁ displays significantly enhanced binding to RAP74₄₃₆₋₅₁₇ relative to the mock-treated FCP1₈₇₉₋₉₆₁. RAP74₄₃₆₋₅₁₇ did not bind to either the GST mock or GST CK2 reaction (Figure 4A lanes 2 and 3, respectively).

Next, we tested whether the S942A or S944A mutations would have any effect on the enhanced binding of FCP1₈₇₉₋₉₆₁ to RAP74 following CK2 phosphorylation. First, purified GST-FCP1₈₇₉₋₉₆₁, GST-FCP1₈₇₉₋₉₆₁ (S942A), and GST-FCP1₈₇₉₋₉₆₁ (S944A) were bound to GSH-Sepharose. The resin-bound proteins were then incubated for 0, 10, and 60 min with recombinant CK2. Following the CK2 treatment, the resin-bound GST-FCP1₈₇₉₋₉₆₁, GST-FCP1₈₇₉₋₉₆₁ (S942A), and GST-FCP1₈₇₉₋₉₆₁ (S944A) proteins were used in binding reactions with purified RAP74₄₃₆₋₅₁₇ (Figure 4B). Surprisingly, both the S942A mutation and the S944A mutation caused a significant reduction in the enhanced binding of FCP1₈₇₉₋₉₆₁ to RAP74 following CK2 phosphorylation. These results in combination with the previously described phosphorylation studies suggest that both S942 and S944 must be phosphorylated by CK2 *in vitro* for the enhanced binding of RAP74 to FCP1₈₇₉₋₉₆₁.

S942, S943, and S944 are phosphorylated by CK2 in a highly ordered fashion.

In order to demonstrate that both S942 and S944 of FCP1 could be phosphorylated by CK2, we prepared a series of synthetic peptides for phosphorylation experiments *in vitro* (Figure 4C). In the initial experiments, the peptides were incubated with recombinant CK2 in the presence of [γ -³²P]-ATP, and the phosphorylation reaction was monitored by checking for incorporation of the ³²P-label as described above for the longer FCP1 fragments. We tested four synthetic peptides as potential CK2 substrates corresponding to FCP1₉₃₇₋₉₆₁, FCP1₅₇₉₋₆₀₀, FCP1₉₄₁₋₉₆₁, and FCP1₅₈₄₋₆₀₇ (Figure 4C). Interestingly, the peptides corresponding to FCP1₉₃₇₋₉₆₁ and FCP1₅₇₉₋₆₀₀ are phosphorylated *in vitro* by CK2, but the peptides corresponding to FCP1₉₄₁₋₉₆₁ and FCP1₅₈₄₋₆₀₇ are not. In the case of the peptide corresponding to FCP1₉₄₁₋₉₆₁, we were rather surprised that there was no apparent phosphorylation at S944. These results were also

supported by HPLC analysis of the incubations, where we failed to see any new species formed in the presence of CK2 (data not shown).

Next, we used the synthetic peptide corresponding to FCP1₉₃₇₋₉₆₁ to demonstrate that both S942 and S944 could be phosphorylated by CK2 *in vitro* and to determine if there was a preference for phosphorylation at S942 or S944. The FCP1₉₃₇₋₉₆₁ peptide was incubated with recombinant CK2 for varying periods of time (0-48 hours) followed by addition of acetic acid to terminate the reaction. The reactions were then analyzed by HPLC to verify the presence of phosphorylated peptides. The HPLC analysis (data not shown) showed that following the addition of CK2, a large percentage (~ 60-70%) of the peptide was converted into two new species within the first four hours. Longer incubations (21-48 hours) with CK2 resulted in a third new species. Our initial hypothesis was that first two new species corresponded to mono-phosphorylated FCP1₉₃₇₋₉₆₁ and di-phosphorylated FCP1₉₃₇₋₉₆₁, while the third new species corresponded to the tri-phosphorylated FCP1₉₃₇₋₉₆₁ peptide.

In order to verify this hypothesis and to determine which serines were phosphorylated in the FCP1₉₃₇₋₉₆₁ peptide, we performed extensive analysis using mass spectrometry (MS). Following incubation with CK2 for 4 hours, both IRMPD and ECD FT-ICR MS/MS analysis was performed on the CK2 treated FCP1₉₃₇₋₉₆₁ peptide. Using the IRMPD method, the location of the phosphorylations could not be determined because of limited sequence information and dominant losses of H₃PO₄ (data not shown). Using the ECD method, the spectrum demonstrated the presence of (N₉₃₇EDEGSSSEADEMAKALEAELNDLM₉₆₁ + HPO₃ + 3H)³⁺ ions (10 ms electron irradiation, sum of 50 time-domain transients) following 20 s of mass-selective external ion accumulation (Figure 5). The detection of the [c₆ + HPO₃ + H]⁺ ions, combined with the detection of [c₇ + HPO₃]⁺ and [c₈ + HPO₃]⁺ ions, establishes that the site of phosphorylation is at

serine 942. Additional FT-ICR MS/MS analyses of the FCP1₉₃₇₋₉₆₁ peptide after incubation with CK2 for varying lengths of time (2 h to 21 h) indicate the presence of just a mono-phosphorylated FCP1₉₃₇₋₉₆₁ peptide phosphorylated at S942 (data not shown).

The ESI mass spectrum obtained from the analysis of the FCP1₉₃₇₋₉₆₁ peptide following incubation with CK2 clearly demonstrated the presence of the un-phosphorylated peptide, and the mono-phosphorylated peptide. However, HPLC analysis of an equivalent sample indicated the presence of additional species that we attributed to be di- and tri-phosphorylated FCP1₉₃₇₋₉₆₁ peptide. The inability to detect the presence of either di or tri-phosphorylated FCP1₉₃₇₋₉₆₁ peptide ions could be due to their low ionization efficiencies.

To further investigate the presence of the di and tri-phosphorylated peptide species, we attempted MS analysis specifically designed for detection of phosphopeptides (54). In this method, the charged phosphate groups are specifically removed and replaced with ethanethiol. This results in an increase in the ionization efficiency of the peptide during positive ion mode ESI (when compared to the native phosphopeptide). ESI FT-ICR MS analysis of derivatized FCP1₉₃₇₋₉₆₁ peptide following incubation with CK2 for 4 h revealed the presence of both the di- and tri phosphorylated FCP1₉₃₇₋₉₆₁ peptide (Figure 6).

Once the presence of the di- and tri-phosphorylated FCP1₉₃₇₋₉₆₁ peptides was established, tandem MS was used to identify the exact sites phosphorylated. This was possible because the mass of the de-phosphorylated and modified serine residue is unique among all common amino acids (except methionine), and the sequence of the peptide is known. The mass spectrum obtained following AI-ECD of $(N_{937}EDEGSSSEADEMAKALEAELNDLM_{961} + 2(C_2H_6S) + 3H)^{3+}$ demonstrates that S942 and S944 are specifically modified and not S943 (Figure 7). This was determined from the fact that the mass difference between the c_5-c_6 and c_7-c_8 ions was 131

Da whereas the mass difference between the c_6 and c_7 ions was 87 Da (mass of an unmodified serine residue). If the CK2 phosphorylation of FCP1₉₃₇₋₉₆₁ peptide was non-specific, precursor ion population could be comprised of a heterogeneous mixture of di-phosphorylated FCP1₉₃₇₋₉₆₁ peptide ions (942* + 943*, 943* + 944* or 942* + 944*). Although these isobaric peptides could not be distinguished based on mass alone, tandem mass spectrometry allows us to determine if the precursor ion population is homogeneous or not. The results clearly demonstrate the first site of phosphorylation is S942, followed by subsequent phosphorylation of S944 to yield the di-phosphorylated species and finally by phosphorylation at S943 to give the tri-phosphorylated species.

Mapping the binding site of S942PO₄ and S944PO₄ on the structure of RAP74₄₃₆₋₅₁₇.

In order to map the interaction site for the two phosphorylated serines in the carboxyl-terminal of FCP1 (S942PO₄ and S944PO₄) on the solution structure of RAP74₄₃₆₋₅₁₇, we have recorded 2D ¹H-¹⁵N HSQC spectra of three complexes: FCP1₉₃₇₋₉₆₁ / ¹⁵N-labeled RAP74₄₃₆₋₅₁₇, FCP1₉₃₇₋₉₆₁ (S942PO₄) / ¹⁵N-labeled RAP74₄₃₆₋₅₁₇, and FCP1₉₃₇₋₉₆₁ (S942PO₄/S944PO₄) / ¹⁵N-labeled RAP74₄₃₆₋₅₁₇. As expected, the 2D ¹H-¹⁵N HSQC spectrum of FCP1₉₃₇₋₉₆₁ bound to ¹⁵N-labeled RAP74₄₃₆₋₅₁₇ is virtually identical to that previously reported for FCP1₈₇₉₋₉₆₁ bound to ¹⁵N-labeled RAP74₄₃₆₋₅₁₇ (33) (data not shown). To identify the binding site for both S942PO₄ and S944PO₄ from FCP1 on the solution structure of RAP74₄₃₆₋₅₁₇, we first computed the differences in ¹H and ¹⁵N chemical shifts of the RAP74₄₃₆₋₅₁₇ / FCP1₉₃₇₋₉₆₁ complex and the RAP74₄₃₆₋₅₁₇ / FCP1₉₃₇₋₉₆₁ (S942PO₄/S944PO₄) complex. This analysis reveals that 6 out of 82 amino acid residues of RAP74₄₃₆₋₅₁₇ displayed significant chemical shift changes ($\Delta\delta > 0.07$; $\Delta\delta = [(0.17\Delta N_H)^2 + (\Delta H_N)^2]^{1/2}$) upon formation of the complex with the di-phosphorylated peptide

(S942PO₄/S944PO₄). These residues correspond to positions K475, K476, F477, T479, K480, and S485 of RAP74₄₃₆₋₅₁₇. When these significantly shifted residues are mapped on the NMR solution structure of RAP74₄₃₆₋₅₁₇ (Figure 8A), we see that the region near the end of helix H2 and the beginning of helix H3, and the flexible loop (L2) that connects H2 to H3 appears to be crucial for the binding of S942PO₄ and the S944PO₄. Interestingly, this area is extremely rich in lysine residues (K475, K476, K480, K481), and thus it is possible that several of these lysine residues form an ion pair with either S942PO₄ or S944PO₄. In fact, three of these lysine residues (K475, K476 and K480) displayed significant changes in chemical shifts upon formation of the complex with the di-phosphorylated peptide.

In an attempt to specifically identify the lysine residues that interact with either the S942PO₄ or the S944PO₄, we first computed the differences in ¹H and ¹⁵N chemical shifts of the RAP74₄₃₆₋₅₁₇ / FCP1₉₃₇₋₉₆₁ complex and the RAP74₄₃₆₋₅₁₇ / FCP1₉₃₇₋₉₆₁ (S942PO₄) complex. This analysis reveals that 5 out of 82 amino acid residues of RAP74₄₃₆₋₅₁₇ displayed significant changes ($\Delta\delta > 0.07$; $\Delta\delta = [(0.17\Delta N_H)^2 + (\Delta H_N)^2]^{1/2}$) upon formation of the complex with the mono-phosphorylated peptide. These residues correspond to positions K475, K476, F477, T479, and S485 of RAP74₄₃₆₋₅₁₇. We then computed the differences in ¹H and ¹⁵N chemical shifts of the RAP74₄₃₆₋₅₁₇ / FCP1₉₃₇₋₉₆₁ (S942PO₄) complex and the RAP74₄₃₆₋₅₁₇ / FCP1₉₃₇₋₉₆₁ (S942PO₄/S944PO₄) complex. This analysis reveals that when comparing the complexes containing the mono-phosphorylated and the di-phosphorylated peptides none of the residues of RAP74₄₃₆₋₅₁₇ displayed significant differences ($\Delta\delta > 0.07$; $\Delta\delta = [(0.17\Delta N_H)^2 + (\Delta H_N)^2]^{1/2}$). However, we did observe that 2 amino acid residues of RAP74₄₃₆₋₅₁₇ displayed intermediate chemical shift changes ($0.07 > \Delta\delta > 0.04$; $\Delta\delta = [(0.17\Delta N_H)^2 + (\Delta H_N)^2]^{1/2}$) when comparing the two complexes, and these residues correspond to K476 and K480 of RAP74₄₃₆₋₅₁₇. Based on

these results, we postulate that S942PO₄ makes an ion pair with K475 of RAP74 while S944PO₄ makes an ion pair with K480 of RAP74.

Ion pairs in the RAP74₄₃₆₋₅₁₇ / FCP1₈₇₉₋₉₆₁ complex.

We then modeled the K475/ S942PO₄, and the K480/ S944PO₄ ion pairs within our NMR structure of the RAP74₄₃₆₋₅₁₇ / FCP1₈₇₉₋₉₆₁ complex. Analysis of our NMR structure of the RAP74₄₃₆₋₅₁₇ / FCP1₈₇₉₋₉₆₁ complex verified that K475 was the closest lysine residue of RAP74 to S942 of FCP1 at a distance of 13 Å, whereas K480 is the closest lysine of RAP74 to S944 of FCP1 at 18 Å. Although these distances are too large for an ion pair, the S942 and S944 are located in an extremely flexible portion of FCP1 and K480 is located in the highly flexible L2 loop of RAP74₄₃₆₋₅₁₇. Thus, it is highly conceivable that there is enough mobility to make ion pairs with the phosphorylated serines, and K475 is the closest positive ion to S942 and K480 is the closest positive ion to S944. This analysis supports our chemical shift mapping with the phosphorylated peptides of FCP1 binding to RAP74. To demonstrate that these two ion pairs are in fact possible, we modeled them into the existing NMR structure of the RAP74₄₃₆₋₅₁₇ / FCP1₈₇₉₋₉₆₁ complex. In the modeling calculations, S942 and S944 were replaced with S942PO₄ and S944PO₄. In addition, we included a distance constraint of 2.5 Å between the side chain HZ of K475 and the oxygen attached to phosphorus of S942PO₄, and a distance constraint of 2.5 Å between the side chain HZ of K480 and the oxygen attached to phosphorus of S944PO₄. In the new structure calculations, the two new distant constraints were not violated, and the resulting structure demonstrated the feasibility of the K475/ S942PO₄, and the K480/ S944PO₄ ion pairs (Figure 9).

Mapping the binding site of T584PO₄ on the structure of RAP74₄₃₆₋₅₁₇.

In order to map the interaction site for the phosphorylated threonine in the central domain of FCP1 (T584PO₄) on the solution structure of RAP74₄₃₆₋₅₁₇, we computed the differences in ¹H and ¹⁵N chemical shifts of the RAP74₄₃₆₋₅₁₇ / FCP1₅₇₉₋₆₀₀ complex and the RAP74₄₃₆₋₅₁₇ / FCP1₅₇₉₋₆₀₀ (T584PO₄) complex. This analysis reveals that 2 out of 82 amino acid residues of RAP74₄₃₆₋₅₁₇ displayed significant chemical shift changes ($\Delta\delta > 0.07$; $\Delta\delta = [(0.17\Delta N_H)^2 + (\Delta H_N)^2]^{1/2}$) upon formation of the complex. These residues correspond to positions T470 and K498 of RAP74₄₃₆₋₅₁₇. In addition, we observe 3 other amino acid residues of RAP74₄₃₆₋₅₁₇ that displayed intermediate, but less dramatic chemical shift changes $0.07 > \Delta\delta > 0.04$; $\Delta\delta = [(0.17\Delta N_H)^2 + (\Delta H_N)^2]^{1/2}$ upon formation of the complex. These residues correspond to K471, Q495 and I496 of RAP74₄₃₆₋₅₁₇. When mapped onto the NMR solution structure of RAP74₄₃₆₋₅₁₇, we see that T470 and K471 are located at the beginning of the H2 helix, and residues Q495, I496 and K498 are located at the end of the H3 helix (Figure 8B). These two points are located on the opposite side of the hydrophobic groove. Based on these results and analysis of the NMR structure of the RAP74₄₃₆₋₅₁₇ / FCP1₈₇₉₋₉₆₁ complex, we postulate that T584PO₄ could either make an ion pair with K471, a hydrogen bond with T470, or both. Although the chemical shift data suggests a possible involvement of K498 this appears unlikely since based on homology with the structure of the RAP74₄₃₆₋₅₁₇ / FCP1₈₇₉₋₉₆₁ complex. The chemical shift changes at Q495, I496 and K498 are most likely the result of a rearrangement of the hydrophobic groove between the H2 and H3 helices of RAP74₄₃₆₋₅₁₇ to accommodate the aromatic sidechains present in the FCP1₅₇₈₋₆₀₀ (T584PO₄) peptide (Y592). There are no aromatic side chains in FCP1₉₃₇₋₉₆₁, and this makes interpretation of these results more difficult. We are currently determining the

NMR structure of the RAP74₄₃₆₋₅₁₇ / FCP1₅₇₉₋₆₀₀T584PO₄ complex to better define the role of the T584PO₄.

DISCUSSION

In this manuscript, we examined the role of phosphorylation of human FCP1 by CK2 *in vivo* and *in vitro*. FCP1 contains numerous potential CK2 phosphorylation sites, and previous studies have clearly shown that CK2 is capable of phosphorylating FCP1 *in vitro* (34,35). In the first study, it was shown that xFCP1 exists as a phosphoprotein and co-purifies with CK2 (35). It was also demonstrated that phosphorylation of xFCP1 by CK2 enhances binding to RAP74, and this leads to a stimulation of FCP1 CTD phosphatase activity. Furthermore, a serine to alanine mutation at residue 457 in xFCP1 lowers overall phosphorylation of xFCP1 and leads to a significant inhibition of CK2 enhanced FCP1 CTD phosphatase activity. Although these studies established a clear role of S457 in CK2 enhancement of FCP1 CTD phosphatase activity, the CK2 phosphorylation site in FCP1 responsible for the increased RAP74 binding was not identified. Furthermore, S457 is not contained within either of the two previously identified RAP74 binding sites of FCP1, and therefore it is unlikely that it would enhance binding of RAP74 to FCP1. In the second study, human FCP1 isolated from baculovirus was also shown to exist in a phosphorylated state, and CK2 was isolated as an FCP1-specific kinase from HeLa cell nuclear extracts (34). In addition, TFIIF (RAP30 / RAP74) was shown to be capable of activating the CTD phosphatase activity of the phosphorylated form of FCP1 but not the unphosphorylated form, and the phosphorylated form of FCP1 was shown to be more active in stimulating transcription elongation reactions than the unphosphorylated form. Interestingly, elongation reactions conducted in the presence of CK2 were inhibited. This inhibitory effect was

FCP1-specific, since CK2 added to elongation reactions in the absence of FCP1 displayed no such inhibitory effect. In addition, human FCP1 isolated from the baculovirus system was shown to be phosphorylated at S575 and S740 by mass spectral analysis. However, it was not determined if CK2 or an alternative kinase were responsible for phosphorylating these serine residues, and it is not known if phosphorylation of these sites has an important role in either enhancing FCP1 CTD phosphatase activity or enhancing FCP1 binding to RAP74.

T584 is in the central domain of FCP1, and is located within the FCP1 binding site for both Tat and for RAP74. Phosphorylation of T584 could potentially play an important role in regulating FCP1-activities associated with RAP74 and Tat interactions. T584 is surrounded by acidic amino acids, with acidic amino acids located at n-4, n-3, n-2, n-1, n+1, n+2, n+3 and n+4. Thus, it is almost an ideal site for CK2 phosphorylation. We have demonstrated that CK2 is capable of phosphorylating FCP1 specifically at residue T584 *in vitro* and that T584 is the predominant CK2 phosphorylation site within the central domain of GST-FCP1₅₆₂₋₆₁₉. In addition, phosphorylation of T584 by CK2 leads to enhanced binding of FCP1₅₆₂₋₆₁₉ to RAP74₄₃₆₋₅₁₇. Preliminary NMR studies with a RAP74₄₃₆₋₅₁₇ / FCP1₅₇₉₋₆₀₀ (T584PO₄) complex indicate that this enhanced binding appears to be the result of T584PO₄ making a salt bridge and/or hydrogen bond with the H2 helix of RAP74₄₃₆₋₅₁₇. The hydrogen bond would form with residue T470 of RAP74₄₃₆₋₅₁₇, whereas the salt bridge would be with K471 of RAP74₄₃₆₋₅₁₇. Our previous NMR structure of the FCP1₈₇₉₋₉₆₁ / RAP74₄₃₆₋₅₁₇ complex contains a motif in the last 22 amino acids, that resembles the motif found in FCP1₅₇₉₋₆₀₀. In fact, it is possible to align the LXXLL-like hydrophobic motif found in FCP1₅₇₉₋₆₀₀ (L593-L597), with the LXXLL-like hydrophobic motif found in FCP1₈₇₉₋₉₆₁ (L957-M961). When the two motifs are aligned, D583 aligns to D947 and T584PO₄ aligns to E948. In the NMR structure of the FCP1₈₇₉₋₉₆₁ /

RAP74₄₃₆₋₅₁₇ complex, D947 of FCP1 makes a salt bridge with K471 of RAP74, E954 of FCP1 makes a hydrogen bond with T470 of RAP74, whereas E948 of FCP1 is exposed to the solvent. Therefore, the binding of FCP1₅₇₉₋₆₀₀ (T584PO₄) to RAP74 is not identical to what is seen in the FCP1₈₇₉₋₉₆₁ / RAP74₄₃₆₋₅₁₇ complex. Our hypothesis is that the helix formed by FCP1₅₇₉₋₆₀₀ (T584PO₄) when complex with RAP74₄₃₆₋₅₁₇ is shorter and more flexible at the amino-terminus than the helix formed by FCP1 in the FCP1₈₇₉₋₉₆₁ / RAP74₄₃₆₋₅₁₇ complex. This would allow for T584PO₄ to make the salt bridge with K471. In addition, the repositioning of the ion pair may be required to allow the bulkier hydrophobic aromatic residue (Y592) present in FCP1₅₇₉₋₆₀₀ (T584PO₄) to be positioned in the hydrophobic groove of RAP74₄₃₆₋₅₁₇. There are no aromatic amino acids in FCP1₈₇₉₋₉₆₁.

Based on these studies, it appears that S575 is not a major phosphorylation target for CK2 *in vitro* and it is difficult to explain how phosphorylation at this site could result in enhanced binding to RAP74. The fact that this residue is phosphorylated when FCP1 is expressed in baculovirus is clear (34). However it does not appear that this phosphorylation is carried out by CK2 unless, the *in vivo* conditions in baculovirus with the full length FCP1 are significantly different than our *in vitro* conditions with FCP1₅₆₂₋₆₁₉. S575 is surrounded by acidic residues at n-2, n+2, n+3, n+4, n+5, n+6, n+7 and there is the crucial acidic residue at the n+3 position. Therefore under the simplest definition, S575 meets the minimal requirements to be considered a potential CK2 site. However, it is also clear from our results where we fail to see phosphorylation of the FCP1₉₄₁₋₉₆₁ and FCP1₅₈₄₋₆₀₇ peptides (Figure 5A), that other factors are very important for the CK2 phosphorylation site selection in addition to acidic residues at position n+3.

S942, S943, and S944 are located in the carboxyl-terminal domain of FCP1, and this domain has been previously shown to be important for interaction with RAP74 and TFIIB (30,33). We have demonstrated that S942, S943, and S944 can be phosphorylated by CK2 *in vitro*, and that phosphorylation of S942 and S944 is crucial for the enhanced binding of FCP1₈₇₉₋₉₆₁ to RAP74₄₃₆₋₅₁₇ following phosphorylation with CK2. Mass spectral analysis using a FCP1₉₃₇₋₉₆₁ peptide revealed that the three phosphorylations occur in a highly ordered fashion, and we do not observe a mixed phosphorylation pattern. Furthermore, NMR analysis clearly indicates that when FCP1₈₇₉₋₉₆₁ is complexed to RAP74₄₃₆₋₅₁₇, both S942PO₄ and S944PO₄ are in position to interact with lysine residues in RAP74₄₃₆₋₅₁₇. Based on changes in chemical shift mapping, analysis of the NMR structure of the RAP74₄₃₆₋₅₁₇ / FCP1₈₇₉₋₉₆₁ complex, and modeling, we postulate that S942PO₄ makes a salt bridge with K475 of RAP74₄₃₆₋₅₁₇, and S944PO₄ makes a salt bridge with K480 of RAP74₄₃₆₋₅₁₇. It should be emphasized that the region of RAP74 is rich in lysine residues (K475, K476, K480, K481) and the loop between the H2 and H3 helices is highly flexible. Thus, it is highly possible that all four lysines are important for the enhanced binding of FCP1 following phosphorylation either through formation of direct ion pairs or simply through electrostatic effects.

Although a detailed kinetic analysis was not performed, it appears that following the initial phosphorylation at S942, there is a rapid phosphorylation at S944, and finally slower phosphorylation at S943. Successive phosphorylations have been observed in several other CK2 sites containing multiple serine residues, but there is currently no detailed analysis of the order of phosphorylation in these multi-phosphorylated sites (53). In the case of the carboxyl-terminal domain of FCP1, it is still not completely clear as to what determines this highly ordered sequence of phosphorylation reactions. If acidic residues were the only factors used to rank

these three sites in order of preference for CK2 phosphorylation (53), S944 would be considered the preferred site since it has acidic residues at n+1 and n+3, in addition to acidic residues at n-4 and n+4. S942 would be the considered the second best site, since this site has an acidic residue at n+3 in addition to acidic residues at n-4, n-3, n-2, n+5 and n+6. S943 would be the least likely site to be phosphorylated by CK2, since it lacks an acidic residue at n+3, but has acidic residues at n-4, n-3, n+2, n+4, and n+5. Based on these results, it appears that factors in addition to acidic residues at n+1, n+2 and n+3 are important. As mentioned above, we are currently unaware of another example of three adjacent serine residues being phosphorylated by CK2 in such a highly ordered fashion. It does appear that acidic residues located on the amino-terminal sides of the three serine residues may be crucial for determining the substrate specificity. If the three acidic amino acids located on the amino-terminal side of the S942-S944 site are removed, the peptide FCP1₉₄₁₋₉₆₁ is no longer a substrate for CK2 *in vitro*. Additional experiments are needed to precisely define the specificity, and to explain why the three phosphorylations occur in such a highly-ordered fashion.

In most cases where multiple phosphorylation occur in a defined segment of a protein, it has been hypothesized that the phosphorylated protein actually serves as a better substrate for subsequent phosphorylations by CK2 as a result of the presence of the additional acidic phosphate group (53,55). This appears to be partially the case for S944. This residue is not phosphorylated in the FCP1₉₄₁₋₉₆₁ peptide and it is phosphorylated in the FCP1₉₃₇₋₉₆₁ peptide after the initial phosphorylation of S942. . However, in the S942A mutant we still see phosphorylation and we assume that this phosphorylation is occurring at S944. So it does appear to be a CK2 site, but S942 is preferred over S944.

An important example of multiple serine phosphorylation by CK2 has been observed in the carboxyl-terminal PEST domain of the transcriptional regulatory protein IκBβ(56-58). This domain contains a stretch of seven amino acids consisting of five serine and two aspartic acid residues (³¹²SSSSDSD³¹⁸). Following UV irradiation, IκBβ is heavily phosphorylated by CK2 at a minimum of three serine residues in the PEST domain and these multiple phosphorylation are critical for IκB control of UV-induced NF-κB activation (56). It has been clearly demonstrated that S313 and S315 are two of the three phosphorylation sites, and that the initial phosphorylation is essential for the subsequent phosphorylation (58). However, there have been no detailed studies conducted to determine if the S313 and S315 sites are phosphorylated in a random fashion, or in a highly ordered fashion as we have observed with FCP1. Our detailed mass spectral analysis of the CK2 phosphorylation of FCP1₉₃₇₋₉₆₁ following various reaction times allowed us to clearly establish that FCP1 was phosphorylated at all three sites in a highly ordered fashion *in vitro*.

Post-translational changes such as phosphorylation can cause dramatic changes in protein conformation or very distinct localized changes in a ligand-binding site. We have demonstrated that both RAP74-binding sites of FCP1 contain conserved CK2 sites that when phosphorylated lead to enhanced binding. The exact biological reason for two RAP74-binding sites in FCP1 is still unknown. The two sites may recruit RAP74 at temporally distinct events within the cell for distinct purposes. The additive effects of multiple phosphorylations within the binding site could serve to strengthen such a macromolecular complex. The NMR results show that phosphorylation within both the RAP74-binding site of FCP1 leads to distinct interactions with RAP74₄₃₆₋₅₁₇, most likely involving the formation of ion pairs and/or electrostatic effects. CK2 phosphorylation of FCP1 may provide a selective advantage for binding RAP74 over other

potential targets, and this may vary under different conditions. Interestingly, we have shown in an accompanying paper that Tat can both block CK2 phosphorylation of T584, and inhibit binding of RAP74 to the central domain of FCP1 (Abbott et al. unpublished). This enhanced binding of FCP1 to RAP74 following phosphorylation is believed to stimulate FCP1 phosphatase activity contributing to CTD dephosphorylation. We have made initial attempts to determine if T584, S942, S943 and S943 are phosphorylated *in vivo* within human cells. In our preliminary MS studies using FCP1 isolated from human cells, we were not able identify peptides from the central domain containing T584, or peptides from the carboxyl-terminal domain containing S942-S944 following proteolytic digestion. We believe that our failure to detect phosphorylation at these sites *in vivo* is a technical problem since we were unable to locate peptides containing these sites that were in the unphosphorylated form. However, it is certainly possible that these sites are not phosphorylated to a significant extent under normal cellular conditions, but may be heavily phosphorylated following insult such as DNA damage as is the case with IκB (56). Future studies are needed to address the *in vivo* physiological consequences and possible regulation of FCP1 by CK2 as well as other kinases.

ACKNOWLEDGEMENTS.

We thank Dr. Jack Greenblatt for the FCP1 antibody, Dr. Claiborne Glover, Dr. Mark Emmett, Dr. Christopher Hendrickson and Dr. John Quinn for valuable discussions. This work was supported by the National Institutes of Health Grant RO1 GM60298-01(to J.G.O and P.L.), and by NSF Grant CHE-99-09502 (AGM), Florida State University, and the National High Magnetic Field Laboratory in Tallahassee, FL.

REFERENCES

1. Corden, J. L., Cadena, D. L., Ahearn, J. M., and Dahmus, M. E. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 7934-7938
2. Corden, J. L. (1990) *Trends Biochem. Sci.* **15**, 383-387
3. Bensaude, O., Bonnet, F., Casse, C., Dubois, M. F., Nguyen, V. T., and Palancade, P. (1999) *Biochem. Cell Biol.* **77**(4), 249-255
4. West, M. L., and Corden, J. L. (1995) *Genetics* **140**, 1223-1233
5. Lu, H., Flores, O., Weinmann, R., and Reinberg, D. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 10004-10008
6. O'Brien, T. S., Hardin, S., Greenleaf, A., and Lis, J. T. (1994) *Nature* **370**, 75-77
7. Proudfoot, N. J., Furger, A., and Dye, M. J. (2002) *Cell* **108**, 501-512
8. Feaver, W. J., Gileadi, O., Li, Y., and Kornberg, R. D. (1991) *Cell* **67**, 1223-1230
9. Lu, H., Zawel, L., Fisher, L., Egly, J. M., and Reinberg, D. (1992) *Nature* **358**, 641-645
10. Serizawa, H. R., Conaway, R. C., and Conaway, J. W. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 7476-7480
11. Liao, S. M., Zhang, J., Jeffery, D. A., Koleske, A. J., Thompson, C. M., Chao, D. M., Viljoen, M., van Huren, H. J., and Young, R. A. (1995) *Nature* **374**, 193-196
12. Marshall, N. F., Peng, J., Xie, Z., and Price, D. H. (1996) *J. Biol. Chem.* **271**, 27176-27183
13. Yeo, M., Lin, P. S., Dahmus, M. E., and Gill, G. N. (2003) *J. Biol. Chem.* **278**, 26078-26085
14. Chambers, R. S., and Dahmus, M. E. (1994) *J. Biol. Chem.* **269**, 26243-26248

15. Archambault, J., Chambers, R. S., Kobor, M. S., Ho, Y., Cartier, M., Bolotin, D., Andrews, B., Kane, C. M., and Greenblatt, J. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 14300-14305
16. Archambault, J., Pan, G., Dahmus, G. K., Cartier, M., Marshall, N., Zhang, S., Dahmus, M. E., and Greenblatt, J. (1998) *J. Biol. Chem.* **273**, 27593-27601
17. Chambers, R. S., and Dahmus, M. E. (1996) *J. Biol. Chem.* **271**, 24498-24504
18. Cho, H., Kim, T. K., Mancebo, H., Lane, W. S., Flores, O., and Reinberg, D. (1999) *Genes & Dev.* **13**, 1540-1552
19. Kobor, M. S., Archambault, J., Lester, W., Holstege, F. C. P., Gileadi, O., Jansma, D. B., Jennings, E. G., Kouyoumdjian, F., Davidson, A. R., Young, R. A., and Greenblatt, J. (1999) *Molecular Cell* **4**(1), 55-62
20. Schroeder, S. C., Schwer, B., Shuman, S., and Bentley, D. (2000) *Genes & Dev.* **14**(19), 2435-2440
21. Mandal, S. S., Cho, H., Kim, S., Cabane, K., and Reinberg, D. (2002) *Mol. Cell Biol.* **22**(21), 7543-7552
22. Costa, P. J., and Arndt, K. M. (2000) *Genetics* **156**(2), 535-547
23. Lindstrom, D. L., and Hartzog, G. A. (2001) *Genetics* **159**(2), 487-497
24. Cho, E.-J., Kobor, M. S., Kim, M., Greenblatt, J., and Buratowski, S. (2001) *Genes & Dev.* **15**(24), 3319-3329
25. McCracken, S., Fong, N., Yankulov, K., Bothers, G., Siderovski, D., Hessel, A., Forster, S., Program, A. E., Shuman, S., and Bentley, D. L. (1997) *Genes & Dev.* **11**, 3306-3318
26. Cho, E.-J., Takagi, T., Moore, C. R., and Buratowski, S. (1997) *Genes & Dev.* **13**, 3319-3326

27. Licciardo, P., Amente, S., Ruggiero, L., Monti, M., Pucci, P., Lania, L., and Majello, B. (2003) *Nucleic Acids Research* **31**(3), 999-1005
28. Collet, J. F., Stroobant, V., Pirard, M., Delpierre, G., and Schaftingen, E. V. (1998) *J. Biol. Chem.* **273**, 14107-14112
29. Marshall, N. F., Dahmus, G. K., and Dahmus, M. E. (1998) *J. Biol. Chem.* **273**, 31726-31730
30. Chambers, R. S., Wang, B. Q., Burton, Z. F., and Dahmus, M. E. (1995) *J. Biol. Chem.* **270**, 14962-14969
31. Licciardo, P., Ruggiero, L., Lania, L., and Majello, B. (2001) *Nucleic Acids Research* **29**(17), 3539-3545
32. Kobor, M. S., Simon, L. D., Omichinski, J., Zhong, G., Archambault, J., and Greenblatt, J. (2000) *Mol. Cell. Biol.* **20**(20), 7438-7449
33. Nguyen, B. D., Chen, H.-T., Kobor, M. S., Greenblatt, J., Legault, P., and Omichinski, J. G. (2003) *Biochemistry* **42**(6), 1460-1469
34. Friedl, E. M., Lane, W. S., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2003) *Proc. Natl. Acad. Sci. USA* **100**(5), 2328-2333
35. Palancade, B., Dubois, M.-F., and Bensaude, O. (2002) *J. Biol. Chem.* **277**(39), 36061-36067
36. Nguyen, B. D., Abbott, K. L., Potempa, K., Kobor, M. S., Archambault, J., Greenblatt, J., Legault, P., and Omichinski, J. G. (2003) *Proc. Natl. Acad. Sci. USA* **100**(10), 5688-5693
37. Quinn, J. P., Emmett, M. R., and Marshall, A. G. (1998) in *In 46th ASMS Conference on Mass Spectrometry and Allied Topics.*, pp. 1388
38. Molloy, M. P., and Andrews, P. C. (2001) *Anal Chem* **73**, 5387-94

39. Hakansson, K., Chalmers, M. J., Quinn, J. P., McFarland, M. A., Hendrickson, C. L., and Marshall, A. G. (2003) *Anal Chem* **75**, 3256-62
40. Senko, M. W., Canterbury, J. D., Guan, S., and Marshall, A. G. (1996) *Rapid Commun. Mass Spectrom.* **10**, 1839-1844
41. Chowdhury, S. K., Katta, V., and Chait, B. T. (1990) *Rapid Commun. Mass Spectrom.* **4**, 81-87
42. Wilcox, B. E., Hendrickson, C. L., and Marshall, A. G. (2002) *J. Am. Soc. Mass Spectrom.* **13**, 304-1312
43. Comisarow, M. B., and G., M. A. (1974) *Chem. Phys. Lett.* **26**, 489-490
44. Marshall, A. G., and Roe, D. C. (1980) *J. Chem. Phys.* **73**, 1581-1590
45. Marshall, A. G., and Verdun, F. R. (1990) *Fourier Transforms in NMR, Optical, and Mass Spectrometry: A User's Handbook.*, Elsevier, Amsterdam
46. Ledford, E. B., Jr., Rempel, D. L., and Gross, M. L. (1984) *Anal. Chem.* **56**, 2744-2748
47. Hendrickson, C. L., Quinn, J. P., Emmett, M. R., and Marshall, A. G. (2000) in *In 48th ASMS Conference on Mass Spectrometry and Allied Topics*, pp. CD-ROM, Long Beach, CA
48. Marshall, A. G., Wang, T.-C. L., and Ricca, T. L. (1985) *J. Am. Chem. Soc.* **107**, 7893-7897
49. Guan, S., and Marshall, A. G. (1996) *Int. J. Mass Spectrom. Ion Processes* **157/158**, 5-37
50. Kay, L. E., Keifer, P., and Saarinen, T. (1992) *J. Am. Chem. Soc.* **114**, 10663-10665
51. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR* **6**, 277-293
52. Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graphics* **14**, 51-55

53. Meggio, F., and Pinna, L. A. (2003) *FASEB J.* **17**, 349-368
54. Jaffe, H., Veeranna, H., and Pant, H. C. (1998) *Biochemistry* **37**, 16211-16224
55. Litchfield, D. W. (2003) *Biochem. J.* **369**, 1-15
56. Kato, T. J., Delhase, M., Hoffmann, A., and Karin, M. (2003) *Mol. Cell* **1**, 829-839
57. Chu, Z.-L., McKinsey, T. A., Liu, L., Qi, X., and Ballard, D. W. (1996) *Mol. Cell. Biol.* **16**, 5974-5984
58. McKinsey, T. A., Zhi-Liang, C., and Ballard, D. W. (1997) *J. Biol. Chem.* **272**, 22377-22380

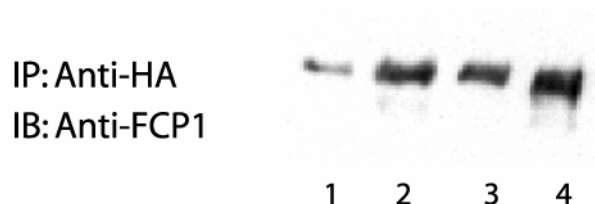


Figure 3.1. FCP1 isolated from human cells exist as a phosphoprotein. HEK(293) cell lysates from cells transfected with HA-FCP1₁₋₉₆₁ were immunoprecipitated with an anti-HA monoclonal antibody (lanes 2-4). Lane 1 represents 20% of the lysate input. Immunoprecipitated proteins were captured on protein A/G conjugated agarose resin and subjected to no further treatment (lane 2), phosphorylation in the presence of recombinant CK2 and ATP (lane 3), or alkaline phosphatase treatment (lane 4). Proteins were resolved on a 7% SDS polyacrylamide gel, and FCP1 was detected by immunoblotting using an FCP1-specific polyclonal antibody (16).

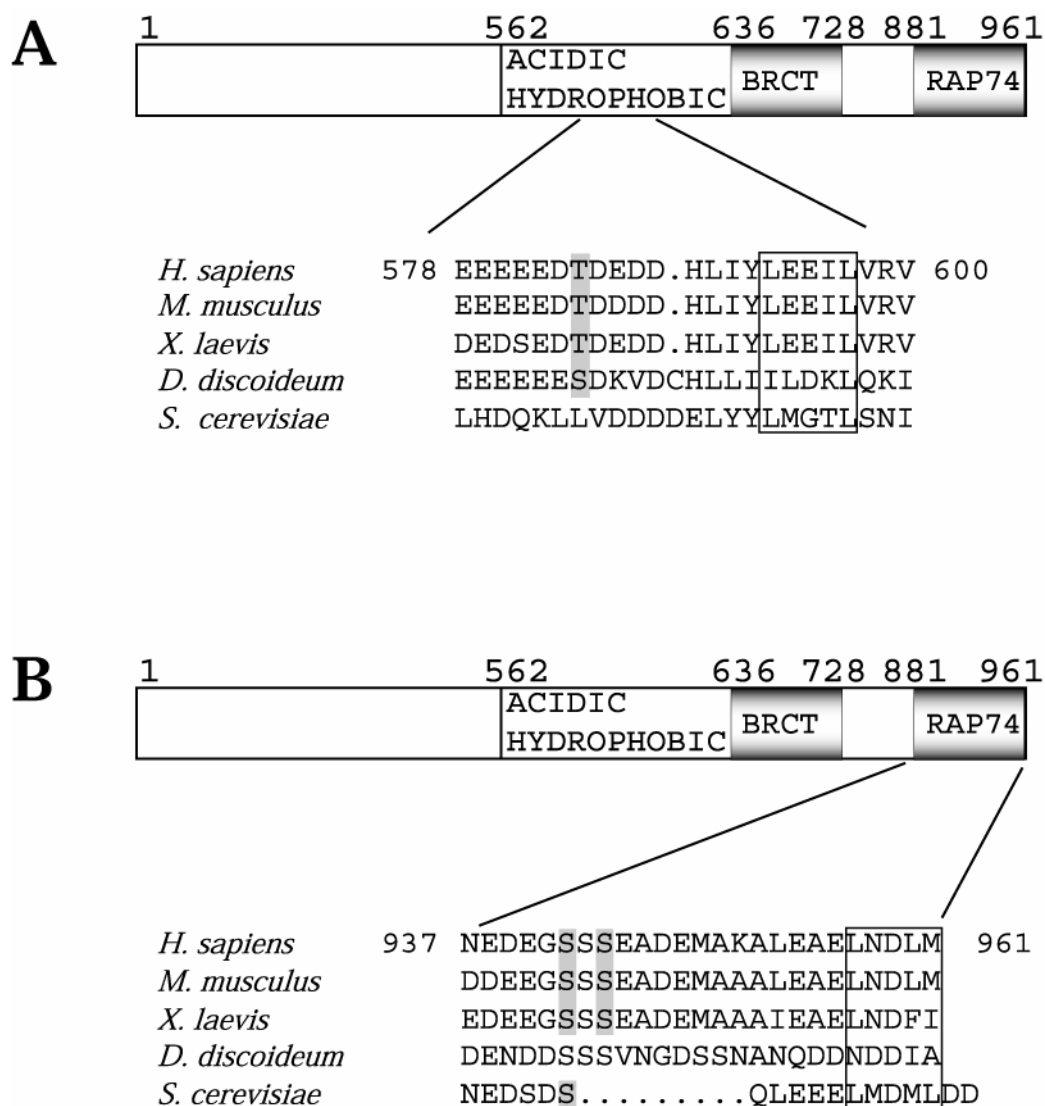


Figure 3.2. FCP1 has two RAP74-binding sites with adjacent conserved CK2 phosphorylation sites. Sequences from the (A) central regulatory domain and (B) carboxyl-terminal regulatory domain of FCP1 from *Homo sapiens* (NP_004706), *Mus musculus* (AAH53435), *Xenopus laevis* (Q98SN2), *Dictyostelium discoideum* (AAM33170), and *Saccharomyces cerevisiae* (NP_014004) were aligned using BLAST (Altschul et al. 1990) followed by manual refinement. The conserved small hydrophobic motif required for RAP74 binding is boxed, and the conserved predicted sites of CK2 phosphorylation based on presence of an acidic residue at the n+3 position are highlighted (53).

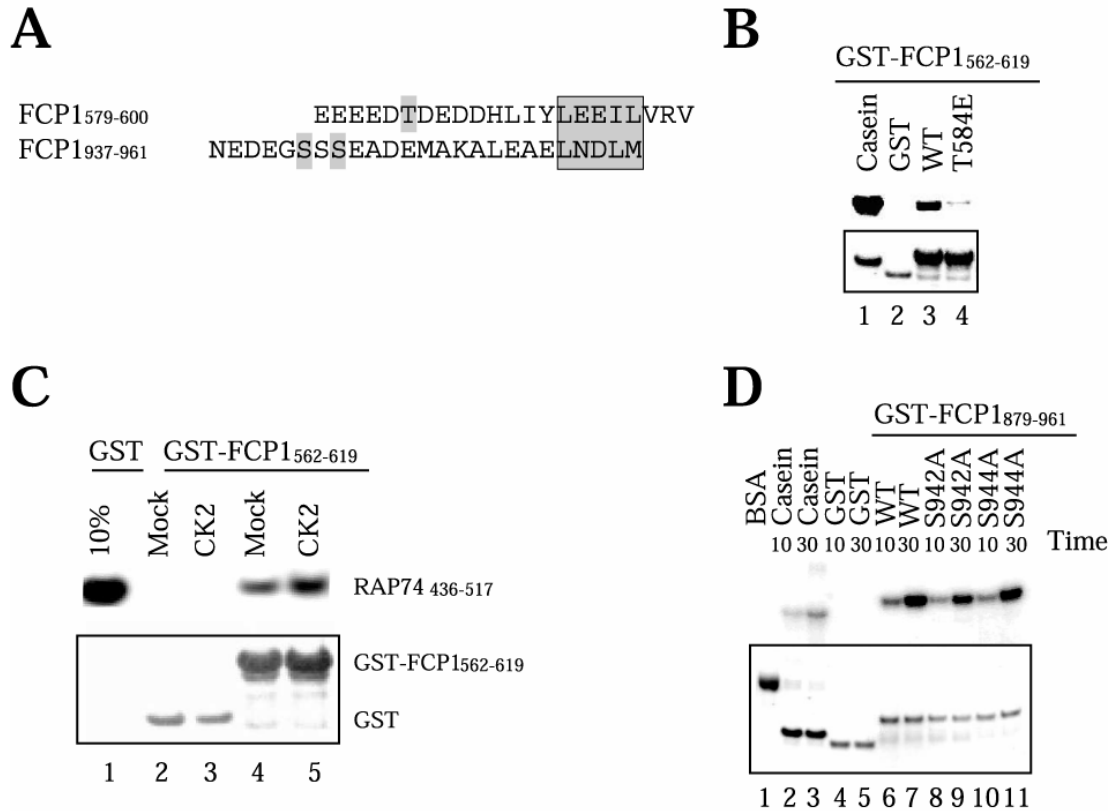


Figure 3.3 Identification of CK2 phosphorylation site located within the FCP1 central and carboxyl-terminal domains. (A) Sequences of peptides FCP1⁵⁷⁹⁻⁵⁹⁸ and FCP1⁹³⁷⁻⁹⁶¹ showing the location of the CK2 sites (shaded) and small hydrophobic motifs (boxed and shaded). The predicted sites of phosphorylation are T584, S942 and S944. (B) Phosphorylation of 1) Casein 2) GST 3) GST-FCP1⁵⁶²⁻⁶¹⁹ (WT) and 4) GST-FCP1⁵⁶²⁻⁶¹⁹T584E (T584E) mutant in the presence of [γ -³²P]-ATP (4500 Ci/mmol) and recombinant CK2 (New England Biolabs, MA). Each reaction contained approximately 25 μ g protein input; the ³²P-labeled proteins were detected by PhosphorImager analysis (Molecular Dynamics-Amersham Pharmacia, NJ). (C) Captured GST (lanes 2-3) or GST-FCP1⁵⁶²⁻⁶¹⁹ (lanes 4-5) proteins were mock-phosphorylated (all components except enzyme; lanes 2 and 4) or CK2-phosphorylated (lanes 3 and 5), as described in materials and methods. Each binding reaction contained GST or GST-FCP1⁵⁶²⁻⁶¹⁹ each at (2 μ M) and RAP74₄₃₆₋₅₁₇ (10 μ M). The bound RAP74₄₃₆₋₅₁₇ protein was detected using the RAP74 polyclonal antibody (C-18). 10% input of RAP74₄₃₆₋₅₁₇ is shown as a control (Lane 1). (D) Phosphorylation of Casein (lanes 2-3), GST (Lanes 4-5), GST-FCP1⁸⁷⁹⁻⁹⁶¹ (WT) (lanes 6-7), GST-FCP1⁸⁷⁹⁻⁹⁶¹ (S942A) (lanes 8-9), or GST-FCP1⁸⁷⁹⁻⁹⁶¹ (S944A) (lanes 10-11) for 10 minutes or 30 minutes in the presence of [γ -³²P]-ATP (4500 Ci/mmol) and recombinant CK2 (New England Biolabs, MA). Each reaction contained approximately 25 μ g protein input; the ³²P-labeled proteins were detected by PhosphorImager analysis (Molecular Dynamics-Amersham Pharmacia, NJ). 10% input of RAP74₄₃₆₋₅₁₇ is shown as a control (Lane 1). For (B), (C) and (D) equivalent input of GST and GST-FCP1 fusion protein was verified by performing a Coomassie blue stain of the membrane (boxed).

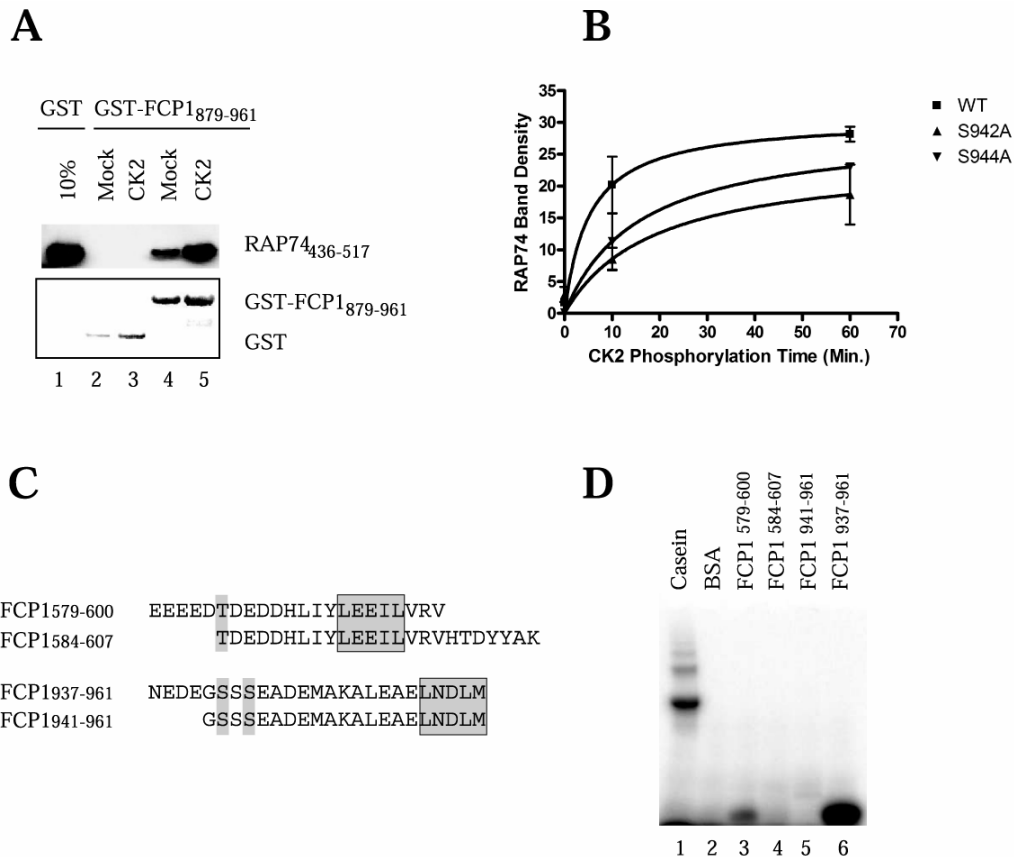


Figure 3.4 Identification of a multiple serine CK2 site located within the carboxyl-terminus of FCP1 that is phosphorylated by CK2 *in vitro* and mediates enhanced RAP74 binding.

(A) Captured GST (lanes 2-3) and GST-FCP1₈₇₉₋₉₆₁ were mock-phosphorylated (all components except enzyme; Lanes 2 and 4) and CK2-phosphorylated (lanes 3 and 5) as described in materials and methods. Each binding reaction contained (2 μ M) GST or GST-FCP1₈₇₉₋₉₆₁ and purified RAP74₄₃₆₋₅₁₇ (10 μ M). The bound RAP74₄₃₆₋₅₁₇ protein was detected using the RAP74 polyclonal antibody (C-18). 10% input of RAP74₄₃₆₋₅₁₇ is shown as a control (Lane 1). Equivalent input of GST and GST-FCP1 fusion protein was verified by performing a Coomassie blue stain of the membrane (boxed). (B) Graphical representation of the level of RAP74 binding after 0, 10, and 60 min of CK2 phosphorylation reactions with wild-type GST-FCP1₈₇₉₋₉₆₁, GST-FCP1₈₇₉₋₉₆₁ (S942A), and GST-FCP1₈₇₉₋₉₆₁ (S944A). CK2 phosphorylation reactions were stopped by the addition of 100 mM EDTA followed by washing in binding buffer 2-3 times. Concentrations of binding reactions are the same as described for A. RAP74 binding is expressed as measured band density on ECL Hyperfilm (Amersham Biosciences, NJ) scanned on a Bio-Rad Fluor S Multi-Imager (Bio-Rad, CA). Data shown is averaged data from three separate experiments with error bars. (C) The synthetic peptides FCP1₅₇₉₋₅₉₈, FCP1₅₈₄₋₆₀₄, FCP1₉₄₁₋₉₆₁, and FCP1₉₃₇₋₉₆₁ used for *in vitro* CK2 phosphorylation reactions. (D) Reactions containing 25 μ g of protein input were incubated for 30 min in the presence of [γ -³²P]-ATP (4500 Ci/mMol) and recombinant CK2 (New England Biolabs). ³²P-labeled proteins were detected by PhosphorImager analysis (Molecular Dynamics- Amersham Pharmacia, NJ).

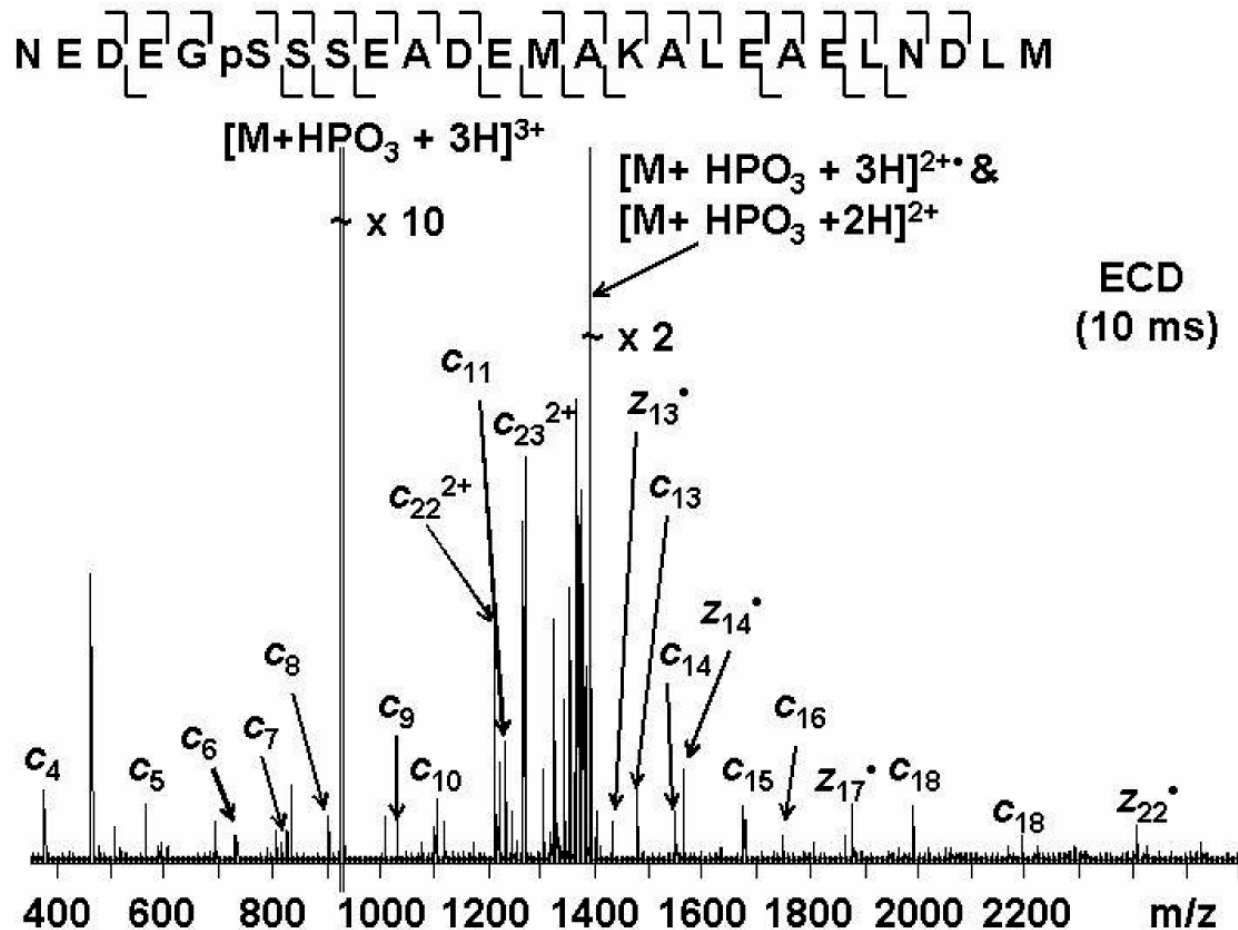


Figure 3.5. FT-ICR MS/MS spectrum of FCP1₉₃₇₋₉₆₁ synthetic peptide after 4 h CK2 incubation. Product ion spectrum obtained from ESI ECD FT-ICR MS/MS of a population of quadrupole- and SWIFT-isolated $[N_{937}EDEGSSSSEADEMAKALEAELNDLM_{961} + HPO_3 + 3H]^{3+}$ FCP1 peptide ions. Twenty-one (out of twenty-four) peptide backbone bonds are broken. The location of the phosphorylation site is identified by the observation of the $[c_6 + HPO_3 + H]^+$ ions, as well as the $[z_{19} + H]^+$ and $[z_{22} + HPO_3 + H]^+$ ions.

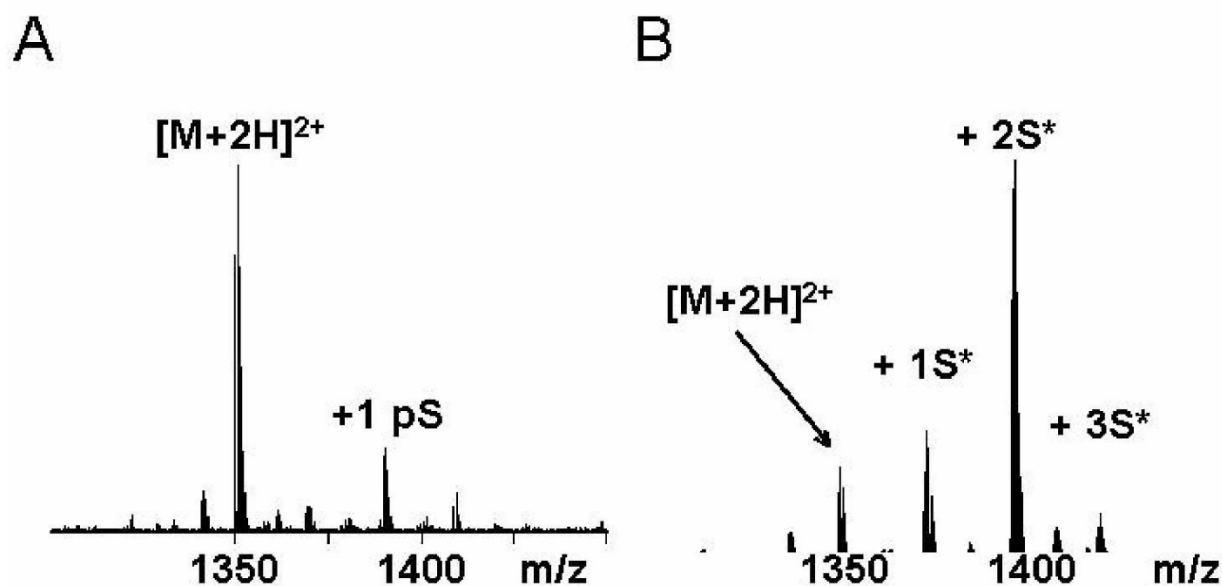


Figure 3.6 β -elimination / ethanethiol modification of FCP1 peptide increases the positive ion mode ESI response factor of multiply phosphorylated FCP1 ions. (A) Expansion of the positive ion ESI FT-ICR MS spectrum of the phosphorylated (incubation with CK2 and ATP 4 h) FCP1 fragment. $[M + 2H]^{2+}$ and $[M + HPO_3 + 2H]^{2+}$ ions were detected with good signal / noise ratios, however no ions corresponding to di- or tri- phosphorylated FCP1 were detected. (B) Positive ion ESI FT-ICR MS spectrum of an equivalent amount of FCP1 fragment after 30 min β -elimination and Michael addition. S-ethylcysteine modified peptide ions (e.g. + S*) confirm the presence of the di- and tri- phosphorylated forms of the FCP1 peptide.

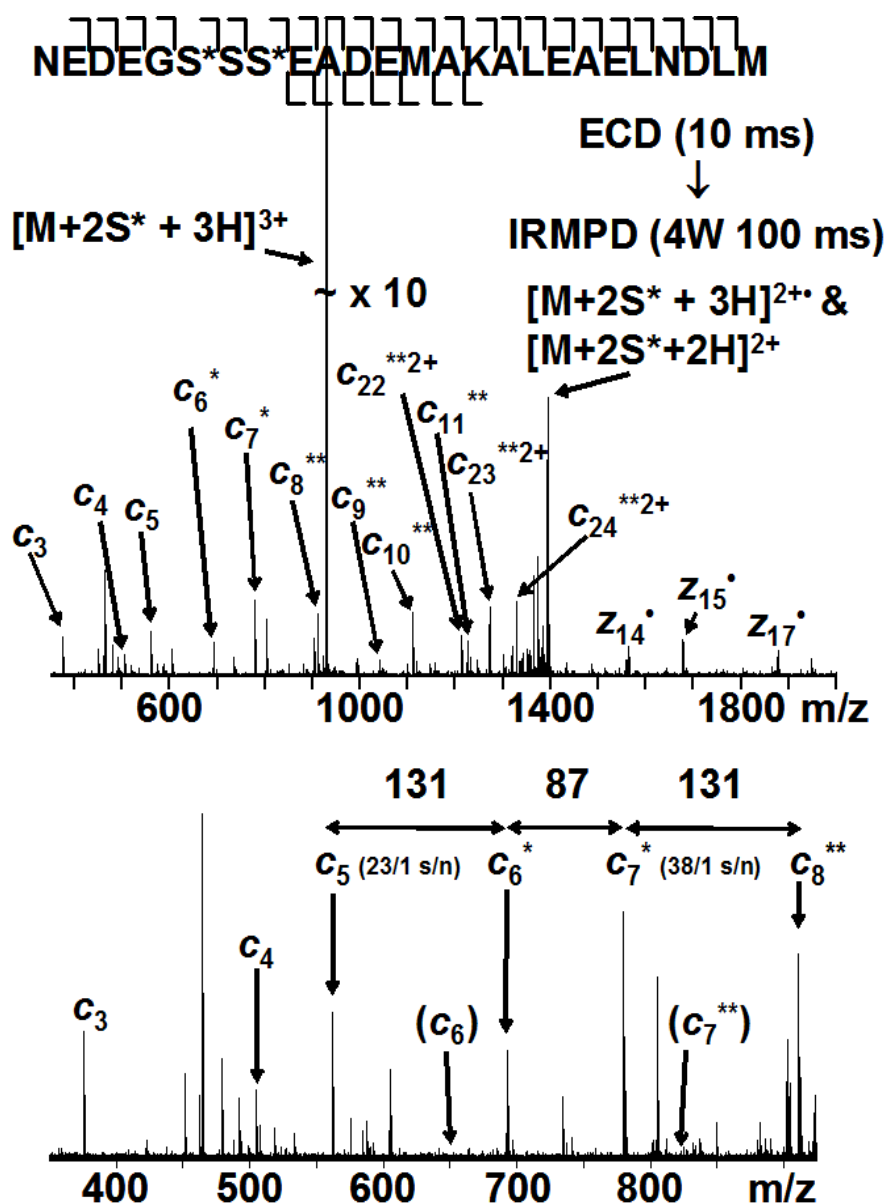


Figure 3.7. ESI FT-ICR MS/MS spectrum of the S-ethylcysteine modified FCP1₉₃₇₋₉₆₁ fragment. Top: Product ion spectrum obtained from the AI-ECD FT-ICR MS/MS of a population of quadrupole- and SWIFT-isolated $[N_{937}EDEGSSSEADEMAKALEAELNDLM_{961} + 2S^* + 3H]^{3+}$ FCP1 ions. Twenty-three (out of twenty-four) peptide backbone bonds are broken. Bottom: Sites of modification are identified by the observation of c_6^* and c_8^{**} ions and the absence of c_6 and c_7^{**} ions (in parentheses).

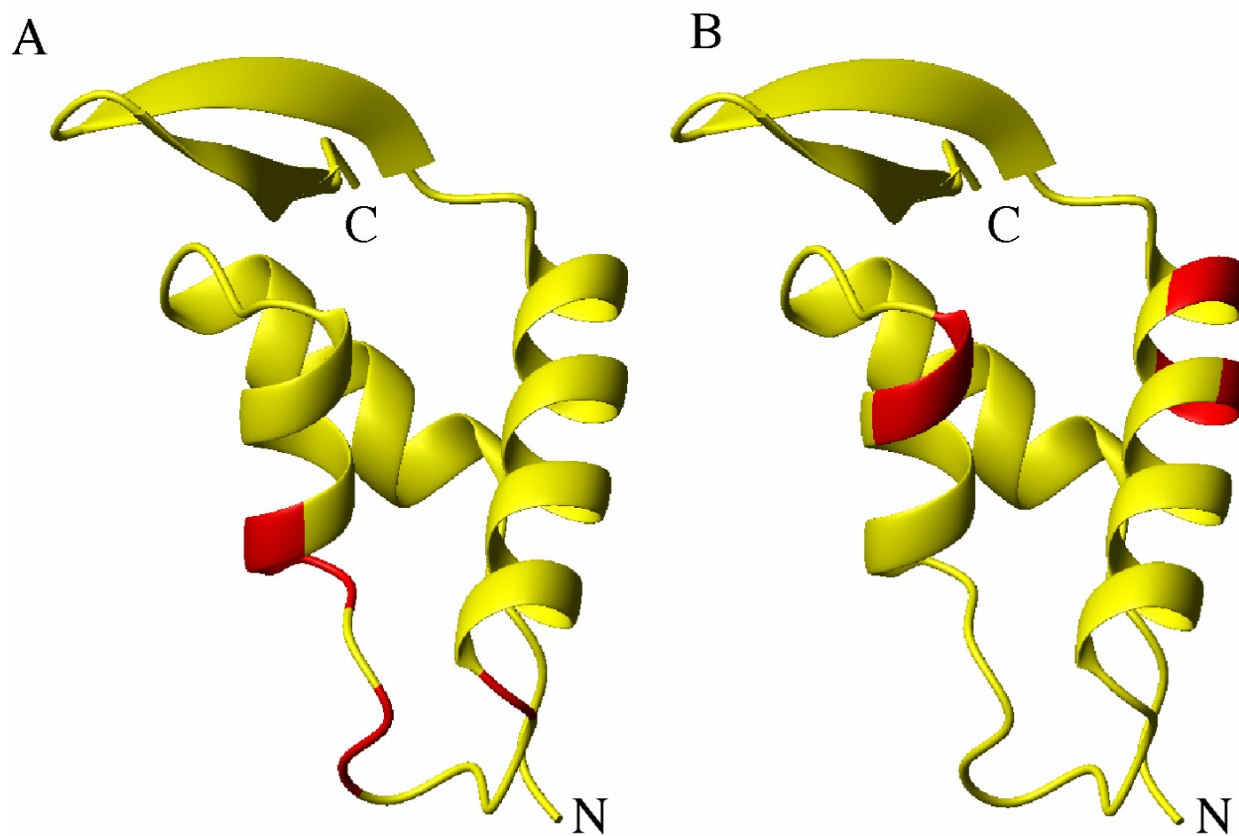


Figure 3.8 Signals which show a significant chemical shift change ($\Delta\delta > 0.07$; $\Delta\delta = [(0.17\Delta N_H)^2 + (\Delta H_N)^2]^{1/2}$) upon comparison of the 2D ^1H - ^{15}N HSQC spectra from (A) the RAP74436-517/FCP1937-961 complex with the RAP74436-517 / FCP1937-961(S942PO₄/S944PO₄) complex and (B) the RAP74436-517 / FCP1579-600 complex with the RAP74436-517 / FCP1579-600(T584PO₄) complex. The signals are mapped onto the NMR structure of RAP74436-517 (yellow) and the locations of significantly shifted signals are indicated (red).

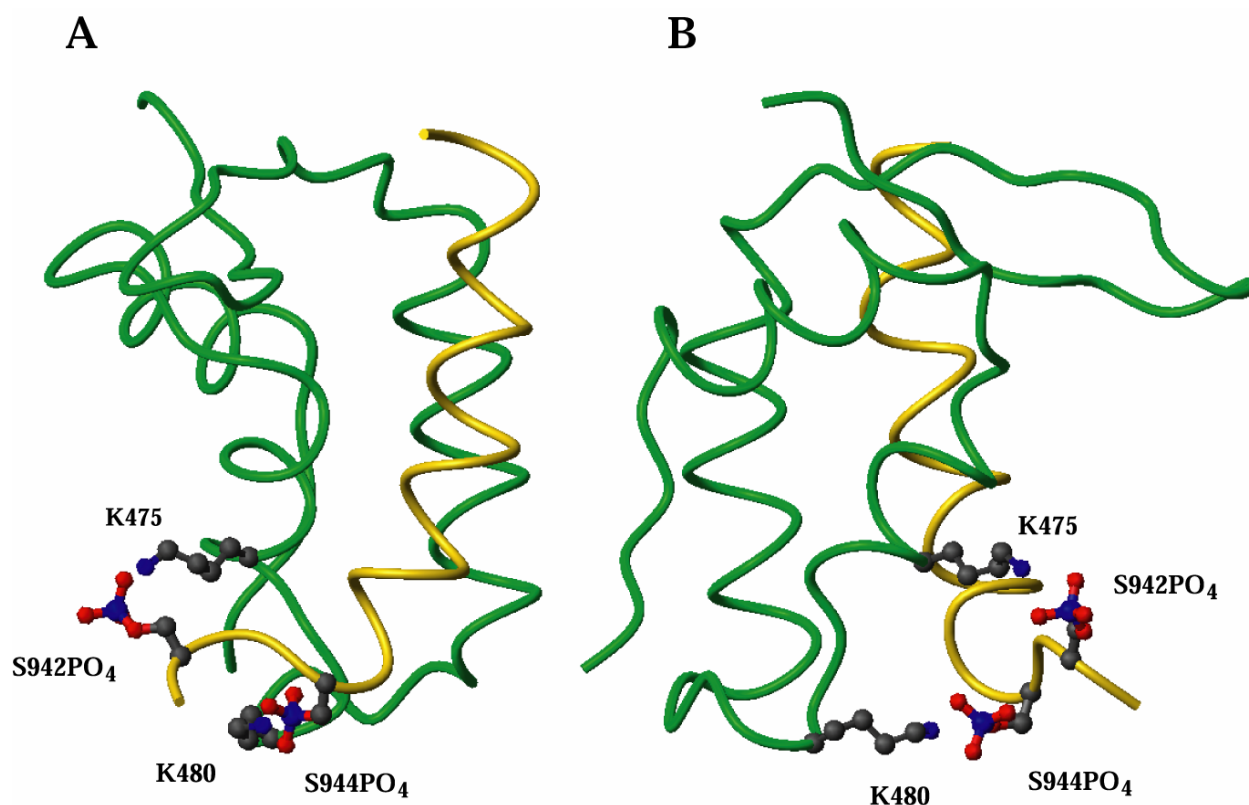


Figure 3.9 Model of the interactions between RAP74₄₃₆₋₅₁₇ and FCP1₉₃₇₋₉₆₁(S942PO₄/S944PO₄). Models were generated using the NMR structure and constraints of the RAP74₄₃₆₋₅₁₇/FCP1₈₈₁₋₉₆₁ complex as a starting point. In the models, the serines at position 942 and 944 of FCP1 were substituted with phosphoserines and constraints were added for the interactions between the lysines of RAP74 and the phosphoserines as described in materials and methods. **(A)** Backbone trace of the RAP74₄₃₆₋₅₁₇ / FCP1₉₃₇₋₉₆₁(S942PO₄/S944PO₄) model complex highlighting the interactions between the lysine amino acids of RAP74₄₃₆₋₅₁₇ (K475 / K480) and the phosphoserine residues of FCP1₉₃₇₋₉₆₁ (S942PO₄ / S944PO₄) at the interface. **(B)** This model is a 90° rotation of A.

CHAPTER 4

NOVEL ENDONUCLEASE ACTIVITY ASSOCIATED WITH THE BRCT DOMAIN OF
THE CTD PHOSPHATASE FCP1³

³Karen L. Abbott, James G. Omichinski, Pascale Legault. To be submitted to Nature.

ABSTRACT

The C-terminal domain (CTD) of RNA polymerase II (RNAP II) has been referred to as a “processing platform”¹. The phosphorylation state of the CTD determines interactions with accessory proteins that influence RNAP II transcription, mRNA processing, and transcription termination events. FCP1, the first identified CTD phosphatase²⁻⁴, is an essential component of the RNAP II transcription complex whose role in transcription is widening⁵⁻⁹. Here we report a previously undetected nucleic acid binding activity and single-strand specific endoribonuclease activity associated with the BRCT-containing central regulatory domain of FCP1. FCP1 specifically cleaves the trinucleotide bulge of the HIV-1 TAR RNA *in vitro*, and HIV-1 Tat can inhibit FCP1 endonuclease activity. The ability of this essential component of the RNAP II holoenzyme to perform endonucleolytic cleavage of single-stranded nucleic acid suggests that FCP1 functions in transcription termination, processing, or repair mechanisms. These results provide new exciting directions for anti-HIV-1 drug discovery.

RESULTS

HIV-1 Tat has been shown to interact directly with the central domain (CD) (562-738) of human FCP1 *in vitro*¹⁰. We wanted to address if FCP1 could form a complex with HIV-1 Tat and the HIV-1 transactivation response element RNA TAR₁₇₋₄₅ (Fig.4.1a). The addition of purified HIV-1 Tat₁₋₈₆ or FCP1₅₆₂₋₇₃₈ to 5'-³²P-labeled TAR₁₇₋₄₅ produced a concentration-dependent shift in TAR RNA (Fig. 4.1b) when resolved on a non-denaturing polyacrylamide gel. To assess whether Tat and FCP1 can bind TAR RNA simultaneously, increasing amounts of purified Tat₁₋₈₆ were added to an FCP1₅₆₂₋₇₃₈/TAR₁₇₋₄₅ complex. Tat was found to effectively

inhibit FCP1 from binding to TAR RNA (Fig. 4.1c). Next, we further investigated the nucleic acid binding activity of FCP1 on various substrates. The FCP1 CD was capable of binding single-stranded RNA and DNA as well as single-stranded regions of structured RNA (S4.1, S4.2). However, FCP1 CD did not bind duplex RNA (S4.3).

During our RNA binding studies we noted the consistent presence of a smaller band seen at the bottom of the native gel that increased in intensity with increasing concentrations of the FCP1 CD. To investigate the possibility of a nucleic acid cleavage activity, we added purified FCP1₅₆₂₋₇₃₈ or baculovirus purified human FCP1₁₋₉₆₁ in increasing concentrations to a 5'-³²P-labeled single-stranded RNA (Fig. 4.2a) designed based on the HIV-1 TAR RNA stem we have termed RNA f. Possible cleavage products were resolved on a denaturing polyacrylamide gel. The results showed that both the purified FCP1 CD and the full-length purified FCP1 could cleave the RNA in a concentration-dependent manner (Fig. 4.2b). This activity observed in the presence of 5 mM MgCl₂, could be inhibited by the addition of 25 mM EDTA, indicating a requirement for divalent metal ions (Fig. 4.2b lanes 5 and 10). Cleavage of single-strand RNA f by FCP1 does not seem to be sequence-specific, as we have tested the complementary strand of RNA f and find that FCP1 cleaves it similarly (S4.1c). Also, post-translational modifications that are present in the baculovirus-isolated human FCP1₁₋₉₆₁ may influence the formation of the active site as evidenced by the differing size of the final cleavage product on unstructured RNA. Another possibility is that amino acids lying outside of the FCP1 CD contribute to nuclease activity.

Next we set out to determine if the BRCT domain located within the FCP1 CD was sufficient for nuclease activity. The BRCT domains of TopBP1, XRCC1, and BRCA1 have previously been reported to participate in DNA end-binding activities and are speculated to be

involved in recognizing DNA breaks^{11,12}. We purified to homogeneity the BRCT domain of FCP1 from the inclusion body as described previously for the FCP CD and tested this domain for nuclease activity. Purified FCP1₆₂₆₋₇₃₈ and purified FCP1₅₆₂₋₇₃₈ were incubated for various times with 5'-³²P-labelled TAR RNA₁₇₋₄₅. FCP1₅₆₂₋₇₃₈ but not FCP1₆₂₆₋₇₃₈ showed a time-dependent cleavage of the RNA (Fig. 4.3a). To our surprise the cleavage of HIV-1 TAR₁₇₋₄₅ was specific to the bulge region. This result was unexpected based on our previous studies using unstructured RNAs indicating that FCP1 cleaves RNA without sequence specificity as the purine rich RNA f and complementary pyrimidine rich RNA g are both cleaved efficiently by FCP1₅₆₂₋₇₃₈ (S4.1C). The failure of FCP1 to produce 5'-labeled cleavage products emanating from the loop region indicates some type of preference for the bulge region of TAR. A band (shown by an asterisk in Fig. 4.3a) that does not align with the hydrolysis ladder or RNase T1 digest accumulates over time. This band may represent a reaction product that accumulates over time such as a conversion from a 2'-3' cyclic phosphate end to a 2' or 3' monophosphate end. To further define the amino acids within FCP1₅₆₂₋₇₃₈ required for nuclease activity, we used GST-FCP1 truncated fragments described previously¹⁰. These experiments (Fig. 4.3b) revealed that GST-FCP1₅₆₂₋₆₉₆ was sufficient for nuclease activity. Our experiments indicate that the region previously described as the acidic/hydrophobic (562-635) region located adjacent to the BRCT domain (636-738)¹⁰ and the amino-terminal portion of the BRCT domain are required for nuclease activity. Several conserved histidine and aspartic acid residues located within the acidic/hydrophobic region and the BRCT domain of FCP1 may be required to coordinate divalent metal ions. Mutation of conserved histidines within the acidic/hydrophobic region inhibits endonuclease activity (S4.4a and S4.4b).

HIV-1 Tat binds directly to the acidic/hydrophobic region and the BRCT domain of FCP1¹⁰. We have already shown that Tat inhibits FCP1 from binding to TAR. However, the mechanism of this competition is unknown. Tat may bind to the RNA binding motif of FCP1 blocking access to RNA, Tat and FCP1 may compete for binding to the bulge region of TAR, or Tat binding may cause a conformational change in FCP1 that inhibits RNA binding. Another unknown question is whether the RNA binding site and the endonuclease active site of FCP1 are the same. To address some of these questions we chose to use the single-stranded RNA substrate RNA f shown in (Fig. 4.2a) since Tat does not bind this RNA. RNA f was 5'-³²P-labeled and incubated with purified FCP1₅₆₂₋₇₃₈ in the presence of increasing concentrations of purified Tat₁₋₈₆. As a control, purified Tat₁₋₈₆ was added in increasing concentrations with 5'-³²P-labeled RNA f. Products were separated on non-denaturing polyacrylamide gels so that any complexes that may form would be visualized as a shift. Results shown in Fig. 4.3b demonstrate that Tat inhibits FCP1 endonuclease activity. From this experiment we can deduce that Tat does not bind with FCP1 when FCP1 binds RNA as no super-shift is visible. Increasing concentrations of Tat inhibit nuclease activity and FCP1 RNA binding activity. Since Tat is not competing with FCP1 for this RNA substrate, it must be binding to FCP1 and resulting in inhibition of RNA binding. Tat either blocks access to the active site or changes the conformation of FCP1 which inhibits RNA binding.

Having identified this novel endonuclease activity within the FCP1 CD, we wanted to know if there was any homology in this region to known nuclease domains. BLAST¹³ searches using FCP1 sequence from 562-696, the minimal amino acid sequence required for nuclease activity, found homology with a known nuclease domain. Shown in Fig. 4.3c a portion of the FCP1 BRCT domain (627-696) contains a region with 37 % identical sequence and 53 %

homologous sequence to the *Caenorhabditis elegans* polyribonucleotide nucleotidyltransferase RNase PH_C domain (496-560). The RNase PH family is a large family of ribonucleases that includes PNPase and members of the exosome¹⁴. Ribonucleases within this family display distinctly diverse modes of action¹⁴.

HIV-1 Tat enhances transcription elongation by promoting the hyperphosphorylation of the RNAP II CTD. Tat stimulates CTD kinases and inhibits FCP1 phosphatase activity. This study illustrates the possibility that the RNase PH_C homology region of FCP1 may be important for the endonuclease activity and the phosphatase activity. Data to support this theory are i) amino acids (562-696) required for nuclease activity are also required for phosphatase activity, ii) Tat inhibits both activities, iii) Mg⁺⁺ is required for both activities. In fact, there is a conserved DxxDR phosphoesterase motif¹⁵ found in several Ser/Thr phosphatases located within the area of FCP1 showing high homology to RNase PH_C domains. FCP1 endonuclease activity may play a role in mRNA 3' end processing and or transcription termination. Therefore, Tat could be enhancing transcription of full-length viral transcripts through an anti-termination mechanism involving FCP1.

METHODS

Protein Purification

FCP1₁₋₉₆₁ was expressed in baculovirus and purified as described previously¹⁶. FCP1₅₆₂₋₇₃₈ and FCP1₆₂₆₋₇₃₈ were expressed as fusion proteins with thioredoxin using the pET32(c) vector (Novagen, WI). Inclusion body pellets were isolated as described previously¹⁰. FCP1 was cleaved from the thioredoxin by thrombin digestion for 2 hrs at RT followed by column chromatography using a Q-sepharose Fast Flow (Amersham Biosciences, NJ) column and a

preparative Superdex 75 gel filtration (Amersham Biosciences, NJ) column. GST-Tat₁₋₈₆ was expressed from pGEX-2T¹⁷ as described previously¹⁰. The supernatant obtained following a 30 min high-speed centrifugation (100,000 x g) was aliquoted, snap-frozen in liquid nitrogen, and stored at -70° C. For each experiment, 1 ml of the lysate was quickly thawed and purified on glutathione sepharose as recommended by the manufacturer (Amersham Biosciences, NJ). Tat₁₋₈₆ was cleaved from GST by thrombin digestion in 1X PBS at room temperature for 2 h. The supernatant containing Tat₁₋₈₆ was incubated in benzamidine sepharose for 30 min and used directly for gel-shift assays.

Substrates and labeling

TAR RNA was transcribed using T7 RNA polymerase prepared in house¹⁸. TAR RNA was heated and snap cooled in 0.5X TEN (20 mM Tris-HCl pH 8.0, 75 mM NaCl, 0.5 mM EDTA) prior to each experiment. RNA f shown in Fig. 2a was purchased from Dharmacon (Boulder, CO). All RNA's were gel purified prior to experiments. RNA substrates were dephosphorylated using calf intestine alkaline phosphatase (NEB) and labeled at the 5' end using T4 polynucleotide kinase (NEB, MA) and [γ -³²P] ATP (4500 Ci/mmol) at 37 °C for 1 h.

Reactions

Gel mobility shift reactions (20 μ l) were carried out in GMSA buffer (30 mM Tris-HCl pH 8.0, 12 % glycerol, 70 mM KCl, 1 mM DTT, 0.05 % NP-40) at 30° C for 10 minutes. Each reaction contained 2 nM ³²P-labeled RNA substrate and concentrations of FCP1₅₆₂₋₇₃₈ or HIV-1 Tat₁₋₈₆ as indicated in the figure or legend. Complexes were resolved on a pre-run 8 % polyacrylamide gel (40:1 acrylamide:bisacrylamide) in 1X Tris-glycine buffer for 2 h at 4° C. Cleavage reactions

(20 μ l) were performed in the buffer described above supplemented with 5 mM MgCl_2 .

Reactions were performed at 30° C for 30 min unless otherwise indicated. Cleavage products were analyzed on 20 % polyacrylamide gel (19:1 acrylamide:bisacrylamide) in 1X Tris-borate EDTA with 7 M urea unless stated otherwise. Time course reactions (100 μ l) were initiated by the addition of 1 μ M purified FCP₁₅₆₂₋₇₃₈. At indicated time points 10 μ l aliquots were removed and immediately quenched with EDTA (100 mM) and 95 % formamide gel loading buffer.

REFERENCES

1. Hirose, Y. & Manley, J. L. RNA polymerase II and the integration of nuclear events. *Genes Dev* **14**, 1415-29 (2000).
2. Chambers, R. S. & Dahmus, M. E. Purification and characterization of a phosphatase from HeLa cells which dephosphorylates the C-terminal domain of RNA polymerase II. *J. Biol. Chem.* **269**, 26243-26248 (1994).
3. Archambault, J. et al. An essential component of a C-terminal domain phosphatase that interacts with transcription factor IIF in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **94**, 14300-14305 (1997).
4. Archambault, J. et al. FCP1, the RAP74-interacting subunit of a human protein phosphatase that dephosphorylates the carboxy-terminal domain of RNA polymerase II. *J. Biol. Chem.* **273**, 27593-27601 (1998).
5. Cho, H. et al. A protein phosphatase functions to recycle RNA polymerase II. *Genes & Dev.* **13**, 1540-1552 (1999).
6. Cho, E. J., Kobor, M. S., Kim, M., Greenblatt, J. & Buratowski, S. Opposing effects of Ctk1 kinase and FCP1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. *Genes & Dev.* **15**, 3319-3329 (2001).
7. Mandal, S. S., Cho, H., Kim, S., Cabane, K. & Reinberg, D. FCP1, a phosphatase specific for the heptapeptide repeat of the largest subunit of RNA polymerase II, stimulates transcription elongation. *Mol. Cell Biol.* **22**, 7543-7552 (2002).

8. Costa, P. J. & Arndt, K. M. Synthetic lethal interactions suggest a role for the *Saccharomyces cerevisiae* Rtf1 protein in transcription elongation. *Genetics* **156**, 535-547 (2000).
9. Lindstrom, D. L. & Hartzog, G. A. Genetic interactions of Spt4-Spt5 and TFIIS with the RNA polymerase II CTD and CTD modifying enzymes in *Saccharomyces cerevisiae*. *Genetics* **159**, 487-97 (2001).
10. Abbott, K. L., Archambault, J.A., Xiao, H., Nguyen B.D., Roeder R.G., Greenblatt J., Omichinski J.G., and Legault P.L.. Interactions of the HIV-1 Tat and RAP74 Proteins with the RNAPII CTD Phosphatase FCP1. *To be submitted*.
11. Yamane, K. & Tsuruo, T. Conserved BRCT regions of TopBP1 and of the tumor suppressor BRCA1 bind strand breaks and termini of DNA. *Oncogene* **18**, 5194-5203 (1999).
12. Yamane, K., Eisaku, K. & Tsuruo, T. The BRCT regions of tumor suppressor BRCA1 and of XRCC1 show DNA end binding activity with a multimerizing feature. *Biochem. Biophys. Res. Commun.* **279**, 678-684 (2000).
13. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403-410 (1990).
14. Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M. & Tollervey, D. The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'-->5' exoribonucleases. *Cell* **91**, 457-66 (1997).
15. White, D. J., Reiter, N. J., Sikkink, R. A., Yu, L. & Rusnak, F. Identification of the high affinity Mn²⁺ binding site of bacteriophage lambda phosphoprotein phosphatase: effects

- of metal ligand mutations on electron paramagnetic resonance spectra and phosphatase activities. *Biochemistry* **40**, 8918-29 (2001).
16. Kobor, M. S. et al. An unusual eukaryotic protein phosphatase required for transcription by RNA polymerase II and CTD dephosphorylation in *S. cerevisiae*. *Molecular Cell* **4**, 55-62 (1999).
 17. Rhim, H., Echeteu, C. O. & Rice, A. P. Wild type and mutant HIV-1 and HIV-1 Tat proteins expressed in *Escherichia coli* as fusions with glutathione S-transferase. *J. Acquired Immune Defic. Syndr.* **7**, 1116-1121 (1994).
 18. Milligan, J. F. & Uhlenbeck, O. C. Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol.* **180**, 51-62 (1989).

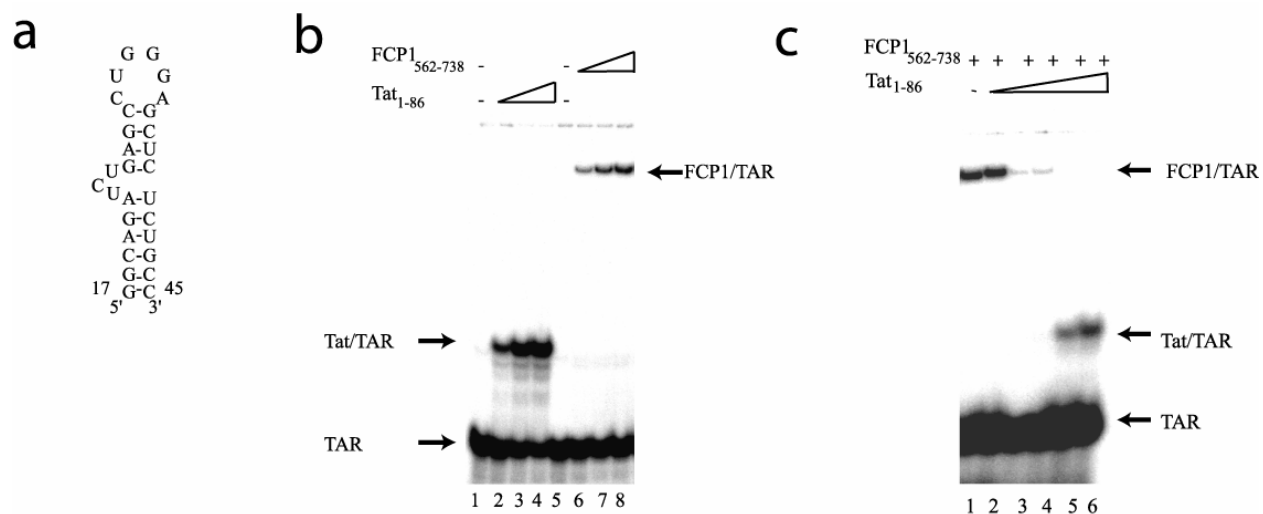


Figure 4.1 RNA binding activity of the FCP1 CD. **a**, Structure of HIV-1 TAR₁₇₋₄₅. **b**, Native gel shift assay using 5'-³²P-labeled TAR with purified Tat₁₋₈₆ (1, 10, and 100 nM) or purified FCP1₅₆₂₋₇₃₈ (0.2, 0.6, and 1.2 μM). **c**, Native gel shift showing FCP1₅₆₂₋₇₃₈/TAR₁₇₋₄₅ complex (1.2 μM FCP1₅₆₂₋₇₃₈ and 2 nM ³²P-labeled TAR₁₇₋₄₅) challenged with increasing concentrations of purified Tat₁₋₈₆ (0.1, 0.5, 1, and 10 nM).

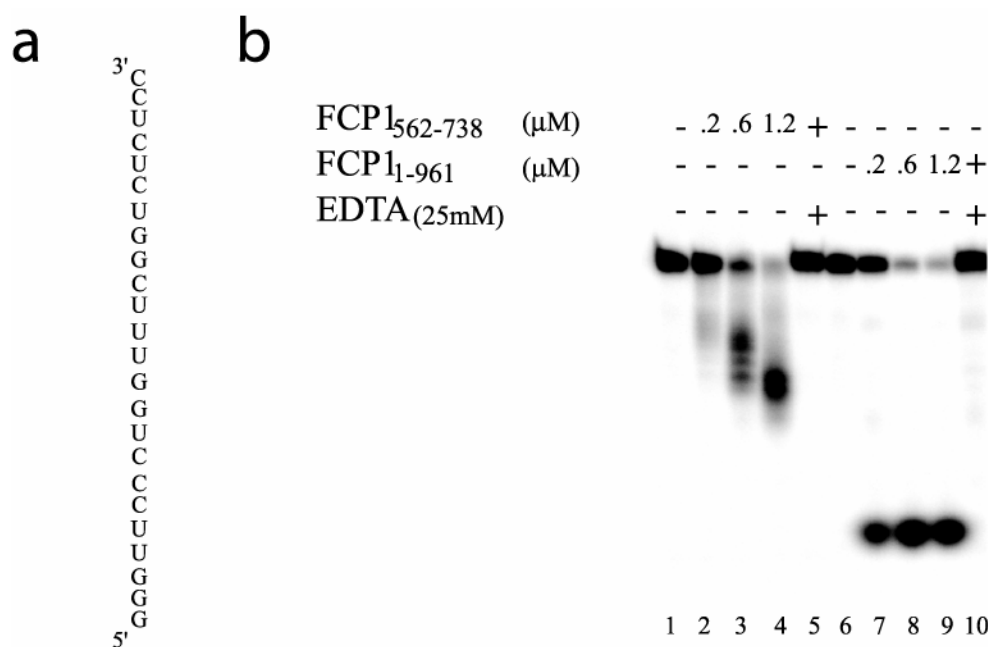
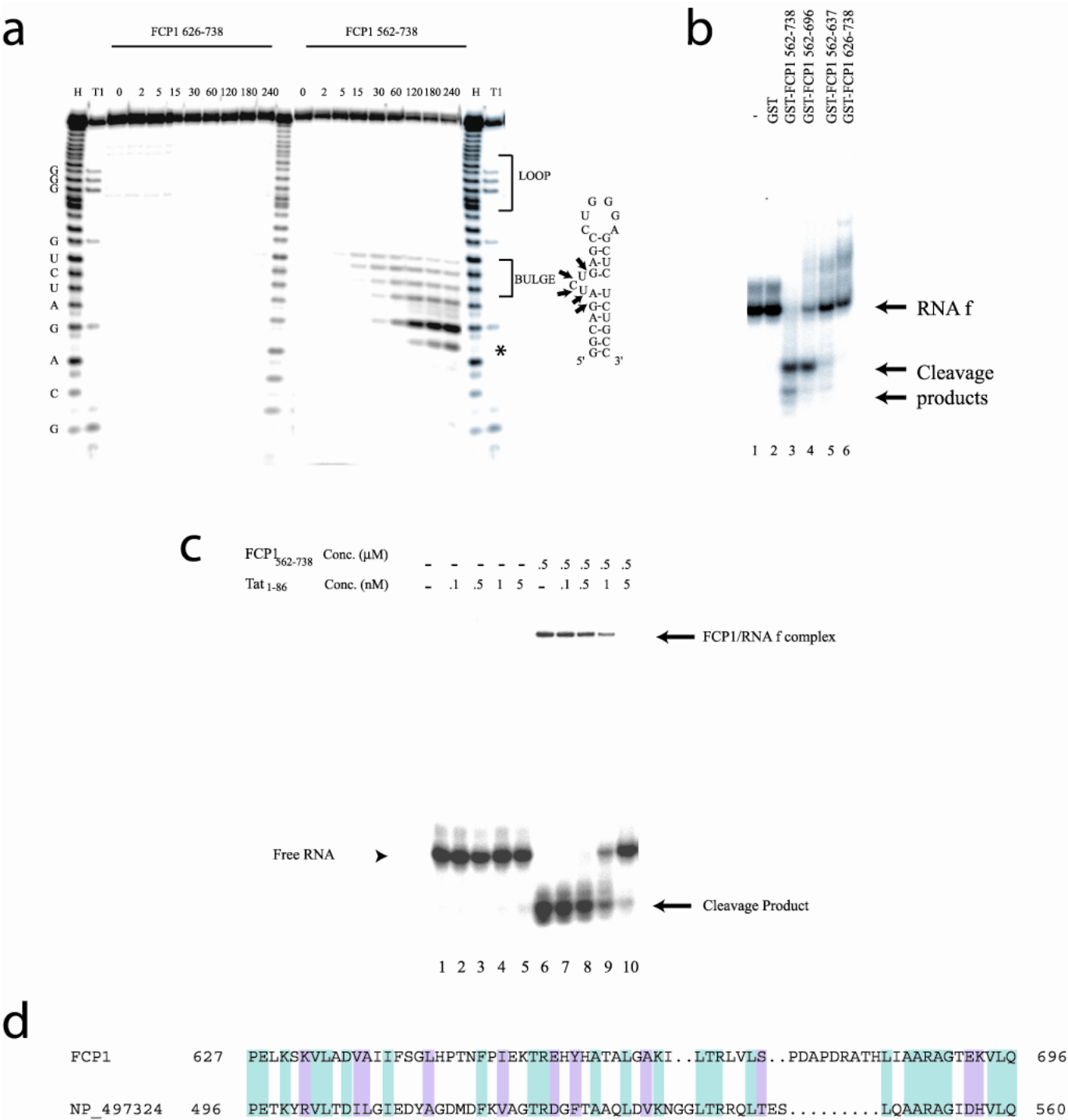


Figure 4.2 Identification of a single-strand, EDTA sensitive, nonsequence-specific endonuclease activity associated with purified human FCP1₅₆₂₋₇₃₈ and purified human FCP1₁₋₉₆₁. **a**, Sequence of the TAR-derived single-strand RNA. **b**, Denaturing gel electrophoresis of nuclease reactions containing purified FCP1₅₆₂₋₇₃₈ (lanes 2,3,4,5) or purified FCP1₁₋₉₆₁ (lanes 7,8,9,10).

Figure 4.3. FCP1 CD has bulge-specific cleavage activity on HIV-1 TAR RNA. **a**, Time course (numbers indicate minutes) of 5'-³²P-TAR₁₇₋₄₅ digested with purified FCP1₆₂₆₋₇₃₈ or purified FCP1₅₆₂₋₇₃₈. Arrows indicate the cleavage points, and the asterisk highlights a product that accumulates over time. **b**, Cleavage activity of various GST-FCP1 fragments with 5'-³²P-labeled RNA. Products were run on a native gel. **c**, Nuclease assay with increasing concentrations of Tat₁₋₈₆ in the absence (lanes 2-5) or presence (lanes 7-10) of purified FCP1₅₆₂₋₇₃₈. Reaction products were resolved on a native polyacrylamide gel. **d**, BLAST¹³ alignment of FCP1 (627-696) to the RNase PH_C domain of polyribonucleotide nucleotidyltransferase (NP_497324) from *Caenorhabditis elegans*. Blue highlights indicate sequence identity; pink highlights indicate sequence homology. Bold amino acids show a conserved DxxDR phosphoesterase motif¹⁹ inserted within this RNase PH_C homologous region of FCP1.



SUPPLEMENTARY MATERIAL

Substrates

DNA oligonucleotide h was purchased from IDT (Integrated DNA Technologies). RNA f was purchased from Dharmacon (Boulder, CO). The following RNA's were transcribed using T7 RNA polymerase prepared in house and single-stranded oligonucleotide templates purchased from IDT :RNA a ,RNA b, RNA c, RNA e, and RNA g¹⁸. RNA d was purchased from Dharmacon (Boulder, CO). All RNA's transcribed for this study were purified from 20% polyacrylamide gels containing 7M urea in 1X Tris-borate EDTA buffer. Alkaline phosphatase dephosphorylated, ³²P-labeled stem-loop RNA and DNA substrates were dissolved in 0.5X TEN (20 mM Tris-HCl pH 8.0, 75 mM NaCl, and 0.5 mM EDTA), heated to 95 °C for 1 min and cooled quickly on ice prior to each experiment.

Figure S4.1. FCP1 binds and cleaves single-stranded nucleic acid substrates without sequence specificity. (A) Sequence of single-stranded nucleic acid substrates. (B) Native gel mobility shift assay using 5' ³²P-labeled substrates (2 nM). Concentrations of purified FCP1₅₆₂₋₇₃₈ used are (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.8, 1.0, and 1.2 μM). (C) Graphical representation of time course cleavage reactions performed using 5' ³²P-labeled substrates. Reactions were quenched at the indicated time points by the addition of 1% SDS and products were resolved on 20% denaturing polyacrylamide gels. The percent cleavage was determined by quantification of the uncleaved and cleaved products using ImageQuant (Molecular Dynamics-Amersham Pharmacia, NJ).

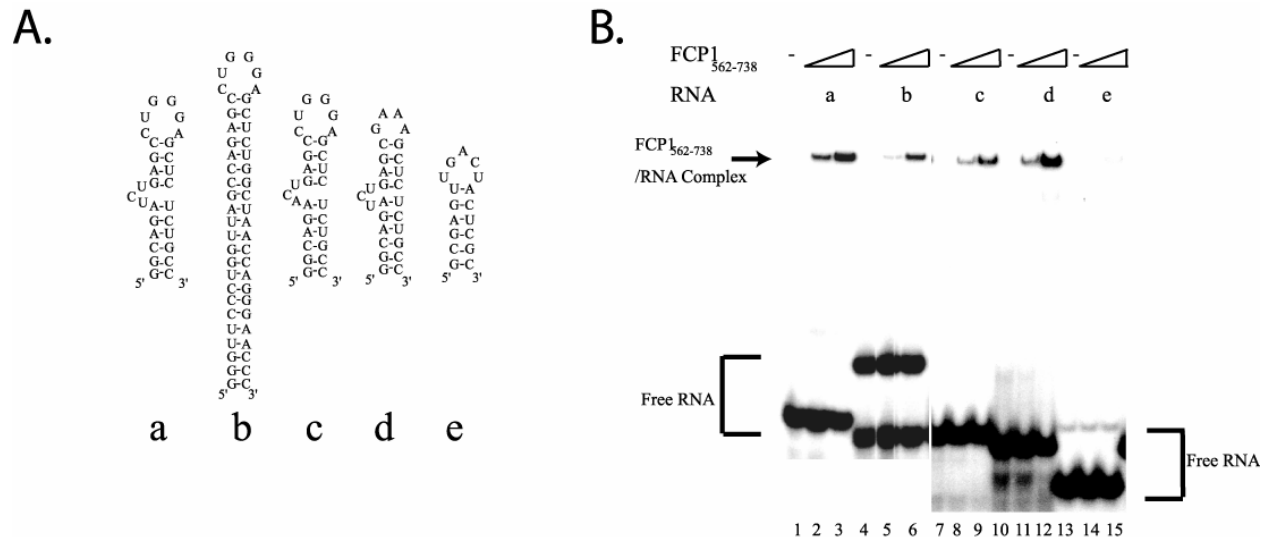


Figure S4.2. FCP1 binds a variety of structured RNA's. (A) Secondary structure and sequence of TAR₁₇₋₄₅ (a), dsTAR (b), Δ bulge TAR (c), loop mutant TAR (d), and stem loop V from *Neurospora* (e). (B) Native gel shift assay using 5' ³²P-labeled stem loop RNA 2 nM and purified FCP1₅₆₂₋₇₃₈ 0.1 and 1.0 μ M.

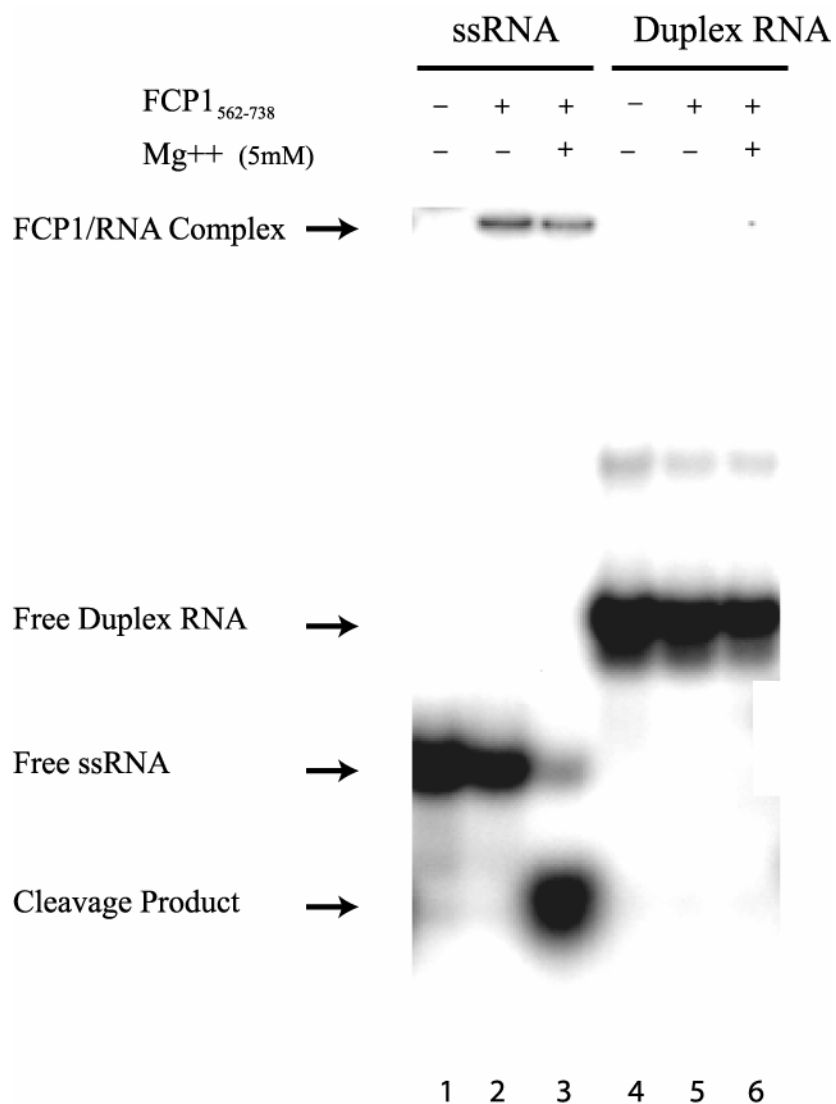


Figure S4.3. RNA binding and cleavage activity of FCP1₅₆₂₋₇₃₈ with ssRNA and duplex RNA. Native gel shift assay using 5'-³²P-labeled RNA f (2 nM) alone or complexed with unlabeled RNA g. Purified FCP1₅₆₂₋₇₃₈ (1.2 μM) was added in the presence and absence of Mg⁺⁺ as indicated.

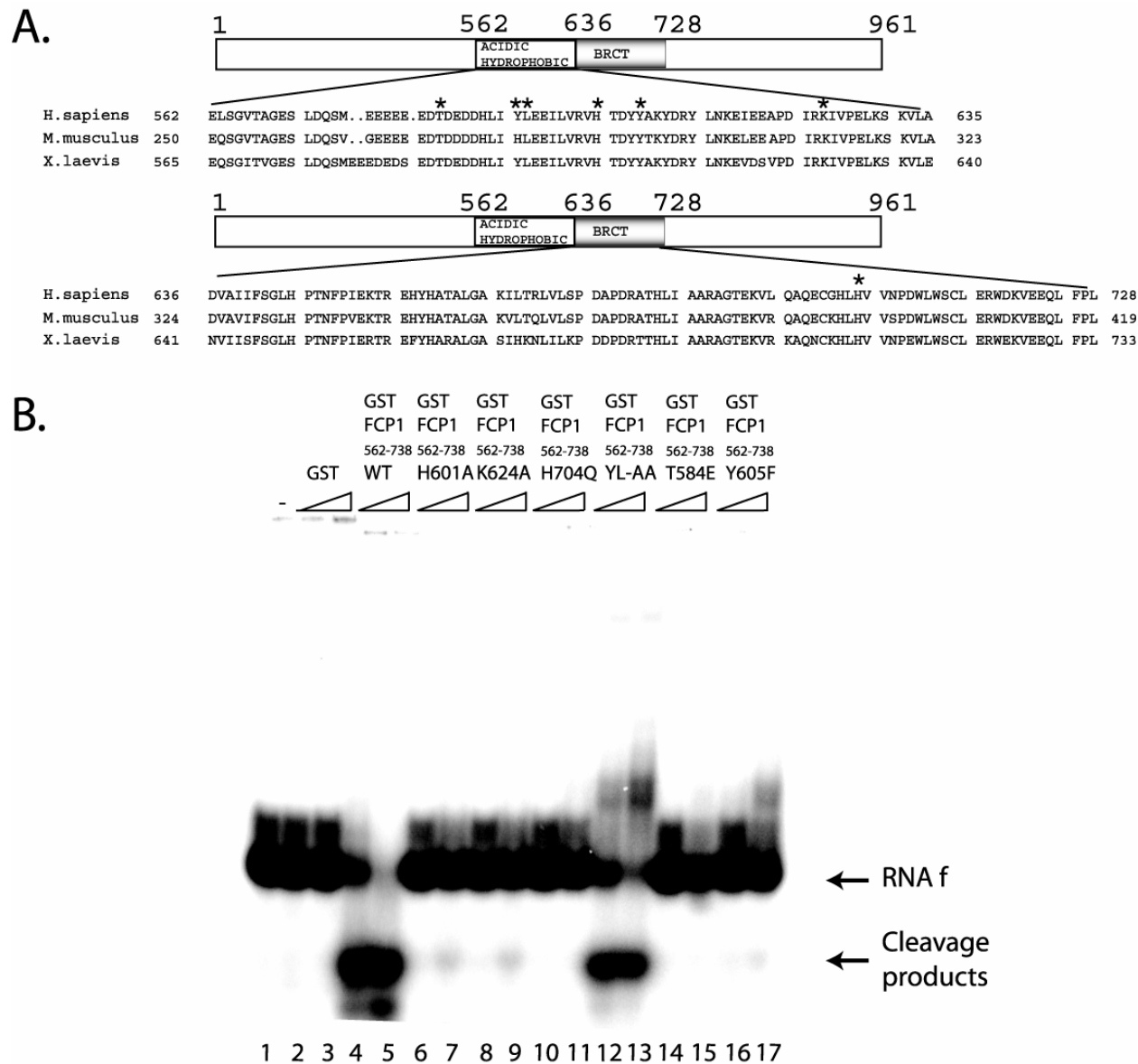


Figure S4.4. Effect of mutations on FCP1 nuclease activity. (A) Alignment of the FCP1 central domain from vertebrates (human, mouse, and frog). Residues that are mutated are marked by an asterisk. (B) Cleavage activity of various wild-type and mutant GST-FCP1 fragments with 5' 32 P-labeled RNA f. Products were resolved on a native 11 % polyacrylamide gel (40:1 acrylamide:bisacrylamide).

CHAPTER 5

¹H, ¹⁵N, AND ¹³C ASSIGNMENTS AND CSI ANALYSIS OF A COMPLEX FORMED BY
THE CARBOXYL-TERMINAL DOMAIN OF HUMAN RAP74 AND A
PHOSPHORYLATED PEPTIDE FROM THE CENTRAL DOMAIN OF FCP1⁴

⁴Karen L. Abbott, James G. Omichinski, and Pascale Legault.

BIOLOGICAL SIGNIFICANCE

Human FCP1 (TFIIF-associated CTD phosphatase) is an essential component of the RNA polymerase II (RNAP II) transcription machinery. FCP1 dephosphorylates the C-terminal domain (CTD) of the largest subunit of RNAP II (Chambers and Dahmus 1994). Phosphorylation of RNAP II CTD induces conformational changes in the polymerase that triggers a switch from transcription initiation at the promoter to an active elongating form of the polymerase (Lu et al. 1991; O'Brien et al. 1994).

Dephosphorylation of RNAP II is required for recycling of RNAP II to allow new initiation events (Dahmus 1996). FCP1 is the first CTD-specific phosphatase identified. FCP1 has three main domains based on sequence conservation (Archambault et al. 1997).

The amino-terminal FCP1 homology domain contains a conserved motif ($\Psi\Psi\Psi\text{DXD}$ X(T/V) $\Psi\Psi$ where Ψ is a hydrophobic residue) found in a class of small molecule phosphohydrolases and phosphotransferases (Collet et al. 1998; Kobor et al. 1999).

Purified recombinant yeast FCP1 carrying a mutation within this conserved phosphotransferase motif (D180E) failed to dephosphorylate the CTD of RNAP II *in vitro* (Kobor et al. 1999). The central regulatory domain of FCP1 has been shown to be essential for cell viability and phosphatase activity in yeast (Kobor et al. 2000). This region contains a conserved region rich in acidic and hydrophobic amino acids adjacent to a single BRCT (BRCA1-C-terminus) domain (Abbott et al. 2004a). The carboxyl terminus of FCP1 is highly acidic and has similarities structurally and functionally to transcription activation domains (Nguyen et al. 2003a). The carboxyl-terminus of the large subunit of the general transcription factor TFIIF (RAP74) interacts with both the

central and carboxyl domains of FCP1 at a small conserved hydrophobic motif (Nguyen et al. 2003a) (Abbott et al. 2004a). Interaction of RAP74 with FCP1 stimulates phosphatase activity (Chambers et al. 1995). High resolution structures have been determined for RAP74₄₃₆₋₅₁₇ alone as well as the complex of RAP74₄₃₆₋₅₁₇/FCP1₈₇₉₋₉₆₁ (Kamada et al. 2001; Kamada et al. 2003; Nguyen et al. 2003a; Nguyen et al. 2003b). Published structures of RAP74₄₃₆₋₅₁₇ in a complex with the carboxyl-terminal region of FCP1 illustrate that FCP1 adopts an H1' helix upon binding with RAP74 (Nguyen et al. 2003a; Kamada et al., 2003). This helix nestles into a hydrophobic pocket formed by helices H2 and H3 of the RAP74 winged-helix domain. Since these structures were completed, it has been discovered that phosphorylation of FCP1 by protein kinase CK2 (CK2) enhances the binding of RAP74 and increases phosphatase activity (Palancade et al. 2002). Both the central and carboxyl RAP74 binding sites within FCP1 contain conserved CK2 sites (Abbott et al. 2004b). The central domain of human FCP1 contains a 21 amino acid region (578-598) conserved in vertebrates that contains a CK2 site upstream of the small hydrophobic motif required for RAP74 binding. Peptide studies have confirmed that this peptide is the minimal RAP74 binding site within the central regulatory domain of FCP1 (Abbott et al. 2004a). This NMR study reports the ¹H, ¹⁵N, and ¹³C resonance assignments of human RAP74₄₃₆₋₅₁₇ in a complex with a peptide (579-600 phosphorylated at position T584) from the central domain of human FCP1.

METHODS

Expression and purification of RAP74 has been described previously (Nguyen et al. 2003b). Uniform $^{15}\text{N}/^{13}\text{C}$ labeling was obtained by growing cells in minimal media containing ^{15}N -labeled NH_4Cl and ^{13}C -labeled glucose as the sole sources of nitrogen and carbon, respectively. The FCP1 peptide (579-600, phosphorylated at T584) was chemically synthesized at the Medical College of Georgia. The peptide was purified to homogeneity on a C-4 reverse phase HPLC column using a 30-50% acetonitrile gradient. Purified peptide was stored lyophilized in aliquots at -20°C . NMR samples prepared by titration of unlabeled peptide into labeled RAP74, consisted of 1 mM [^{13}C , ^{15}N]RAP74₄₃₆₋₅₁₇/ 1 mM unlabeled FCP1₅₇₉₋₆₀₀T584PO₄) in 20 mM sodium phosphate pH 6.5 and 1 mM EDTA. All experiments were collected on a sample containing 90% H₂O: 10% D₂O.

NMR spectra were acquired on a Varian Inova Unity 600 MHz NMR spectrometer equipped with a z pulsed-field gradient unit and HCN triple resonance probes. The temperature of all probes was calibrated to 27°C using a methanol standard. Backbone and aliphatic side chain chemical shifts were assigned using a combination of experiments including 2D ^1H - ^{15}N HSQC (Kay et al. 1992), 3D HNCACB (Grzesiek et al. 1992; Wittekind and Mueller 1993; Muhandiram and Kay 1994), 3D (HB)CBCA(CO)NNH (Grzesiek et al. 1992; Muhandiram and Kay 1994), 3D HNCO (Grzesiek et al. 1992), and 3D H(CCO)NNH-TOCSY and 3D C(CO)NNH-TOCSY (Grzesiek et al. 1993; Muhandiram and Kay 1994). Chemical shifts of all proton, carbon, and nitrogen resonances were referenced to an external DSS standard at 0 ppm. NMR data was processed with NMRPipe / NMRDraw (Delaglio et al. 1995) and analyzed using

NMRView (Johnson and Blevins 1994). Secondary structure elements were identified using a chemical shift index (CSI) consensus plot (Wishart and Sykes 1994).

EXTENT OF ASSIGNMENTS

100% of ^{15}N , $^1\text{H}_\text{N}$, $^{13}\text{C}_\alpha$, and $^{13}\text{C}_\beta$ resonances for RAP74₄₃₆₋₅₁₇ have been assigned excluding prolines and the $^1\text{H}_\text{N}$ of serine 436. 98% of the $^{13}\text{C}'$ resonances were assigned excluding prolines. 89% of the $^1\text{H}_\alpha$ resonances have been assigned for RAP74₄₃₆₋₅₁₇ (all residues). There are six proline residues in RAP74₄₃₆₋₅₁₇, and the ^{15}N and $^{13}\text{C}'$ resonances have not been assigned. The global chemical shift index calculated from the $^{13}\text{C}_\alpha$, $^1\text{H}_\alpha$, $^{13}\text{C}_\beta$, and $^{13}\text{C}'$ chemical shifts for both free RAP74₄₃₆₋₅₁₇ (Nguyen et al. 2003b) and RAP74₄₃₆₋₅₁₇ complexed to FCP1₅₇₉₋₆₀₀(T584PO₄) are displayed in Figure 5.1. The overall secondary structure predicted from the chemical shift data does not change significantly between free RAP74₄₃₆₋₅₁₇ and RAP74₄₃₆₋₅₁₇ complexed with FCP1₅₇₉₋₆₀₀(T584PO₄).

For RAP74₄₃₆₋₅₁₇ complexed with FCP1₅₇₉₋₆₀₀(T584PO₄), 90% of side chain resonances (^1H -bound ^{13}C , ^1H -bound ^{15}N , and ^1H , except for ^1H - ^{15}N of Lys and Arg) were completely assigned, 8% were partially assigned, and 2% (P440 and E517) were unassigned. The complete chemical shift assignments are shown in Table 5.1. Figure 5.2 shows a 2D ^1H - ^{15}N HSQC spectrum overlay of RAP74₄₃₆₋₅₁₇ complexed with FCP1₅₇₉₋₆₀₀ and RAP74₄₃₆₋₅₁₇ complexed with FCP1₅₇₉₋₆₀₀(T584PO₄) showing the amino acid assignments.

ACKNOWLEDGEMENTS:

We would like to thank Frank Delaglio, Daniel Garrett, and Bruce Johnson for NMR data processing and analysis programs. We would like to thank Lewis Kay for NMR pulse sequences. This work was supported by National Institutes of Health Grant RO1 GM60298-01 (J.G.O. and P.L.).

REFERENCES

- Abbott, K.L., J. Archambault, H. Xiao, B.D. Nguyen, R.G. Roeder, J. Greenblatt, J.G. Omichinski, and P. Legault. 2004a. Interactions of the HIV-1 Tat and RAP74 Proteins with the RNAPII CTD Phosphatase FCP1. *(To be submitted to J. Biol.Chem.)*
- Abbott, K.L., M.B. Renfrow, M.J. Chalmers, B.D. Nguyen, J. Greenblatt, A.G. Marshall, P. Legault, and J.G. Omichinski. 2004b. Investigations into the enhanced binding of the RNAPII CTD phosphatase FCP1 to RAP74 following CK2 phosphorylation. *(To be submitted to J. Biol. Chem.)*
- Archambault, J., R.S. Chambers, M.S. Kobor, Y. Ho, M. Cartier, D. Bolotin, B. Andrews, C.M. Kane, and J. Greenblatt. 1997. An essential component of a C-terminal domain phosphatase that interacts with transcription factor IIF in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **94**: 14300-14305.
- Chambers, R.S. and M.E. Dahmus. 1994. Purification and characterization of a phosphatase from HeLa cells which dephosphorylates the C-terminal domain of RNA polymerase II. *J. Biol. Chem.* **269**: 26243-26248.
- Chambers, R.S., B.Q. Wang, Z.F. Burton, and M.E. Dahmus. 1995. The activity of COOH-terminal domain phosphatase is regulated by a docking site on RNA polymerase II and by the general transcription factors IIF and IIB. *J. Biol. Chem.* **270**: 14962-14969.
- Collet, J.F., V. Stroobant, M. Pirard, G. Delpierre, and E.V. Schaftingen. 1998. A new class of phosphotransferases phosphorylated on an aspartate residue in an amino-terminal DXDX(T/V) motif. *J. Biol. Chem.* **273**: 14107-14112.
- Dahmus, M.E. 1996. Reversible phosphorylation of the C-terminal domain of RNA polymerase II. *J. Biol. Chem.* **271**: 19009-19012.

- Delaglio, F., S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, and A. Bax. 1995. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**: 277-293.
- Grzesiek, S., J. Anglister, and A. Bax. 1993. Correlation of backbone amide and aliphatic side-chain resonances in $^{13}\text{C}/^{15}\text{N}$ -enriched proteins by isotropic mixing of ^{13}C magnetization. *J. Magn. Res.* **101**: 114-121.
- Grzesiek, S., M. Ikura, G.M. Clore, A.M. Gronenborn., and A.Bax. 1992. A 3D triple-resonance NMR technique for qualitative measurement of carbonyl- H_β J couplings in isotopically enriched proteins. *J. Magn. Res.* **96**: 215-221.
- Johnson, B.A. and R.A. Blevins. 1994. *J. Biomol. NMR* **4**: 603-614.
- Kamada, K., J.D. Angelis, R. Roeder, and S. Burley. 2001. Crystal structure of the C-terminal domain of the RAP74 subunit of human transcription factor IIF. *Proc. Natl. Acad. Sci. U.S.A.* **98**: 3115-20.
- Kamada, K., R.G. Roeder, and S.K. Burley. 2003. Molecular mechanism of recruitment of TFIIF-associating RNA polymerase C-terminal domain phosphatase (FCP1) by transcription factor IIF. *Proc. Natl. Acad. Sci. U.S.A.* **100**: 2296-2299.
- Kay, L.E., P. Keifer, and T. Saarinen. 1992. Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc.* **114**: 10663-10665.
- Kobor, M.S., J. Archambault, W. Lester, F.C.P. Holstege, O. Gileadi, and J. Greenblatt. 1999. An unusual eukaryotic protein phosphatase required for transcription by RNA Polymerase II and CTD dephosphorylation in *S. cerevisiae*. *Molecular Cell* **4**: 55-62.

- Kobor, M.S., L.D. Simon, J. Omichinski, G. Zhong, J. Archambault, and J. Greenblatt. 2000. A motif shared by TFIIF and TFIIB mediates their interaction with the RNA polymerase II carboxyl-terminal domain phosphatase Fcp1p in *Saccharomyces cerevisiae*. *Mol. Cell Biol.* **20**: 7438-7449.
- Lu, H., O. Flores, R. Weinmann, and D. Reinberg. 1991. The nonphosphorylated form of the RNA polymerase II preferentially associates with the preinitiation complex. *Proc. Natl. Acad. Sci. USA* **88**: 10004-10008.
- Muhandiram, D. and L. Kay. 1994. *J. Magn. reson., Ser.B* **103**: 203-216.
- Nguyen, B.D., K.L. Abbott, K. Potempa, M.S. Kobor, J. Archambault, J. Greenblatt, P. Legault, and J.G. Omichinski. 2003a. NMR structure of a complex containing the TFIIF subunit RAP74 and the RNA polymerase II carboxyl-terminal domain phosphatase FCP1. *Proc. Natl. Acad. Sci. USA* **100**: 5688-5693.
- Nguyen, B.D., H. Chen, A.S. Kobor, J. Greenblatt, P. Legault, and J.G. Omichinski. 2003b. Solution structure of the carboxyl-terminal domain of RAP74 and NMR characterization of the FCP1-binding sites of RAP74 and Human TFIIB. *Biochemistry* **42**:1460-1469.
- O'Brien, T.S., S. Hardin, A. Greenleaf, and J.T. Lis. 1994. Phosphorylation of RNA polymerase II C-terminal domain and transcriptional elongation. *Nature* **370**: 75-77.
- Palancade, B., M.F. Dubois, and O. Bensaude. 2002. FCP1 phosphorylation by casein kinase 2 enhances binding to TFIIF and RNA polymerase II carboxyl-terminal domain phosphatase activity. *J. Biol. Chem.* **277**: 36061-36067.
- Wishart, D.S. and B.D. Sykes. 1994. *J. Biomol. NMR* **4**: 171-180.
- Wittekind, M. and L. Mueller. 1993. *J. Magn. Reson. Ser. B* **101**: 201-205.

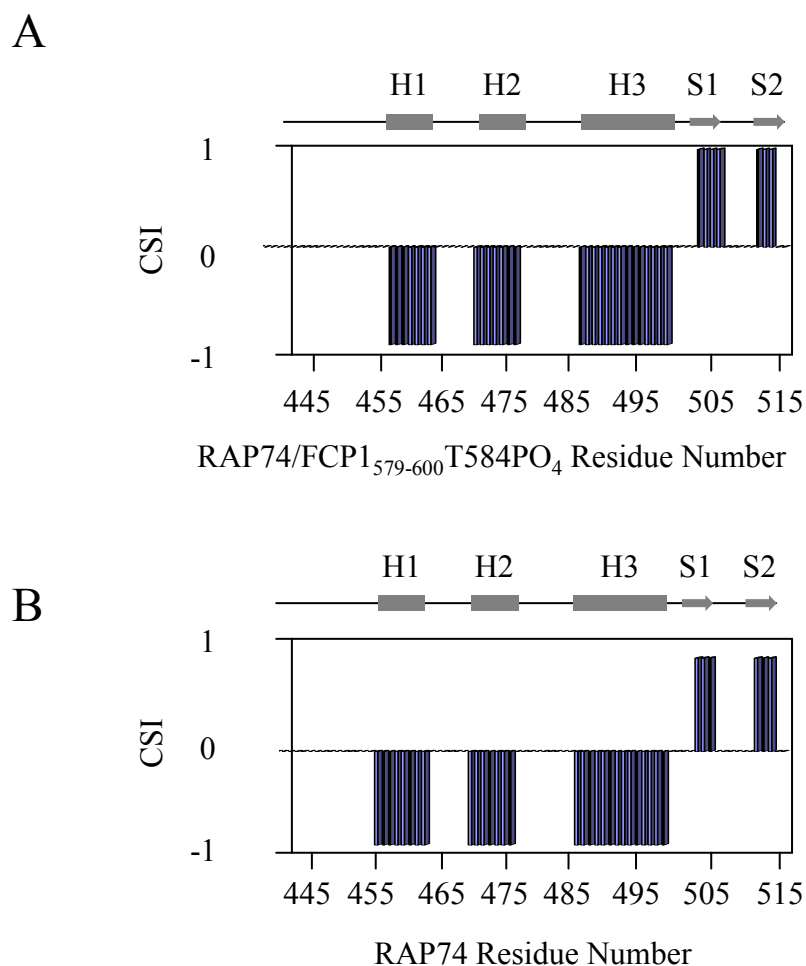


Fig. 5.1. Chemical shift index (CSI) consensus plot based on $C\alpha$, $H\alpha$, $C\beta$, and C' chemical shifts (Wishart and Sykes, 1994) identifying regular secondary structure elements for RAP74₄₃₆₋₅₁₇ in a complex with FCP1₅₇₉₋₆₀₀PO₄ (A) and RAP74₄₃₆₋₅₁₇ alone (B). Residue numbering for secondary structure elements shown in A are: H1 (457-463), H2 (470-476), H3 (486-499), S1 (504-507), and S2 (513-515). Residue numbering for secondary structure elements shown in B are: H1 (456-463), H2 (470-476), H3 (486-499), S1 (504-506), and S2 (513-515).

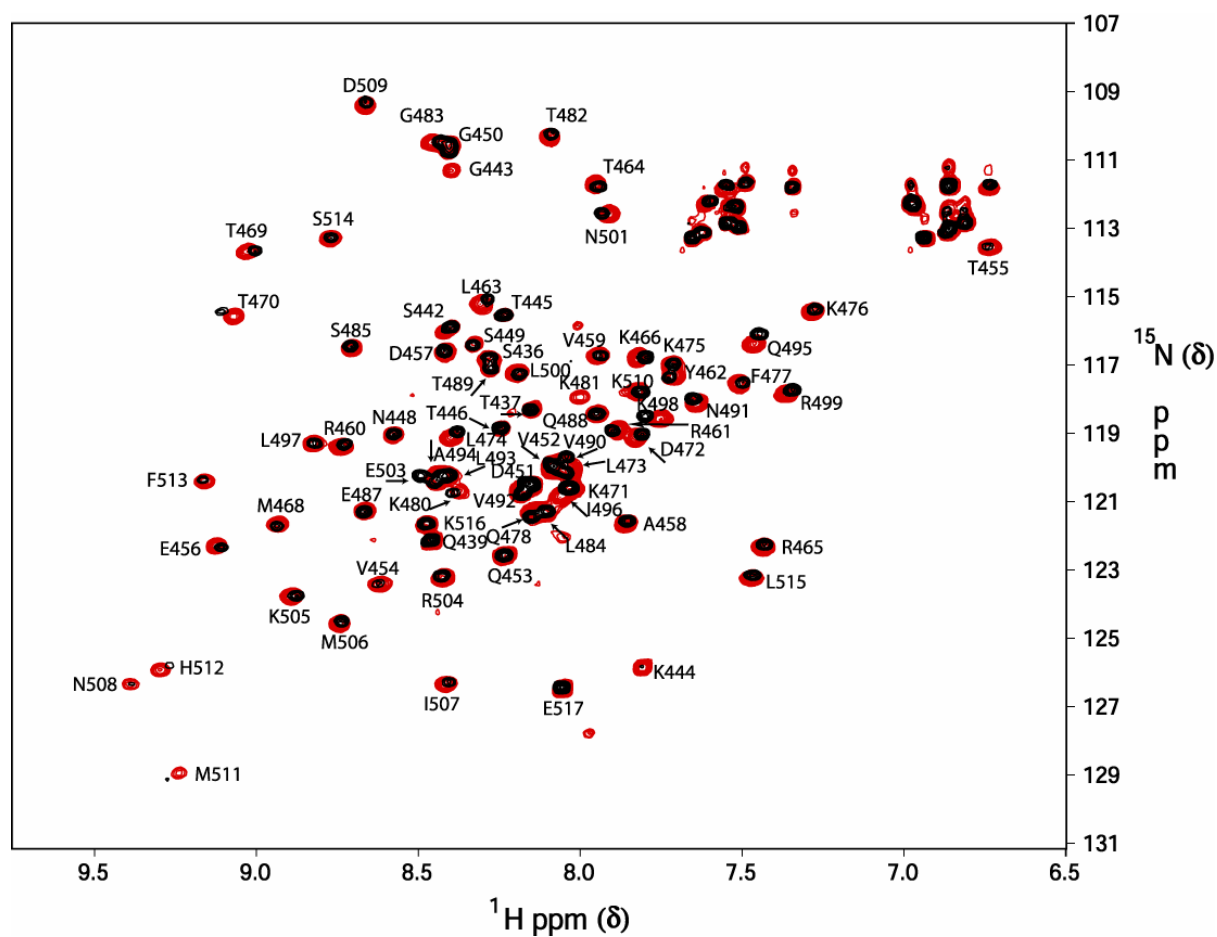


Figure 5.2. 2D $[^{15}\text{N}, ^1\text{H}]$ -HSQC spectrum overlay of RAP74₄₃₆₋₅₁₇/FCP1₅₇₉₋₆₀₀ (black) and RAP74₄₃₆₋₅₁₇/FCP1₅₇₉₋₆₀₀(T584PO₄) (red) recorded on a 600 MHz Varian Inova spectrometer. Peaks are identified using the single letter amino acid code.

Table 5.1. Chemical shift assignments for RAP74 in the RAP74₄₃₆₋₅₁₇/FCP1₅₇₉₋₆₀₀ (T584PO₄) complex.

#	RAP74 #	AA	N (± 0.2) ^a	HN (± 0.1) ^a	C α (± 0.2) ^b	H α (± 0.2) ^c	C β (± 0.2) ^b	H β (± 0.2) ^c	C' (± 0.3) ^d	C γ (± 0.2) ^b	H γ (± 0.2) ^c	C δ (± 0.2) ^b	H δ (± 0.2) ^c	C ϵ (± 0.2) ^b	H ϵ (± 0.2) ^c
1	436	S	116.9		58.2	4.49	63.9	3.85	174.27						
2	437	T	118.3	8.15	60.2	4.50	69.7	3.83	174.20						
3	438	P	122.1		62.9	4.41	34.5	2.29	176.7	24.7	1.99		HD1 3.86 HD2 3.67		
4	439	Q	122.1	8.45	53.5	4.44	28.7	1.90	176.69		2.27				
5	440	P			63.6		32.1								
6	441	P			63.1	4.47	32.1	2.32	177.10	27.5	2.04	50.5	HD1 3.84 HD2 3.64		
7	442	S	116.0	8.42	58.4	4.43	64.0	3.88	174.90						
8	443	G	110.5	8.46	45.4	3.96			173.16						
9	444	K	125.9	7.81	56.2	4.42	33.3	1.86	176.90	24.8	1.41	29.1	1.71	42.2	2.98
10	445	T	115.5	8.23	61.7	4.49	70.0	3.86	176.84	21.8	0.93				
11	446	T	118.9	8.24	60.2		69.7		174.47						
12	447	P			63.4	4.45	32.3	2.31	176.66	27.5	2.00	51.2	HD1 3.88 HD2 3.75		
13	448	N	119.0	8.58	53.3	4.71	39.0	2.84	175.48						
14	449	S	116.4	8.33	58.4	4.40	64.0	3.87	175.30						
15	450	G	110.6	8.40	45.3	3.98			173.99						
16	451	D	120.6	8.16	54.6	4.36	41.3	2.92	176.40						
17	452	V	120.6	8.16	62.5	4.12	32.9	2.02	175.39		HG1 1.30 HG2 0.99				
18	453	Q	122.6	8.23	55.2	4.46	30.7	2.37	175.57	34.0	2.37				
19	454	V	123.4	8.62	62.4	3.54	30.0	1.87	174.40		0.61				
20	455	T	113.6	6.74	59.0	4.60	72.5	4.60	174.60	21.9	1.12				

#	RAP74 #	AA	N (± 0.2) ^a	HN (± 0.1) ^a	C α (± 0.2) ^b	H α (± 0.2) ^c	C β (± 0.2) ^b	H β (± 0.2) ^c	C' (± 0.3) ^d	C γ (± 0.2) ^b	H γ (± 0.2) ^c	C δ (± 0.2) ^b	H δ (± 0.2) ^c	C ϵ (± 0.2) ^b	H ϵ (± 0.2) ^c
21	456	E	122.3	9.12	60.2	4.46	30.0	2.06	177.65	37.0	2.15				
22	457	D	116.6	8.42	57.4	4.30	41.2	2.56	178.05						
23	458	A	121.6	7.86	55.1	4.01	19.1	1.67	176.75						
24	459	V	116.7	7.95	67.6	3.69	31.6	2.02	177.43		0.79				
25	460	R	119.4	8.74	60.2	3.92	30.3	2.26	178.54	28.0	1.94	43.1			
26	461	R	118.9	7.88	59.7	3.92	29.9	1.71	179.70	27.6	0.86	43.5	3.06		
27	462	Y	117.4	7.71	61.5	4.53	37.9	HB2 2.99 HB1 3.41	178.41						
28	463	L	115.2	8.31	56.6	4.60	42.5	1.89	178.60	26.5	1.37		0.98		
29	464	T	111.7	7.95	63.5	4.55	69.7	4.21	175.88	22.0	1.22				
30	465	R	122.3	7.43	58.5	4.23	31.2	1.87	179.69	27.3	1.72	43.6	3.27		
31	466	K	116.8	7.82	54.4		33.5		176.72						
32	467	P			62.4	5.05	31.9	1.99	178.22	28.0	1.74	50.5	3.55		
33	468	M	121.7	8.93	54.6	5.24	38.3	2.28	176.34	32.3	2.81				1.89
34	469	T	113.7	9.03	60.3	4.73	71.4	4.70	175.79	20.6	0.47				
35	470	T	115.6	9.07	54.4	4.65	67.6	4.25	176.47	23.4	0.92				
36	471	K	120.7	8.03	60.1	3.87	32.4	1.83	179.40	25.3	1.40	29.4	1.69	42.1	2.94
37	472	D	119.1	7.83	54.8	4.14	41.5	3.10	173.87						
38	473	L	120.1	8.03	58.0	3.88	43.0	HB2 1.81 HB1 2.03	178.42	27.3	1.58		0.99		
39	474	L	119.1	8.40	58.3	3.97	41.6	1.73	179.44	27.2	1.73		0.85		
40	475	K	117.0	7.72	58.7	4.02	32.4	1.87	178.66	25.5	1.48	29.4	1.66	42.1	2.93
41	476	K	115.4	7.28	55.6	3.89	31.2	1.38	177.31	23.1	1.08	27.4	1.24	42.0	2.44
42	477	F	117.6	7.51	56.6	4.66	38.6	3.18	175.45						
43	478	Q	121.3	8.15	56.3		29.7	2.68	175.30		2.86				
44	479	T	117.9	8.52	63.5	4.70	66.2	4.27	176.44	21.6					
45	480	K	120.7	8.37	57.6	4.29	32.6	1.88	177.18	25.1	1.46	29.2	1.68	42.1	2.99
46	481	K	117.9	8.00	56.8	4.43	33.1	1.93	178.63	25.4	1.47	29.0	1.71	42.2	2.99

#	RAP74 #	AA	N (± 0.2) ^a	HN (± 0.1) ^a	C α (± 0.2) ^b	H α (± 0.2) ^c	C β (± 0.2) ^b	H β (± 0.2) ^c	C' (± 0.3) ^d	C γ (± 0.2) ^b	H γ (± 0.2) ^c	C δ (± 0.2) ^b	H δ (± 0.2) ^c	C ϵ (± 0.2) ^b	H ϵ (± 0.2) ^c
47	482	T	110.3	8.09	62.1	4.52	70.3	4.30	176.60	21.0	1.25				
48	483	G	111.3	8.40	45.3	3.96			174.09						
49	484	L	121.3	8.12	54.0	4.59	43.7	1.58	177.49	25.3	1.58		0.83		
50	485	S	116.5	8.71	57.3	4.15	64.9	3.90	177.42						
51	486	S	121.2	9.39	62.4	4.15	62.4	3.90	176.07						
52	487	E	121.3	8.67	59.8	3.90	29.4	1.94	178.80	36.5	2.24				
53	488	Q	118.4	7.95	58.6	4.45	29.3	2.11	178.77	32.2	2.60				
54	489	T	117.2	8.28	58.3	3.86	67.7	4.21	179.30	21.3	1.17				
55	490	V	120.0	8.09	67.1	3.51	31.6	2.15	179.41		1.00				
56	491	N	118.1	7.64	56.6	4.45	38.6	2.83	178.10						
57	492	V	120.7	8.18	66.2	3.86	32.5	1.97	176.56		0.90				
58	493	L	120.3	8.44	58.0	3.80	42.0	1.82	176.96	25.2	1.46		0.69		
59	494	A	120.3	8.43	55.6	3.86	18.3	1.51	180.15						
60	495	Q	116.4	7.46	58.7	4.66	28.4	2.60	179.20	33.8	2.69				
61	496	I	120.8	8.06	65.3	3.56	38.6	1.81	178.08		0.79	14.4	0.54		
62	497	L	119.3	8.82	58.4	3.67	41.1	1.75	174.83	27.2	1.56		0.95		
63	498	K	118.6	7.75	59.3	4.00	32.3	1.89	178.85	25.3	1.42	29.4	1.68	42.2	2.98
64	499	R	117.8	7.36	58.7	4.08	30.6	1.96	178.51	27.7	1.76	43.6	3.20		
65	500	L	117.2	8.20	56.2	4.06	43.5	1.86	176.60	26.6	1.36		0.81		
66	501	N	112.6	7.91	53.5		38.2		176.59						
67	502	P			63.5	4.39	32.1	2.31	176.72	27.4	2.01	50.8	3.80		
68	503	E	120.3	8.44	57.0	4.28	30.4	2.21	175.98	36.9	2.36				
69	504	R	123.2	8.42	54.8	5.05	31.9	1.70	175.96	27.7	1.54	44.0	3.22		
70	505	K	123.8	8.89	54.7	4.53	36.2	1.61	174.90	24.4	1.23	29.3	1.53	42.0	2.84
71	506	M	124.6	8.74	54.1	4.90	31.1	1.99	176.16	31.6	2.55				1.87
72	507	I	126.3	8.42	61.2	4.05	40.3	1.06	176.10	CG1 26.5 CG2 16.8	0.65	13.6	0.24		

#	RAP74 #	AA	N (± 0.2) ^a	HN (± 0.1) ^a	C α (± 0.2) ^b	H α (± 0.2) ^c	C β (± 0.2) ^b	H β (± 0.2) ^c	C' (± 0.3) ^d	C γ (± 0.2) ^b	H γ (± 0.2) ^c	C δ (± 0.2) ^b	H δ (± 0.2) ^c	C ϵ (± 0.2) ^b	H ϵ (± 0.2) ^c
73	508	N	126.3	9.39	54.9	4.28	37.2	HB2 2.85 HB1 3.00	174.31						
74	509	D	109.4	8.66	56.2	4.18	40.0	2.89	174.80						
75	510	K	117.8	7.83	54.6	4.77	36.4	1.75	177.10	24.6	1.51	29.1	1.63	42.1	3.06
76	511	M	129.0	9.24	56.9	4.51	31.6	1.91	175.11	31.6	2.08				1.91
77	512	H	125.9	9.30	55.5	4.49	32.3	2.86	173.53						
78	513	F	120.4	9.16	56.2		42.5	2.87	175.79						
79	514	S	113.3	8.77	57.6	4.90	65.9	3.52	175.78						
80	515	L	123.2	7.47	53.3	4.90	46.4	1.58	175.60	27.6	1.32		1.03		
81	516	K	121.7	8.47	56.4	4.34	33.7	1.74	175.60	24.8	1.49	29.2	1.59	42.2	3.01
82	517	E	126.5	8.05	58.1		31.6		175.91						

^a=HSQC^b=C(CC)TOCSY-NNH^c=H(CC)TOCSY-NNH^d=HNCO

CHAPTER 6

CONCLUSIONS

The biochemical and structural studies of the FCP1 domains reported herein have led to several novel discoveries that expand the possible biological roles of FCP1. Amino acid sequence analysis has revealed a conserved acidic and hydrophobic region adjacent to the BRCT domain that had not previously been characterized (Chapter 2). Binding studies with HIV-1 Tat have highlighted the importance of this region as well as the conserved BRCT domain. The HIV-1 Tat binding domain of FCP1 was shown to be extensive, and Tat binding inhibits the binding of the RAP74 subunit of TFIIF. The mutually exclusive interaction that HIV-1 Tat and RAP74 have for binding to FCP1 provides a possible molecular model for the previously described inhibition and stimulation (respectively) of CTD phosphatase activity.

Structural studies linked with binding studies have defined a new LXXLL-like motif within both the central and carboxyl FCP1 domains (Chapter 3). LXXLL motifs are also present within the co-activators and co-repressors of nuclear steroid hormone receptors (Heery et al. 1997). The discovery of these motifs within FCP1 suggests that FCP1 could play a regulatory role in nuclear steroid hormone receptor-mediated transcription. Future *in vivo* studies are needed to investigate the possible biological role of FCP1 and RAP74 in steroid hormone receptor regulation.

The biological role of the BRCT domain is still very mysterious. This domain was initially proposed to be a protein-protein interaction module. It is clear, based on the work presented in this dissertation and elsewhere, that BRCT domains can interact with nucleic acid as

well as protein (Taylor et al. 1998; Soulier and Lowndes 1999; Yamane and Tsuruo 1999). The biological role of the BRCT domain within FCP1 is still unknown; however, studies reported in this dissertation have identified a nuclease activity associated in part with the BRCT domain of FCP1 (chapter 4). This is the first report of an enzymatic activity associated with a BRCT domain within a eukaryotic protein. Sequence homology searches using the minimal region of FCP1 required for nuclease activity *in vitro* have identified a stretch of amino acids (627-696) that are highly homologous to the RNase PH_C domain. The RNase PH_C domain is predicted by sequence annotation to function as an exoribonuclease, however our data reveals that FCP1 cleaves RNA internally. Furthermore, prokaryotic homologues of RNase PH_C domains require phosphate in the catalytic mechanism and yield products that are nucleoside diphosphates. Our initial work characterizing FCP1 activity indicates that this domain may have evolved in eukaryotes utilizing a different mechanism. Preliminary analysis of cleavage products indicates that FCP1 activity does not require phosphate and yields monophosphate nucleosides (unpublished results). Further biochemical studies isolating the reactions products will be necessary to substantiate these findings. Future studies will also be necessary to establish whether FCP1 activity is distributive (on and off the substrate) or processive (locking onto the substrate). *In vivo* functional studies using knockouts and mutants will be necessary to elucidate the role of FCP1 nuclease activity within the cell. Structural information will also provide valuable information on the catalytic mechanism. The presence of amino acids homologous to an RNase domain located within the BRCT domain of FCP1 suggests that a domain fusion event may have occurred in the evolutionary history of FCP1. This finding further expands the possible roles of FCP1 within the cell linking it to RNA metabolism. Several proteins with RNase PH_C domains are components of the exosome, and recent research in this field has linked the activity

of the exosome to transcription through interactions with RNA polymerase II (Andrulis et al. 2002).

The possible roles of FCP1 nuclease activity during transcription are many. FCP1 could be performing a proofreading function. FCP1 binds directly to the RNA polymerase II and therefore would be poised to perform such a function. In fact, the only other reported case of a ribonuclease domain linked to a BRCT domain occurs in the proofreading epsilon subunit (DNAQ) of the replicative DNA polymerase III in *Mycobacterium tuberculosis*. Also, FCP1 could be the predicted and as yet unidentified endonuclease that releases the polymerase following transcription termination (Proudfoot et al. 2002). The fact that HIV-1 Tat inhibits FCP1 nuclease activity may point to a role for FCP1 in transcription termination. Through inhibition of FCP1 nuclease activity Tat would be acting as an anti-terminator allowing transcription of full-length viral transcripts. Another possibility is that the nuclease activity of FCP1 could be acting to enhance the efficiency of transcription by RNAP II. FCP1 binds directly to RNAP II and is known to immunoprecipitate with splicing-associated proteins (Licciardo et al. 2003). Messenger RNA is co-transcriptionally capped and spliced. The resulting RNA products of these reactions could clog the active site of RNAP II and stall transcription. FCP1 nuclease activity may be degrading the RNA by-products of these reactions, allowing enhanced transcription.

The BRCT domain of FCP1 is required for CTD phosphatase activity as well as nuclease activity. This raises the interesting question of whether or not the phosphoesterase and phosphodiesterase activities of FCP1 are related. What activates FCP1 phosphatase activity and FCP1 nuclease activity *in vivo*? *In vivo* studies using wild-type and mutant forms of purified FCP1 will help define the regulation and connection between these activities. The fact that both

Tat and RAP74, two proteins that regulate phosphatase activity *in vitro*, bind directly to the region of FCP1 responsible for nuclease activity supports the notion that these activities are interrelated.

During the course of this study, our lab and others have found that FCP1 is a phosphoprotein. Although the biological significance of FCP1 phosphorylation *in vivo* is not known, phosphorylation of FCP1 *in vitro* enhances RAP74 binding and phosphatase activity. The biochemical and structural studies of the two conserved CK2 sites located within the FCP1 central and carboxyl-terminal domains presented in this thesis have illustrated the importance of this post-translational modification for mediating RAP74 binding. NMR studies are ideal for examining changes in protein-protein interactions upon phosphorylation. $^1\text{H}^{15}\text{N}$ HSQC spectra have highlighted how two phosphorylated peptides from the central and carboxyl-terminal regions of FCP1 bind RAP74. The location of the phosphorylated residue within each FCP1 peptide helps to anchor the helix into the hydrophobic pocket of RAP74. HIV-1 Tat was found to block CK2 phosphorylation at a conserved threonine CK2 site within the central domain of FCP1. This discovery further strengthens the importance of CK2 phosphorylation for mediating FCP1 biological activity, since HIV-1 Tat has been shown to inhibit FCP1 phosphatase activity. HIV-1 Tat also inhibits FCP1 nuclease activity; therefore, phosphorylation may play a role in regulating the nuclease activity of FCP1. Future studies evaluating potential cell cycle-induced or DNA damage-induced changes in FCP1 phosphorylation will further enhance our understanding of how FCP1 phosphatase and nuclease activities are regulated by phosphorylation within the cell.

Although we have met our goal of characterizing the molecular interactions of FCP1 with HIV-1 Tat and RAP74, this study has fostered several future projects. We are currently working

on purifying $^1\text{H}^{15}\text{N}$ labeled samples of the FCP1 central domain for NMR studies while we continue working on the three-dimensional structure of the FCP1 central LXXLL motif peptide complexed to RAP74. Many questions about FCP1 still remain to be answered. What is the role of FCP1 nuclease activity during transcription? Is the FCP1 BRCT domain involved in repair mechanisms or cell cycle regulation as other BRCT domain containing proteins are? Does FCP1 have substrates other than the RNAP II CTD? Are the LXXLL-like motifs of FCP1 involved in nuclear steroid receptor regulation? Is FCP1 phosphorylation by CK2 regulated? Transcription by RNAP II is an intriguing area of study that will captivate me for years to come.

REFERENCES

- Andrulis, E.D., J. Werner, A. Nazarian, H. Erdjument-Bromage, P. Tempst, and J.T. Lis. 2002. The RNA processing exosome is linked to elongating RNA polymerase II in *Drosophila*. *Nature* **420**: 837-41.
- Heery, D.M., E. Kalkhoven, S. Hoare, and M.G. Parker. 1997. A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature* **387**: 733-736.
- Licciardo, P., S. Amente, L. Ruggiero, M. Monti, P. Pucci, L. Lania, and B. Majello. 2003. The FCP1 phosphatase interacts with RNA polymerase II and with MEP50 a component of the methylosome complex involved in the assembly of snRNP. *Nucleic Acids Res.* **31**: 999-1005.
- Proudfoot, N.J., A. Furger, and M.J. Dye. 2002. Integrating mRNA processing with transcription. *Cell* **108**: 501-512.
- Soulier, J. and N.F. Lowndes. 1999. The BRCT domain of the *S. cerevisiae* checkpoint protein Rad9 mediates a Rad9-Rad9 interaction after DNA damage. *Curr. Biol.* **9**: 551-554.
- Taylor, R.M., B. Wickstead, S. Cronin, and K.W. Caldecott. 1998. Role of a BRCT domain in the interaction of DNA ligase III-alpha with the DNA repair protein XRCC1. *Curr. Biol.* **8**: 877-880.
- Yamane, K. and T. Tsuruo. 1999. Conserved BRCT regions of TopBP1 and of the tumor suppressor BRCA1 bind strand breaks and termini of DNA. *Oncogene* **18**: 5194-5203.