This is to certify that the
dissertation entitled

Characterization of the transcriptional coactivator proteins
ADA2 and GCN5 in *Arabidopsis*

presented by

Yaopan Mao

has been accepted towards fulfillment
of the requirements for the

Doctoral degree in Genetics
Cellular and Molecular Biology

Major Professor's Signature

5-15-03

Date

CHARACTERIZATION OF THE TRANSCRIPTIONAL COACTIVATOR PROTEINS ADA2 AND GCN5 IN *ARABIDOPSIS*

By

Yaopan Mao

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Genetics Graduate Program
and
Graduate program in Cell and molecular biology

2003

ABSTRACT

CHARACTERIZATION OF THE TRANSCRIPTIONAL COACTIVATOR PROTEINS ADA2 AND GCN5 IN *ARABIDOPSIS*

By

Yaopan Mao

Chromatin modifications play important roles in regulating transcriptional activation in eukaryotes. The GCN5-ADA2 protein complexes are core histone acetylation modifiers that have been described in yeast and metazoans, where the GCN5 protein is a histone acetyltransferase (HAT). The GCN5-ADA2 complexes are involved in the activation of many genes in these organisms.

This study has been focused on the GCN5 and ADA2 homologous proteins in the plant *Arabidopsis thaliana*. *Arabidopsis* has two ADA2 homologs termed ADA2a and ADA2b and one GCN5 homolog.

The *Arabidopsis*GCN5 is a HAT that acetylates histone H3 primarily. GCN5 interacts with the ADA2 proteins. The regions of each protein involved in GCN5-ADA2 interaction have been defined. The ADA2 proteins enhance GCN5 HAT activity on both free core histones and nucleosomal histones.

The ADA2 proteins interact with the protein CBF1, a cold stress response activator in Arabidopsis. In addition, the interaction domain of CBF1 with the ADA2 proteins has been defined to the DNA binding domain of CBF1. These findings suggest that GCN5-ADA2 complexes may play roles in plant cold acclimation.

The *Arabidopsis*ADA2 proteins can be acetylated by GCN5 *in vitro*. The acetylation site has been identified as lysine 215 in the ADA2b protein. The sequence around lysine

215 in ADA2b is conserved in plant ADA2 proteins and is similar to that of the N-terminal tail of histone H3. Mutant ADA2b that can no longer be acetylated by GCN5 can complement a T-DNA disruption allele of ADA2b for most growth phenotypes. This leaves the significance of ADA2 acetylation an open question.

to Lin Mei

ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Steve Triezenberg for his excellent and diligent guidance and encouragement as well as his patience through my time in his laboratory. His support and help are appreciated with my whole heart forever. I would like to thank my committee members for their support and assistance. Their time and effort are greatly appreciated. I would also like to thank Dr. Michael Thomashow for his advice on my research.

I would like to acknowledge the members of the Triezenberg lab for their companionship, friendship and scientific help and advice. Their participation in creating a friendly and active atmosphere is appreciated very much. I would also like to thank Dr. Kostas Vlachonasios for his help in establishment of transgenic plants.

I would like to thank my wife, family and friends. They have provided a constant source of support through my life. Their support makes me go through all kinds difficulties.

List of Fig

List of tab

List of Sy

Chapter
Introduct
P
A
C
T
M
T
C

Chapt
The Gr

Chap
The

TABLE OF CONTENTS

LIST OF FIGURES

Images in this dissertation are presented in color.

Chapter Two

Chapter Three

Chapter Four

Chapter Five

Chapter Six

Chapter Seven

CONTENTS OF TABLES

SD
PM
AG
BBD
AT
TG
CK
ET
ED
DS-PAGE
AF TBP
FE

LIST OF SYMBOLS AND ABBREVIATIONS

AD	activation domain
bp	base pair
CPM	counting per minute
GAD	Gal4 activation domain
GDBD	Gal4 DNA binding domain
HAT	histone acetyltransferase
IPTG	isopropyl-2-D-thiolgalactopyranoside
kDa	kiloDalton
MET	methyltransferase
OD	optical density
SDS-PAGE	sodium-dedecyl-sulfate polyacrylamide gel dlectropheresis
TAF TBP	associated factor
TFII	transcriptional factor, RNA polymerase II

Chapter One

Introduction and Literature Review

Pol II and the general transcriptional factors

The transcription of protein-coding genes in eukaryotic nuclei is carried out by RNA polymerase II (pol II). Pol II consists of 10-12 subunits, which are well-conserved across species (Hampsey, 1998, Woychik and Hampsey, 2002). One eminent feature of pol II is the carboxyl-terminal domain (CTD) of the largest subunit, a heptapeptide repeat ranging from 26 repeats in yeast to 52 repeats in human. The CTD is under extensive and differential phosphorylation regulation during transcription and thus provides a switch to regulate pol II activity during transcription (Carlson, 1997, Hampsey, 1998, Majello and Napolitano, 2001). For example, hyper-phosphorylation of the CTD accompanies the transition of pol II from transcriptional initiation to elongation, and hypo-phosphorylation of the CTD is associated with the termination of elongation.

The general transcriptional factors (GTFs) of RNA pol II include TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH (Hampsey, 1998, Orphanides, et al., 1996) TFIID can also be categorized as a co-activator (see below). TFIID contains the TATA binding factor TBP, and is the first GTF to bind to promoter to initiate the assembly of transcriptional machinery (Burley and Roeder, 1996). These GTFs are identified through *in vitro* studies of transcription constitution, where the adding of these GTFs enabled pol II to initiate transcription precisely from the promoter in a DNA template (Roeder, 1996).

The GTFs recruit pol II to the core promoter and participate in promoter melting, DNA template unwinding and pol II dispatch from initiation to elongation (Orphanides,

al., 1996).

One form

holoenzyme c

TFIA and TF

I Lee and Y

espond to ac

holoenzyme

This is suppo

ranscription

protein Srb4

1998).

Similar

interestingly

including th

Wilson et

situation in

SMCC TR-

and CRSP

transcriptio

Activators

The transcr

activators b

Transcripti

et al., 1996).

One form of pol II is the so-called holoenzyme. The biochemically purified yeast holoenzyme consists of some GTFs (including TFIIB, TFIIE, TFIIH, TFIIF, but not TFIIA and TFIID) and the Srb/Mediator complex which associates with the CTD of pol II (Lee and Young, 2000). The holoenzyme, unlike the simpler form of pol II, is able to respond to activators to initiate transcription *in vitro* (Koleske and Young, 1994). The holoenzyme is globally required for transcription *in vivo* (Thompson and Young, 1995). This is supported by the microarray data (Holstege, et al., 1998), in which the transcription of more than 93% of the 5361 tested yeast genes were dependent on the Srb protein Srb4. Srb4 is known to be part of pol II holoenzyme exclusively (Myers, et al., 1998).

Similar pol II holoenzymes have been purified from human and murine cells. Interestingly, a broader range of proteins are found in mammalian pol II holoenzyme, including the chromatin modifiers CBP (Neish, et al., 1998), the SWI/SNF complex (Wilson, et al., 1996) and DNA repair proteins (Maldonado, et al., 1996). Like the situation in yeast, human cells have several forms of Srb/Mediator complexes, including SMCC/TRAP (Ito, et al., 1999), ARC (Naar, et al., 1999), DRIP (Rachez, et al., 1999), and CRSP (Ryu, et al., 1999). These complexes are the interacting targets of numerous transcriptional activators and nuclear hormone receptors (Rachez and Freedman, 2001).

Activators

The transcriptional activation of protein-coding genes usually starts with transcriptional activators binding to promoters or/and enhancer elements (Bryant and Ptashne, 2003). Transcriptional activators are typically modular having both a DNA-binding domain,

which lead activators to specific genes, and an activation domain (AD), which activates transcription (Ptashne and Gann, 1997, Triezenberg, 1995). Multiple activators generally coordinate the transcriptional activation of a specific gene (Carey, 1998), and one specific activator can participate in activation of multiple genes.

Coactivators

Co-activators are broadly defined as transcriptional factors that facilitate transcriptional activation but lack cognate DNA binding activities. The studies of eukaryotic protein-coding gene transcription could be roughly divided into two ages: the DNA age, and the chromatin age which began nearly a decade ago. Both ages have uncovered many co-activators. Co-activators uncovered in the DNA age include TFIID and Srb/Med complexes, which interact with the pol II transcriptional machinery. Co-activators found in the chromatin age are chromatin modification/remodeling complexes, such as SWI/SNF and SAGA complexes.

Their functions can be broadly divided into two categories: directing activator recruitment of the transcriptional machinery and chromatin remodeling or modifying (Naar, et al., 2001). This might reflect that chromatin remodeling and general transcription machinery assembly is an integrated business during transcription initiation. Co-activator TFIID and Mediator are discussed in more detail below, and chromatin modifiers are discussed later.

TFIID

TFIID is composed of TBP and about 14 TAFs (for TBP associated factors) (Burley and Roeder, 1996). These components are well conserved across species. The

multifunctional activities of TFIID include core promoter binding, activator interaction, histone acetyltransferase (HAT) activity, kinase activity and ubiquitin ligase activity (Burley and Roeder, 1996).

Although TBP, a component of TFIID, is considered a GTF, the TAF components of TFIID have the characteristic of co-activators (Albright and Tjian, 2000). TFIID binds promoter DNA. TBP binds the TATA element, and some TAFs bind the promoter element of Initiator (INR) and down stream element (DPE). Because of its promoter binding capacity, TFIID is generally believed the first recruited GTF to nucleate the assembly of transcriptional initiation machinery in target genes (Naar, et al., 2001). It has been proposed that TFIID is globally required for gene transcription (Pugh and Tjian, 1992). Recently, Mencia et al proposed that TFIID is required for transcription from promoters with weak or nonexistent TATA elements including most ribosomal protein genes (Mencia, et al., 2002).

The largest TAF in humans ($TAF_{II}250$ or TAF1) has two tandem bromodomains that can bind acetylated histone H4 tail at lysine 5 and 12 *in vitro* (Jacobson, et al., 2000). This suggests that TFIID can bind to, besides DNA, the histone tail at promoters *in vivo*. Remarkably, the acetylation state of histone tails at promoters marks the transcription competence of genes. How TFIID incorporates the acetylation status of promoters through bromodomain binding into transcription regulation is not known.

Not only does TFIID bind to chromatin DNA and histone tails, one component of TFIID has been proposed to mimic DNA structure. The largest TAF inhibits TBP binding to the TATA box *in vitro* (Kokubo, et al., 1998). The N-terminus of the *Drosophila* largest TAF ($TAF_{II}230$) has been found to resemble the minor groove surface of the

TATA box in b

inhibitory effec

a VP16 to rec

The larges

activities *in vit*

tissues *in viv*

which has redu

TAF$_{II}$250, sug

by regulating

The *Dros*

2000). H1 is t

Although the

ubiquitinylati

Pham and Sa

transcriptiona

In vitro d

indicating dir

2001). Direct

protein prom

All toge

promoters, w

histone bindi

may then hel

TATA box in binding to the concave side of TBP protein (Liu, et al., 1998). The auto-inhibitory effect in TFIID is proposed to provide a "hand-off" switch for activators such as VP16 to recruit TFIID in transcriptional activation of genes (Nishikawa, et al., 1997).

The largest TAF in TFIID also has HAT (histone acetyltransferase) and kinase activities *in vitro* (Wassarman and Sauer, 2001). Whether TFIID acetylates promoter histones *in vivo* is unknown, but clearly this HAT activity is important. Mutant $TAF_{II}250$ which has reduced HAT activity can not complement a temperature sensitive mutation of $TAF_{II}250$, suggesting that acetyltransferase activity is required for cell cycle progression by regulating the expression of essential proliferative control genes (Dunphy, et al., 2000).

The *Drosophila* $TAF_{II}230$ mono-ubiquitinylates histone H1 *in vivo* (Pham and Sauer, 2000). H1 is the linker histone which is important for higher order structure of chromatin. Although the function of H1 ubiqintinylation is not known, TAFII250 defective in mono-ubiquitinylation reduces expression of genes targeted by the maternal activator Dorsal (Pham and Sauer, 2000). This suggests the importance of H1 ubiquitinylation in transcriptional activation.

In vitro data have found many transcriptional activators interact with various TAFs, indicating directed recruitment of TFIID in transcriptional activation *in vivo* (Naar, et al., 2001). Direct evidence *in vivo* is the recruitment of TFIID by activator Rap1 to ribosomal protein promoters recently reported by the Struhl group (Mencia, et al., 2002).

All together, a scenario is emerging for TFIID. Activators target TFIID to core promoters, where TFIID binds precisely to specific promoters through its DNA and histone binding capacity. The histone acetylation and ubiquitinylation activities of TFIID may then help to decompress chromatin structure. TFIID then initiates the assembly of

transcriptional machinery.

Mediators

The yeast Mediator was originally isolated biochemically for its ability to stimulate activator-dependent transcription activity in reconstituted transcriptional reactions (Malik and Roeder, 2000). There are three protein groups in the yeast Mediator: Srb proteins, Med proteins, and a diverse group of proteins (Myers and Kornberg, 2000). The Srb proteins are genetically identified as the suppressors of pol II CTD truncation mutants (Myers, et al., 1998). The yeast Mediator exists either as s free complex or associated with pol II through physical interactions with the pol II CTD. Yeast Mediator is generally required for transcriptional regulation *in vivo* (Lee and Young, 2000).

Mediator-like complexes have also been isolated from mammalian cells. Interestingly, mammalian Mediators are more diverse (Myers and Kornberg, 2000, Naar, et al., 2001), including complexes such as SMCC/TRAP, NAT, CRSP, ARC, and DRIP. All these mammalian Mediators share a subset of common components (Rachez and Freedman, 2001). Therefore, they might be the isoforms of a core Mediator dependent on the status of cells or purification methods (Naar, et al., 2001).

Exactly how Mediators work in transcriptional activation is not well known. Yeast Mediator stimulates TFIIH kinase activity up to 30 fold (Kim, et al., 1994). This suggests that Mediator stimulates pol II to escape from promoters and proceed to elongation. Because of their interaction with both activators and pol II, Mediator proteins might act as an interface between activators and the general transcriptional machinery (Myers and Kornberg, 2000, Naar, et al., 2001).

Interestingly, the Mediator protein Nut1 from yeast is found to be able to acetylate

histones, and it belongs to the GCN5-related GNAT HAT family (Lorch, et al., 2000). This might indicate that Mediator might also modify chromatin during the regulation of transcription.

The mechanisms of transcription activation by activators

Activators activate transcription by exploiting the facts that (i) pol II can not initiate transcription automatically from DNA template, (ii) GTFs can guide pol II to genes' promoters to initiate transcription (Woychik and Hampsey, 2002), and (iii) the chromatin generally possesses inhibitory effect on pol II transcription (Struhl, 1999). A few mechanisms are currently understood for activators to work.

First, activators activate transcription by stimulating assembly of transcriptional initiation machinery at core promoters (Ptashne and Gann, 1997). This could be achieved by the stepwise and concerted assembly of GTFs and pol II at core promoters. The stimulated assembly can also be accomplished by one-step loading of pol II holoenzyme at core promoters. It is possible that *in vivo* the activator-stimulated assembly of transcriptional initiation machinery at core promoters is a multiple-step procedure, which is in between the two extremes described above (Lee and Young, 2000).

This mechanism of activation is suggested by the *in vitro* observations that activators interact physically with proteins from GTFs, pol II and the Srb/Mediator (Barberis and Gaudreau, 1998, Burley and Roeder, 1996, Lee and Young, 2000, Ptashne and Gann, 1997). In support of this mechanism, when the components of GTFs, pol II and the Srb/Mediator are artificially brought to the core promoter by fused DNA binding domains, the requirement of activators is no longer necessary (Keaveney and Struhl, 1998). The

tested factors for artificial recruitment have included TBP (Chatterjee and Struhl, 1995, Klages and Strubin, 1995), Gal11 (Farrell, et al., 1996, Gaudreau, et al., 1997), TBP associated factor (TAFIIs) (Gonzalez-Couto, et al., 1997, Keaveney and Struhl, 1998), and TFIIB (Gonzalez-Couto, et al., 1997).

Secondly, activators recruit chromatin-remodeling complexes to facilitate transcription. Various chromatin remodeling complexes, including ATP-dependent nucleosomal modifiers (such as SWI/SNF complex) and histone covalent modifiers (such as SAGA complex) have been reported to be recruited by numerous activators. For example, the HAT complex of SAGA can be recruited by Gal4 (Larschan and Winston, 2001), VP16 (Utley, et al., 1998), Gcn4 (Swanson, et al., 2003), and Pho4 (Barbaric, et al., 2003); the SWI/SNF complex is recruited by Swi5 (Cosma, et al., 1999), EKLF (Kadam, et al., 2000), VP16, Gcn4 and Hap4 (Neely, et al., 1999); and co-activator CBP by nuclear hormone receptors, CREB, E2F, and p53 (Chan and La Thangue, 2001).

The importance of co-activator chromatin modifiers in transcription is strongly supported by two facts: (i) the existing physical interaction between activators and the components of chromatin modifiers, (ii) dysfunctional chromatin modifiers leading to defective transcriptional activation of target genes. For instance, reporter gene transcription from *in vitro* assembled chromatin is dependent on the SAGA complex and acetyl-CoA (Utley, et al., 1998). *In vivo*, the GCN4 dependent activation of *HIS3* in yeast is dependent on SAGA HAT activity (Kuo, et al., 1998, Wang, et al., 1998).

Activators can also use other mechanisms to activate transcription. For instance, the HIV Tat activates transcription by stimulating the processivity of pol II (Garber and Jones, 1999, Jeang, et al., 1999, Karn, 1999). Activators also stimulate promoter escape

and elongation rate (Blau, et al., 1996, Brown, et al., 1998, Krumm, et al., 1995). Activators might also induce changes in structure of co-activators to enable them to activate transcription. For instance, the co-activator ARC becomes a smaller co-activator CRSP when induced by activator VP16 or SREBP-1a, and the induced CRSP is contrastingly different from ARC in three dimensional conformations (Taatjes, et al., 2002). Interestingly, only the CRSP is able to embark on activated transcription while ARC does not (Taatjes, et al., 2002).

These mechanisms described above are not mutually exclusive for a specific activator. For instance, the typical acidic activator VP16 is able to recruit TBP (Hall and Struhl, 2002), histone acetyltransferase complex SAGA (Utley, et al., 1998) and SWI/SNF complex (Neely, et al., 1999).

Many activators can recruit multiplicity of co-activators for transcriptional activation. For instance, the yeast GCN4 recruits as many as seven co-activators to activate transcription of various genes, including co-activator SAGA, SWI/SNF and Mediator (Swanson, et al., 2003). Gcn4 seems able to recruit all these co-activators directly, because direct physical interaction with these co-activators exists (Natarajan, et al., 1999, Swanson, et al., 2003). Interesting, Swanson et al propose that not all recruited co-activators are necessary for a specific promoter, only a subset of co-activators are required. This multitude of recruitment might be explained by the fact that Gcn4 is required for hundreds of gene promoters and for each promoter only some co-activators are needed.

In summary, the action of activators is to influence transcription either thermodynamically by recruitment of transcriptional factors or kinetically by alteration of

the activity of transcriptional factors.

On the other hand, the activation of a specific gene usually makes use of a combination of the mechanisms, where activators bring in chromatin modifiers and the transcriptional initiation machinery in an orchestral manner (Orphanides and Reinberg, 2002). This is well illustrated in the activation of yeast *HO* gene (Cosma, et al., 2001, Cosma, et al., 1999). In the HO promoter, the activator Swi5 recruits SWI/SNF complex first, which then facilitates the recruitment of SAGA histone acetylase complex in the absence of Swi5. The remodeled chromatin is then able to bind a second activator SBF which orchestrates assembly of the transcriptional initiation machinery.

One theme of the concerted recruitment is the requirement of multitude of functions. In fact, many co-activators are multifunctional. For instance, the yeast SAGA complex has at least two activities: chromatin remodeling and TBP binding (Sterner, et al., 1999). Some co-activators such as CBP/p300 possess these functions in one molecule (Goodman and Smolik, 2000). This design, presumably through the fusion of multiple proteins during evolution, simplifies the multiple steps of recruitments.

Chromatin and chromatin modifiers

In the eukaryotic nucleus, the genomic DNA is packaged into a well-organized nucleoprotein structure known as chromatin. The basic building unit of chromatin is the nucleosome, which consists of a core histone octamer (two copies of each of H2A, H2B, H3 and H4) wrapped by 147 base pairs of DNA (Luger, et al., 1997). Although the nucleosome structure has been resolved (Luger, et al., 1997), the higher ordered structure of chromatin is still poorly understood. There are linker histone H1 and many non-histone proteins incorporated into chromatin.

It is

subject

known

ground

nucleos

al., 199

transcri

Ea

termin

These

nucleos

order s

interac

regulat

Genet

al., 19

T

and re

of the

which

covala

phosp

1999,

It is now clear that the chromatin is not just a way of storaging genomic DNA, it is subject to dynamic re-organization and plays crucial role in regulating transcription (Jenuwein and Allis, 2001). The chromatin maintains the transcriptional restrictive ground state by preventing some transcriptional factors from binding (Struhl, 1999). The nucleosome has a general inhibitory effect on transcriptional initiation *in vitro* [(Lorch, et al., 1987, Workman and Roeder, 1987), and loss of nucleosomes *in vivo* increases transcriptional initiation (Han and Grunstein, 1988).

Each histone tail (including the N-termini of all four core histones and the C-terminus of H2A) is strikingly conserved, and very basic in amino acid composition. These tails comprise about 25% of the core histone mass. They are not required for the nucleosome assembly, but they are, together with linker histone H1, critical for higher order structure of chromatin (Wolffe and Hayes, 1999). Indeed, various tail-DNA interactions have been observed *in vitro* for proper chromatin assembly and transcription regulation (Lee and Hayes, 1997, Loyola, et al., 2001, Tse and Hansen, 1997). Genetically, all the four tails are required for basal transcriptional repression (Lenfant, et al., 1996).

The sophisticated organization of chromatin provides a platform for modification and remodeling in regulating transcription. There are two major enzymatic modifications of the highly compact chromatin, (i) ATP-dependent remodeling of the nucleosomes which is carried out by the SWI/SNF like remodelers (Peterson and Workman, 2000), (ii) covalent modification of the histone proteins, including acetylation, methylation, phosphorylation, ADP- and ubiquitinylation (Allen, et al., 1998, Spencer and Davie, 1999).

Histon

ntification

nsequen

ached to

nsequen

2006).

The func

It ha

nscrip

ually a

al., 1988

Nu

nscrip

nscri

of chic

acetyla

he reg

2004).

activa

region

Histone acetylation has been the most extensively investigated among all the histone modifications. Histone acetylation happens in the N-tails of all four core histones. The consequence of histone acetylation is the acetyl moiety from acetyl-CoA is covalently attached to the amide group of lysine residue. This attachment provides the platform for subsequent signaling which eventually leads to the remodeling of targets (Fischle, et al., 2003).

The function of chromatin acetylation

It has been known for decades that histone acetylation has intimate connection with transcriptional regulation (Grunstein, 1997, Struhl, 1998), and acetylated histones are usually a hallmark of transcriptionally active or competent chromatin region (Hebbes, et al., 1988).

Numerous studies have revealed that chromatin acetylation is related to active transcription. Kuo et al illustrated that H3 and H4 are hyper-acetylated at transcriptionally active promoters and enhancers (Kuo, et al., 1998). At the ß-globin locus of chicken and mammalian erythroid cells, there is general background low level of acetylation throughout the active locus, hyper-acetylation of H3 and H4 occurs only at the regulatory site and the actively transcribed genes (Bulger, et al., 2002, Litt, et al., 2001). Histone acetylation also maintains chromatin competent for transcriptional activation. For instance, the chicken β-globin HS4 insulator keeps insulated chromatin region at high acetylation state (Mutskov, et al., 2002).

At promoters, HAT recruitment and hyper-acetylation with transcriptional activation initiation is well documented (Hassan, et al., 2001). For instance, the transcription from a

chromatin template *in vitro* is stimulated by recruited local acetylation (Utley, et al., 1998). *In vivo*, hyperacetylation precedes transcriptional activation at yeast *PHO8* promoter (Reinke, et al., 2001), and *HIS3* promoter (Kuo, et al., 2000).

Acetylation of histones at promoters is generally believed to facilitate the binding of other transcriptional factors, such as GTFs and other chromatin modifiers. *In vitro*, local acetylation by SAGA helps to recruit SWI/SNF complex (Hassan, et al., 2001). *In vivo*, hyperacetylation at promoter improves TBP binding in human estrogen-responsive pS2 promoter (Sewack, et al., 2001), improves the recruitment of transcriptional activator Adr1 in yeast ADH2 promoter (Verdone, et al., 2002), improves the recruitment of CBP-Pol II holoenzyme complex (Agalioti, et al., 2000), and improves the recruitment of SWI/SNF or TFIID (Agalioti, et al., 2002). Interestingly, differential acetylation of histones seems to have different meanings. Agalioti et al have shown that, in the promoter of IFN-beta, acetylation of histone H4 K8 mediates recruitment of the SWI/SNF complex whereas acetylation of K9 and K14 in histone H3 is critical for the recruitment of TFIID (Agalioti, et al., 2002).

Recently, the Horz group found that, during the activation of the *PHO5* gene in yeast, the histones altogether dissociate from the promoter region of *PHO5* gene following hyperacetylation by the recruited SAGA (Reinke, et al., 2001). This finding suggests additional functions of promoter hyperacetylation.

Compared to the hyperacetylation at regulatory elements, the importance of histone hyperacetylation at genes' coding regions is relatively less clear. There is high level genomic wide chromatin acetylation in yeast, and this acetylation is severely affected by mutations in HAT enzymes GCN5 and Esa1 (Kristjuhan, et al., 2002). *In vitro*,

acetylation of H3 and H4 tails relieves chromatin repressive effect on T7 transcription elongation on nucleosomal template (Protacio, et al., 2000). This suggests that the progressive transcription along the coding regions require acetyaltion. A HAT complex known as elongator has been found to associated with elongating pol II in yeast (Wittschieben, et al., 1999), and similar human elongator has been shown to facilitate pol II transcription through chromatin *in vitro* (Kim, et al., 2002). Severe histone H3 hypoacetylation in the coding region can be achieved by double mutation of Gcn5 and Elp3. This hypoacetylation coincides with transcription inhibition, suggesting hyperacetylation in coding regions is important for sufficient elongation (Kristjuhan, et al., 2002).

The GCN5-containing histone modifiers

Numerous histone acetylation modifiers have been identified from budding yeast to human to plants (Roth, et al., 2001, Stockinger, et al., 2001, Tanner, et al., 2000). These histone acetylation proteins and complexes cover a wide range of activity in transcription, such as TFIID which is a basic component of transcriptional machinery, Mediators which are either an independent complex or associated with pol II holoenzyme, and nuclear receptor co-factors such as CBP/p300 which are involved in many aspects of transcriptional activation.

Most of the histone modifiers are highly conserved among species. For instance, CBP/p300 and their HAT activity are conserved from *C. elegans* to human and to plants (Bordoli, et al., 2001, Calvo, et al., 2001). Another feature of these histone modifiers is that they are mostly multi-peptide complexes. The most complicated one might be yeast

Mediator which contains more than 20 subunits (Woychik, 1998).

Based on their acetyltransferase domain sequence similarity and function, five groups of histone acetyltransferases have been described (Roth, et al., 2001, Sterner and Berger, 2000), including CBP/p300, TAF1/TAF_{II}250, ACTR/SRC, the GNAT family, and the MYST family that contains enzymes such as Esa1, Sas2 and Sas3 in yeast.

GCN5 and its homologues such as PCAF form the HAT family of GNAT which is found in all eukaryotes (Roth, et al., 2001, Sterner and Berger, 2000). Recombinant yeast Gcn5 acetylates preferentially histone H3 lysine 14 and H4 lysines 8 and 16, but it has no apparent activity on nucleosomal histones (Kuo, et al., 1996).

There are at least three distinct GCN5-containing complexes in yeast, namely ADA SALSA, and SAGA (Eberharter, et al., 1999, Grant, et al., 1997, Kuo, et al., 1996, Sterner, et al., 2002). The ADA complex (0.8 MDa) contains proteins Ada2, Ada3, Ahc1 and Gcn5. The SAGA complex (for Spt-Ada-Gcn5-acetyltransferase) (1.8 MDa) is composed of four classes of proteins, transcriptional adapter proteins (Ada1, Ada2, Ada3, Ada4/Gcn5, and Ada5/Spt20), Spt proteins (Spt3, Spt7, Spt8), TBP associated factors (Taf_{II}90, -68/61, -60, -25/23, and -20/17), and Tra1. The SALSA complex is very similar to SAGA except it lacks Spt8 and contains a truncated version of Spt7. There is also a forth Gcn5-containing complex reported in yeast called A2 that consist of proteins Ada2, Ada3 and Gcn5 (Sendra, et al., 2000).

Recombinant Gcn5 alone can not acetylate nucleosomal histones *in vitro*. But ADA and SAGA complexes are able to acetylate nucleosomal histones. This indicates the functional participation of the additional components in both communication with nucleosome and regulating Gcn5 activity (Balasubramanian, et al., 2002). In these

complexes, Gcn5 acetylates more lysine residues on histone H3, and also acetylates H2B and H4 (Eberharter, et al., 1999, Grant, et al., 1997, Kusch, et al., 2003).

The Ada proteins were originally identified genetically as suppressors of toxicity caused by transcriptional activator VP16 in yeast (Berger, et al., 1992, Pina, et al., 1993). The Ada2 protein from yeast and metazoans has two conserved domains at its N-terminus, the ZZ finger domain and the SANT/Myb domain. The SANT domain is also found in proteins Swi3, TFIIIB and nuclear co-repressor NCo-R (Aasland, et al., 1996). The SANT domain is required for Ada2 to support transcriptional activation (Candau, et al., 1997) and recent findings have indicated that SANT domain is essential for Gcn5 acetylation activity on nucleosomal histones (Boyer, et al., 2002, Sterner, et al., 2002). In yeast, the ZZ finger is essential for interaction with Ada3 and Gcn5 *in vitro* (Candau, et al., 1997).

The Spt proteins were originally isolated in genetic screens for mutants that compensate for the defects caused by transposon element insertion at yeast promoters. They are functionally connected with TBP, and Spt3 interacts with TBP (Fassler and Winston, 1988, Winston and Sudarsanam, 1998).

The largest protein in SAGA is the Tra1 protein which has more than 2000 amino acids. Protein Tra1 belongs to ATM (ataxia telangiectasia mutated)/DNA-PK/phosphatidylinositol 3-kinase family including human TRRAP protein, implicated in cell cycle checkpoint signaling and cellular response to DNA damage (Abraham, 2001, Grant and Berger, 1999, Khanna, et al., 2001, Vassilev, et al., 1998).

There are at least two functions associated with SAGA: the HAT activity to acetylate histones and the interaction with TBP to regulate the assembly of transcriptional

machinery. Genetic and biochemical analyses in yeast also reveal distinct roles of SAGA proteins (Berger, et al., 1992, Grant, et al., 1997, Sterner, et al., 1999). Null mutation in *ADA1*, *SPT7* or *ADA5/SPT20* result in severe growth and transcription phenotype with loss of SAGA complex, while null mutations in *GCN5*, *ADA2*, *ADA3*, *SPT3* and *SPT8* lead to moderate phenotype and the SAGA complex still remain intact. Loss of both Gcn5/Spt3 or Gcn5/Spt8 produces similar phenotype as that of Ada1, Spt7 or Spt20. These results demonstrate that there are two distinct protein groups in SAGA complex.

Similar SAGA complexes have been purified from human cells and from *Drosophila* (Kusch, et al., 2003). The human complexes include PCAF, STAGA and TFTC (Brand, et al., 2001, Martinez, et al., 2001, Ogryzko, et al., 1998). The situation of these complexes in higher eukaryotes is more complicated, indicating more cellular functions carried out by the complexes. First, human has two GCN5 homologues, GCN5 and PCAF. Human GCN5 and PCAF also have intimate association with CBP/p300 independently. Secondly, two ADA2 proteins are found in human, mouse, *Drosophila*, as well as plant *Arabidopsis* (Kusch, et al., 2003, Muratoglu, et al., 2003, Stockinger, et al., 2001). In *Drosophila*, only one ADA2 is a component of SAGA complex. The other ADA2 variant, ADA2a, can associate with GCN5 and ADA3 (Kusch, et al., 2003). More intriguing, the ADA2a protein in *Drosophila* is part of GCN5-independent complexes, which are concentrated at transcriptionally active regions in polytene chromosomes (Kusch, et al., 2003). This suggests that the two ADA2s have distinct function in higher eukaryotes.

This is also revealed by the distinct effect of mutations in the two *Arabidopsis* ADA2 genes. ADA2b-null plants have severe phenotype whereas ADA2a-null plants seem not to have any defects (Vlachonasios, et al., 2003).

The functional divergence among SAGA-like complexes is also reflected by the observations that PCAF is dispensable for mouse development, while GCN5 null mutation leads to death early in embryogenesis (Xu, et al., 2000, Yamauchi, et al., 2000).

Histone methylation

The role of histone methylation in transcription regulation has emerged as an important issue in the last few years. Like the situation of histone acetylation, this explosion was triggered by the identification of histone methylation enzymes CARM1 (Chen, et al., 1999) and human Suv39h1 (Rea, et al., 2000). The availability of antibodies that specifically recognize methylated lysine or arginine residues also has facilitated the research.

Histone methylation plays important role in both transcriptional activation and repression. Histone methylation is implicated in many biological processes such as transcription, cell cycle regulation, development and even tumor genesis (Kouzarides, 2002, Lachner and Jenuwein, 2002, Zhang and Reinberg, 2001).

In contrast to other histone modifications, there are a few unique features to histone methylation (Zhang and Reinberg, 2001). First, this covalent modification is known to occur on lysine and arginine residues of histone H3 and H4 N-termini, including K4, K9, K27, K36, K79, R2, R17 and R26 in H3, and K20 and R3 in H4. Secondly, the modification can be mono-, di- or tri-methylation on lysine, and mono- or di-methylation on arginine either symmetrically or asymmetrically. How these variant forms of methylation differ in regulating chromatin function is not well known.

Thirdly, the modification is currently believed to be permanent. That is, no enzymes

have been identified that can remove the methyl group from histones. This modification is removed only after DNA replication. This permanency property might explain that histone methylation is intimately associated with long-term epigenetic marks (Jenuwein, 2001). However, recent clues indicate that transient methylation might be possible. Some genes that are rapid turned on and off are controlled partially by promoter methylation, such as the cyclin E gene and nuclear hormone receptor regulated gene pS2 (Bannister, et al., 2002). In addition, The N-terminal domain of yeast HAT enzyme Elp3 is suggested to be a demethylase (Chinenov, 2002), implicating the existence of possible histone de-methylation enzyme.

Fourthly, methylation of histones has dual effects on chromatin. It is connected with both gene activation and repression (see more description below). This is in contrast with histone acetylation that is generally associated with gene activation (Fischle, et al., 2003).

Finally, most histone methyltransferases (HMTs) function as single polypeptide to methylate histones. The only known exception is the Enhancer of Zeste, a Polycomb-group transcriptional repressor, which requires three other proteins to be enzymatically active (Kuzmichev, et al., 2002).

Two groups of HMT enzymes can methylate H3 and H4 tails (Zhang and Reinberg, 2001). Group one methylates arginine residues, including enzymes PRMT1, PRMT2, etc. Group two methylates lysine residues.

In group one, the best-characterized are PRMT1 and PRMT4 (also known as CARM1) (Schurter, et al., 2001). PRMT1 is a nuclear hormone receptor co-activator. It specifically methylates histone H4 R3 both *in vitro* and *in vivo* (Strahl, et al., 1999, Wang, et al., 2001). PRMT1 stimulates transcription from chromatin in *Xenopus* oocyte system

a HMT depe

ranscription (V

actor YY1 to

CARM1 i

eceptor signa

nd 26 (Schur

Illustrated rec

H3 acetylatio

promoter wa

uggesting th

The sec

share a conv

particular hi

SET domai

2002) distir

SET domai

and *Droso*

least SET

Trithorax

SET2 fam

including

The

methylati

in a HMT dependent manner, directly illustrating the involvement of R3 methylation in transcription (Wang, et al., 2001). PRMT1 can be recruited directly by transcriptional factor YY1 to facilitate transcriptional activation (Rezai-Zadeh, et al., 2003).

CARM1 is associated with GRIP, a p160 family co-activator in nuclear hormone receptor signaling (Chevillard-Briet, et al., 2002). *In vitro*, CARM1 methylates H3 R2, 17 and 26 (Schurter, et al., 2001). Direct *in vivo* methylation of R17 by CARM1 was illustrated recently (Daujat, et al., 2002). Interestingly, this methylation depends on prior H3 acetylation by CBP. In the activation of pS2 gene, histone H3 R17 methylation at the promoter was observed to coincide with the association of CARM1 (Bauer, et al., 2002), suggesting that CARM1 methylation of H3 R17 plays roles in transcriptional activation.

The second group of HMTs methylates specifically lysine residues. These enzymes share a common motif termed the SET domain. The specificity of these enzymes for particular histone substrates is suggested to be dependent on the pre-SET or/and post-SET domain (Rea, et al., 2000, Zhang and Reinberg, 2001). Kouzarides (Kouzarides, 2002) distinguished four families in this group based on the detailed sequences of the SET domains. (i) The Suv39 family, including human Suv39h1 and fission yeast Clr4 and *Drosophila* Su(var)3-9, is able to methylate H3 K9. (ii) The SET1 family, including yeast SET1, Polycomb proteins EZH1 and EZH2 (containing SANT domain), and Trithorax protein MLL (containing bromodomain), is able to methylate H3 K4. (iii) The SET2 family, including *Drosophila* ASH1, is able to methylate H3. (iv)The RIZ family, including RIZ, BLIMP and PFM1, does not contain pre-SET and post-SET domains.

The best characterized methylated lysine residues are histone H3 K4 and K9. K9 methylation is generally associated with transcriptional repression while K4 methylation

s generally a

Mamma

ediate the r

HP1, a prote

this specifi

HP1 also bir

heterochrom

the homolo

model Jenu

more Suv39

heterochron

process for

H3 K9

is known to

recruited to

showed tha

and HP1 to

nucleosom

in euchrom

sites are ob

range of us

In con

transcriptio

is generally associated with transcriptional activation.

Mammalian HMT Suv39h1 specifically methylates H3 K9. The HP1 protein may mediate the repression of K9 methylation in heterochromatin (Jenuwein and Allis, 2001). HP1, a protein universally associated with heterochromatin (Eissenberg and Elgin, 2000), binds specifically to H3 methylated at K9 (Bannister, et al., 2001, Lachner, et al., 2001). HP1 also binds with Suv39h1, and null Suv39h1 leads to the disassociation of HP1 from heterochromatin (Lachner, et al., 2001). In fission yeast, crippled HMT activity of CLR4 (the homologue of mammalian Suv39) results in decreased HP1 in heterochromatin. On model (Jenuwein and Allis, 2001) proposes that initial K9 methylation brings HP1, and more Svu39h1 and HP1 are recruited until this propagation creates locally repressive heterochromatin. The chromodomain of HP1 is probably play important role in this process for it binds specifically to H3 tail with methylated K9 (Lachner, et al., 2001).

H3 K9 methylation is also evident in local repression of euchromatic genes. Suv39h1 is known to cooperate with Rb protein to repress E2F activity and Suv39h1 could be recruited to E2F1 through its interaction with Rb (Vandel, et al., 2001). Neilsen el al showed that Rb, which is recruited by E2F1 at the cyclin E promoter, recruits Suv39h1 and HP1 to repress transcription of cyclin E, and that the local K9 methylation (only one nucleosome) is Rb dependent (Nielsen, et al., 2001). The involvement of HP1 repression in euchromatic genes probably is not restricted to Rb regulated promoters, for many HP1 sites are observed in *Drosophila* euchromatin (Kouzarides, 2002). This suggests a wider range of usage of HP1 in gene repression.

In contrast to H3 K9 methylation, H3 K4 methylation is associated with transcriptionally active chromatin (Lachner and Jenuwein, 2002) (Jenuwein and Allis,

2001). H3 K4 methylation is highly conserved among Tetrahymena, yeast and human, and H3 K4 methylation in Tetrahymena is associated with active transcription (Strahl, et al., 1999). At the mating-type locus (47 kb) in fission yeast, H3 K4 methylation is strictly found in active euchromatin, while H3 K9 methylation is highly enriched in flanking silent heterochromatin (Noma and Grewal, 2002). In another case of the chicken ß-globin locus and its flanking loci, H3 K9 methylation is associated with constitutive condensed chromatin and inactive globin genes during erythropoiesis, while H4 K4 methylation correlates with active chromatin which is hyper-acetylated (Litt, et al., 2001).

But this rule is not universal. H3 K4 methylation by Set1 in budding yeast is required for rDNA silencing (Briggs, et al., 2002, Bryk, et al., 2002). On the other hand, K9 and K4 methylation are not exclusive from each other in chromatin. For instance, Ash1, the epigenetic activator of the Trithorax group, methylates H3 K4 and K9, and H4 K20 simultaneously. Transcriptional activation by Ash1 coincides with the methylation of these three lysine residues at the promoter of Ash1 target genes (Beisel, et al., 2002). This reveals that combinatorial K methylations might have different meanings.

Recently, H3 K79 was also identified to be methylated by the enzyme Dot1 (Ng, et al., 2002). Both the enzyme and K79 methylation are found in yeast and human cells. Curiously, Dot1 does not have a SET domain, and thus represents a distinct class of lysine methylation enzymes. K79 methylation is important for telomere and mating-type silencing in yeast, but K79 hyper-methylation is also a mark of euchromatin in both yeast and mammalian cells (Ng, et al., 2003), suggesting complicated roles played by K79 methylation. Interestingly, efficient K79 methylation is dependent on ubiquitinylation of H2B K123 (Briggs, et al., 2002, Ng, et al., 2002).

In summa

markers are re

connected to h

Histone phos

Histone p

characterized

chromosome

as well as for

et al., 1999).

upon ionizin

phosphoryla

(Cheung., et

Proteins

mammalian

al., 2001). I

with HAT G

S10 by Snf1

happen in a

Histone co

At leas

ubiquitinyla

In summary, distinct patterns of methylation have different meaning. But how these markers are read differently is still not clear. In addition, histone methylation is closely connected to histone acetylation (see below).

Histone phosphorylation

Histone phosphorylation, compared to histone acetylation and methylation, is less characterized (Cheung, et al., 2000). The H3 S10 phosphorylation is required for chromosome condensation and segregation during mitosis and meiosis (Wei, et al., 1998), as well as for transcriptional activation for certain genes (Lo, et al., 2000, Sassone-Corsi, et al., 1999). The phosphorylation of H2A S139 is connected with DNA break repair upon ionizing radiation and apoptosis (Thomson, et al., 2001). The H2B S14 is also phosphorylated during apoptosis and the kinase identified (Mst1) is caspase activated (Cheung, et al., 2003).

Proteins Rsk and Msk have been identified as potential H3 S10 kinases in mammalian cells in response to mitogen stimuli (Sassone-Corsi, et al., 1999, Thomson, et al., 2001). In yeast, Snf1 was identified as an H3 S10 kinase (Lo, et al., 2001). Together with HAT GCN5, this kinase regulates gene *INO*1 expression. The phosphorylation of S10 by Snf1 and acetylation of K14 by GCN5 on histone H3 at the *INO1* promoter happen in a sequential manner (Lo, et al., 2001).

Histone code and the cross talk of histone tail modification

At least five histone tail modifications are known, including acetylation, methylation, ubiquitinylation, phosphorylation and ADP-ribosylation. There can be many covalent

attachment patterns on single histone tail, on single nucleosome, and on nucleosomal clusters in a sequential and combinatorial manner. The histone code hypothesis (Strahl and Allis, 2000, Turner, 2002) proposes that various modifications exert different downstream functions by either recruiting various proteins or affecting chromatin structure, thereby to execute various cellular functions such as transcriptional regulation (Turner, 2002).

Based on the histone code concept, one specific modification may have different functions, dependent on the modification context. One example is H3 S10 phosphorylation (Strahl and Allis, 2000, Turner, 2002). This S10 phosphorylation plays dual roles: it is required for chromatin condensation during mitosis as well as for certain immediate early genes activation in response to mitogen stimulus. The coupled acetylation of H3 K14 might make S10 phosphorylation a gene activation marker (Cheung, et al., 2000).

Another prediction is that various proteins or complexes bind to specific patterns of modifications and carry out the specific downstream functions. There are a few known proteins or domains that bind to particular covalent tags on histone tails. The chromodomain of HP1 binds to K9 methylated H3 tail as described above (Jacobs and Khorasanizadeh, 2002). Bromodomains bind acetylated H3 and H4 tails (Dhalluin, et al., 1999). Bromodomains are found in many transcriptional factors and complexes, such as SAGA, TAF250, CBP/p300, SWI/SNF, and RSC (Jeanmougin, et al., 1997), as well as some histone methylases (Kouzarides, 2002). One function of bromodomains in binding to acetylated H3 and H4 tails might be to keep HAT complexes at local promoters in order to maintain local hyperacetylation status. Hassen et al showed that *in vitro* retention

of SAGA and SWI/SNF on acetylated nucleosome template is dependent on their bromodomain, and this is also true *in vivo* at promoters targeted by SAGA and SWI/SNF (Hassan, et al., 2001).

The double bromodomain in human $TAF_{II}250$ binds preferentially to H4 with a specific acetylation pattern that is most frequently found in actively transcribed genes (Jacobson, et al., 2000). Crystallographic information shows that the two binding pockets of the tandem bromodomains prefer acetylated lysine residues that are seven amino acids apart. In fact, these two bromodomains in $TAF_{II}250$ bind best to H4 tails with double methylations of either K5-K12 or K8-K16. This case well exemplifies the histone code idea that specific patterns of modification are recognized by specific proteins.

The histone code also predicts that one modification can inhibit or enhance the other modification. For example, NuRD is a histone deacetylase complex associated with transcription repression. Its preferential binding to the H3 tail is inhibited by K4 methylation but not by K9 methylation (Nishioka and Reinberg, 2001, Zegerman, et al., 2002). H2B K123 ubiquitinylation in budding yeast is a prerequisite for H3 K4 methylation. H3 K4 methylation is abolished in the H2B K123R mutant (Dover, et al., 2002, Nishioka and Reinberg, 2001, Sun and Allis, 2002, Zegerman, et al., 2002). The H2B K123R mutation perturbs silencing at the telomere, illustrating functional links between Rad6-mediated H2B (K123) ubiquitinylation, SET1-mediated H3 (Lys 4) methylation, and transcriptional silencing. H2B K123 ubiquitinylation is also required for efficient methylation of H3 K79 (Briggs, et al., 2002, Ng, et al., 2002). Mutation of H2B K123R also results in severe loss of H3 K79 methylation. Interestingly, these two sites (H3 K79 and H2B K123) are proximate to each other in the nucleosome. Both K4 and

K79 methylation are involved in gene silencing. H2B ubiquitination might act as a switch that controls the site-selective histone methylation patterns responsible for this silencing (Briggs, et al., 2002).

Besides cross talk *in trans* as described above, there are also intra-tail communications in histone modifications. H3 K14 acetylation by GCN5, PCAF and p300 is facilitated by prior S10 phosphorylation (Cheung, et al., 2000, Lo, et al., 2000). The crystal structure reveals close contact between the arginine residue (R164) near the yeast Gcn5 reaction center and H3 S10. Indeed, the mutation of Gcn5 (R164A) resulted in acetylation of H3 K14 independent of S10-phosphorylation, and this mutation led to decreased Gcn5 acetylation activity in target promoters (Lo, et al., 2000).

This meaningful cross talk between histone tail modifications *in cis* has also been described in a few more cases. Upon epidermal growth factor (EGF) treatment in mammalian cells, H3 S10 phosphorylation precedes K14 acetylation and this di-modification is tightly associated with target promoters after EGF treatment (Cheung, et al., 2000). In yeast *INO*1 promoter, as mentioned above Snf1 phosphorylation at H3 S10 and Gcn5 acetylation at H3 K14 occur in a strict sequential manner to enhance transcription of *INO1* gene (Lo, et al., 2001). Coupled H3 S10 phosphorylation and H3 K14 acetylation is also seen in thyroid hormone receptor (TR) targets where reduction in both S10 phosphorylation and K14 acetylation coincides with ligand-unbound TR status whereas increase of both modifications coincides with ligand-bound state, indicating these coupled-modifications are a mark of active gene (Li, et al., 2002).

Not only does H3 S10 phosphorylation cross talk with K14 acetylation, but also with lysine methylation. H3 S10 phosphorylation inhibits H3 K9 methylation by Suv39h1 *in*

vitro, and *vice versa* (Rea, et al., 2000).

Instances of cross-talk between histone acetylation and methylation have also been reported. *In vitro*, H3 K9 and K14 acetylation improves H3 K4 methylation by MLL, the human homolog of *Drosophila* trithorax which is a SET-domain-containing HMT (Milne, et al., 2002). This linked modification might also occur *in vivo*, since histones at MLL regulated promoters have these modifications. In another case, H4 R3 methylation by PRMT1 improves H4 K8 and K12 acetylation by p300, but acetylated H4 inhibits R3 methylation by PRMT1 (Wang, et al., 2001).

H3 K4 methylation and K9 methylation were also reported to be mutually inhibitory (Wang, et al., 2001). Methylated H3 at K4 improves subsequent p300 acetylation on H3, while H3 K9 methylation inhibits H3 K14, K18 and K23 acetylation by p300. As discussed in histone methylation section, H3 K4 methylation is associated with active genes while H3 K9 methylation with repressive chromatin, and histone acetylation by p300 is well documented with gene activation (Sterner and Berger, 2000), indicating that H3 K4 and K9 methylation have very different functions.

Moreover, the cross-talk *in cis* between histone acetylation and arginine methylation was also reported recently (Daujat, et al., 2002). After estrogen treatment, p300 was observed at the pS2 promoter and H3 K18 was acetylated, followed by H3 K23 acetylation. CARM1 was then recruited, leading to the methylation of R17. This sequential acetylation and methylation was abolished if CBP HAT activity was crippled, clearly demonstrating that H3 K18 and K23 acetylation is prerequisite for R17 methylation.

Histone modifications can also crosstalk to pol II. Two groups report that H3 K36

methylation was dependent on pol II CTD phosphorylation in yeast (Li, et al., 2002, Xiao, et al., 2003). Partial deletion or CTD kinase knock-out resulted in total abolition of K36 methylation, whereas H3 K4 and K79 methylation is not affected.

In summary, the covalent tags on chromatin histones establish various patterns of chromatin modification that may serve as marks for specific downstream functions in regulating genes' expression. These patterns can be dynamic depends on multiple cellular stimuli and the cross talks between these modifications. How these specific modifications are interpreted in detail is the subject of intense study currently.

The plant world

Most knowledge about the regulation of transcription activation has been learned from the studies of yeast and metazoans. Relatively less has been known in the plant kingdom.

The genome sequence of *Arabidopsis thaliana* [the *Arabidopsis* initiative, Nature 408, 796 - 815 (2000)] reveals that the genome encodes nuclear DNA-dependent RNA polymerase I, II and III that are typical of eukaryotes. This suggests that plants use the common mechanisms of gene transcription.

Plants, like yeast and metazoans, also have various chromatin remodeling modification factors (Li, et al., 2002) and some of these factors have been found to play crucial roles in plant gene expression (Verbsky and Richards, 2001), such as Snf2-type chromatin remodeling ATPases, histone acetyltransferase that include the GNAT family, the MYST family, and TAF250 or CBP-like proteins, and histone deacetylases that include Rpd3 or Sir2-like proteins. *Arabidopsis* CBP-like protein PCAT2 has been shown

to be able to acetylate histones *in vitro* although it lacks the bromodomain (Bordoli, et al., 2001). *Arabidopsis* Snf5-like protein BSH can complement partially the *snf5* mutant in budding yeast (Brzeski, et al., 1999), and the Swi2-like protein MOM has been found important in gene silencing (Amedeo, et al., 2000). Meanwhile, limited studies available todate indicate that chromatin modification plays important roles in regulating gene expression in plants. Hyperacetylation of promoter histones of the pea plastocyanin gene correlated directly with gene activation (Chua, et al., 2001). All these findings indicate that plants make use of similar chromatin remodeling mechanisms in regulating gene expression.

But plants, due to their vital difference in life cycle from animals, might have evolved plant-specific mechanisms to regulate gene transcription (Lusser, 2002). For instance, *Arabidopsis* lacks most transcriptional factors for polymerase I, and it has a fourth class of the largest subunit and the second-largest subunit for nuclear RNA polymerase which is not found in other eukaryotes [the *Arabidopsis* initiative]. Moreover, the *Arabidopsis* genome encodes about 1500 transcriptional factors whereas *Drosophila melanogaster* and *C. elegans* have about 650 (Riechmann, et al., 2000). Among these transcriptional factors, at least five families are not found in yeast and metazoans (Riechmann and Ratcliffe, 2000), such as the AP2/EREBP family (~140 members) implicated in flower development, cell proliferation, stress response, ABA and ethylene response, the NAC (~100 members) family implicated in pattern formation and organ development, and the WRKY family (~70 members) implicated in defense response.

In addition, owing to its sessile nature, plants are affected greatly by both biotic and abiotic environmental signals, such as drought, high salt, cold temperature and pathogen

attacks. These aboitic and biotic stresses are the major losses in crop productivity worldwide (Singh, et al., 2002). Transcriptional regulations in plants in response to these stressful conditions take part in vital roles in plants' life (Singh, et al., 2002). This makes the studies of transcription regulation of plants in response to stressful condition very significant.

One example is the transcriptional regulation in plants in response to cold temperature. In *Arabidopsis*, many cold regulated (COR) genes are transcriptionally activated that enable the plant to accommodate cold temperature (Thomashow, 1998). One key transcriptional activator of COR genes CBF1, belongs to the AP2/EREBP family of to plant transcriptional factors (Stockinger, et al., 1997). Constitutive expression of CBF1 induced COR gene expression and increased the freezing tolerance of non-acclimated *Arabidopsis* plants (Jaglo-Ottosen, et al., 1998). Very intriguingly, the transcriptional activity of CBF1 was found to depend on the Ada2, Ada3 and Gcn5 proteins in yeast (Stockinger, et al., 2001).

This initial finding of CBF1 dependence on yeast co-activator proteins led to the hypothesis that *Arabidopsis* has similar Ada2 and Gcn5 proteins and that these proteins might modify chromatin in facilitating plant COR genes' activation. In addition, the limited knowledge about plant histone acetyltransferases, the indicated importance of transcriptional regulation in plant life, as well as the indications of the divergence of transcriptional activation in plants altogether led to this thesis study of *Arabidopsis* ADA2 and GCN5 homologous proteins.

Chapter Two

The GCN5 Gene and protein of *Arabidopsis*

Introduction

The Gcn5 protein was first identified genetically from budding yeast. Under the growth condition of amino acid starvation, yeast genes related to amino acid synthesis are turned on by activator Gcn4. The efficient activation activity of Gcn4 required protein Gcn5 (Georgakopoulos, et al., 1995). Gcn5 was also required by chimerical activator Gal4-VP16 for efficient activity in yeast (Marcus, et al., 1994). Mutation in GCN5 reduced the transcriptional activity of Gcn4 and Gal4-VP16, suggesting that Gcn5 might function in facilitating transcription ad co-activator protein. Meanwhile, Gcn5 was found to function together with protein Ada2 and Ada3 as a unit. Budding yeast with mutation in GCN5 showed similar phenotypes as the mutation in ADA2 or ADA3 as well as double mutations in ADA2 and GCN5, and each mutation could relieve the toxicity caused by the over-expression of Gal4-VP16 (Berger, et al., 1992, Georgakopoulos, et al., 1995).

The mechanism of yeast Gcn5 in facilitating transcription remained elusive until the milestone discovery in 1996 that the yeast Gcn5 was found to be a histone acetyltransferase (HAT) that acetylated histone H3 (Brownell, et al., 1996). Chromatin acetylation had been observed to connect with gene activation (Turner, 1991). This finding shed light on the mechanism of Gcn5 in facilitating transcription by modifying chromatin. The GCN5 protein is conserved in yeast and metazoan with two conserved domains, the catalytic HAT domain and the bromodomain (Xu, et al., 1998).

But very little was known about GCN5 proteins in plant world. A cDNA from plant *Arabidopsis thaliana* was previously cloned at our laboratory that encodes a protein homologous to GCN5 with a conserved HAT domain and bromodomain, as well as a unique N-terminus (200 amino acids). Recombinant GCN5 protein from Tetrahymena, yeast and human can acetylate free histone H3 *in vitro* (Brownell, et al., 1996, Xu, et al., 1998, Yang, et al., 1996). It was fundamental for us to determine whether *Arabidopsis* GCN5 possessed histone acetyltransferase activity *in vitro*.

Research work was also performed to determine the genomic sequence of *Arabidopsis GCN5* gene. It was not clear whether the known N-terminal extension of *Arabidopsis* GCN5 was complete for mammalian GCN5 homologues have longer N-terminus extension (about 400 amino acids), although the predicted N-terminal extension of *Arabidopsis* GCN5 shared no obvious sequence similarity to that of mammalian GCN5 homologues. The GCN5 genomic sequence could also reveal the boundaries of exons and introns of *GCN5* gene that would help in the future plan in our lab of screening *Arabidopsis* lines with T-DNA insertion in the *GCN5* gene.

Experimental Methods

Determining the genomic sequence of *GCN5* gene

A genomic DNA library of *Arabidopsis thaliana* was provided kindly by Dr. Christopher Benning of our department, in which the genomic DNA was digested partially with Hind III and subcloned into cosmic vector pBIC20. The library had an average insert size of 20 kB. The cosmids had been already introduced into *E. coli* strain NM554.

The library was plated on LB medium with 10 μg/ml tetracycline at a density of about 100 colonies per plate (100 x 15 mm). The colonies were transferred to nitrocellulose. Standard colony hybridization procedures were followed using the GCN5 cDNA (full length) as probe. Cosmid DNA prepared from the positive colonies was screened again after digested with Hind III by standard Southern blotting procedure using GCN5 (full length) cDNA probe. Two positive DNA fragments were identified and sub-cloned to vector pBS (KS+). The two fragments were DNA sequenced by the sequencing center at Michigan State University using various primers that ensured that every DNA region was read in both 5' and 3' directions.

Arabidopsis suspension cell lysate

Arabidopsis leaf suspension cell culture was originally provided by Dr. Natasha Raikhel's laboratory from Plant Research Laboratory at Michigan State University. The cell culture was maintained in Gamborg's B-5 (GibcoBRL) liquid medium with 1.8 μg/ml 2, 4-dichloro-phenoxyacetic acid. The suspension cells were passaged every 10 days in 50 ml medium at 25°C and 24-hour light cycle.

To make cell lysates, the cells were resuspended in Tris buffer (50 mM Tris pH8.0, 1 mM EDTA, 100 mM KCl, 10% glycerol, 2 mM MgCl2, 5 mM DTT) with protease inhibitor cocktail (Roche) and 1mM PMSF. The cells were treated with French Press once. The broken cells were either used directly as "crude lysate", or centrifuged at 7,600 x g for 15 min, and the supernatant was collected as "supernatant cell lysate".

Western blotting

Two rabbit antisera were used. Antiserum #84-6 was raised with GCN5 peptide antigen (amino acids 383-400) and the antiserum was collected after boosting for the sixth time. Antiserum #91-3 was raised with recombinant GCN5 antigen (amino acids 155-568) and the antiserum was collected after boosting for the third time.

1μl cell lysate or supernatant cell lysate was resolved in 7.5% SDS-PAGE gel (20 mA current for 3 hours), and then transferred to nitrocellulose membrane. The membrane was incubated with antiserum #84-6 (1:10000 dilutions) or #91-3 (1:10000 dilutions). Then the membrane was incubated with goat against rabbit second antibody (1:4000 dilutions, BioRad company, affinity purified and HRP conjugated). The signal was detected using Lumi-light Western Blotting Substrate kit (Roche).

Purification of recombinant GCN5

For purifying His6-GCN5 (155-568), cDNA fragments encoding GCN5 (155–568) was subcloned into His6-tagged vector pET-28c (Novagen), and the constructed plasmid was transformed into *E. coli* strain BL21(DE3). 10 ml overnight culture of transformed *E. coli* was transferred to 500 ml LB medium with 50 μg/ml kanamycin, and the culture was induced with 90 μM IPTG for 16 hours at room temperature. The cells were collected by centrifugation and washed once with Na-PO_4 buffer (50 mM Sodium phosphate, pH7.8, 500 mM NaCl, 10% glycerol and 5 mM 2-mercaptoethanol).

The cell pellet was resuspended in 15 ml Na-PO_4 buffer with protease inhibitor cocktail and 12 mM imidazole, and treated with French Press once to make cell lysate. The supernatant was collected from the lysate after centrifugation at 20,000 x g for 20 minutes. Appropriate amount of lysate supernatant was mixed with Ni-NTA beads

(Qiagen) at cold room for 1 hour, and then the beads was washed with Na-PO_4 buffer with 0.1% NP-40 for several times. The beads were washed again with Na-PO_4 buffer with 40 mM imidazole once. The His6-GCN5 (155-568) was eluted with Na-PO_4 buffer with 250 mM imidazole.

For purification of recombinant GCN5 (1-568, 205-568, or 205-384), DNA fragments encoding the respective GCN5 proteins were subcloned into GST-tagged vector pGEX-6p (Pharmacia) and transformed into *E. coli* strain BL21(DE3) Codon Plus. 10 ml overnight culture of transformed *E. coli* was transferred to 1000 ml LB culture, and the culture was induced with 70 μM IPTG for about 14 hours at room temperature. The cells were collected by centrifugation and washed once with PBS buffer (140 mM NaCl, 2.7 mM KCl, pH7.3, 10 mM Na_2PO4, 1.8 mM KH_2PO4, 10% glycerol, and 7 mM 2-mercaptoethanol).

The cell pellet was resuspended in about 20 ml PBS buffer with protease inhibitor cocktail and 1 mM EDTA, and treated with French Press once to make lysate. The supernatant was collected from the lysate after centrifugation at 20,000 x g for 20 minutes. Appropriate amount of lysate supernatant was mixed with GST Glutathione Sepharose beads at cold room for 1 hour, and then the beads were washed with PBS buffer for several times. To recover GCN5 proteins from the beads, the beads were cut with PreScision (Pharmacia) protease at cold room for 4 hours and the supernatant was collected.

HAT assay

The HAT assays were carried out in 30 μl reactions containing 50 mM Tris–HCl pH

8.0, 10% glycerol, 1 mM DTT, 1 mM PMSF, 0.1 mM EDTA, 50 mM NaCl, 10 mM butyric acid, 20 μg calf thymus histones (Sigma), 0.15 or 0.2 μCi [^{3}H]acetyl-CoA (NEN Life Science Products, 0.1 μCi/μl) and about 2.9 μg GCN5(155-568), or 250 ng GCN5(1-568), or 70 ng GCN5(205-568), or 70 ng GCN5(205-384) or 2.8 μg purified recombinant human TAF31 (as a negative control). The reaction mixture was incubated at 27 or 30°C for 30 min or 40 min, after which the reactions were terminated by adding 10 or 20 μl of SDS–PAGE loading buffer. 20 μl from each terminated reaction mixture was electrophoresed on 15% or 18% SDS–polyacrylamide gels. Gels were fixed (40% methanol, 10% acetic acid solution), soaked in autoradiographic enhancer (NEN Life Science Products), vacuum dried and exposed to X-ray film.

Results

The *Arabidopsis* GCN5 protein is a HAT

The predicted amino acid sequence of *Arabidopsis* GCN5 protein based on the cDNA sequence (Stockinger, et al., 2001) suggests that it is a HAT enzyme. An *in vitro* HAT assay was carried out to test directly whether the *Arabidopsis* CGN5 protein is a HAT.

Recombinant GCN5 proteins as either full length or truncated versions were purified from expressed *E. coli* cells. As shown in Figure 1B, relatively pure recombinant proteins of GCN5 were obtained, including GCN5 (1-568, full length), GCN5 (155-568) that still contained part of the unique N-terminus, GCN5 (203-568) that contained both the HAT domain and the bromodomain, and GCN5 (203-384) that contained the HAT domain only.

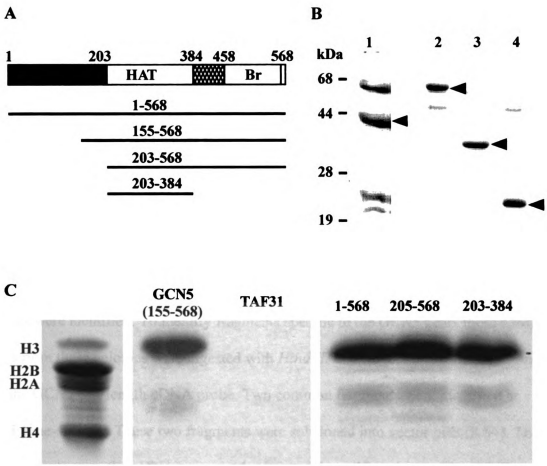


Figure 1. ***Arabidopsis*** **GCN5 protein is a histone acetyltransferase..**
(**A**) Diagram of the GCN5 protein. The predicted HAT domain and bromodomain (Br) are indicated. The bars show the GCN5 proteins in various lengths that were purified as recombinant for HAT assays. (**B**) The purified recombinant proteins of GCN5 in various lengths. The recombinant proteins were electrophoresed in 10% SDS-PAGE gel and were either Coomassie Blue stained (lane 2, 3, and 4) or silver stained (lane 1). The respective GCN5 proteins were indicated by the arrows. 1: His6-GCN5 (155-568), 2: GCN5 (1-568), 3: GCN5 (203-568), 4: GCN5 (203-384). Note that the protein mark applies only to lane 1 and that the unindicated bands in lane 1 were contaminated proteins. (**C**) The results of the HAT assays with recombinant GCN5 proteins in various lengths. The HAT reactions were electrophoresed in SDS-PAGE gel. The histones in gel were first visualized by Coommasie Blue staining as represented on the left panel, then the acetylation signals were detected by enhanced fluorography as shown on the right two panels. Note that the histones between H4 and H2A on the left panel were also histone H4.

The HAT assays were then performed using the four recombinant GCN5 proteins, free core histones and [^{3}H] acetyl CoA. The HAT assay results (Figure 1C) showed that purified recombinant GCN5 proteins were able to acetylate free core histons, and core histone H3 was the major acetylated substrate. Moreover, the HAT domain alone was sufficient to acetylate core histone H3.

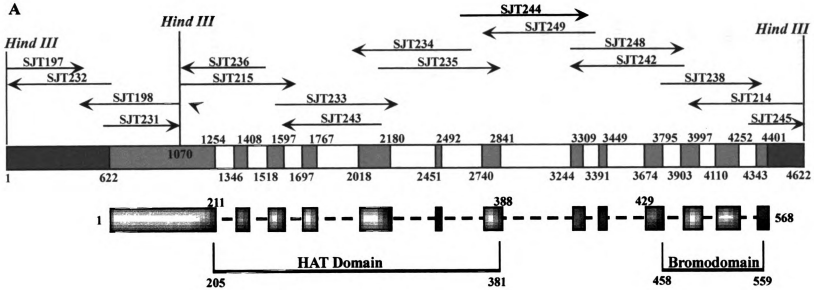
The genomic sequence of *GCN5* gene

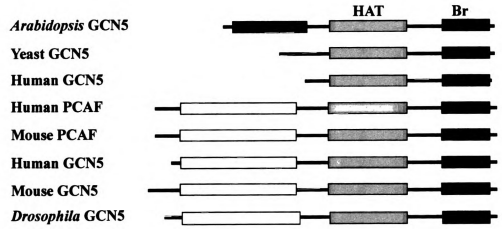
An *Arabidopsis* genomic library (provided by Dr. Christopher Benning) was screened with GCN5 cDNA probe to identify the *GCN5* gene. Seven positive library colonies were identified. To identify fragments specific to the *GCN5* gene, these cosmid DNAs from these colonies were digested with *Hind III* and subject to Southern blotting with the GCN5 full-length cDNA probe. Two common fragments were identified as *GCN5* gene-specific. These two fragments were subcloned into vector pBS (KS+). The two fragments were then DNA sequenced using serial specific primers. The *GCN5* gene sequence was then assembled using the computer program Sequencher. The genomic sequence of *GCN5* gene was deposit in the GenBank (accession number gi:13591701).

By matching the genomic sequence to that of the GCN5 cDNA, the boundaries of exons and introns were determined. Detailed map information and the location of specific primers used for sequencing are shown in Figure 2. The result revealed no apparent open reading frame at the 5' end besides the previous determined cDNA sequence of GCN5. But interestingly, the N-terminus of GCN5 (amino acids 1-211) protein is encoded by a single exon.

The endogenous GCN5 protein

Figure 2. Diagram shows GCN5 gene and its encoded protein GCN5.
(A) Diagram of GCN5 gene. The lightly shaded areas in upper panel represent exons and the open areas represent introns, with boundaries indicated by the sequence numbers. Various primers and the sequences read by them are shown at the top where the arrows indicate the relative sequence length and direction obtained using each primer. Three *Hind III* restriction enzyme sites defined the two genomic DNA fragments identified by Southern blotting of cosmid DNA. The lower panel indicates the encoded GCN5 protein corresponding to the exons in the upper panel.
(B) Schematic alignment of *Arabidopsis* GCN5 against its homologues from other species. The HAT domain and the bromodomain (Br) are indicated. The unique N-terminus of *Arabidopsis* GCN5 is indicated by blue box, while the N-termini of mammalian and *Drosophila* homologues are indicated by the unfilled box.


A
Hind III
Hind III
Hind III
SJT244
SJT249
SJT234
SJT248
SJT197
SJT236
SJT235
SJT242
SJT232
SJT215
SJT238
SJT198
SJT233
SJT214
SJT231
SJT243
SJT245
1254 1408 1597 1767 2180 2492 2841 3309 3449 3795 3997 4252 4401
1070
1 622 1346 1518 1697 2018 2451 2740 3244 3391 3674 3903 4110 4343 4622
1
211
388
429
568
HAT Domain
205
381
Bromodomain
458
559


B
HAT
Br
Arabidopsis GCN5
Yeast GCN5
Human GCN5
Human PCAF
Mouse PCAF
Human GCN5
Mouse GCN5
Drosophila GCN5

Both the genomic sequence and the cDNA sequence revealed that *Arabidopsis* GCN5 protein comprises 568 amino acids in full length. The calculated molecular weight of GCN5 protein is 63 kDa, and the apparent molecular weight of recombinant GCN5 protein is about 75 kDa in SDS-PAGE gel (Figure 3A).

To check the expression of endogenous GCN5 protein, cell lysates from *Arabidopsis* cell suspension culture were analyzed by Western blotting. The antibodies used were raised against either purified recombinant GCN5 antigen (amino acids 155-568) or peptide antigen (amino acids 383-400) which was right after the HAT domain. The peptide antigen sequence is not inside the conserved HAT domain and is specific to *Arabidopsis* GCN5. Therefore the antiserum against this peptide antigen was more specific than the one raised against recombinant GCN5 antigen which contained both the conserved HAT domain and bromodomain.

The Western blotting detected a predominant protein by both antibodies with approximate molecular weight of 75 kDa (Figure 3B). This size is consistent with the migration of recombinant full-length GCN5 protein expressed from the cloned cDNA.

Curiously, the antiserum against recombinant GCN5 antigen also detected a larger protein that is unlikely a longer version of GCN5 for it was not detected by the more specific peptide antiserum. It might be a protein that also contains the conserved HAT domain or bromodomain.

Interestingly, a protein, with a little bit smaller size, was weakly detected by peptide antiserum but not by antiserum against recombinant GCN5 antigen (aa 155-568) (Figure 3B).

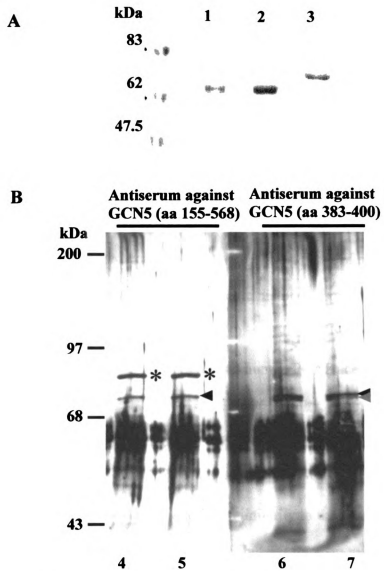


Figure 3. Detection of endogenous GCN5 protein.
(A) Recombinant GCN5 protein in 10% SDS-PAGE gel. All the proteins were Coomassie Blue stained. 1: 0.5 μg BSA, 2: 1μg BSA, 3: recombinant GCN5 protein. **(B)** The endogenous GCN5 protein was detected by Western blotting. Crude lysate of *Arabidopsis* leaf suspension cells (lane 4 and 6) and the supernatant of the crude lysate (lane 5 and 7) were resolved in 7.5% SDS-PAGE gel. The proteins in gel were then transferred to nitrocellulose membrane and detected by antiserum against either recombinant GCN5 antigen (aa 155-568) or GCN5 peptide antigen (aa 383-400). The detected GCN5 protein is indicated by the arrows. There are two close bands in lane 6 and 7 as indicated by the black and gray arrows. An unknown protein cross reacted with antiserum against recombinant GCN5 antigen, as indicated by the asterisk marker in lane 4 and 5.

Discussion

The results described in this chapter indicate that the predicted GCN5 homologue of *Arabidopsis* cDNA has histone acetyltransferase activity. This GCN5 protein preferentially acetylates histone H3. The predicated HAT domain of *Arabidopsis* GCN5 is sufficient for this activity. In addition, the genomic locus of *Arabidopsis* GCN5 was cloned and sequenced. The GCN5 gene contains thirteen exons and 12 exons.

Recombinant *Arabidopsis* GCN5 acetylates mainly core histone H3 and weakly H2B and H4 (Figure 1), a shared enzymatic activity characteristic of recombinant protein of GCN5 and PCAF from other species when core histone mixture is provided as the substrate. Which lysine residue on histone H3 is acetylated by *Arabidopsis* GCN5 was not determined. Recombinant yeast Gcn5 and human PCAF acetylate mainly lysine residue 14 (K14) on histone H3 (Kuo, et al., 1996, Schiltz, et al., 1999). It is reasonable to infer that recombinant *Arabidopsis* GCN5 protein also acetylates mainly H3 K14 *in vitro*, based on the facts that (i) the highly sequence conservation in the *Arabidopsis* GCN5 HAT domain (60% identity and 72% similarity to the yeast Gcn5 HAT domain), (ii) the conservation of the A, B and D motifs but not the C motif in the *Arabidopsis* HAT domain, like that of other species' GCN5 and PCAF protein (Neuwald and Landsman, 1997), (iii) the sufficiency of the *Arabidopsis* GCN5 HAT domain for enzymatic activity (Figure 3B).

The sufficiency of the HAT domain of *Arabidopsis* GCN5 to acetylate histones suggests that the unique N-terminus, the conserved bromodomain and the region between the bromodomain and the HAT domain (aa 384-458, designated as H-B) are not important for GCN5 HAT activity *in vitro*. The H-B regions in GCN5 and PCAF from

different species also show sequence conservation. The *Arabidopsis* H-B has 50% similarity to that of yeast Gcn5. However, the function of H-B is not known.

The bromodomain was also not necessary for yeast Gcn5 HAT activity *in vitro* (Candau, et al., 1997). However, efficient acetylation of nucleosomal substrate by the yeast SAGA *in vitro* requires the Gcn5 bromodomain (Sterner, et al., 2002). Similar situation was also observed for HAT protein CBP. CBP does not belong to the GNAT HAT family where GCN5 proteins belong (Deng, et al., 2003). I*n vivo*, the bromodomain was required for the transcriptional activation of some genes in yeast (Georgakopoulos, et al., 1995, Sterner, et al., 2002). These findings reveal that, although there is a general dispensability of this bromodomain *in vitro* for GCN5 to acetylate free core histones, it is important for GCN5 function *in vivo*.

Evidence suggests that the bromodomain is more involved in communicating with acetylated histone H3 and H4 tails. Yeast Gcn5 bromodomain bound to acetylated H3 and H4 tail (Ornaghi, et al., 1999), and NMR structural research revealed that the human PCAF bromodomain formed a hydrophobic pocket to which acetylated lysine residue from H3 or H4 tail bound (Dhalluin, et al., 1999). More recently, bromodomains within the catalytic subunits of SAGA and SWI/SNF has been shown to anchor these complexes to acetylated promoter nucleosomes (Hassan, et al., 2001). The bromodomain of human PCAF is also bind to acetylated lysine residue of the HIV Tat protein, and this interaction synergizes Tat and PCAF action on transcriptional activation of HIV promoter (Dorr, et al., 2002). This binding to non-histone protein of bromodomain proposes that the general role of bromodomain is to recognize acetylated lysine residues, in analogue to the SH2 domain in signal transduction pathways (Horn and Peterson, 2001).

Our genomic sequence of GCN5 gene is identical to that released more than one year later by the "the *Arabidopsis* initiative" [the *Arabidopsis* initiative, Nature 408, 796-815 2000] which determined the whole genomic DNA sequence of *Arabidopsis*. There are twelve introns in the GCN5 gene. Interestingly, there is no intron in the first exon which encodes the first 211 amino acids that encompass the whole unique N-terminus.

The peptide antiserum detected two closely positioned proteins in the Western blotting of endogenous GCN5 proteins (Figure 3B). It is not known whether the smaller-sized one is a shorter version of GCN5, or the bigger-sized one is a covalently modified version of GCN5, or one of detected proteins is not GCN5 protein.

Chapter Three

The Interaction between GCN5 and ADA2

Introduction

In budding yeast, the GCN5 protein is an integral component of at least two distinct complexes, termed ADA and SAGA, that acetylate nucleosomal histone H3 and H2B (Eberharter, et al., 1999, Grant, et al., 1997). The 1.8 MDa SAGA contains Ada adapter (or co-activator) proteins (Ada1, Ada2, Ada3, and Gcn5), Spt proteins (Spt3, Spt7, Spt8, and Spt20/Ada5), TBP associated proteins (TAF_{II}) proteins (TAF5/yTAF_{II}90, TAF6/yTAF_{II}60, TAF9/TAF_{II}17/20, TAF10/yTAF_{II}25, and TAF12/yTAF_{II}61/68), and the Tra1 protein that is related to the members of the ATM/DNA-PK/phosphatidylinositol 3-kinase family (Abraham, 2001, Khanna, et al., 2001, Vassilev, et al., 1998) . The smaller ADA complex (approximately 0.8 MDa) contains the Ada2, Ada3, Gcn5 and Ahc1 proteins. SAGA-like complexes are conserved in yeast and metazoan organisms. At least three distinct SAGA like complexes have been identified from human cells, including PCAF (Ogryzko, et al., 1998), TFTC (Brand, et al., 2001), and STAGA (Martinez, et al., 2001).

The yeast Ada2 and Gcn5 proteins can interact with each other (Candau and Berger, 1996, Candau, et al., 1996, Candau, et al., 1997). Biochemically purified ADA and SAGA complexes contain both ADA2 and GCN5. ADA2 is lost from these two complexes in yeast lacking Gcn5 protein (Grant, et al., 1997).

At present, no direct biochemical evidence indicates whether *Arabidopsis* has similar GCN5-containing HAT complex (es). Chapter II shows that *Arabidopsis* GCN5 is a histone acetyltransferase (HAT). Previously, our lab cloned cDNAs from *Arabidopsis*

encoding proteins homologous to ADA2 protein. *Arabidopsis* has two ADA2 homologues, namely ADA2a and ADA2b. We hypothesized these two ADA2 proteins interact with GCN5 based on the fact that yeast Ada2 interacts with Gcn5. The experiments described in this chapter test the physical interaction of the *Arabidopsis* GCN5 protein with the ADA2a and ADA2b proteins.

Experimental Methods

GST pull-down assay

For constructing plasmids expressing GCN5 GST-fusion proteins, the cDNAs encoding various fragments of GCN5 protein (amino acids 1–568, 1–150 or 155–568) were suncloned into GST vector pGEX (Pharmacia). GST-fusion proteins were expressed in *E. coli* strains XA90 or BL21(DE3) Codon Plus following induction with 40–60 μg/ml IPTG at 20°C for 20 h. The E. coli cells were then lysed by French press in Tris buffer (20 mM Tris pH 7.9, 150 mM NaCl, 1 mM EDTA and 10% glycerol) with protease inhibitor cocktail (Roche) and 1 μg/ml pepstatin. Triton X-100 and DTT were added to 1% (v/v) and 10 mM final concentration, respectively, and the lysates were rocked at 4°C for 40 min. The lysates were centrifuged at 20,000 x *g* for 15 min and the supernatants were collected. To check the expression of GST-fusion protein, 1 ml of the supernatant was shaken together with 60 μl glutathione–Sepharose beads (Pharmacia) for 90 min at 4°C. The beads were then washed eight times with 1 ml Tris buffer containing 10 mM ß-mercaptoethanol, 0.2% Triton X-100, 1 μg/ml pepstatin, 50 μg/ml PMSF, 1 μg/ml leupeptin and 2 μg/ml aprotinin. The bead-bound protein was eluted with 2% SDS and visualized by Coomassie Brilliant Blue staining after fractionation by SDS–PAGE.

To express the *Arabidopsis* ADA2a/ADA2b proteins *in vitro*, the corresponding cDNAs were cloned into His6-tagged vector pET-28 (Novagen). The ADA2a/ADA2b proteins were then expressed using CsCl gradient purified plasmids templates and the TNT T7 transcription/translation system (Promega). [^{35}S] methionine was included in the system to radiolabel the expressed proteins.

Depending on the specific experiment, GST or GST fusion protein (about 1-2 μg) bound to glutathione–Sepharose beads was incubated with 4 μl of radiolabeled ADA2a or ADA2b translation mixture for 1 h at 25°C in 200 μl of binding buffer (40 mM HEPES pH 7.6, 100 mM NaAc, 5 mM $MgCl_2$, 1 mM EDTA, 10% glycerol, 0.2% Triton X-100, 1 μg/ml pepstatin, 50 μg/ml PMSF, 1 μg/ml leupeptin and 2 μg/ml aprotinin) plus 8% bovine serum. After washing six times with 800 μl of binding buffer (with 300 or 400 mM NaCl) at 4°C or room temperature, the bound proteins were eluted with 2% SDS and detected by autoradiography after SDS–PAGE.

Yeast two-hybrid assay

GAL4 System 3 (Clontech) was used for the yeast two-hybrid assays, including prey vector pGADT7 that expresses the Gal4 activation domain (GAD) and has Leu2 and Amp^r markers, and bait vector pGBKT7 that expresses Gal4 DNA binding domain (GDBD) and has Trp1 and Kan^r markers. cDNAs encoding the full length or different parts of GCN5 were inserted into bait vector pGBKT7 in frame with that of GDBD. cDNAs encoding the full length or different parts of ADA2a or ADA2b were inserted into prey vector pGADT7 in frame with that GAD. All the cDNA encoding different

parts of ADA2a, ADA2b and GCN5 were made by PCR from full length cDNA templates, and the DNA sequences were checked to make sure that there were no mutations introduced during PCR reaction.

The constructed bait plasmids were prefixed by "pGDBD-". For instance, pGDBD-GCN5(fl) meant bait plasmid with the cDNA encoding GCN5 in full length. The constructed prey plasmids were "pGAD-". For instance, pGAD-ADA2b meant prey plasmid with cDNA encoding ADA2b protein.

To test protein-protein interaction, the transformed yeast cells (strain AH109) with respective bait and prey plasmids were first grown on solid synthetic dropout (SD) medium lacking leucine and tryptophan (Leu- Trp-). Colonies from SD (Leu- Trp-) medium were then streaked on SD medium (Ade- His- Leu- Trp-) containing X-α-gal for testing protein-protein interaction, as well as SD medium (Leu- Trp-) to make sure the introduced ADA2 and GCN5 proteins were not toxic to yeast cells. The yeast colonies were allowed to grow for three to four days at 30° C before the pictures were taken. Three selective markers were used simultaneously as criteria to judge the positive interaction between prey and bait proteins, the expression of *ADE2* and *HIS3* genes, as well as the expression of *MEL1* gene whose protein product was secreted to medium and catalyzed substrate X-α-gal into blue color product.

Results

ADA2 interaction with GCN5

The initial approach to investigate possible physical interactions between GCN5 and ADA2a or ADA2b was to use a GST pull-down assay. In this assay, GCN5 proteins

(amino acids 1-568, 1-150, and 155-568) were expressed in *E. coli* as GST-fusions. Approximately the same amount of each fusion protein was purified from *E. coli* and immobilized on glutathione-Sepharose beads. The ADA2a and ADA2b proteins were expressed and [^{35}S] methionine labeled *in vitro* using the TNT T7 transcription/translation system (Promega). The GCN5 fusion proteins bound to beads were incubated with the ADA2a and ADA2b proteins. After washing to remove non-specific proteins, the ADA2a and ADA2b proteins bound to the beads were eluted with SDS-PAGE loading buffer and detected by autoradiography.

As shown in Figure 1, both the full length GCN5 (aa 1-568) and the N-terminally truncated version of GCN5 (aa 155-568) pulled down both ADA2a and ADA2b. Comparable amounts of the N-terminal protein of GCN5 (aa 1-150) and the control GST protein did not pull down the ADA2 protein. We concluded that *Arabidopsis* GCN5 can directly interact with both ADA2a and ADA2b, and that the N-terminal region of GCN5 is neither necessary nor sufficient for these interactions.

To corroborate and extend this result, a yeast two-hybrid assay was employed. The cDNAs encoding full length or truncated GCN5 proteins as bait were inserted into vector pGBKT7 in frame with that of Gal4 DNA binding domain (DBD). cDNAs encoding ADA2 proteins as prey were subcloned into vector pGADT7 in frame with that of Gal4 activation domain (AD). Yeast strain AH109 cells transformed with respective bait and prey plasmids were grown on selective SD medium for three to four days at 30° C to test protein-protein interaction.

As shown in Figure 2, both ADA2a and ADA2b interacted with the full length of GCN5, but not the unique N-terminus of GCN5 (amino acids 1-210). This data confirmed

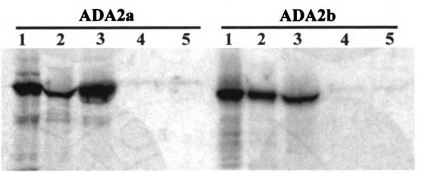


Figure 1. GCN5 interacts with ADA2a and ADA2b.
GST pull-down assay was performed to test GCN5 interaction with ADA2a and ADA2b. ADA2a and ADA2b proteins, expressed and [^{35}S]-methionine radiolabeled *in vitro*, were pulled down by equivalent amount of GST-GCN5 and GST proteins, that were bound to glutathione-Sepharose beads. The pull-down ADA2a and ADA2b proteins were visualized by autoradiography after being analyzed in SDS-PAGE gel. 1: 50% Input, 2: GST--GCN5(1-568), 3: GST--GCN5(155-568), 4: GST--GCN5(1-150), 5: GST.

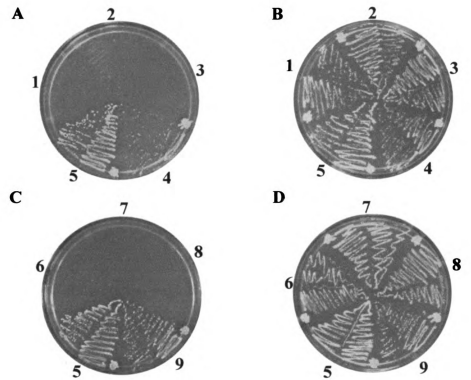


Figure 2. GCN5 interacts with ADA2a and ADA2b.
Yeast two-hybrid assays were performed to test the interaction. The interaction of GCN5 with ADA2a is shown in panel **A** and **B**, and ADA2b in panel **C** and **D**. The pictures were taken after the yeast cells had been grown on selective solid medium (SD Trp- Leu- His- Ade- /X-a-gal) (**A** and **C**) or on plasmid maintaining solid medium (SD Trp- Leu-) (**B** and **D**) for 4 days at 30°C. 1, 2, 3, 4: yeast co-transformed with prey plasmid pGAD-ADA2a and bait plasmid pGDBD-GCN5 (1-210), pGDBD-Lamin, pGBKT7, and pGDBD-GCN5(fl) respectively. 6, 7, 8, 9: yeast co-transformed with prey plasmid pGAD-ADA2b and bait plasmid pGDBD-GCN5 (1-210), pGDBD-Lamin, pGBKT7, and pGDBD-GCN5(fl) respectively. 5: yeast co-transformed with plasmid pGDBD-p53 and pGAD-LT as positive control, where LT stands for SV40 large T antigen. Plasmid pGBKT7 was the empty bait vector. In the plasmid nomenclature, GDBD stands for Gal4 DNA binding domain and GAD for Gal4 activation domain.

the result of GST pull-down experiment of GCN5 interaction with ADA2a and ADA2b, and further found that the N-terminus of GCN5 (amino acids 1-210) was not sufficient for interaction with ADA2a and ADA2b. Based on yeast colony growth, the ADA2b protein seemed to have stronger interaction with the GCN5 protein than the ADA2a protein.

Defining the interaction domains/regions between GCN5 and ADA2

Yeast two-hybrid assays were performed to further define the interacting regions or domains between GCN5 and ADA2a or ADA2b. GCN5 has distinguishable domains including the bromodomain (aa 451-568) and HAT domain (aa 203-384), and its unique N-terminus (aa 1-200). For the yeast two hybrid assay, the GCN5 protein was split into the N-terminus (aa 1-210), the HAT domain (aa 203-384), the bromodomain (aa 451-568) and the region between the HAT and the bromodomain (H-B, aa 361-468). The cDNAs of these region of GCN5 (except the N-terminus) were PCR amplified and subcloned into prey plasmid pGADT7 in frame with that of Gal4 AD.

These plasmids were transformed into yeast strain AH109 together with ADA2a or ADA2b plasmids, and assayed in selective plates as described above.

The results in Figure 3 showed that the HAT domain of GCN5 (amino acids 203-384) was sufficient to interact with ADA2a or ADA2b. Neither the bromodomain nor the H-B region showed interaction. GCN5 (aa 230-468) failed to show interaction with ADA2a/2b (Figure 3), suggesting that amino acids 203-230 in the HAT domain is necessary for interaction with ADA2a or ADA2b.

Similar experiments were set up to test which region in ADA2a and ADA2b

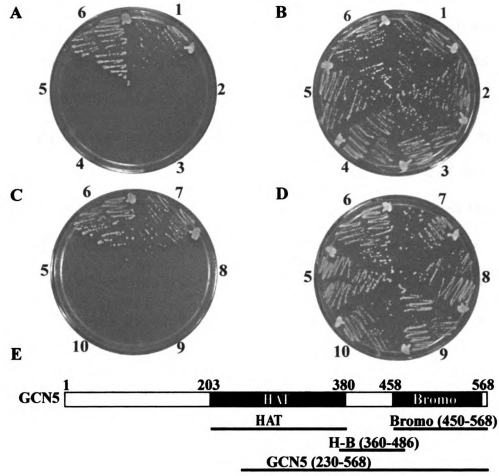


Figure 3. The GCN5 HAT domain interacts with ADA2a and ADA2b.
Yeast two-hybrid assays were performed to test the interaction. The interaction of GCN5 with ADA2a is shown in panel **A** and **B**, and ADA2b in panel **C** and **D**. The pictures were taken after the yeast cells had been grown on selective solid medium (SD Trp- Leu- His- Ade- /X-a-gal) (**A** and **C**) or on plasmid maintaining solid medium (SD Trp- Leu-) (**B** and **D**) for 4 days at 30°C. 1, 2, 3, 4: yeast co-transformed with prey plasmid pGAD-ADA2a and bait plasmid pGDBD-GCN5 (HAT), pGDBD-GCN5(H-B), pGDBD-GCN5 (230-468), and pGDBD-GCN5 (Bromo) respectively. 7, 8, 9, 10: yeast co-transformed with prey plasmid pGAD-ADA2b and bait plasmid pGDBD-GCN5 (HAT), pGDBD-GCN5 (H-B), pGDBD-GCN5 (230-468), and pGDBD-GCN5 (Bromo) respectively. 5: yeast co-transformed with plasmid pGDBD-GCN5 (HAT) and plasmid pGADT7. 6: yeast co-transformed with plasmid pGDBD-p53 and pGAD-LT as positive control, where LT stands for SV40 large T antigen. Plasmid pGADT7 was the empty prey vector. In the plasmid nomenclature, GDBD stands for Gal4 DNA binding domain and GAD for Gal4 activation domain. (**E**) Diagram shows the respective GCN5 regions tested in the yeast two-hybrid assays.

interacted with GCN5 and its respective domains. ADA2 proteins, including *Arabidopsis* ADA2a and ADA2b, have three recognizable conserved regions, respectively named as the N-terminus, the middle part and the C-terminus. In order to test their detailed interaction with GCN5, ADA2a and ADA2b were divided accordingly into three conserved parts as illustrated in Figure 4A, the N-terminus (ADA2a aa 2-267, ADA2b aa 2-229), the middle part (ADA2a aa 267-418, ADA2b aa 226-377) and the C-terminus (ADA2a aa 415-548 ADA2b aa 354-486). The cDNAs of these regions were PCR amplified and subcloned into prey plasmid pGADT7 in frame with that of Gal4 AD.

The results of yeast two-hybrid assays, as shown at Figure 4B, indicate that the middle part of ADA2a or ADA2b was found sufficient to interact with the HAT domain of GCN5. The C-terminal of ADA2a and ADA2b did not interact with GCN5. Curiously, the N-terminus of both ADA2a (ADA2aN) and ADA2b (ADA2bN) showed interaction only to the full length of GCN5. Neither the bromodomain nor the HAT domain of GCN5 was sufficient for interaction with the N-terminus of ADA2a and ADA2b. The interaction between ADA2aN and GCN5 seemed weak because the corresponding yeast colonies were pretty small in growth compared to the ADA2bN-GCN5 yeast colonies (Figure 4B).

The regions of ADA2a and ADA2b that interact with GCN5 were more narrowly defined by testing additional deletion constructs. These interactions were confirmed again as shown in Figure 5, where the N-terminus of ADA2 (ADA2a: aa 2-250, 2-267, and 2-272; ADA2b: aa 2-208, 2-225, 2-229) showed interaction with the full length of GCN5, and the middle part of ADA2 (ADA2a: aa 267-418, 273-418; ADA2b: aa 226-377, 230-377) showed interaction with the HAT domain of GCN5.

The results shown thus far indicate that the interaction pattern of ADA2a with GCN5

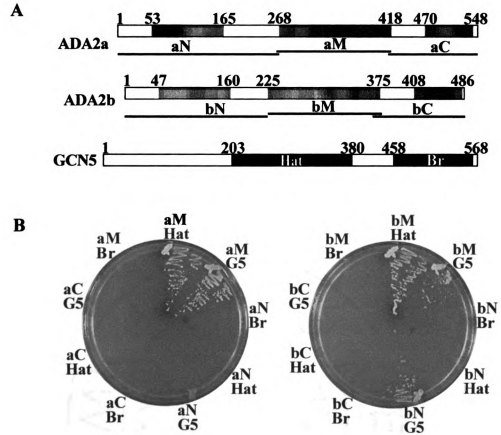


Figure 4. The middle part of ADA2a and ADA2b is sufficient to interact with the HAT domain of GCN5.

Yeast two-hybrid assays were performed to test the interactions. The different parts of GCN5 as bait protein and the different parts of ADA2a/b as prey protein are shown schematically in panel **A**. The GCN5 protein was tested for the full length (G5), the HAT domain and the bromodomain (Br). The ADA2a/b proteins were tested for the three conserved parts, the amino-terminus (aN or bN), the middle part (aM or bM) and the carboxyl-terminus (aC or bC). The yeast two-hybrid assays results are shown at panel **B**, where the yeast cells have been grown on selective solid medium (SD Trp- Leu- His- Ade- /X-α-gal) for 4 days at 30°C. Note: there is tiny colony growth in the group "aN G5" which is not visible in the picture

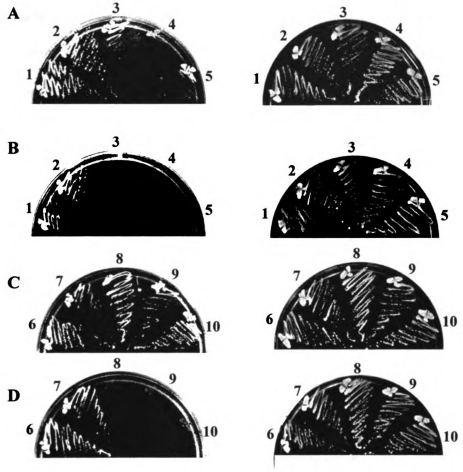


Figure 5. The N-terminus of ADA2 interacts with GCN5 and the middle part of ADA2 interacts with the HAT domain of GCN5.

Yeast two-hybrid assays were performed to test the interactions. Shown are the interaction of GCN5 (full length, fl) with various regions of ADA2a (**A**) or ADA2b (**C**), and the HAT domain of GCN5 with various regions of ADA2a (**B**) or ADA2b (**D**). The pictures were taken after the yeast cells had been grown on selective solid medium (SD Trp- Leu- His- Ade- /X-a-gal) (the left picture for each panel) or on plasmid maintaining solid medium (SD Trp- Leu-) (the right picture for each pane) for 4 days at 30°C. 1, 2, 3, 4, 5: yeast co-transformed with prey plasmid pGAD-ADA2a (273-418, 261-418, 1-272, 1-267 or 1-250 respectively) and bait plasmid pGDBD-GCN5 (fl) (**A**) or pGDBD-GCN5 (HAT) (**B**). 6, 7, 8, 9, 10: yeast co-transformed with prey plasmid pGAD-ADA2b (230-377, 225-377, 1-229, 1-225 or 1-208 respectively) and bait plasmid pGDBD-GCN5 (fl) (**C**) or pGDBD-GCN5 (HAT) (**D**). In the plasmid nomenclature, GDBD stands for Gal4 DNA binding domain and GAD for Gal4 activation domain.

was the same as that of ADA2b with GCN5, except that the N-terminus of ADA2a showed weaker interaction with the full length GCN5. The following experiments identified three exceptions to this pattern, indicating subtle distinctions in the interactions of GCN5 with ADA2a and ADA2b.

As shown in Figure 6B and C, the GCN5 HAT domain (aa 203-384) was repeatedly observed to interact with the middle part of both ADA2a and ADA2b. However, a short version of the GCN5 HAT domain (aa 203-368), interacted only with the middle part of ADA2a (Figure 6B), but not the middle part of ADA2b (Figure 6C).

As shown in Figure 7, the N-terminus of ADA2b (1-229), but not ADA2a (1-267), retained the ability to interact with the truncated GCN5 (aa 369-568) that contained the H-B region and the bromodomain. But neither the bromodomain nor the H-B region alone was sufficient for interaction with the N-terminus of ADA2b.

Moreover, the extended N-terminus of GCN5 (amino acids 1-250, compared to previous 1-210 in Figure 2) showed interaction with the full length of ADA2a but not the full length of ADA2b (Figure 8). Amino acids 210-250 of GCN5 are inside the HAT domain which had previously found to interact with the middle part of ADA2 (Figure 5). But the new yeast two-hybrid result show that the middle part of ADA2a was not sufficient for interaction with the extended N-terminus of GCN5 (Figure 8).

Discussion

The experiments described in this chapter indicate that the *Arabidopsis* GCN5 protein can interact with both ADA2a and ADA2b both *in vitro* and in yeast interaction assays. The HAT domain of GCN5 interacts with the middle part of ADA2a and ADA2b,

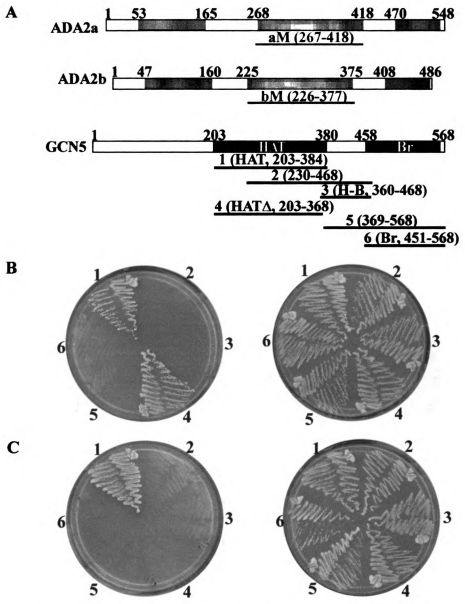


Figure 6. Differential interaction of the middle part of ADA2a and ADA2b with the HAT domain of GCN5.

Yeast two-hybrid assays were performed to test the interactions. The different parts of GCN5 (1, 2, 3, 4, and 5) as bait protein and the middle conserved part of ADA2a (aM) and ADA2b (bM) as prey protein are shown schematically in panel **A**. The yeast two-hybrid assays results are shown at panels **B** (for aM) and **C** (for bM), where the yeast cells have been grown on selective solid medium (SD Trp- Leu- His- Ade- /X-α-gal) (the left picture for each panel) or on plasmid maintaining solid medium (SD Trp- Leu-) (the right picture for each panel) for 4 days at 30°C.

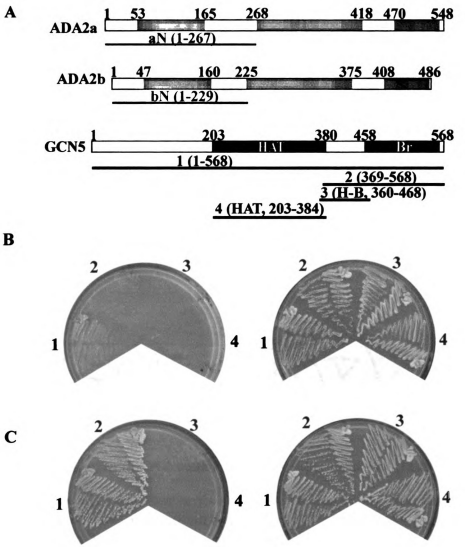


Figure 7. The N-terminus of ADA2b but not ADA2a interacts with GCN5 (369-568). Yeast two-hybrid assays were performed to test the interactions. The different parts of GCN5 (1, 2, 3, and 4) as bait protein and the middle conserved part of ADA2a (aN) and ADA2b (bN) as prey protein are shown schematically in panel **A**. The yeast two-hybrid assays results are shown at panels **B** (for aN) and **C** (for bN), where the yeast colonies have been grown on selective solid medium (SD Trp- Leu- His- Ade- /X-α-gal) (the left picture for each panel) or on plasmid maintaining solid medium (SD Trp- Leu-) (the right picture for each panel) for 4 days at 30°C.

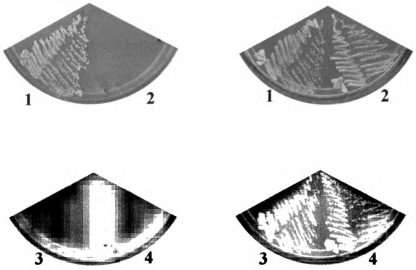


Figure 8. The extended N-terminus of GCN5 interacts with ADA2a, but not ADA2b.
Yeast two-hybrid assays were performed to test the interactions. The growth of the yeast colonies with respective bait and prey plasmids on selective solid medium (SD Trp- Leu- His- Ade- /X-a-gal) is shown at the left, and on plasmid maintaining solid medium (SD Trp- Leu-) at the right. The pictures were taken after the colonies had been growing at 30°C for 4 days. 1, 2, 3, 4: yeast co-transformed with bait plasmid pGDBD-GCN5 (1-250) and prey plasmid pGAD-ADA2a (full length), pGAD-ADA2a (267-418), pGAD-ADA2b (full length), or pGAD-ADA2b (226-377) respectively.

and the N-termini of ADA2a and ADA2b interact with the full length GCN5 protein.

These interactions can be taken as preliminary evidence for the presence of ADA2-GCN5 complex(es) in *Arabidopsis in vivo*. To what extent the *in vivo* complex(es) is similar to the yeast SAGA and human PCAF complexes is unknown. The genomic sequence of *Arabidopsis* reveals homologous proteins of yeast Tra1 protein and the TAFs (TAF5, 6, 9, 10 and 12) which are components of yeast SAGA. Spts. But curiously, the *Arabidopsis* genome encodes no apparent homologues components of yeast SAGA complex including Ada1, Ada3, Ada5, Spt3, Spt7 and Spt8, suggesting the GCN5-ADA2 complex in *Arabidopsis* might be different in composition.

It is unlikely that both ADA2a and ADA2b are present in the same complex *in vivo* based on the observation that both ADA2a and ADA2b bind to a common part of GCN5. Rather, *Arabidopsis* likely contains two GCN5-ADA2 complexes: one containing ADA2a and GCN5 and the other containing ADA2b and GCN5. Supporting evidence comes from the characterization of ADA2 proteins in *Drosophila*. *Drosophila*, like *Arabidopsis*, has two ADA2s designated as ADA2a and ADA2b. Biochemical purification of *Drosophila* ADA2 associated proteins illustrates that ADA2b is in a SAGA-like complex but ADA2a is not (Kusch, et al., 2003, Muratoglu, et al., 2003), although ADA2a also shows interaction with *Drosophila* GCN5. The enrooted phylogenetic tree shows that the plant ADA2s have diverged into a separate branch different from that of the metazoan ADA2s (Kusch, et al., 2003). This suggests that the *Arabidopsis* ADA2-GCN5 complex(es) might also have its own uniqueness in compositions.

In addition, genetic data from our lab also shows functional distinctions between

ADA2a and ADA2b in *Arabidopsis*. *Arabidopsis* mutants with homozygous disruption in the GCN5 gene has very similar defective phenotype to that of ADA2b disruption mutants (Vlachonasios, et al., 2003), while no obvious defective phenotype has been traced to plant with homozygous disruption of ADA2a gene.

Although the *Arabidopsis* ADA2a and ADA2b proteins interact with GCN5, the details of those interaction differ in three ways, (i) ADA2a, but not ADA2b binds to GCN5 (aa 1-250), (ii) the N-terminus of ADA2b, but not the N-terminus of ADA2a binds to truncated GCN5 (aa 369-568), (iii) the middle part of ADA2a, but not ADA2b binds to GCN5 (aa 203-368). Whether these differences have biological relevance is not known.

In budding yeast, the conserved N-terminus of Ada2 protein interacts with Gcn5 protein and the HAT domain of Gcn5 interacts with the ADA2 protein (Candau and Berger, 1996). These interactions are similar to what have been found in the interactions of *Arabidopsis* GCN5 with ADA2 and ADA2b. But there are some specific interactions of *Arabidopsis* GCN5 with ADA2 and ADA2b in detail as discussed below.

The GCN5 C-terminal (aa 369-568) is sufficient to interact with the N-terminus of ADA2b (Figure 7). This interaction has not reported from other system. This C-terminal of GCN5 includes the conserved H-B region and the bromodomain. However, neither H-B nor bromodomain is sufficient to interact with the N-terminus of ADA2b (Figure 7 and Figure 4). One explanation of these observations is that the H-B region can not fold correctly unless bromodomain is connected or that the N-terminus of ADA2b needs to contact both H-B and bromodomain for sufficient interaction.

The physical interaction of the middle part of ADA2a and ADA2b with GCN5 HAT has also not been reported from other model organisms. This specific interaction is also

confirmed in separate yeast two-hybrid screening experiment. When GCN5 in full length was used as bait to screening for interacting proteins expressed from *Arabidopsis* seedling cDNA library, the major prey proteins identified were truncated ADA2a and ADA2b protein that converge on the conserved middle part (see Chapter V for more information). A very recent report (Sterner, et al., 2002) showed that in yeast partial deletion of the conserved middle part of the Ada2 protein results in disintegration of Gcn5 and Ada3 in purified SAGA complex. This observation might suggest indirectly that the middle part of Ada2 is involved in the interaction with Gcn5 in budding yeast.

Curiously, the middle part of ADA2b loses interaction with GCN5 HAT domain when 16 amino acids (369-364) are deleted from the HAT domain. This is reminiscent of what has been observed in budding yeast (Candau, et al., 1997). In budding yeast, the HAT activity of Gcn5 is defined to amino acids 1-261. The region right after the HAT domain (aa 262-280) in yeast Gcn5 is necessary for interaction with Ada2 in an *in vitro* co-expression and co-immunoprecipitation assay. This interacting region in Gcn5 with Ada2 is critical for the *in vivo* function of GCN5. In *gcn5* null mutant, GCN5 (1-280) is able to complement for growth in minimum medium, able to potentiate the activator function of Gal4-VP16 in low copy plasmid, able to restore the toxicity of Gal4-VP16 in high copy plasmid to yeast. All these functions are lost when GCN5 (1-280) is truncated to GCN5 (1-261). In protein sequence alignment, the region of amino acids 262-280 in yeast Gcn5 overlaps with the region of amino acids 369-384 of *Arabidopsis* GCN5, and both regions are located right at the end of HAT domain. This overlapping suggests that there might be common motif important in the GCN5 protein for interaction with ADA2.

That the N-terminus interacts with GCN5 and the middle part of ADA2 interacts

with the HAT domain of GCN5 suggests that the ADA2 and GCN5 interaction involves multi-site contacts. One of the significances of these physical interactions is illustrated by the fact that both the N-terminus and the middle part of ADA2 enhance GCN5 HAT activity (see Chapter IV for more detail).

Chapter Four

Stimulation of GCN5 Enzyme Activity by ADA2b

Part I. Purification of recombinant protein of GCN5, ADA2a and ADA2b

Part II. Both ADA2a and ADA2b enhance the HAT activity of GCN5.

Introduction

The role of ADA2 in GCN5-ADA2 complex is not yet well understood. In yeast, ADA2 is not an essential gene. The SAGA complex can be purified intact in yeast strains with null mutations in ADA2 (Grant, et al., 1997). Although early models proposed that ADA2 serves as a handle for transcriptional activators to recruit SAGA complex to target promoters (Barlev, et al., 1995), more recent work indicates that found that transcriptional activators contact SAGA complex by its component, instead of Ada2, ADA1, TAF6, TAF12 and Tra1 (Brown, et al., 2001, Hall and Struhl, 2002). On the other hand, the GCN5 transcriptional activation activity has been shown to depend on Ada2 interaction in yeast (Candau, et al., 1997).

Recombinant yeast Gcn5 protein is active on free histone but not nucleosomal histones (Kuo, et al., 1996), while SAGA is able to acetylate nucleosomal histones (Grant, et al., 1997) *in vitro*. Therefore, other components in SAGA must enable Gcn5 to acetylate nucleosomal substrates. All these observations led us to test the hypothesis that *Arabidopsis* ADA2 might regulate GCN5 enzyme activity.

Part I. Purification of recombinant protein of GCN5, ADA2a and ADA2b

Experimental Methods

Expressing plasmid construction

GST-tagged expression vector pGEX-6P-1 (Amersham Biosciences) was used to construct all GST-fusion plasmids. The GST tag in the expressed GST fusion proteins can be cut off by PreScission protease which cuts the amino acid sequence Leu-Glu-Val-Leu-Phe-Gln /Gly-Pro.

cDNA fragments of the open reading frame of GCN5, ADA2a and ADA2b were inserted in frame into vector pGEX-6P-1. The cDNA fragments included those encoding GCN5 (aa 1-568, full length), ADA2b (the full length of wild type, alanine or arginine point mutation at K215), the three individual parts of ADA2b (the N-terminus, aa 1-225; the middle part, aa 226-377; and the C-terminus, aa 354-486), and the three individual parts of ADA2a (the N-terminus, aa 1-267; the middle part, aa 267-418; and the C-terminus, aa 415-548). For ADA2b, the cDNA encoding the full-length protein was also inserted to His6-tagged vector pET-28 (Novagen) to express the His6-tagged ADA2b protein.

Recombinant protein purification

All recombinant proteins were expressed in *E. coli* strain BL21(DE3) Codon Plus (Stratagene). The purification procedure was carried out at 4°C unless indicated.

Transformed *E. coli* cells with respective plasmids were first grown in 10 ml liquid LB with ampicillin (100 μg/ml) and chloramphenicol (40 μg/ml) at room temperature overnight. The 10 ml culture was transferred to 1000 ml liquid LB with the same concentration of antibiotics. The cells were induced at room temperature (around 20°C)

for 18 hours with 60-90 μM IPTG. The cells were then collected by centrifugation, and washed once with cold HEPES buffer (40 mM HEPES, 50 mM KCl, 10% glycerol, 7 mM 2-mercaptoethanol). The washed cells were resuspended in about 20 ml of HEPES buffer plus 150 mM KCl with protease inhibitor cocktail tablets (Roche). Cell lysates were made by French press treatment. In order to improve the solubility of the expressed fusion protein, Triton X-100 was usually added to the lysates to final concentration of 0.2% and the lysate was rocked at 4°C for 30 minutes. Cleared supernatant was recovered after centrifugation at 20,000 x g for 20 min at 4°C and stored at -80°C.

To purify GST-fusion proteins, the supernatant was incubated with appropriate volume of glutathione-Sepharose beads (Amersham Biosciences) in cold room for about one hour. The beads were then washed with an excess volume of HEPES buffer plus 150 mM KCl and 0.001% Triton X-100, and then washed once with HEPES buffer.

GST fusion proteins were eluted from the Sepharose beads with HEPES buffer containing 20 mM glutathione. The directly eluted GST fusions included GST-ADA2b (the full length of wild type, the K215A and K215R mutants), GST-ADA2bN (the N-terminus), GST-ADA2bM (the middle part), GST-ADA2bC (the C-terminus), GST-ADA2aN, GST-ADA2aM and GST-ADA2aC. GST protein alone as control was purified in the same way.

For purifying recombinant GCN5, the GCN5 protein was cut off from GST fusion protein directly while it was still bound to the Sepharose beads by PreScission protease (Amersham Biosciences) in HEPES buffer for four hours at cold room.

For the purification of His6-ADA2b, the procedure of the expression and purification of GST fusions described above was followed with the following modifications: (i) Ni-

NTA beads (Qiagen) were used instead of Sepharose beads, (ii) buffer A (20 mM HEPES, 25 mM KCl, 5% glycerol, 7 mM 2-mercatpoethanol) was used, (iii) 20 mM imidazole was included in the cell lysates, (iv) 200 mM KAc included in the washing solutions and 40 mM imidazole was also included in the final washing step, (v) the His6-ADA2b protein was eluted from the Ni-NTA beads with 225 mM imidazole in buffer A. The eluted His6-ADA2b was either used directly or dialyzed in HEPES buffer at cold room.

Because of the considerable contamination of *E. coli* proteins in purified His6-ADA2b protein, the same purification procedure was followed to purify the contaminated proteins from *E. coli* cells transformed with the vector pET-28.

Western blotting

Protein samples of His6-ADA2b and GST-ADA2b were electrophoresed in 7.5% SDS-PAGE gel. The proteins were detected by standard Western blotting. The first antibody was either ADA2b specific IgG (1:4000 dilutions) purified from rabbit antiserum (rabbit #108) against ADA2b peptide antigen (aa 392-405), GST antibody (BioRad, 1:2000 dilutions) and His6-tag monoclonal antibody (Sigma, 1:3000 dilutions). The signal was detected using Lumi-light Western Blotting Substrate kit (Roche).

Experimental results

In order to test the possible regulation role of ADA2 on GCN5 enzyme activity, the *Arabidopsis* GCN5, ADA2a, and ADA2b proteins were first expressed and purified from *E. coli* either as His6-tagged or as GST-tagged fusion protein.

The GCN5 protein was purified as a GST fusion protein and then the GST tag was

subsequently cut off with protease. As shown in Figure 1A, relatively pure GCN5 protein was obtained. Acquire

The ADA2 protein has been very hard to express and purify. After many trials, reasonable amounts of recombinant ADA2b protein were obtained using low concentrations of IPTG (60-90μM) and prolonged induction time (about 18 hours) at room temperature. The expressed recombinant ADA2b, as GST-tagged or His6-tagged fusion protein, was then purified as shown in Figure 1B and Figure 1C. Recombinant ADA2a protein in full length could not be successfully purified regardless of the modification of expression and purification methods.

There was considerable contamination in the isolated ADA2b protein either in GST tagged or His6-tgged form. The presence of GST-ADA2b or His6-ADA2b protein was confirmed by Western blotting assay using either ADA2b peptide antibody, His6 monoclonal antibody or GST antibody (Figure 2). The Western blotting result also detected some truncated ADA2b proteins in the purified preparation.

Recombinant ADA2a and ADA2b protein were also purified in their truncated forms. Based on the amino acid sequence similarity as discussed in chapter III, ADA2b protein was divided into three regions, the N-terminus (2bN, amino acids 1-225), the middle part (2bM, amino acids 226-377) and the C-terminus (2bC, amino acids 353-486). Relatively pure GST-tagged proteins of all these ADA2b regions were isolated (Figure 1D).

Similarly, the three regions of ADA2a (2aN: 1-267, 2aM: 267-418, and 2aC: 418-548) were also obtained as GST-tagged fusions in relative purity (Figure 1E).

For the three regions of both ADA2a and ADA2b, in addition to the relatively high purity, the yield was also relatively high: GST-ADA2bN (up to7 μg/μl), GST-ADA2bM

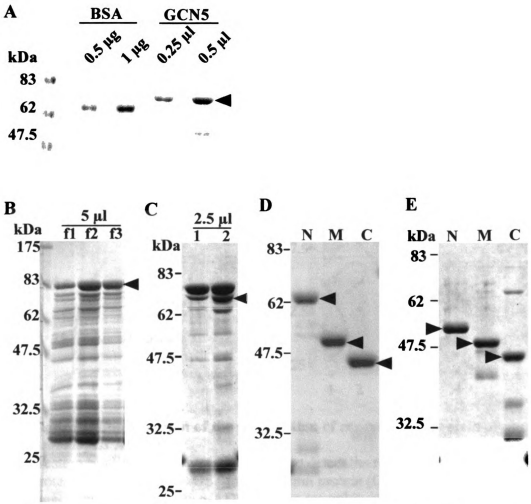


Figure 1. The purified recombinant proteins of GCN5, ADA2a and ADA2b. The recombinant proteins were analyzed in SDS-PAGE gel and Coomassie Blue stained. Shown are purified recombinant proteins GCN5 (**A**) where BSA protein was included to indicate the relative quantity of the GCN5 protein, GST-ADA2b (**B**), His6-ADA2b (**C**, lane 2) where proteins purified from expressed *E. coli* transformed with His6 vector only is also shown (lane 1), the three parts as GST fusions of ADA2a (**D**) and ADA2b (**E**). The positions of the respective proteins in the gels are indicated by the arrows. (**B**) f1, f2, f3: elution fraction 1, 2, and 3 respectively from glutathione Sepharose beads. (**D**, **E**) N: the N-terminus, M: the middle part, C: the C-terminus.

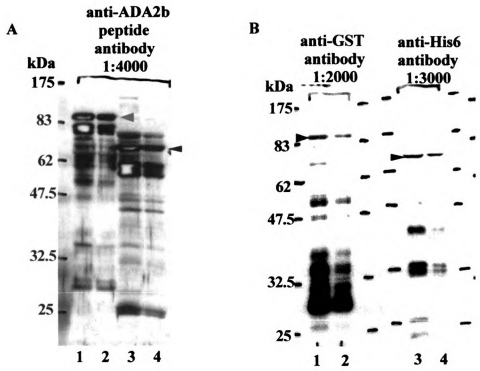


Figure 2. Confirmation of the purification of recombinant protein of *Arabidopsis* ADA2b.

Western blotting assays were performed to detect the recombinant ADA2b protein purified either as GST-tagged fusion protein (GST-ADA2b) or His6-tagged fusion protein (His6-ADA2b). The recombinant ADA2b proteins were electrophoresed in 7.5% SDS-PAGE gel, and detected by antibodies against either ADA2b peptide antigen (**A**) or GST tag and His6 tag antigens (**B**). The detected His6-ADA2b protein is indicated by the black arrow while the GST-ADA2b protein by the blue arrow. For the GST-ADA2b protein preparation, either 0.75 μl (lane 1) or 0.38 μl (lane 2) was loaded in the SDS-PAGE gel. For His6-ADA2b protein preparation, either 1.5 μl (lane 3) or 0.75 μl (lane 4) was loaded.

(about 0.5 μg/μl), GST-ADA2bC (up to 7 μg/μl), GST-ADA2aN (up to 2 μg/μl), GST-ADA2aM (267-418) (about 0.5 μg/μl), and GST-ADA2aC (up to 20 μg/μl).

Discussion for Part I

The purification ADA2 proteins of other species has been hard. For instance, reasonable amount of yeast ADA2 was purified only when yeast GCN5 is co-expressed in E. coli (Balasubramanian, et al., 2002).

Reasonable amount of recombinant fusion proteins GST-ADA2b and His6-ADA2b were obtained from *E. coli* cells under the mild expression conditions of low temperature (around 20 °C) and prolonged induction (around 18 hours) with low amount of IPTG (60-90 μM), although the purity is low. However, expression and purification of full-length ADA2a protein has not been successful.

The preparations of GST-ADA2b or His-ADA2b had much contamination (Figure 1). No further purifications were applied to remove the contamination. But the existence of purified His6-ADA2b or GST-ADA2b in the contaminated mixture was confirmed by Western blotting, using antibodies against the tag of His6 or GST or ADA2b specific antibodies. The Western blotting showed that the purified ADA2b protein preparation contained the full-length ADA2b protein and some truncated ADA2b proteins which were probably the products of incomplete translation or proteolysis.

For unknown reasons, early experiments fount that phosphate buffer (140 mM NaCl, 2.7mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.3, for GST-ADA2b; and 50 mM NaH_2PO_4, 300 mM NaCl, for His6-ADA2b) was found not very compatible with ADA2b protein, which sometimes could lead to either poor elution of ADA2b protein from

glutathione Sepharose and Ni-NTA beads, or precipitation of the fusion proteins after dialysis. The chelating reagent EDTA also caused similar problems. This incompatibility might be due to the putative zinc finger at the ADA2b N-terminus, for EDTA sequesters zinc ions and phosphate precipitates zinc ions. Therefore, HEPES buffer (40 mM HEPES, pH 7.3, 50 mM KCl, 10% glycerol) was used subsequently in purification of all ADA2 and GCN5 recombinant proteins.

Using HEPES buffer, however, only a very small fraction of GST-ADA2b or His6-ADA2b could be eluted from glutathione Sepharose or Ni-NTA beads. But relatively concentrated amounts of the three parts of protein ADA2b could be eluted from glutathione Sepharose beads. The GST tag could also be readily removed by protease digestion from both GST-ADA2bN and GST-ADA2bC proteins.

For ADA2a, it was unable to purify more than trace amounts of either His6-tagged or GST-tagged protein regardless of using phosphate or HEPES buffer. However, like the three parts of ADA2b, very clean and concentrated purification of the three regions of ADA2a were achieved using HEPES buffer (Figure 1).

Part II. ADA2 enhances the HAT activity of GCN5.

Experimental Methods

Free core histones

Core histone mixture (H3, H2B, H2A, and H4) was purchased from Sigma (Cat. No. H-9250, purified from calf thymus). The mixture was dissolved in HEPES buffer (40 mM HEPES, 50 mM KCl, 10% glycerol, 7 mM 2-mercaptoethanol) to final concentration of 10 mg/ml.

Free histone acetylation assay

For testing the enhancing activity of GST-ADA2b, the reaction was carried out in 51 µl HEPES buffer, containing about 4 pM (0.25 µg) GCN5, 15 µg core histone mixture, and 0.1 µCi [^{3}H] Ac-CoA. Varying amounts of GST-ADA2b ranging from about 1.2 to 6 µg were added. The GST protein (ranging from about 2 to 10 µg) was added in the controls. The reaction proceeded for 5 or 10 min and was quenched with 12 ul of 2N HCl. 15.5µl from each reaction was spotted onto P81 Whatman filter paper (half of filter paper disc with the diameter of 2.5 cm). The paper discs were washed in 500 ml of 50 mM $NaHCO_3$ (pH adjusted to 9.1) for three times. Each disc was then soaked in 5 ml of Ecolume liquid scintillation cocktail (ICN company). Scintillation activity was read for one minute three days later by scintillation counter (LKB, 1208 rackbeta). For each 63 µl reaction (51 µl reaction plus 12 µl quenching HCl), three aliquots of 15.5 µl were counted.

The remaining 16.5 µl of each reaction was analyzed in 15% SDS-PAGE gel. The gel was first stained with Coomassie Blue to visualize the proteins, and then the gel was soaked in autoradiographic enhancer (NEN Life Science Products), vacuum dried and exposed to X-ray film for three days for visualization of acetylated histones and ADA2 proteins.

For testing the stimulatory activity of the three parts of ADA2b or ADA2a, the same procedure was followed varying amounts of the three parts.

Nucleosome purification

Mono and oligonucleosomes were purified from turkey erythrocytes, following the

protocol (Cote, et al, Methods de Molecular Genetics 1995, 108-127) with some modification. All procedure was done at 4°C or on ice. About 60 ml blood was collected from turkey, and the erythrocytes were collected after centrifugation at 3,000 x g. The cells were suspended in cell lysis buffer [20 mM HEPES, pH 7.5, 0.25 M sucrose, 3 mM $MgCl_2$, 0.2% (v/v) Nonidet P-40, 3 mM 2-mercaptoethanol, 0.4 mM phenylmethylsulfonyl fluoride (PMSF), 1 μM pepstatin A, 1 μM leupeptin]. The nuclei were pelleted at 3,000 x g and washed with the lysis buffer and then with buffer B (20 mM HEPES, pH 7.5, 3 mM $MgCl_2$, 0.2 mM EGTA, 3 mM 2-mercaptoethanol, 0.4 mM PMSF, 1 μM pepstatin A, 1 μM leupeptin). The nuclei were then resuspended in 5 ml buffer B, to which an additional 3 ml buffer B containing 0.6 M KCl and 10% glycerol was added dropwise while stirring. The nuclei were pelleted again at 17,500 x g for 30 min, and washed with buffer MSB (20 mM HEPES, pH7.5, 0.4 M NaCl, 1 mM EDTA, 1 mM 2-mercaptoethanol, 0.5 mM PMSF, 5% glycerol).

The nuclei were then resuspended in buffer HSB (20 mM HEPES, pH7.5, 0.65 M NaCl, 1 mM EDTA, 1 mM 2-mercaptoethanol, 0.5 mM PMSF, 0.34 M sucrose). The nuclei were broken by gentle Dounce homogenization (B pestle, stroked about 100 times) to release high molecular weight chromatin. The homogenate was centrifuged at 10,000 x g for 20 min. The supernatant was dialyzed overnight in buffer LSB (20m M HEPES, pH7.5, 0.1 M NaCl, 1 mM EDTA, 1 mM 2-mercaptoethanol, 0.5 mM PMSF). 0.03 x volumes of 100 mM $CaCl_2$ were added to the dialyzed chromatin (final volume was 8.0 ml). 4.0 ml chromatin was digested with 105 units of micrococcal nuclease (Amersham Bioosciences) at 37°C for 6 min, and the digestion was quenched with 440 μl 0.5 M EGTA. 4 M NaCl was added to final concentration of 0.6 M. The supernatant was

collected after centrifugation at 150,000 x g for 30 min.

3.5 ml of the supernatant was applied to Sepharose CL-6B (Amersham Biosciences) column (65 x 2.5 cm) which was equilibrated in HSB buffer without sucrose. The column was run at 12 ml/hr and 1.1 ml fractions were collected. Fractions were checked for nucleosome integrity by monitoring DNA length and core histone composition. For checking the DNA length, the fraction samples were first extracted with phenol/chloroform and then 40 µl from each extraction was loaded in 1% agarose gel. For checking the histone composition, 5 µl of each fraction was mixed with 4 x SDS-PAGE loading buffer and boiled for 10 minutes before loading to 15% SDS-PAGE gel. The proteins in gel were visualized with Coomassie Blue staining.

Fractions 51-54 were pooled as nucleosome B and fractions 58-61 as nucleosome C. Nucleosomes B and C were dialyzed against buffer C (40 mM HEPES, pH 7.4, 20m M KCl, 2% glycerol, 1 mM EDTA, 0.7 µg/ml pepstatin, 1 µg/ml leupeptin, 1µg/ml aproptinin, 0.5 mM PMSF, 7 mM 2-mercaptoethanol) overnight. The dialyzed nucleosome solutions, while still inside dialysis bags, was concentrated by soaking in dry polyethylene glycol (Sigma, molecular weight 15,000-20,000 Da) for four hours. The concentrated nucleosome solutions were kept at -80°C until used.

Nucleosomal histone acetylation assay

The yeast HAT complexes SAGA and NuA4 were included as positive controls, which were kindly provided by Dr. Shelley Berger at Wistar Institute in Philadelphia.

For comparing the HAT activity of *Arabidopsis* GCN5 with that of SAGA and NuA4, similar free histone acetylation assays as that above were performed. The assays were

carried out in 50 µl HEPES buffer with or without 6 mM sodium butyrate, 15 µg free core histone mixture, 0.1 µCi [^{3}H] Ac-CoA. 4 µl SAGA or NuA4, or 4 picomole GCN5, or 4 picomole GCN5 plus about 10 picomole GST-ADA2b was included. The reaction was incubated at 25°C for 15 minutes before quenched with 12 µl 2N HCl. Three aliquots of 15 µl from the reaction were read for acetylation activity by the method of scintillation counting on filter paper as described above. The remaining 17 µl reaction was electrophoresed in 15% SDS-PAGE gel for fluorographic picture.

For nucleosomal histone acetylation assay, the reaction was performed in 24 µl HEPES buffer containing 6 mM sodium butyrate, 8 µl nucleosome preparation, 0.05 µCi [^{3}H] Ac-CoA, about 10 pM GST-ADA2b or His-ADA2b, and 96 pM, 24 pM or 6 pM GCN5. The mixture reacted for 30 min at 25°C and was stopped by adding 16 µl of 6 x SDS-PAGE loading buffer and boiled for 5 min. As control, 2 µl of SAGA or NuA4 was added to the reaction mixture instead of GCN5 and GST-ADA2 or His6-ADA2b. The reactions were boiled and analyzed in 15% SDS-PAGE gel. Then the gel was soaked in autoradiographic enhancer (NEN Life Science Products), vacuum dried and exposed to X-ray film for visualization of acetylated histones and ADA2 proteins.

For testing mutant ADA2b effect of GCN5 activity on nucleosomal substrate, the reaction was carried out in 27 µl HEPES buffer containing 6mM sodium butyrate, 8µl nucleosome preparation, 0.2 µC tritiated Ac-CoA, about 8 pM GCN5 and about 4 pM or 8pM GST-ADA2b (wild type, K215R and K215A mutant). The reaction proceeded for 30 min at 25°C and was stopped by adding 16 µl of 6x SDS-PAGE loading buffer and boiled for 5 min. equivalent amount of GST was included instead of GST-ADA2b as negative control. The reactions were resolved in 15% SDS-PAGE gel, and the acetylated

histones were detected by fluorography for 4 months at -80°C.

Results

ADA2b enhanced GCN5 HAT activity.

To test whether ADA2b could enhance the enzymatic activity of GCN5, HAT assays were performed using the purified recombinant proteins as described in Part I.

For a typical HAT assay, the reaction mixture included about 4 pM (0.25 μg) GCN5, 15 μg core histone mixture, [^{3}H] Ac-CoA, and varying amounts of GST-ADA2b protein or His6-ADA2b protein which had a different repertoire of contaminated proteins (Figure 1C) from that GST-ADA2b proteins. GST or protein preparation from *E. coli* cells transformed with His6-tag vector only that contained identical contamination proteins found in His6-ADA2b preparation (Figure 1C) was used as a negative control in parallel reactions. The mixture was incubated for 10 minutes at 25°C. The acetylation signals of core histones were measured by liquid scintillation counting on filter paper. As shown in Figure 3, using free core histone mixture (H2A, H2B, H3 and H4) as the substrate, the HAT activity of GCN5 was enhanced up to 5 fold by the presence of GST-ADA2b or His6-ADA2b.

The filter paper scintillation counting method counted all acetylated proteins, and did not indicate whether the acetylation was enhanced on one core histone or an expanded substrate spectrum. To address this uncertainty, aliquots of the HAT assay reactions were electrophoresed in a 15% SDS-PAGE gel. The proteins were first stained with Coomassie blue to show their position, and the acetylation signals were visualized by enhanced fluorography. As shown in Figure 3C, GCN5 alone (with GST protein

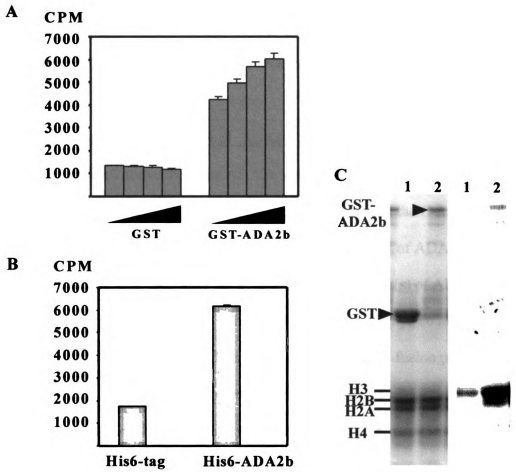


Figure 3. ADA2b enhances the HAT enzyme activity of GCN5 on free core histones.

HAT assays were carried out with recombinant GCN5, [^{3}H] Ac-CoA, free core histones (H2A, H2B, H3 and H4) and the recombinant ADA2b protein or controls. The HAT activity was either quantitatively measured by liquid scintillation counting as shown by CPM readings (**A** and **B**) or visualized by fluorography (**C**). (**A**) The results of HAT assays using recombinant GST-ADA2b (ranging from about 0.6 to 1. 5 μg) or equivalent amounts of control GST. (**B**) The results of HAT assays using recombinant His6-ADA2b (about 0.4 μg) or equivalent amounts of control His6-tag which was purified from expressed *E. coli* cells that contained His6-tag vector only (see figure 1C lane 1). (**C**) Fluorogram of the HAT assays. Parellel assays of panel **A** were analyzed in 15% SDS-PAGE gel and the proteins were first visualized by Coomassie Blue staining as shown at the left, and then the acetylation signals were detected by enhanced fluorography as shown at a the right. 1: GST, 2: GST-ADA2b.

present as control) acetylated histone H3, H2B and H4 weakly. This activity was enhanced by the addition of GST-ADA2b protein. The pattern of histone acetylation did not change. Therefore, the enhancement activity mainly came from stronger acetylation of current substrate histones. ADA2b did not change GCN5 substrate specificity.

The N-terminus and the middle part of ADA2a/b enhance GCN5 HAT activity.

As alluded in Part I above, both the ADA2a and ADA2b protein could be divided into the three conserved parts, the N-terminus, the middle part and the C-terminus (C). As discovered at Chapter III, both the N-terminus and the middle part of ADA2a and ADA2b interact with GCN5. A following question was which part(s) of ADA2a and ADA2b are necessary and sufficient to enhance GCN5 HAT activity. To answer this question, similar HAT assays were performed using purified GST-fusion proteins of the three parts. As shown on Figure 4A, both the N-terminus and the middle part of ADA2a and ADA2b enhanced GCN5 HAT activity, whereas the C-terminus of both ADA2a and ADA2b did not (the result of the C-terminus of ADA2b is not shown). The N-terminus and the middle part of ADA2b had slightly better enhancement activity than their counterparts of ADA2a.

To test whether the pattern of histone acetylation was altered by these two parts of ADA2a and ADA2b, core histones were checked for their acetylation by fluorography. As shown in Figure 4B, both the N-terminus and the middle part enhanced GCN5 acetylation activity on core histones H3, H2B and H4, but did not change GCN5 substrate specificity. The N-terminus of ADA2b is acetylated by GCN5 (see Chapter VII), but like the situation in ADA2b (wild type, full length), the enhanced acetylation signal by

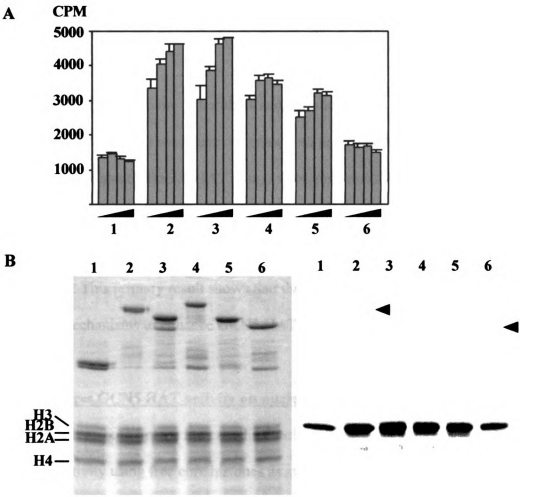


Figure 4. The N-terminus and the middle part of both ADA2a and ADA2b enhance the HAT enzyme activity of GCN5 on free core histones.

HAT assays were carried out with recombinant GCN5, [^{3}H] Ac-CoA, free core histones (H2A, H2B, H3 and H4) and the three regions of protein ADA2a or ADA2b. Control GST was also included. The HAT activity for each assay was either quantitatively measured by liquid scintillation counting as shown by CPM readings (**A**) or visualized by fluorography (**B**). 1: GST, 2: GST-ADA2bN, 3: GST-ADA2bM, 3: GST-ADA2aN, 4: GST-ADA2aM,5: GST-ADA2aM, 6: GST-ADA2aC: ADA2a, N: the N terminus, M: the middle part, C: the C terminus. (**A**) The results of HAT assays using recombinant GST (about 20-160 pM), GST-ADA2bN (about 8 to 60 pM), GST-ADA2bM (about 12 -90 pM), GST-ADA2aN (about 9-70 pM), GST-ADA2aM (about 13-100 pM), and GST-ADA2bC (about 13-100 pM). (**B**) Fluorogram of the HAT assays. Parellel assays of panel **A** were analyzed in 15% SDS-PAGE gel and the proteins were first visualized by Coomassie Blue staining as shown at the left, and then the acetylation signals were detected by enhanced fluorography as shown at a the right.The acetylated GST-ADA2bN and ADA2aC are indicated by the arrows.

ADA2bN mainly came from the acetylation of core histones. Interestingly, the C-terminus of ADA2a but not ADA2b was acetylated by GCN5 (Figure 4B, and data not shown).

Having found that both the N-terminus and the middle-part of ADA2b enhanced GCN5 HAT activity, it was not known whether both parts together enhanced GCN5 HAT activity better than alone. Similar HAT assays were performed with the two parts included in the assays simultaneously. Figure 5 showed an additive effect on GCN5 HAT activity when both the N-terminus and the middle-part of ADA2b were added simultaneously. This primary result shows that the N-terminus and the middle part might use different mechanisms to enhance GCN5 HAT activity.

ADA2b enhances GCN5 HAT activity on nucleosomal substrate

The preceding experiments indicate that *Arabidopsis* ADA2 proteins can enhance the GCN5 HAT activity using free core histones as substrate. However, the native substrate for GCN5 is nucleosomal histones. Therefore, studies were followed to investigate whether ADA2 proteins can influence GCN5 HAT activity on nucleosomal histones.

The nucleosomes were purified from turkey erythrocytes. The purified chromatographic nucleosomes fractions contained equal amounts of the four core histones (Figure 6A) and contained DNA fragments with length ranging from 200 base pairs to a few kilobase pairs (Figure 6B). These nucleosomes were devoid of linker histone H1. Fractions of mononucleosome or dinucleosomes (based on the DNA length) from the gel-filtration were pooled and used in all subsequent acetylation assays.

As a positive control for the nucleosome acetylation assay, yeast HAT complexes

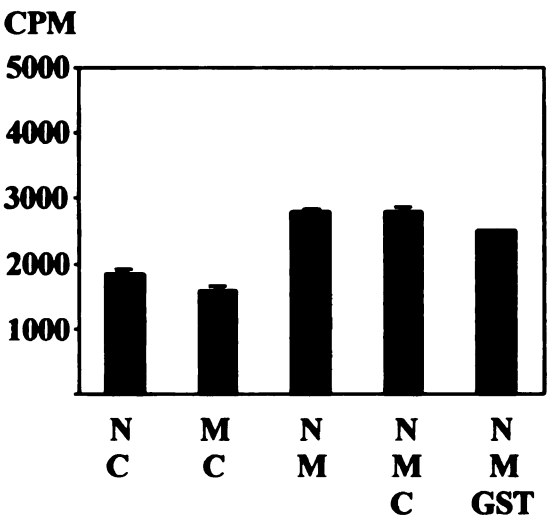


Figure 5. The combinatorial effect of the N-terminus and the middle part of ADA2b in enhancing the HAT enzyme activity of GCN5 on free core histones.

HAT assays were carried out with recombinant GCN5, [^{3}H] Ac-CoA, free core histones (H2A, H2B, H3 and H4) and the various combinations of the three regions of protein ADA2b as well as the control GST protein. Equivalent amount of each protein was included in each assay. The HAT activity for each assay was quantitatively measured by liquid scintillation counting as shown by the CPM readings. N: GST-ADA2bN, M: GST-ADA2bM, C: GST-ADA2bC.,

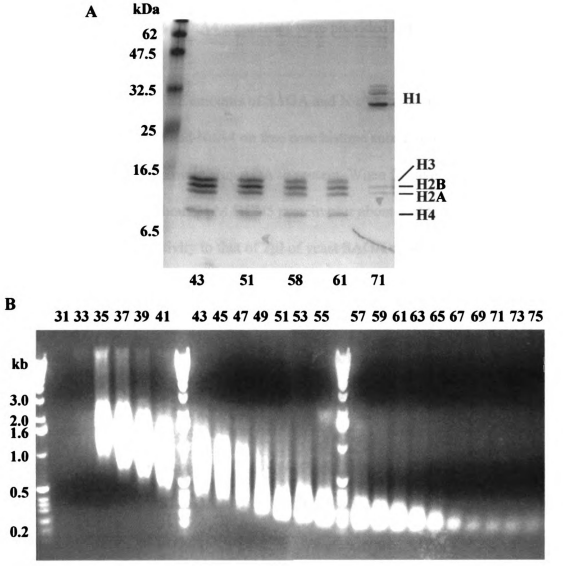


Figure 6. Checking the purified nucleosomes from turkey blood erythrocytes. Different chromatographic fractions of nucleosome preparations as indicted by the numbers were either analyzed for the composition of histone proteins (**A**) or for the length of genomic DNA fragments (**B**). (A) The fractions were analyzed in 15% SDS-PAGE gel and stained with Coomassie Blue in gel. (**B**) The fractions were electrophoresed in 1% agarose gel after phenol/chloroform extraction. The DNA fragments were stained by ethidium bromide.

SAGA and NuA4 were included, which are known to acetylate nucleosomal histone H3, H2B and H4. The SAGA and NuA4 complexes were provided kindly by Dr. Shelley Berger from the Wistar Institute.

In order to use appropriate amounts of SAGA and NuA4 in the nucleosomal HAT assay, the activity of SAGA and NuA4 on free core histone substrate was compared to that of *Arabidopsis* GCN5 protein plus ADA2b protein. When free core histone mixture was used as the substrate, about 2 pM GCN5 protein plus about 5 pM GST-ADA2b protein had similar HAT activity to that of 2μl of yeast SAGA or NuA4 (Figure 7).

Nucleosomal HAT assays were then carried out using the purified turkey nucleosomes and the recombinant GCN5 protein and ADA2b protein as well as the positive controls. As shown in Figure 8, the prepared turkey nucleosomes were readily acetylated by 2 μl of yeast HAT complex SAGA or NuA4. Core histones H3, H2B, H4 were acetylated. GCN5 alone did not acetylate the core histones in nucleosomes, even using as much as about 96 picomole of GCN5 enzyme. However, in the simultaneous presence of recombinant GST-ADA2b protein in the reaction, core histone H3 from the nucleosomes was acetylated, even at the lowest amount of GCN5 enzyme tested (about 6 picomoles). Because of the concentration limitation of recombinant GST-ADA2b protein, it was impractical to add more ADA2b to the HAT assay reaction. Since the purified GST-ADA2b protein preparation had numerous contaminating proteins, a parallel assay was performed using recombinant His6-ADA2b instead of GST-ADA2b. The purified His6-ADA2b preparation had a very different contamination protein profile from that of GST-ADA2b (Figure 1). As shown in Figure 8, GCN5 was also stimulated to acetylate nucleosomal histone H3 when recombinant His6-ADA2b protein was added,

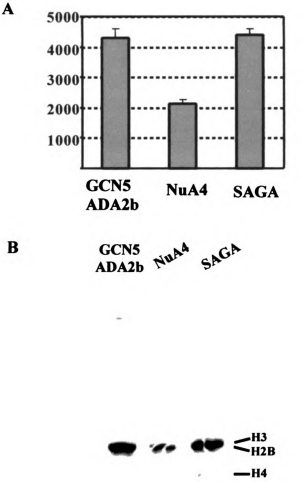


Figure 7. Comparing the HAT activity on free core histones of *Arabidopsis* GCN5-ADA2 with that of yeast HAT complexes SAGA and NuA4.
HAT assays were carried out with [^{3}H] Ac-CoA, free core histones (H2A, H2B, H3 and H4) and *Arabidopsis* recombinant proteins GCN5 (~ 2 pM) and GST-ADA2b (~5pM), or purified yeast HAT complex SAGA (2 µl) or NuA4 (~2 µl). The HAT activity for each assay was either quantitatively measured by liquid scintillation counting as shown by the CPM readings (**A**) or visualized by fluorography (**B**).

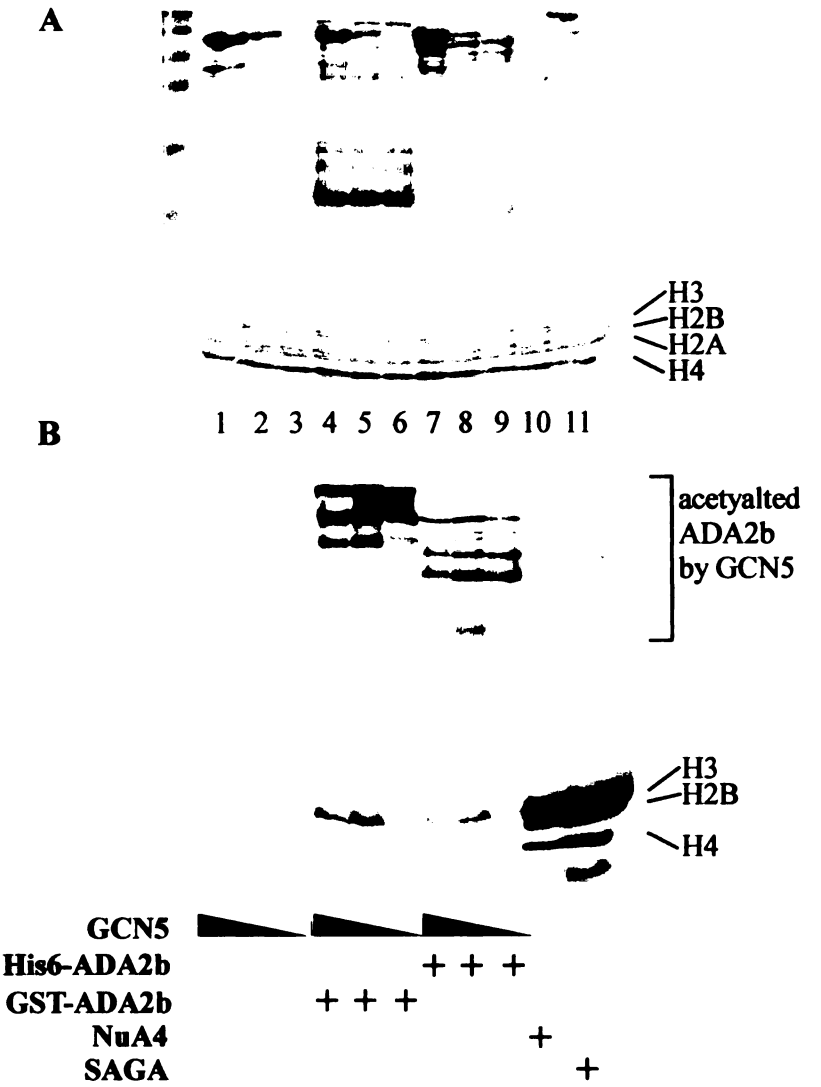


Figure 8. *Arabidopsis* GCN5 does not acetylate nucleosomal core histone while GCN5 plus ADA2b does weakly.

Nucleosomes obtained from turkey erythrocytes were incubated, in the presence of [^{3}H]-Ac-CoA,with recombinant *Arabidopsis* GCN5 protein (lane 1-9, ~6, 24 and 96 pM) supplemented with GST-ADA2b (lane 4-6, ~10 pM) or His6-ADA2b (lane 7-9), or yeast HAT complex NuA4 (lane 10, 2 µl) or SAGA (lane 11, 2 l). Reacted mixtures were analyzed by SDS-PAGE, stained by Coomassie Blue (panel **A**) and then subjected to enhanced fluorography for 30 days (panel **B**).

demonstrating the effect was ADA2b protein specific.

Nevertheless, compared to positive control SAGA which strongly acetylated nucleosomal histone H3, 2B and H4, *Arabidopsis* GCN5 together with ADA2b acetylated nucleosomal histone weakly and only acetylated H3.

The mechanisms of ADA2b to enhance GCN5 HAT activity

The mechanisms by which ADA2 enhances GCN5 HAT activity on free core histones are unknown. One possibility is that ADA2b improves GCN5 affinity for its substrate histones or acetyl CoA. To test this hypothesis, similar HAT assays were performed with constant amount of GCN5 enzyme, GST-ADA2b protein or control GST protein, and varying amount of core histone or acetyl CoA.

As shown in Figure 9A, the GCN5 HAT activity increased in the presence of control GST or GST-ADA2b with increasing amount of acetyl CoA and constant amount of core histones, and the fold of enhancement of GST-ADA2b over control GST remained stable. In contrast, as shown in Figure 9B, with increasing amount of core histones and constant amount of acetyl CoA, the GCN5 HAT activity decreased slightly in the presence of GST-ADA2b, but the HAT activity increased in the presence of control GST. These preliminary results indicate that ADA2b protein might enhance GCN5 HAT activity by reducing the GCN5 Km for core histones.

Discussion for Part II

The experiments described in this chapter indicate that both *Arabidopsis* proteins ADA2a and ADA2b can enhance the histone acetyltransferase activity of GCN5 on free

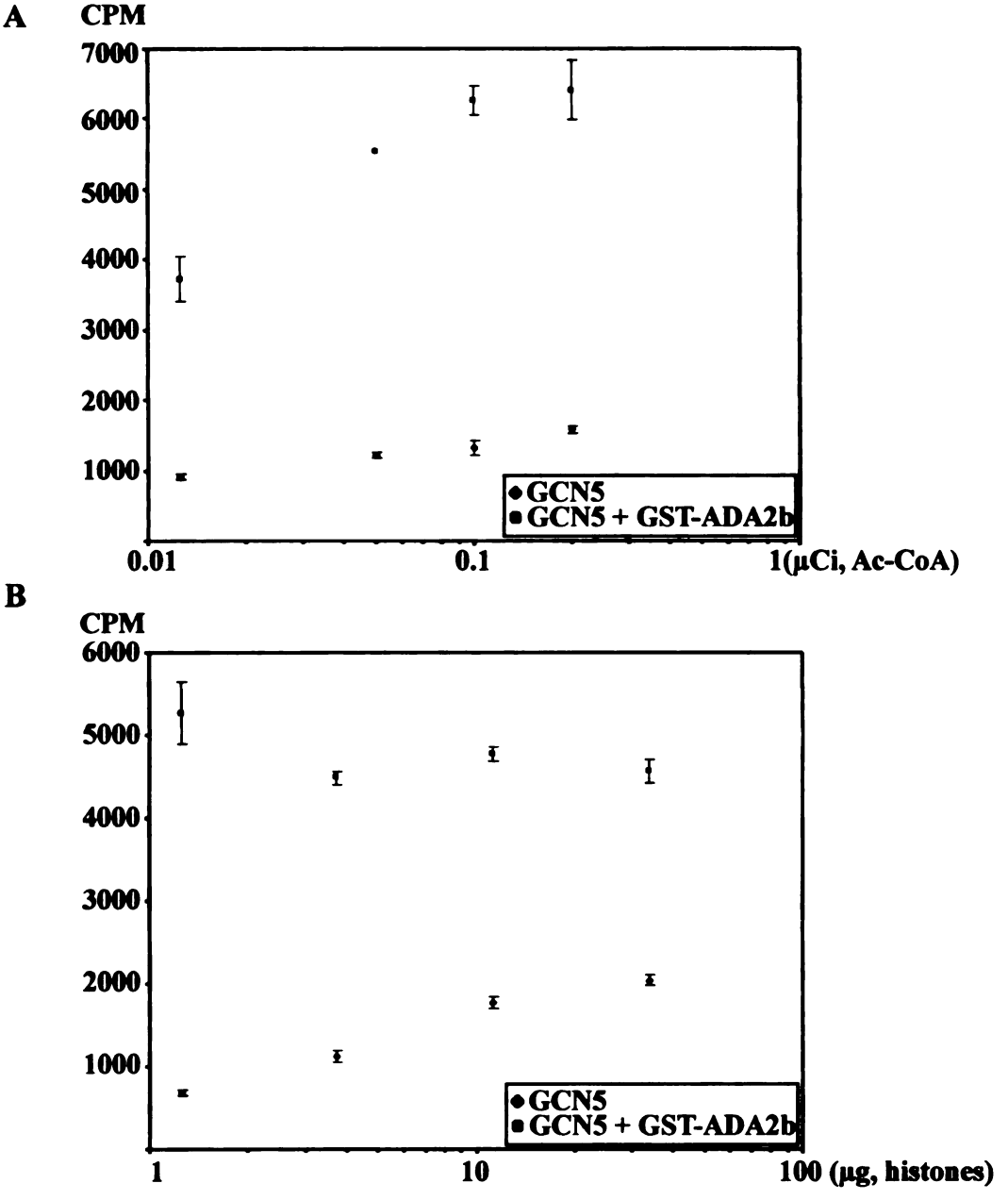


Figure 9. Differential HAT activities of GCN5-ADA2b to varying amount of Ac-CoA and core histotons.

HAT assays were carried out with recombinant GCN5 protein in the presence or absence of recombinant GST-ADA2b protein, and either constant amount of free core histones (H2A, H2B, H3 and H4) and varying amount of [^{3}H] Ac-CoA (**A**) , or constant [^{3}H] Ac-CoA with varying amount of core histones (**B**). The HAT activities were quantitatively measured by liquid scintillation counting as shown by the CPM readings.[^{3}H] Ac-CoA ranged from 0.0125 μCi to 0.2 μCi, and core histones from 1.25 μg to 33.75 μg.

core histones *in vitro*, and ADA2b stimulates GCN5 acetylation on nucleosomal histone H3. The enhancing activity on free histones is attributed to the conserved N-terminus and the middle part of ADA2a and ADA2b proteins. This functionality helps interpret the earlier finding that both ADA2a and ADA2b interact with GCN5 *in vitro* and that the N-terminus and the middle part of ADA2a and ADA2b interact with GCN5 (chapter III). We propose that *in vivo* ADA2a and ADA2b regulate GCN5 enzyme activity in *Arabidopsis*.

Compared to yeast SAGA complex that strongly acetylates core histone H3, H2B and H4 from the purified turkey nucleosomes, the *Arabidopsis* protein GCN5 together with ADA2b acetylates nucleosomal histones weakly and is limited to histone H3 (Figure 8). This suggests that *in vivo* the *Arabidopsis* GCN5-ADA2 complex, like the complex of yeast SAGA or human PCAF, has more components that make it capable of efficient acetylation of nucleosomal substrate.

In support of this suggestion, Song Tan group (Balasubramanian, et al., 2002) reported last year that in yeast, recombinant protein complex containing Ada2 and Gcn5, or Ada2 and Gcn5 and Ada3 had about 5-fold HAT activity over Gcn5 alone on free core histone substrate, and Ada2 enhanced GCN5 HAT activity on nucleosomal histones about 15 fold. Moreover, the recombinant complex Ada2-Ada3-Gcn5, compared to SAGA complex, had similar HAT activity on nucleosomal substrate as well as the same trend of acetylation spectrum of lysine residues on H3 tail (Lys-18 > Lys-14 > Lys-9 > Lys-23). This similar observation in yeast Ada2 and Gcn5 suggests the general mechanistic conservation among species. The weak acetylation activity of *Arabidopsis* GCN5-ADA2 on nucleosomes suggests that more components may affect this GCN5-

ADA2 complex *in vivo*, although *Arabidopsis* genomic sequence reveals there is no apparent homolog of yeast Ada3 protein.

Arabidopsis GCN5 protein alone acetylated mainly core histone H3, and acetylated H2B as well as H4 weakly when free core histones were assayed. ADA2b as well as the N-terminus and middle part of ADA2a and ADA2b enhanced GCN5 HAT activity on all these three core histones (Figure 3 and 4). There was no striking change of substrate repertoire with the addition of these ADA2 proteins. The apparent identical enhancement effect on GCN5 of the N-terminus and the middle part of ADA2a or ADA2b was surprising in considering that the N-terminus had different interaction nature with GCN5 from that of the middle part (Chapter III). For instance, the N-terminus of ADA2b interacts the c-terminal of GCN5 (amino acids 369-568, including the H-B and the bromodomain), while the middle part of ADA2b interacts the HAT domain of GCN5 (amino acids 203-384).

Although there was no substrate change in enhanced HAT activity of GCN5, it was not known for a specific substrate, such as H3, whether the improved signal of acetylation was due to more acetylation of one specific residue or due to expansion of acetylated residues on the specific tail. This question might be answered by Western blotting using specific antibodies against specific acetylated residues of histone tails or by Edman degradation sequencing.

Another general trend was that the N-terminus and middle part of ADA2b had better enhancement activity than their counterparts of ADA2a, as revealed in Figure 5C. The significance of this difference is not known

In the related MYST family HATs, most members such as Esa1 and Sas3 have a

cysteine rich, zinc-binding module at the N-terminus of their HAT domain. The module is essential for the HAT activity (Akhtar and Becker, 2001, Marmorstein, 2001). The HAT activity of Esa1 is reduced more than 90% if the cysteine module is mutated (Takechi and Nakayama, 1999). The HAT domain of CBP/p300 also has a cysteine rich, zinc-binding module. The module is not present in the HAT domain of GNAT family members including GCN5. Curiously, the conserved N-terminus of ADA2 has a similar cysteine rich, zinc-binding module. The fact that the N-terminus of *Arabidopsis* ADA2 enhances GCN5 HAT activity suggests that the module might be important for GCN5 proper activity.

Meanwhile, the enhancement of GCN5 activity by the N-terminus of ADA2a and ADA2b may also involve the SANT domain. The SANT domain is also conserved in the N-terminus of ADA2. It locates right after the cysteine rich, zinc-binding module. In yeast, purified SAGA with SANT domain deletion in Ada2 is inactive in nucleosomal HAT assays (Boyer, et al., 2002). *In vitro*, the SANT domain in yeast ADA2 has also been demonstrated to be required for acetylation of histone H3 (Grant, et al., 1997).

Very interestingly, our preliminary data (Figure 9) suggests that ADA2b enhances GCN5 HAT activity by lowering GCN5 Km for histone substrate but not acetyl CoA. This prediction is in agreement with observations that yeast ADA2 improves GCN5 HAT activity on free histone H3 tail by decreasing Km for H3 tail (3.8 fold) and increasing Vmax (2.6 fold), but yeast ADA2 does not change GCN5 affinity for acetyl-CoA (Boyer, et al., 2002).

The following discussion describes possible ways that the GCN5 enzyme activity could be altered.

Crystal structures of the HAT domains of GNAT family members (yeast HAT1, yeast Gcn5, Tetrahymena GCN5, human PCAF) and the distantly related MYST family member Esa1 reveal a striking conserved core domain, consisting of 4 β-sheets, a loop and a α-helix (Marmorstein, 2001, Roth, et al., 2001). This core domain is responsible for Ac-CoA binding and catalysis. Variable N- and C-termini, juxtaposed at the periphery of the core domain, form two hydrophobic clefts together with the core domain. The tail of H3 and Ac-CoA bind to the clefts. The variable N- and C- termini mediate histone substrate binding (Marmorstein, 2001, Roth, et al., 2001). The constancy of the core domain and the variability of the N- and C-termini suggest that proteins, by influencing the structure of the N- and C-termini, may change GCN5 enzyme catalytic activity and/or substrate specificity. Indeed, subtle conformational changes in the HAT domain are observed upon substrate Ac-CoA binding (Marmorstein, 2001).

Biochemical studies show an ordered substrate binding and product release in yeast Gcn5 and human PCAF (Tanner, et al., 2000, Tanner, et al., 2000), in which Ac-CoA binds first and then H3 tail. The acetylation reaction is carried out by an ordered Bi-Bi mechanism, where the acetyl moiety is transferred directly to the ε-amino group of lysine 14 residue on histone H3 (H3 K14). After the reaction, acetylated H3 tail is released first followed by CoA. This strictly ordered mechanism suggests that any facilitation by other protein in substrate binding or product release may enhances GCN5 HAT activity, although which step is limiting is not known.

In the GCN5 histone acetylation reaction, the acetyl moiety is transferred directly to the ε-amino group of H3 K14 via nucleophilic attach. Under physiological pH (7-8), the ε-amide group of H3 K14 (pKa~10) is protonated, which hinders the nucleophilic attach

by acetyl-CoA. Therefore, any influential action on the protonation of the ε-amide group will improve the HAT enzyme activity by changing Kcat. A critical glutamic acid residue is conserved in HAT domain of the GCN5 and PCAF proteins. In yeast Gcn5 protein, mutation of the critical glutamic acid (residue 173) to glutamine almost abolishes the enzyme activity (Kuo, et al., 1998, Wang, et al., 1998). This mutation does not change the Km for either histone tail or acetyl CoA substrate, but Kcat drops nearly 360 fold (Tanner, et al., 1999). The crystal structure reveals the critical Glu residue in human PCAF to be closely positioned to the ε-amino group of H3 k14, suggesting its action of abstracting a proton from the ε-amino group (Clements, et al., 1999). Therefore, any proteins (such as ADA2) which improve the deprontonation process will enhance GCN5 HAT activity.

On the other hand, the histone tails are notoriously non-structural random coils (Luger, et al., 1997), whereas the crystal structure of PCAF HAT domain with bound histone H3 tail illustrates that the histone H3 tail adapts a distinct path (Clements, et al.,). This suggests that proteins such as ADA2 might help to load the histone tail to the catalytic site of HAT domain to facilitate HAT enzyme activity. In support of this suggestion, ADA2 is indeed observed to lower GCN5 Km for histone substrate (Boyer, et al., 2002) (and my preliminary data in Figure 9).

In summary, substrate binding, direct attack at the ε-amino group by acetyl moiety, and products release could all be the possible mechanisms for regulating *Arabidopsis* GCN5 activity on free histone substrate. These possible ways might be carried out by ADA2 protein through direct or indirect modulating of the variable N- and C-termini of GCN5 HAT domain.

As to the histone tail substrate from nucleosomes, another layer of possible regulation may be carried out by *Arabidopsis* ADA2b. The histone tails are notoriously resistant to being acetylated by GCN5 alone. One possible hindrance might be the inaccessibility of histone tails that might caused by the interaction between the positive tail and the negative DNA. Therefore, ADA2 protein might, as suggested by John Denu and Craig Petersom (Boyer, et al., 2002), helps present histone tail to GCN5 for acetylation.

Chapter Five

The Function of the Unique N-terminus of the *Arabidopsis* GCN5 Protein

Introduction

The transcriptional coactivator protein GCN5, originally identified in yeast, has well conserved homologs across broad phylogenetic distances, including Tetrahymena, *Drosophila*, mouse and human. Mammals have two closely related proteins GCN5 and PCAF. They all contain a HAT (histone acetyltransferase) domain and a bromodomain. An N-terminal region (about 400 amino acids) is highly conserved among mouse and human GCN5/PCAF as well as *Drosophila* GCN5, but the N-terminus is not present in yeast GCN5. In contrast, *Arabidopsis* GCN5 has a unique N-terminus (amino acids 1-203), which shares no sequence similarity to that of human GCN5 (as illustrated in Figure 1 in Chapter II).

GCN5 homologues from rice, maize and alfalfa also have N-terminal regions that are similar in sequence to the N-terminus of *Arabidopsis* GCN5, suggesting some specific function in plant world. One notable feature of the N-terminus of *Arabidopsis* GCN5 is the repetitive stretches of serine residues (residue frequency 22 %).

The unique N-terminus of *Arabidopsis* GCN5 is not required for its HAT activity on free core histones (Chapter II), and it does not interact with ADA2a or ADA2b (Chapter III). Virtually no function of the N-terminus of *Arabidopsis* GCN5 has been known. Intriguingly, the human and mouse GCN5 N-terminus interacts with CBP/p300 *in vitro* and the *Arabidopsis* genome also encodes CBP/p300 homologues (Bordoli, et al., 2001, Xu, et al., 1998, Yang, et al., 1996). The experiments described in this chapter were

undertaken to explore the potential function of the unique N-terminus of *Arabidopsis* GCN5.

Experimental methods

The yeast two-hybrid system

The MATCHMAKER Two-Hybrid System 3 (Clontech) was used for yeast two-hybrid screening. This system is Gal4-based, in which the bait gene is expressed as a fusion to Gal4 DNA-binding domain. The system features yeast strain AH109, which reduces the frequency of false positives by using three reporters—*ADE2*, *HIS3*, and *MEL*1. *MEL*1 encodes α-galactosidase which is secreted into medium and turns X-α-gal into blue color. The system also includes vector pGBKT7 for construction of bait plasmid and vector pGADT7 for construction of prey plasmid.

Vector plasmid pGADT7 is derived from vector pACT, which is the vector used to construct the two cDNA libraries used in the screening. pGADT7 and pACT both have the 2 micron replication origin, leucine auxotrophic and ampicillin resistant markers, and express the Gal4 activation domain driven by ADH1 promoter. The major difference is that pGADT7 had more restriction enzyme sites in the multiple cloning sites, integration of HA epitope tag and incorporation of *in vitro* translation start sequence right after the Gal4 activation domain. Vector pGBKT7 has tryptophan auxotrophic and kanamycin resistant markers and expresses Gal4 DNA binding domain.

Construction of plasmids

To construct bait plasmid pGDBD-GCN5 (1-568, the full length; or 1-210, the N-

terminus), cDNA fragments were inserted into plasmid pGBKT7 in frame with that of Gal4 DNA binding domain. The cDNA fragment of the N-terminus of GCN5 was PCR amplified first from relevant plasmid. The two bait plasmids were named as pGDBD-GCN5 and pGDBD-GCN5 (1-210).

To construct plasmid pGAD-MSI4 (full length, 1-507), pGAD-ΔMSI4 (missing the 50 amino acids in the beginning, 51-507), PCR fragment (*Nco* I - *Sca* I fragment) of cDNA encoding MSI4 (aa 1-70, with the first 50 amino acids encoded) or MSI4 (aa 51-70, without the first 50 amino acids encoded) were sub-cloned into vector pGADT7, together with the rest cDNA of MSI4 which was cut from candidate #133 plasmid (cut by *Sca* I and *Xho* I). Both the PCR fragments in the constructed plasmids were DNA sequenced to make sure no mutation was introduced. The encoding cDNAs in the constructed plasmids were then cut out with *Bgl* II and *Xho* I and subcloned into vector pGADT7 again which was cut with *BamH* I and *Xho* I. The reason for the second subcloning was to introduce longer spacer between Gal4 activation domain and MSI4 proteins.

Western blotting of bait protein GCN5 (1-210) expressed in yeast

Yeast AH109 transformed with plasmid pGBKT7 expressing Gal4 DNA binding domain GDBD or pGDBD-GCN5(1-210) were grown in 5 ml liquid synthetic dropout medium (SD Trp-) overnight, and then the 5 ml culture was transferred to fresh 40 ml SD Trp- medium for another four hours. The yeast cells were collected by centrifugation, and washed with chilled water. The cells were then mixed with 4-fold volume of glass beads (Sigma,#8772-6, 425-600 μm, acid treated) and 3-fold cracking buffer (8M urea, 5%

SDS, 50mM Tris, pH6.8, 1mM EDTA, 10% glycerol), and vortexed vigorously in the cold room for about 8 min. The supernatants from the vortexed cells were collected as cell lysates. 5 µl of each sample was resolved in 10% SDS-PAGE gel, and the proteins in gel were transferred to nitrocellulose membrane. The Gal4 and Gal4-GCN5 proteins were detected by standard Western blotting using first antibody LA2-3(1:100000) (rabbit antiserum). Antibody LA2-3 had been raised against antigen Gal4 DNA binding domain in fusion with VP16 activation domain, therefore the antibody could detect Gal4 DNA-binding domain.

cDNA libraries

The first cDNA library was made using polyA RNA from 3-day-old etiolated *Arabidopsis* (Kim, et al., 1997). Samples of this library as plasmids were provided by Dr. Peter Quail.

The CD4-10 phage library (LACT) was provided by *Arabidopsis* Biological Research Center at Ohio State University (ABRC order # 8485). This library was constructed using random primed RNA from mature *Arabidopsis* leaves and roots. The cDNA fragments bigger than 300 bp were inserted into the T-filled *Xho* I site of phage vector LACT.

To convert the phage library to plasmid library, 0.5 ml *E. coli* strain BNN132 from 50 ml LB/0.2% maltose overnight culture was transferred to 50ml LB/0.2% maltose and cultured for three hours until OD600 reached 0.620. For BNN132, one OD600 contained about 8.6×10^8 cell/ml. 30 ml culture of the above culture was spun down and resuspended in 10 ml 10 mM $MgSO_4$. About 3.2×10^9 *E. coli* cells were mixed with

$2.4x10^8$ PFU of CD4-10 library phage, and the mixture was incubated for 30 min at 30°C. Liquid LB medium was then added to the infected cells to final volume of 6 ml. A 20 µl sample was taken to estimate how many *E. coli* cells were infected, and the result was 2.6 x 10^7 cells were infected by the library phage (the original transformants were 9 $x10^6$ when the library was made). Then the phage infected cells were incubated at 30°C for 70 min with slow shaking. The cell culture was then spread on 24 plates (150 x 15mm) with LB/0.2% glucose/Amp (50 µg/ml) and the plates were incubated for 22 hours at 37°C.

E. coli lawn was scraped from all the plates and the plates were washed twice with LB/Amp (50µg/ml). All the *E. coli* cells were added into 6 liters of LB/Amp (50µg/ml) and cultured for 5 hours. Library plasmid was prepared from the 6-liter culture by plasmid DNA mega-preparation method. The library plasmid DNA was purified twice with CsCl/ ethidium bromide gradient centrifugation. The library DNA was finally resuspended in 3 ml TE (pH 7.5) after ethanol precipitation. The final DNA concentration was 1µg/µl. The library cDNA fragments sized from 300bp to 2kb.

Yeast cDNA library transformation

The transformation methods followed the protocols recommended by the MATCHMAKER Two-Hybrid System 3. Yeast AH109 (Clontech) carrying bait plasmid pGDBD-GCN5 (1-210) was grown in 280 ml liquid SD Trp- medium for 23 hours (OD600=1.18). The culture was transferred to 1000 ml YPDA medium (YPD medium plus 0.002% adenine) and cultured for 3 hours and 10 min at 30°C (OD600=0.666). The cells were collected by centrifugation and washed with 700 ml water. The cells were suspended in 7 ml LiAc/TE buffer (100 mM LiAc, 1 x TE, pH 7.5). The cells were then

mixed with 100-200μg library DNA, 20 mg carrier DNA and 30 ml PEG/LiAc/TE buffer (PEG, final concentration 40%, MW3350). The mixture was incubated for 30 min at 30°C with gentle rocking. Then 7 ml DMSO was added and mixed gently. The mixture was incubated at 42°C for 20 min with gentle rocking occasionally. The mixture was then shocked in ice for 10 min. The cells were spun down and resuspended in 5 ml YPDA with 6 ml TE buffer. 25 μl was taken from the resuspension to plate on SD (Trp- Leu-) solid medium plates to determine the transformation efficiency. The rest was spread on solid medium of SD (Trp- Leu- His- Ade-). Totally about one million transformants were obtained. The plates were incubated at 30°C for two weeks, and grown positive colonies were streaked on SD (Trp- Leu- His- Ade-) plates. Then these positive colonies were picked up and re-streaked on SD Trp- Leu- His- Ade-/X-α-gal plates to isolate single colony.

The subsequent library transformation for low stringency screening, for bait GCN5 in full length or for CD4-10 library followed the same protocol.

Western blotting to check if MSI4 has the 50 amino acids extension

To construct *in vitro* transcription/translation plasmids, PCR fragment (*Nco* I-*Sca* I fragment) of cDNA encoding MSI4 (aa 1-70, with the first 50 amino acids encoded) or MSI4 (aa 51-70, without the first 50 amino acids encoded) was sub-cloned into vector pET-28(b) (Novagen), together with the rest cDNA of MSI4 which was cut from candidate #133 plasmid by *Sca* I and *Xho* I. The PCR fragments were inserted directly into *Nco* I site in vector pET-28(b) which is right after the ribosomal binding sequence, so that His and T7 tags were eliminated to make sure that the translated products only

contained MSI4 protein. Both the PCR fragments in the constructed plasmids were DNA sequenced later to make sure no mutation was introduced.

The *in vitro* translated proteins were made by using TNT T7 Quick Coupled Transcription/Translation system (Promega) and the constructed plasmids. The cell lysate of *Arabidopsis* leaf suspension cell culture was prepared as Chapter II.

For performing Western blotting, 1 µl or 4 µl cell lysate 1.25 µl from translated TNT system were resolved in 10% SDS-PAGE gel and transferred to nitrocellulose membrane. The MSI4 proteins were detected by standard Western blotting method. The first antibody was provided by Dr. William Folk at the Department of Biochemistry in University of Missouri-Columbia. The antiserum was raised in rabbit against antigen MSIA4 (with the first 50 amino acids missing) and affinity purified.

Prey plasmid (library cDNA plasmid) DNA preparation

Positive colony candidates were picked up from the initial screening plates. They were all patch-streaked on solid medium SD (Trp-Leu-His-Ade-) plates. Then each grown candidate was streaked again on SD (Trp-Leu-His-Ade-) plates to isolate single colony. Single colony from each streaked candidate was grown in liquid medium SD (Trp- Leu- His- Ade-), and the plasmid was extracted from the liquid culture. For plasmid extraction, the yeast cells was spun down from 1.5 ml overnight culture, and then the cells were mixed with 200 µl QAD buffer (2% Triton X-100, 1%SDS, 100 mM NaCl, 10 mM Tris-HCl, pH 8.0, 1 mM EDTA), about 100 µl glass beads and 100 µl phenol/chloroform. The cell mixture was vortexed for 2 min. The supernatant was collected after 5 min centrifugation at highest speed on bench centrifuge. Plasmid DNA

was precipitated by 2 x volume of ethanol, and dissolved in 70 µl 1x TE.

To recover the prey plasmid, *E. coli* (DH5α strain) was transformed by the plasmid extraction obtained above and grown on solid LB/Ampicillin plate. Single colony from solid LB/Ampicillin plate was grown in 5 ml liquid LB/Ampicillin medium. The prey plasmid was then miniprepared from the *E. coli* culture.

Confirmation of positive candidates

To confirm the candidates, AH109 yeast cells were transformed again with the prey plasmid and empty plasmid pGBKT7 (only the Gal4 DNA-binding domain) or bait plasmid. The transformed yeast cells were grown on SD Trp- Leu- solid medium first to select positive transformants and then the grown colonies were streaked onto SD Trp- Leu- His- Ade- /X-α-gal solid medium and allowed to grow for 4 days at 30°C.

Yeast cell plasmid transformation (mini-scale)

All the procedure was carried out at room temperature unless indicated. AH109 cells were grown in 10 ml liquid YPDA medium overnight at 30°C. The YPDA medium was YPD medium (20 g/l Difco peptone and 10 g/l yeast extract) supplemented with 0.003% adenine hemisulfate (Sigma #A-9126). The overnight culture was then transferred to fresh 60 ml YPDA for about 3 hours until OD600 reached about 0.6. The cells were then spun down at 1000 x g and washed with 20 ml H_2O. The cell then was resuspended in 1x TE buffer (10 mM Tris, 1 mM EDTA, pH 7.5, 100 mM LiAc). About 0.5 µg DNA each of prey or/and bait plasmid was mixed gently with 100 µg Herring testes carrier DNA, and then was vortexed together with the 100 µl prepared yeast suspension. 600 µl

PEG/LiAc/TE buffer (TE buffer plus 40% PEG MW4000) was then added to the cell mixture by vortexing. The mixture was incubated at 30°C for 30 min at gentle shaking. 70 µl DMSO was added to the cell mixture and gentle inversion was performed to mix DMSO evenly. The cell mixture was then heat shocked at 42°C water bath for 15 min and then cold shocked on ice for 5 min. The cells were the recovered by brief centrifugation and resuspended in 200-300 µl 1x TE buffer. About 150 µl suspended cells were spread on appropriate selective solid medium.

GST pull-down assay.

The detailed procedure was described in chapter II.

Results

Yeast two-hybrid screening for proteins interacting with GCN5 (1-210)

To understand the function of the N-terminus of *Arabidopsis* GCN5, one convenient way is to identify proteins interacting with the N-terminus. The yeast two-hybrid screening strategy was employed in search for interacting proteins with the N-terminus.

The MATCHMAKER Two-Hybrid System 3 (Clontech) was used for yeast two-hybrid screening. This system is based on the Gal4 activation domain and DNA-binding domain. The system includes bait plasmid pGBKT7 (which encodes the Gal4 DNA binding domain GDBD, and carries the Trp1 selective marker) and prey plasmid pGADT7 (which encodes the Gal4 activation domain GAD, and carries the Leu2 marker). The yeast host strain AH109 has four auxotrophic mutations (tryptophan, leucine, histidine and adenine), and carries three Gal4 responsive integrated reporter genes to

detect bait-prey protein interaction: *ADE2*, *HIS3*, and *MEL*1. *MEL*1 encodes α-galactosidase which is secreted into the medium and turned X-α-gal into blue color. These three selection criteria have an advantage of eliminating most false positive interactions between bait and prey proteins.

The cDNA encoding the amino acids 1-210 of the N-terminus of GCN5 was inserted into the plasmid pGBKT7 in frame with that of GDBD, creating the bait plasmid pGDBD-GCN5 (1-210). The cDNA library was provided by Dr. Peter Quail at the Department of Plant and Microbial Biology, University of California, Berkeley. The cDNA library was derived from 3-day-old etiolated *Arabidopsis* seedlings, in which the first strand cDNA was synthesized using oligo (dT) primer and afterwards the library cDNA was inserted into plasmid pACT. Plasmid pACT was an early version of plasmid pGADT7. The major difference is that GADT7 has more restriction enzyme choices at the multiple cloning sites and that there is a HA epitope tag at the end of Gal4 activation domain.

Before starting the two-hybrid screening, the bait protein GCN5(1-210) was first checked in host AHA109 for its authentic expression, non-toxicity to host strain, and incapability to activate reporter genes. Strain AH109 was transformed with plasmid pGDBD-GCN5 (1-210) to check the expression of bait protein. The transformed cells were grown in liquid synthetic dropout medium lacking tryptophan minus (SD Trp-) overnight, and then the cells were lysed using glass beads. Standard Western blotting was performed to detect bait protein using the Gal4 DNA binding domain specific antibody, LA2-3. Control cell lysate from AH109 cells transformed with the empty plasmid pGBKT7 was also included in the Western blotting experiment to differentiate bait

protein size. As shown in Figure 1, the Western blotting result showed that bait protein GCN5 (1-210) in fusion with Gal4 DNA-binding protein was expressed in host strain of AH109.

To test whether the bait protein GDBD-GCN5(1-210) could not activate reporter genes, yeast cells transformed with bait plasmid pGDBD-GCN5(1-210) were grown on solid medium SD (Trp-, SD Trp- His-) and SD (Trp- Ade-). The transformed yeast cells grew only on SD (Trp-), but not SD (Trp- His-) and SD (Trp- Ade-) solid media after 4 days incubation at 30°C (Figure 2). This result confirmed that GCN5 (1-210) could not activate reporter genes. On the other hand, yeast cells with plasmid pGDBD-GCN5 (1-210) on SD (Trp-) solid medium appeared similar in size to that of yeast colonies with plasmid pGBKT7 on SD (Trp-) solid medium, confirming that the bait protein was not toxic to the yeast cells (Figure 2).

After it was clear that bait protein GCN5 (1-210) was not toxic, nor did it activate reporter genes in yeast host strain AH109, yeast two-hybrid screening was performed. The yeast AH109 cells, with the bait plasmid pGDBD-GCN5(1-210), was transformed with the cDNA library following the protocol of library transformation recommended for the MATCHMAKER Two-Hybrid System 3. The transformants were spread on solid medium of SD (Trp- Leu- His- Ade-), in which markers Trp- Leu- were to keep the prey and bait plasmids and markers His- Ade- were to select positive candidates. A small portion of transformed cells was spread on non-selective solid medium SD (Trp- Leu-) to estimate the total transformants. About one million transformed yeast cells were screened on selective solid medium. The cells on medium SD (Trp- Leu- His- Ade-) were allowed to grow for two weeks at 30°C.

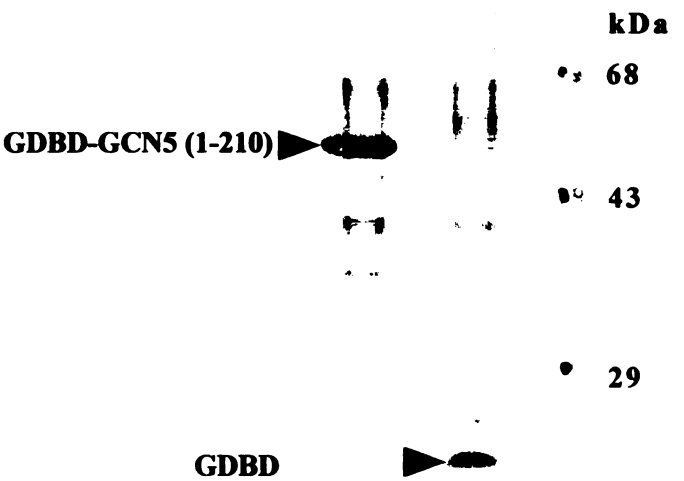


Figure 1. Bait protein GDBD-GCN5 (1-210) was expressed in yeast host strain AH109.
Cell lysates from yeast cells transformed with plasmid pGBKT7 or plasmid pGDBD-GCN5(1-210) were resolved in 10% SDS-PAGE gel, and the proteins in gel were transferred to nitrocellulose membrane. The protein GDBD and fusion protein GDBD-GCN5 (1-210) were detected by Western blotting using rabbit antiserum LA2-3 which had been raised against antigen GDBD in fusion with VP16 activation domain. The positions of both GDBD and pGDBD-GCN5 (1-210) are indicated by the filled triangles. GDBD: Gal4 DNA binding domain.

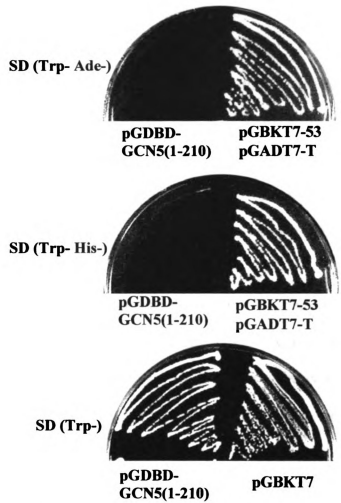


Figure 2. The protein GCN5(1-210) was not toxic, and did not activate reporter genes in yeast host strain AH109.
The transformed yeast cells with respective plasmids were grown on solid SD medium of Trp-, or Trp- Ade-, or Trp- His- for 4 days at 30°C. Plasmid pGDBD-GCN5(1-210) expressed bait protein GCN5 (amino acids 1-210) in fusion with Gal4 DNA binding domain (GDBD), plasmid pGBKT7 expressed GDBD only, and plasmid pair pGBKT7-53 and pGADT7-LT, as positive control, expressed p53 protein and SV40 large T antigen. Trp- medium was to maintain plasmids pGBKT7 and pGDBD-GCN5 which had a $Trp1^{+}$ marker.

Since both selective markers His- and Ade- were included, the screening was under high stringent selection. Totally 102 positive candidate colonies were picked up from the selective medium during the 2-week incubation period.

The cDNA prey plasmids were recovered from the positive yeast colonies. First, the plasmids (including bait and cDNA prey plasmids) were prepared from the 102 candidate yeast cells by standard plasmid extraction method. Then, *E. coli* cells were transformed with the prepared plasmids and were grown to single colony on solid LB medium with ampicillin to select cDNA prey plasmid. The cDNA prey plasmids were prepared from the transformed *E. coli* by minipreparation.

To ensure that the cDNA encoded prey protein interacted only with bait protein GCN5 (1-210) but not the Gal4 DNA binding domain, the prey plasmids were transformed again to yeast AH109 cells which carried either plasmid pGBKT7 or bait plasmid pGDBD-GCN5 (1-210). The transformed yeast cells were grown on SD (Trp- Leu-) solid medium first to keep the bait and prey plasmids and subsequently were streaked on solid medium SD (Trp- Leu- His- Ade- /X-a-gal) to test bait-prey interaction. However, the test result showed that proteins from all 102 candidate prey plasmids showed interaction with both the Gal4 DNA binding domain and GDBD-GCN5 bait. This result indicated that either these 102 candidates interacted with GDBD only or they alone could activate reporter genes. None of them had specific interaction with bait GCN5 (1-210).

The screening markers above used two reporters (Ade- His-) in the initial screening. This high screening stringency might have missed proteins which interacting weakly with

bait GCN5 (1-210). Therefore, a second round of yeast two-hybrid screening was repeated with lower selective stringency using only the His- selective marker.

About 2 million yeast transformants were screened on SD (Trp- Leu- His-) solid medium. Transformed yeast cells were allowed to grow for 10 days at 30°C. About 1500 positive colonies were picked up from the SD (Trp- Leu- His-) solid medium and patch-streaked on SD (Trp- Leu- His- Ade- /X-α-Gal) solid medium. Apparently, lower selective stringency retrieved more initial positive candidates. But only about 80 candidates grew well and turn blue color on high stringent selective medium SD (Trp- Leu- His- Ade- /X-α-Gal). For these 80 candidates, prey plasmids were prepared from the yeast cells, and they were checked, as done before, for the interaction with GAL4 DNA binding domain only. The result showed that all of the 80 candidates interacted with both the Gal4 DNA binding domain and the bait protein GCN5 (1-210). This result again indicated that no candidates interacting specifically with GCN5 (1-210) were identified when the initial screening stringency was lowered.

Yeast two-hybrid screening for proteins interacting with full-length GCN5

The two genetic screens described above uncovered no potential positive candidate when the N-terminus of GCN5 was used as the bait. One possible defect using the N-terminus of GCN5 as the bait was that the N-terminus might not be in its native conformation when the rest of GCN5 was absent. Therefore, the full length of GCN5 was employed as bait to repeat the yeast two-hybrid screening. There was another advantage of using the full length GCN5 as bait. GCN5 in full length was known to interact with proteins ADA2a and ADA2b as shown in chapter III. ADA2a and ADA2b candidates

should be screened out from the library when GCN5 full length was used as the bait. Therefore, successful recovery of ADA2a and ADA2b candidates from the cDNA library served as positive control, in making sure that the two-hybrid screening had no systemic problem.

The cDNA of full length GCN5 was inserted into plasmid pGBKT7 in frame with that of Gal4 DNA binding domain to construct bait plasmid pGDBD-GCN5. As done with the N-terminus above, the full length of GCN5 protein bait was checked first for non-toxicity in the host strain as well as its incapability to activate reporter genes. Yeast AH109 cells were transformed with bait plasmid pGDBD-GCN5 and prey plasmid pACT1or pACT2. These two prey plasmids were purified randomly from the cDNA library with unknown library cDNA insertions. The transformed yeast cells, compared with that of positive controls, grew well on SD Trp- Leu- solid medium after 4 days culture at 30°C, confirming the bait GCN5 protein was not toxic to host yeast cells. Meanwhile, the transformants failed to grow on selective medium of SD (Trp- Leu- His), SD (Trp- Leu- Ade-), and SD (Trp- Leu- His -Ade-) solid medium while the positive control did, confirming that GCN5 protein did not activate reporter genes (Figure 3).

After clarifying that bait GCN5 protein was neither toxic nor capable of activating reporter genes, a third round of library screening was carried out. Yeast AH109 cells with bait plasmid pGDB-GCN5 were transformed with 200 μg library DNA. The transformed yeast cells were spread on SD (Trp- Leu- His-) solid medium (low stringency) and incubated for 13 days at 30°C. Totally about 1.8 million transformants were screened, and about 550 initial positive candidates were picked up and patch-streaked on SD (Trp- Leu- His- Ade- /X-α-Gal) solid medium to confirm their candidacy. Based on the criteria

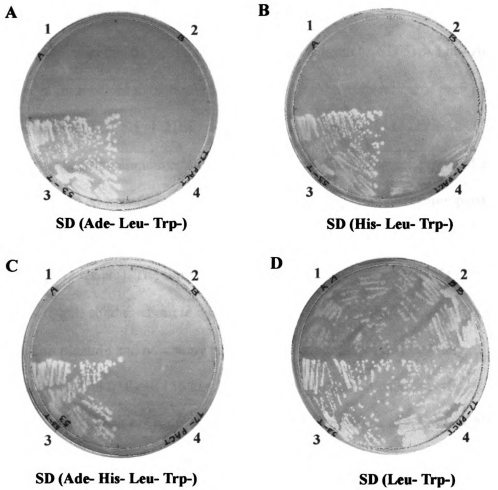


Figure 3. The protein GCN5 (full length) was not toxic, and did not activate reporter genes in yeast host strain AH109.
Yeast AH109 cells co-transformed with respective plasmids were grown on on the medium as indicated at 30 °C for 4 days before the pictures were taken. 1: pGDBD-GCN5+pACT1, 2: pGDBD-GCN5+pACT2, 3: pGBKT7-p53+pGADT7-LT, 4: pGBKT7+pACT1. Bait plasmid pGDBD-GCN5 expressed bait protein GCN5 in fusion with Gal4 DNA binding domain (GDBD). Prey plasmid pACT1and pACT2 were purified randomly from the cDNA library, which were empty pACT vectors with possible unknown library cDNA insertions, therefore expressed Gal4 activation domain GAD in fusion with possible unknown protein. The plasmids pGBKT7-p53 (expressing p53 protein) and pGADT7-LT (expressing SV40 large T antigen) were positive control, and plasmids pGBKT7 (expressing GDBD) and pACT1 were negative control.

of good colony growth and capability of turning blue color on SD (Trp- Leu- His- Ade- /X- α -Gal) medium, totally 156 potential candidates were recovered.

These156 potential candidates were colony-streaked again on SD Trp- Leu- Ade- His- /X-α-gal solid medium to isolate single colonies for the preparation of prey plasmids as done for that of bait GCN5 (1-210).

To confirm specific interaction with GCN5, the prey plasmids were prepared as described above and co-transformed into yeast AH109 cells again with empty plasmid pGBKT7 (expressing Gal4 DNA-binding domain only) or bait plasmid pGDBD-GCN5 (expressing GDBD-GCN5). The transformed yeast cells were streaked on SD (Trp- Leu- Ade- His- /X- α-gal) solid medium to confirm real candidacy by the criterion of colony growth with bait plasmid and no colony growth with empty plasmid after 4-day incubation at 30°C. Out of these 156 candidates, 28 candidates showed specific interaction with GCN5 bait only, the rest showed either interaction with both GCN5 bait and GDBD or no interaction with either GCN5 bait or GDBD.

Among these 28 candidates, candidates #2, 3, 4, 5, 7, 8, 9, 10, 11, 13, 16, 20, 26, 30, 73, 77, 104 and 133 showed big-size-colony growth and bright blue colonies with bait GCN5 (some of them were shown on Figure 4). Candidates #81, 145, 137, 141, 155, and 161 showed big-size-colony growth and moderate blue colonies with bait GCN5. Candidates 82 showed moderate colony growth and moderate blue colonies with bait. Candidate #38, 46, 106 showed small colony growth and little blue colonies with bait.

DNA sequencing indicated that candidates #2, 3, 4, 5, 7, 9, 10, 11, 13, 16, 20, 26, 30, 104, were cDNAs of ADA2b, and # 8 was that of ADA2a. Details of these cDNAs related to the full length of *Arabidopsis* ADA2a or ADA2b are shown in Figure 5. This

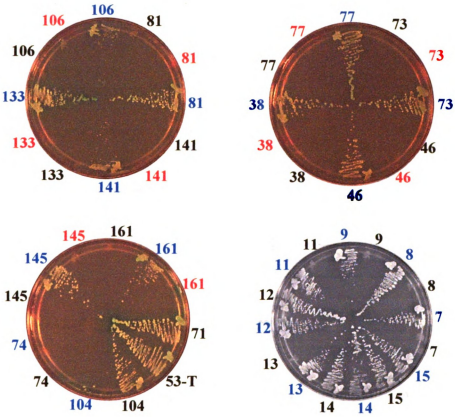


Figure 4. Confirmation of the candidates from the two-hybrid screening using the full-length GCN5 bait.

These candidate plasmids (as indicated by the numbers in the pictures) were co-transformed to yeast strain AH109 cells with plasmid pGDBD-GCN5 (expressing GCN5 full length in fusion with Gal4 DNA binding domain), pGDBD-GCN5 (1-210) (expressing the N-terminus of GCN5 in fusion with Gal4 DNA binding domain) or pGBKT7 (expressing Gal4 DNA binding domain), and the transformed cells were grown on selective solid medium SD Ade- His -Leu- Trp- /x-a-Gal for 4 days at 30°C. The candidates numbers are colored differently to indicated the co-transformed plasmids where the red color is for pGDBD-GCN5 (1-210), blue color for pGDBD-GCN5, and black color for pGBKT7. Candidate #74 did not interacted with GCN5 and Gal4 DNA binding domain, and #104 interacted only with the Gal4 DNA binding domain. 53-T was a positive control of p53 protein and SV40 large T antigen.

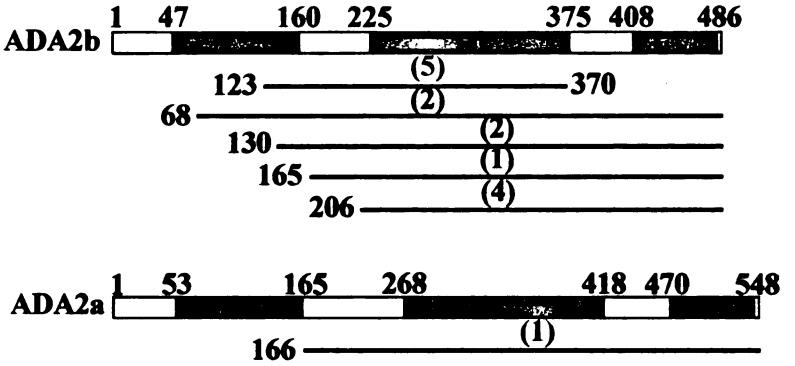


Figure 5. The isolated ADA2a and ADA2b candidates by GCN5 (fl) bait in yeast two hybrid screening.
Each line represents the ADA2a or ADA2b isolate from yeast two-hybrid screening with the GCN5 full length as the bait. The length of each line indicates the relative region encoded by the candidate cDNA in respective of the full length of protein ADA2a or ADA2b. The numbers are the amino acid positions for respective proteins. The numbers in the brackets indicate the number of isolates of each prey plasmid. The conserved N-terminus, the middle part and the C-terminus of ADA2a and ADA2b protein are represented by the shaded areas inside the two bars.

result illustrates that the performed yeast two-hybrid screening was technically working.

Candidates #38, 46, 73, 77, 81, 106, 133, 141, 145, 161 were checked again for their specific interaction with GCN5. The candidate plasmids were co-transformed into yeast AH109 cells with plasmid pGDBD-GCN5, pGDBD-GCN5 (1-210) and pGBKT7. The transformants were grown on SD (Trp- Leu- Ade- His- /X- α -gal) solid medium. They showed interaction only with GCN5 in full length. None of them interacted with the N-terminus of GCN5 (Figure 4).

These remaining candidates were DNA sequenced. A summary of these candidates is listed in Table 5-1, including the putative proteins and their functions from Blast search of GenBank. Among all the candidates, candidates #133 and #81 were most interesting in considering the apparent function of *Arabidopsis* GCN5 as histone tail modifier.

Candidate #133

The #133 cDNA encoded a 507 amino acid protein. The cDNA contained the whole open reading frame as well as the polyA tail and upstream untranslated 5' sequences. The protein encoded by candidate #133 cDNA was identical to the *Arabidopsis* protein named MSI4 (Kenzior and Folk, 1998) except that #133 had additional 50 amino acids at its N-terminus. *Arabidopsis* MSI4 is homologous to human protein RbAp46/48 (for Rb associated protein), which belongs to the Msi1-like family with a featured domain called WD-40.

Many Msi1-like proteins are components of complexes involved in the assembly and modification of chromatin (Vignali, et al., 2000). Msi1, RbAp48 and related the related protein p55 are components of chromatin assembly complexes of yeast, humans and

Table 5-1 Summary of the Candidates

candidate	GenBank ID	candidate encoded potein size	GenBank protein, size	related protein or domain
#46	15231656	122-287 aa	pirin-like protein, 287 aa	pirin[A]
#73	15081612	356-579 aa	Pyruvate kinase	
#77	15234187	28-315 aa	bHLH protein, 315 aa	helix-loop-helix DNA-binding domain
#81	30686609	1-~260 aa [&]	411 aa	modified RING finger domain immediate-early fungal elicitor protein CMPG1
#106	15234136	72-262 aa	ribosomal protein L5p family, 262 aa	
#133	30680701	1-507 aa [§]	MSI4. 507 aa	WD-40 repeat. RbAp48/46 like proteins
#137	11288388	70-~500 aa	expressed, 576 aa	chromosome segregation ATPases (Cell division and chromosome partitioning
#141	7486967	190-501 aa	ethylene-responsive RNA/DNA helicase, putative. 501 aa	Srmb, DEAD, DEAD/DEAH box helicase. members of this family include the DEAD and DEAH box helicases.
#145, #155	15221998	867-~1120 aa	hypothetical protein 1396 aa	tripeptidyl peptidase II, peptidase [B]
#82, #161 #38	166688 99742	113-525 aa 208-525 aa	3-deoxy-D-arabino-heptulosonate y-phosphate synthase. 525 aa	class-II DAHP synthetase family. Members of this family are aldolase enzymes that catalyse the first step of the shikimate pathway. [C]

A: Pirin. This family consists of Pirin proteins from both eukaryotes and prokaryotes. The function of Pirin is unknown but the coding for this protein is known to be expressed in all tissues in the human body although it is expressed most strongly in the liver and heart. Pirin is known to be a nuclear protein, exclusively localised within the nucleoplasma and predominantly concentrated within dot-like subnuclear structures. A tomato homologue of human Pirin has been found to be induced during programmed cell death. B: Subtilase family. Subtilases are a family of serine proteases. They appear to have independently and convergently evolved an Asp/Ser/His catalytic triad, like that found in the trypsin serine proteases. Structure is an alpha/beta fold containing a 7-stranded parallel beta sheet. C: Helicases are involved in unwinding nucleic acids. The DEAD box helicases are involved in various aspects of RNA metabolism, including nuclear transcription, pre mRNA splicing, ribosome biogenesis, nucleocytoplasmic transport, translation, RNA decay and organellar gene expression. &: aa 1-36 encoded by 5' cDNA out of ORF were necessary for interaction with GCN5. §: aa 1-20 encoded by 5' cDNA out of ORF were necessary for interaction with GCN5.

Drosophila respectively. The *Drosophila* p55 protein also is a subunit of NURF, an ATP-dependent nucleosome remodeling complex. In humans, Rbap48 is a retinoblastoma interacting protein that acts as an apparent histone H4 chaperone (Roth and Allis, 1996) and a subunit of human chromatin assembly factor CAF-1(Verreault, et al., 1996).

As mentioned, candidate #133 had additional 50 more amino acids at its N-terminus compared to the reported *Arabidopsis* MSI4 (Kenzior and Folk, 1998). The first 60 amino acids of candidate #133 were MESDEAAAVS PQATTPSGGT GASGPKKRGR KPKTKEDSQT PSSQQQSDVK MKESGKKTQQ.

To test whether the endogenous MSI4 had the 50 amino acids extension, MSI4 protein from *Arabidopsis* suspension cell lysate was compared to that of *in vitro* translated MSI4 with or without this 50 amino acids extension by Western blotting. To construct *in vitro* expression plasmids, PCR products from candidate #133 encoding protein MSI4 with or without the 50 amino acids were subcloned into His-tag expressing vector pET28 where the His tag was removed so that the expressed protein contained no extra amino acids. The two proteins were expressed *in vitro* using the constructed plasmids and the TNT T7 transcription/translation system (Promega). The *Arabidopsis* suspension cell lysate was prepared previously (Chapter II). The Western blotting result, using MSI4 specific antibody, showed that endogenous MSI4 did have this 50 amino acids extension (Figure 6A), where protein size of endogenous MSI4 was the same as that of *in vitro* translated MSI4 with the 50 amino acids. The protein size of *in vitro* translated MSI4 without the 50 amino acids was smaller than that of endogenous MSI4.

Candidate #133 cDNA included 60 nucleotides 5' to the start ATG codon of MSI4. These 60 nucleotides encoded 20 amino acids in frame following the cDNA of Gal4

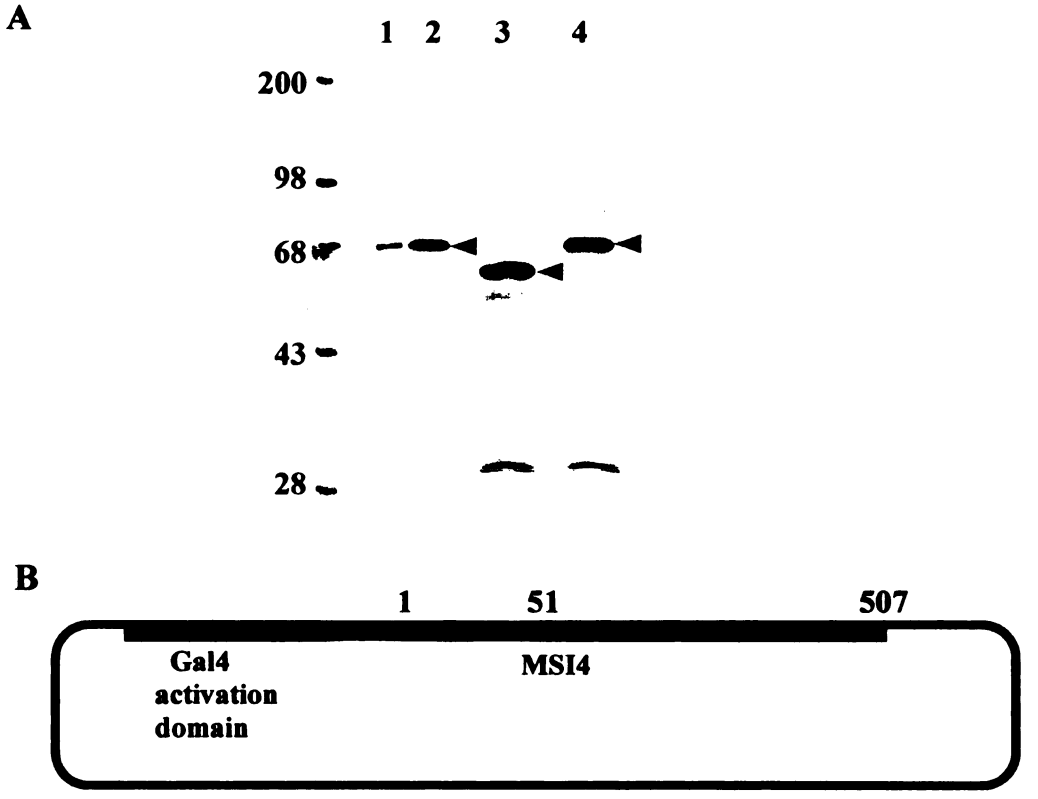


Figure 6. Native protein of *Arabidopsis* MSI4 contains the 50 amino acids in its N-terminus.
(**A**) Western blotting was carried out to confirm that the native MSI4 protein had the 50 amino acids extension. The native endogenous MSI4 protein was from the *Arabidopsis* leaf suspension cell lysates. The *in vitro* MSI4 proteins (amino acids 1-507 or 51-507) were made in TNT T7 Quick Coupled Transcription/Translation system (Promega) using their respective expressing plasmids. The protein samples were resolved in 7.5% SDS-PAGE gel and the proteins were detected by Western blotting using specific antibody against MSI4 (51-107). 1: 1ul of Arabidopsis leaf suspensin cell lysate, 2: 4ul of *Arabidopsis* leaf suspensin cell lysate, 3: in vitro translated MSI4 protein (aa 51-507), 4: in vitro translated MSI4 protein (1-507). The blue arrow points to the MSI4 proteins. **(B)** The diagram of candidate #133 prey plasmid. Various encoding regions are indicated in the plasmid, including the regions encoding the first 50 amino acids of native MSI4 protein (gray area) and the extra leading 20 amino acids (blue area). The numbers refer to the amino acids positions in the MSI4 protein.

activation domain (Figure 6B). Therefore, it was legitimate to determine if these 20 amino acids were necessary for the interaction with GCN5 bait protein.

To test this necessity, a yeast two-hybrid assay was performed. The cDNA of MSI4 (full length or missing the first 50 amino acids) were subcloned into yeast prey vector pGADT7 in frame with that of the Gal4 activation domain. Yeast AH109 cells were transformed by bait plasmid pGDBD-GCN5 and respective constructed MSI4 plasmids, and grown on selective solid medium SD (Trp- Leu- His- Ade- /X-α-gal) for 4 day at 30°C. Surprisingly, as shown in Figure 7, candidate #133 lost interaction with GCN5 when these 20 amino acids were not present. The control #133 with these 20 amino acids still showed interaction with GCN5. Therefore the 20 amino acids were necessary for the interaction with protein GCN5, and very likely that protein MSI4 does not interact with GCN5, although the MSI4 protein expression in yeast cells was not checked. The 20 amino acids sequence were RYRDTESKDQKRKIKKRERK. (Note: The 5 amino acids encoded by the adaptor sequence for Xho I, which was used to construct the cDNA library, were not included in these 20 amino acids.)

Candidate #81

Based on the experience with candidate #133, other were checked for the existence of extra 5' sequence. Only candidate #81 was found to have 5'untranslated sequence, which encoded 41 amino acids in frame with Gal4 activation domain and prior to the legitimate start codon.

As for candidate 133, a yeast two-hybrid assay was performed to check the involvement of these 41 amino acids interaction with GCN5. Indeed these 40 amino acids

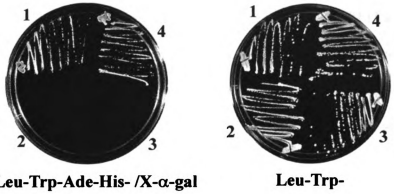


Figure 7. Candidate #133 lost interaction with GCN5 when the extra leading 20 amino acids were deleted.

Yeast two-hybrid assay was performed to check the interaction of GCN5 with the different lengths of candidate #133. Yeast AH109 cells were co-transformed with bait plasmid pGDBD-GCN5 and the respective constructed #133 candidate plasmids. The transformed yeast cells were grown on selective solid medium SD (Trp- Leu-His- Ade- /X-α-gal) for 4 day at 30°C.The transformed cells were also grown on non-selective solid medium SD (Trp- Leu-)to confirm proteins were not toxic to the host cells.

1: pGAD-ADA2b+pGDBD-GCN5 (positive control),
2: pGAD-MSI4 (1-507)+pGDBD-GCN5,
3: pGAD-MSI4 (50-507)+pGDBD-GCN5.
4: #133+pGDBD-GCN5,
#133: the original prey plasmid from the library screening.

were necessary for interaction with GCN5 bait protein (data not shown), and candidate #81 also did not interact with GCN5 legitimately. The 41 amino acids were PRDLPPNPTTTTCIKKKIYSLSFFDQYIRQGSRQSPQHLHH.

In conclusion, candidate #133 and #81 were false positives. Curiously, there is a common sequence in the cDNA encoded extra peptide of candidate #133 and #81: IKKK in #81 and IKKR in #133. It is not known if this common sequence is responsible for the interaction with GCN5 bait protein. However, none of the other candidates have this common sequence. But curiously, *Arabidopsis* ADA2a does have IKKK at its C-terminus.

Candidate #77

Candidate #77 encodes a putative bHLH transcriptional activator. The full length cDNA in GenBank encodes a protein of 315 amino acids (termed as 77fl). The candidate isolated by the two-hybrid screen encoded residues 29-315.

To test which part of GCN5, except the N-terminus, interacted with protein 77fl, an *in vitro* GST pull-down assay was performed. Recombinant GST fusion proteins of various GCN5 parts bound to Sepharose beads were tested to pull down protein 77fl which was *in vitro* translated and radiolabeled by [S^{35}] methionine. As shown in Figure 8, protein 77fl showed interaction with GCN5, the HAT domain and the H-B region but not the bromodomain. A yeast two-hybrid assay also was performed to test various GCN5 parts' interaction with 77fl, and two overlapping regions s of GCN5 (1-250, and the HAT domain) were found to interact with 77fl (Figure 10). These results suggested that the HAT domain of GCN5 interacted with 77fl.

Since 77fl had a putative DND binding domain, it was also tested for being a

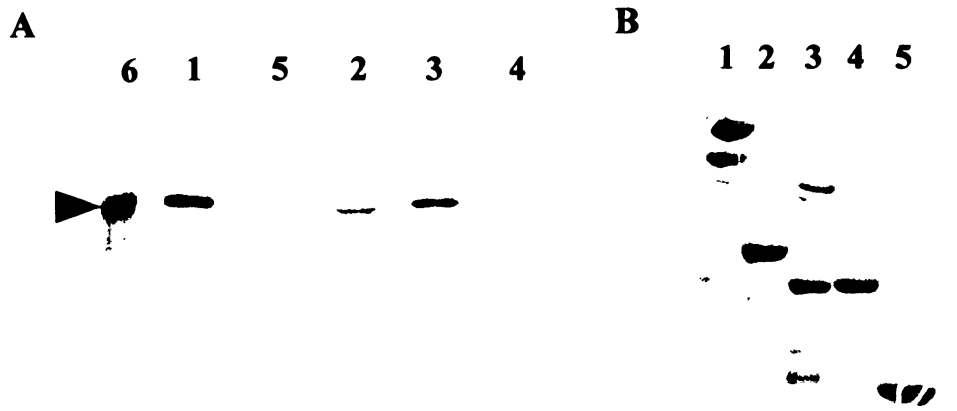


Figure 8. The full length of candidate #77 protein interacted with the HAT domain and the the H-B region of GCN5.
An *in vitro* GST pull-down assays were performed. Various recombinant GST-GCN5 proteins bound to Sepharose beads were tested to pull down candidate #77 in full length which was translated *in vitro* and radiolabeled by [S^{35}] methionine using the TNT T7 Translation/Transcription Couple system (Promega). The pulled-down proteins were eluted from the beads by boiling in SDS-PAGE loading buffer. Then the eluates were electrophoresed in 10% SDS-PAGE. The gel was then dried to expose to X-ray film for autoradiogram, as shown in (**A**). The amount of GST-GCN5 proteins as well as control GST protein used in each pull-down assays were shown in (**B**) that were stained with Coomassie Blue. The arrow points to the #77 protein in full length. 1: GST-GCN5, 2: GST-GCN5 (HAT, 203-384), 3: GST-GCN5 (H-B, 360-486),4: GST-GCN5 (Br, 451-568), 5: GST, 6: 50% input.

transcriptional activator. Yeast one-hybrid assay was performed where truncated versions of 77fl were fused to Gal4 DNA binding domain to test their potentiality activating reporter genes. As shown in Figure 9, protein 77 could activate reporter genes in yeast one-hybrid assays. The N-terminus (1-82) was sufficient to activate reporter genes. This indicated that the N-terminus contained an activation domain. Based on yeast colony growth, the full length and the part including the helix-loop-helix DNA-binding domain (1-173) did not activate as well as that of N-terminus (1-82). This might be attributed to the variation of protein expression level, the presence of repression domain(s), or that the DNA binding domain sequestered the expressed protein elsewhere in the yeast genome. Further research of candidate 77 was not pursued.

Screening library CD4-10 using GCN5 (full length) as the bait

After these three rounds of two-hybrid screening, no potential proteins interacting with the N-terminus of GCN5 were identified. The mRNAs to make the cDNA library used above were made from 3-day-old etiolated *Arabidopsis* seedlings, and the first strand cDNA was synthesized using oligo-T primer. This library might not include cDNAs for some genes in other growth conditions, such as light which is very important for plant growth. Therefore, a second cDNA library called CD4-10 was screened. This library was derived from adult *Arabidopsis* plant roots and leaves grown in light, and the first strand cDNA was synthesized using random primers. The phage library was provided by *Arabidopsis* Biological Research Center at Ohio State University. The plasmid cDNA library was prepared. The plasmid vector was pACT, the same as that of the cDNA library used above.

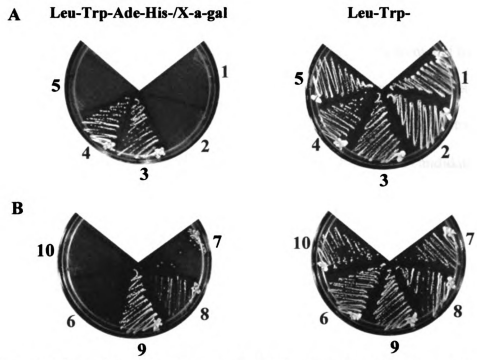


Figure 9. Candidate #77 was a potential activator and cinteracted with the HAT domain GCN5. It is a potential activator.

The candidate #77 cDNA encoded amino acids 29-315 of the full length protein of amino acids 1-315. The full length cDNA was cloned from the screened library by PCR, and various cDNA lengths as indicated were subcloned into vector pGBKT7 or vector pGADT7. Yeast AH109 cells co-transformed with respective plasmids were grown on selective solid medium SD (Trp- Leu- His- Ade- /X-α-gal) for 4 day at 30°C. The transformed cells were also grown on non-selective solid medium SD (Trp- Leu-) to check the expressed proteins were not toxic to the host cells.1, 2, 3, 4, 5, and 6 tested the interaction of GCN5 with candidate #77, and 7, 8, 9 and 10 tested candidate #77 as a potential transcriptional activator where the co-transformation of plasmid pGADT7 (carried the marker of Leu2+) was to make the transformed yeast cells grow on Leu- solid medium. 1: pGAD-77(FL)+pGDBD-GCN5(bromo), 2: pGAD-77(FL)+pGDBD-GCN5(230-468), 3: pGAD-77(FL)+pGDBD-GCN5(HAT, 203-384), 4: pGAD-77(FL)+pGDBD-GCN5(1-250), 5: pGAD-77(FL)+pGDBD-GCN5 (1-210), 6: pGAD-77(FL)+pGDBD-GCN5(H-B, 360-486), 7: pGADT7+pGDBD-77(FL), 8: pGADT7+pGDBD-77(1-173), 9: pGADT7+pGDBD-77 (1-82), 10: pGADT7+pGBKT7

Similar yeast two-hybrid screening was carried out. Yeast AH109 cells with bait plasmid pGDB-GCN5 were transformed with 600 μg CD4-10 library DNA. The transformed yeast cells were spread on SD (Trp- Leu- His-) solid medium and incubated for 11 days at 30°C to identify candidates. Totally about 2 million transformants were screened, and about 260 initial positive candidates were picked up and patch-streaked on SD (Trp- Leu- His- Ade- /X-α-Gal) solid medium. Totally 58 potential candidates were recovered from the second screening of patch-streaking based on their good growth and turning to blue color.

These 58 potential candidates were colony-streaked again on SD (Trp- Leu- Ade- His- /X-α-Gal) solid medium to isolate single colony, and prey plasmids were then prepared as described before. To confirm these candidates' specific interaction with GCN5 bait, the prepared prey plasmids were co-transformed into yeast AH109 cells again with empty plasmid pGBKT (expressing Gal4 DNA-binding domain, GDBD) or bait plasmid pGDBD-GCN5 (expressing GDBD-GCN5). The transformed yeast cells were streaked on SD (Trp- Leu- Ade- His- /X-α-Gal) solid medium. Out of these 58 candidates, all showed interaction with both the GCN5 bait and GDBD. This result illustrated none of the 58 candidates was real.

Testing interaction with PCATs

Arabidopsis has four CBP/p300 orthologs namely PCAT1, PCAT2, PCAT3 and PCAT4 (Bordoli, et al., 2001). These PCATs have about 600 amino acids at their C-terminus highly homologous to the C-terminus of CBP/300 (amino acids 1201-1807), including the C/H2 and C/H3 domains. The corresponding C/H2 and C/H3 regions are

amino acids 1078-1350 and 1596-1743 in PCAT1, and 986-1256 and 1507-1654 in PACT2. The homologous 600 amino acids region in PCAT2, like that of CBP/p300, was shown to have HAT activity and E1A binding property (Bordoli, et al., 2001).

As mentioned in the introduction, the N-terminus of mammalian GCN5 interacted with CBP/p300 *in vitro*. This led to the hypothesis that the N-terminus of *Arabidopsis* GCN5, although very different from its animal counterpart in sequence, interacts with PCATs. Yeast two-hybrid assay was performed to test the possible interaction, in which the full length of GCN5 was used as the bait.

The cDNA clones encompassing the full length of PCAT2 and C/H3 domain of PCAT1 were provided kindly by Dr. Richard Eckner from the Institute for Molecular Biology, University of Zurich, Zurich, Switzerland. Various cDNA fragments were amplified by PCR and inserted into vector pGADT7 in frame with the Gal4 activation domain. These cDNA fragments encoded PCAT2 regions of 1-704, 611-994, 977-1362, 1323-1654, and the C/H3 domain of PCAT1 (amino acids 1588-1745).

Yeast AH109 cells were co-transformed with GCN5 bait plasmid pGDBD-GCN5 together with the plasmids containing various parts of PCAT1 or PCAT2. The transformed cells were grown on SD (Trp- Leu- Ade- His- /X-α-Gal) solid selective medium for 4 days at 30°C as well as non-selective medium SD (Trp- Leu-). As shown in Figure 10, no interaction was found between GCN5 and any parts of PCAT1 or PCAT2.

Discussion

In summary, no proteins were identified that interact with the N-terminus of GCN5 protein through the strategy of yeast two-hybrid screening, leaving the function of GCN5

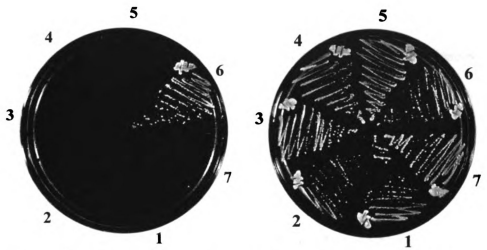


Figure 10. Testing the interaction of GCN5 with PCAT1 and PCAT2, the Arabidopsis orthologs of CBP/p300.
Yeast AH109 cells were co-transformed with GCN5 bait plasmid pGDBD-GCN5 (fl) together with the plasmids expressing various parts of PCAT1 or PCAT2. The transformed cells were grown on SD (Trp- Leu- Ade- His- /X-a-gal) solid selective medium for 4 days at 30°C as shown at the left, as well as on solid medium SD (Trp- Leu-) as shown at the right.
1: pGAD-PCAT1(C/H3) + pGDBD-GCN5,
2: pGAD-PCAT2(1-704) + pGDBD-GCN5,
3: pGAD-PCAT2(611-994) + pGDBD-GCN5,
4: pGAD-PCAT2(977-1362) + pGDBD-GCN5,
5: pGAD-PCAT2(1323-1654) + pGDBD-GCN5,
6: pGAD-ADA2b(fl) + pGDBD-GCN5 (positive control),
7: pGADT7 + pGDBD-GCN5 (negative control).

N-terminus an open question. The N-terminus of GCN5 also does not interact with the CBP/P300-like proteins from *Arabidopsis*.

For the first cDNA library, the screening was saturated. First, the same ADA2b candidates were isolated several times. Secondly, about 2 million yeast colonies were screened, whereas *Arabidopsis* has about 25,000 genes. For the CD4-10 library, the screening was very likely saturated but the library might not be comprehensive because no ADA2 candidates have been fished out. Attempt to amplify ADA2 cDNA by PCR directly from the library also failed (data not shown).

All the yeast two-hybrid screenings were carried out at temperature of 30°C. This could be one reason for failing to detect interacting proteins with GCN5, because some *Arabidopsis* proteins might not like this temperature in regard that the favorite ambient temperature for *Arabidopsis* plants to grow is below this temperature. Meanwhile, if the N-terminus is supposedly involved in plant response to abiotic and biotic stresses, the targets' cDNAs may not be present in the two cDNA library which were not made under stressful growth conditions. This might also lead to the failure to identify interacting protein with the N-terminus of GCN5.

Some function learned from the N-terminus of mammalian GCN5 might be interesting to mention here. The mouse GCN5 alone is able to acetylate nucleosomal substrates, suggesting that the unique N-terminal domains of mammalian P/CAF and GCN5 may provide additional functions important to recognition of chromatin substrates and the regulation of gene expression (Xu, et al., 1998). The same group has found that TRF2 (telomere related factor 2) can interact with the N-terminus of mouse GCN5 through yeast two-hybrid screening (personal communication).

Some potential candidates were identified that interact with the full length of GCN5, as listed in Table 6. Since the most two interesting candidates #133 and #81 turned out to be false positive, further investigation was not pursued for the rest of the candidates. However, discussion of some of the candidates is revealing.

Candidate #77 encodes a transcriptional activator with putative a helix-loop-helix DNA binding domain. The future identification of the cognate DNA sequence to which the native protein binds may reveal some roles played by the GCN5-ADA2 complex.

Candidate #141 encodes a putative ethylene-responsive DNA/RNA helicase protein. The whole protein has 501 amino acids, where the candidate #141 recovered from yeast two-hybrid encodes about 300 amino acids, beginning at amino acid 182. The protein contains a domain called DEAD/DEAH located at amino acids 101-305, and another overlapping domain called SrmB (from Superfamily II DNA and RNA helicases) that is located at amino acids 72-430. The protein is implicated as an ATP-driven RNA or DNA helicase participating in a variety of cellular processes such as DNA replication and recombination and repair, transcription, translation, ribosomal structure and biogenesis, and pre-RNA splicing.

At first, it is hard to understand why GCN5 might interact with this functionally remote protein. However, Roeder and colleagues have recently reported that the human SAGA-like complex STAGA has a loose association with protein SAP130 (Martinez, et al., 2001). Protein SAP130 is a component of the splicing factor SF3b, a U2 snRNP-associated protein complex that is essential for spliceosome assembly (Martinez, et al., 2001). They also find that STAGA associates with components of UV-damaged DNA binding complex (UV-DDB). UV-DDB recognizes UV-damaged DNA and links to DNA

damage repair. In addition, two transcriptional HAT complexes, the human TIP60 and yeast INO80, both contain proteins related to DNA helicase and both complexes have DNA helicase activity (Ikura, et al., 2000, Legube, et al., 2002, Shen, et al., 2000). Cells with impaired HAT activity in TIP60 have defective double-strand break repair (Ikura, et al., 2000), and yeast with null INO80 shows hypersensitivity to DNA damage (Shen, et al., 2000). Quite curiously, loss of GCN5 protein in mouse embryo results in increased apoptosis that is intimately connected to DNA damage (Xu, et al., 2000). Actually, the huge protein of Tra1 in SAGA belongs to the protein family of ATM (ataxia telangiectasia mutated)/DNA-PK/phosphatidylinositol 3-kinase implicated in cell cycle checkpoint signaling and cellular response to DNA damage (Abraham, 2001, Khanna, et al., 2001, Vassilev, et al., 1998). Taken together, these results implicate that GCN5-containg complex might also participate in DNA damage repair. All these functional connections might suggest that candidate #141 is a legitimate interacting protein with GCN5. The interaction of GCN5 with the candidate #141 is weak based on yeast colony growth in yeast two-hybrid assay. This might reflect that the interaction is transient.

Candidate #73 is a chloroplast pyruvate kinase isoenzyme. It seems unlikely that *Arabidopsis* GCN5, a histone acetyltransferase, also works in the chloroplast. Interestingly, an analysis of subcellular localization of the yeast proteins (Kumar, et al., 2002) shows that the majority of yeast GCN5 (70%) is in nucleus, while a small fraction (12%) is associated in mitochondria.

Candidate #137 encodes a protein related to chromosome segregation ATPase and membrane-bound metallopeptidase that are involved in cell division and chromosome partitioning. Candidate #145 and #155 also encodes protein that is similar to tripeptidyl

peptidase. It is curious why peptidases were isolated three times.

Candidate #46 encodes a Pirin-like protein. Pirin is a nuclear protein and the N-terminal half of Pirin is significantly conserved between mammals, fungi and plants (Wendler, et al., 1997). Human Pirin stabilizes the formation of complexes containing Bcl-3, the anti-apoptotic transcription factor NF-kappaB and its DNA target sequences *in vitro*. Moreover, and Bcl-3 mediated transcription is stimulated by HAT protein Tip60 (Dechend, et al., 1999), suggesting a role for HATs in Pirin-associated biological activities.

Chapter Six

The interaction between CBF1 and ADA2

Introduction

The expression of many genes is induced in *Arabidopsis* plants under cold stress. COR genes (for cold response) are among these genes which are actively transcribed when plant *Arabidopsis* is exposed to cold temperature, and are quickly turned off when the plant is shifted back to warm temperature. One of the transcriptional activators of COR genes is CBF1 (Stockinger, et al., 1997), (for C-repeat Binding Factor).

When protein CBF1 was assayed for transcriptional activity in budding yeast, its activity was dependent on yeast proteins Ada2, Ada3 and Gcn5 (Stockinger, et al., 2001) (Figure 1). These yeast proteins are components of HAT complexes ADA and SAGA (Grant, et al., 1997). Since CBF1 is a transcriptional activator of *Arabidopsis* and evidence suggests that *Arabidopsis* has similar ADA2 and GCN5 containing complex (es) *in vivo* (as discussed in Chapter III and IV), the experiments performed in this chapter were conducted to test whether there is functional interaction between *Arabidopsis* protein CBF1 and ADA2 or/and GCN5.

Experimental Methods

Plasmid construction

To construct GST-CBF1 expressing plasmids, cDNAs encoding CBF1 (amino acids 1-213, fl; 1-115, N; 115-213, C) were PCR amplified and inserted into GST vector pGEX-6P in frame with that of GST. The constructed plasmids were correspondingly

named pGEX-6P-CBF1(fl, 1-115 or N, 115-213 or C).

To constructed His6-tagged expressing plasmids, cDNAs fragments of respective proteins were inserted into vector pET-28 in frame.

To construct yeast two-hybrid plasmids of CBF1, cDNA fragments (encoding amino acids 1-78, 24-115, 48-115, 97-213, 78-213, and 48-78) were amplified by PCR and inserted into prey plasmid pGADT7 in frame to that of Gal4 activation domain. All the cDNAs fragments in constructed plasmids were checked by DNA sequencing to make sure no mutation was introduced. Other yeast two-hybrid plasmids were described in Chapter III.

For constructing prey plasmids expressing TINY (aa1-95), ERF4 (1-85), ERF5 (1-219), or DRAB2a (1-139), the corresponding DNA fragments were amplified directly from genomic DNA by PCR using specific primer pairs and subcloned into yeast two-hybrid prey vector pGADT7 in frame with Gal4 activation domain. All the cloned DNA fragments encodes AP2 domain. The sequences of the DNA fragments were confirmed to ensure the accuracy of PCR products.

Recombinant protein expression and purification

The expression and purification of GST-tag and His6-tag proteins of ADA2b, GCN5 and CBF1 were essentially as described in chapter IV. The GST tag was subsequently cut off with protease thrombin (for CBF1) or PreScission protease (for GCN5 and ADA2b). PBS buffer with 10% glycerol and 1 mM DTT or HEPES buffer (40 mM HEPES, 50 mM NaCl, pH 7.3, 10 % glycerol and 7 mM 2-mercaptoethanol) was used dependent on individual purification procedure.

Recombinant His6-CBF1 and His6-ADA2b were purified using Ni-NTA beads (Qiagen). The eluted protein solution was dialyzed against HEPES buffer or otherwise indicated specifically.

GST pull-down assay

To express *in vitro* radiolabeled proteins, about 0.5-1 μg of the corresponding pET-28 vector-derived plasmids was mixed with Transcription/Translation coupled TNT T7 system (Promega) in a final volume of 50 μl that included [^{35}S] methionine. The mixture was allowed to react for two hours at 30°C.

Depending on the specific experiment, about 1-2 μg GST or GST fusion protein was bound to glutathione Sepharose beads by incubating the beads with appropriate amount expressed *E. coli* cell lysate for about one hour and subsequent washing with HEPES buffer (40 mM HEPES pH 7.4, 100 mM NaAc, 5 mM $MgCl_2$, 1 mM EDTA, 10% glycerol, 0.02% Triton X-100). Then the beads were incubated with 4 μl of radiolabeled translation mixture of target protein for 1 h at 25°C in 200 μl of binding buffer (HEPES buffer with 1 μg/ml pepstatin, 50 μg/ml PMSF, 1 μg/ml leupeptin and 2 μg/ml aprotinin) plus 8% bovine serum. After washing six times with 800 μl of binding buffer (with salt raised to 300 or 450 mM NaCl) at room temperature, the bound proteins were eluted by boiling in 50 μl 4x SDS-PAGE loading buffer. The eluates were electrophoresed in 10 or 12% SDS-PAGE gel. Pulled down proteins were detected by autoradiography after exposure to X-ray film for 1-2 days.

Yeast two-hybrid assay

The yeast two-hybrid system MACTCHMAKER Gal4 system 3 (Clontech) was used for two hybrid assays. Yeast AH109 cells were transformed with respective bait and prey plasmids. The transformed yeast cell were grown on selective solid medium SD (Ade- His- Leu- Trp-) to find protein-protein interaction, as well as on non-selective solid medium SD (Leu- Trp-) to check the expressed foreign proteins were not toxic to yeast cells.

Gel mobility shift assay

The DNA probe was a double 17-mer oligonucleotide derived from the promoter sequence of gene COR15a. The wild type sequence was ATTTCATGG CCGACCTGCTTTTT. To make wild type probe, synthetic oligonucleotides (ST567 and ST568) were annealed, and were end labeled with γ-[^{32}P]-ATP and T4 nucleotide kinase (Gibico 18004-010). The mutant sequence was ATTTCATGGtatgtCTGCTTTTT. The corresponding synthetic oligonucleotide primers ST567 and ST568 were also annealed for competition assay.

The DNA-CBF1 complex formation was carried out in 20 µl buffer (20 mM HEPES, 50 mM KCl, 2.5 mM MgCl2, 1 mM EDTA, 10% glycerol, 1mM DTT) containing 4 µg BSA, 1 µg poly (dI-dC), 1 µl probe and about 2 pM CBF1. The reactions were electrophoresed in native polyacrylamide gel (40:1, in 0.5 x TBE buffer). The gel was pre-run at 150V for 30 minutes at cold room before adding the samples. The samples were run at 150V for 4 hours at cold temperature. The TBE buffer was changed to new buffer after sample running for 2 hours. After gel running, the signal was visualized by auto-radiography.

For ADA2b and GCN5 interacting with CBF1-DNA complex, 1, 2, 4 μl of recombinant ADA2b (5 μl contained about 0.5 μg) or 0.5 μl GCN5 (about 0.5 μg) was added to the 20 μl reaction at the beginning of the assembly. 0.15 μCi [H3] labeled acetyl-CoA was also included in a duplicated assay with both GCN5 and ADA2b in the beginning of reaction, and this assembly was allowed to react for 40 min at 27°C before adding probe. BSA 2μg was also included in one control assay.

Results

Test the physical interaction between CBF1 and ADA2a/b or GCN5

In order to test the physical interaction between CBF1 and ADA2 and/or GCN5, GST pull-down assay was performed. Purified recombinant fusion protein GST-CBF1 was attached to glutathione Sepharose beads and was used to pull down *Arabidopsis* ADA2a, ADA2b and GCN5 proteins. ADA2a, ADA2b and GCN5 proteins were expressed and radiolabeled by [^{35}S] methionine using *in vitro* transcription/translation coupled TNT T7 system (Promega). The bound proteins were detected by autoradiography after electrophoresis in SDS–PAGE gel.

As shown in Figure 1, recombinant GST-CBF1 pulled proteins ADA2a, ADA2b and GCN5, but not control protein luciferase. The same amount of protein of control GST pulled down much less of the target proteins. A later experiment was repeated where more stringent washing condition was applied: the salt NaCl was raised to 450 mM from original 150 mM, 0.2% Triton X-100 was replaced with 1% NP-40 and the result was the same. The results suggest that CBF1 interacts with ADA2a, ADA2b and GCN5.

Test the transcriptional activation domain of CBF1.

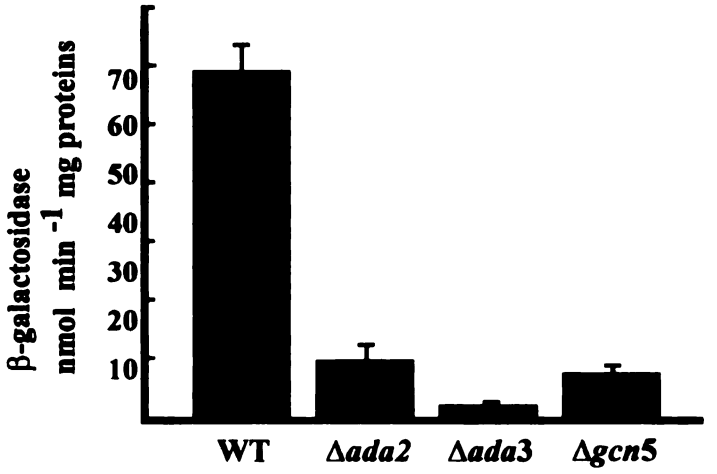


Figure 1. The transcriptional activation of protein CBF1 is dependent on yeast protein ADA2, ADA3, and GCN5.
β-galactosidase assays of yeast cell lysates were performed to detect CBF1 transcriptional activity. Protein CBF1 was expressed in yeast strains (wild type WT, Δ*ada*2, Δ*ada*3, and Δ*gcn*5) under the control of ADC1 promoter. The yeast strains contained an integrated reporter gene *LacZ* which had two direct repeat copies of the CRT/DRE sequence at its promoter to which CBF1 bound. This experiment was performed by Dr. Eric Stockinger at Dr. Michael Thomashow laboratory of Plant Research Laboratory [Stockinger et al, Transcriptional adaptor and histone acetyltransferase proteins in Arabidopsis and their interactions with CBF1, a transcriptional activator involved in cold-regulated gene expression. Nucleic Acids Res. 2001 29:1524-33.]

After finding that CBF1 interacting with ADA2a, ADA2b and GCN5, investigation was carried out further to find which part(s) of CBF1 was responsible for interaction with ADA2 and GCN5 protein. Protein CBF1 (full length of 213 amino acids), according its sequence homology, is composed of two domains: the activation domain (amino acids 115-213, the C-terminus) and the DNA binding domain (amino acids 1-115, the N-terminus). The N-terminus contains a well conserved DNA domain called AP2 (AP for apetala) (Allen, et al., 1998).

To confirm this domain boundary prediction, the N-terminus and the C-terminus of CBF1 were fused separately to the Gal4 DNA binding domain and the fusion proteins were tested for their ability to activate the reporter genes in yeast. The results indicate that the C-terminus activated the transcription of reporter genes, whereas the N-terminus of CBF1 (aa 1-115) did not (Figure 3). This confirmed that the C-terminus (115-213) is a transcriptional activation domain. Gel mobility shift assays (see below) also confirm that the N-terminus has the capacity of binding to specific DNA motif.

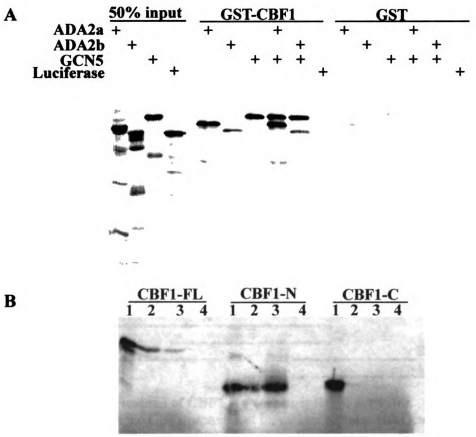
Test the interacting region of CBF1 with ADA2a/b and GCN5

After roughly defining the boundaries of the N- and C- terminus of CBF1 protein, study was carried out to determine which parts of CBF1(N or/and C) could interact with GCN5 and ADA2a/b using the method of GST pull-down assays. Recombinant protein GST-ADA2b and GST-GCN5 were used to pull down the full length, the N-terminus and the C-terminus of CBF1 which were translated and radiolabeled *in vitro* using the TNT T7 system.

As shown in Figure 2B, GST-ADA2b and GST-GCN5 pulled down the full length of

A

B



1: 20% input, 2: GST-GCN5 3: GST-ADA2b 4: GST

Figure 2. Interaction of CBF1 with ADA2a, ADA2b and GCN5. GST pull-down assays were performed to detect the interactions.**(A)** The *Arabidopsis* ADA2a, ADA2b and GCN5 proteins as well as luciferase (as a negative control) were expressed and [^{35}S] methionine radiolabeled using the *in vitro* transcription/translation coupled TNT T7 system (Promega). Equivalent amounts of GST or the GST–CBF1 fusion protein bound to glutathione Sepharose beads were incubated with ADA2a, ADA2b and GCN5 solutions. After washing, the bound proteins were eluted with SDS-PAGE loading buffer by boiling. The eluates were electrophoresed in SDS-PAGE gel, and the radiolabeled proteins were detected by autoradiography. **(B)** Similar GST pull-down assays to find out the N -terminus of CBF1 interact with ADA2b and GCN5. The major difference from (A) was that ADA2b and GCN5 were in the form of GST fusion protein while CBF1 proteins were expressed and radiolabeled in the TNT T7 system. CBF1-N: amino acids 1-115, CBF1-FL: amino acids 1-213, CBF1-C: amino acids 115-213.

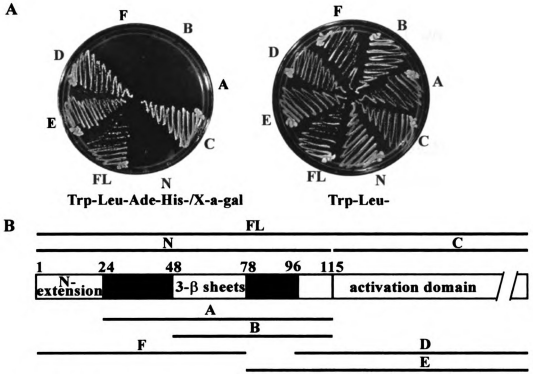


Figure 3. The C-terminus of CBF1 is a transcriptional activator.
(**A**) Diagram shows the CBF1 protein and its various regions. (**B**) Determine the transcriptional activation domain of CBF1 in yeast cells. cDNAs encoding different regions of CBF1 (as shown in panel **A**) were subcloned into vector pGBKT7 in frame with that of Gal4 DNA binding domain. CBF1 plasmids were co-transformed into yeast AH109 cells with plasmid pGADT7, and the transformed cells were grown on selective solid medium SD (Ade- His- Leu- Trp- /X-α-gal) to test transcriptional activation activity for various regions of CBF1, and on non-selective solid medium SD Leu- Trp- to check the expressed CBF1 proteins were not toxic to yeast cells. The co-transformation of plasmid of pGADT7 (with Leu+ marker) enable the yeast cells to grow on Leu- solid medium.

CBF1, confirming the earlier result that the full length CBF1 interacted with ADA2b and GCN5. Surprisingly, in contrary to the general model that co-activators usually contact the activation domain, the N-terminus of CBF1 (1-115) but not the C-terminus interacted with ADA2b and GCN5 (Figure 2B).

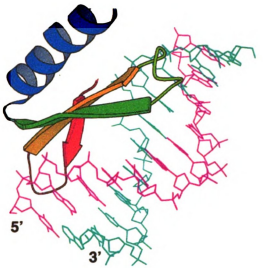
Determining the interacting motif in the N-terminus of CBF1 with ADA2

CBF1 is a member of the AP2 domain-containing transcriptional factors in *Arabidopsis*, such as ERF1, ERF4, ERF5, DREB2a, and TINY. The AP2 (apetala) domain is a DNA binding domain found in many transcriptional factors and is unique to plants (Riechmann, et al., 2000). The NMR structure of ERF1 AP2 domain reveals it is composed of an α-helix and a three β-sheets, where the three β-sheet bundle contacts the DNA GCC-box motif (Allen, et al., 1998) (and see Figure 4A). Based on CBF1's sequence homology to the DNA binding domain of ERF1 and CBF1's binding to similar GCC-box DNA motif (Stockinger, et al., 1997), the CBF1 N-terminus was divided into distinct regions as shown in Figure 4B.

GST pull-down assays were performed to test which part of the N-terminus of CBF1 was necessary for interaction with GCN5 and ADA2b. As showed in Figure 5, recombinant fusion proteins GST-GCN5 and GST-ADA2b pulled down the three overlapping regions of CBF1 N-terminus, amino acids 24-115, 48-115 and 1-78. Control GST alone did not pull down any of these three regions. This result suggested that the minimal interaction region of CBF1 was in the region of amino acids 48-78, the 3-β sheets.

The above date indicted that the putative three β-sheet region of CBF1 was sufficient

A



B

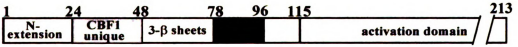


Figure 4. Structure of the AP2 DNA-binding domain

(A) Shown is the NMR structure of GCC-box binding domain of *Arabidopsis* protein ERF1 (for ethylene response factor). The domain has about 60 amino acids, its three-β-sheet bundle contacts the DNA motif (AGCCGCC) underneath. The picture is modified from M. D. Allen et al EMBO J. 17, 5484-9 1998.

(B) The defined regions of protein CBF1. The region (amino acids 48-96) in CBF1 is highly homologous to the GCC-box binding domain of *Arabidopsis* protein ERF1, and CBF1 protein has been shown to bind to similar GCC-box (GGCCGACC).

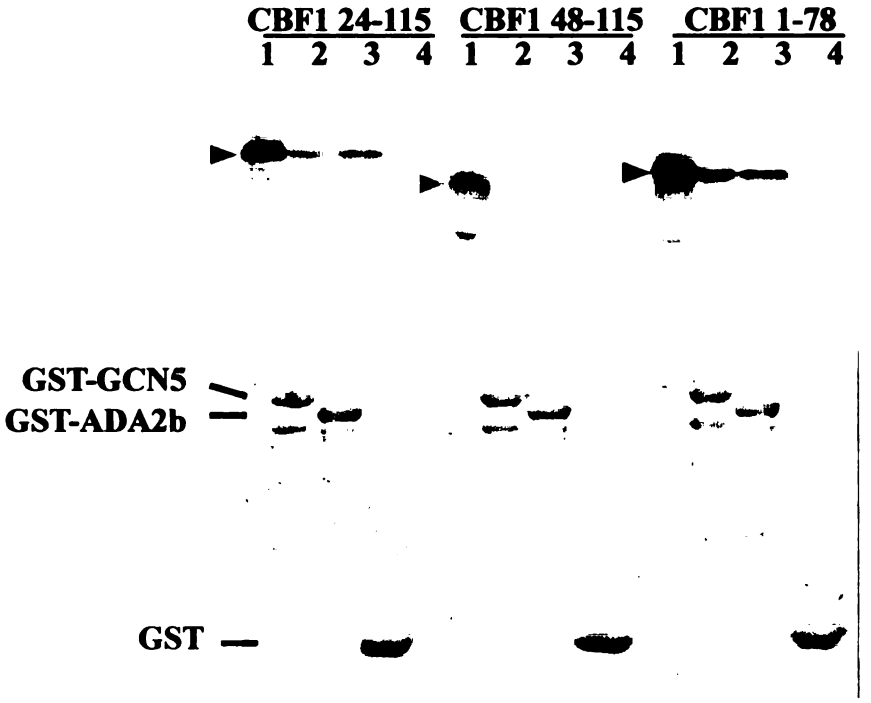


Figure 5. The N-terminus of CBF1 interacts with ADA2b and GCN5. GST pull-down assays were performed to detect the interactions. Recombinant GST–ADA2 and GST-GCN5 fusion proteins and GST alone were bound to glutathione Sepharose beads. The three N-terminal regions of CBF1 aa 24-115, 48-115 and 1-78 were expressed and [S^{35}] methionine radiolabeled using the *in vitro* transcription/translation coupled TNT T7 system (Promega). GST, GST-ADA2b or GST-GCN5 fusion protein bound to beads was incubated with the solutions of three CBF1 proteins. After washing, the bound proteins were eluted with SDS-PAGE loading buffer. The eluates were electrophoresed in SDS-PAGE gel, and the pulled down proteins were detected by autoradiography, as shown in the upper panel. The protein amounts of GST-ADA2b, GST-GCN5 and GST bound to beads tested were shown in the lower panel that were stained with Coomassie Blue. 1: 25% Input, 2: GST-ADA2b, 3: GST-GCN5, 4: GST.

for in

furthe

into p

variou

and 7

plasm

transf

X-α-

conta

intera

down

teste

ADA

neces

suffi

inter

was

GCN

teste

and p

for interacting with either GCN5 or ADA2b. A yeast two-hybrid assay was carried out to further confirm the result. cDNA fragments encoding various CBF1 regions were inserted into plasmid pGADT7 (encoding Gal4 activation domain) to make prey plasmids. These various domains and regions included CBF1 amino acids 1-78, 24-115, 48-115, 96-213 and 78-213. Yeast strain AH109 cells were co-transformed with respective CBF1 prey plasmids together bait plasmid encoding the full-length GCN5, ADA2a or ADA2b. The transformed yeast cells were grown on selective solid medium SD (Trp- Leu- His- Ade- /X-α-gal) for 4 days at 30°C.

As shown in Figure 6B, ADA2b showed interaction with the CBF1 regions that contained the putative three β-sheet region. The C-terminus of activation domain did not interact with ADA2b. This result correlated with earlier result obtained from GST pull-down assays. The yeast two-hybrid assay result also reveals that ADA2a, which was not tested in the previous GST pull-down assays, had similar interaction properties as that of ADA2b.

The results in Figure 6B showed that three-β-sheet bundle of CBF1 (aa 48-78) was necessary for interaction with ADA2a and ADA2b. To find whether this bundle was sufficient for interaction with ADA2a and ADA2b, the bundle alone was tested for its interaction in yeast two-hybrid assay. As shown in Figure 6C, the three β-sheet region was sufficient for interaction with ADA2a and ADA2b.

Surprisingly, the yeast two-hybrid assays did not indicate any interaction between GCN5 and CBF1, regardless which part of CBF1 (N-terminus and C-terminus) was tested (Figure 6D). This failure of interaction was not due to the lack of expression of bait and prey proteins in transformed yeast cells, for GCN5 bait protein is known to be

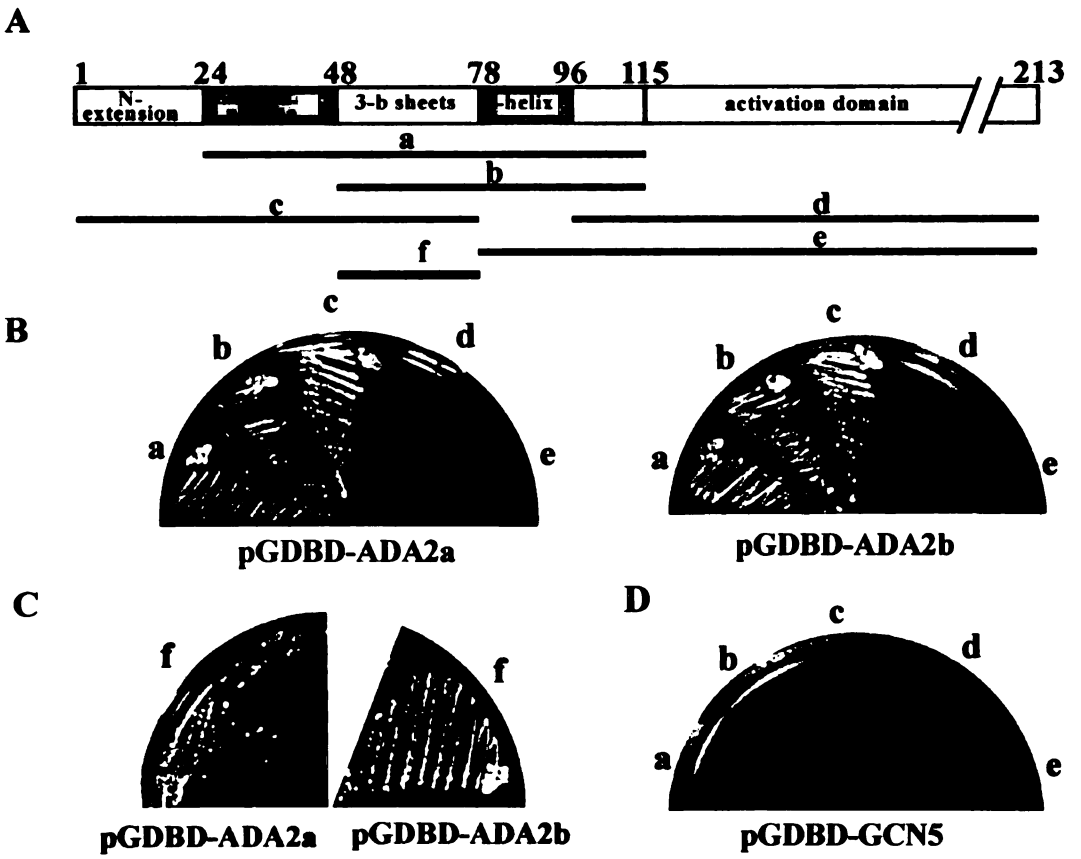


Figure 6. Defining the minimal interaction domain of CBF1 with ADA2a and ADA2b.

Yeast two-hybrid assays were performed to detect the interactions. cDNA encoding the various CBF1 regions (as indicated in panel A) were subcloned into vector pGADT7 to create prey plasmids that expressed the respective CBF1 protein regions in fusion with Gal4 activation domain.Yeast strain AH109 cells were co-transformed with CBF1 prey plasmids together with bait plasmids of GCN5, ADA2a or ADA2b. The transformed cells were grown on selective solid medium (SD Trp- Leu- His- Ade-/X-α-gal) for 4 days at 30°C. (**A**) Diagrm shows the CBF1 protein in full length and the various regions (a, b, c, d, e) tested for interaction with ADA2 and GCN5 proteins. c: amino acid 1-78, a: 24-115, b: 48-115, d: 96-213, e: 78-213, and f: 48-78.

(B, C, D) Shown is the growth of transformed yeast colonies on the selective medium.

expressed and free to interact with other proteins (see chapter III) and all the CBF1 regions and domains were known to express in the yeast cells (Figure 6B and Figure 3B).

The AP2 domain of TINY interacts with ADA2a and ADA2b.

There are about 140 AP2 domain-containing proteins in *Arabidopsis* that are implicated in variety of biological activities (Riechmann, et al., 2000). The interaction of the AP2 domain of BCF1 with ADA2a and AD2b raises the question whether the interaction is specific to CBF1 or is common to all AP2 domain-containing proteins.

To address whether the interaction of CBF1 domain with ADA2 is specific, four transcription factor proteins from *Arabidopsis*, ERF4, ERF5, TINY and DREB2a were tested for their interaction with ADA2, ADA2b and GCN5 by yeast two-hybrid assay. The ERF4/5 proteins are ethylene response factors, DREAb2a is implicated in *Arabidopsis* response to drought (Liu, et al., 1998, Sakuma, et al., 2002), and over expression of TINY leads to tiny plants (Wilson, et al., 1996). These four proteins all contained very similar AP2 domains as that of CBF1, as shown in Figure 7A. DNAs encoding AP2 domains of TINY (amino acids 1-95), ERF4 (1-85), ERF5 (1-219) and DREB2a (1-139) were amplified from *Arabidopsis* genomic DNA. Since these genomic DNA fragments did not have any introns, they were directly subcloned into vector pGADT7 to construct prey plasmids. Yeast two-hybrid assays were performed as before. Interestingly, only TINY (1-95) showed interaction with ADA2a and ADA2b (Figure 7B), demonstrating that ADA2 proteins interact with only specific AP2 domain. As observed for CBF1, none of them including TINY showed any interaction with GCN5 (Figure 7B).

Figure 7. Differential interaction of AP2 domain with ADA2 proteins.
(A) Amino acids sequence alignment shows the high homology among the AP2 domains of *Arabidopsis* transcriptional factors TINY, ERF4, ERF5, DREB2a and CBF1. The dash lines indicate the identical residues and the dots indicate where residues are absent for the respective AP2 domain. **(B)** Yeast two hybrid assays were performed to detect protein-protein interactions. DNA fragments encoding respective AP2-domain containing regions of TINY (amino acids 1-95), ERF4 (1-85), ERF5 (1-219) and DREB2a (1-139) were inserted in to prey vector pGADT7 in frame with that Gal4 activation domain. Yeast cells were co-transformed with the corresponding prey plasmids and the bait plasmids of ADA 2a (pGDBD-ADA2a), ADA2b (pGDBD-ADA2b), or GCN5 (pGDBD-GCN5), and were grown on selective medium for 4 days at 30°C.
1: pGAD-ERF4 (1-85),
2: pGAD-ERF5 (1-219),
3: pGAD-DREB2a(1-139),
4: pGAD-TINY(1-95),
5: pGAD-GCN5 (HAT, 203-384),
6: pGADT7, 7: pGAD-ADA2b.

A

```
CBF1(46-106)    PIYRGVRQRNSGKWVSEVREPNKK.TRIWLGTFQTAEMAARAHDVAALALRGRSAC-NFADS
TINY(34-94)     PV-----K--W------I---R--.S-------PSP-----------SIK-A--I---PDL
ERF4(23-83)     IR-----K-PW-RYAA-I-D-G--.--V-----D---E----Y-T--RDF--AK-KT--PTF
DREB2a(77-137)CSF------IW----A-I----R.GS-L-----P--QE--S-Y-E--K-MY-PL-R---PRS
ERF5(154-215)  KH-------PW--FAA-I-D---RGS-V-----D--IE----Y-E--FR---SK-I---PLE
ERF1(144-205)  KH-------PW--FAA-I-D-A-NGA-V-----E---D--L-Y-R--FRM--SR-L---PLR
```

B

pGDBD-ADA2a pGDBD-ADA2b pGDBD-GCN5

Detect ADA2b-CBF1-DNA tertiary complex

The observation of physical interaction between ADA2a/b and CBf1 suggests a model that when CBF1 activates the transcription of COR genes, it recruits the ADA2-GCN5 chromatin modifier to facilitate transcription.

The data described above show that the putative three β-sheet bundle of CBF1 interacts with proteins ADA2a and ADA2b. On other hand, this three β-sheet bundle, based on the NMR structure of the homologous AP2 domain of EFR1 with bound DNA GCC-box, likely binds directly to DNA motifs of CRE element from CBF1 responsive genes such as COR genes. It is not known whether CBF1 could interact with its cognate DNA motif and ADA2 proteins simultaneously, or if there exists competitive binding to CBF1 between ADA2 proteins and DNA motif.

To address this question, gel mobility shift assays were carried out to test whether adding of ADA2b protein results in either supershift or disappearing of CBF1-DNA complex. In the gel mobility shift assays, the DNA probe was 17-mer oligonucleotides which contained the CBF1-binding motif from the COR15a gene. A mutant probe was also introduced as the control. The probes were radiolabeled by [^{32}P]. Recombinant His6-ADA2b and His6-CBF1(aa 1-213) proteins were purified as described in Chapter IV, using HEPES buffer (40 mM HEPES, 50 mM KCl, 10% glycerol, 7mM 2-mercaptoethanol). The proteins of His6-CBF1 and His6-ADA2b were eluted from Ni-NTA beads by 220 mM imidazole in HEPES buffer. The proteins solution was then dialyzed in HEPES buffer to remove the imidazole. The purified His6-ADA2b and His6-CBF1 proteins are shown in Figure 8A. There was considerable contamination in the purified His6-ADA2b. The presence of His-ADA2b in the purified preparation was

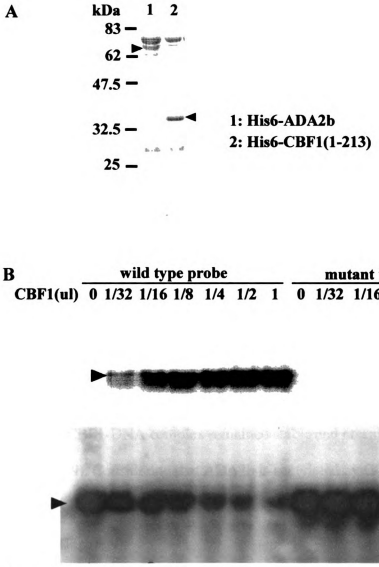


Figure 8. Testing the CBF1 protein binding to DNA probe.
(**A**) The purified recombinant proteins CBF1 and ADA2b as His6-tagged fusions. The protein preparations were electrophoresed in 10% SDS-PAGE gel and visualized by Coomassie Blue staining. The protein position of CBF1 and ADA2b is indicated by the filled triangle. These two recombinant proteins were used in gel shift mobility assays in panel B and Figure 9. (**B**) Gel mobility shift assay to detect the binding of CBF1 with DNA probe. The [32P] radiolabeled DNA probe (wild type or mutant) was incubate with recombinant His-CBF1 protein, and then electrophoresed in native polyacrylamide gel. The positions of the DNA probe in gel were detected by autoradiography. The free probe and CBF1-probe are indicated by the arrows.

confir

Chapt

T

probe

forme

bindi

mobil

ADA

S

protei

recon

IV, w

show

detec

acety

same

Disc

T

intera

or GS

is nec

sugge

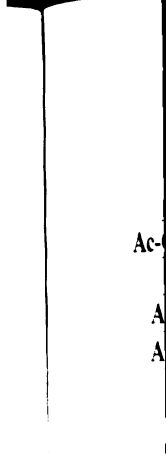
confirmed by Western blotting and by its ability to enhance GCN5 HAT activity (sees Chapter IV).

The gel mobility shift assays were carried out in HEPES buffer with the labeled probes and recombinant His6-CBF1 protein. As shown in Figure 8B, the CBF1 proteins formed complexes with wild type probe but not the mutant one, illustrating the specific binding of CBF1 with the wild type probe. Adding of the ADA2b protein alters the mobility of the radiolabeled DNA probe, as shown in Figure 9. However, the adding of ADA2 protein also did not disrupt the CBF1-DNA complex (Figure 9).

Since the GCN5 protein interacts with the ADA2b protein, the presence of the GCN5 protein might change ADA2b properties in binding to other partners. Therefore, the recombinant GCN5 protein, which was purified in HEPES buffer as described in Chapter IV, was also added together with ADA2b in the gel mobility shift assays. The result showed that the CBF1-DNA complex remained stable and no supershift complexes were detected (Figure 9). Considering that ADA2b is acetylated by GCN5 (see Chapter VII), acetyl-CoA was also included in another two parallel reactions. The result was again the same as above (Figure 9).

Discussion

The experiments described in this chapter show that *Arabidopsis* protein CBF1 can interact with co-activator proteins ADA2a and ADA2b using the yeast two-hybrid assay or GST pull-down assay. The putative three β-sheet bundle in the AP2 domain of CBF1 is necessary and sufficient for interacting with ADA2a and ADA2b. This interaction suggests a model. When plant *Arabidopsis* is under cold stress, activator CBF1 activates



	included in the GMSA reaction in µl														
Ac-CoA (60 nM/µl)														2	2
GCN5										1/2	1/2	1/2	1/2	1/2	1/2
ADA2b (batch 1)						1	2	4			2	4			4
ADA2b (batch 2)			1	2	5				2	5			5		
CBF1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1

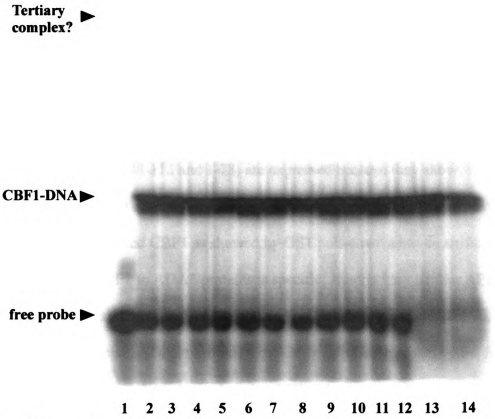


Figure 9. Detect ADA2b-CBF1-DNA tertiary complex.
Gel mobility shift assays were performed to detect the potential tertiary complex. [32P] radiolabeled DNA probe (wild type) was incubate with the CBF1 protein in combination with the ADA2b protein (lane 3-8) or the ADA2b protein plus the GCN5 protein (lane 9-12). The last two lanes also contained Acetyl-CoA in the reaction in order to make the ADA2b protein acetylated. The incubated mixtures were then electrophoresed in native polyacrylamide gel. Two different purified batches of the ADA2b protein (batch 1 and batch 2) were tested. Both the recombinant proteins CBF1 (1-213) and ADA2bwere purified as His-tgged fusions, while recombinant GCN5 was purified as GST fusion and the GST tag was removed subsequently. The free probe in the last two lanes seemed to diffuse for unknown reason.

the transcription of cold response genes in recruiting GCN-ADA2 complex that acetylates histones at the local promoter for facilitating transcriptional activation. In support of this model, *Arabidopsis* plants with crippled ADA2b or GCN5 protein show delayed transcription of COR genes in response to cold temperature treatment (Vlachonasios, et al., 2003).

GST pull-down assays showed that GCN5 can also contact the CBF1 N-terminus. But yeast two hybrid does not show this interaction. This lack of interaction of GCN5 with any region and domains of CBF1 in yeast two-hybrid assay was not due to the failed expression of bait or prey proteins in yeast cells. Bait protein GCN5 in fusion of Gal4 DNA binding domain is expressed in yeast (see Figure 3 at chapter III). Prey protein CBF1 (amino acids 24-115, 48-115 and 1-78) are expressed because they show interaction with ADA2a and ADA2b in the parallel assays. We conclude that the interaction between GCN5 and CBF1 as showed by GST pull-down assay is artificial.

The interaction of CBF1 to ADA2 proteins is specific. Among very similar AP2 domains from *Arabidopsis* transcriptional factors TINY, ERF4, ERF5 and DREB2a, only TINY shows interaction with ADA2a and ADA2b. This might be explained by the fact that putative three β-sheets and a α-helix domain of TINY have the highest similarity in amino acids sequence to that of CBF1 (see Figure 7). A less likely explanation is that the yeast cells failed to express these AP2 domain-containing factors.

Regardless, the interaction of ADA2 adaptor protein with AP2 DNA binding domain suggests that ADA2-GCN5 complex might play multiple roles in plants. The AP2 domain is unique to plant world, and is found in a wide range of proteins in *Arabidopsis*, including important regulators of several developmental processes involved in floral

organ identity determination or control of leaf epidermal cell identity, and many key responding factors to various types of biotic and environmental stress such as cold stress, dehydration and pathogen attack (Chang and Shockey, 1999, Memelink, et al., 2001, Riechmann, et al., 2000, Shinozaki and Yamaguchi-Shinozaki, 2000). This possible broad range of action is also supported by the finding that disruption of the *ADA2b* or *GCN5* gene *Arabidopsis* resulted in severe phenotypes such as dwarf and sterility (Vlachonasios, et al., 2003). Interestingly, a recent report shows that the ADA2b protein in *Arabidopsis* is involved in cell differentiation in response to auxin and cytokinin concentration variation (Sieberer, et al., 2003), and a AP2 domain containing factor is involved in shoot differentiation in combination with cytokinin action (Banno, et al., 2001).

The failure of detecting an ADA2b-CBF1-DNA complex by gel mobility shift assays could be a technical problem such as the buffer conditions are not appropriate, and does nor necessary mean that *in vivo* CBF1 can not recruit the ADA2-GCN5 complex.

The AP2 domain of CBF1, which binds DNA motif of COR genes' promoter directly, contacts the co-activator ADA2 protein. This observation is contrary to the standard model in which the activation domains of transcriptional activators contact and recruit transcriptional co-activators. But increasing evidence supports that the DNA binding domain recruitment of co-activator is also widely employed. One example is EKLF, a transcriptional activator in erythroid cell differentiation, in which the DNA binding domain of EKLF contacts SWI/SNF complex directly and the recruited Swi/Snf complex sufficiently remodels local chromatin (Kadam, et al., 2000). Interestingly, for transcription to start, the activation domain of EKLF is additionally required. Another case is the DNA binding domain of nuclear receptors RXR/RAR heterodimer which

contacts histone modifier PCAF directly (Blanco, et al., 1998).

On the other hand, the contact of the ADA2 proteins by transcriptional activator in *Arabidopsis* also suggests that similar complex in other species such as yeast SAGA might also be recruited by activators through interaction of the ADA2 protein in addition to the known handles of GCN5 and Tra1 proteins (Brown, et al., 2001, Kulesza, et al., 2002).

Chapter Seven

Acetylation of ADA2a and ADA2b by GCN5

Part I Acetylation of ADA2a/b by GCN5 *in vitro*

Part II Test the significance of ADA2b acetyaltion *in vivo*

Introduction

Although discovered initially as histone acetyltransferses, the mammalian CBP/p300 and PCAF proteins can also acetylate other proteins including general transcriptional factors and regulatory factors such as TFIIE, TFIIF, p53, EKLF, GATA-1, HMG I (Y), HMG17, ACTR, Tat, MyoD and E2F1(Spilianakis, et al., 2000). The biological significance of acetylation for some non-histone proteins has been reported, including the regulation of transcriptional factor activity and the regulation of protein cellular location (Freiman and Tjian, 2003, Kouzarides, 2000, Sterner and Berger, 2000).

For instance, acetylation affects protein cellular localization. Protein POP in *Caenorhabditis elegans* is a homologue of LEF/TCF, a nuclear effector of the Wnt signaling pathway that controls cell fate decisions and embryonic patterning in animal development. Recently, the Shi group found that POP from *Caenorhabditis elegans* embryos is acetylated by CBP/P300. Acetylated POP enhances its nuclear retention through blocking nuclear export pathway (Gay, et al., 2003). CIITA is a tansactivator of MHC class II genes. CIITA associates with CBP/PCAF and is acetylated by CBP/PCAF. Interestingly, the acetylated lysines are within a nuclear localization signal (NLS). The acetylated NLS enhances CIITA nuclear import (Spilianakis, et al., 2000).

Another case of protein acetylation affecting cellular localization is adenovirus 12 S

E1A (Madison, et al., 2002). E1A is acetylated at lysine 239 by CBP in transfected cells. E1A with acetylation at lysine 239 or substitution mutation at lysine 239 resides predominantly in cytoplasm. Unacetylated E1A has much higher affinity for importin-α3 than acetylated E1A. This might be the first case indicating that acetylation of protein results in disruption of protein-protein interaction.

Acetylation of tumor suppressor protein p53 is another case in which acetylation regulates protein function. Human p53 is acetylated at lysines 373 and 382 within its C-terminal domain known to regulate p53 DNA binding activity. Acetylated p53 enhances its DNA binding activity *in vitro* (Gu and Roeder, 1997). *In vivo*, acetylation of p53 promotes recruitment of co-activators such as CBP and TRRAP and therefore potentiates p53 activity as a transcriptional activator (Barlev, et al., 2001).

A similar case is HIV Tat acetylation. Acetylation of Tat at lysine 50 by p300/PCAF results in its dissociation from TAR RNA and promotes formation of a multiprotein complex comprised of Tat, p300/CBP, and PCAF (Deng, et al., 2000). This event enables the stalled transcriptional machinery complex to elongate efficiently on the HIV DNA template (Garber and Jones, 1999).

Given that many transcriptional activators are acetylated, we hypothesized that the *Arabidopsis* activator CBF1 may be acetylated by GCN5. *In vitro* acetylation assays found no support for this hypothesis. Serendipitously, ADA2b, which was included originally to facilitate acetylation, was itself acetylated by GCN5. This unexpected finding resulted in detailed study of the *Arabidopsis* ADA2 acetylation *in vitro* and its significance *in vivo* in this chapter.

Experimental Methods

Purification of recombinant GCN5, ADA2b and CBF1 proteins

Recombinant full length an truncated GCN5, ADA2b and ADA2a were prepared as described in Chapter IV. CBF1 was also expressed as a GST fusion protein from a plasmid named pGEX-CBF1(fl). The GST-CBF1 fusion protein was affinity purified by glutathione Sepharose beads and then the fusion protein was digested with protease thrombin for 5 hours in the cold room in PBS buffer, and the thrombin was subsequently absorbed away by incubating the digestion supernatant with Streptavidin agarose (Novagen) at room temperature for 30 min.

Purification of recombinant yADA2N

The cDNA encoding the yeast ADA2 N-terminus (amino acids 1-176) was subcloned into pGEX vector to express GST fusion protein GST-yADA2(1-176). As control, GST tagged ADA2b N-terminus GST-ADA2bN(1-225) as well as GST proteins were purified. The expression and purification procedure for GST, GST-ADA2b (1-225) and GST-yADA2 (1-1760 was the same as that of purification of ADA2bN (see chapter IV) except that the GST fusion proteins bound to Sepharose beads were eluted with GSH Tris buffer (20 mM glutathione, 50 mM Tris, pH 8.0, 50 mM NaCl, 1 mM DTT).

Purification of ADA2b (1-179, 28-225, 28-179, 1-208)

cDNA encoding the deleted versions of ADA2b N-terminus (amino acids 1-179, 28-225, 28-179, 1-208) were amplified by PCR and subcloned into pGEX-6P vector in frame with that of GST. The expression condition and purification procedure was the

same as described in Chapter IV except that induction condition was 80 μM IPTG for 17 hours. The GST tag was removed by protease PreScission (digestion for 3 and half hours).

Acetylation assay for ADA2b

The acetylation was carried out in Tris buffer (50 mM Tris, pH 7.8 or 7.4 or 7.0, 40 mM NaCl 10 % glycerol, 0.1 mM EDTA, and 1 mM DTT) or HEPES buffer (40 mM HEPES, 50 mM KCl, pH 7.3, 10% glycerol and 1 mM DTT). About 0.3-0.5 μg GCN5 was used in the assay together with appropriate amount of ADA2b proteins as needed. The mixture was allowed to react for 30 min at 27°C or otherwise specified. After the reaction was finished, the mixture was analyzed in SDS-PAGE gel. The proteins in gel were first stained with Coomassie Blue and the acetylation signals were then detected by enhanced fluorography. For lower pH assay, the same assay was repeated except that the Tris pH was adjusted to 7.0, or 7.4.

Acetylation assay for ADA2bN, ADA2bM, ADA2bC and ADA2bN deletions

About 0.5 μg GCN5, 0.3 μg ADA2b, 2 μg ADA2bN, 0.5 μg ADA2bM, 2 μg ADA2bC, and 0.15 μCi H3-Ac-CoA were added as needed in 30 μl Tris buffer (50 mM Tris, pH7.4, 40 mM NaCl 10 % glycerol, 0.1 mM EDTA, and 1 mM DTT). The acetylation reaction proceeded for 40 min at 27°C. In some experiments, about 0.5 or 1μg ADA2bM (225-377) was also included improve acetylation efficiency. The reaction products were resolved in a 12% SDS-PAGE gel and the proteins were Coomassie stained. The protein gel was then soaked in autoradiographic enhancer (NEN Life Science Products), vacuum dried and exposed to X-ray film at -70°C for fluorographic

pictur

Muta

S

codo

GST

K25

(K.F

mix

The

min

proc

puri

DH

or A

enz

frag

pla

app

DN

picture.

Mutations of putative acetylation site in ADA2a and ADA2b

Site-directed mutagenesis was used to change lysine codons to arginine or alanine codons in ADA2a and ADA2b expressing plasmids. Plasmid pGEX-ADA2a encoding GST fused to ADA2a (full length) was mixed with primer pairs ST660-ST666 (K/A at K257), ST661-ST667 (K/R at K257), ST662-ST668 (K/A at K265), or ST663-ST669 (K/R at K265). Plasmid pGEX-ADA2b encoding GST-ADA2b (1-486, full length) was mixed with primer pairs ST664-ST670 (K/A at K215) or ST665-ST671 (K/R at K215). The PCR reaction was cycled for 16 times at settings of 95°C 1 min, 55°C 1 min, 72°C 6 min and 5 unit PFU (Turbo) polymerase enzyme in 50 μl reaction volume. Then the PCR product was cut with 20 units of *Dpn* I for 80 min to digest the templates. DNA was purified from the digestion by ethanol precipitation and used to transform *E. coli* strain DH5α.

Positive clones were identified by digestion with restriction enzyme *Hae* III (GG/CC) or *Hha* I (GC/GC), for the mutated plasmids contained new restriction sites for these enzymes.

To make mutant GST fusion expressing plasmid for ADA2b, the *Nco* I-*EcoR* I fragment of ADA2b (aa172-288, mutated site K215) was cut from the positive mutated plasmid and used to replace the same region in wild type plasmid pGEX-ADA2b. This approach avoided introducing unnecessary other mutated sites and reduced the work of DNA sequencing---only the *Nco* I-*EcoR* I region was sequence checked.

To make mutant GST fusion expressing plasmid for ADA2a, the *Hind* III-*Pst* I

fragment of ADA2a (aa210-313, mutated site K265 or K257) was cut from the positive mutated plasmid and used to replace the same region in wild type plasmid pET-ADA2a which expressed His6-tagged ADA2a. Then cDNA encoding the mutated ADA2a (1-273) was subcloned into GST vector to make GST fusion expressing plasmid. The DNA sequences were checked.

In summary, for ADA2aN(1-273), wild type codon AAG at K257 was changed to codon GCC for alanine or CGC for arginine, codon AAA at K265 was changed to codon GCC for alanine or CGC for arginine. For ADA2b (fl), wild type codon AAG at K215 was changed to codon GCC for alanine or CGC for arginine.

Acetylation assays for free and nucleosomal histones with mutant ADA2b

Recombinant GST-ADA2b (K215R or K215A) was purified essentially the same as that of GST-ADA2b (wild type) (see chapter IV). GST-ADA2b (wild type) was prepared in chapter IV. The free core histones' acetylation assays were also performed the same way as described in chapter IV.

For testing mutant ADA2b effect of GCN5 activity on nucleosomal substrate, the reactions were carried out in 27 µl HEPES buffer containing 6mM sodium butyrate, 8µl nucleosome preparation (prepare in chapter IV), 0.2 µC [^{3}H] Ac-CoA, about 8 pM GCN5 and about 4 pM or 8pM GST-ADA2b (wild type, K215R and K215A mutant). The reaction proceeded for 30 min at 25°C and was stopped by adding 16 µl of 6 x SDS-PAGE loading buffer and boiled for 5 min. Equivalent amount of GST was included as negative control. The reaction products were resolved in a 15% SDS-PAGE gel, and the acetylated histones were detected by fluorography for 4 months at -80°C.

ADA2bN acetylation for mass spectrometry

The acetylation reaction was carried out in 16 µl buffer (40 mM Tris, 50 mM KCl 10% glycerol, 7 mM 2-mercaptoethanol, pH 7.4) containing ~2µg ADA2bN (1-225), ~1.5 µg ADA2bM (225-377), ~0.5µg GCN5, and 60 nM Ac-CoA. The reaction was allowed to proceed for 30 min at 25°C. The acetylated protein was resolved in 12% SDS-PAGE gel. The protein band was cut out from the gel after being stained with Coomassie Blue and sent to the mass spectrometry facility at our department. There the acetylated protein was digested in gel by trypsin. The recovered trypsin-digested faction was subjected to mass spectrometry (LC/MS/MS). This was done by Dr. Brett Phinney.

Construction of plasmid pGEX6P-2bNR which expresses ADA2bNR

A cDNA fragment encoding ADA2b(1-172) cut from plasmid pGAD-ADA2b (1-208) by *EcoR* I and *Nco* I, and cDNA encoding ADA2b 173-221 (cut with *EcoR* I and *Xho* I from PCR product amplified from ADA2b full length cDNA by primer pairs ST486-ST707) were subcloned into GST vector pGEX-6P-1 by *EcoR* I and *Xho*I. The encoded GST fusion protein sequence is read as GST-ADA2b(1-221)-SGRIVTD, including the 7 amino acids encoded by the plasmid vector. The acetylation site of K215 reads as •••RSFVDRSFGGK^{215}KPVSTSSGRIVTD. The new ADA2b (1-221) plus tailed vector sequence was named ADA2bNR.

ADA2bNR acetylation for mass spectrometry

The protocol was basically the same as that of ADA2bN except the reactions were performed in HEPES buffer (40 mM HEPES, pH 7.4, 50 mM NaCl, 10 % glycerol and 7

mM 2-mercaptoethanol).

Constructing transformation plasmid for plant

Binary plasmid pKVA31was used to construct transformation plasmids with mutant cDNAs of ADA2b. Plasmid pKVA31, constructed by Dr. Kostas Vlachonasios, is a derivative of plasmid pCAMBIA 3301 (CAMBIA, Canberra, Australia) and contains the cDNA of wild type ADA2b under the control of the strong promoter 35S, and the bar1 (herbicide basta) selection marker. The cDNAs of ADA2b (K215R) and ADA2b (K215A) were subcloned into pKVA31 to the wild type ADA2b cDNA.

***Arabidopsis* plant transformation**

The constructed plasmids with mutated ADA2b cDNAs were introduced into bacterium *Agrobacterium tumefaciens* strain GV3101. To make transgenic plants, the transformed *Agrobacteria* were used to transform *Arabidopsis* plants which were in the early floral stage by the method floral dip (Clough SJ, Bent AF). The genotype of the transformed plants was heterozygous for T-DNA insertion at the *ADA2b* gene locus.

Four plants were transformed by each constructed plasmid containing mutated cDNA of ADA2b (K215R), ADA2b (K215A) or wild type ADA2b.

T1 seeds were collected from the transformed plants. The seeds collected from plants transformed by wild type ADA2b plasmid were named as A1, A2, A3, and A4, and by that of ADA2b (K215A) as B1, B2, B3 and B4, and by that of ADA2b (K215R) as C1, C2, C3 and C4.

Gro

The

20-3

6°C

med

whi

con

wat

Ger

Wi

gen

pro

tha

PC

AD

the

pla

ST

1.3

Growth of transgenic plants

The collected seeds were first sterilized in 30% bleach (Clorox) with 0.5 % Tween 20 for 20-30 min, and washed with water for 4 times. The seeds were then left in cold room (4-6°C) for a week to synchronize germination time. The seeds were then put on solid medium of Gamborg's B-5 with Basta (4 mg/L) to germinate. 7 to 10 days old seedlings which were Basta resistant were planted in soil for future screening under the growth condition of 24-hour light cycle. The plant in soil pots were watered every 8 days with water, and with nutrient-solution in the middle of every 8 days.

Genotyping the offspring of the transformed plants

DNA samples were prepared from the growing plant rosette leaves in soil using Wizard Genomic DNA Purification Kit (Promega). To identify the plants' genotype, the genomic DNA was PCR amplified with specific primers and afterwards the PCR products were analyzed by agarose gel electrophoresis. The primers were designed such that the genomic DNA fragment of the *ADA2b* gene across the T-DNA insertion site was PCR amplified. The primers were SJT780 and STQ29. The PCR product from the *ADA2b* gene with T-DNA insertion was about 8.2 kb long, whereas PCR product from the wild type *ADA2b* gene was about 2.7 kb long.

To check the transgenic integration of the ADA2b cDNA by transformation, the plant genomic DNA was amplified by PCR with another pair of primers, SJT780 and STQ29, where SJT780 was a 35S promoter specific primer. The PCR product was about 1.3 kb.

RNA analysis

Total RNA was prepared from the plant rosette leaves using the RNeasy Plant Mini kit (Qiagen). About 5 μg of total RNA was electrophoresed in 1% agarose gel and transferred to Hybond N+ membrane (Amersham Biosciences) by the standard method of capillary transfer (Frederick M Ausubel et al, Current Protocols in Molecular Biology, 4.9.5-4.9.6). The RNA on the membrane was cross-linked by UV light. The membrane was hybridized with [^{32}P] labeled cDNA probe of ADA2b (cDNA nucleotides 1-434). Hybridization was performed in PerfectHyb™ (SIGMA) at 68°C overnight with 100 μg/ml sheared herring testis single strand DNA as non-specific blocker. The hybridized membrane was washed subsequently in 2 x SSC, 0.1% SDS and 0.01% sodium pyrophosphate briefly at room temperature, and washed in 0.5 x SSC, 0.1% SDS and 0.01% sodium pyrophosphate at 65°C for 30 min. The signal of hybridization was visualized with phosphor imager.

Results

Part I Acetylation of ADA2 by GCN5 *in vitro*

ADA2b is acetylated by GCN5 *in vitro*

Triggered by the observation that many transcriptional activators can be acetylated by PCAF , a sibling of GCN5, we tested if the *Arabidopsis* transcriptional activator CBF1 was acetylated by GCN5. *In vitro* acetylation assays were performed to test this hypothesis.

Recombinant proteins CBF1, ADA2b and GCN5 of *Arabidopsis* were expressed as

GST fusion

GST tag wa

and GST-A

purified pro

The quality

of this Cha

The ac

proteins. A

that ADA2

bridges CB

assay was

acetylated

Surpr

Control bo

acetylatio

This

could be

al., 1996)

that the p

by GCN5

acetylatio

the reacti

As d

GST fusions in *E. coli*. After affinity purification on glutathione Sepharose beads, the GST tag was subsequently removed by proteases, where GST-CBF1 was cut by thrombin and GST-ADA2b and GST-GCN5 were cut by the Prescission protease. The three purified proteins were stored in PBS buffer (pH 7.4, with 10% glycerol and 1 mM DTT). The quality of the purified proteins is documented in Figure 9 of Chapter VI and Figure 1 of this Chapter.

The acetylation assays were performed in Tris buffer (pH 7.8) with the recombinant proteins. As shown in Figure 1A, CBF1 alone was not acetylated by GCN5. Considering that ADA2b enhances GCN5 HAT on histone substrate activity and that ADA2b protein bridges CBF1 and GCN5 interaction by its virtue of contacting both proteins, the HAT assay was also performed with the addition of ADA2b protein. But again CBF1 was not acetylated by GCN5 (Figure 1A).

Surprisingly, the added ADA2b was found serendipitously to be acetylated by GCN5. Control bovine serum albumin was not acetylated, indicating the specificity of GCN5 acetylation on the ADA2b protein.

This preliminary acetylating assay was carried out at pH 7.8 in Tris buffer. There could be GCN5-independent chemical acetylation activity at this high pH value (Kuo, et al., 1996). Therefore, the assay was repeated using the same reaction conditions except that the pH was lowered to7.4 or 7.0. The results showed that ADA2b was still acetylated by GCN5 at the lower pH (Figure 1C). In addition, GCN5 was found necessary for the acetylation. No acetylation occurred to ADA2b alone without the presence of GCN5 in the reaction.

As discussed in Chapter III, *Arabidopsis* GCN5 protein is composed of the unique

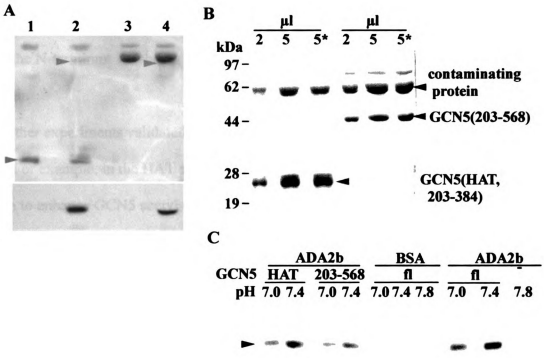


Figure 1. ADA2b is acetylated by GCN5.
(A) ADA2b but not CBF1 was acetylated by GCN5. The acetylation assays were performed with GCN5, [^{3}H] Ac-CoA and respective proteins CBF1, ADA2b and BSA at pH 7.8. The reaction products were analyzed in a 10% SDS-PAGE gel. The proteins in gel were stained by Coomassie Blue as shown at the upper where the ADA2b and CBF1 proteins are indicated by the red and green arrows respectively, and the acetylation signals were detected by enhanced fluorography as shown at the bottom. 1: CBF1, 2: CBF1+ADA2b, 3: BSA, 4: BSA+ADA2b. **(B)** Shown are purified recombinant GCN5 (203-384, HAT domain), or GCN5(203-568) which contains the HAT domain and the bromodomain. The proteins were stained with Coomassie Blue after SDS-PAGE. These GCN5 proteins were used as acetylation enzyme in panel C. The GCN5 proteins are indicated by the blue arrows *****: the sample was incubated at room temperature for one hour to check the protein stability. **(C)** Acetylation assays of ADA2b at lower pH and with truncated GCN5 versions. Similar acetylation assays as in panel A were performed at pH 7.0 and 7.4 with full-length (fl) and truncated versions of GCN5. The black arrow points to the acetylated ADA2b.

N-terminus, HAT domain and the bromodomain. To test if the acetylation of ADA2b required the intact GCN5 protein or if the HAT domain was sufficient, acetylation assays were repeated using purified recombinant GCN5 HAT domain or GCN5 (203-568) which lacks the N-terminus. As shown in Figure 1C, the HAT domain sufficed to acetylate ADA2b.

Other experiments validated the finding of the acetylation of ADA2b by GCN5 *in vitro*. For example, in the HAT assays described in Chapter IV, where ADA2b was shown to enhance GCN5 acetylation activity on core histones, ADA2b was repeatedly found acetylated simultaneously (see Figure 3 of Chapter IV).

The N-terminus of ADA2b is acetylated by GCN5 *in vitro*.

Having found that ADA2b is acetylated by GCN5 *in vitro*, experiments were performed to determine which region of ADA2b was acetylated. The ADA2b protein was split into three regions as described in Chapter III, ADA2bN (amino acids 1-225), ADA2bM (aa 225-377), and ADA2bC (aa 353-486). Recombinant proteins of these three different regions were purified from *E. coli* as GST fusion proteins, and the GST tag was subsequently removed by protease PreScission. Acetyaltion assays were performed as before in which the three regions were assayed either alone or in combination. The result showed that ADA2bN, but not ADA2bM and ADA2bC, was acetylated by GCN5 (Figure 2). Interestingly, the HAT assays also found that ADA2bM could improve the acetylation efficiency of ADA2bN by GCN5 (Figure 2B, compare lane N to lane N plus M).

No ADA2 from other organisms has been reported to be acetylated by its partner GCN5 and or other HAT enzymes. Since the N-terminus is highly conserved among

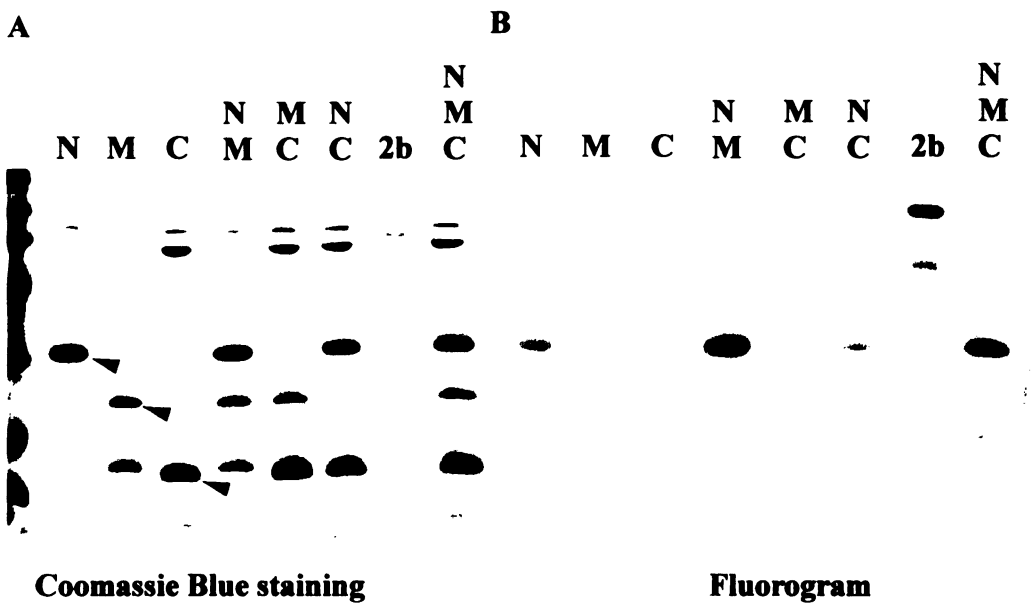


Figure 2. The N-terminus of ADA2b is acetylated by GCN5.
In the acetylation assays, about 2 μg ADA2bN, 0.5 μg ADA2bM, or 2 μg ADA2bC alone or in combination were acetylated by about 0.3 μg GCN5 with 0.15 μCi [^{3}H] Ac-CoA. The acetylation reactions were resolved in a 12% SDS-PAGE gel and the proteins in gel were Coomassie blue stained, as shown at panel A. The proteins ADA2bN, ADA2bM and ADA2bC are indicated by the blue arrows. The acetylated proteins in gel were then visualized by enhanced fluorography as shown at panel B. ADA2b in full length (2b) was also included as a positive control. N: ADA2b (aa 1-225), M: ADA2b (aa 225-377), C: ADA2b (aa 353-486).

ADA2 from different species, we tested whether the acetylation of the N-terminus of ADA2b was specific. The corresponding N-terminus of *Arabidopsis* ADA2a and budding yeast ADA2 was tested for acetylation by *Arabidopsis* GCN5. Recombinant proteins of ADA2a (amino acids 1-267) and GST-yAda2 (yeast Ada2, amino acids 1-176, as GST fusion) were purified. Similar acetylation assays found that the N-terminus of ADA2a, but not that of the budding yeast ADA2, was acetylated by *Arabidopsis* GCN5 (Figure 3). Again, the acetylation of ADA2aN was improved with the presence of ADA2bM (see Figure3). Based on the facts that the yeast GCN5 was capable to acetylate ADA2b (see below) and that yeast GCN5 HAT domain is very similar to that of *Arabidopsis* GCN5, it is very likely that yeast GCN5 can not acetylate the N-terminus of yeast Ada2. Therefore, the acetylation is limited to the *Arabidopsis* ADA2 proteins.

K215 of ADA2b is necessary for acetylation

The previous results show that the acetylation site of ADA2a or ADA2b was located at the N-terminus. The results also suggested that the acetylated residue(s) might not reside in the conserved sequences of ADA2 N-terminus. Indeed, there are extensions at the amino and carboxyl ends of *Arabidopsis* ADA2 N-terminus that are not found in yeast ADA2. Therefore, serial deletions at the both ends of ADA2bN were assayed to test the necessity of these extensions in acetylation. GST fusion plasmids expressing serial deletions of the N-terminus of ADA2b (amino acids 1-179, 28-225, 28-179, 1-208) were constructed, and the recombinant proteins were expressed and purified from *E. coli* (Figure 4).

Acetylation assays were performed as before with similar amounts of GCN5 and

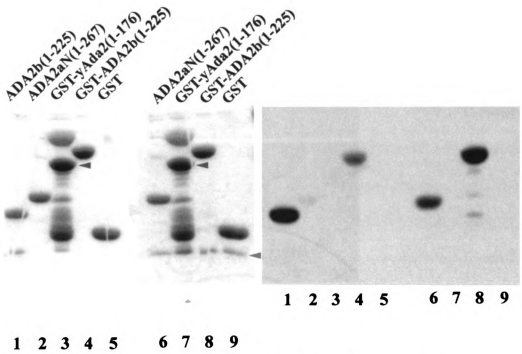
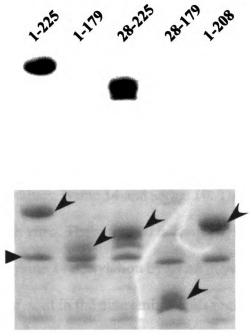


Figure 3. The N-terminus of *Arabidopsis* ADA2a is acetylated by GCN5. Acetylation assays were carried out using about 5 μg of each protein substrate and about 0.3 μg GCN5 plus 0.15 μCi [^{3}H] Ac-CoA. The acetylation reactions were analyzed in a 12% SDS-PAGE gel and the proteins in gel were Coomassie Blue stained first, as shown in the left panel. The acetylated proteins in gel were then visualized by enhanced fluorography as shown in the right panel. The protein positions of yeast N-terminus as GST fusion [GST-yADA2a (1-176)] is indicated by blue arrows. Lane 6, 7, 8, and 9 were actually the repeat of lane 2, 3, 4 and 5, except that lane 6, 7, 8, and 9 had the middle part of Ada (ADA2bM, amino acids 225-377) included in the reaction to improve the acetylation by GCN5. The ADA2bM is indicated by the red arrow in the protein gel in the right panel. Positive control of lane 1 also contained ADA2bM.

A



B

ADA2b	FV DRSFGGK^{215}KPVSTSVNN^{225}SLVELSNYN
Histone H3	TA RKSTGGK^{14} APRK
ADA2a	QS DRSVGEK^{257}KLRLPGEK^{265}VPLVTELYGYN
Histone H3	TA RKSTGGK^{14} APRK

Figure 4. Defining the acetylation site of ADA2b N-terminus.
(**A**) Acetylation assay of various deletions of the N-terminus of ADA2b. About 3 μg ADA2b (1-225, 28-225, 28-179, 1-208) and 2μg ADA2b(1-179) were acetylated by about 0.3 μg GCN5 with 0.15 μCi [^{3}H] Ac-CoA. About 1μg ADA2bM (amino acids 225-377) was also included in the reactions to enhance the acetylation. The acetylation reactions were resolved in a 12% SDS-PAGE gel and the proteins in gel were Coomassie Blue stained first, as shown in the lower panel. The included ADA2bM is indicated by the filled triangle. The truncated proteins of ADA2bN are indicated by the arrows. The acetylated proteins in gel were then visualized by enhanced fluorography as shown at the upper panel. (**B**) Sequence alignment of the partial ADA2a and ADA2b N-termini against the N-tail of histone H3. The partial sequence of ADA2b N-terminus includes the region of amino acids 208-225 required for acetylation. The serine and lysine residues are colored.

acetyl CoA. The results showed that only two deleted version of ADA2b (amino acids 1-225, and 28-225) could be acetylated (Figure 4). This suggested that the peptide region at amino acids 208-225 was necessary for acetylation by GCN5. Interestingly, this region is not present in ADA2 of yeast or metazoans, suggesting the uniqueness of ADA2 acetylation in *Arabidopsis*. Moreover, a stretch of nine amino acids within this region was found very similar to that of histone H3 N-tail (Figure 4B). The histone H3 N-tail contains two noteworthy residues, lysine 14 and serine 10. The lysine14 is the major site acetylated by yeast GCN5 *in vitro*. The serine 10 is phophorylated and its phosphorylation enhances lysine 14 acetylation by Gcn5 in yeast (Lo, et al., 2000). Corresponding residues also exist in the nine amino acids region of the N-terminus of ADA2b. A similar sequence is also present in the corresponding region of the N-terminus of ADA2a, but the ADA2aN seemed to have two similar stretches (Figure 4B).

The GCN5/PCAF protein does not recognize a precise motif to acetylate, but a general consensus sequence of GKXP has been suggested (Marmorstein, 2001), where the lysine is the acetylated residue. Sequence checking of the histone H3 tail-like region of ADA2a and ADA2b suggested that K215 might be the potential acetylation sites in ADA2b, and K257 and K265 in ADA2a (Figure 4B).

An alternative approach was also carried out to support this prediction. A parallel HAT assay showed that Gcn5 of budding yeast was able to acetylate *Arabidopsis* ADA2b (Figure 5), although the yeast Gcn5 was less efficient in acetylating *Arabidopsis* ADA2b than *Arabidopsis* GCN5. This revealed a common acetylation theme in protein ADA2b by GCN5 family protein.

In order to test the prediction, mutations were introduced into full-length ADA2b

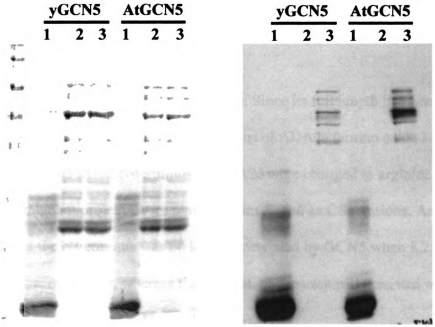


Figure 5. Yeast Gcn5 acetylates *Arabidopsis* ADA2b.
The HAT assay was performed in 26 µl HEPES buffer with about 0.5 µg GST-ADA2b (wt,and mutant K215A) or 20 µg core histones, 0.15 µCi [^{3}H] Ac-CoA, 0.5 µg *Arabidopsis* GCN5 (AtGCN5) or about 1 µg purified recombinant yeast Gcn5 (yGCN5). The reaction proceeded for 30 min at 26°C. Then the reactions were analyzed by 10% SDS-PAGE gel. The proteins in gel were first stained with Coomassie Blue as shown in the left. Then the acetylation signals of GST-ADA2b and histones were detected by enhanced fluorography, as shown in the right. 1: core histones, 2: GST-ADA2b (K215A), 3: GST-ADA2b (wt).

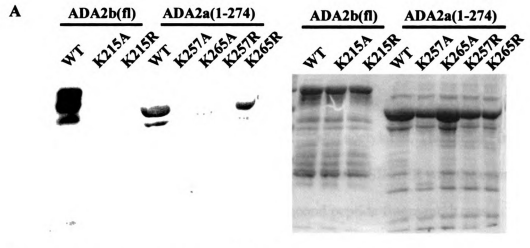
such that K215 was changed to arginine or alanine. The mutant proteins were assayed for acetylation by GCN5. The results showed that this point mutation of K215 either to arginine or alanine abolished acetylation of ADA2b by GCN5, while control wild type of ADA2b remained acetylated (Figure 6). Since arginine is very similar to lysine, there should not be big conformational change of ADA2b protein when K215 was changed to arginine.

A similar strategy was also applied to ADA2a. Since its full length had been hard to purify, the assay was still focused on the N-terminus of ADA2a (amino acids 1-272). Residues K257 or K265 in the N-terminus of ADA2a were changed to arginine or alanine. The mutant recombinant proteins were purified from *E. coli* as GST fusions. Acetylation assay showed that the N-terminus was no longer acetylated by GCN5 when K257 was changed to arginine or alanine, whereas the acetylation capacity still persisted when K265 was changed to arginine or alanine (Figure 6).

This mutational analysis suggested that K215 in ADA2b and K257 in ADA2a are the sites of acetylation by GCN5 *in vitro*. Curiously, about 20 amino acids around the lysine residue are also found to exist in ADA2 homologous from plant rice and maize, but not in ADA2s from budding yeast or any know metazoan organisms (Figure 6B).

Determination of the acetylation site of ADA2bN by mass spectrometry

The HAT assay result of mutant ADA2b demonstrated only the necessity of K215 for acetylation. It did not prove the sufficiency of K215 for acetylation. To directly determine which lysine residue(s) in ADA2b was acetylated by GCN5, the method of mass spectrometry was approached. To simplify the analysis, only the N-terminus of



B

athADA2a	QSDRSVGEK[257]	KLRLPGEK[265] VPLVTELYGYN
athADA2b	FVDRSFGGK[215]	KPVSTSVNN SLV ELSNYN
zmaADA2	HVDRSVGVK[263]	KPRYSADEG P SL TELSGYN
OsaADA2	HVDRSFGVK	KPRYSADEG P SL TELSGYN
Histone H3	TARKSTGGK[14]	APRK

Figure 6. Lysine residues in ADA2a (1-274) and ADA2b necessary for acetylation. (**A**) Testing the necessary lysine residues in ADA2a and ADA2b. The recombinant GST fusion proteins GST-ADA2a (1-274) and GST-ADA2b were either wild type (wt) or mutant where lysine residue at position 215 (for ADA2b), or 257 and 265 (for the N-terminus of ADA2a) was changed to arginine or alanine. These wild type and mutant recombinants bound to Sepharose beads were incubated with about 0.5 μg GCN5, 0.2 μCi [^{3}H] Ac-CoA for acetylation. After the acetylation was finished, the bound GST fusion proteins were eluted by boiling in SDS-PAGE loading buffer and electrophoresed in a 10% SDS-PAGE gel. The resolved protein in gel were stained with Coomassie Blue, as shown in the right. Then the acetylated signals were visualized by enhanced fluorography, as shown in the left. WT: wild type, K215R: ADA2b with lysine residue at position 215 mutated to arginine, and so on. (**B**) Sequence alignment shows that a stretch of amino acids around the lysine residues required for acetylation are found in ADA2 homologs from plants. athADA2: *Arabidopsis* ADA2, zmaADA2: corn ADA2, osaADA2: rice ADA2.

ADA2b (1-225, ADA2bN) was subjected to mass spectrometry.

In the mass spectrometry assay, ADA2bN was acetylated by GCN5 using cold acetyl CoA. Then the acetylated ADA2bN was SDS-PAGE purified and sent to the mass spectrometry facility at our department for further analysis, where the ADA2bN was first digested in gel by trypsin and analyzed by LC/MS/MS mass spectrometry.

The mass spectra showed that two peptide fragments were singly acetylated. One was the fragment of $SK^{137}EQCLEH$, where the K^{137} was acetylated. The second fragment was $SFGGK^{215}KPVSTSVNNS^{225}$ GSIELER, in which the GSIELER was the expressed peptide from the GST vector sequence. For the second peptide fragment, it was ambiguous from the data as to whether K215 or K216 was acetylated because of the fragment was too big.

Since the uncertainty of K215 or K216 acetylation was due to the large-sized peptide, smaller recombinant ADA2bN was prepared for mass spectrometry again. To do so, the original ADA2bN (1-225) expressing plasmid was re-subcloned so that 4 amino acids of $VNNS^{225}$ at the C-terminal end were deleted. The shorted ADA2bN was named ADA2bNR and the amino acids at the C-terminal end from expressed plasmid were ...$RSFVDRSFGGK^{215}KPVSTSSGRIVTD$ where peptide of GRIVTD was from expressed vector sequence. Now the trypsin digestion peptide from ADA2bNR flanking K215 was shorter: $SFGGK^{215}KPVSTSSGR$. The recombinant protein of ADA2bNR was expressed and purified as for ADA2bN. For mass spectrometry assay, the recombinant ADA2bNR was acetylated by GCN5 using tritium-labeled acetyl CoA. The reason for using tritium-labeled Ac-CoA was to label the acetylated fragment that could be purified after HPLC treatment. As done for ADA2bN, the acetylated ADA2bNR was SDS-PAGE

gel purified and sent to the mass spectrometry facility of our department, where the sample was digested with trypsin in gel and fractionated in two separate HPLC runs.

The HPLC fractions with peak tritium activity were subjected to MALDI mass spectrometry. MALDI results of the sample identified three peptides with possible acetylation. M/Z 1335.17 Da corresponded to *Arabidopsis* ADA2b aa 45-52 with one acetylated residue; M/Z 1350.08 Da corresponded to aa 178-187 with two acetylated residues; M/Z 1436.69 Da corresponded to aa 211-224 with one acetylated residue. aa 211-224 is the sequence of SFGGK^{215}KPVSTSSGR.

Since the MALDI data did not tell which K was acetylated in peptide aa 211-224, one saved HPLC fraction with peak tritium activity was further subjected to Q-TOF analysis. From Mascot Search, ADA2b was identified as the top hit with a score of over 75 by the Q-TOF mass data. M/Z 1350 Da was found to contain no acetylation sites, (for the real mass plus two acetylations should be 1333.4. It is not known how it was picked up as the peptide aa 178-187 with two acetylations in previous MALDI assay). MS/MS of 1437 Da identified the predicted peptide SFGGK^{215}KPVSTSSGR with one acetylation site. The detailed calculated masses for this peptide are listed in Table 7-1. The highlighted data in the table were the identified masses from Q-TOF. The Q-TOF mass data only matched the prediction that the K5 (i.e. K215 of ADA2b) was acetylated. The data did not match acetylation at K6, or null acetylation (see the y and b ions data in Table 7-1 below). Therefore it was certain that K215 of ADA2b was the acetylation site. The key mass data to determine that K5 was acetylated rather than K6 were the mass data underlined in the table.

In summary, the mass spectrometry data identified that K215 was acetylated in

Table 7-1 Q-TOF mass data for ADA2bNR

		Avg 1+			Mono 1+					Avg 2+				
Seq	#	b	y	#	Seq	#	b	y	#	Seq	#	b	y	#
S	1	88.1	1437.6	14	S	1	88.0	1436.7	14	S	1	44.5	719.3	14
F	2	235.3	1350.5	13	F	2	**235.1**	1349.7	13	F	2	118.1	**675.7**	13
G	3	292.3	1203.3	12	G	3	292.1	**1202.6**	12	G	3	146.7	**602.2**	12
G	4	349.4	1146.3	11	G	4	349.2	1145.6	11	G	4	175.2	**573.6**	11
K	**5**	519.5	1089.2	**10**	**K**	**5**	**519.2**	1088.6	**10**	**K**	**5**	260.3	**545.1**	**10**
K	6	647.7	919.0	9	K	6	**647.3**	**918.5**	9	K	6	324.4	**460.0**	9
P	7	744.8	790.9	8	P	7	744.4	**790.4**	8	P	7	372.9	395.9	8
V	8	844.0	693.7	7	V	8	843.5	693.4	7	V	8	422.5	347.4	7
S	9	931.0	594.6	6	S	9	930.5	**594.3**	6	S	9	466.0	297.8	6
T	10	1032.1	507.5	5	T	10	1031.5	507.3	5	T	10	516.6	254.3	5
S	11	1119.2	406.4	4	S	11	1118.6	406.2	4	S	11	560.1	203.7	4
S	12	1206.3	319.3	3	S	12	1205.6	319.2	3	S	12	603.7	160.2	3
G	13	1263.4	232.3	2	G	13	1262.6	**232.1**	2	G	13	632.2	116.6	2
R	14	1419.5	175.2	1	R	14	1418.7	175.1	1	R	14	710.3	88.1	1

ADA2b. Some evidence suggests acetylation on lysine residues such as K137 (in peptide SK^{137}EQCLEH). But it seems unlikely that K137 is acetylated for the deletion and mutant acetylation assays described earlier does not support this finding.

Part II Test the significance of ADA2b acetyaltion *in vivo*

Although the previous section clearly shows that ADA2a and ADA2b proteins can be acetylated by GCN5 *in vitro*, it is not known whether ADA2 is acetylated *in vivo*, nor is the biological significance of acetylation known. Experiments were carried out to test whether the mutant ADA2b is able to enhance GCN5 HAT activity and to be functional *in vivo* to understand some of the biological function of ADA2 acetylation.

Mutant ADA2b is able to enhance GNC5 HAT activity.

Recombinant GST-ADA2b (K/A or K/R) proteins were purified the same way as that of wild type GST-ADA2b (wild type, see chapter IV). HAT assays were carried out using the same conditions as for wild type GST-ADA2b (see chapter IV). As shown in Figure 7A, these two ADA2b mutants enhanced GCN5 HAT activity on free core histones as good as wild type ADA2b did, and the mutants did not change GCN5 substrate specificity (Figure 7B). Meanwhile, nucleosomal HAT assays also showed that mutant ADA2b proteins behaved like wild type ADA2b in stimulating GCN5 HAT (Figure 8). Therefore, lysine 215 in ADA2b is not important for the enhancement of GCN5 HAT activity.

Dwarf and sterile *ADA2b-/- Arabidopsis* plants

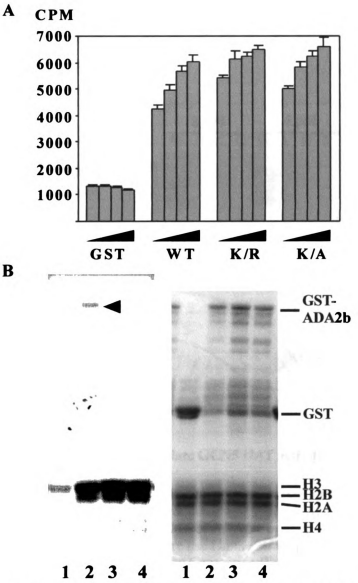


Figure 7. The ADA2b mutants enhance the HAT activity of GCN5 on free core histones.

HAT assays were carried out with purified recombinant GCN5, [H^3] labeled Ac-CoA, free core histones and the purified recombinant ADA2b proteins as GST fusion in wild type (WT), or alanine (K/A) or arginine (K/R) substitution at lysine 215. The the HAT activities wer quantitated by liquid scintillation counting on filter paper as shown by CPM readings in chart A. The reactions were also analyzed in a 15% SDS-PAGE gel. The proteins in gel were stained with Coomassie Blue, as shown in the right of panel B and then the acetylation signals were detected by enhanced fluorography. The acetylated wild type ADA2b is indicated by the arrow in panel B.

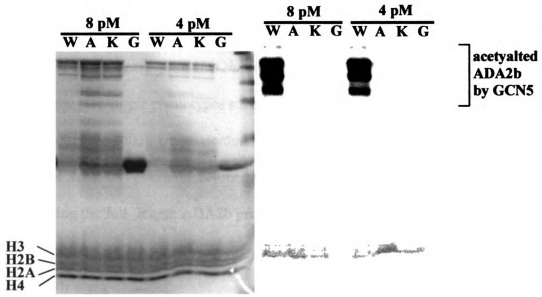


Figure 8. Mutant ADA2b proteins stimulate GCN5 HAT activity on nucleosomal histone.

The HAT assays were performed with purified turkey erythrocyte nucleosomes, [^{3}H] labeled Ac-CoA, GCN5 and GST fusion protein of wild type ADA2b (W), or alanine (A) or arginine (R) substitution at lysine K215 of ADA2b, or control GST (G). About 8 pm and 4 pm ADA2b as GST fusions were used in the assays]. The reaction products were analyzed by SDS-PAGE, and stained by Coomassie Blue (as shown in the left panel) and then subjected to enhanced fluorography for 4 months (as shown in the right panel).

Our lab has an *Arabidopsis* line which has an insertion of T-DNA at the *ADA2b* gene. The insertion site is at the fifth intron, and the ADA2b gene can only express the first four exons that encode approximately 160 amino acids of the N-terminal of ADA2b protein (Vlachonasios, et al., 2003). Phenotypically, plants with heterozygous T-DNA insertion at *ADA2b* loci (designated as *ADA2b+/-*) are wild type, while plants with homozygous T-DNA insertion at *ADA2b* loci (designated as *ADA2b-/-*, or *ada2b-1* in the literature) show pleiotropic phenotypes such as dwarf and sterility, as shown in Figure 9 (Vlachonasios, et al., 2003). Under normal laboratory growth conditions, the two severe phenotypes of *ADA2b-/-* plants are rescued by transgenic overexpression of ADA2b cDNA encoding the full- length ADA2b protein.

Transgenic strategy

Therefore, the *ADA2b-/-* plants provide a good system to test the function of mutant ADA2b which can no longer be acetylated by GCN5 *in vitro*. The approach was to express mutant ADA2b (K215R and K215A) in *ADA2b-/-* plants to investigate the degree of phenotype rescue.

To do so, *ADA2b+/-* plants (*ADA2b-/-* plants are sterile) were transformed with a specific binary plasmid that contained the mutated cDNA encoding ADA2b (K215R) or ADA2b (K215A). The transformation resulted in random insertion of the mutated cDNAs into *ADA2b+/-* plant genome together with an herbicide resistant marker bar1. The expression of both the mutated cDNA and the bar1 gene were under the control of the strong 35S promoter derived from cauliflower mosaic virus. Therefore, positive transgenic plant offspring could be screened by the feature of being resistant to herbicide

Figure 9. The dwarf phenotype of *Arabidopsis* with homozygous T-DNA insertion at the *ADA2b* locus.
The big plant is wild type with heterozygous T-DNA insertion at the *ADA2b* locus (genotype: *ADA2b*+/-). The three dwarf plants have homozygous T-DNA insertion at the *ADA2b* locus (genotype: *ADA2b*-/-).
Picture is provided by Dr. Kostas Vlachonasios.

Basta. A positive control encoding wild type ADA2b was also included in parallel transformants.

Totally 4 *ADA2b+/-* plants were transformed each by the plasmids containing cDNA of protein ADA2b (K215R), ADA2b (K215A), and ADA2b (wt, wild type). Seeds were collected from these plants as stage T1. These T1 seeds were allowed to germinate on solid Gamborg's B-5 medium with Basta selection. Resistant seedlings (about 7-10 days) were transferred to soil to grow for both genotype identification and growth phenotype examination.

Plant offspring with the genotype of *ADA2b-/-* were identified by the method of PCR screening. Single plant genomic DNA was extracted from one rosette leaf, and the *ADA2b* gene fragment spanning the T-DNA insertion site was PCR amplified using primers flanking the site of T-DNA insertion. The *ADA2b* gene with T-DNA insertion gave a PCR product of about 8.2 kb in size and the *ADA2b* gene without T-DNA insertion gave a PCR product of about 2.7 kb. An example of determining plant genotypes is shown in Figure 10A. After determining plant genotype, the integration of transgenic cDNA of ADA2b was also checked by PCR method using an ADA2b cDNA and 35S promoter specific primers which gave a PCR product of about 1.3 kb (Figure 10B).

ADA2b (K215R) can complement *ADA2b-/-* plant phenotype

To check the function of protein ADA2b (K215R) *in vivo*, plants with genotype of *ADA2b-/-* and integration of ADA2b (K215) cDNA [designated as *ADA2b-/-*:ADA2b (K215R)] were first identified and then checked for their growth phenotype.

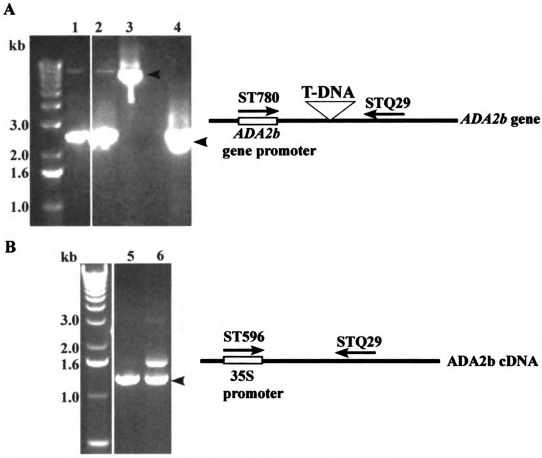


Figure 10. Genotyping the offspring of transformed plants.
The plants' genomic DNA was amplified by PCR and the products were analyzed in 1% agarose gel and the PCR products were visualized by ethidium bromide staining. (**A**) PCR analysis to determine the plants' genotypes. The used PCR primers' positions (ST780 and STQ29) in *ADA2b* gene in respective of the T-DNA insertion site are indicated in the right diagram. The 8.2 kb and 2.7 kb PCR products are indicated by the arrows. The genotypes of 1, 2, 3, and 4 plants are listed as the followings, 1, 2: *ADA2b*+/-, 3: *ADA2b*-/-, 4: *ADA2b*+/+. (**B**) Determining the integration of ADA2b cDNA into the genome in the transformed plants. The PCR primers (ST596 and STQ29) are 35S promoter and ADA2b cDNA specific as indicated in the right diagram. The 1.3 kb PCR products are indicated by the arrow.

Three such plants were identified, designated as R3, R5 and R12. Line R3 and R5 showed the mutant phenotypes, whereas R12 displayed a near-normal growth phenotype with reduced seed yields (Figure 11). This result indicated that the K215R mutant protein can suffice for most of the discernable biological activity of ADA2b protein, and thus the acetylation of this K215 is not essential for those biological activities.

One explanation for the lack of complement in line R3 and R5 might be that the transgene was integrated but not expressed. To test this hypothesis, the steady-state levels of ADA2b mRNA were assayed by Northern blotting. As shown in Figure 12, the mRNA levels of ADA2b mutant in line R5 are comparable to that in line K61 (genotype: *ADA2b-/-*:ADA2b (wt)), which carried the wild type transgene and showed normal growth phenotype. Expression of the K215R transgene in line R3 is much lower, but still comparable to the levels of that of endogenous *ADA2b*. These results suggest that levels of expression of the transgene might important in performing complementation.

To further characterize the ability of the K215R transgene to rescue the *ADA2b-/-* phenotype, nine T1 lines were selected with genotype of *ADA2b*+/-:ADA2b (K215R). T2 progeny of these lines were screened by PCR to identify plants with genotype of *ADA2b-/-*:ADA2b (K215R). For two T1 lines (designated as R45 and R48), 8 out of 8 dwarf T2 plants had the genotype of *ADA2b-/-*:ADA2b (K215R), whereas none of the 28 normal T2 plants examined had the genotype of *ADA2b-/-*:ADA2b (K215R). Expression of the K215R transgene was low or undetectable in these dwarf T2 plants as shown in Figure 13, reinforcing the conclusion that the expression level of the transgene might be critical for complementation.

Another T1 line R44 yielded T2 progeny with the genotype of *ADA2b-/-*:ADA2b

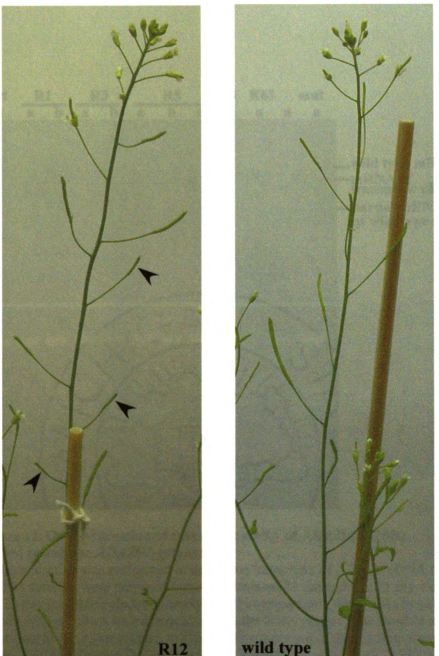


Figure 11. Comparison of the silique length of R12 plant to that of wild type plant. Plant R12 seeds were in T2 stage from parents that had the genotype of *ADA2b*-/- with the integration of the cDNA of mutant ADA2b (K215R) [designated as *ADA2b*-/-: ADA2b (K215R)]. The wild type and R12 plants seeds were geminated on Gamborg B-5 solid medium and planted in soil. The pictures were taken when the plants had been growing in the soil for 23 days. The genotype of plant R12 was *ADA2*b-/-:ADA2b (K215R). The short siliques in plant R12 are indicated by the arrows. Both figures are shown at the same scale.

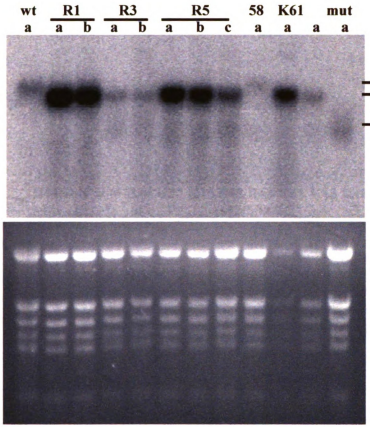


Figure 12. Over-expression of transgenic cDNA of ADA2b(K215R) did not rescue the *ADA2b-/-* phenotype.

Nothern blotting was performed to detect the transcripts of ADA2b cDNA that was in RNA samples were prepared from the plants on 21 day old (a), 23 day old (b) and 25 day (c). After electrophoresis in a 1% agarose gel , the rRNAs in gel were stained with ethidium bromide as shown in the lower picture. The Norther results are shown in the upper picture. mRNA of The genotype of these plants were listed above. R1: *ADA2b+/-*:ADA2b (K215R), R3: *ADA2b-/-*:ADA2b (K215R), R5: *ADA2b-/-*:ADA2b (K215R), mut: *ADA2b-/-*, 58: *ADA2b+/-*, K61: *ADA2b-/-*:ADA2b (wild type).

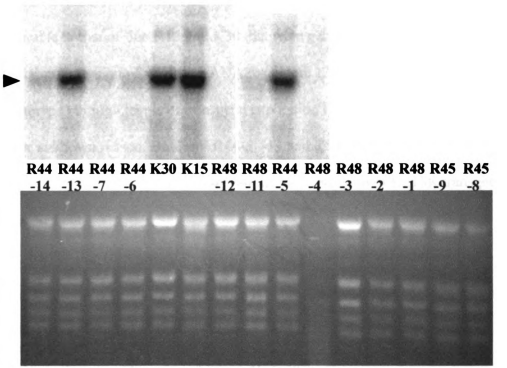


Figure 13. Over expression of transgene is required to rescue the *ADA2b*-/- phenotype.

Northern blotting was performed to detect the expression level of transgenic cDNA. About 5 µg total RNA sample, that was extracted from rosette leaf at day 20 after the seedling was transferred to soil, was loaded for each plant as indicted by the rRNAs in the lower panel. The detected mRNA levels of transgenic cDNA either as wild type (plant K15 and K30) or mutant (K215R, the rest plants) are shown in the upper panel as indicated by the arrow. All plants were dwarf except plant K15 and K30 which were normal in size, plant R44-13 which was small but not dwarf, and plant R44-5 which was dwarf and died prematurely. The genotype of the plants are listed as the followings, K15, K30: *ADA2b*-/-:ADA2b(wil type), the rest: *ADA2b*-/-:ADA2b(K215R).

(K215R). Some of these, displayed a near normal phenotype although the siliques were often reduced in size and number (such as line R44-21), while five others showed the mutant phenotypes (designated as line R44-5, -6, -7, -13 and -14). The mRNA levels of the transgene in line R44-6, 77 and -14 were low, while the mRNA level in line R44-13 was comparable to that of line K15 and K30 that were genotype *ADA2b*-/- and rescued by transgenic overexpression of wild type ADA2b (Figure 13). Line R44-5, which had a dwarf phenotype but high levels of transgene expression (Figure 13), died before maturity for unknown reason. But lines R44-5, -13 and R5 as described above suggest that there is exception to the general observation that expression of K215R transgene is able to complement.

In another T1 line (designated as R16), three T2 plants from all examined 12 plants were identified as *ADA2b*-/-:ADA2b (K215R) (the rest had wild type *ADA2b* gene). These three plants (designated as R16-32, R16-44, and R16-48) showed normal growth and seed production, although one plant had somewhat smaller roseate leaves, and all the three plants had fewer siliques than the control with the genotype *ADA2b*+/-. This results showed that expression of K215R in T2 can complement in *ADA2b*-/- plants. The growth state of the three plants after 22 days in soil was shown in Figure 14.

For lines R12, R44-21 and R16-48 (as described above), general growth and transgene expression were examined again in their T3 plants in comparison to that of plants phenotypically complemented by wild type transgene (line K51 and K61, genotype of *ADA2b*-/-:ADA2b(wt)). As shown in Figure 15, these plants grew to normal size, and produced good amount of seeds (comparable to that of wild type). The steady state mRNA levels of the transgenes in these plants were also similar and they all were

Figure 14. The phenotype rescue of plant *ADA2b-/-* by the expression of transgenic cDNA of ADA2b (K215R) in line R16.
These plants were at T2 stage from T1 line R16. The growth status of the rescued plant R16-48, R16-45 and R16-32 is compared to that of the wild type plants which are the rest plants. The picture was taken when the plants had been growing in the pot for 22 days. The genotype of R16-48, R16-45 and R16-32 plants are *ADA2b-/-*:ADA2b (K215R).

Figure 15. The acetylation of ADA2b at lysine 215 *in vivo* is not required for plant development.

These plant seeds germinated on solid Gamborg's B-5 medium without Basta selection. The germinated seedlings were transferred to soil pots. Pictures were taken after the plant had been growing in the pot for 10 days. The plants genotype as well as T stage are indicated. The last figure shows the severe dwarf phenotype of plant with homozygous T-DNA insertion at *ADA2b* locus.

overexpressed (Figure 16). The most obvious distinction from wild type plants was that these plants (including K51 and K61) had some shorter siliques along the shoots (see Figure 11 for illustration) and these plants were delayed in flowering by 2-3 days. Since this defect also happened in plants K51 and K61, it was not unique to plant R12, R44-21, and R16-48. The expression of transgene in R12, R44-21 and R16-48 was as the same level as that of line K51 and K61, but they all had much higher expression than endogenous *ADA2b* gene.

ADA2b (K215A) can complement *ADA2b-/-* plant phenotype

The described data above suggest that acetylation of ADA2b at K215 *in vivo* is not required for normal function. Similar complementation experiments were performed to test whether the K215 is important for ADA2b by changing the residue to alanine.

As performed for K215R screening, T1 plants from transformed parents with K215A cDNA were screened by PCR as described above to identify genotype of *ADA2b-/-*:ADA2b (K215A). In one screening, 34 T1 plants (17 normal size and 17 dwarves) were screened. Of these, 5 dwarf plants were genotyped as *ADA2b-/-*:ADA2b (K215A), none of the normal size plant had this genotype.

In another screening, 24 normal-sized T1 plants were screened (dwarf plants were not screened). One plant, designated as A14, had the genotype of *ADA2b-/-*:ADA2b (K215A). Plant A14 had normal size and produced similar seeds as that of plant R16 identified above with genotype of *ADA2b-/-*:ADA2b(K/R).

To assess the growth of A14, T2 stage plants of A14 and T4 plants of R16-48 and K51 as well as wild type were grown for comparison. These plant seeds were germinated

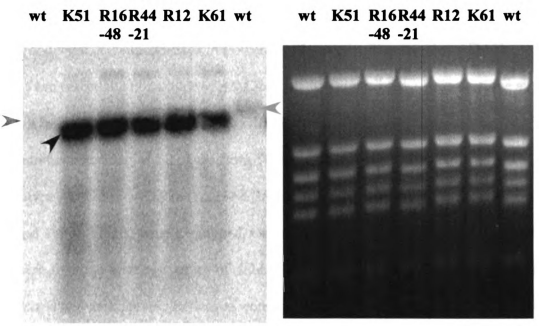


Figure 16. Over expression of transgenic cDNA is required to rescue *ADA2b-/-* phenotype.

Northern blotting was performed to show the mRNA level of either wild type transgenic cDNA in plant K61 and K51, or mutated transgenic cDNA (K215R) in plants R12, R16-48 and R44-21. All the plants were normal in size. Two wild type plants were also included. About 5 μg total RNA, that was extract from the plant rosette leaf at day 20 after the seedling was transferred to soil, was loaded in the agarose gel for each plant. The right panel shows the rRNAs for each sample after electrophoresis in agarose gel, indicating the equal loading for each sample. The result of Northern blotting by auto-radiography is shown in the left panel. The black arrow indicates the probed mRNA of ADA2b transgenic cDNA. The probed mRNA of wild type endogenous *ADA2b* gene is indicated by blue arrow. The genotype for plants K61 and K51 was *ADA2b-/-*:ADA2b (wild type), and for plants R12, R16-48, R44-21 is *ADA2b-/-*:ADA2b(K215R).
wt: wild type.

in Gamborg's medium without Basta selection. For A14, 132 T2 seeds were germinated, 39 yielded dwarf plants indicating heterozygous status of the transgene. The T2 plants of A14 growing in soil were very similar to that of R16-48 and K51 as shown in Figure 17. All plants of R16-48, K51 and K14, compared to wild type control, had some shorter siliques, and were delayed 2-3 days in flowering. The results illustrate that ADA2b (K215A) can rescue the phenotype of *ADA2b-/-* plant, suggesting the lysine residue at 215 is not important for plant normal growth under laboratory conditions.

In another screening of T2 plants from T1 parent plant (designated as A4) that had the genotype *ADA2b+/-*:ADA2b (K215A), 25 T2 plants (including 4 dwarves, 2 small plants and 3 short plants) were genotyped. Only the three short plants and the two small plants were found to be the genotype of *ADA2b-/-*:ADA2b (K/A). These short plants had normal-sized and dark leaves, the leaves were curved, and these three plants were a few days delayed in inflorescent shooting. These three plants never produced seeds. The two small plants were about the double size of dwarf plant of genotype of *ADA2b-/-*. These small and short plants were all sterile. The growth of these small and short plants was shown in Figure 18 and Figure 19 in comparison with wild type plants.

Similarly, in another screening of T2 plants from T1 parent plant (designated as A21) that had genotype of *ADA2b+/-*:ADA2b (K215A), 24 T2 plants (including 2 dwarf and 2 small and 4 short plants) were screened. All the dwarf, small and short plants were identified to be the genotype of *ADA2b-/-*:ADA2b (K215A). These 4 short plants had the same phenotype as that of the short plants described above for A4.

To examine the expression of the transgene in these dwarf, short and small plants, Northern blotting was performed. As shown in Figure 20, they all had overexpression of

plant K51 T4
***ADA2b-/-*: ADA2b (wild type)**

plant R16-48 T4
***ADA2b-/-*: ADA2b (K215R)**

plant A14 T2
***ADA2b-/-*: ADA2b (K215A)**

wild type and dwarf

Figure 17. Expression of transgenic cDNA (K215A) rescues plant *ADA2b-/-* phenotype.

These plant seeds germinated on solid Gamborg's B-5 medium without Basta selection. The germinated seedlings were transferred to soil pots. Pictures were taken after the plant had been growing in the pot for 10 days. The plants genotype as well as the T stage are indicated. The right bottom figure also shows a dwarf plant with genotype of *ADA2b-/-*.

Figure 18. Partial rescue of phenotype of plant *ADA2b*-/- by over-expression of transgenic cDNA of ADA2b (K215A).
The plants were all at T2 stage from A4 transformed parent. The growth status of three short plants (A, B and C) and two small plants (D and E) were compared to that of the three wild type plants (X,Y and Z). The short plants had normal leaf size but short inflorescent shoots. The small plants were about the double size of *ADA2b*-/- dwarf plants. The picture was taken when the plants had been growing in the pot for 18 days. The genotype of the small and short plants was *ADA2b*-/-:ADA2b (K215A).

Figure 19. Partial rescue of the phenotype of plant *ADA2b*-/- by over-expression of transgenic cDNA of ADA2b (K215A).
The growth status of the short plant C was compared to that of the wild type plant Y in later growth stage. Their growth in earlier stage was shown in Figure 18. The insert shows the whole pot with all the plants shown in Figure 18. The short plant had shorter and bushy shoots, and was sterile although it did bear some short siliques. The wild type plant bore mature siliques. The picture was taken when the plants had been growing in the pot for 30 days. The genotype of the small C plant was *ADA2b*-/-:ADA2b (K215A).

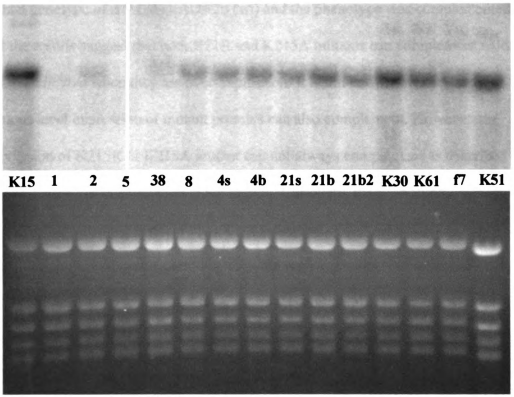


Figure 20. Complementation variation with over-expression of transgenic cDNA to rescue *ADA2b-/-* phenotype.

Northern blotting was performed to detect the mRNA levels of the transgenic cDNA in the *ADA2b-/-* plants. The RNA samples were prepared when the plants were 27 day old in soil. The RAN samples were loaded in equal total amount for each plant, as indicated by the rRNAs in lower panel which were stained by ethidium bromide. The results of the Northern blotting is shown in the upper panel. All plants were in T2 stage except K15 and K30 which were in T1 stage. Plant 4s and 4b were the "small" and "short" plant respectively referred in the results part for line A4, and the same case for plant 21s and 21b. Plant 21b2 was a second "short" plant of A21 line. The rest plants were identified in parallel screenings unless specified. The genotypes of all plants were *ADA2b-/-* except plant 38 (*ADA2b+/-*), and the transgenic cDNA for each plant and its growth phenotypes is listed in the following. 1: :ADA2b(K215R), dwarf and sterile, 2: :ADA2b(K215A), dwarf and sterile, 5: :ADA2b(K215R), dwarf and sterile, 8: :ADA2b(K215R), small size, sterile, 4s, 21s: :ADA2b(K215A), about the double size of dwarf plant and sterile, 4b, 21b, 21b2: :ADA2b(K215A), wild type leaf, short and delayed fluorescent shoots, K51, K15, K30: :ADA2b(wt), wild type, K61: :ADA2b(wt), wild type size, some shorter siliques, K61: :ADA2b(wt), wild type size, some shorter siliques, f7: :ADA2b(wt), wild type size leaf, delayed shooting and half normal length of shoot, and sterile.

the transgene, and the levels were even to that of plant K61, which, as mentioned above, had genotype of *ADA2b-/-*:ADA2b (wt) and the phenotype was complemented.

All the results suggest that both K21R and K215A mutants can complement wild type ADA2b *in vivo* when they are over-expression. It is not known whether similar endogenous level expression of mutant proteins can also complement. However, the overexpression of K215R or K215A mutant can not always complement as described above. The reason for this variation is not known. Curiously, the phenotype complementation variation is not unique to the transgene of K215R and K215A, but also happened during the complementation by the wild type transgene. For example, one screening was performed to identify *ADA2b-/-* plants complemented by the wild type transgene. Totally 117 normal plants and three short T2 plants were screened by PCR from 11 lines that were genotyped as *ADA2b+/-*:ADA2b (wt). Only these three short plants (designated as a3, e9 and f7) were genotyped as *ADA2b-/-*:ADA2b (wt), the remaining 117 plants were genotyped as either *ADA2b+/-* or *ADA2b+/+*. Plant a3 had small leaf (but not dwarf) and dwarf inflorescent shoots. Plants e9 and f7 had normal size leaf but darker in color, dwarf shoots in early stage, and their shoots grew to wild type in later stage (42 days after the seedlings transferred to soil), but had shorter and sterile siliques. The mRNA level of the wild type transgene in plant f7 was examined. As shown in Figure 20, plant f7 had similar level of transgene expression as that of plant K61.

Discussion

The major conclusions of this work are that (i) *Arabidopsis* protein ADA2, including ADA2a and ADA2b, can be acetylated by GCN5 *in vitro*, and this acetylation does not

happen to yeast ADA2 protein, (ii) the acetylation site was identified to be residue of lysine 215 in ADA2b, (iii) the peptide sequence surrounding the acetylation site is similar to the N-tail of histone H3 and is unique to plant ADA2, (iv) the acetylation of lysine 215 is not essential for ADA2b protein function *in vivo* for normal plant growth under laboratory conditions.

That the mutant ADA2b protein (K215R and K215A) behaves like wild type *in vivo* raises a few questions. First, the acetylation of *Arabidopsis* ADA2 protein may be just an *in vitro* artificial effect. Secondly, acetylation of ADA2 may occur *in vivo* but its importance may not be demonstrated in the assays we have used. For instance, ADA2 acetylation might happen only in specific biological responses in plants such as pathogen attach response, UV damaged DNA repair, and so on. Since only the normal laboratory growth condition has been tested to check the function of ADA2b mutants, this can not exclude the possibility that in other stressful growth conditions the acetylation of ADA2 is required. Interestingly, very preliminary data indicate that plant seedlings of *ADA2b-/-*:ADA2b (K215A) is weaker in resisting to micro-organisms attach than seedlings of *ADA2b-/-*:ADA2b (wt). More careful testing is needed to support this finding. Curiously, many stress response transcriptional factors in *Arabidopsis* implicated in pathogen attack, dehydration and high salt treatment contain AP2 domain (Park, et al., 2001, Riechmann and Meyerowitz, 1998, Stockinger, et al., 1997), and as shown in Chapter VI, ADA2 proteins can interact with some AP2 domain- containing transcriptional activators.

Another possible explanation for the complementation by K215 mutants *in vivo* is that the over-expression of transgene of ADA2b mutant might conceal the acetylation importance, although it is not known whether the protein level of mutant protein is also

over-expressed in transgenic plants or not. Although ADA2a can not complement ADA2b-/- plants, but the ADA2b acetylation per se might be compensated for by ADA2a acetylation. A recent research in budding yeast (Gu, et al., 2003) indicates that, for some duplicated genes, the genetic defect can only been seen when both the duplicated genes are not functional.

Regardless, the observation that K215 mutants can function *in vivo* raises the key of question whether ADA2 proteins *in vivo* are really acetylated. Theoretically, this question can be answered in the following way. ADA2 protein could be first isolated from plant cell lysates by immuno-precipitation, perhaps in combination with other biochemical method such as chromatography. Then the purified protein could be subjected to mass spectrometry to determine if ADA2 is acetylated. Alternatively, the purified ADA2b protein acetylation status could be determined by specific antibodies. The antibodies could be raised against synthetic ADA2b peptide along the acetylation site in which the lysine residue is either acetylated or unacetylated (as a negative control). Technically, there are a few good sides of this approach. First, ADA2a or ADA2b antigen to make the antibody is not limiting for, as shown at Chapter IV, the recombinant N-terminus and the C-terminus of ADA2a and ADA2b could be purified in very good purity and quantity. Secondly, the current mass spectrometry techniques are able to detect picogram scale protein, therefore no huge amount of *Arabidopsis* plant material required. Finally, specific antibodies against acetylated and unacetylated peptides have been successfully used in many other similar occasions. The major concern of the above approach is that the acetylation of ADA2 might occur only in unknown special biological responses.

The *ADA2b-/-* phenotype is generally rescued by mutant ADA2b protein, but some

phenotypic distinctions remained. Some siliques are shorter (as shown in Figure 11), and flowering is delayed by about 2-3 days. These defects are also observed in *ADA2b-/-* plants rescued by wild type ADA2 (such as in K61 and K51 plants). These defects are not likely to arise due to simply the overexpression of ADA2b protein, since overexpression of wild type or mutant ADA2b (K215R) does not produce these problems in plants with wild type genotype.

That mutant ADA2b (K215R) behaves like ADA2b (K215A) is somewhat surprising. Arginine is very similar to lysine, so substitution of K215 by arginine should not change ADA2b protein conformation very much, while alanine substitution might have dramatic effect on the overall structure of ADA2b protein. However, two arguments support the notion that the region flanking K215 in ADA2b may not take a defined structure. First, the histone H3 N-tail, the native substrate of GCN5 protein, is a random coil (Luger, et al., 1997). Secondly, the region flanking K215 in ADA2b is only found in plant ADA2, and is not inside the conserved parts of ADA2 protein. These two facts might argue that the region in ADA2b dose not take a specific conformation, therefore alanine substitution to K215 will not affect the overall conformation of ADA2b protein.

As mentioned in the introduction, HIV Tat protein is acetylated by PCAF. The acetylated sites in Tat are at K28 and K50 (Deng, et al., 2000). The sequence surrounding K50 is very similar to that of ADA2b. In Tat the sequence is SYGRKK, and in ADA2b the sequence is SFGGKK. Recently, acetylated Tat at K50 has been shown to bind specifically to the bromodomain of PCAF and that this interaction competes effectively against HIV-1 TAR RNA binding to the acetylated Tat (Mujtaba, et al., 2002). Bromodomains are known to bind acetylated lysine residue (Zeng and Zhou, 2002) and

this domain is present in nearly all nuclear histone acetyltransferases including *Arabidopsis* GCN5.

Chapter Eight

Perspective and Hypothesis

The GCN5-ADA2 complexes in *Arabidopsis*

This study show that the *Arabidopsis* GCN5 protein is a histone acetyltransferase and associates with *Arabidopsis* ADA2a or ADA2b proteins *in vitro*. Based on these two facts, it is hypothesized that there are GCN5-ADA2 complexes in *Arabidopsis*, and one complex contains GCN5 and ADA2a proteins and the other contains GCN5 and ADA2b proteins.

Recombinant GCN5 and ADA2b proteins together acetylate nucleosomal histone H3 weakly in comparison to the budding yeast HAT complex SAGA that has more than ten subunits. The *Arabidopsis* GCN5-ADA2 complexes are hypothesized to contain unidentified protein components in addition to the proteins GCN5 and ADA2. One function of the additional proteins is hypothesized to enable GCN5 to acetylate nucleosomal histones effectively. The *Arabidopsis* genome sequence [http://www.tigr.org/tdb/e2k1/ath1/ath1.shtml] encodes putative proteins homologous to the proteins TAF5, TAF6, TAF9, TAF10, TAF12, and Tra1 found in budding yeast SAGA complex, but the genome does not encode apparent homologues of Ada1, Ada3, Ada5/Spt20, Spt3, Spt7 and Spt8 also found in budding yeast SAGA complex (Grant, et al., 1997) and Ahc1 that is the component of budding yeast ADA complex (Eberharter, et al., 1999). Therefore, it is likely that the *Arabidopsis* GCN5-ADA2 complexes have distinct components. It is eminent in the next step to investigate what the other protein components are in the *Arabidopsis* GCN5-ADA2 complexes.

The elucidation of the components of *Arabidopsis* GCN5-ADA2 complexes will be beneficial in many aspects. One key goal in the future research is to find out what the biological roles the complexes are exactly play *in vivo*. The identification of the components might shed light on some biological roles the complexes play *in vivo*. In addition, the SAGA complex is well conserved in metazoans (Kusch, et al., 2003, Ogryzko, et al., 1998). It is not known whether that *Arabidopsis* GCN5-ADA2 like complexes are conserved in plants. GCN5 and ADA2 homologs have been found in maize (Bhat, et al., 2003), implicating those other plants might have similar GCN5-AAD2 complexes. The identification of components of *Arabidopsis* GCN5-ADA2 complexes might also pave the way for characterization of related GCN5-ADA2 complexes in other plants.

On the other hand, it will be very interesting to know how similar the GCN5-ADA2a complex is to the GCN5-ADA2b complex in terms of the protein components. The thesis study finds differences in the detailed interactions between GCN5-ADA2a and GCN5-ADA2b. Genetics studies discovered that *Arabidopsis* plants with homozygous disruption of the ADA2b gene have pleiotropic phenotypes while plants with homozygous disruption of the ADA2a gene dud not have any visible defects (Vlachonasios, et al., 2003). Moreover, *Arabidopsis* plants with homozygous disruption of GCN5 gene show pleiotropic phenotypes that are similar but do not overlap completely with the phenotypes in plants with homozygous disruption of the ADA2b gene (Vlachonasios, et al., 2003). All these facts lead to the hypothesis that the GCN5-ADA2a complex is very likely different from the GCN5-ADA2b complex. This hypothesis is also supported by the recent description that the GCN5-ADA2a complex is different from the GCN5-ADA2b

in fly (Kusch, et al., 2003). Interestingly, *Arabidopsis* has two homologues for each of budding yeast Tra1, TAF6, and TAF12. If these homologous proteins are components of GCN5-ADA2 complexes, it is not known whether one group of these homologues is exclusively in one GCN5-ADA2 complex. Therefore, during the endeavor of identifying the components of *Arabidopsis* GCN5-ADA2 complexes, it is equally important to distinguish the components of GCN5-ADA2a complex from that of the GCN5-ADA2b complex.

The targets of GCN5-ADA2 complexes

Although the *Arabidopsis* GCN5-ADA2 complexes are featured as chromatin covalent modifiers, the specific targets for the modification by the complexes are largely unknown. *Arabidopsis* plants with homozygous disruption of ADA2b or GCN5 show multiple growth defects. Microarray data indicated that the transcription of many genes involved in a variety of cellular processes is up- or down-regulated in these disrupted mutants (Vlachonasios, et al., 2003). This suggests that the GCN5-ADA2 complexes are targeted to many biological activities, but it is unlikely that all these effects found in microarray data are primary.

The thesis study shows that two *Arabidopsis* proteins TINY and CBF1 physically interact with ADA2 proteins *in vitro*. Both TINY and CBF1 proteins are AP2 domain containing transcriptional factors Jaglo-Ottosen, 1998 #96] (Wilson, et al., 1996), The AP2 domain is a DNA binding domain only found in plants. There are about 140 AP2 domain-containing proteins in *Arabidopsis* (Riechmann, et al., 2000). Only five AP2 domain-containing proteins (ERF4, ERF5, DREb2a, TINY and CBF1) were tested for

their interaction with ADA2 proteins, and two show interaction. It is hypothesized that more AP2 proteins may interact with the GCN5-ADA2 complexes. The AP2 transcriptional factors are involved in a variety of biological functions in plants such as response to abiotic and biotic stresses and hormonal regulation (Riechmann and Meyerowitz, 1998). The proteins CBF1 a transcriptional activator involved in *Arabidopsis* response to cold stress (Stockinger, et al., 1997), and over expression of protein TINY results in tiny plants (Wilson, et al., 1996). Therefore it hypothesized that the GCN5-ADA2 complexes participate in many cellular activities associated with AP2 proteins, and the study of searching for more interacting AP2 proteins with GCN5-ADA2 complex promise to be fruitful.

In addition, the two candidates isolated by yeast two-hybrid screening with GCN5 as bait are interesting for further investigation (see chapter V). One is candidate #77, a basic helix-loop-helix transcriptional activator. It is not known what genes this activator activates. It is also curious whether other basic helix-loop-helix transcriptional activators interact with the GCN5-ADA2 complexes. Another candidate #141 implicated as ATP-driven RNA or DNA helicase participating in variety of cellular processes such as DNA recombination and repair. Increasing evidences show that nuclear HAT complexes might participate in DNA-damage repair. The human SAGA-like complex STAGA associates with components of UV-damaged DNA binding complex (UV-DDB) (Martinez, et al., 2001). And two transcriptional HAT complexes, the human TIP60 and yeast INO80, both contain proteins related to DNA helicase and both complexes have DNA helicase activity (Ikura, et al., 2000, Legube, et al., 2002, Martinez, et al., 2001, Shen, et al., 2000). Human cells with impaired HAT activity in TIP60 have defective double-strand break

repair (Ikura, et al., 2000), and yeast with null INO80 shows hypersensitivity to DNA damage (Shen, et al., 2000). Moreover, loss of GCN5 protein in mouse embryos results in increased apoptosis that is intimately reconnected to DNA damage (Xu, et al., 2000). The huge Tra1 protein in yeast SAGA belongs to the ATM family implicated in cell cycle checkpoint signaling and cellular response to DNA damage (Abraham, 2001, Khanna, et al., 2001, Vassilev, et al., 1998). Therefore, the *Arabidopsis* GCN5-ADA2 complexes are very likely involved in DNA damage repair, and this might be a potential area for future study as addition to the biological functions of the complexes. Approachable experiments might be to investigate the capacity of DNA-damage repair in the mutant ADA2b or GCN5 plants as described above.

The potential function of the unique N-terminus of the GCN5 protein

The N-terminus of *Arabidopsis* GCN5 (aa 1-200) is completely different to the N-terminus of other animal GCN5/PCAF proteins. The function of this N-terminus remained unknown. It is not required for GCN5 HAT activity *in vitro* and no interacting proteins have been identified by the strategy of yeast two-hybrid screening. Similar N-terminal regions also exists in GCN5 proteins of rice and maize, suggesting the importance of this N-terminus to plant GCN5 proteins. And curiously, the N-terminus in *Arabidopsis* is encoded on exon (chapter I), suggesting the N-terminus might function as unit.

The identification of other components of GCN5-ADA2 complexes might provide clues as to the possible role of the N-terminus of GCN5, and perhaps it has to wait to until the identification of the components of GCN5-ADA2 complexes before there is a

glimpse of the function of the N-terminus. However, since plants with disrupted GCN5 gene show pleiotropic phenotypes that can be rescued by transgenic expression of GCN5 protein (Vlachonasios, et al., 2003), it is practical to investigate whether GCN5 protein with the N-terminus deleted will behave like wild type GCN5 in the mutant GCN5 plants.

The acetylation of ADA2 proteins by GCN5

The serendipitous finding that both ADA2a and ADA2b can be acetylated by GCN5 *in vitro* is novel, and no similar cases have been reported in yeast or metazoan ADA2 proteins. Since the sequence along the acetylation site is only conserved in plant ADA2 proteins, it is hypothesized that the acetylation event is unique to plant ADA2 proteins. As discussed in chapter VII, one major question is whether ADA2 proteins are acetylated *in vivo*, and this has to be answered first to investigate the significance of ADA2 acetylation. Nevertheless, the conservation of the sequence flanking the acetylation site in plant ADA2 proteins suggests that there be functions associated with the region. One approach to investigate the significance of this region might be to test the function of the ADA2b protein that has multiple mutations at the conserved region, such as simultaneous mutations at serine 210 and lysine 215 for both residues are conserved in similar histone H3 N-tail.

Although mutant ADA2b that is not acetylated by GCN5 behaves like wild type ADA2b in enhancing GCN5 HAT activity on nucleosomal histones, it is still not known whether this residue is important in native GCN5-ADA2 complexes that are hypothetically have more components. On the other hand, the significance of ADA2 proteins acetylation *in vivo* is still an open question. First, only overexpression (in the

level of mRNA) of ADA2b mutants was tested, and overexpression might conceal the defects associated with mutant ADA2b if the protein level is also overexprssed. It is obligate to test whether similar level expression of mutant to that of endogenous wild type gene can perform complementation in ADA2b-mutated plants. But, as the positive control, it is fundamental to find out first whether similar level of transgenic expression of wild type ADA2b cDNA is able to complement. Secondly, only the normal laboratory growth conditions were tested and only gross phenotypes were examined for the complementary function of the mutant ADA2b. It is not known whether the mutant ADA2b protein can complement in other growth conditions and whether there are subtle differences in complemented plants in microscopic detail. Therefore, it is meaningful to investigate the complement ability of mutant ADA2b under other growth conditions such UV-light exposure and pathogen attack.

Bibliography

1. **Aasland, R., A. F. Stewart, and T. Gibson** 1996. The SANT domain: a putative DNA-binding domain in the SWI-SNF and ADA complexes, the transcriptional co-repressor N-CoR and TFIIIB Trends Biochem Sci. **21:**87-8.

2. **Abraham, R. T.** 2001. Cell cycle checkpoint signaling through the ATM and ATR kinases Genes Dev. **15:**2177-96.

3. **Agalioti, T., G. Chen, and D. Thanos** 2002. Deciphering the transcriptional histone acetylation code for a human gene Cell. **111:**381-92.

4. **Agalioti, T., S. Lomvardas, B. Parekh, J. Yie, T. Maniatis, and D. Thanos** 2000. Ordered recruitment of chromatin modifying and general transcription factors to the IFN-beta promoter Cell. **103:**667-78.

5. **Akhtar, A., and P. B. Becker** 2001. The histone H4 acetyltransferase MOF uses a C2HC zinc finger for substrate recognition EMBO Rep. **2:**113-8.

6. **Albright, S. R., and R. Tjian** 2000. TAFs revisited: more data reveal new twists and confirm old ideas Gene. **242:**1-13.

7. **Allen, M. D., K. Yamasaki, M. Ohme-Takagi, M. Tateno, and M. Suzuki** 1998. A novel mode of DNA recognition by a beta-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA Embo J. **17:**5484-96.

8. **Amedeo, P., Y. Habu, K. Afsar, O. M. Scheid, and J. Paszkowski** 2000. Disruption of the plant gene MOM releases transcriptional silencing of methylated genes Nature. **405:**203-6.

9. **Balasubramanian, R., M. G. Pray-Grant, W. Selleck, P. A. Grant, and S. Tan** 2002. Role of the Ada2 and Ada3 transcriptional coactivators in histone acetylation J Biol Chem. **277:**7989-95.

10. **Bannister, A. J., R. Schneider, and T. Kouzarides** 2002. Histone methylation: dynamic or static? Cell. **109:**801-6.

11. **Bannister, A. J., P. Zegerman, J. F. Partridge, E. A. Miska, J. O. Thomas, R. C. Allshire, and T. Kouzarides** 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain Nature. **410:**120-4.

12. **Banno, H., Y. Ikeda, Q. W. Niu, and N. H. Chua** 2001. Overexpression of Arabidopsis ESR1 induces initiation of shoot regeneration Plant Cell. **13:**2609-18.

13. **Barbaric, S., H. Reinke, and W. Horz** 2003. Multiple mechanistically distinct functions of SAGA at the PHO5 promoter Mol Cell Biol. **23:**3468-76.

14. **Barberis, A., and L. Gaudreau** 1998. Recruitment of the RNA polymerase II holoenzyme and its implications in gene regulation Biol Chem. **379:**1397-405.

15. **Barlev, N. A., R. Candau, L. Wang, P. Darpino, N. Silverman, and S. L. Berger** 1995. Characterization of physical interactions of the putative transcriptional adaptor, ADA2, with acidic activation domains and TATA-binding protein J Biol Chem. **270:**19337-44.

16. **Barlev, N. A., L. Liu, N. H. Chehab, K. Mansfield, K. G. Harris, T. D. Halazonetis, and S. L. Berger** 2001. Acetylation of p53 activates transcription through recruitment of coactivators/histone acetyltransferases Mol Cell. **8:**1243-54.

17. **Bauer, U. M., S. Daujat, S. J. Nielsen, K. Nightingale, and T. Kouzarides** 2002. Methylation at arginine 17 of histone H3 is linked to gene activation EMBO Rep. **3:**39-44.

18. **Beisel, C., A. Imhof, J. Greene, E. Kremmer, and F. Sauer** 2002. Histone methylation by the Drosophila epigenetic transcriptional regulator Ash1 Nature. **419:**857-62.

19. **Berger, S. L., B. Pina, N. Silverman, G. A. Marcus, J. Agapite, J. L. Regier, S. J. Triezenberg, and L. Guarente** 1992. Genetic isolation of ADA2: a potential transcriptional adaptor required for function of certain acidic activation domains Cell. **70:**251-65.

20. **Bhat, R. A., M. Riehl, G. Santandrea, R. Velasco, S. Slocombe, G. Donn, H. H. Steinbiss, R. D. Thompson, and H. A. Becker** 2003. Alteration of GCN5 levels in maize reveals dynamic responses to manipulating histone acetylation Plant J. **33:**455-69.

21. **Blanco, J. C., S. Minucci, J. Lu, X. J. Yang, K. K. Walker, H. Chen, R. M. Evans, Y. Nakatani, and K. Ozato** 1998. The histone acetylase PCAF is a nuclear receptor coactivator Genes Dev. **12:**1638-51.

22. **Blau, J., H. Xiao, S. McCracken, P. O'Hare, J. Greenblatt, and D. Bentley** 1996. Three functional classes of transcriptional activation domain Mol Cell Biol. **16:**2044-55.

23. **Bordoli, L., M. Netsch, U. Luthi, W. Lutz, and R. Eckner** 2001. Plant orthologs of p300/CBP: conservation of a core domain in metazoan p300/CBP acetyltransferase-related proteins Nucleic Acids Res. **29:**589-97.

24. **Boyer, L. A., M. R. Langer, K. A. Crowley, S. Tan, J. M. Denu, and C. L. Peterson** 2002. Essential role for the SANT domain in the functioning of multiple chromatin remodeling enzymes Mol Cell. **10:**935-42.

25. **Brand, M., J. G. Moggs, M. Oulad-Abdelghani, F. Lejeune, F. J. Dilworth, J. Stevenin, G. Almouzni, and L. Tora** 2001. UV-damaged DNA-binding protein in the TFTC complex links DNA damage recognition to nucleosome acetylation Embo J. **20:**3187-96.

26. **Briggs, S. D., T. Xiao, Z. W. Sun, J. A. Caldwell, J. Shabanowitz, D. F. Hunt, C. D. Allis, and B. D. Strahl** 2002. Gene silencing: trans-histone regulatory pathway in chromatin Nature. **418:**498.

27. **Brown, C. E., L. Howe, K. Sousa, S. C. Alley, M. J. Carrozza, S. Tan, and J. L. Workman** 2001. Recruitment of HAT complexes by direct activator interactions with the ATM-related Tra1 subunit Science. **292:**2333-7.

28. **Brown, S. A., C. S. Weirich, E. M. Newton, and R. E. Kingston** 1998. Transcriptional activation domains stimulate initiation and elongation at different times and via different residues Embo J. **17:**3146-54.

29. **Brownell, J. E., J. Zhou, T. Ranalli, R. Kobayashi, D. G. Edmondson, S. Y. Roth, and C. D. Allis** 1996. Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation Cell. **84:**843-51.

30. **Bryant, G. O., and M. Ptashne** 2003. Independent recruitment in vivo by Gal4 of two complexes required for transcription Mol Cell. **11:**1301-9.

31. **Bryk, M., S. D. Briggs, B. D. Strahl, M. J. Curcio, C. D. Allis, and F. Winston** 2002. Evidence that Set1, a factor required for methylation of histone H3, regulates rDNA silencing in S. cerevisiae by a Sir2-independent mechanism Curr Biol. **12:**165-70.

32. **Brzeski, J., W. Podstolski, K. Olczak, and A. Jerzmanowski** 1999. Identification and analysis of the Arabidopsis thaliana BSH gene, a member of the SNF5 gene family Nucleic Acids Res. **27:**2393-9.

33. **Bulger, M., T. Sawado, D. Schubeler, and M. Groudine** 2002. ChIPs of the beta-globin locus: unraveling gene regulation within an active domain Curr Opin Genet Dev. **12:**170-7.

34. **Burley, S. K., and R. G. Roeder** 1996. Biochemistry and structural biology of transcription factor IID (TFIID) Annu Rev Biochem. **65:**769-99.

35. **Calvo, D., M. Victor, F. Gay, G. Sui, M. P. Luke, P. Dufourcq, G. Wen, M. Maduro, J. Rothman, and Y. Shi** 2001. A POP-1 repressor complex restricts inappropriate cell type-specific gene transcription during Caenorhabditis elegans embryogenesis Embo J. **20:**7197-208.

36. **Candau, R., and S. L. Berger** 1996. Structural and functional analysis of yeast putative adaptors. Evidence for an adaptor complex in vivo J Biol Chem. **271:**5237-45.

37. **Candau, R., P. A. Moore, L. Wang, N. Barlev, C. Y. Ying, C. A. Rosen, and S. L. Berger** 1996. Identification of human proteins functionally conserved with the yeast putative adaptors ADA2 and GCN5 Mol Cell Biol. **16:**593-602.

38. **Candau, R., J. X. Zhou, C. D. Allis, and S. L. Berger** 1997. Histone acetyltransferase activity and interaction with ADA2 are critical for GCN5 function in vivo Embo J. **16:**555-65.

39. **Carey, M.** 1998. The enhanceosome and transcriptional synergy Cell. **92:**5-8.

40. **Carlson, M.** 1997. Genetics of transcriptional regulation in yeast: connections to the RNA polymerase II CTD Annu Rev Cell Dev Biol. **13:**1-23.

41. **Chan, H. M., and N. B. La Thangue** 2001. p300/CBP proteins: HATs for transcriptional bridges and scaffolds J Cell Sci. **114:**2363-73.

42. **Chang, C., and J. A. Shockey** 1999. The ethylene-response pathway: signal perception to gene regulation Curr Opin Plant Biol. **2:**352-8.

43. **Chatterjee, S., and K. Struhl** 1995. Connecting a promoter-bound protein to TBP bypasses the need for a transcriptional activation domain Nature. **374:**820-2.

44. **Chen, D., H. Ma, H. Hong, S. S. Koh, S. M. Huang, B. T. Schurter, D. W. Aswad, and M. R. Stallcup** 1999. Regulation of transcription by a protein methyltransferase Science. **284:**2174-7.

45. **Cheung, P., K. G. Tanner, W. L. Cheung, P. Sassone-Corsi, J. M. Denu, and C. D. Allis** 2000. Synergistic coupling of histone H3 phosphorylation and acetylation in response to epidermal growth factor stimulation Mol Cell. **5:**905-15.

46. **Cheung, W. L., K. Ajiro, K. Samejima, M. Kloc, P. Cheung, C. A. Mizzen, A. Beeser, L. D. Etkin, J. Chernoff, W. C. Earnshaw, and C. D. Allis** 2003. Apoptotic phosphorylation of histone H2B is mediated by mammalian sterile twenty kinase Cell. **113:**507-17.

47. **Chevillard-Briet, M., D. Trouche, and L. Vandel** 2002. Control of CBP co-activating activity by arginine methylation Embo J. **21:**5457-66.

48. **Chinenov, Y.** 2002. A second catalytic domain in the Elp3 histone acetyltransferases: a candidate for histone demethylase activity? Trends Biochem Sci. **27:**115-7.

49. **Chua, Y. L., A. P. Brown, and J. C. Gray** 2001. Targeted histone acetylation and altered nuclease accessibility over short regions of the pea plastocyanin gene Plant Cell. **13:**599-612.

50. **Clements, A., J. R. Rojas, R. C. Trievel, L. Wang, S. L. Berger, and R. Marmorstein** 1999. Crystal structure of the histone acetyltransferase domain of the human PCAF transcriptional regulator bound to coenzyme A Embo J. **18:**3521-32.

51. **Cosma, M. P., S. Panizza, and K. Nasmyth** 2001. Cdk1 triggers association of RNA polymerase to cell cycle promoters only after recruitment of the mediator by SBF Mol Cell. **7:**1213-20.

52. **Cosma, M. P., T. Tanaka, and K. Nasmyth** 1999. Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle- and developmentally regulated promoter Cell. **97:**299-311.

53. **Daujat, S., U. M. Bauer, V. Shah, B. Turner, S. Berger, and T. Kouzarides** 2002. Crosstalk between CARM1 Methylation and CBP Acetylation on Histone H3 Curr Biol. **12:**2090-7.

54. **Dechend, R., F. Hirano, K. Lehmann, V. Heissmeyer, S. Ansieau, F. G. Wulczyn, C. Scheidereit, and A. Leutz** 1999. The Bcl-3 oncoprotein acts as a bridging factor between NF-kappaB/Rel and nuclear co-regulators Oncogene. **18:**3316-23.

55. **Deng, L., C. de la Fuente, P. Fu, L. Wang, R. Donnelly, J. D. Wade, P. Lambert, H. Li, C. G. Lee, and F. Kashanchi** 2000. Acetylation of HIV-1 Tat by CBP/P300 increases transcription of integrated HIV-1 genome and enhances binding to core histones Virology. **277:**278-95.

56. **Deng, Z., C. J. Chen, M. Chamberlin, F. Lu, G. A. Blobel, D. Speicher, L. A. Cirillo, K. S. Zaret, and P. M. Lieberman** 2003. The CBP bromodomain and nucleosome targeting are required for Zta-directed nucleosome acetylation and transcription activation Mol Cell Biol. **23:**2633-44.

57. **Dhalluin, C., J. E. Carlson, L. Zeng, C. He, A. K. Aggarwal, and M. M. Zhou** 1999. Structure and ligand of a histone acetyltransferase bromodomain Nature. **399:**491-6.

58. **Dorr, A., V. Kiermer, A. Pedal, H. R. Rackwitz, P. Henklein, U. Schubert, M. M. Zhou, E. Verdin, and M. Ott** 2002. Transcriptional synergy between Tat and PCAF is dependent on the binding of acetylated Tat to the PCAF bromodomain Embo J. **21:**2715-23.

59. **Dover, J., J. Schneider, M. A. Tawiah-Boateng, A. Wood, K. Dean, M. Johnston, and A. Shilatifard** 2002. Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6 J Biol Chem. **277:**28368-71.

60. **Dunphy, E. L., T. Johnson, S. S. Auerbach, and E. H. Wang** 2000. Requirement for TAF(II)250 acetyltransferase activity in cell cycle progression Mol Cell Biol. **20:**1134-9.

61. **Eberharter, A., D. E. Sterner, D. Schieltz, A. Hassan, J. R. Yates, 3rd, S. L. Berger, and J. L. Workman** 1999. The ADA complex is a distinct histone acetyltransferase complex in Saccharomyces cerevisiae Mol Cell Biol. **19:**6621-31.

62. **Eissenberg, J. C., and S. C. Elgin** 2000. The HP1 protein family: getting a grip on chromatin Curr Opin Genet Dev. **10:**204-10.

63. **Farrell, S., N. Simkovich, Y. Wu, A. Barberis, and M. Ptashne** 1996. Gene activation by recruitment of the RNA polymerase II holoenzyme Genes Dev. **10:**2359-67.

64. **Fassler, J. S., and F. Winston** 1988. Isolation and analysis of a novel class of suppressor of Ty insertion mutations in Saccharomyces cerevisiae Genetics. **118:**203-12.

65. **Fischle, W., Y. Wang, and C. D. Allis** 2003. Histone and chromatin cross-talk Curr Opin Cell Biol. **15:**172-83.

66. **Freiman, R. N., and R. Tjian** 2003. Regulating the regulators: lysine modifications make their mark Cell. **112:**11-7.

67. **Garber, M. E., and K. A. Jones** 1999. HIV-1 Tat: coping with negative elongation factors Curr Opin Immunol. **11:**460-5.

68. **Gaudreau, L., A. Schmid, D. Blaschke, M. Ptashne, and W. Horz** 1997. RNA polymerase II holoenzyme recruitment is sufficient to remodel chromatin at the yeast PHO5 promoter Cell. **89:**55-62.

69. **Gay, F., D. Calvo, M. C. Lo, J. Ceron, M. Maduro, R. Lin, and Y. Shi** 2003. Acetylation regulates subcellular localization of the Wnt signaling nuclear effector POP-1 Genes Dev. **17:**717-22.

70. **Georgakopoulos, T., N. Gounalaki, and G. Thireos** 1995. Genetic evidence for the interaction of the yeast transcriptional co-activator proteins GCN5 and ADA2 Mol Gen Genet. **246:**723-8.

71. **Gonzalez-Couto, E., N. Klages, and M. Strubin** 1997. Synergistic and promoter-selective activation of transcription by recruitment of transcription factors TFIID and TFIIB Proc Natl Acad Sci U S A. **94:**8036-41.

72. **Goodman, R. H., and S. Smolik** 2000. CBP/p300 in cell growth, transformation, and development Genes Dev. **14:**1553-77.

73. **Grant, P. A., and S. L. Berger** 1999. Histone acetyltransferase complexes Semin Cell Dev Biol. **10:**169-77.

74. **Grant, P. A., L. Duggan, J. Cote, S. M. Roberts, J. E. Brownell, R. Candau, R. Ohba, T. Owen-Hughes, C. D. Allis, F. Winston, S. L. Berger, and J. L. Workman**

1997. Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex Genes Dev. **11:**1640-50.

75. **Grunstein, M.** 1997. Histone acetylation in chromatin structure and transcription Nature. **389:**349-52.

76. **Gu, W., and R. G. Roeder** 1997. Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain Cell. **90:**595-606.

77. **Gu, Z., L. M. Steinmetz, X. Gu, C. Scharfe, R. W. Davis, and W. H. Li** 2003. Role of duplicate genes in genetic robustness against null mutations Nature. **421:**63-6.

78. **Hall, D. B., and K. Struhl** 2002. The VP16 activation domain interacts with multiple transcriptional components as determined by protein-protein cross-linking in vivo J Biol Chem. **277:**46043-50.

79. **Hampsey, M.** 1998. Molecular genetics of the RNA polymerase II general transcriptional machinery Microbiol Mol Biol Rev. **62:**465-503.

80. **Han, M., and M. Grunstein** 1988. Nucleosome loss activates yeast downstream promoters in vivo Cell. **55:**1137-45.

81. **Hassan, A. H., K. E. Neely, M. Vignali, J. C. Reese, and J. L. Workman** 2001. Promoter targeting of chromatin-modifying complexes Front Biosci. **6:**D1054-64.

82. **Hassan, A. H., K. E. Neely, and J. L. Workman** 2001. Histone acetyltransferase complexes stabilize swi/snf binding to promoter nucleosomes Cell. **104:**817-27.

83. **Hebbes, T. R., A. W. Thorne, and C. Crane-Robinson** 1988. A direct link between core histone acetylation and transcriptionally active chromatin Embo J. **7:**1395-402.

84. **Holstege, F. C., E. G. Jennings, J. J. Wyrick, T. I. Lee, C. J. Hengartner, M. R. Green, T. R. Golub, E. S. Lander, and R. A. Young** 1998. Dissecting the regulatory circuitry of a eukaryotic genome Cell. **95:**717-28.

85. **Horn, P. J., and C. L. Peterson** 2001. The bromodomain: a regulator of ATP-dependent chromatin remodeling? Front Biosci. **6:**D1019-23.

86. **Ikura, T., V. V. Ogryzko, M. Grigoriev, R. Groisman, J. Wang, M. Horikoshi, R. Scully, J. Qin, and Y. Nakatani** 2000. Involvement of the TIP60 histone acetylase complex in DNA repair and apoptosis Cell. **102:**463-73.

87. **Ito, M., C. X. Yuan, S. Malik, W. Gu, J. D. Fondell, S. Yamamura, Z. Y. Fu, X. Zhang, J. Qin, and R. G. Roeder** 1999. Identity between TRAP and SMCC

complexes indicates novel pathways for the function of nuclear receptors and diverse mammalian activators Mol Cell. **3:**361-70.

88. **Jacobs, S. A., and S. Khorasanizadeh** 2002. Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail Science. **295:**2080-3.

89. **Jacobson, R. H., A. G. Ladurner, D. S. King, and R. Tjian** 2000. Structure and function of a human TAFII250 double bromodomain module Science. **288:**1422-5.

90. **Jaglo-Ottosen, K. R., S. J. Gilmour, D. G. Zarka, O. Schabenberger, and M. F. Thomashow** 1998. Arabidopsis CBF1 overexpression induces COR genes and enhances freezing tolerance Science. **280:**104-6.

91. **Jeang, K. T., H. Xiao, and E. A. Rich** 1999. Multifaceted activities of the HIV-1 transactivator of transcription, Tat J Biol Chem. **274:**28837-40.

92. **Jeanmougin, F., J. M. Wurtz, B. Le Douarin, P. Chambon, and R. Losson** 1997. The bromodomain revisited Trends Biochem Sci. **22:**151-3.

93. **Jenuwein, T.** 2001. Re-SET-ting heterochromatin by histone methyltransferases Trends Cell Biol. **11:**266-73.

94. **Jenuwein, T., and C. D. Allis** 2001. Translating the histone code Science. **293:**1074-80.

95. **Kadam, S., G. S. McAlpine, M. L. Phelan, R. E. Kingston, K. A. Jones, and B. M. Emerson** 2000. Functional selectivity of recombinant mammalian SWI/SNF subunits Genes Dev. **14:**2441-51.

96. **Karn, J.** 1999. Tackling Tat J Mol Biol. **293:**235-54.

97. **Keaveney, M., and K. Struhl** 1998. Activator-mediated recruitment of the RNA polymerase II machinery is the predominant mechanism for transcriptional activation in yeast Mol Cell. **1:**917-24.

98. **Kenzior, A. L., and W. R. Folk** 1998. AtMSI4 and RbAp48 WD-40 repeat proteins bind metal ions FEBS Lett. **440:**425-9.

99. **Khanna, K. K., M. F. Lavin, S. P. Jackson, and T. D. Mulhern** 2001. ATM, a central controller of cellular responses to DNA damage Cell Death Differ. **8:**1052-65.

100. **Kim, J., K. Harter, and A. Theologis** 1997. Protein-protein interactions among the Aux/IAA proteins Proc Natl Acad Sci U S A. **94:**11786-91.

101. **Kim, J. H., W. S. Lane, and D. Reinberg** 2002. Human Elongator facilitates RNA polymerase II transcription through chromatin Proc Natl Acad Sci U S A. **99:**1241-6.

102. **Kim, Y. J., S. Bjorklund, Y. Li, M. H. Sayre, and R. D. Kornberg** 1994. A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II Cell. **77:**599-608.

103. **Klages, N., and M. Strubin** 1995. Stimulation of RNA polymerase II transcription initiation by recruitment of TBP in vivo Nature. **374:**822-3.

104. **Kokubo, T., M. J. Swanson, J. I. Nishikawa, A. G. Hinnebusch, and Y. Nakatani** 1998. The yeast TAF145 inhibitory domain and TFIIA competitively bind to TATA-binding protein Mol Cell Biol. **18:**1003-12.

105. **Koleske, A. J., and R. A. Young** 1994. An RNA polymerase II holoenzyme responsive to activators Nature. **368:**466-9.

106. **Kouzarides, T.** 2000. Acetylation: a regulatory modification to rival phosphorylation? Embo J. **19:**1176-9.

107. **Kouzarides, T.** 2002. Histone methylation in transcriptional control Curr Opin Genet Dev. **12:**198-209.

108. **Kristjuhan, A., J. Walker, N. Suka, M. Grunstein, D. Roberts, B. R. Cairns, and J. Q. Svejstrup** 2002. Transcriptional inhibition of genes with severe histone h3 hypoacetylation in the coding region Mol Cell. **10:**925-33.

109. **Krumm, A., L. B. Hickey, and M. Groudine** 1995. Promoter-proximal pausing of RNA polymerase II defines a general rate-limiting step after transcription initiation Genes Dev. **9:**559-72.

110. **Kulesza, C. A., H. A. Van Buskirk, M. D. Cole, J. C. Reese, M. M. Smith, and D. A. Engel** 2002. Adenovirus E1A requires the yeast SAGA histone acetyltransferase complex and associates with SAGA components Gcn5 and Tra1 Oncogene. **21:**1411-22.

111. **Kumar, A., S. Agarwal, J. A. Heyman, S. Matson, M. Heidtman, S. Piccirillo, L. Umansky, A. Drawid, R. Jansen, Y. Liu, K. H. Cheung, P. Miller, M. Gerstein, G. S. Roeder, and M. Snyder** 2002. Subcellular localization of the yeast proteome Genes Dev. **16:**707-19.

112. **Kuo, M. H., J. E. Brownell, R. E. Sobel, T. A. Ranalli, R. G. Cook, D. G. Edmondson, S. Y. Roth, and C. D. Allis** 1996. Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines Nature. **383:**269-72.

113. **Kuo, M. H., E. vom Baur, K. Struhl, and C. D. Allis** 2000. Gcn4 activator

targets Gcn5 histone acetyltransferase to specific promoters independently of transcription Mol Cell. **6:**1309-20.

114. **Kuo, M. H., J. Zhou, P. Jambeck, M. E. Churchill, and C. D. Allis** 1998. Histone acetyltransferase activity of yeast Gcn5p is required for the activation of target genes in vivo Genes Dev. **12:**627-39.

115. **Kusch, T., S. Guelman, S. M. Abmayr, and J. L. Workman** 2003. Two Drosophila Ada2 homologues function in different multiprotein complexes Mol Cell Biol. **23:**3305-19.

116. **Kuzmichev, A., K. Nishioka, H. Erdjument-Bromage, P. Tempst, and D. Reinberg** 2002. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein Genes Dev. **16:**2893-905.

117. **Lachner, M., and T. Jenuwein** 2002. The many faces of histone lysine methylation Curr Opin Cell Biol. **14:**286-98.

118. **Lachner, M., D. O'Carroll, S. Rea, K. Mechtler, and T. Jenuwein** 2001. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins Nature. **410:**116-20.

119. **Larschan, E., and F. Winston** 2001. The S. cerevisiae SAGA complex functions in vivo as a coactivator for transcriptional activation by Gal4 Genes Dev. **15:**1946-56.

120. **Lee, K. M., and J. J. Hayes** 1997. The N-terminal tail of histone H2A binds to two distinct sites within the nucleosome core Proc Natl Acad Sci U S A. **94:**8959-64.

121. **Lee, T. I., and R. A. Young** 2000. Transcription of eukaryotic protein-coding genes Annu Rev Genet. **34:**77-137.

122. **Legube, G., L. K. Linares, C. Lemercier, M. Scheffner, S. Khochbin, and D. Trouche** 2002. Tip60 is targeted to proteasome-mediated degradation by Mdm2 and accumulates after UV irradiation Embo J. **21:**1704-12.

123. **Lenfant, F., R. K. Mann, B. Thomsen, X. Ling, and M. Grunstein** 1996. All four core histone N-termini contain sequences required for the repression of basal transcription in yeast Embo J. **15:**3974-85.

124. **Li, G., T. C. Hall, and R. Holmes-Davis** 2002. Plant chromatin: development and gene control Bioessays. **24:**234-43.

125. **Li, J., Q. Lin, H. G. Yoon, Z. Q. Huang, B. D. Strahl, C. D. Allis, and J. Wong** 2002. Involvement of histone methylation and phosphorylation in regulation of transcription by thyroid hormone receptor Mol Cell Biol. **22:**5688-97.

126.
Correla
chicker

127.
2001.
neighb

128.
Kay, Y
compl
TBP C

129.
and K
EREB
droug
Plant

130.
Shiek
the hi

131.
Mar
funct
Cell.

132.
Med

133.
initia
Cell.

134.
reco
chro

135.
199
389:

136.
regu

126. **Litt, M. D., M. Simpson, M. Gaszner, C. D. Allis, and G. Felsenfeld** 2001. Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus Science. **293:**2453-5.

127. **Litt, M. D., M. Simpson, F. Recillas-Targa, M. N. Prioleau, and G. Felsenfeld** 2001. Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci Embo J. **20:**2224-35.

128. **Liu, D., R. Ishima, K. I. Tong, S. Bagby, T. Kokubo, D. R. Muhandiram, L. E. Kay, Y. Nakatani, and M. Ikura** 1998. Solution structure of a TBP-TAF(II)230 complex: protein mimicry of the minor groove surface of the TATA box unwound by TBP Cell. **94:**573-83.

129. **Liu, Q., M. Kasuga, Y. Sakuma, H. Abe, S. Miura, K. Yamaguchi-Shinozaki, and K. Shinozaki** 1998. Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in Arabidopsis Plant Cell. **10:**1391-406.

130. **Lo, W. S., L. Duggan, N. C. Tolga, Emre, R. Belotserkovskya, W. S. Lane, R. Shiekhattar, and S. L. Berger** 2001. Snf1--a histone kinase that works in concert with the histone acetyltransferase Gcn5 to regulate transcription Science. **293:**1142-6.

131. **Lo, W. S., R. C. Trievel, J. R. Rojas, L. Duggan, J. Y. Hsu, C. D. Allis, R. Marmorstein, and S. L. Berger** 2000. Phosphorylation of serine 10 in histone H3 is functionally linked in vitro and in vivo to Gcn5-mediated acetylation at lysine 14 Mol Cell. **5:**917-26.

132. **Lorch, Y., J. Beve, C. M. Gustafsson, L. C. Myers, and R. D. Kornberg** 2000. Mediator-nucleosome interaction Mol Cell. **6:**197-201.

133. **Lorch, Y., J. W. LaPointe, and R. D. Kornberg** 1987. Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones Cell. **49:**203-10.

134. **Loyola, A., G. LeRoy, Y. H. Wang, and D. Reinberg** 2001. Reconstitution of recombinant chromatin establishes a requirement for histone-tail modifications during chromatin assembly and transcription Genes Dev. **15:**2837-51.

135. **Luger, K., A. W. Mader, R. K. Richmond, D. F. Sargent, and T. J. Richmond** 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution Nature. **389:**251-60.

136. **Lusser, A.** 2002. Acetylated, methylated, remodeled: chromatin states for gene regulation Curr Opin Plant Biol. **5:**437-43.

137. **Madison, D. L., P. Yaciuk, R. P. Kwok, and J. R. Lundblad** 2002. Acetylation of the adenovirus-transforming protein E1A determines nuclear localization by disrupting association with importin-alpha J Biol Chem. **277:**38755-63.

138. **Majello, B., and G. Napolitano** 2001. Control of RNA polymerase II activity by dedicated CTD kinases and phosphatases Front Biosci. **6:**D1358-68.

139. **Maldonado, E., R. Shiekhattar, M. Sheldon, H. Cho, R. Drapkin, P. Rickert, E. Lees, C. W. Anderson, S. Linn, and D. Reinberg** 1996. A human RNA polymerase II complex associated with SRB and DNA-repair proteins Nature. **381:**86-9.

140. **Malik, S., and R. G. Roeder** 2000. Transcriptional regulation through Mediator-like coactivators in yeast and metazoan cells Trends Biochem Sci. **25:**277-83.

141. **Marcus, G. A., N. Silverman, S. L. Berger, J. Horiuchi, and L. Guarente** 1994. Functional similarity and physical association between GCN5 and ADA2: putative transcriptional adaptors Embo J. **13:**4807-15.

142. **Marmorstein, R.** 2001. Structure of histone acetyltransferases J Mol Biol. **311:**433-44.

143. **Martinez, E., V. B. Palhan, A. Tjernberg, E. S. Lymar, A. M. Gamper, T. K. Kundu, B. T. Chait, and R. G. Roeder** 2001. Human STAGA complex is a chromatin-acetylating transcription coactivator that interacts with pre-mRNA splicing and DNA damage-binding factors in vivo Mol Cell Biol. **21:**6782-95.

144. **Memelink, J., R. Verpoorte, and J. W. Kijne** 2001. ORCAnization of jasmonate-responsive gene expression in alkaloid metabolism Trends Plant Sci. **6:**212-9.

145. **Mencia, M., Z. Moqtaderi, J. V. Geisberg, L. Kuras, and K. Struhl** 2002. Activator-specific recruitment of TFIID and regulation of ribosomal protein genes in yeast Mol Cell. **9:**823-33.

146. **Milne, T. A., S. D. Briggs, H. W. Brock, M. E. Martin, D. Gibbs, C. D. Allis, and J. L. Hess** 2002. MLL targets SET domain methyltransferase activity to Hox gene promoters Mol Cell. **10:**1107-17.

147. **Mujtaba, S., Y. He, L. Zeng, A. Farooq, J. E. Carlson, M. Ott, E. Verdin, and M. M. Zhou** 2002. Structural basis of lysine-acetylated HIV-1 Tat recognition by PCAF bromodomain Mol Cell. **9:**575-86.

148. **Muratoglu, S., S. Georgieva, G. Papai, E. Scheer, I. Enunlu, O. Komonyi, I. Cserpan, L. Lebedeva, E. Nabirochkina, A. Udvardy, L. Tora, and I. Boros** 2003. Two different Drosophila ADA2 homologues are present in distinct GCN5 histone acetyltransferase-containing complexes Mol Cell Biol. **23:**306-21.

149. **Mutskov, V. J., C. M. Farrell, P. A. Wade, A. P. Wolffe, and G. Felsenfeld** 2002. The barrier function of an insulator couples high histone acetylation levels with specific protection of promoter DNA from methylation Genes Dev. **16:**1540-54.

150. **Myers, L. C., C. M. Gustafsson, D. A. Bushnell, M. Lui, H. Erdjument-Bromage, P. Tempst, and R. D. Kornberg** 1998. The Med proteins of yeast and their function through the RNA polymerase II carboxy-terminal domain Genes Dev. **12:**45-54.

151. **Myers, L. C., and R. D. Kornberg** 2000. Mediator of transcriptional regulation Annu Rev Biochem. **69:**729-49.

152. **Naar, A. M., P. A. Beaurang, S. Zhou, S. Abraham, W. Solomon, and R. Tjian** 1999. Composite co-activator ARC mediates chromatin-directed transcriptional activation Nature. **398:**828-32.

153. **Naar, A. M., B. D. Lemon, and R. Tjian** 2001. Transcriptional coactivator complexes Annu Rev Biochem. **70:**475-501.

154. **Natarajan, K., B. M. Jackson, H. Zhou, F. Winston, and A. G. Hinnebusch** 1999. Transcriptional activation by Gcn4p involves independent interactions with the SWI/SNF complex and the SRB/mediator Mol Cell. **4:**657-64.

155. **Neely, K. E., A. H. Hassan, A. E. Wallberg, D. J. Steger, B. R. Cairns, A. P. Wright, and J. L. Workman** 1999. Activation domain-mediated targeting of the SWI/SNF complex to promoters stimulates transcription from nucleosome arrays Mol Cell. **4:**649-55.

156. **Neish, A. S., S. F. Anderson, B. P. Schlegel, W. Wei, and J. D. Parvin** 1998. Factors associated with the mammalian RNA polymerase II holoenzyme Nucleic Acids Res. **26:**847-53.

157. **Neuwald, A. F., and D. Landsman** 1997. GCN5-related histone N-acetyltransferases belong to a diverse superfamily that includes the yeast SPT10 protein Trends Biochem Sci. **22:**154-5.

158. **Ng, H. H., D. N. Ciccone, K. B. Morshead, M. A. Oettinger, and K. Struhl** 2003. Lysine-79 of histone H3 is hypomethylated at silenced loci in yeast and mammalian cells: a potential mechanism for position-effect variegation Proc Natl Acad Sci U S A. **100:**1820-5.

159. **Ng, H. H., Q. Feng, H. Wang, H. Erdjument-Bromage, P. Tempst, Y. Zhang, and K. Struhl** 2002. Lysine methylation within the globular domain of histone H3 by Dot1 is important for telomeric silencing and Sir protein association Genes Dev. **16:**1518-27.

160. N
H2B by
J Biol C

161. N
O'Carr
2001. R

162. N
Drosopl
TATA
94:85-9

163.
regulat

164.
by Set1
Acad S

165.
H. Ho
acetyla

166.
bromo
histon

167.
facto

168.
Cell

169
Ov
fac
Ce

17
re

17
TA
60

160. **Ng, H. H., R. M. Xu, Y. Zhang, and K. Struhl** 2002. Ubiquitination of histone H2B by Rad6 is required for efficient Dot1-mediated methylation of histone H3 lysine 79 J Biol Chem. **277:**34655-7.

161. **Nielsen, S. J., R. Schneider, U. M. Bauer, A. J. Bannister, A. Morrison, D. O'Carroll, R. Firestein, M. Cleary, T. Jenuwein, R. E. Herrera, and T. Kouzarides** 2001. Rb targets histone H3 methylation and HP1 to promoters Nature. **412:**561-5.

162. **Nishikawa, J., T. Kokubo, M. Horikoshi, R. G. Roeder, and Y. Nakatani** 1997. Drosophila TAF(II)230 and the transcriptional activator VP16 bind competitively to the TATA box-binding domain of the TATA box-binding protein Proc Natl Acad Sci U S A. **94:**85-90.

163. **Nishioka, K., and D. Reinberg** 2001. Transcription. Switching partners in a regulatory tango Science. **294:**2497-8.

164. **Noma, K., and S. I. Grewal** 2002. Histone H3 lysine 4 methylation is mediated by Set1 and promotes maintenance of active chromatin states in fission yeast Proc Natl Acad Sci U S A. **99:**16438-45.

165. **Ogryzko, V. V., T. Kotani, X. Zhang, R. L. Schiltz, T. Howard, X. J. Yang, B. H. Howard, J. Qin, and Y. Nakatani** 1998. Histone-like TAFs within the PCAF histone acetylase complex Cell. **94:**35-44.

166. **Ornaghi, P., P. Ballario, A. M. Lena, A. Gonzalez, and P. Filetici** 1999. The bromodomain of Gcn5p interacts in vitro with specific residues in the N terminus of histone H4 J Mol Biol. **287:**1-7.

167. **Orphanides, G., T. Lagrange, and D. Reinberg** 1996. The general transcription factors of RNA polymerase II Genes Dev. **10:**2657-83.

168. **Orphanides, G., and D. Reinberg** 2002. A unified theory of gene expression Cell. **108:**439-51.

169. **Park, J. M., C. J. Park, S. B. Lee, B. K. Ham, R. Shin, and K. H. Paek** 2001. Overexpression of the tobacco Tsi1 gene encoding an EREBP/AP2-type transcription factor enhances resistance against pathogen attack and osmotic stress in tobacco Plant Cell. **13:**1035-46.

170. **Peterson, C. L., and J. L. Workman** 2000. Promoter targeting and chromatin remodeling by the SWI/SNF complex Curr Opin Genet Dev. **10:**187-92.

171. **Pham, A. D., and F. Sauer** 2000. Ubiquitin-activating/conjugating activity of TAFII250, a mediator of activation of gene expression in Drosophila Science. **289:**2357-60.

172. **Pina, B., S. Berger, G. A. Marcus, N. Silverman, J. Agapite, and L. Guarente** 1993. ADA3: a gene, identified by resistance to GAL4-VP16, with properties similar to and different from those of ADA2 Mol Cell Biol. **13:**5981-9.

173. **Protacio, R. U., G. Li, P. T. Lowary, and J. Widom** 2000. Effects of histone tail domains on the rate of transcriptional elongation through a nucleosome Mol Cell Biol. **20:**8866-78.

174. **Ptashne, M., and A. Gann** 1997. Transcriptional activation by recruitment Nature. **386:**569-77.

175. **Pugh, B. F., and R. Tjian** 1992. Diverse transcriptional functions of the multisubunit eukaryotic TFIID complex J Biol Chem. **267:**679-82.

176. **Rachez, C., and L. P. Freedman** 2001. Mediator complexes and transcription Curr Opin Cell Biol. **13:**274-80.

177. **Rachez, C., B. D. Lemon, Z. Suldan, V. Bromleigh, M. Gamble, A. M. Naar, H. Erdjument-Bromage, P. Tempst, and L. P. Freedman** 1999. Ligand-dependent transcription activation by nuclear receptors requires the DRIP complex Nature. **398:**824-8.

178. **Rea, S., F. Eisenhaber, D. O'Carroll, B. D. Strahl, Z. W. Sun, M. Schmid, S. Opravil, K. Mechtler, C. P. Ponting, C. D. Allis, and T. Jenuwein** 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases Nature. **406:**593-9.

179. **Reinke, H., P. D. Gregory, and W. Horz** 2001. A transient histone hyperacetylation signal marks nucleosomes for remodeling at the PHO8 promoter in vivo Mol Cell. **7:**529-38.

180. **Rezai-Zadeh, N., X. Zhang, F. Namour, G. Fejer, Y. D. Wen, Y. L. Yao, I. Gyory, K. Wright, and E. Seto** 2003. Targeted recruitment of a histone H4-specific methyltransferase by the transcription factor YY1 Genes Dev. **17:**1019-29.

181. **Riechmann, J. L., J. Heard, G. Martin, L. Reuber, C. Jiang, J. Keddie, L. Adam, O. Pineda, O. J. Ratcliffe, R. R. Samaha, R. Creelman, M. Pilgrim, P. Broun, J. Z. Zhang, D. Ghandehari, B. K. Sherman, and G. Yu** 2000. Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes Science. **290:**2105-10.

182. **Riechmann, J. L., and E. M. Meyerowitz** 1998. The AP2/EREBP family of plant transcription factors Biol Chem. **379:**633-46.

183. **Riechmann, J. L., and O. J. Ratcliffe** 2000. A genomic perspective on plant transcription factors Curr Opin Plant Biol. **3:**423-34.

184.
RNA p

185.
single

186.
Rev B

187.
cofac
Natur

188.
Yam
Arab
gene

189.
A. H
activ

190.
Nak
coac

191
A. I
Met
Bio

192
A2
275

193
pro
Bio

194
rem

195
deh
patl

184. **Roeder, R. G.** 1996. The role of general initiation factors in transcription by RNA polymerase II Trends Biochem Sci. **21:**327-35.

185. **Roth, S. Y., and C. D. Allis** 1996. Histone acetylation and chromatin assembly: a single escort, multiple dances? Cell. **87:**5-8.

186. **Roth, S. Y., J. M. Denu, and C. D. Allis** 2001. Histone acetyltransferases Annu Rev Biochem. **70:**81-120.

187. **Ryu, S., S. Zhou, A. G. Ladurner, and R. Tjian** 1999. The transcriptional cofactor complex CRSP is required for activity of the enhancer-binding protein Sp1 Nature. **397:**446-50.

188. **Sakuma, Y., Q. Liu, J. G. Dubouzet, H. Abe, K. Shinozaki, and K. Yamaguchi-Shinozaki** 2002. DNA-binding specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in dehydration- and cold-inducible gene expression Biochem Biophys Res Commun. **290:**998-1009.

189. **Sassone-Corsi, P., C. A. Mizzen, P. Cheung, C. Crosio, L. Monaco, S. Jacquot, A. Hanauer, and C. D. Allis** 1999. Requirement of Rsk-2 for epidermal growth factor-activated phosphorylation of histone H3 Science. **285:**886-91.

190. **Schiltz, R. L., C. A. Mizzen, A. Vassilev, R. G. Cook, C. D. Allis, and Y. Nakatani** 1999. Overlapping but distinct patterns of histone acetylation by the human coactivators p300 and PCAF within nucleosomal substrates J Biol Chem. **274:**1189-92.

191. **Schurter, B. T., S. S. Koh, D. Chen, G. J. Bunick, J. M. Harp, B. L. Hanson, A. Henschen-Edman, D. R. Mackay, M. R. Stallcup, and D. W. Aswad** 2001. Methylation of histone H3 by coactivator-associated arginine methyltransferase 1 Biochemistry. **40:**5747-56.

192. **Sendra, R., C. Tse, and J. C. Hansen** 2000. The yeast histone acetyltransferase A2 complex, but not free Gcn5p, binds stably to nucleosomal arrays J Biol Chem. **275:**24928-34.

193. **Sewack, G. F., T. W. Ellis, and U. Hansen** 2001. Binding of TATA binding protein to a naturally positioned nucleosome is facilitated by histone acetylation Mol Cell Biol. **21:**1404-15.

194. **Shen, X., G. Mizuguchi, A. Hamiche, and C. Wu** 2000. A chromatin remodelling complex involved in transcription and DNA processing Nature. **406:**541-4.

195. **Shinozaki, K., and K. Yamaguchi-Shinozaki** 2000. Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways Curr Opin Plant Biol. **3:**217-23.

196. **Sieberer, T., M. T. Hauser, G. J. Seifert, and C. Luschnig** 2003. PROPORZ1, a Putative Arabidopsis Transcriptional Adaptor Protein, Mediates Auxin and Cytokinin Signals in the Control of Cell Proliferation Curr Biol. **13:**837-42.

197. **Singh, K., R. C. Foley, and L. Onate-Sanchez** 2002. Transcription factors in plant defense and stress responses Curr Opin Plant Biol. **5:**430-6.

198. **Spencer, V. A., and J. R. Davie** 1999. Role of covalent modifications of histones in regulating gene expression Gene. **240:**1-12.

199. **Spilianakis, C., J. Papamatheakis, and A. Kretsovali** 2000. Acetylation by PCAF enhances CIITA nuclear accumulation and transactivation of major histocompatibility complex class II genes Mol Cell Biol. **20:**8489-98.

200. **Sterner, D. E., R. Belotserkovskaya, and S. L. Berger** 2002. SALSA, a variant of yeast SAGA, contains truncated Spt7, which correlates with activated transcription Proc Natl Acad Sci U S A. **99:**11622-7.

201. **Sterner, D. E., and S. L. Berger** 2000. Acetylation of histones and transcription-related factors Microbiol Mol Biol Rev. **64:**435-59.

202. **Sterner, D. E., P. A. Grant, S. M. Roberts, L. J. Duggan, R. Belotserkovskaya, L. A. Pacella, F. Winston, J. L. Workman, and S. L. Berger** 1999. Functional organization of the yeast SAGA complex: distinct components involved in structural integrity, nucleosome acetylation, and TATA-binding protein interaction Mol Cell Biol. **19:**86-98.

203. **Sterner, D. E., X. Wang, M. H. Bloom, G. M. Simon, and S. L. Berger** 2002. The SANT domain of Ada2 is required for normal acetylation of histones by the yeast SAGA complex J Biol Chem. **277:**8178-86.

204. **Stockinger, E. J., S. J. Gilmour, and M. F. Thomashow** 1997. Arabidopsis thaliana CBF1 encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit Proc Natl Acad Sci U S A. **94:**1035-40.

205. **Stockinger, E. J., Y. Mao, M. K. Regier, S. J. Triezenberg, and M. F. Thomashow** 2001. Transcriptional adaptor and histone acetyltransferase proteins in Arabidopsis and their interactions with CBF1, a transcriptional activator involved in cold-regulated gene expression Nucleic Acids Res. **29:**1524-33.

206. **Strahl, B. D., and C. D. Allis** 2000. The language of covalent histone modifications Nature. **403:**41-5.

207. **Strahl, B. D., R. Ohba, R. G. Cook, and C. D. Allis** 1999. Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in Tetrahymena Proc Natl Acad Sci U S A. **96:**14967-72.

208. **Struhl, K.** 1999. Fundamentally different logic of gene regulation in eukaryotes and prokaryotes Cell. **98:**1-4.

209. **Struhl, K.** 1998. Histone acetylation and transcriptional regulatory mechanisms Genes Dev. **12:**599-606.

210. **Sun, Z. W., and C. D. Allis** 2002. Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast Nature. **418:**104-8.

211. **Swanson, M. J., H. Qiu, L. Sumibcay, A. Krueger, S. J. Kim, K. Natarajan, S. Yoon, and A. G. Hinnebusch** 2003. A multiplicity of coactivators is required by Gcn4p at individual promoters in vivo Mol Cell Biol. **23:**2800-20.

212. **Taatjes, D. J., A. M. Naar, F. Andel, 3rd, E. Nogales, and R. Tjian** 2002. Structure, function, and activator-induced conformations of the CRSP coactivator Science. **295:**1058-62.

213. **Takechi, S., and T. Nakayama** 1999. Sas3 is a histone acetyltransferase and requires a zinc finger motif Biochem Biophys Res Commun. **266:**405-10.

214. **Tanner, K. G., M. R. Langer, and J. M. Denu** 2000. Kinetic mechanism of human histone acetyltransferase P/CAF Biochemistry. **39:**11961-9.

215. **Tanner, K. G., M. R. Langer, Y. Kim, and J. M. Denu** 2000. Kinetic mechanism of the histone acetyltransferase GCN5 from yeast J Biol Chem. **275:**22048-55.

216. **Tanner, K. G., R. C. Trievel, M. H. Kuo, R. M. Howard, S. L. Berger, C. D. Allis, R. Marmorstein, and J. M. Denu** 1999. Catalytic mechanism and function of invariant glutamic acid 173 from the histone acetyltransferase GCN5 transcriptional coactivator J Biol Chem. **274:**18157-60.

217. **Thomashow, M. F.** 1998. Role of cold-responsive genes in plant freezing tolerance Plant Physiol. **118:**1-8.

218. **Thompson, C. M., and R. A. Young** 1995. General requirement for RNA polymerase II holoenzymes in vivo Proc Natl Acad Sci U S A. **92:**4587-90.

219. **Thomson, S., A. L. Clayton, and L. C. Mahadevan** 2001. Independent dynamic regulation of histone phosphorylation and acetylation during immediate-early gene induction Mol Cell. **8:**1231-41.

220. **Triezenberg, S. J.** 1995. Structure and function of transcriptional activation domains Curr Opin Genet Dev. **5:**190-6.

221. **Tse, C., and J. C. Hansen** 1997. Hybrid trypsinized nucleosomal arrays: identification of multiple functional roles of the H2A/H2B and H3/H4 N-termini in chromatin fiber compaction Biochemistry. **36:**11381-8.

222. **Turner, B. M.** 2002. Cellular memory and the histone code Cell. **111:**285-91.

223. **Turner, B. M.** 1991. Histone acetylation and control of gene expression J Cell Sci. **99:**13-20.

224. **Utley, R. T., K. Ikeda, P. A. Grant, J. Cote, D. J. Steger, A. Eberharter, S. John, and J. L. Workman** 1998. Transcriptional activators direct histone acetyltransferase complexes to nucleosomes Nature. **394:**498-502.

225. **Vandel, L., E. Nicolas, O. Vaute, R. Ferreira, S. Ait-Si-Ali, and D. Trouche** 2001. Transcriptional repression by the retinoblastoma protein through the recruitment of a histone methyltransferase Mol Cell Biol. **21:**6484-94.

226. **Vassilev, A., J. Yamauchi, T. Kotani, C. Prives, M. L. Avantaggiati, J. Qin, and Y. Nakatani** 1998. The 400 kDa subunit of the PCAF histone acetylase complex belongs to the ATM superfamily Mol Cell. **2:**869-75.

227. **Verbsky, M. L., and E. J. Richards** 2001. Chromatin remodeling in plants Curr Opin Plant Biol. **4:**494-500.

228. **Verdone, L., J. Wu, K. van Riper, N. Kacherovsky, M. Vogelauer, E. T. Young, M. Grunstein, E. Di Mauro, and M. Caserta** 2002. Hyperacetylation of chromatin at the ADH2 promoter allows Adr1 to bind in repressed conditions Embo J. **21:**1101-11.

229. **Verreault, A., P. D. Kaufman, R. Kobayashi, and B. Stillman** 1996. Nucleosome assembly by a complex of CAF-1 and acetylated histones H3/H4 Cell. **87:**95-104.

230. **Vignali, M., A. H. Hassan, K. E. Neely, and J. L. Workman** 2000. ATP-dependent chromatin-remodeling complexes Mol Cell Biol. **20:**1899-910.

231. **Vlachonasios, K. E., M. F. Thomashow, and S. J. Triezenberg** 2003. Disruption mutations of ADA2b and GCN5 transcriptional adaptor genes dramatically affect Arabidopsis growth, development, and gene expression Plant Cell. **15:**626-38.

232. **Wang, H., R. Cao, L. Xia, H. Erdjument-Bromage, C. Borchers, P. Tempst, and Y. Zhang** 2001. Purification and functional characterization of a histone H3-lysine 4-specific methyltransferase Mol Cell. **8:**1207-17.

233. **Wang, H., Z. Q. Huang, L. Xia, Q. Feng, H. Erdjument-Bromage, B. D. Strahl, S. D. Briggs, C. D. Allis, J. Wong, P. Tempst, and Y. Zhang** 2001. Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor Science. **293:**853-7.

234. **Wang, L., L. Liu, and S. L. Berger** 1998. Critical residues for histone acetylation by Gcn5, functioning in Ada and SAGA complexes, are also required for transcriptional function in vivo Genes Dev. **12:**640-53.

235. **Wassarman, D. A., and F. Sauer** 2001. TAF(II)250: a transcription toolbox J Cell Sci. **114:**2895-902.

236. **Wei, Y., C. A. Mizzen, R. G. Cook, M. A. Gorovsky, and C. D. Allis** 1998. Phosphorylation of histone H3 at serine 10 is correlated with chromosome condensation during mitosis and meiosis in Tetrahymena Proc Natl Acad Sci U S A. **95:**7480-4.

237. **Wendler, W. M., E. Kremmer, R. Forster, and E. L. Winnacker** 1997. Identification of pirin, a novel highly conserved nuclear protein J Biol Chem. **272:**8482-9.

238. **Wilson, C. J., D. M. Chao, A. N. Imbalzano, G. R. Schnitzler, R. E. Kingston, and R. A. Young** 1996. RNA polymerase II holoenzyme contains SWI/SNF regulators involved in chromatin remodeling Cell. **84:**235-44.

239. **Wilson, K., D. Long, J. Swinburne, and G. Coupland** 1996. A Dissociation insertion causes a semidominant mutation that increases expression of TINY, an Arabidopsis gene related to APETALA2 Plant Cell. **8:**659-71.

240. **Winston, F., and P. Sudarsanam** 1998. The SAGA of Spt proteins and transcriptional analysis in yeast: past, present, and future Cold Spring Harb Symp Quant Biol. **63:**553-61.

241. **Wittschieben, B. O., G. Otero, T. de Bizemont, J. Fellows, H. Erdjument-Bromage, R. Ohba, Y. Li, C. D. Allis, P. Tempst, and J. Q. Svejstrup** 1999. A novel histone acetyltransferase is an integral subunit of elongating RNA polymerase II holoenzyme Mol Cell. **4:**123-8.

242. **Wolffe, A. P., and J. J. Hayes** 1999. Chromatin disruption and modification Nucleic Acids Res. **27:**711-20.

243. **Workman, J. L., and R. G. Roeder** 1987. Binding of transcription factor TFIID to the major late promoter during in vitro nucleosome assembly potentiates subsequent initiation by RNA polymerase II Cell. **51:**613-22.

244. **Woychik, N. A.** 1998. Fractions to functions: RNA polymerase II thirty years later Cold Spring Harb Symp Quant Biol. **63:**311-7.

245. **Woychik, N. A., and M. Hampsey** 2002. The RNA polymerase II machinery: structure illuminates function Cell. **108:**453-63.

246. **Xiao, T., H. Hall, K. O. Kizer, Y. Shibata, M. C. Hall, C. H. Borchers, and B. D. Strahl** 2003. Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast Genes Dev. **17:**654-63.

247. **Xu, W., D. G. Edmondson, Y. A. Evrard, M. Wakamiya, R. R. Behringer, and S. Y. Roth** 2000. Loss of Gcn5l2 leads to increased apoptosis and mesodermal defects during mouse development Nat Genet. **26:**229-32.

248. **Xu, W., D. G. Edmondson, and S. Y. Roth** 1998. Mammalian GCN5 and P/CAF acetyltransferases have homologous amino-terminal domains important for recognition of nucleosomal substrates Mol Cell Biol. **18:**5659-69.

249. **Yamauchi, T., J. Yamauchi, T. Kuwata, T. Tamura, T. Yamashita, N. Bae, H. Westphal, K. Ozato, and Y. Nakatani** 2000. Distinct but overlapping roles of histone acetylase PCAF and of the closely related PCAF-B/GCN5 in mouse embryogenesis Proc Natl Acad Sci U S A. **97:**11303-6.

250. **Yang, X. J., V. V. Ogryzko, J. Nishikawa, B. H. Howard, and Y. Nakatani** 1996. A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A Nature. **382:**319-24.

251. **Zegerman, P., B. Canas, D. Pappin, and T. Kouzarides** 2002. Histone H3 lysine 4 methylation disrupts binding of nucleosome remodeling and deacetylase (NuRD) repressor complex J Biol Chem. **277:**11621-4.

252. **Zeng, L., and M. M. Zhou** 2002. Bromodomain: an acetyl-lysine binding domain FEBS Lett. **513:**124-8.

253. **Zhang, Y., and D. Reinberg** 2001. Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails Genes Dev. **15:**2343-60.