DEVELOPMENT AND APPLICATION OF MASS SPECTROMETRY BASED PROTEOMICS TECHNOLOGIES TO DECIPHER KU70 FUNCTIONS

MENG WEI

(B.S, NAN KAI UNIVERSITY)

A THESIS SUBMITTED

FOR THE DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF BIOLOGICAL SCIENCES

NATIONAL UNIVERSITY OF SINGAPORE

2007

ACKNOWLEDGEMENTS

My deepest gratitude goes to my supervisor, Assistant Professor Sze Siu Kwan for his patience, encouragements and professional guidance during the last two years. My project will not be completed well without his help and understanding.

My heartfelt thanks also go to my co-supervisors Dr. Ni Binhui and Dr. Walter Stunkel. I was able to finish part of my project in S*Bio Pte Ltd because of their help and valuable suggestions.

In addition, I would like to thank the persons who helped me a lot when I was in Genome Institute of Singapore: Dr.Liu jining, who always discussed with me for my project and gave me a lot of advice; Dr. Hua lin and Dr. Low Teck Yew, who helped me to run mass spectrometry analysis.

I also want to thank Institute for Systems Biology and Dr. J. Donald Capra who provided me the plasmids.

Special thanks should go to Associate Professor Chung Ching Ming and Assistant Professor Dr. Mok Yu-Keung for their invaluable suggestions about my project during the pre-submission seminar.

Especially, I would like my parents know that without their love, support and understanding, this would not have been possible. I really appreciated their trust.

Lastly, I thank National University of Singapore for awarding me a research scholarship and thank Genome Institute of Singapore and S*Bio Pte Ltd providing enough funding for my research.

Table of Contents

Summary

Ku70 is a protein with multiple biological functions. It is well-known that Ku70 forms heterodimer with Ku80 and is essential for the repair of nonhomologous DNA double-strand breaks. Recent studies showed that the acetylation of Ku70 is a master switch in the apoptotic pathway. Therefore, Ku70 might be a therapeutic target for cancer treatment, and it is important to thoroughly study the Ku70 protein complex to unravel its biological functions. We employed mass spectrometry based proteomic methods to characterize the Ku70 protein complex. As proteomics is still in its infancy stage, we have developed novel effective peptide purification and concentration method for in-gel digestion sample in order to drill down to the details of the Ku70 complex by identification of both strong and weak binding partners. In addition, we have adapted one step FLAG-tag purification and two steps tandem affinity purification (TAP) methods to purify the Ku70 protein complex. The FLAG-tag and TAP-tag were fused in-frame to Ku70 gene and the tagged-Ku70 fusion proteins expressed in 293F cell were used as bait to pull down its interacting partners. The pulled-down complex was analyzed by both SDS-PAGE coupled to MALDI-TOF/TOF-MS and shotgun LC-MS/MS proteomics approaches. Epitope-tag based purification strategies enable protein complex to be isolated with exceptional purity and eliminated background of non-specific binding proteins. As a result, it has significantly improved the outcome of mass spectrometry-based protein complex

characterization and enables identification of weak interaction partners. Consequently, 151 proteins were characterized in Ku70 protein complexes. These proteins play a diverse range of biological functions from DNA repair to transcriptional regulation to cellular signal mediator. Among these, 20 are known Ku70 interacting proteins, they function mainly in DNA repairs and telomeric maintenance. Others are mainly cytosolic proteins that are classified to be apoptotic regulatory proteins or signal transduction proteins by Panther gene ontology database. These are consistent with the recent reported Ku70 functions in regulating cellular apoptosis. As Ku70 acetylation has been reported to be a pivotal post translational modification that regulates Ku70 activities, we finally identified the potential acetylation sites of Ku70 by both *in vivo* and *in vitro* acetylation analysis. We are working to further characterize the Ku70 complex by biochemical assays to understand its role in cancer development and other human diseases.

List of Tables

List of Figures

List of Abbreviations

two-dimensional gel electrophoresis	2D-PAGE
ammonium bicarbonate	ABB
acetyl-coenzyme A	AcCoA
acetonitrile	ACN
bovine serum albumin	BSA
octadecyl	C18
octyl	C8
collisionally activated or aided dissociation	CAD
chitin-binding domain	CBD
calmodulin-binding peptide	CBP
α-cyano-4-hydroxycinnamic acid	CHCA
collision-induced dissociation	CID
dimethyl pimelimidate	DMP
dimethylsulfoxide	DMSO
DNA-dependent serine/threonine protein kinase	DNA-PK
double-strand breaks	DSB
dithiothreitol	DTT
1-ethyl-3-(3-dimethylaminopropyl) carbodiimide	EDC
electrospray ionization	ESI
histone acetyltransferase	HAT
histone deacetylase	HDAC
histone deacetylase inhibitors	HDACI
high-performance liquid chromatography	HPLC
iodoacetamide	IAA
immunoblotting	IB
immunoprecipitation	IP
Luria Bertani	LB
matrix-assisted laser desorption/ionization	MALDI
mass spectrometry	MS
tandem mass spectrometry	MS/MS
nonhomologous end joining	NHEJ
polyacrylamide gel electrophoresis	PAGE
poly(ADP-ribose) polymerase	PARP
phosphate-buffered saline	PBS
p300-CBP-associated factor	PCAF
peptide mass fingerprinting	PMF

phenlmethylsulfonyl fluoride	PMSF
post-source decay	PSD
polytetrafluoroethylene	PTEE
post-translational modification	PTM
suberoylanilide hydroxamic acid	SAHA
strong-cation-exchange	SCX
sodium dodecyl sulfate	SDS
solid phase extraction	SPE
staurosporine	STS
tandem affinity purification	TAP
tobacco etch virus	TEV
trifluoroacetate	TFA
time-of-flight	TOF
trichostatin A	TSA

Chapter 1

Introduction

The completion of the genome sequences of the human and many other organisms is undoubtedly a big step towards the full understanding of biological sciences. The availability of enormous amount of data in the genome and the expressed sequence tag (EST) databases opens new avenues to analyze protein functions and leads us to the post-genomic era. Proteomics, the global analysis of proteins, is a new field of research in the post-genomics era. As proteins are the functional molecules that control the living process, from structural elements, to catalysts, to signaling messengers and gene regulatory transcription factors, proteomics data will provide rich information to unravel how living systems work at the molecular level.

Proteins do not act alone. They usually form protein complexes to exert different functions and transmit different signals. All the biological activities depend upon direct physical interaction of specific cellular proteins. Each protein in living matter functions as part of an extended web of interacting molecules. The classic view of protein function focuses on the action of a single protein molecule. However, to truly understand a multifunctional protein, the isolated study of individual protein is not enough and thorough. The pull down of protein complex in a given cell or tissue at a defined condition will then provide a detailed information about the interaction partners of a specific protein in a certain state. The interaction partners will reveal the functions of the target protein. With the availability of various databases such as Gene Ontology (GO) or Panther which collect all the known molecular functions of proteins and group them to pathways and

biological processes, the functions of target protein can be classified to functional pathways and biological processes by interrogating the set of pull down proteins with these databases.

In characterizing binding partners for a molecule of interest, the quality of the purified protein complex plays a critical role in the ultimate success of the experiment. Classical biochemical purification methods rely heavily on the biophysical properties of a given protein, for example, a typical immunoprecipitation (IP). However, most antibodies usually cross-react with many irrelevant proteins, thus generate significant background. Although it will not interfere with the traditional biochemical assays such as Western blotting which only probes for the presence of one specific protein, it will generate significant false positive results when the pull down sample is analyzed by mass spectrometry method which unselectively identifies all proteins in the sample. To circumference the problem of nonspecific binding contaminants, affinity purification procedures utilizing epitoge-tag or two consecutive epitope tags - tandem affinity purification (TAP) steps have been developed and proven to be highly effective in the identification of protein complexes. In this tandem affinity purification (TAP) method, two affinity tags with orthogonal purification properties such as IgG-binding domain of protein A of *Staphylococcus aureus* (ProtA) and calmodulin-binding peptide (CBP) are inserted into a vector. The protein of interest is fused to the TAP tag and expressed in the organism or cell line under investigation. The two TAP tags are separated by spacer regions and a cleavage site for tobacco etch virus

(TEV) protease. The fusion protein and associated components are first recovered from cell extracts by affinity selection on an IgG matrix and eluted by incubation with TEV protease. The eluates which contained protein complexes were further purified by incubation with calmodulin coated beads in the presence of calcium. Highly purified protein complexes were eluted by depleting of calcium ions with EGTA. The two steps purification with two different affinity chromatographic methods minimizes the nonspecific binding contaminants. The whole procedure is carried out under mild, nondenaturing conditions that maximize the chance of isolating an intact and functional protein complex.

Detecting trace amount of TAP tag purified protein complex is challenging. Compared with DNA analysis, trace DNA in a sample can be detected by amplifying the DNA using polymerase chain reaction (PCR). A method for amplifying protein has not yet been developed. Mass spectrometry is the method of choice for protein identification and analysis because of its sensitivity and accuracy. Recent development in ionization methods and improvement on instrumentation have significantly enhanced its performance in protein characterization. It has rapidly become a standard method for protein analysis.

Molecular weight is a unique property of each molecule. Mass spectrometry is an instrument designed to measure molecular weight accurately, and thus to characterize the molecule through its molecular weight. The application of mass spectrometry for biomolecules analysis is triggered by

the recent advent of electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI) methods. These soft ionization methods can bring the fragile biomolecule into the gas phase with charges. The molecular weight of the intact protein can be measured accurately by different types of mass analyzers such as time-of-flight (TOF), quadrupole, quadrupole ion trap, Orbitrap and Fourier transform ion cyclotron resonance (FTICR, also known as FTMS). Coupling different ionization methods to different mass analyzers provides a versatile set of mass spectrometry methods with surprising sensitivity and accuracy for protein and other biomolecular analysis.

The true power of mass spectrometry for protein characterization is the utilization of tandem mass spectrometry (MS/MS) and proteolytic digestion for protein sequence analysis and for the localization of post-translational modifications. Trypsin is by far the most widely used protease in proteomic analysis. Trypsin cleaves proteins at lysine and arginine residues, unless either of these is followed by a proline residue in the C-terminal direction. The spacing of lysine and arginine residues in many proteins is distributed relatively evenly. The resulting tryptic peptides are of a length well-suited to MS analysis.

The most common MS/MS methods currently for protein fragmentation involve low energy fragmentation by collisionally-activated dissociation (CAD) or post-source decay (PSD). The sequence information is obtained from the generation of fragments by the cleavage of the protein backbone. The post translational modification of protein can be determined by

the shift in mass of fragments with the molecular weight of the modified functional group.

The detection of post-translational modification (PTM) is another important application of modern mass spectrometry. PTM represents an important mechanism for regulating protein function. Lysine acetylation, or the transfer of an acetyl group from acetyl coenzyme A to the ε-amino group of a lysine residue, is among the most important post-translational modifications. It was initially discovered on histones. But later many nonhistone proteins were also found to be acetylated indicating the critical role of this modification. This project focuses on Ku70, a protein with multiple functions and can be acetylated by histone acetyltransferase. The TAP purification and mass spectrometry analysis of Ku70 complex will be described and discussed in Chapter 3. The *in vivo* and *in vitro* acetylation of Ku70 will be discussed in Chapter 4.

As proteomics is still in its infancy stage, technology development is the key to advance the field to address more sophisticated biological questions. Since the success of a mass spectrometry based proteomic analysis critically depend on the samples quality, an optimal sample preparation is one of the most important factors in proteomics analysis. Particularly, the interface between protein digestion and mass spectrometric analysis has a large influence on the overall quality and sensitivity of the analysis. As salt and detergent are detrimental to ionization process, a typical sample preparation involves in-gel digestion, sample extraction and concentration, desalting and

elution. Solid phase extraction (SPE) is a widely used technique for the purification and concentration of analytes from liquid samples to improve sensitivity in the analytical process. Commercial tips containing chromatographic material polymerized into pipette tips (ZipTips, Millipore) have also been introduced and are widely used. However, both of them have some shortcomings. New methods which have large binding capacity, small elution volume, minimum sample lost are still under development. Here we developed a novel method for concentration and purification of proteolysis products by using small pieces of Empore Disk (3M, Minneapolis, MN). The Empore Disk is a particle-loaded membrane that incorporates tightly solid phase sorbent particles within an inert matrix of polytetrafluoroethylene. The high density of the particles increases the extraction efficiency and reduces the elution volumes. Chapter 2 will describe this newly developed method.

Overview

This dissertation describes development and application of mass spectrometry based proteomics technologies to decipher Ku70 biological functions. We first improve the proteomics sample preparation method in order to enhance the sample recovery after tryptic in gel digestion. Chapter 2 describes this newly developed method for concentration and purification of proteolysis products by using small pieces of Empore Disk (3M, Minneapolis, MN). This method has been applied to maximally recovery peptides from in-gel and in-solution digestion. Then we generated two cell lines with stable

transfected FLAG and TAP tagged Ku70 in HEK293F cell respectively. Chapter 3 describes the tandem affinity purification and mass spectrometry analysis of Ku70 complex. The *in vivo* and *in vitro* acetylation analysis of Ku70 is discussed in Chapter 4.

Chapter 2

Empore Disk Extraction of Peptides

From In Solution and In Gel Digestion

2.1 Introduction

In the post-genomic era, proteomics research attracts much effort to identify new biomarkers or targets for drug discovery. Protein identification by mass spectrometry plays an important role in proteomics research because of the sensitivity, specificity and speed of mass spectrometer. Since the sensitivity and accuracy of the analysis critically depend on the quality of sample to be analyzed by mass spectrometry, optimal sample preparation is one of the most critical factors for positive result in proteomics analysis. Particularly, the interface between protein digestion and mass spectrometric analysis usually determines the overall quality and sensitivity of the analysis.

A proteome can be characterized by either shotgun proteomics approach with LC-MS/MS or 2D gel approach couple to mass spectrometry for gel spot identification. In shotgun approach, the protein mixture is first digested in solution to a complex peptide mixture. The peptide mixture is then separated by multi-dimensional liquid chromatography coupled to on-line electrospray ionization (ESI). Peptide ions generated by ESI are subsequently subjected to collision induced dissociation in tandem mass spectrometer to sequence the peptides. In gel approach, the protein mixture is separated by polyacrylamide gel electrophoresis. Proteins are detected by staining methods. Relevant protein bands or spots are then cut out, digested with trypsin “in gel”, and the resulting peptide mixtures are analyzed by MALDI-TOF-MS or LC-MS/MS to identify the protein.

The peptide samples obtained by proteolysis are usually not directly analyzed, because they are in buffer with high concentration of salts or detergents that are not compatible with high-sensitivity mass spectrometric analysis, or the concentration of the samples are too diluted to be directly detected. Concentration and purification are routinely achieved by binding of the peptides to reversed-phase material in microcolumns or microtips. The peptides are then eluted off-line in one or several steps for subsequent analysis by mass spectrometry.

Solid phase extraction (SPE) is a widely used technique for the purification and concentration of analytes from liquid samples to achieve increased sensitivity in the analytical process. Bonded silica sorbents are commonly used for the solid phase extraction of analytes from complex samples. A variety of functional groups, such as octadecyl (C18) and octyl (C8) can be bonded to the silica surface to provide non polar interactions. Each of these sorbents exhibits unique properties of retention and selectivity for a particular analyte.

Commercial tips containing chromatographic material polymerized into pipette tips (ZipTips, Millipore) have been introduced and are widely used for concentration and purification of femtomoles to picomoles of biomolecules for MS analysis. The ZipTips pipette tips are simple and easy to use for single or few samples. However, the operation is tedious for large amount of samples. The need to pass the same sample through the material multiple times manually introduces repetitive pipetting steps. Besides their high material

costs, the binding properties of these tips and the required elution volumes have been found unsatisfactory by some researchers (Larsen et al., 2002). ZipTips are also not readily coupled to nanoelectrospray needles.

Here we developed a method for concentration and purification of proteolysis products by using small piece of Empore Disk (3M, Minneapolis, MN). The Empore Disk is a particle-loaded membrane. This membrane can be secured into a variety of devices that offer advantages. The Empore membrane technology incorporates tightly solid phase sorbent particles within an inert matrix of polytetrafluoroethylene (90% sorbent: 10% PTFE, by weight). The high density of the particles increases the extraction efficiency and reduces the elution volumes. At the same time, the PTFE fibrils do not interfere with the activity of the particles in any way.

Many C18 microspin columns were developed in recent years for the desalting and concentration of peptides (Rappsilber et al., 2003; Naldrett et al., 2005; Ishihama et al., 2006). These devices are expensive and low throughput. We introduce here a simple, inexpensive and efficient method for peptide extraction from digestion buffer by C18 Empore Disk. A small piece of C18 Empore material can be reproducibly corked out by a hollow tool, such as a blunt tipped hypodermic needle or a 200 µl pipette tip in exactly the same way that cookies are cut. One disk is sufficient for producing thousands of such membrane disks. Peptides in aqueous solution bind to C18 materials through hydrophobic interactions, allowing small interfering molecules (salts,

detergents) to be washed off. The peptides are then eluted with organics solvent that are compatible with mass spectrometry analysis.

2.2 Materials and Methods

2.2.1 In Solution Empore Disk Extraction

Samples with 1 picomol, 100 femtomol, 50 femtomol and 10 femtomol commercial digested BSA peptides (MICHROM Bioresources Inc.) in 10µl 100mM ABB/0.1% TFA were prepared. Empore Disk C18 (3M) was cut manually by a 200 µl pipette tip in cookie cutter fashion. The cut-out Empore Disks were first wet by 100% ACN and conditioned by 100% methanol and then put inside the BSA peptides solution incubating for 30 minutes at 60°C and 3 hours at room temperature. The Empore Disk was then washed with 200µl 0.1% TFA for 15 minutes and eluted by 2µl elution buffer (90% ACN, 0.1% TFA) for 1 h. In order to elute the peptides completely, the small piece of Empore Disk was inserted into the tip of a 200µl gel loading pipette tip by a capillary. The elution buffer was pressed through Empore Disk for 15 times by a 200µl Eppendorf pipette. The eluted peptides were mixed with same amount of matrix solution containing 10 mg/ml of α-cyano-4-hydroxycinnamic acid (CHCA) in 0.1%TFA/50%ACN, and spotted onto a 384-well stainless-steel MALDI target plate (Applied Biosystems).

2.2.2 SDS-polyacrylamide Gel Electrophoresis (SDS-PAGE)

Different amounts of standard protein BSA (Sigma) were prepared and run in different lanes of 7.5% polyacrylamide gel from 10 picomol, 1 picomol, 500 femtomol, 100 femtomol, 50 femtomol, 10 femtomol, 5 femtomol to 1 femtomol with a Mini protein II electrophoresis apparatus (Bio-Rad Laboratories, California, USA) or on 4-12% (w/v) gradient pre-cast NuPAGE® Novex Bis-Tris gels (Invitrogen) with a XCell *SureLock*™ Mini-Cell (Invitrogen).

2.2.3 Silver Staining

For low amount of BSA, after electrophoresis, the gels were fixed in fixing buffer (40% Ethanol, 10% Acetic acid and 50% H_2O) and stained using SilverQuest Silver Staining Kit (Invitrogen) according to the manufacturer's instructions.

2.2.4 Simply Blue Staining

For high amount of BSA, after electrophoresis, the gels were stained in Simply Blue solution (Invitrogen) by microwave method according to the manufacturer's instructions. The gel was then distained by MilliQ water until the background of the gel was no longer bluish.

2.2.5 In-gel Empore Disk Extraction

Gel bands were directly excised from the gel and cut into small pieces. Gel bands of small amount of BSA which cannot be seen by eyes were cut according to their molecular weight. After destaining, washing by water followed by 100% ACN to dehydration, gel bands were completely dried to "glass" state by freeze drier. The dried gels were broken into powder by a homogenizer and then reduced by 10mM DTT at 60°C for 1h and alkylated by 55mM IAA at room tempeture in dark for 1h. The gels were then washed by 100mM ABB and 100% ACN and dehydrated in a Speed-Vac. 15µl of Trypsin (Promega) was employed to digest BSA at the ratio of 1:50 (Trypsin:BSA) overnight (12-16h) at 37°C. After digestion, the reaction was terminated by adding TFA to a final concentration of 0.1%. Small pieces of Empore Disk were first wet by 100% ACN and equilibrated by 100% methanol and then put inside the gel bands solution. The extraction procedure was the same as in solution Empore Disk extraction which was described above. The eluted peptides were mixed with same amount of matrix solution containing 10 mg/ml of α-cyano-4-hydroxycinnamic acid (CHCA) in 0.1%TFA/50%ACN, and spotted onto a 384-well stainless-steel MALDI target plate (Applied Biosystems).

2.2.6 MALDI TOF/TOF MS/MS Analyses

An ABI 4800 Proteomics Analyzer MALDI TOF/TOF mass

spectrometer (Applied Biosystems) was used to analyze the samples on the MALDI target plates. For MS analyses, typically 1800 shots were accumulated for each sample. MS/MS analyses were performed using post-source decay (PSD) fragmentation.

2.2.7 Data Analysis

MASCOT search engine (version 2.0; Matrix Science) was used to search all of the tandem mass spectra. GPS Explorer™ software version 3.6 (Applied Biosystems) was used to search files with the MASCOT search engine for peptide and protein identifications. Swissprot database was used for the search. The MS and MS/MS spectra were combined for the search. Cysteine carbamidomethylation and methionine oxidation were selected as variable modifications. Three missing cleavages were allowed. Precursor error tolerance was set to <0.3 Da and MS/MS fragment error tolerance < 0.3 Da.

2.3 Results

2.3.1 In Solution Empore Disk Extraction

Samples with different amount of commercial digested BSA peptides were prepared in 10µl 100mM ABB/0.1% TFA. The samples were concentrated by Speedvac to about 1 µl and mixed (1:1) with matrix in 50% acetonitrile/0.1%TFA. All the samples were spotted on a stainless steel

MALDI plate. But only one or two spots gave signals. Most of the spots did not produce any precursor ions suggesting that salt will greatly suppress the ionization of peptides. No ion precursors then can be analyzed by MS/MS. Thus, desalting is an important step before samples being submitted to MALDI-TOF/TOF.

Same samples were desalted by cut-off Empore Disk as described above. After mixing with matrix solution, the final eluates were spotted to MALDI plate to be analyzed by mass spectrometer. The obtained peak lists were used for Mascot's mass fingerprint search combined with MS/MS data search. BSA in all samples were detected by MS with reasonable Mascot scores (Table 2.1). Down to 10 femtomol BSA can be detected. We also used C18 Ziptip to desalt the same samples. Only 1 picomole, 100 femtomol and 50 femtomol BSA could be detected. The sample with 10 femtomole BSA were not detected because of high background noise and low signal intensity.

The amount of ACN in elution buffer was optimized. We tried 50% ACN, 70% ACN, 90% ACN and 100% ACN. By comparing the signal intensity, the Mascot scores and sequence coverage, 90% ACN/0.1%TFA was found to be the best elution buffer.

2.3.2 In Gel Empore Disk Extraction

Large amount of BSA from 100 picomole down to 1 picomole were loaded and run by SDS-PAGE and then stained by Simply Blue. Gel bands

were directly excised from the gel and cut into small pieces. The gel was digested and desalted by Empore Disk as described above. All of the bands with large amount of BSA could be detected with very high scores.

Low amount of BSA from 1 picomole down to 1 femtomole were also run by SDS-PAGE and stained by silver staining. Gel bands with less than 100 femtomole BSA are too week to be seen by eyes (Figure 2.1). So they were cut according to the visible bands with sufficient amount of BSA in other lanes of the same gel. The bands were digested and desalted by Empore Disk. The bands with one picomole and 500 femtomole of BSA are consistently detected by MALDI-TOF/TOF every time. However, the detectabilities of samples with less than 500 femtomole of BSA are varies from experiment to experiment. Sometimes, 10 femtomole BSA loaded on gel can be detected with good score!

We also compared with normal in gel digestion protocol desalting by C18 Ziptip. The difference lies in the procedure after digestion. Normal protocol sequesters trypsin and extracts peptide by adding 50% ACN/5% formic acid. Then the solution with extracted peptides was desalted by Ziptip. But only band with 1 picomole BSA can be detected.

2.4 Discussion

The quality of MALDI mass spectra is highly dependent on the sample preparation step prior to MS analysis, especially for in gel sample

Table 2.1 MS analysis of in solution and in gel digestion

In solution digestion

Amount of peptides	peptide matched	MASCOT score	Peptide coverage(%)	MS cluster area matched
1 picomole	20	716	36	54.156
100 femtomole	23	534	44	40.268
50 femtomole	22	378	40	30.966
10 femtomole	17	145	30	14.486

In gel digestion

Amount of proteins	peptide matched	MASCOT score	Peptide coverage(%)	MS cluster area matched
1 picomole	41	553	69	36.973
500 femtomole	13	400	19	9.911

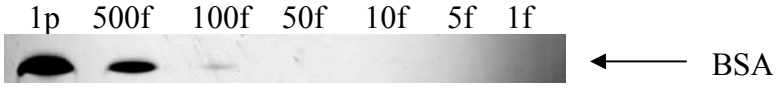


Figure 2.1 Silver staining of different amount of BSA
1 picomol, 500 femtomol, 100 femtomol, 50 femtomol, 10 femtomol, 5 femtomol to 1 femtomol BSA (From left to right)were loaded and run by SDS-PAGE and then stained by sliver.

preparation. In gel digestion is much trickier than in solution digestion. The peptides may crosslink with gel matrix and hard to be extracted from it. The recovery of peptides sometimes is quite low which makes it hard to be detected by mass spectrometer. Different staining methods also affect the efficiency of in gel digestion. Silver staining is not as compatible as Simply Blue or Commassie blue staining with MS. The formaldehyde in most of silver staining reagents is possible to crosslink proteins to gel matrix and reduce the ratio of peptides from diffusing to solution. Our method is not only efficient in in-solution digestion, but also may improve the low recovery of in-gel digestion. After proteins in gel are digested by trypsin, some of peptides will diffuse from the gel into solution. We hypothesize that a equilibrium of peptide distribution between gel and solution was established. When Empore Disk was put into the solution, peptides in solution were concentrated to it so that the concentration of free peptides in solution were decreased. To keep the equilibrium, more peptides in gel were diffused out to solution. The binding of peptides to Empore Disk shift the equilibrium. As a result, peptides were gradually migrated from gel to Empore Disk.

Empore Disk has high capacity. For low amount of peptides, only small piece is needed. The small material allows small volume of elution buffer which makes the avoidance of Speedvac possible. Speedvac dry down and recovery were reported to lead to adsorptive losses (Stewart et al., 2001). Minimizing losses is very important especially for low amount of proteins. Losses can occur during every step: being extracted from gel, binding to

membrane, elution from the membrane, Speedvac dry down and recovery. It can also occur due to binding to the plastics the samples are contained in, such as tubes, tips etc. Trying to avoid or reduce the losses in every step became the key element for low amount samples. The use of Empore Disk can reduce many of the losses as explained above. We also used low-binding tubes to further prevent the losses of binding to plastics.

Our result showed that Empore Disk extraction can detect proteins down to 10 femtomole in in-solution digestion and at least 500 femtomole in in-gel digestion. Interestingly, very low amount of peptide can be detected sometimes. We got good signals from the samples with 1 femtomole BSA or 5 femtomole BSA in in-gel digestion while nothing can be detected in the samples with 50 femtomole and 100 femtomole. Such situation happened several times. One possible reason we can provide is that the mixing of peptides and matrix solution and distribution of the mixture on MALDI spots are not equal. The peptides in low amount samples are few. When beam of Laser fires on the spot, it's random. Only part of the matrix-sample mixture will be induced into ions. If the beam happens to fire the area where is full of peptides, it will get strong signals.

Compared to desalting with Ziptip, the efficiency of Empore Disk was slightly better. Considering their cheap material costs and easy handling for batch samples, they have the potential to be widely used.

2.5 Conclusion

We have presented a new method for desalting and extraction of peptides from in-solution an in-gel digestion. It is not only suitable for sample preparation prior to MALDI-MS, but also LC/MS analysis. The Empore Disk extraction minimizes the losses of samples and offers a very effective and convenient tool for sample preparation. The procedure is simple, reproducible and economical. Its handling is easy and convenient for batch samples. These properties make it useful in proteomics research.

Chapter 3

Proteomics Studies on Ku70 Protein Complex by Tandem Affinity Purification (TAP) Tag Pull Down and Mass Spectrometry

3.1 Introduction

3.1.1 Mass Spectrometry (MS)

In the post-genome era, mass spectrometry (MS) has emerged as a powerful tool to quickly and efficiently identify proteins in biological samples, placing MS at the forefront of technologies in proteomics research. MS is a highly sensitive tool capable of analyzing samples ranging in size from small molecules to whole viruses.

During the last decade, great advances in MS instrumentation and techniques revolutionized protein chemistry and fundamentally make it possible to investigate the proteome of a cell, an organ or even an organism. How to generate ions from large, nonvolatile analytes such as proteins and peptides without significant analyte fragmentation is the major problem when applying mass spectrometry to biological research. This difficult problem is solved by two technical breakthroughs in the late 1980s: the development of the two soft ionization methods - electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) (Karas and Hillenkamp, 1988; Fenn et al., 1989; Hillenkamp and Karas, 1990) which catalysed the quick development of MS instruments later. Now, commercial MS instruments can offer routine picomole to attomole analytical sensitivity of a large variety of compounds, including: proteins, peptides, carbohydrates, oligonucleotides, natural products, drugs and drug metabolites. Perhaps most exciting is that the developmental stage of mass spectrometry has not stopped; innovations such

as nanoelectrospray, curved reflectrons and electrospray with orthogonal spraying continue to expand its capability.

The five basic parts of any mass spectrometer are: a vacuum system; a sample introduction device; an ionization source; a mass analyzer; and an ion detector. Combining these parts, a mass spectrometer determines the molecular weight of molecules by ionizing, separating, and measuring molecular ions according to their mass-to-charge ratio (m/z). First, an ionization source ionizes the molecule of interest, then a mass analyzer differentiates the ions according to their mass-to-charge ratio and finally, a detector measures the ion beam current. Each of these elements exists in many forms and is combined to produce a wide variety of mass spectrometers with specialized characteristics.

Both Electrospray ionization (ESI) and Matrix-assisted laser desorption ionization (MALDI) mass spectrometry are sensitive ionization sources for mass measurement and can provide surprisingly large amount of other information as well. The ability to analyze complex mixtures has made ESI and MALDI very useful for the examination of proteolytic digests, an application otherwise known as peptide mass fingerprinting (PMF). Through the application of sequence specific proteases, mass analysis is performed on the resulting proteolytic fragments thus yields mass information of the fragments. The specific protease fragmentation pattern is then compared with the patterns predicted for all proteins within a database and matches are statistically evaluated.

Electrospray ionization (ESI) is one of the most exciting ionization techniques for biomolecules. It generates ions directly from solution (usually an aqueous or aqueous/organic solvent system) by creating a fine spray of highly charged droplets in the presence of a strong electric field. As the droplets pass through a heated chamber, the buffer is evaporated, sending desolvated peptide ions to the mass analyzer. As the peptide ions enter the mass analyzer, their mass/charge ratios are measured in real time and spectrum is generated. Since ESI generates ions directly from solution, it is especially useful for the shotgun analysis which refers to strategies that involve the proteolytic digestion of a mixture of proteins and the subsequent analysis of the resulting even more complex mixture of peptides (McCormack et al., 1997). This shotgun technique usually detects more low abundant proteins in a medium complex sample than the gel-MS approach. Thus we have adapted mainly this shotgun proteomics approach to character the pulled-down Ku70 protein complex.

Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS), first introduced in 1988 by Tanaka and independently by Hillenkamp and Karas, has become a widespread analytical tool for peptides, proteins and most other biomolecules. MALDI provides for the nondestructive vaporization and ionization of both large and small biomolecules. In MALDI analysis, the analyte is first co-crystallized with a large molar excess of a matrix compound, usually a UV-absorbing weak organic acid, after which pulse UV laser radiation of this analyte-matrix mixture results in the

vaporization of the matrix which carries the analyte with it. The matrix therefore plays a key role by strongly absorbing the laser light energy and causing, indirectly, the vaporization and ionization of the analyte. The matrix also serves as a proton donor and receptor, acting to ionize the analyte in both positive and negative ionization modes, respectively.

To obtain more structural information on the molecular ions generated in the electrospray ionization and MALDI ionization sources, it has been necessary to apply tandem mass spectrometry (MS/MS) to induce fragmentation. Fragmentation of precursor ions is accomplished either through application of 'slow heating' methods or of more recently developed gas-phase ion-ion reactions or ion-electron reactions (Horn et al., 2000; Shukla and Futrell, 2000; Zubarev et al., 2000; Sleno and Volmer, 2004). The most commonly used slow heating fragmentation method is low-energy CID (also known as collisionally activated or aided dissociation, or CAD). It involves slow transfer of energy by collision with a neutral molecule to the precursor ions such that the internal energy of the ions exceeds the energy needed for fragmentation. Collision-induced dissociation is accomplished by selecting an ion of interest with a mass analyzer and introducing that ion into a collision cell. The selected ion then collides with a collision gas (typically argon, helium or nitrogen) leading to the disruption of peptide bonds. The fragments are then analyzed to obtain a fragment ion spectrum, which is rich in sequence information. Finally, protein database search engine (such as SEQUEST, Mascot, etc.) compares the acquired MS/MS spectrum to theoretical spectra

generated from protein sequences in an appropriate database and identifies the proteins in the original samples.

Immediately following ionization, gas phase ions enter a region of the mass spectrometer known as the mass analyzer. The mass analyzer is used to separate ions within a selected range of mass-to-charge (m/z) ratios. The analyzer is an important part of the instrument because of the role it plays in the instrument's accuracy and mass range. Ions are typically separated by magnetic fields, electric fields, or by measuring the time it takes an ion to travel a fixed distance. Recently, the most commonly used mass analyzers include quadrupole, ion trap, time-of-flight (TOF) mass analyzers. Coupled to ESI and MALDI, they provide surprisingly sensitive and accurate result (Aebersold and Goodlett, 2001).

State-of-the-art proteomic workflow involves multiple interdependent steps: sample preparation, protein-peptide separation, mass spectrometry and data analysis. Improvements in any of these steps can increase the accuracy and sensitivity of analysis. In order to reduce the complexity of the samples, the proteins or proteolytic peptides are usually separated before submitted to mass spectrometry. Both one-dimensional sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (SDS-PAGE) and, providing even greater separative capacity, two-dimensional gel electrophoresis (2D-PAGE) are powerful tools for the separation and analysis of proteins. One common workflow in proteomics involves 1D-PAGE or 2D-PAGE followed by mass spectrometry-based identification. For interface with electrospray ionization

(ESI), some of the most popular and effective separation techniques use high-performance liquid chromatography (HPLC), a liquid interface allows for coupling the separation step directly to the ionization source of the mass spectrometer. Reverse-phase (RP) chromatography is the most commonly used technique for separating peptides. For the most complex mixtures of proteins or peptides, it is necessary to add another dimension of separation. This is most commonly accomplished using a strong-cation-exchange (SCX) column for the first dimension, followed by multiple RP separations. Other separation strategies that have been utilized in proteomics workflows include other types of HPLC, isoelectric focusing, capillary electrophoresis and chip-based separations (Goodlett and Yi, 2002; Issaq et al., 2002; Shen and Smith, 2002; Delahunty and Yates, 2005).

Due to its sensitivity and specificity, mass spectrometry is widely applied to identify protein complex in combination with a proper biochemical purification of proteins. This strategy is becoming an important tool to define relations existing among gene products.

3.1.2 Tandem Affinity Purification Strategy

The transmission of a wide range of biological signals and the proceeding of all the biological activities depend upon direct physical interaction of specific cellular proteins. A more thorough understanding of the precise mechanisms of these signal transduction pathways and biological

processes relies on the identification and biochemical characterization of the specific molecules involved. Protein does not act alone. Instead, they usually interact with one another to exert their cellular function in concert. The resulting protein complexes are highly ordered dynamic structures that assemble, store and transduce biological information.

The availability of complete genome sequence of several organisms provides an exceptional opportunity to analyze the different functions governed by these genes. Many researchers are looking for high throughput technique to systematically study protein-protein interaction networks. This would open the way for a chain of predictions: open reading frame function, complex formation, and supramolecular organization of the proteome. Large-scale two-hybrid screening has been used for this purpose (FromontRacine et al., 1997; Ito et al., 2000; Uetz et al., 2000). However, the results are often false-positive and false-negative, especially when performed in heterologous organisms, and restricted to the description of pairs of proteins. Moreover, the limited set of conditions makes it desirable to use additional strategies to easily detect protein interactions.

In characterizing binding partners for a molecule of interest using MS, the major challenge is to identify *bona fide* interacting partners *versus* sample contaminants. Because of the sensitivity of the mass spectrometry, detection and identification of protein by MS is not the bottle neck, while the quality of the purification method plays a more and more critical role in the ultimate success of an experiment.

Classical biochemical purification methods rely heavily on the biophysical properties of a given protein. For example, a typical immunoprecipitation (IP). However, antibodies usually cross-react with one or more irrelevant proteins; Moreover, many proteins bind non-specifically to the solid support resin and antibody. A typical IP will therefore contain not only cross-reactive and other non-specific contaminating proteins, but will also bring down their interacting partners, resulting in the identification of multiple proteins that are unrelated to the target complex, thus generate significant background. Since each protein is unique, designing a purification protocol for a given protein is often empirical. The lack of a generic protein complex purification protocol hampers the progress in elucidating the functional interactions between cellular proteins.

Among a plethora of protein purification techniques, affinity purification through epitope-tag appears to be the most efficient and gentle discriminatory separation technique for the retrieval of protein complexes. Well-characterized small epitope-tags such as Flag, 6XHis or GST have been widely used in affinity purification of recombinant proteins or protein complexes. However, nonspecific binding is the biggest problem for these single-step purification. How to reduce nonspecific contaminaton becomes the key for the purification of protein complexes. Two generic affinity purification procedures utilizing two consecutive affinity purification steps have been developed and proven to be highly effective in the identification of protein complexes in yeast (Rigaut et al., 1999; Puig et al., 2001) with minimal

background. Rigaut *et al.* described the tandem affinity purification (TAP) method (Rigaut et al., 1999). They chose IgG-binding units of protein A of *Staphylococcus aureus* (ProtA) and calmodulin-binding peptide (CBP) as the two tags among several widely used tags, eg. FLAG tag, the Strep tag, the His tag, the chitin-binding domain (CBD) etc. by selecting the yeast SmX4p protein.

In this method, the protein of interest must be fused to the TAP tag and expressed in the organism or cell line under investigation. The TAP tag contains two epitope tags: a ProtA tag and a CBP tag. Both epitope types are separated by spacer regions and a cleavage site for tobacco etch virus (TEV) protease. Extracts prepared from cells expressing the TAP-tagged protein are subjected to two successive purification steps. The first purification step involved binding of the ProtA tag to an IgG column and eluting bound materials by incubation with TEV protease. This elution method significantly reduced the contaminating proteins bound nonspecifically to the column. The eluates which contained protein complexes were further purified by incubation with calmodulin coated beads to remove TEV protease and contaminants. Highly purified protein complexes were eluted by adjusting concentration of calcium ions with EGTA.

Compared to the conventional chromatographic methods, TAP method has great advantages. It is efficient, yielding pure protein complex in sufficient amounts for mass spectrometric (MS) identification from small sample size. It can produce an extremely pure preparation of the protein of

interest with associated proteins. The two specific elutions prevent most of contaminants. Furthermore, the whole procedure including the final elution step can be carried out under mild, nondenaturing conditions under which the purified protein complex usually retains functionality.

Although the TAP method has been originally developed in the budding yeast *Saccharomyces cerevisiae*, this approach has been successfully applied to other organisms including bacterial (Gully et al., 2003; Kumar et al., 2004; Butland et al., 2005), mammalian (Westermarck et al., 2002; Knuesel et al., 2003; Bouwmeester et al., 2004; Brajenovic et al., 2004) , *Drosophila* (Forler et al., 2003), plant (Rohila et al., 2004) and trypanosomatids (Aphasizhev et al., 2003a; Aphasizhev et al., 2003b; Panigrahi et al., 2003).

Other tandem-affinity tags include a modified version of the TAP-tag in which the CBP part is replaced by the S-tag (Cheeseman et al., 2001) and combinations of multi-histidine-tags with FLAG (Denison et al., 2005; Hannich et al., 2005; Kim, S. C. et al., 2006), myc-tags (Graumann et al., 2004) or bio-tag (Tagwerker et al., 2006) have also been developed. A similar technique using three different tags (CBP, 6XHis, HA) called MAFT for "multiple affinity purification" has also been developed and used for purification of Cyclin-Cdk complexes in yeast (Honey et al., 2001).

3.1.3 Ku70

In this project, we have applied epitope-tag purification and mass spectrometry to isolate and identify Ku70 protein complex to elucidate the roles of Ku70 in biological processes.

Ku70 is a protein with multiple functions. It was originally recognized as an autoantigen as one component of the Ku antigen by autoantibodies from certain Japanese patients with scleroderma-polymyositis overlap syndrom (Mimori et al., 1986) and was subsequently found to be an autoantigen in patients with several other autoimmune diseases, including systemic lupus erythematosis, Graves disease, and Sjogren syndrome (Reeves, 1992). Ku antigen is a complex composed of two subunits of 70 and 80 kDa which are designated Ku70 and Ku80, respectively (Mimori et al., 1986). The *Ku70* and *Ku80* genes are distributed in a wide variety of eukaryotic species, ranging from yeast to human (Dynan and Yoo, 1998).

Ku70 is essential for the repair of DNA double-strand breaks (DSB), which can be induced directly by a variety of damaging agents, or indirectly as a result of endogenous damage arising from the cell's own metabolism by nonhomologous end joining (NHEJ) and the rearrangement of antibody and T cell receptor genes via V(D)J recombination which is a physiologic DNA double strand breakage and rejoining process restricted to lymphoid cells (Rathmell and Chu, 1994; Featherstone and Jackson, 1999). In this role, Ku70 complexes with Ku80 to form the DNA-binding component of the DNA-dependent serine/threonine protein kinase (DNA-PK) (Leesmiller et al., 1990;

Anderson, 1993; Gottlieb and Jackson, 1993; Weaver et al., 1995). The crystal structure of Ku bound to DNA was reported to show the mechanism of Ku-end recognition (Walker et al., 2001) and the three-dimensional structure of the human DNA-PKcs/Ku70/Ku80 complex was also reported (Spagnolo et al., 2006). Ku protein is also important in activating DNA-PK (Jin and Weaver, 1997).

The Ku70/80 heterodimer also has important roles in telomere maintenance and transcriptional regulation (Tuteja, R. and Tuteja, 2000). Ku might function as part of a telomeric length sensing system protecting chromosomal termini from nucleolytic attack, as was shown in yeast (Boulton and Jackson, 1996) and mammals (Bailey et al., 1999; Bianchi and de Lange, 1999; Hsu et al., 1999; Samper et al., 2000). Ku70 and Ku80 heterodimer can also bind to telomeric sequence independent of the DNA-PKcs (Hsu et al., 2000) to prevent chromosomes end-to-end fusions.

Although Ku protein and DNA-PKcs always function together, several studies showed that Ku protein may play some role independent of DNA-PK activity. As expected, Ku protein knockout mice have a deficiency in DNA DSB repair, but in addition, they show a retardation in growth (Nussenzweig et al., 1996; Gu, Y. S. et al., 1997). However, the retardation in growth was not shown in DNA-PKcs-knockout mice (Gao et al., 1998). These knockout experiments suggest that there is an important link between growth control and Ku protein.

Although Ku70 always forms a heterodimer with Ku80 and involves biological process together, it may have its own functions that are independent of Ku80. There are some differences in the phenotype between Ku70- and Ku80-knockout mice (Nussenzweig et al., 1996; Zhu, C. M. et al., 1996; Gu, Y. S. et al., 1997; Li et al., 1998). For example, Ku70-knockout mice have small populations of mature T lymphocytes and a significant incidence of thymic lymphoma, but Ku80-knockout mice do not. Ku70 has been reported to show Ku80-dependent and -independent DNA binding, whereas Ku80 requires association with Ku70 for DNA binding (Wang, J. S. et al., 1998).

Altered expression of Ku70 was found in many tumors, eg. oral cavity carcinomas(Korabiowska et al., 2006b), colorectal carcinoma (Komuro et al., 2005; Korabiowska et al., 2006a), nasopharyngeal carcinoma (Lee et al., 2005) etc., suggesting that Ku70 may involve in the tumorigenesis and may become the target for anti-tumor drugs.

Recent studies suggest that Ku70 has a Bax inhibitory activity that is independent of previously recognized DNA damage repair function. Consistent with this hypothesis, it was reported that Ku70_/_ knockout mice are hypersensitive to ionizing radiation (Ouyang et al., 1997) and have increased apoptotic neuronal death during embryonic development (Gu, Y. S. et al., 2000). Fibroblasts derived from Ku70-deficient mice become hypersensitive to several anti-cancer drugs that do not induce DNA damage (Kim, S. H. et al., 1999). Chechlacz *et al.* also reported that cells from Ku70 knockout mice are hypersensitive to agents, such as staurosporine (STS), that

promote apoptosis in the absence of DNA damage (Chechlacz et al., 2001). Furthermore, ectopic expression of Ku70 protects human colon cancer cells (SW480) against curcumin-induced apoptosis (Rashmi et al., 2004).

Ku70 was originally reported to be a nuclear protein, consistent with its functions as a DNA repair protein. However, several studies have revealed the cytoplasmic or cell surface localization of Ku proteins in various cell types (Prabhakar et al., 1990; Dalziel et al., 1992; Grawunder et al., 1996). Manabu Koike *et al.* found that Ku70 is in the nuclei in interphase cells. But in mitotic cells, most of Ku70 were found diffused in the cytoplasm (Koike et al., 1999). It is suspected that the less abundant cytoplasmic pool is responsible for the sequestration of Bax, the protein that will induce mitochondrial-dependent cell death signals in mammalian cells (Sawada et al., 2003a). By screening a human cDNA library to identify inhibitors of Bax, Sawada *et al.* (Sawada et al., 2003b) identified the Bax-binding and Bax inhibiting properties of Ku70. More recent work has shown that Ku70 is also an acetylation-sensitive binding partner for Bax (Cohen et al., 2004). They showed that Ku70 binds to Bax through its C terminus. Five lysine residues within this domain (539, 542, 544, 553, and 556) can be acetylated by CBP and PCAF. They suggested that acetylation of C terminal of Ku70 will regulate the Bax-mediated apoptosis. The similar result was demonstrated in NB cells (Subramanian et al., 2005).

Besides the two antiapoptotic effects-facilitating DSBR and sequestering Bax to prevent its translocation to mitochondria, the most recent research showed that Ku70 also has other antiapoptotic functions in GzmA-

induced cell death, which are blocked when GzmA proteolyses Ku70 (Zhu, P. C. et al., 2006).

Given Ku70's dual role in DNA end joining and suppressing apoptosis, it could conceivably be a central player in coordinating DNA repair with the decision between cell survival and programmed cell death. In this project, we will probe in-depth to the Ku70 protein complex to decipher the role of Ku70 in regulating DNA repair and apoptosis.

3.2 Materials and Methods

3.2.1 Plasmid Constructions

The Ku70 cDNA was amplified using the following oligonucleotides as primers for the PCR: forward primer 5'CTGGATCCGATGTCAGGGTGG GAGTC3' and reverse primer 5'GATCGCTCACGCGGCCGCTCAGTCCTG GAAGTGCT3' which introduce a BamHI cutting site into the 5' terminal of Ku70 cDNA and a NotI cutting site into the 3'terminal of Ku70 cDNA. PCR product was purified from the gel (QIAquick Gel Extraction Kit, Qiagen, Germany) and double digest by BamHI and NotI. The digested PCR product was purified (QIAquick PCR Purification Kit, Qiagen, Germany) and cloned into the BamHI and NotI sites of pcDNA3-NTAP vector (provided by Institute for Systems Biology) to generate pcDNA3-NTAP-Ku70. The plasmid pWay20-GFP-Ku70 was kindly provided by Dr. J. Donald Capra (Rodgers et al., 2002b).

3.2.2 Preparation of Ku70 Mutants

Ku70 site-specific mutants were formed by incorporating mutant oligonucleotides by strand extension reactions. The QuikChange® Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA) was used according to the manufacturer's recommendation. Each mutant was identified by DNA sequencing. The primer set employed were as follows: Ku70-L385R: forward primer 5' GCTCAACCCTGTTCAGTGCTAGGCTCATCAAGTGTCTGG 3', reverse primer 5' CCAGACACTTGATGAGCCTAGCACTGAACAGGGTT GAGC 3', Ku70-K553A: forward primer 5' CGATAATGAAGGTTCTGGA AGCGCAAGGCCCAAGGTGGAG 3', reverse primer 5' CTCCACCTTGG GCCTTGCGCTTCCAGAACCTTCATTATCG 3', Ku70-E548A: forward primer 5' CCAAGAGAAAACACGATAATGCAGGTTCTGGAAGCAAAA GGCC 3', reverse primer 5' GGCCTTTTGCTTCCAGAACCTGCATTATCG TGTTTTCTCTTGG 3'.

3.2.3 Cell Culture

FreeStyle 293-F Cells were bought from Invitrogen. Cells were cultured in FreeStyle 293 Expression Medium in sterile Erlenmeyer flask and maintained in a 37°C incubator containing a humidified atmosphere of 8% CO_2 in air on an orbital shaker platform rotating at 150 rpm. HEK293 cells were cultured in Minimum Essential Medium (GIBCO) supplemented with

10% fetal bovine serum (FBS), 2mM L-glutamine, 1%Non-Essential Amino Acid solution (NEAA), 1mM NaPyravate, 100u/ml Penicillin and 100mg/ml Streptomycin at 37°C in 5% CO2.

3.2.4 Transient and Stable Protein Expression

Plasmids were transient transfected into HEK293 cells grown at 90% confluency using lipofectamine2000 (Invitrogen) according to the manufacturer's instructions. FreeStyle 293-F cells were transfected at the density of 1 x 10^6 cells/ml using 293fectin (Invitrogen) according to the manufacturer's instructions. FreeStyle 293-F cells stably expressing TAP tagged and FLAG tagged Ku70 were obtained by limiting dilution and the selection of 750 μg/ml Geneticin (GIBCO). Correct integration of constructs and expression of TAP-tagged and FLAG-tagged proteins were verified by immunoblotting with normal mouse or rabbit IgG antibodies and Flag antibody (Sigma).

3.2.5 Cell lysis and Quantitative Protein Essay for Cell Lysate

HEK293 cells were harvested when they were at 90% confluency via trypsinization, centrifugation, and washed in phosphate-buffered saline twice. FreeStyle 293-F Cells were harvested when they reached the density of 3 x 10^6 cells/ml and then centrifuged and washed in PBS twice. Cells were then lysed in lysis buffer ((20% glycerol, 420 mM NaCl, 20 mM HEPES pH 7.9, 0.5%

igepal CA-630 (Sigma), 0.2 mM EDTA, 0.5 mM DDT, 0.5 mM PMSF) containing protease inhibitors cocktail (Sigma). After 1h incubation on ice, the lysates were centrifuged for 30 min at 13,000rpm.

The amount of protein dissolved in supernatant was determined by employing the method of Bradford. 10µl of the proteins and a series of BSA (Sigma) solutions with the concentration of 500µg/ml, 250µg/ml, 125µg/ml, 62.5µg/ml, 31.25µg/ml, prepared in PBS, were added into 200µl diluted Bradford protein assay reagent (Bio-Rad Laboratories, California, USA) which was made by mixing Milli-Q water and Bradford protein assay reagent in the ratio of 4:1.

The absorbance of the preparations was determined at 595nm using a ninety-six-well microtiter plate (Greiner, USA) with a spectrophotometer. With the absorbance reading of the BSA solutions as standards, standard curve was plotted. From the standard curve, the concentrations of proteins present in the samples were determined.

3.2.6 SDS-polyacrylamide Gel Eletrophoresis (SDS-PAGE)

The prepared samples were analyzed on different concentrations of SDS-polyacrylamide gels according to the proteins which were focused with a Mini protein II electrophoresis apparatus (Bio-Rad Laboratories, California, USA) or on 4-12% (w/v) gradient pre-cast NuPAGE® Novex Bis-Tris gels (Invitrogen) with a XCell *SureLock*™ Mini-Cell (Invitrogen). .

3.2.7 Immunoprecipitation and Western Blot

Protein G Sepharose beads (Amersham) were washed by 10mM Tris-HCl and blocked by 600μg/ml BSA first in IP buffer (PBS+0.1%Triton X). Clear cell lysate containing 500 μg of total protein were precleared by incubating with 30 μl of protein G-Sepharose beads for 1h at 4 °C. After removal of the beads, the extracts were incubated overnight (~14 h) at 4 °C with 2 μg anti-GFP antibody (Santa Cruz) or monoclonal anti-Ku70 antibody (Santa Cruz). 30μl protein G Sepharose beads were then added into the extracts and incubated for 3h. The beads were collected by centrifugation at 2000rpm for 5min at 4 °C and washed three times with 1ml of IP buffer.

For Western blot analyses, the beads were resuspended in 30 μl of 2× SDS-polyacrylamide gel sample buffer, boiled for 5mins, and proteins were resolved by SDS-PAGE. Proteins were then transferred onto PVDF membranes, blocked with 5% skim milk in TBST (25 mM Tris, pH 7.4, 137 mM NaCl, 2.7 mM KCl, 0.1% Tween20) for 1h at room temperature, and then incubated for 1 h at room temperature or overnight at 4 °C with the appropriate antibody in 2.5% skim milk. After 1 h of incubation at room temperature with the appropriate secondary antibodies, the blots were developed using ECL reagents (Amersham Biosciences).

3.2.8 Purification of Flag-tagged Ku70

Appropriate amount of the ANTI-FLAG M2 affinity gel was washed twice with 1 ml cell lysis buffer. The clear cell lysate containing 1-10mg total protein was diluted 2 times by 20mM HEPE and then added to the washed beads. After incubation with gentle rotation at 4°C overnight, the affinity gel was transferred to a clean Bio-spin column (Bio-Rad) and then washed by 20 column volumes of TBS. Ku70 was eluted by 3X FLAG elution solution (100 ng/µl 3X FLAG peptide in TBS).

3.2.9 TAP Tag Purification

Appropriate amount of IgG sepharose 6 fast flow (Amersham) were washed 3 times with lysis buffer. The clear cell lysate containing 5-20mg total protein was diluted 5 times by 20mM HEPES until the concentration of igepal CA-630 to 0.1% and then added to the washed beads. After incubation with gentle rotation at 4°C for 4 hours, the IgG beads were washed 3 times with 1ml diluted lysis buffer and 3 times with TEV buffer (10mM HEPES-KOH pH8.0, 150mM NaCl, 0.1% igepal CA-630, 0.5mM EDTA, 1mM DTT). The IgG beads were then resuspended in 300µl TEV buffer containing 150 units TEV protease (Invitrogen) and incubated overnight at 4°C. The IgG beads was spined down and the supernatant was transferred to a fresh tube. The post-cleavage IgG beads were then washed 3 times by 300µl calmodulin binding buffer (10mM ß-mercaptoethanol, 10mM HEPES-KOH pH8.0, 150mM NaCl,

0.1% igepal CA-630, 1mM MgOAc, 1mM imidazole, 1mM $CaCl_2$). All the washed supernatants were combined with post-cleavage supernatant and transferred to the tube containing the calmodulin-sepharose 4B (Amersham) which was washed 3 times by 1ml calmodulin binding buffer and incubate 120 minutes with gentle agitation at 4°C. The Calmodulin beads were then transferred to a clean Bio-spin column (Bio-Rad) and washed 3 times with 10-20 volumes calmodulin binding buffer followed by 2 times with 10-20 volumes calmodulin rinsing buffer (50mM ABB pH 8.0, 75mM NaCl, 1mM MgOAc, 1mM imidazole, 2mM $CaCl_2$). Two volumes of calmodulin elution buffer (50mM ABB pH 8.0, 25mM EGTA) was added into the column and incubate for 1h with rotation at 4°C. The final eluate was collected for future analysis.

3.2.10 Crosslinking

Three crosslinkers EDC, DMP, formaldehyde were used to crosslink the protein complex during TAP tag purification before or after the protein complex was cleaved with tobacco etch virus protease (TEV) from IgG beads. Formaldehyde was added at the concentration of 0.1% and incubated at 4°C for 5 min with slow mixing. The reaction was quenched with 125 mM glycine, pH 7.0. Fresh prepared 6.54mM EDC in crosslink buffer (20 mM MES–NaOH pH 6.0, 100 mM NaCl, 2 mM $MgCl_2$, 0.5 mM EDTA pH 8.0, 10% glycerol) was added to the beads to crosslink the protein complex and reacted

for 1 hours at 4°C. 2-mercaptoethanol (final concentration of 20 mM) was added to quench the reaction. DMP was dissolved in TEV buffer at the final concentration of 5.78mM and added to the bead to crosslink the protein complex. After 1h incubation at 4°C, the reaction was stopped by adding Tris or glycine at a 20-50 mM final concentration.

3.2.11 Silver Staining

After electrophoresis, the gels were fixed in fixing buffer (40% Ethanol, 10% Acetic acid, 50% H_2O) and stained using SilverQuest Silver Staining Kit (Invitrogen) according to the manufacturer's instructions.

3.2.12 Cell Cycle Analysis

Cells were synchronized to M phase by being treated with 40ng/ml nocodazole for 12h and then trypsinized and replanted to release cells from nocodazole (0h). The cells were then observed under confocal microscope at different time points and collected for flow cytometry analysis.

3.2.13 Propidium Iodide Staining of Cells for FACS Analysis

Cell pellet was washed with PBS supplemented with 1% serum. Chilled 80% ethanol was added very slowly to the pellet while vortexing to prevent clumbing. Cells were then left in the fixative at 4°C for 24h. After

fixing, cells were spinned down and washed with PBS supplemented with 1% serum. Propidium Iodide/ RNaseA solution was added to resuspend cells and incubated at 37°C for 30 minutes in dark. Cells are then ready for FACS analysis.

3.2.14 MTS Assay

Cells were seeded at 20,000 cells/well in 96-well plates and treated with a varity of HDAC inhibitors in different concentrations for 60h. Cell viability was quantitated with an MTS assay as per the manufacturer's instructions (Promega, Madison, Wis).

3.2.15 Sample Preparation for Mass Spectrometry

3.2.15.1 In Gel Digestion

3.2.15.1.1 Gel Bands

The protein complexes recovered from TAP purification were fractionated on a 4-12% SDS-PAGE gradient gel and stained by silver staining. Gel bands were directly excised from the gel and destained. After wash by water followed by 100% ACN, gel bands were reduced by 10mM DTT at 60°C for 1h and alkylated by 55mM IAA at room tempeture in dark for 1h. The gels were then washed by 100mM ABB and 100% ACN and dehydrated in a speed-Vac or Freeze drier. Trypsin (Promega) was employed (100–200

ng/digestion in 100mM ABB) to digest the sample overnight (12-16h) at 37°C. After digestion, the peptides were extracted from the gel slices by adding 50% acetonitrile, 5% formic acid and sonicating for 30 min. This aqueous samples were concentrated in a speed-Vac and desalted using C18 ZipTips (Millipore, Billerica, MA) according to the manufacturer's instruction.

3.2.15.1.2 Gel Section

The protein complexes recovered from TAP purification were fractionated on a 4-12% SDS-PAGE gradient gel. Gels were divided into several sections according to the molecular weight across the entire separation range of each lane and sliced into small pieces to sample all potential interacting proteins without bias with respect to size and relative abundance. Cut bands were reduced, alkylated, and digested as described previously, and peptides were sequenced by tandem mass spectrometry (LC-MS/MS).

3.2.15.2 Solution Phase Digestion

The protein complexes recovered from TAP purification were freeze-dried and redissolved in 50µl MilliQ water. The concentration of the proteins was estimate by Bradford method. DTT was added into the solution at the final concentration of 10mM and incubate at 60°C for 1h to reduce the sample. IAA was then added to a final concentration of 20mM to alkylate the proteins and incubate for another 1h at room tempeture in dark. Trypsin was employed

at the amount ratio of 1:50 (trypsin : protein) to digest the sample overnight (12-16h) at 37°C. After digestion, 0.1% TFA was added to terminate the reaction. Peptides were sequenced by tandem mass spectrometry (LC-MS/MS).

3.2.15.3 Empore Disk Extraction

Described in chapter 2

3.2.16 Mass Spectrometry Analysis

Tryptic peptides were analyzed with a Finnigan Surveyor HPLC system coupled online to a LTQ ion trap mass spectrometer (Thermo Electron, San Jose, CA) equipped with nano-ESI source. A C18 nanocolumn (i.d. 75 μm, length 100 mm, tip 15 μm, New Objectives, Woburn, MA) was packed with Magic C18 particles (particle size 5 μm, pore size 150 Å, Michrom BioResources, Auburn, CA). The flow rate through the column was 200 nL/min. A water-acetonitrile gradient was employed with both mobile phases containing 0.1% formic acid. The gradient used was 5%-50% acetonitrile over 60 min. The injection volume was 100 μL. The electrospray mass spectra were acquired at a voltage of 1.8 kV, ion transfer tube temperature of 180 °C, collision gas pressure of 0.85 mTorr and normalized collision energy of 35% for MS2. Ion selection threshold was set to 500 counts for initiating MS2 while activation q was set to 0.25 and activation time to 30 ms. The MS scan range was 400 - 2000 Da and mass spectra were collected in data-dependent

MS2 mode in which five of the highest intensity peaks in each MS scan were chosen for CID, with an isolation width of 3 Da.

3.2.17 Data Analysis

MS/MS data were search with IPI human database by MASCOT (Matrix Science, London) database search engine. Database search was performed with carbaminomethylation (Cys) set as fixed modification, oxidized methionine (+16 Da) and acetylatioin as viable modifications. Missed cleavages of trypsin were set as two. Peptide charges were confined to 1+, 2+ and 3+ and mass tolerance were set to 2.0 Da for peptide precursor ion and 0.8 Da for fragment ions, respectively.

3.3 Results

3.3.1 Generation of Stable Mammalian Cell Lines Expressing GFP-tagged, FLAG-tagged and TAP-tagged Ku70

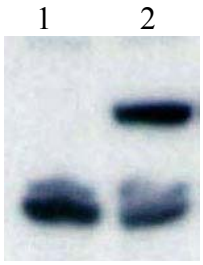
The plasmid pWay20-GFP-Ku70 was kindly provided by Dr. J. Donald Capra. For generating stable transfected cell line, lipofectamine2000 (Invitrogen) was used to transfect the plasmid to HEK293 cell line. Cells were selected by 500 μg/ml Geneticin(GIBCO). The selected clones were observed under fluorescent microscope to check the cell purity and Ku70 expression level. We also constructed the TAP-tagged Ku70 pcDNA3 expression vector

and FLAG-tagged pCMV expression vector. The pcDNA3 vector contains a Human cytomegalovirus immediate-early (CMV) promoter which allows high-level expression in mammalian cells. This plasmid was transfected into FreeStyle 293-F cells by using 293 fectin (Invitrogen). FreeStyle 293-F cell line is a variant of the 293 cell line that has been adapted to suspension growth in FreeStyle 293 expression medium. It can grow in big volume to high density which makes it easy and convenient to harvest a large amount of cell at one time. It also can be cultured in complementary Minimum Essencial Medium (GIBCO) as adherent cells at 37°C in 5% CO2. Normal stable transfection methods for suspension cell lines cannot ensure the purity, eg. diluting and selecting through several generations by appropriate selective medium, or are time consuming, eg. plating single-cell cloning into 96-well plates. To prevent the above shortcomings, we transfected 293-F cells in suspension status and incubate for 72 h to allow for expression of the resistance gene. Then we diluted the cell density and culture them in adhesion status with 750 μg/ml Geneticin. The concentration of Geneticin was decided by a dose-response test of Geneticin for 293-F cells. All normal 293-F cells were killed under the concentration of 750 μg/ml Geneticin. After culture for several weeks, single colonies were picked and tested by Western blot. According to western blot analysis, we found pcDNA3-NTAP-Ku70 vector expressing Ku70 at close to endogenous levels (Figure 3.1).

A

B

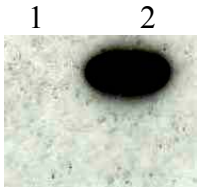
C



← TAP-Ku70

← Ku70 (endogenous)

D



← TAP-Ku70

E

← GFP-Ku70

← Ku70 (endogenous)

F

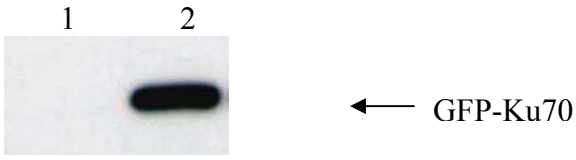


Figure 3.1. Generation of mammalian stable cell lines

A. Schematic diagram of the structure of recombinant NTAP-tagged Ku70.

B. Schematic diagram of the structure of recombinant GFP-tagged Ku70.

C. Western blot analysis of Ku70 expression using anti-human Ku70 antibody (Santa Cruz Biotechnoolgy). Each lane was loaded 30 μg total cell lysates. Both NTAP-tagged Ku70 and endogenous Ku70 were detected. 1, control-293F cells transfected with empoty vector. 2, 293F cells transfected with pcDNA3-NTAP-Ku70.

D. Western blot analysis of NTAP-Ku70 expression using Normal mouse IgG antibody (Upstate) which can recognize Protein A in TAP tag. Endogenous Ku70 can not be recognized by this antibody.

E. Western blot analysis of Ku70 expression using anti-human Ku70 antibody (Santa Cruz Biotechnoolgy). Each lane was loaded 50 μg total cell lysates. Both GFP-tagged Ku70 and endogenous Ku70 were detected. 1, control-293F cells transfected with empoty vector. 2, 293F cells transfected with pWay20-GFP-Ku70..

F. Western blot analysis of GFP-Ku70 expression using anti-GFP antibody (Santa Cruz Biotechnoolgy). Only GFP-tagged Ku70 was detected. Endogenous Ku70 can not be recognized by anti-GFP antibody.

3.3.2 Purification of The Ku70 Complex

Traditional immunoprecipitation (IP) was used to purify Ku70 complex in our first attempt. The whole cell lysate was precleared and incubated with anti-GFP antibody and then incubated with protein G sepharose. After wash, proteins which still bind to beads were eluted by 100mM glycine, pH 2.0. The presence of Ku70 was demonstrated by western blot analysis (Figure 3.2, A). But when the eluted proteins were separated by SDS-PAGE and silver stained, we could observe a lot of non-specific binding proteins. Even the control lane showed many bands that were from non-specific binding proteins to the solid support resin or antibody (Figure 3.2, B). After we increased the concentration of NaCl in washing buffer, the nonspecific binding still had not been reduced (Figure 3.2, C). The result indicated that normal IP is suitable for procedures such as western blotting, where the presence of one or more specific proteins is probed and background is not revealed. Therefore, a more efficient mass spectrometry compatible strategy is needed to purify the Ku70 protein complex.

In this research, we used TAP tag purification strategy to purify the Ku70 complex. The clear cell lysate were prepared from the stable transfected 293F cells and incubated with IgG sepharose 6 fast flow. After washing by diluted lysis buffer and TEV buffer for 3 times respectively, the IgG beads were then incubated with TEV protease overnight at 4°C. The post-cleavage supernatant was further incubated with calmodulin-sepharose 4B for 120 minutes with gentle agitation at 4°C. After washing the beads with 50 volumes

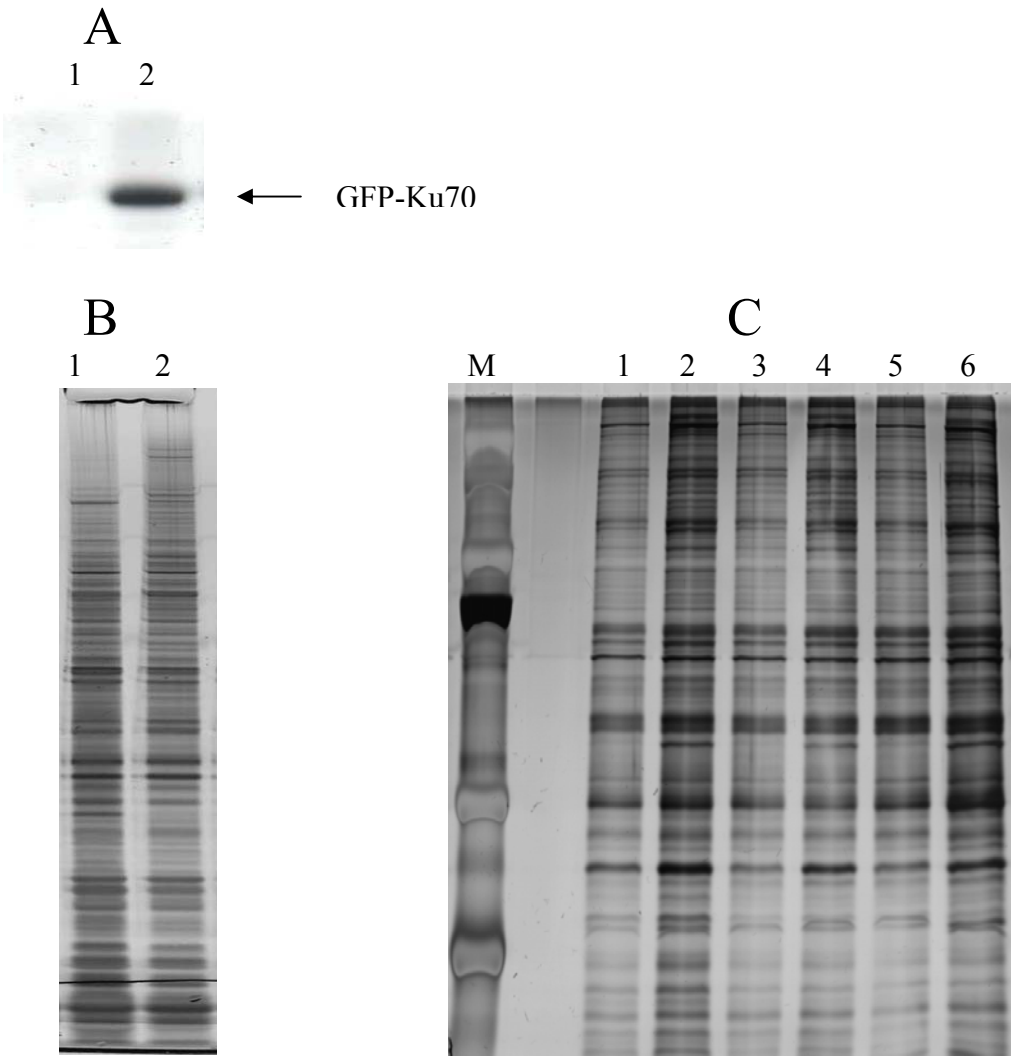


Figure 3.2 Traditional IP pull down

A. Western blot analysis of IP pull down eluate. IP: anti-GFP. IB: anti-Ku70. 1, control-293F cells transfected by empty vector. 2, 293F cells transfected by pWay20-GFP-Ku70.

B. Silver staining of IP pull down eluate. IP: anti-GFP. 1, control-293F cells transfected by empty vector. 2, 293F cells transfected by pWay20-GFP-Ku70.

C. Silver staining of IP pull down eluate. IP: anti-GFP. 1, 3, 5, control-293F cells transfected by empty vector, 300mM, 400mM and 500mM NaCl in washing buffer respectively. 2, 4, 6, 293F cells transfected by pWay20-GFP-Ku70, 300mM, 400mM and 500mM NaCl in washing buffer respectively.

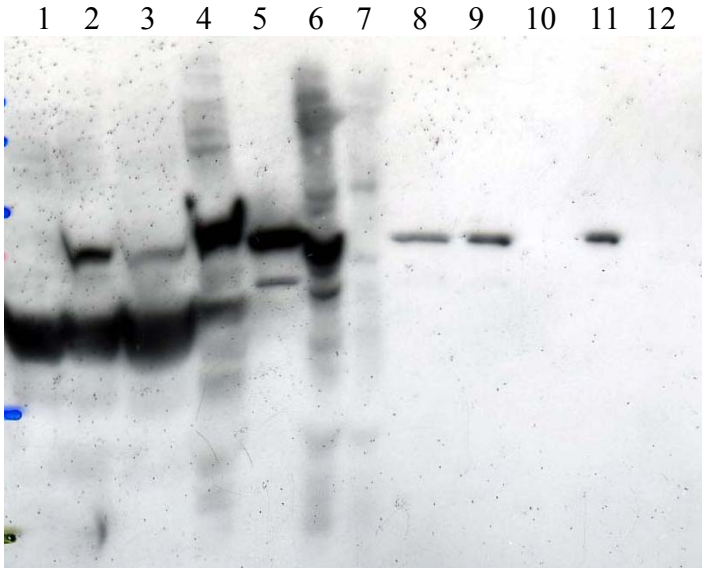
of calmodulin binding buffer followed by 40 volumes of calmodulin rinsing buffer, the proteins bound to the calmodulin beads were eluted by elution buffer. Each step was verified by western blot analysis to monitor the presence of Ku70 in the eulate (Figure 3.3, A). As depicted in Figure 3.3, the first step of TAP tag purification with IgG column reduced the complexity of the protein mixture. As a result, Ku70 protein complex was significantly enriched. Second step purification with calmodulin beads further greatly reduced the nonspecific proteins compared with the eluate from normal IP pull down. The 70 kDa Ku70 was the major protein in the eluates from calmodulin beads, suggesting the two-step tandem purification procedure can efficiently isolate TAP-Ku70 and its associated proteins from cultured mammalian cells.

We also used anti-FLAG M2 affinity gel to purify the FLAG-tagged Ku70 and the associated proteins. This one step affinity purification was also very effective. But the protein profile was a little different from the protein profile of TAP-tag purification (Figure 3.3, D).

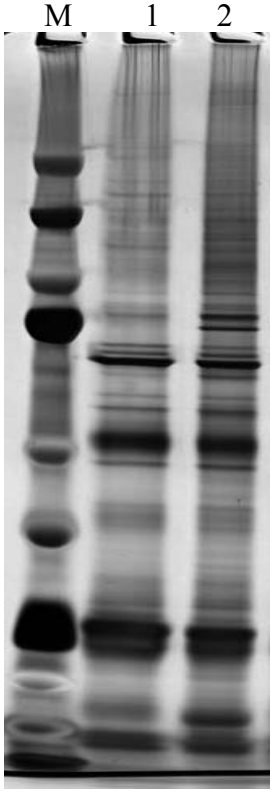
3.3.3 *In Vitro* Closslinking of Ku70 Complex

Although two-step tandem purification greatly reduced the background of protein complex, it also would destroy some weak protein-protein interactions. Many biologically important interactions between proteins are transient or relatively weak. They are possible to be disturbed during the purification procedure because of the high dilution of the protein solutions in

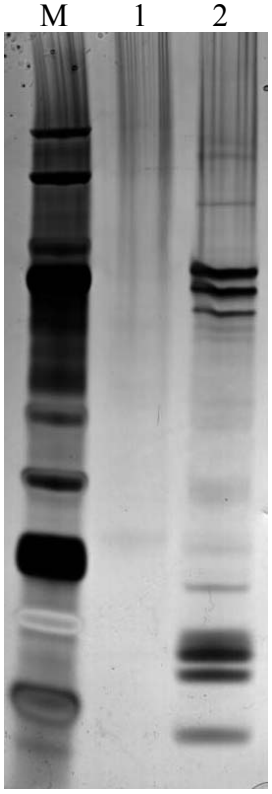
A


1 2 3 4 5 6 7 8 9 10 11 12
TAP-Ku70
Ku70 (endogenous)

B


M 1 2

C


M 1 2
250kD
150kD
75kD
50kD
25kD
10kD

D

Figure 3.3 purification of TAP-tagged and FLAG-tagged Ku70 and its associated proteins

A. A portion of proteins in each step of purification were separated using 4-12% SDS-PAGE and verified by western blot. IB: anti-Ku70
1, Control-293F cells, whole cell lysate
2, 293F cells transfected by pcDNA3-NTAP-Ku70, whole cell lysate
3, Cell lysate after incubating with IgG beads
4, IgG beads after incubating with cell lysate
5, Post-cleavage supernatant after cleavage with TEV protease
6, IgG beads after cleavage with TEV protease
7, Control-IgG beads after cleavage with TEV protease
8, Post-cleavage supernatant after incubating with Calmodulin beads
9, Eluate
10, Control-eluate
11, Calmodulin beads after elution
12, Control-Calmodulin beads after elution

B. Silver staining of post-cleavage supernatant (1-step).
M, Maker. 1, Control. 2, 293F cells transfected by pcDNA3-NTAP-Ku70.

C. Silver staining of TAP tag purification eluate (2-step).
M, Maker. 1, Control. 2, 293F cells transfected by pcDNA3-NTAP-Ku70.

D. Silver staining of the eluate of FLAG-tagged Ku70.
M, Maker. 1. Control. 2. 293F cells transfected by pCMV2B-FLAG-Ku70

the later steps. Transient and weak protein interactions can be preserved by covalently crosslinking protein complexes *in vivo* prior to purification. But crosslinking proteins in the very early stage may increase the chance to pull down nonspecific proteins. We therefore crosslinked the protein complex *in vitro* between the two steps of purification to facilitate the identification of protein complexes that are formed by less stable or transient interactions and reduce the background as many as possible at the meantime.

We tested three crosslinkers, 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC), dimethyl pimelimidate (DMP) and formaldehyde, to crosslink the protein complex. Formaldehyde is a widely used crosslinker. It is usually applied to crosslink protein and DNA in chromatin IP. It crosslinks proteins by forming methelene bridges between reactive groups like amido, guanidino, thiol, phenol, imidazole and indolyl groups. But it crosslinks proteins nonspecifically, it can crosslink protein-protein or protein-DNA within a ~20A radius. EDC is a zero-length crosslinking agent used to couple carboxyl groups to primary amines. This crosslinker has been used in diverse applications such as forming amide bonds in peptide synthesis etc. DMP is also a common crosslinker and is used to crosslink antibody to Protein G or A Sepharose. It possesses two identical groups which can react with primary amine groups to form stable covalent bonds. One advantage of this crosslinker is that it is pH-dependent. It will react rapidly with amines at alkaline pH values (pH 8-10) and is reversible at high pH values.

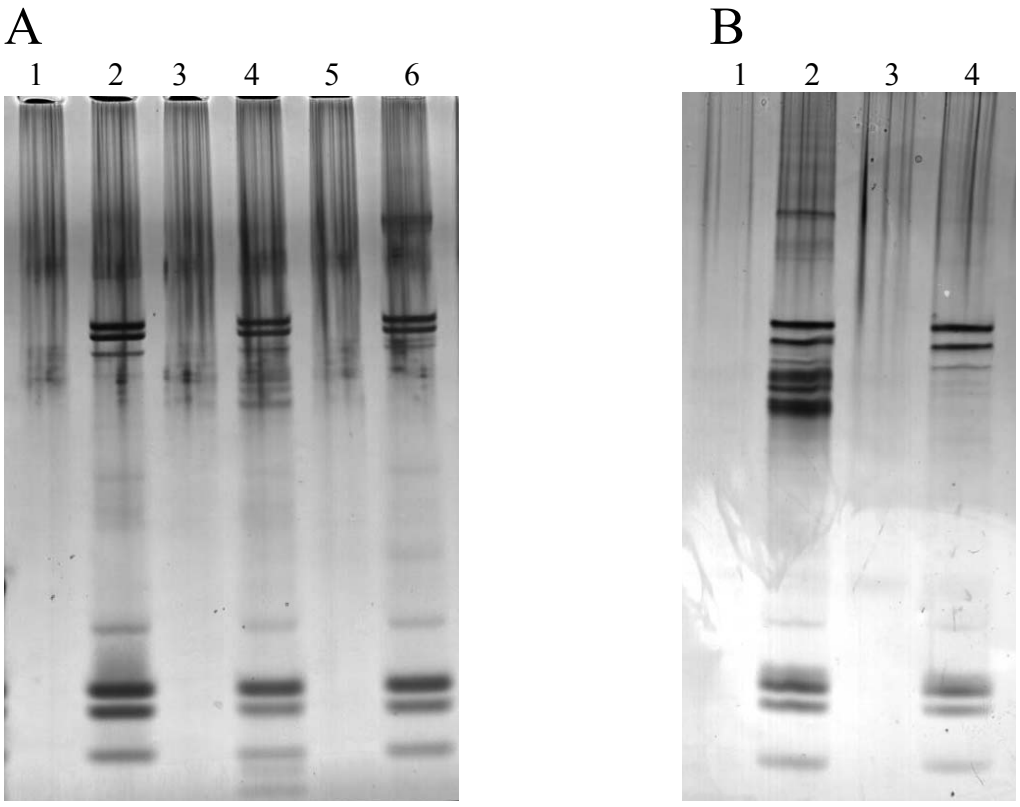


Figure 3.4 Crosslinking of Ku70 complex combined with TAP tag purification.

A. The Ku70 complex was crosslinked after it was cleft from IgG beads.
1, Control. 2, 293F-NTAP-Ku70.
3, Control-293F crosslinked by EDC. 4, 293F-NTAP-Ku70 crosslinked by EDC.
5, Control-293F crosslinked by DMP. 6, 293F- NTAP-Ku70 crosslinked by DMP.

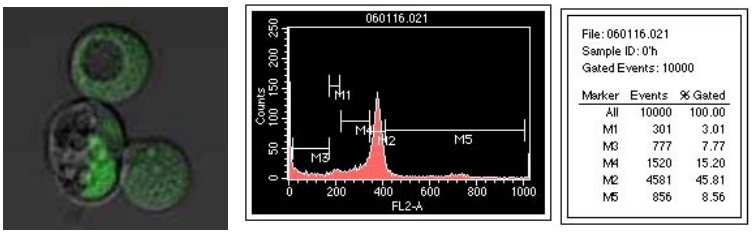
B. The Ku70 complex was crosslinked when it still bound to IgG beads.
1, Control-293F crosslinked by DMP. 2, 293F- NTAP-Ku70 crosslinked by DMP.
3, Control-293F crosslinked by Formaldehyde. 4, 293F- NTAP-Ku70 crosslinked by Formaldehyde.

The Ku70 complex was crosslinked at two consecutive stages respectively. We first corsslinked it after the first-step purification when it still bound to IgG beads. Considering that there were some nonspecific proteins binding to the beads which were possible to be crosslinked to Ku70 complex, we carried out another experiment to crosslink the complex immediately after it was cleaved from IgG beads. The result showed that only DMP worked in both processes. The failure of EDC in first process and formaldehyde in second process may because of the coverage of the TEV cleavage site or calmodulin binding domain by the crosslinked proteins. As shown in Figure 3.4, the backgrounds were still clear and more proteins were pulled down after crosslinking compared with the eluate without crosslinking. But the protein profiles were a little different by using different crosslinkers in different stages. It may be because of the preferential difference of each corsslinker for target proteins according to their own crosslinking mechenisms.

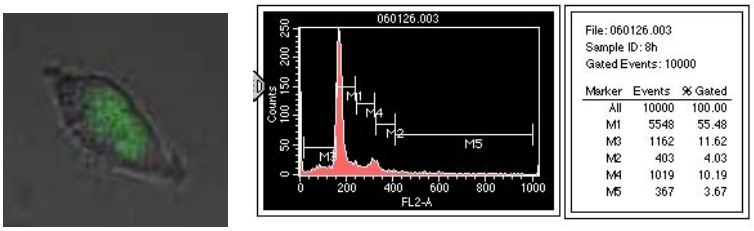
3.3.4 Enrich of Cytoplasmic Pool of Ku70

Ku70 is a predominant nuclear protein. The HEK293 cells stable expressing GFP-tagged Ku70 were observed under confocal microscope in different time points of the cell cycle. Ku70 stays in nucleus during the whole cell cycle. It has been observed to transiently translocated to cytoplasm in M phase (Figure 3.5). Although the cytoplasmic pool of Ku70 is less abundant, it has been found to regulate cellular apoptosis. However, the functions of Ku70

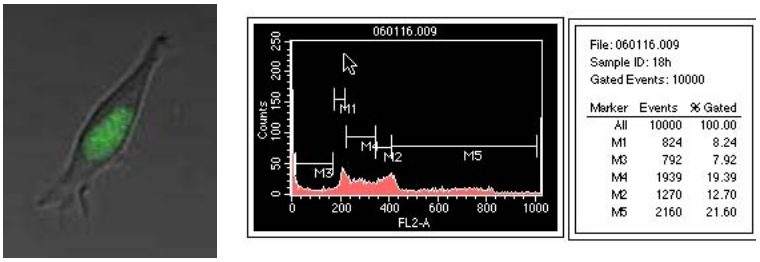
A



B



C



D

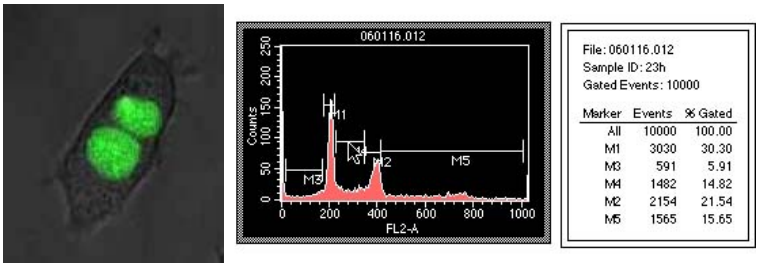


Figure 3.5 Location of Ku70 during Cell cycle

HEK293 cells stablely expressing GFP-tagged Ku70 were synchronized by 40ng/ml nocodazole for 12h and then trypsinized and replanted to release cells from nocodazole (0h). The cells were then observed under confocal microscope at time point 8h, 18h, 23h. Flow cytometry analysis showed that most of cells in each time point were in M phase (0h, A), G1 phase (8h, B), S phase (18h, C) and G2/M phase (23h, D)

in cytoplasm largely remain unknown. In order to pull down more cytoplasmic partners of Ku70, we enriched cytoplasmic Ku70 by separating cytoplasmic proteins from nuclear proteins, arresting cells in G2/M phase or constructing Ku70 mutant K533A to prevent it from translocating to nucleus. The lysates were purified by TAP procedure. Ku70 was demonstrated to exist in cytoplasm by western blot analysis (Figure 3.6, A). The profiles of eluates from cytoplasmic and nuclear proteins were different (Figure 3.6, B). The region composed of 18 amino acid residues (positions 539-556) of Ku70 was reported to be the intrinsic nuclear localization signal (NLS) (Koike, 2002). It will help Ku70 translocate to nucleus by binding to specific NLS receptors. We constructed 2 mutants which replaced glutamic acid in position 548 and lysine in position 553 by alanine. The change of amino acids will alter the configuration of this region so that to prevent the translocation of Ku70 from cytoplasm to nucleus. HEK293 cells transfected by GFP-tagged mutants E548A and K553A were observed under confocal microscope. As shown in Figure 3.7 (B, C), K553A mutant prevented Ku70 translocation to the nucleus while E548A mutation could not stop the translocation of Ku70 to the nucleus.

3.3.5 Reduction of Protein Complexity in Eluate Submitting to Mass Spectrometry

The number of proteins identified in mixture usually depends on the complexity of mixture. In general, the simpler the sample is, the more accurate

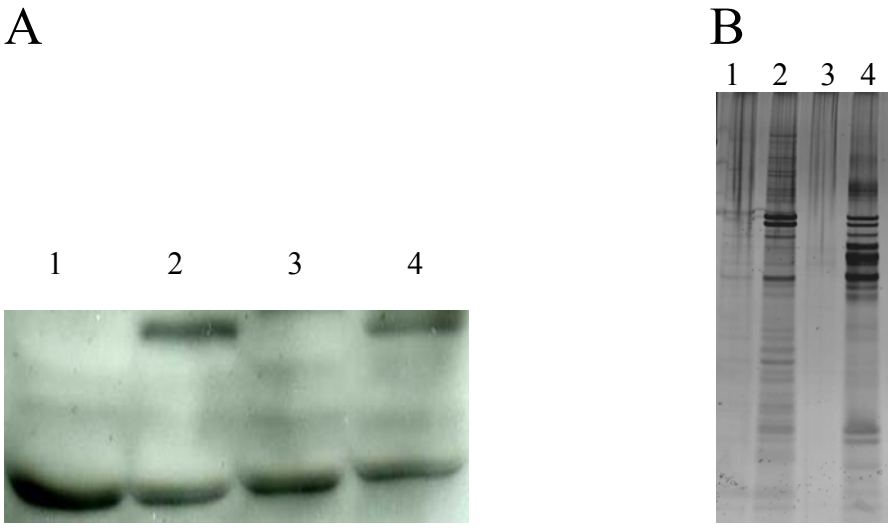


Figure 3.6 Cytoplasmic and nuclear pool of Ku70

A. 293F cells expressing NTAP-tagged Ku70 were lysed. cytoplasmic proteins were separated from nuclear proteins. Ku70 was detected by antibody in raw cell lysate.

B Raw cell lysates containing cytoplasmic proteins or nuclear proteins were purified by TAP procedure. The eluates were separated in 4%-12% SDS-PAGE gel.

1, Control-293F, cytoplasmic fraction. 2, 293F-NTAP-Ku70, cytoplasmic fraction
3, Control-293F, nuclear fraction. 4, 293F-NTAP-Ku70, nuclear fraction.

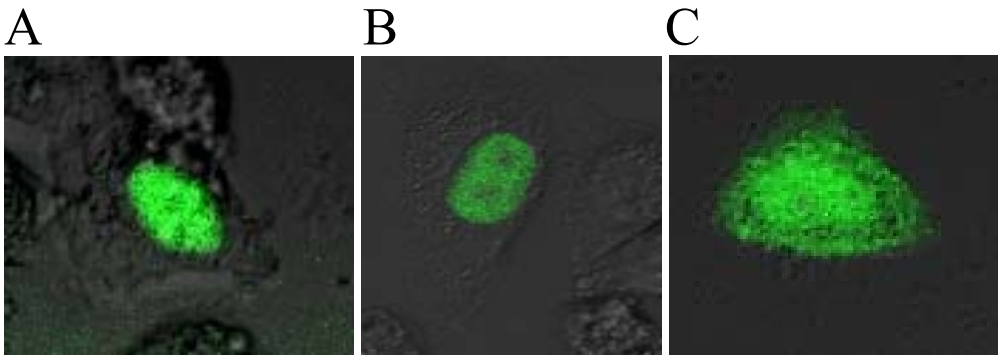


Figure 3.7 Location of Ku70 and its mutants

HEK293 cells transfected by GFP-tagged wild type Ku70 and its mutants E548A and K553A were observed under confocal microscope.
A. wild type Ku70 B. E548A C. K553A

the mass spectrometry analysis will be. Fractionation of a complex protein mixture to multiple simpler fractions is the strategy to improve the detectability of low abundant proteins. In order to reduce the complexity of proteins in eluate, we eluted proteins bound to calmodulin beads by a series of NaCl gradient elution buffers (50mM ABB pH 8.0 with 250mM, 500mM or 750mM NaCl). Proteins were eluted gradually along with the increase of salt concentration (Figure 3.8).

3.3.6 Identification of Ku70 Complex by Mass Spectrometry

The shotgun proteomics approach was applied to analyze the eluate recovered from the TAP purification. The complex was first digested by trypsin. Resulting peptides were then separated by a C18 nanocolumn and sequenced by tandem mass spectrometry. The LC gradient time was either 60 minutes or 180 minutes to improve the separation of peptides. MS/MS data were analyzed by MASCOT with IPI human database. In order to reduce the complexity of the mixture, we also separated the eluate from the TAP purification by 4%-12% SDS-PAGE. The lane was then divided into several equal size sections without any staining to prevent fixing of proteins in gel. Each section was subjected to in-gel digestion. The resulting peptide recovered from in-gel digestion was analyzed by LC-MS/MS.

We had run multiple pull-down experiments with different experimental conditions as described above. A group of proteins has been

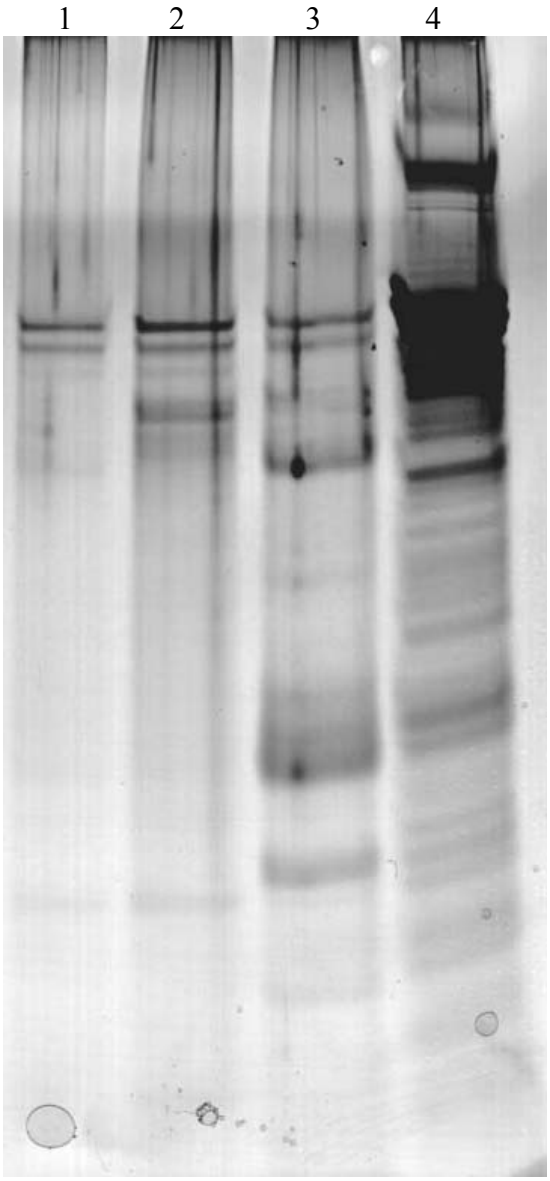


Figure 3.8 Elution of proteins by NaCl gradient elution buffer

1, 250 mM NaCl in elution buffer
2, 500 mM NaCl in elution buffer
3, 750 mM NaCl in elution buffer
4, EGTA elution buffer

identified in every experiment. Ku70 was identified with the highest score. Ku80 was the second most abundant protein identified in every pull down because Ku80 is known to form heterodimer with Ku70. DNA-dependent protein kinase catalytic subunit, another member protein of DNA-PK, was also found among the most abundant proteins in the complex. All histone proteins were detected with very high scores. Many subunits of 40S ribosomal protein and 60S ribosomal protein were also detected with high scores. We also detected some subunits of tubulin, nuclear ribonucleoproteins and heat shock proteins. Besides, 94 other proteins were found with peptides score 50 or above. (Table 3.1, 3.2). In total, we detected 151 proteins interacting with Ku70 under different conditions.

The IPI accession number for each protein was then converted to gene symbol with IPI cross reference table. The 151 genes were then inputted into Panther Database to classify these proteins according to their molecular functions, biological processes or pathways they are involved in.

All the 151 proteins in the complex are associated with 19 molecular functions, 19 biological processes and 23 pathways (Figure 3.9, 3.10, 3.11, Table 3.3, 3.4, 3.5). The proteins under each biological process and pathway were listed (Table 3.6, 3.7). Our result showed that Ku70 interacted with 81 proteins in the biological process of nucleoside, nucleotide and nucleic acid metabolism (Table 3.8). This is consistent with the known Ku70 biological functions in DNA repair, telomere maintenance and transcriptional regulation. As shown in tables, among these proteins, 15 proteins are involved in DNA

repair that is the major function of Ku70. Seventeen Ku70 interacting proteins were found to play roles in apoptosis, this is consistent with recent reports that Ku70 is a regulator of apoptosis. Furthermore, the postulated biological processes and pathways both suggested that Ku70 is an important protein in cell cycle, signal transduction and transcription regulation. Moreover, Ku70 also associated with proteins in p53 pathway. As p53 mutation and deregulation is involved in up to 50% of human tumors and p53 dramatically regulates the cellular response to ionizing radiation and chemotherapeutic drugs, the result strongly evidenced that Ku70 may play a pivotal role in cancer biology.

3.3.7 Validation of Search Result of Mass Spectrometry

Ku80 is a well-known Ku70 binding partner which forms heterodimer with Ku70 to participate in many biological processes. Ku80 was identified to be the second most abundant protein in the complex after Ku70. We had validated the Ku80 by carrying out a western blot analysis on the eluate from TAP purification. As shown in Figure 3.12 (A), Ku80 could be detected by Anti-human Ku80 antibody (Santa Cruz) in the eluate. This demonstrated that TAP tag purification couple to mass spectrometry analysis is a specific and sensitive method for protein complex characterization.

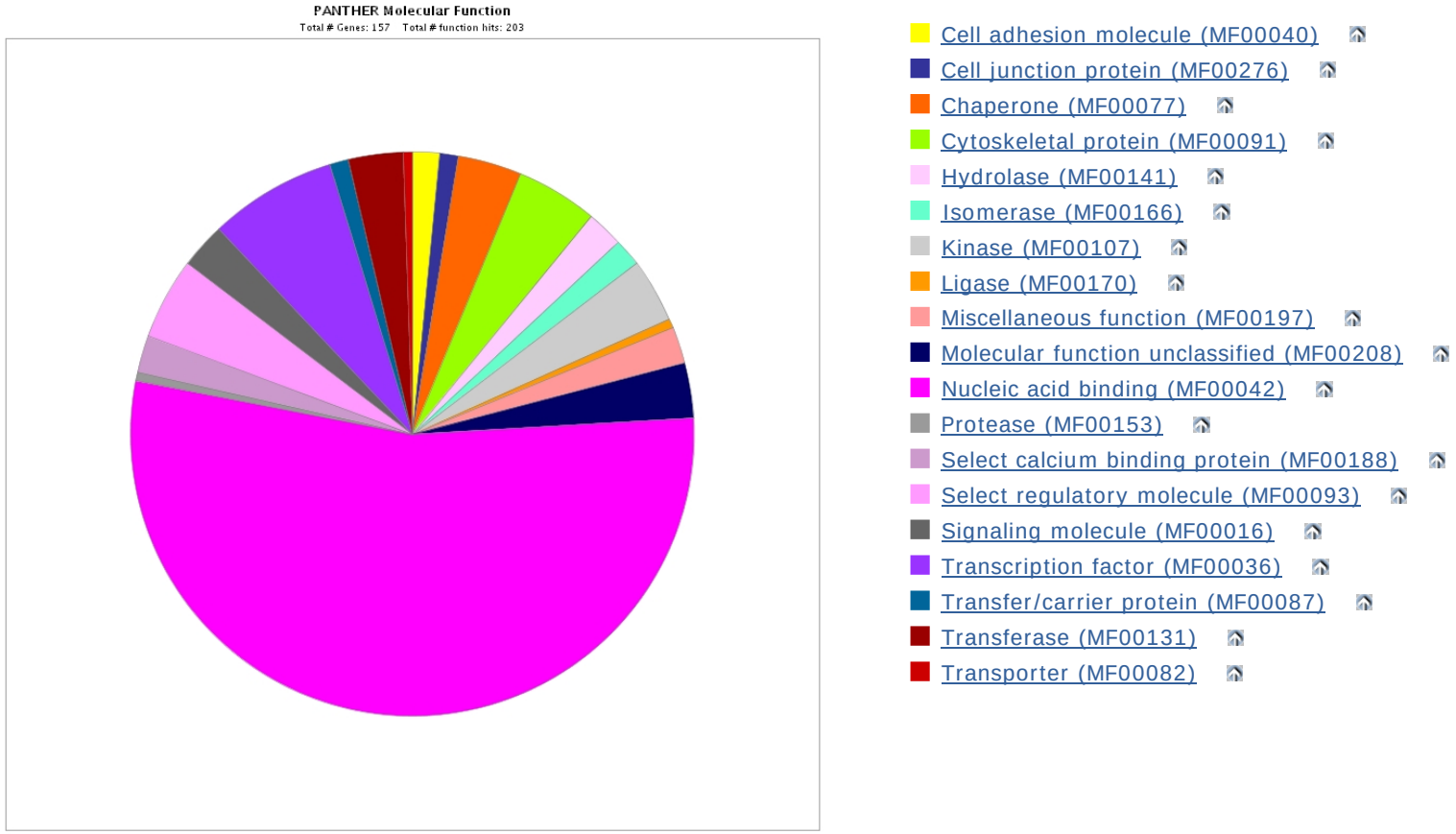


Figure 3.9 Molecular functions modulated by all complex proteins

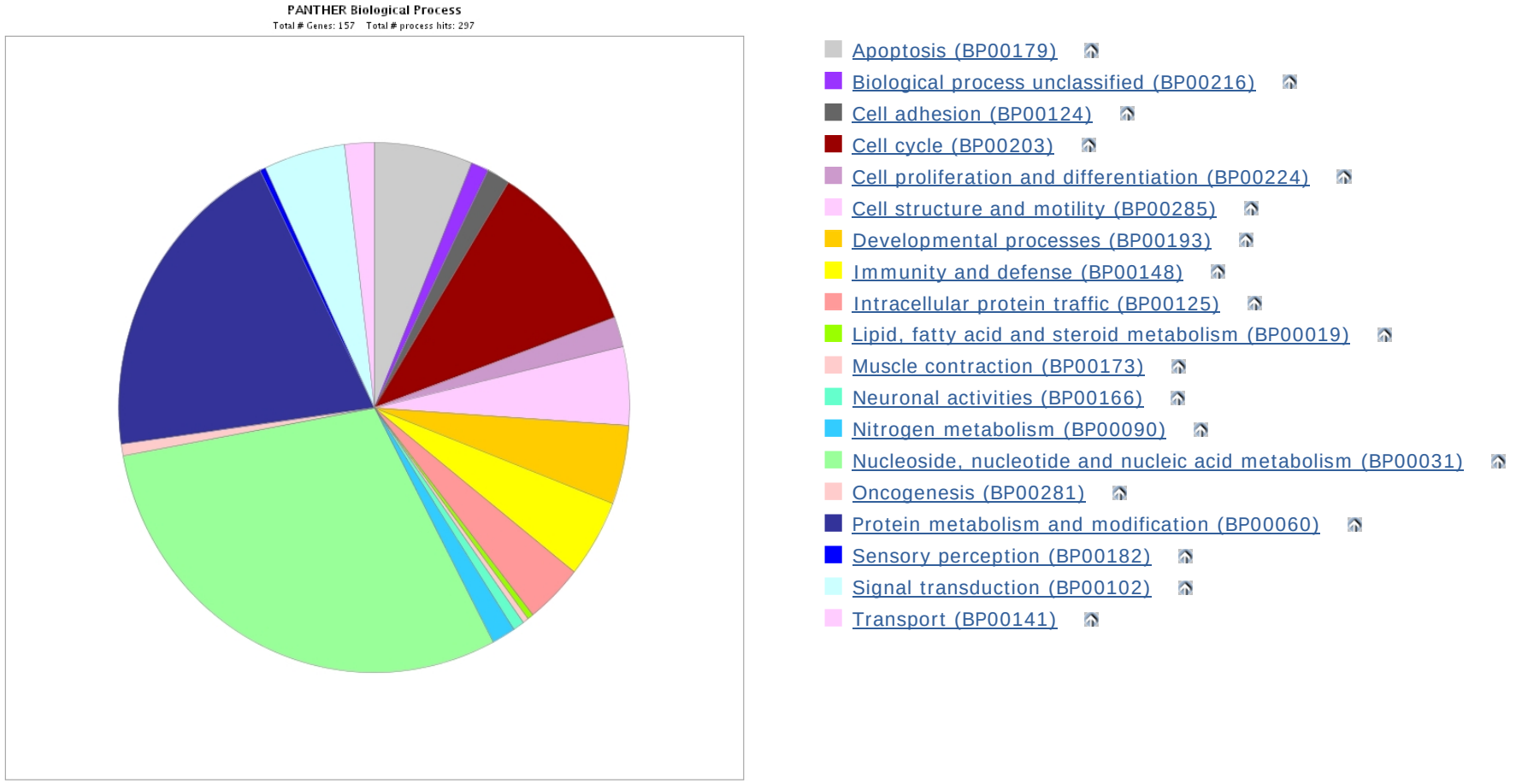
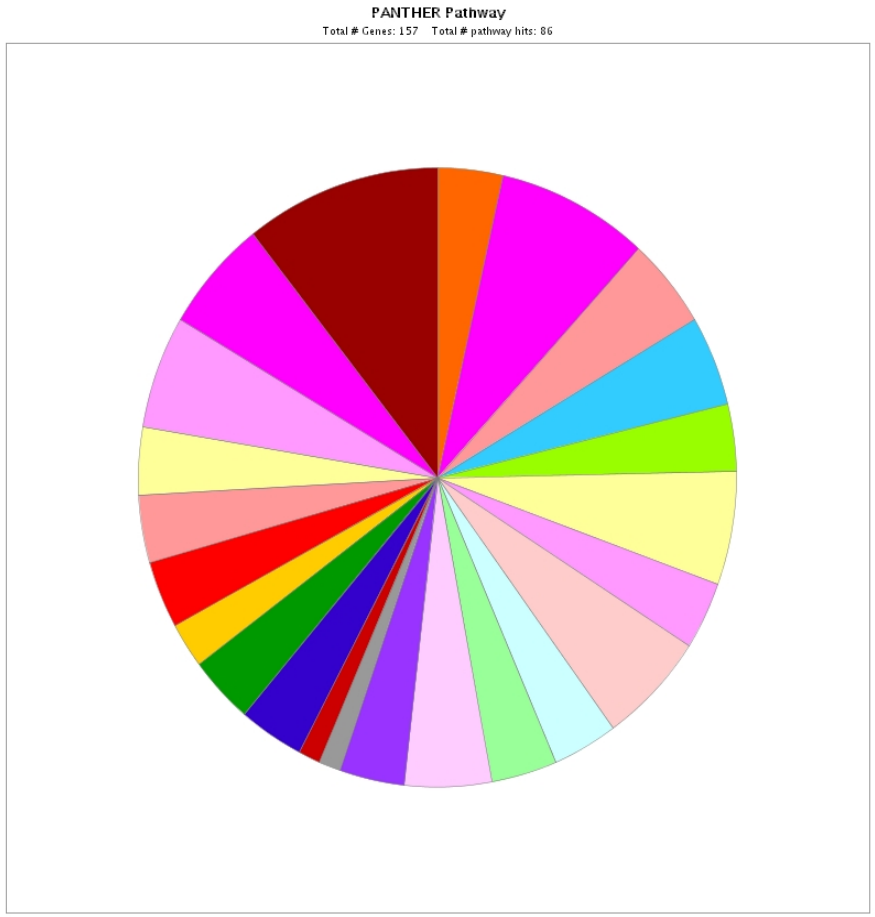


Figure 3.10 Biological processes modulated by all complex proteins



Angiogenesis (P00005)
Apoptosis signaling pathway (P00006)
DNA replication (P00017)
EGF receptor signaling pathway (P00018)
Endothelin signaling pathway (P00019)
FAS signaling pathway (P00020)
FGF signaling pathway (P00021)
Huntington disease (P00029)
Hypoxia response via HIF activation (P00030)
Inflammation mediated by chemokine and cytokine signaling pathway (P00031)
Insulin/IGF pathway-protein kinase B signaling cascade (P00033)
Interleukin signaling pathway (P00036)
Notch signaling pathway (P00045)
Oxidative stress response (P00046)
PDGF signaling pathway (P00047)
PI3 kinase pathway (P00048)
Parkinson disease (P00049)
Ras Pathway (P04393)
T cell activation (P00053)
VEGF signaling pathway (P00056)
Wnt signaling pathway (P00057)
p53 pathway feedback loops 2 (P04398)
p53 pathway (P00059)

Figure 3.11 Pathways modulated by all complex proteins

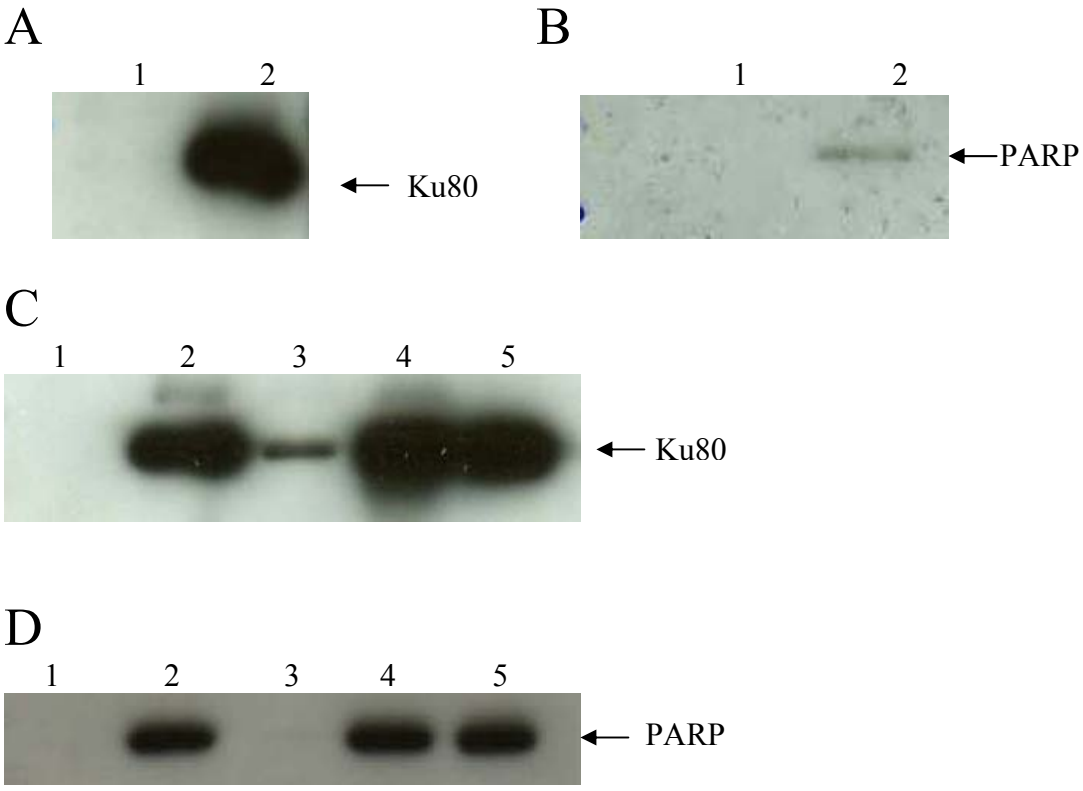


Figure 3.12

The eluate from TAP purification was separated by SDS-PAGE and then was transferred to a PVDF membrane. The exist of complex proteins was tested by antibodies.
A. IB: anti-Ku80. B. IB: anti-parp
1, Control-293F with empty vetctor. 2, 293F-NTAP-Ku70

The eluates from TAP purification of NTAP-tagged wild type Ku70 and three mutants: L385R, E548A and K553A were separated by SDS-PAGE and then were transferred to PVDF membrane.
C. Ku80 binding ability was tested. IB: Ku80.
D. PARP was tested. IB: anti-Parp.
1, Control-293F with empty vetctor. 2, 293F-NTAP-Ku70
3, 293F-NTAP-Ku70-L385R. 4, 293F-NTAP-Ku70- E548A.
5, 293F-NTAP-Ku70- K553A

3.3.8 Poly [ADP-ribose] Polymerase (PARP) May Interact With Ku70 Through Ku80

The poly [ADP-ribose] polymerase (PARP) in the eluate from TAP purification was also detected by anti-PARP antibody (Figure 3.12, B). To further explore how PARP interacts with Ku70, we constructed another Ku70 mutant L385R. The region from residue position 378 to 482 of Ku70 were considered to be the Ku80-binding domain (Koike, 2002). The mutation in L385 significantly impairs its ability to interact with Ku80 (Figure 3.12, C). We tested for PARP in the eluate from three Ku70 mutants: L385R, E548A and K553A and found that PARP could be detected in the eluate of E548A and K553A, but did not exist in the eluate of L385R (Figure 3.12, D). This suggested that PARP may not interact with Ku70 directly. It may interact with Ku80, and pull down by Ku70 through Ku80 because when Ku80 interacting domain of Ku70 was mutated, PARP could not be detected in the Ku70 complex.

3.3.9 Role of Ku70 in Apoptosis

Ku70 was reported as a mediator of Bax which induced apoptotic pathway. We also found several proteins in Ku70 complex function in apoptotic pathway. To investigate the effects of overexpression of Ku70 on cell death, we used HDAC inhibitors, Trichostatin A (TSA), a hydroxyamic acid derivative originally developed as an antifungal agent, and SAHA to treat

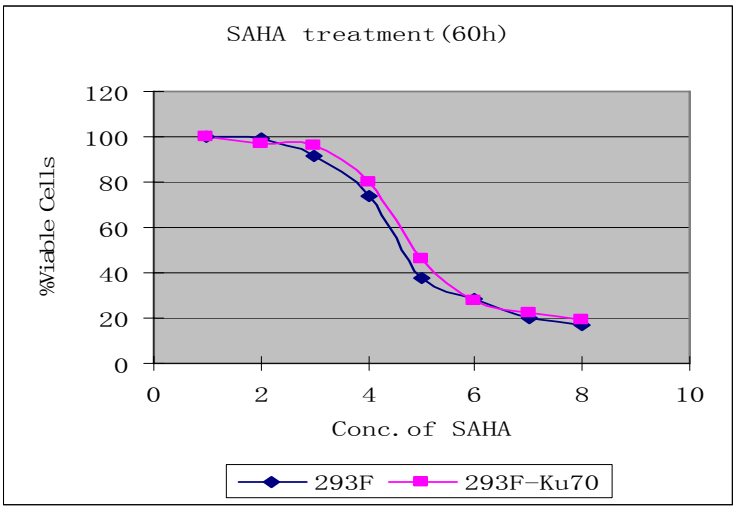
the HEK293F cell that overexpressed Ku70. Both compounds suppress cell growth, induce differentiation and apoptosis in different tumor types. As shown in Figure 3.13, overexpression of Ku70 did not induce cell death when cells were treated by these HDAC inhibitors. Although more Ku70 were acetylated by HDACIs in Ku70 overexpression cells which should release Bax to trigger apoptosis, the amount of nonacetylated Ku70 was also higher abundant than normal cells to sequester Bax. As a result, the ratio of nonacetylated and acetylated Ku70 remained unchange between normal 293F cells and HDACI treated Ku70 overexpressed 293F cells so that their cell death curves nearly overlapped.

3.4 Discussion

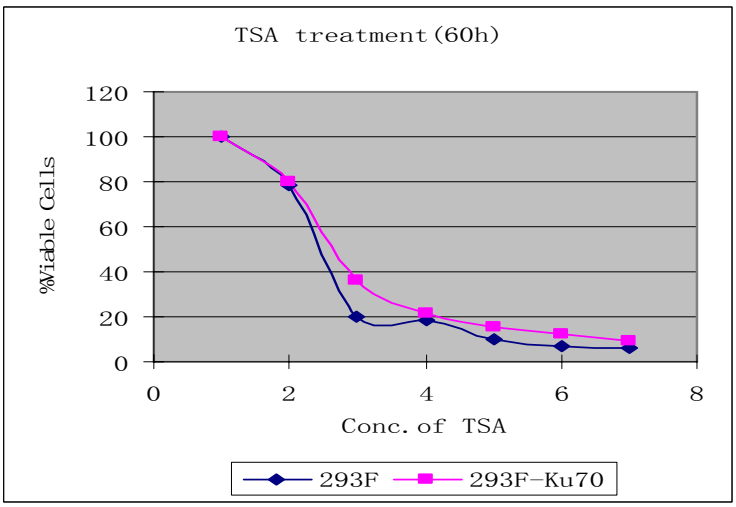
3.4.1 Affinity Purification

Nearly all proteins exert their molecular functions by interacting with one another to form a large protein complex. Elucidation of the composition of these protein complexes will help greatly to clarify protein function or pathways they involve in. It has increasingly become one of the focuses of proteomic research. To detect the protein-protein interactions in higher eukaryotes, methods that are capable of rapidly and efficiently purifying intact protein complexes will be required. In our study, we applied the one step FLAG-tag purification and two steps tandem affinity tag purification method to purity Ku70 complex in HEK293F cells. These epitope tag approaches were

A



B



SAHA(uM)	0.1%DMSO	0.1	0.5	1	2.5	5	7.5	10
TSA(uM)	0.1%DMSO	0.1	0.5	1	2.5	5	10	

Figure 3.13 HDACI induces cell death in 293F cells and Ku70 overexpressed 293F cells.

293F cells with empty vector and 293F cells stablely expressing GFP-tagged Ku70 were treated with various concentrations of SAHA or TSA as indicated for 60h. Cell viability was determined by MTS assay. Results are expressed as the percentage of viable cells compared with vehicle treated controls.

demonstrated to be efficient for Ku70 complex pull down. Nonspecific bindings were greatly reduced because of the specific and mild elution.

This generic tandem affinity tag purification method is not only limited to identify protein-protein interactions. It will also facilitate analysis and identification of post-translational modifications of proteins which we will describe and discuss in chapter 4.

Some papers reported that CBP tag did not work in their systems (Knuesel et al., 2003; Kim, S. C. et al., 2006). The reason may be the binding of endogenous calmodulin to CBP will block its ability to bind to calmodulin beads. It depends on the cell lines which have different expression levels of endogenous calmodulin and also the abundance of the target proteins. Our target protein Ku70 is relatively high abundant in HEK293 cell that reduces the interference of endogenous calmodulin.

Although the TAP tag greatly improved the specificity of purification, we encountered several problems. Some calcium binding proteins besides calmodulin in mammalian cells also bound to CBP. These proteins formed a feint that they were also the partners of Ku70. In order to avoid the interference of these nonspecific bindings, we abtained a protein list from the control sample which was transfected by the empty vector only with TAP tag domain. The detected non-specific proteins were removed from the Ku70 complex list. The interference of immunoglobulin is another problem in this system. High levels of immunoglobulin protein washed out from IgG beads can easily mask the lower abundance peptides derived from low stoichiometry

interacting partners in the MS run. The signal of high levels of Ku70 and Ku80 in the eluates from TAP tag purification also suppressed lower abundant peptides. To solve this problem, we eluted proteins bound to calmodulin beads by a series of NaCl gradient elution buffers. Therefore, the binding proteins were divided into several fractions according to their binding strength to Ku70 complex. Another solution was to separate the proteins by gel according to their molecular weight. As described above, gel lane containing all proteins was divided into several equal size sections and was digested individually.

3.4.2 Core Ku70 Complex And Other Regulated Ku70 Complex

Ku70 widely expresses in various mammalian systems and the Ku complex has been described as various terms such as nuclear factor IV (NFIV), transferrin receptor promoter element binding factor (TREF), proximal sequence element 1 (PSE 1), end binding protein (EBP), and human DNA helicase II (HDHII) (Devries et al., 1989; Roberts et al., 1989; Knuth et al., 1990; Stuiver et al., 1990; Falzon et al., 1993; Tuteja, N. et al., 1994). Many direct and indirect evidences imply that Ku70 is not only a DNA binding protein, but in fact it is a multifunctional protein and a member of a multigene family. The Ku70 interacting proteins from our experiments can be classified into several biological processes, and agrees with previously reported Ku70 functions. However, we have identified many novel Ku70 interacting proteins.

Here we use the term Ku70 complex loosely, which refers to proteins that interact with Ku70 directly or indirectly. However, proteins bind to Ku70 with different affinity. Some proteins bind to Ku70 tightly and are always puried together with Ku70, such as Ku80, DNA-dependent protein kinase catalytic subunit and Poly[ADP-ribose] polymerase. Some proteins bind to Ku70 transiently, indirectly or only in specific conditions. The interactions may be easy to be disrupted during cell lysis or the procedure of purification. Furthermore, one or a set of proteins may appear in multiple complexes and complex will interact with other complexes which make the situation more complicated. In this case, Yi Wang *et al*. have put forward the concept "Core complex" and "regulated complex" (Waksman, 2005). Proteins in core complex are always together in cells and are not disrupted by buffers containing high concentration of salt and detergents. Other proteins can interact with the core protein complex to form the regulated complexes. There is only one core complex, but there are many regulated complexes in which different proteins can join the core complex to assume different functions.

The identified Ku70 interacting proteins depended on experimental conditions. High abundant proteins and stable members were always detected such as Ku70, Ku80, histones, ribosomal proteins etc. but low abundant proteins were varied partly because of the regulated complexes. To ensure pulling down of true positive Ku70 interacting partners, we compared the results from repeated and independent experiments. Proteins which have been

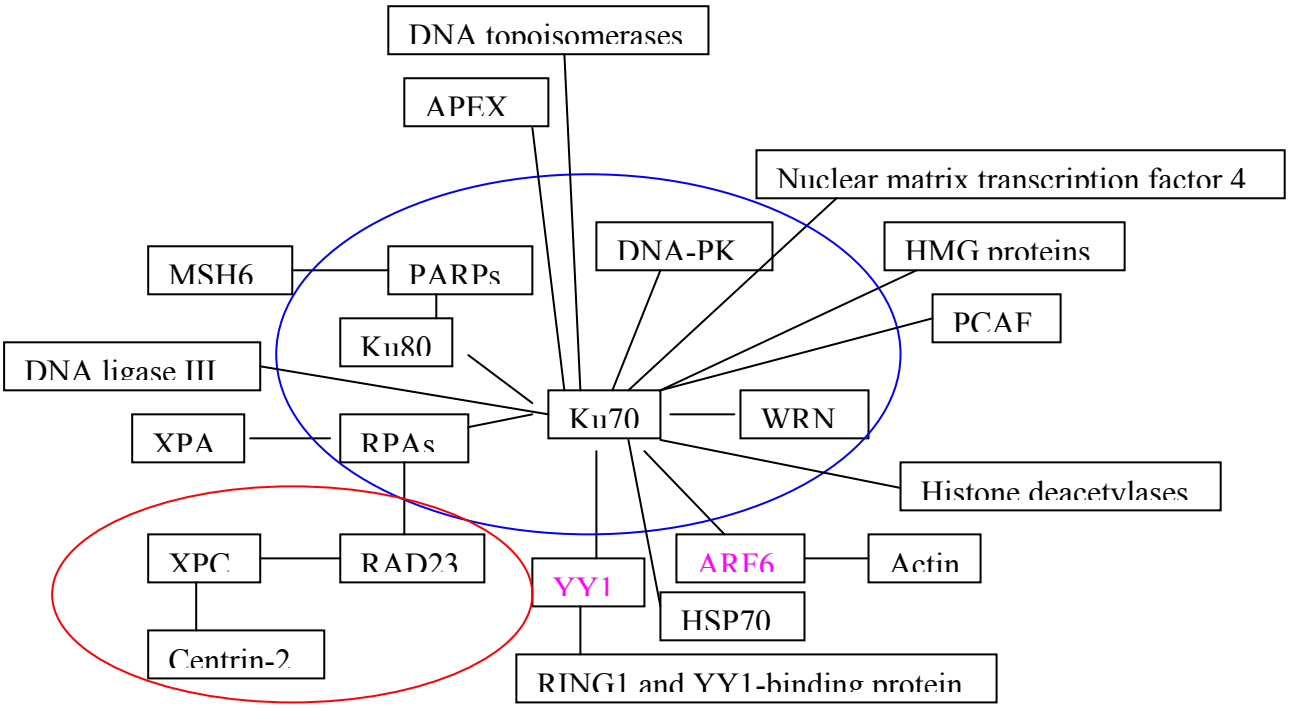
identified in multiple independent experiments but did not appear in control, and with MASCOT score higher more than 50 were treated as true positives.

Based on our data, we propose a hypothesis that Ku70, Ku80, DNA-dependent protein kinase catalytic subunit, PARP1,2, RPA1,2,3 and werner syndrome protein form the core Ku70 complex. All of them have been reported to be the partners of Ku70 by other researchers. They mainly involve in the DSB repair by NHEJ pathway (Critchlow et al., 1997; Weinfeld et al., 1997; Ismail et al., 2004). Other proteins may form regulated complexes with Ku70 transiently or in specific condition to exert particular functions, such as Ku70-PCAF and Ku70-HDACs complexes. PCAF can acetylate Ku70 while HDACs can deacetylate Ku70. These two kinds of proteins exert opposite functions. Thus, they will never present in the complex at the same time.

We have identified 20 known Ku70 interacting proteins which have been reported to interact with Ku70 directly or indirectly, these protein includes Ku80, PARPs, RPA proteins, DNA-PK, HSP70, DNA topoisomerases, PCAF, Werner syndrome protein, DNA ligase III, Histone deacetylase1, 2, APEX2. Nine proteins that haven't been reported to interact with Ku70, but they can form complex with other Ku70 known interacting proteins, including H2B, XPC, XPA, MSH6, RING and YY1-binding protein, actin, nuclear matrix transcription factor 4, high mobility group proteins and RAD23. They are the most potential candidates of functional Ku70 protein complex. Figure 3.14 shows the potential functional network between these proteins.

Besides the core Ku70 complex, additional participants in the DNA-PK mediated NHEJ repair pathway were also found in our complex proteins such as DNA ligase III. Besides, some other DNA repair proteins such as centrin 2, Xeroderma pigmentosum group C complementing protein(p125) (XPC), Xeroderma pigmentosum, complementation group A (XPA) and Structural maintenance of chromosome 1-like 1 protein (Smc1L1) were also detected. The structural maintenance of chromosomes (SMC) family of proteins has been implicated in the repair of DNA double-strand breaks (DSBs) by homologous recombination (HR) (Schar et al., 2004). Pull down of these proteins suggested that Ku70 may involve in other DNA repair processes besides NHEJ and V(D)J recombination repair pathway. XPA can form complex with the core Ku70 complex member protein RPA to initiate the nucleotide excision repair (NER) (Frit et al., 1998). Another nucleotide excision repair factor xeroderma pigmentosum C (XPC) protein is associated with RAD23 (Hsieh et al., 2005) and is required for global genome repair, while both RAD23 and RPA were copurified in NER complex (Rodriguez et al., 1998). Human centrin 2 is a component of the nucleotide excision repair system, as a subunit of the heterotrimer including xeroderma pigmentosum group C protein (XPC) and hHR23B. The pull down of RAD23, XPC and centrin2 complex strongly indicated that Ku70 may play important role in nucleotide excision repair.

Figure 3.14 Protein-Protein interactions between Ku70 and its associated proteins.
The core Ku70 complex is circled in blue. The NER complex is circled in red. Proteins in pink are known Ku70 proteins, but were not detected in our experiment.



MSH6 may interact with Ku70 through PARP since poly(ADP-ribose)-binding site was identified in MSH6 as well as other proteins like p53, DNA ligase III etc (Pleschke et al., 2000). Ku70 was identified as part of a complex with ARF6 which is essential for cytokinesis (Schweitzer and D'Souza-Schorey, 2005). Although we did not detect ARF6, 13 proteins involving in cell structure and motility were identified. ARF6 will also induce changes in actin cytoskeleton. YY1 is another known Ku70 interacting protein which form complex with Ku protein and is regulated in human heart failure and represses α myosin heavy-chain gene expression (Sucharov et al., 2004). The RING1 and YY1-binding protein may interact with Ku70 through YY1. Some papers reported that Ku70 can transiently associate with nuclear matrix and high mobility group proteins during V(D)J recombination (Yu et al., 1998; Rodgers et al., 2002a).

3.4.3 Analysis of Other Regulated Complex

Histones have never been reported to interact with Ku70 or DNA-PK. Only H2B was reported to be down-regulated by DNA-PK in response to DNA damage through modulation of octamer transcription factor 1 (Oct1) (Schild-Poulter et al., 2003). But we found histones were the most abundant proteins in the complex besides Ku80 which is known to form heterodimer with Ku70. Consider the downregulation of H2B by Oct1, other histones may also take part in the process of DNA repair. Ku70 has been reported as a

transcription factor that recruits DNA-PK directly to negative regulatory element 1 specific sequence and represses glucocorticoid-induced MMTV transcription (Giffin et al., 1996). It also has been shown that Ku acts as a potential transcription factor for RNA polymerase II (Kuhn et al., 1993). Binding of Ku70 to promoter sequences could occur via the carboxyl-terminal SAP domain (Aravind and Koonin, 2001). The modification of histones controls the activation and repression of gene transcription. The interaction between Ku70 and histones further suggests that Ku70 may participate in gene transcription. Consistent with this hypothesis, we found 30 proteins in the Ku70 complex are related to mRNA transcription.

Another interesting finding is ribosomal proteins and nuclear ribonucleoproteins in Ku70 complex. The interaction between ribosomal proteins and Ku70, nuclear ribonucleoproteins and Ku70 have not been reported. Nineteen subunits of 60S ribosomal proteins, 10 subunits of 40S ribosomal proteins and 9 subunits of nuclear ribonucleoproteins were detected in our experiments. The major function of ribosomal proteins is protein biosynthesis. The heterogeneous nuclear ribonucleoproteins (hnRNP) are a family of proteins which share common structural domains, and extensive research has shown that they have central roles in DNA repair, telomere biogenesis, cell signaling and in regulating gene expression at both transcriptional and translational levels. The result suggested Ku70 may play roles in protein synthesis and regulating gene expression through hnRNP.

Recent reports showed that Ku70 plays central role in apoptosis pathway. Although our assay did not show significant difference between normal cell and Ku70 overexpressed HEK293 cells under the treatment of HDAC inhibitors, other recent reports provide strong evidence that Ku70 is important in mediating apoptosis (Cohen et al., 2004; Subramanian et al., 2005). Ku70 was reported to bind to the pro-apoptosis protein Bax and sequester Bax-mediated apoptosis by preventing its relocalization to the mitochondria. Our data showed that 17 proteins in the Ku70 complex involving in apoptotic process including pro-apoptosis proteins FMR1, SMG1, LGALS1, FXR1, FXR2 and anti-apoptosis proteins AVEN, AKT1, 2, 3, HDAC1, 2, BAG3; however, Bax was not found in our list suggesting the interaction between Bax and Ku70 is transient and weak. This also shows that the interaction between Ku70 and other apoptosis proteins that have been identified in our pull down is stronger than Bax. Therefore, Ku70 may regulate cellular apoptosis in other ways than simply sequestering Bax.

Ku70 was also proposed to act as either a tumor suppressor or an oncoprotein. It may act with other oncoproteins or tumor suppressors to affect the oncogenic state (Gullo et al., 2006). Altered expression of Ku70 was found in many tumors. Overexpression of Ku proteins promotes oncogenic phenotypes, such as hyperproliferation or resistance to apoptosis. Deficient or low expression of Ku leads to tumorigenesis. Recent studies have also provided some evidence that defects in the two major DNA DSBR pathways are related to the development of cancer (Khanna and Jackson, 2001). The

involvement of Ku70 associated proteins in p53 pathway further demonstrated the relationship between Ku70 and the development of cancer. We also found 2 oncoproteins in the complex: MAGEB18 and FUS. We detected PARP1 and 2 in the complex and demonstrated the interaction between PARP and Ku80. PARP-1 is a nuclear enzyme that catalyses the poly(ADP-ribosyl)ation of target proteins in response to DNA damage. It is suggested to play a vital role in DNA repair and recombination, cell death, proliferation and stabilization of the genome (Lindahl et al., 1995; Herceg and Wang, 2001). Recent report showed that PARP-1 not only plays important role in NHEJ repair pathway, but also in suppression of liver tumorigenesis together with Ku80 (Tong et al., 2002). These findings all strongly suggested Ku70 may be a tumor suppressor or an oncoprotein. More excitingly, the relationship between Ku70 and cancer makes Ku70 an important candidate target for anticancer drug development.

Paillard and Strauss predicted a possible role of Ku70 in replication forks movement during DNA replication (Paillard and Strauss, 1991). Ku was reported to interact with human DNA containing replication origin (Toth et al., 1993). Our result showed that the critical proteins in DNA replication process such as topoisomerase I and topoisomeraseII, replication protein A as well as WRN, MCM3 etc. were associated with Ku70. These findings suggested that Ku70 may have an important role in the initiation of DNA replication.

It was reported that the subcellular localization of Ku70 is cell-cycle dependent which was also demonstrated by our own experiments. Ku70 was distributed in both the nucleus and the cytoplasm in M phase while it only

restricted to nucleus in other phases. The cell cycle was reported to be blocked at G2 phase in yeast Ku70-deficient strain, indicating the potential role of Ku70 in the regulation of cell cycle (Feldmann and Winnacker, 1993). Our result also strongly suggested the involvement of Ku70 in cell cycle regulation. Twenty-nine proteins in the Ku70 complex are related to cell cycle regulation.

Ku70 is a subunit of DNA-PK complex that is reported to play an important role in signal transduction pathway (Hartley et al., 1995). The involvement of 14 Ku70 associated proteins in cell signaling indicated that Ku70 may take part in the pathway directly. In fact, Ku70 was reported as an important component of signal transduction pathways according to its response to different external stimuli (Fewell and Kuff, 1996).

Five proteins in the complex involve in Huntington disease (HD) pathway. Huntington's disease is an autosomal dominantly inherited neurodegenerative disorder caused by a polyglutamine repeat expansion. Recent studies showed that transcriptional dysregulation may be the main reason for HD pathogenesis. The control of eukaryotic gene expression depends on the modification of histone. We have indicated that Ku70 may interact with histones and play roles in gene transcription so that connected to HD. Interestingly, recent studies have shown a therapeutic role for histone deacetylase (HDAC) inhibitors in a number of HD models (Sadri-Vakili and Cha, 2006).

3.5 Conclusion

In conclusion, we applied tandem affinity purification method to pull down Ku70 complex in HEK293F cells. There were 151 proteins or subunits being found in the complex. Most of them involve in DNA repair, apoptosis, cell cycle regulation, gene transcription and signal transduction. Many Ku70 associated proteins participate in p53 pathway indicating the relationship between Ku70 and the development of cancer. PARP was demonstrated to be a real partner of Ku70 and interact with Ku70 through Ku80. We proposed a hypothesis that Ku70, Ku80, DNA-dependent protein kinase catalytic subunit, PARP1,2, RPA1,2,3 and werner syndrome protein form a core protein complex that mainly exert the function of DSB repair through NHEJ pathway. Other proteins form regulated complexes with Ku70 which are not as stable as the core complex or form complex in specific condition. These regulated complexes assume multiple functions of Ku70. Our result provided clues for the roles of Ku70 in biological processes. The interaction between Ku70 and its associated proteins needs further validation to firmly elucidate their biological functions.

Table 3.1 Histones, ribosomal proteins, nuclear ribonucleoproteins, heat shock proteins and tubulins in Ku70 complex,

Protein Name	IPI No.	Gene Symbol
Histone proteins (12)		
Histone H1.1	IPI00217469	HIST1H1A
Histone H1.2	IPI00217465	HIST1H1C
Histone H1.4	IPI00217467	HIST1H1E
Histone H1.5	IPI00217468	HIST1H1B
Histone H2A type 1-A	IPI00045109	HIST2H2AA
Histone H2A type 1-B	IPI00216730	HIST2H2AB
Histone H2A type 1-C	IPI00216456	HIST1H2AC
Histone H2A type 1-D	IPI00255316	HIST1H2AD
Histone H2B type 3-B	IPI00166293	HIST3H2BB
Histone H3.1	IPI00465070	HIST1H3A
Histone H3.4	IPI00216402	HIST3H3
Histone H4	IPI00453473.5	HIST2H4A
Ribosomal proteins (29)		
60S ribosomal protein L4	IPI00003918	RPL4
60S ribosomal protein L5	IPI00000494	RPL5
60S ribosomal protein L6	IPI00329389	RPL6
60S ribosomal protein L7	IPI00030179	RPL7
60S ribosomal protein L8	IPI00012772	RPL8
60S ribosomal protein L11	IPI00219150	RPL11
60S ribosomal protein L12	IPI00024933	RPL12
60S ribosomal protein L15	IPI00470528	RPL15
60S ribosomal protein L17	IPI00413324	RPL17
60S ribosomal protein L18	IPI00215719	RPL18
60S ribosomal protein L19	IPI00025329	RPL19
60S ribosomal protein L24	IPI00306332	RPL24
60S ribosomal protein L26	IPI00027270	RPL26
60S ribosomal protein L27	IPI00219155	RPL27
60S ribosomal protein L29	IPI00419919	RPL29
60S ribosomal protein L30	IPI00219156	RPL30
60S ribosomal protein L31	IPI00026302	RPL31
60S acidic ribosomal protein P1	IPI00008527	RPLP1
60S acidic ribosomal protein P2	IPI00008529	RPLP2
40S ribosomal protein S2	IPI00013485	RPS2
40S ribosomal protein S8	IPI00216587	RPS8

40S ribosomal protein S9	IPI00221088	RPS9
40S ribosomal protein S10	IPI00008438	RPS10
40S ribosomal protein S11	IPI00025091	RPS11
40S ribosomal protein S14	IPI00026271	RPS14
40S ribosomal protein S16	IPI00221092	RPS16
40S ribosomal protein S19	IPI00215780	RPS19
40S ribosomal protein S23	IPI00218606	RPS23
40S ribosomal protein S24	IPI00029750	RPS24
Nuclear ribonucleoprotein (9)		
Heterogenous nuclear ribonucleoprotein U	IPI00025054	HNRPU
Hetogeneous nuclear ribonucleoproteins C1/C2	IPI00216592	HNRPC
Heterogeneous nuclear ribonucleoprotein R	IPI00012074	HNRPR
Heterogeneous nuclear ribonucleoprotein H'	IPI00026230	HNRPH2
Small nuclear ribonucleoprotein associated protein N	IPI00220360	SNRPN
Small nuclear ribonucleoprotein B'	IPI00384173	SNRPB
Heterogeneous nuclear ribonucleoprotein Q	IPI00402182	SYNCRIP
Heterogeneous nuclear ribonucleoprotein K	IPI00640296	HNRPK
Heterogeneous nuclear ribonucleoprotein C-like 1	IPI00027569	HNRPCL1
Heat shock proteins (4)		
Heat shock protein HSP 90-alpha 2	IPI00382470	HSPCA
Heat shock protein HSP 90-beta	IPI00414676	HSPCB
Heat shock cognate 71 kDa protein	IPI00037070	HSPA8
Heat shock 70 kDa protein 1	IPI00304925	HSPA1A
Tubulin and tubulin related proteins (3)		
Tubulin alpha-2 chain	IPI00179709	TUBA2
Tubulin alpha-6 chain	IPI00218343	TUBA6
Tubulin family protein	IPI00646909	TUBA8

Table 3.2 Other 94 proteins in Ku70 complex.

Protein Name	IPI No.	Gene Symbol	Mascot Score
ATP-dependent DNA helicase II, 70 kDa subunit	IPI00465430	G22P1	232043
ATP-dependent DNA helicase II 80 kDa subunit	IPI00220834	XRCC5	207006
Poly [ADP-ribose] polymerase 1	IPI00449049	PARP1	4621
Glial fibrillary acidic protein, astrocyte	IPI00443478	GFAP	4374
K-ALPHA-1 protein	IPI00328163	K-ALPHA-1	3971

Replication protein A 70 kDa DNA-binding subunit	IPI00020127	RPA1	3547
UBC protein	IPI00554749	UBB	2685
Nuclease sensitive element-binding protein 1	IPI00031812	NSEP1	2545
DNA-dependent protein kinase catalytic subunit	IPI00296337	PRKDC	2158
DNA-binding protein	IPI00385699	YBX1	1950
Vimentin	IPI00646867	VIM	1827
BBF2H7/FUS protein (Fragment)	IPI00428056	FUS	1585
Actin, alpha cardiac	IPI00023006	ACTC	1348
Actin, aortic smooth muscle	IPI00008603	ACTA2	1309
Structure-specific recognition protein 1	IPI00005154	SSRP1	1097
Creatine kinase B-type	IPI00022977	CKB	1079
Centrin-2	IPI00215928	CETN2	751
DNA-binding protein A	IPI00219148	CSDA	711
EIF4G3 protein	IPI00328268	EIF4G3	639
DNA topoisomerase 2-alpha	IPI00218753	TOP2A	635
Elongation factor 1-alpha 2	IPI00014424	EEF1A2	633
EEF1A protein (Fragment)	IPI00382804	EEF1A1	592
Replication protein A 14 kDa subunit	IPI00017373	RPA3	555
Y-box-binding protein 2	IPI00250153	YBX2	475
SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A3	IPI00339381	SMARCA3	472
zinc finger protein 265 isoform 2	IPI00183794	ZNF265	472
BAG-family molecular chaperone regulator-3	IPI00639961	BAG3	466
Bcl-2-associated transcription factor 1	IPI00006079	BCLAF1	425
Replication protein A 32 kDa subunit	IPI00646500	RPA2	388
Dynein light chain 1, cytoplasmic	IPI00019329	DNCL1	374
Methyl-CpG-binding protein 2	IPI00418234	MECP2	365
nuclear matrix transcription factor 4	IPI00220697	ZNF384	342
Galectin-1	IPI00219219	LGALS1	341
Histone acetyltransferase PCAF	IPI00022055	PCAF	340
Poly(A) binding protein, cytoplasmic 4	IPI00642944	PABPC4	325
ATP-dependent RNA helicase DDX17	IPI00651653	DDX17	300
Similar to Calcium-dependent protease, small subunit	IPI00025084	CAPNS1	286
Splice Isoform Beta-2 of DNA topoisomerase 2-beta	IPI00027280	TOP2B	279
Splice Isoform 1 of Poly [ADP-ribose] polymerase 2	IPI00026497	PARP2	277

Splice Isoform 1 of Nucleolar RNA helicase 2	IPI00015953	DDX21	268
Calumenin precursor	IPI00014537	CALU	267
high mobility group AT-hook 1 isoform a	IPI00179700	HMGA1	245
High mobility group protein HMGI-C	IPI00005996	HMGA2	243
hepatoma-derived growth factor-related protein 2 isoform 1	IPI00419322	HDGF2	217
Treacle protein	IPI00298696	TCOF1	188
FK506-binding protein 4	IPI00219005	FKBP4	181
programmed cell death 8 isoform 2	IPI00157908	PDCD8	174
SMARCA4 isoform 2	IPI00029822	SMARCA4	169
Cell death regulator Aven	IPI00006904	AVEN	166
PREDICTED: similar to Enhancer of rudimentary homolog	IPI00455070	CLEC2L	157
DNA-repair protein complementing XP-C cells	IPI00156793	XPC	154
UV excision repair protein RAD23 homolog B	IPI00008223	RAD23B	141
DNA replication licensing factor MCM3	IPI00013214	MCM3	140
Splice Isoform 1 of Ubiquitin conjugation factor E4 B	IPI00005715	UBE4B	138
Probable ATP-dependent RNA helicase DDX5	IPI00017617	DDX5	137
Fragile X mental retardation syndrome-related protein 1	IPI00016249	FXR1	134
Splice Isoform 1 of Nuclear autoantigenic sperm protein	IPI00179953	NASP	131
ATP-dependent RNA helicase DDX50	IPI00031554	DDX50	129
desmocollin 1 isoform Dsc1b preproprotein	IPI00007425	DSC1	122
UV excision repair protein RAD23 homolog A	IPI00008219	RAD23A	114
Werner syndrome ATP-dependent helicase	IPI00029107	WRN	112
Fragile X mental retardation syndrome-related protein 2	IPI00016250	FXR2	97
Structural maintenance of chromosome 1-like 1 protein	IPI00291939	SMC1L1	97
Splice Isoform GTBP-alt of DNA mismatch repair protein MSH6	IPI00106847	MSH6	96
Transcription factor SOX-11	IPI00020435	SOX11	95
SWI/SNF-related matrix-associated actin dependent regulator of chromatin subfamily A5	IPI00297211	SMARCA5	91
Splice Isoform 1 of Chromodomain helicase-DNA-binding protein 3	IPI00373870	CHD3	86
ligase III, DNA, ATP-dependent isoform beta precursor	IPI00000156	LIG3	83
Transcription factor GATA-4	IPI00020064	GATA4	77

Nucleosome assembly protein 1-like 1	IPI00023860	NAP1L1	77
Annexin A1	IPI00549413	ANXA1	76
remodeling and spacing factor 1	IPI00290652	HBXAP	75
Transcriptional repressor p66 beta	IPI00103554	GATAD2B	74
Splice Isoform 3 of Enhancer of polycomb homolog 1	IPI00552203	EPC1	71
Histone deacetylase 1	IPI00013774	HDAC1	69
Centaurin-delta 1	IPI00292471	CENTD1	62
Chromodomain-helicase-DNA-binding protein 2	IPI00023109	CHD2	62
Histone deacetylase 2	IPI00289601	HDAC2	62
RAC-alpha serine/threonine-protein kinase	IPI00012866	AKT1	60
RAC-beta serine/threonine-protein kinase	IPI00012870	AKT2	60
Splice Isoform 1 of RAC-gamma serine/threonine-protein kinase	IPI00031747	AKT3	60
Rho GDP-dissociation inhibitor 1	IPI00003815	ARHGDIA	60
Splice Isoform 2 of Mitogen-activated protein kinase kinase kinase 9	IPI00179189	MAP3K9	60
Splice Isoform 1 of Histone-lysine N-methyltransferase, H4 lysine-20 specific	IPI00288890	SET8	58
Melanoma-associated antigen B18	IPI00065350	MAGEB18	57
DNA-(apurinic or apyrimidinic site) lyase 2	IPI00083281	APEX2	56
Splice Isoform 5 of GTPase-activating Rap/Ran-GAP domain-like 1	IPI00456723	GARNL1	55
DNA topoisomerase I	IPI00413611	TOP1	54
Splice Isoform 6 of Fragile X mental retardation 1 protein	IPI00412343	FMR1	53
Splice Isoform 1 of cAMP-specific 3',5'-cyclic phosphodiesterase 4D	IPI00002449	PDE4D	53
RING1 and YY1-binding protein	IPI00296594	RYBP	51
nadrin isoform 1	IPI00064767	ARHGAP17	50
PI-3-kinase-related kinase SMG-1	IPI00556369	SMG1	50
DNA-repair protein complementing XP-A cells	IPI00009694	XPA	50

Table 3.3 Molecular functions modulated by all complex proteins

1	Nucleic acid binding (MF00042)	103	65.60%	50.70%
2	Transcription factor (MF00036)	14	8.90%	6.90%
3	Select regulatory molecule (MF00093)	9	5.70%	4.40%
4	Cytoskeletal protein (MF00091)	9	5.70%	4.40%
5	Chaperone (MF00077)	7	4.50%	3.40%
6	Kinase (MF00107)	7	4.50%	3.40%
7	Transferase (MF00131)	6	3.80%	3.00%
8	Molecular function unclassified (MF00208)	6	3.80%	3.00%
9	Signaling molecule (MF00016)	5	3.20%	2.50%
10	Select calcium binding protein (MF00188)	4	2.50%	2.00%
11	Miscellaneous function (MF00197)	4	2.50%	2.00%
12	Hydrolase (MF00141)	4	2.50%	2.00%
13	Isomerase (MF00166)	3	1.90%	1.50%
14	Cell adhesion molecule (MF00040)	3	1.90%	1.50%
15	Transfer/carrier protein (MF00087)	2	1.30%	1.00%
16	Cell junction protein (MF00276)	2	1.30%	1.00%
17	Transporter (MF00082)	1	0.60%	0.50%
18	Ligase (MF00170)	1	0.60%	0.50%
19	Protease (MF00153)	1	0.60%	0.50%

Table 3.4 Biological processes modulated by all complex proteins

1	Nucleoside, nucleotide and nucleic acid metabolism (BP00031)	81	51.60%	27.30%
2	Protein metabolism and modification (BP00060)	54	34.40%	18.20%
3	Cell cycle (BP00203)	29	18.50%	9.80%
4	Apoptosis (BP00179)	17	10.80%	5.70%
5	Signal transduction (BP00102)	14	8.90%	4.70%
6	Cell structure and motility (BP00285)	13	8.30%	4.40%
7	Developmental processes (BP00193)	13	8.30%	4.40%
8	Immunity and defense (BP00148)	13	8.30%	4.40%
9	Intracellular protein traffic (BP00125)	10	6.40%	3.40%
10	Cell proliferation and differentiation (BP00224)	5	3.20%	1.70%
11	Transport (BP00141)	5	3.20%	1.70%
12	Nitrogen metabolism (BP00090)	4	2.50%	1.30%
13	Cell adhesion (BP00124)	4	2.50%	1.30%
14	Biological process unclassified (BP00216)	3	1.90%	1.00%

15	Neuronal activities (BP00166)	2	1.30%	0.70%
16	Oncogenesis (BP00281)	2	1.30%	0.70%
17	Muscle contraction (BP00173)	1	0.60%	0.30%
18	Lipid, fatty acid and steroid metabolism (BP00019)	1	0.60%	0.30%
19	Sensory perception (BP00182)	1	0.60%	0.30%

Table 3.5 Pathways modulated by all complex proteins

1	p53 pathway (P00059)	9	5.70%	10.50%
2	Apoptosis signaling pathway (P00006)	7	4.50%	8.10%
3	p53 pathway feedback loops 2 (P04398)	5	3.20%	5.80%
4	Huntington disease (P00029)	5	3.20%	5.80%
5	Wnt signaling pathway (P00057)	5	3.20%	5.80%
6	FAS signaling pathway (P00020)	5	3.20%	5.80%
7	Insulin/IGF pathway-protein kinase B signaling cascade (P00033)	4	2.50%	4.70%
8	EGF receptor signaling pathway (P00018)	4	2.50%	4.70%
9	DNA replication (P00017)	4	2.50%	4.70%
10	Angiogenesis (P00005)	3	1.90%	3.50%
11	Interleukin signaling pathway (P00036)	3	1.90%	3.50%
12	Inflammation mediated by chemokine and cytokine signaling pathway (P00031)	3	1.90%	3.50%
13	Hypoxia response via HIF activation (P00030)	3	1.90%	3.50%
14	Endothelin signaling pathway (P00019)	3	1.90%	3.50%
15	PI3 kinase pathway (P00048)	3	1.90%	3.50%
16	PDGF signaling pathway (P00047)	3	1.90%	3.50%
17	Ras Pathway (P04393)	3	1.90%	3.50%
18	VEGF signaling pathway (P00056)	3	1.90%	3.50%
19	T cell activation (P00053)	3	1.90%	3.50%
20	FGF signaling pathway (P00021)	3	1.90%	3.50%
21	Parkinson disease (P00049)	2	1.30%	2.30%
22	Oxidative stress response (P00046)	1	0.60%	1.20%
23	Notch signaling pathway (P00045)	1	0.60%	1.20%

Table is read as: Category name (Accession): # genes; Percent of gene hit against total # genes; Percent of gene hit against total # Pathway hits

Table 3.6 Listing of all complex proteins in each of the 19 biological processes

Nucleoside, nucleotide and nucleic acid metabolism (BP00031) 81		
GeneID:84717	HDGF2	hepatoma-derived growth factor-related protein 2;HDGF2
GeneID:10492	SYNCRIP	synaptotagmin binding, cytoplasmic RNA interacting protein;SYNCRIP
GeneID:27301	APEX2	APEX nuclease (apurinic/apyrimidinic endonuclease) 2;APEX2
GeneID:9406	ZNF265	zinc finger protein 265;ZNF265
GeneID:2626	GATA4	GATA binding protein 4;GATA4
GeneID:1107	CHD3	chromodomain helicase DNA binding protein 3;CHD3
GeneID:50810	HDGF2	hepatoma-derived growth factor, related protein 3;HDGFRP3
GeneID:3980	LIG3	ligase III, DNA, ATP-dependent;LIG3
GeneID:10236	HNRPR	heterogeneous nuclear ribonucleoprotein R;HNRPR
GeneID:5886	RAD23A	RAD23 homolog A (S. cerevisiae);RAD23A
GeneID:8850	PCAF	p300/CBP-associated factor;PCAF
GeneID:7155	TOP2B	topoisomerase (DNA) II beta 180kDa;TOP2B
GeneID:3065	HDAC1	histone deacetylase 1;HDAC1
GeneID:6638	SNRPN	small nuclear ribonucleoprotein polypeptide N;SNRPN
GeneID:3190	HNRPK	heterogeneous nuclear ribonucleoprotein K;HNRPK
GeneID:8243	SMC1L1	SMC1 structural maintenance of chromosomes 1-like 1 (yeast);SMC1L1
GeneID:51773	HBXAP	hepatitis B virus x associated protein;HBXAP
GeneID:23429	RYBP	RING1 and YY1 binding protein;RYBP
GeneID:7486	WRN	Werner syndrome;WRN
GeneID:4673	NAP1L1	nucleosome assembly protein 1-like 1;NAP1L1
GeneID:7153	TOP2A	topoisomerase (DNA) II alpha 170kDa;TOP2A
GeneID:6597	SMARCA4	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4;SMARCA4
GeneID:6664	SOX11	SRY (sex determining region Y)-box 11;SOX11
GeneID:6628	SNRPB	small nuclear ribonucleoprotein polypeptides B and B1;SNRPB
GeneID:1106	CHD2	chromodomain helicase DNA binding protein 2;CHD2
GeneID:5144	PDE4D	phosphodiesterase 4D, cAMP-specific (phosphodiesterase E3 dunce homolog, Drosophila);PDE4D
GeneID:3183	HNRPC	heterogeneous nuclear ribonucleoprotein C (C1/C2);HNRPC
GeneID:1655	DDX5	DEAD (Asp-Glu-Ala-Asp) box polypeptide 5;DDX5
GeneID:8467	SMARCA5	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 5;SMARCA5
GeneID:7520	XRCC5	X-ray repair complementing defective repair in Chinese hamster cells 5 (double-strand-break rejoining; Ku

		autoantigen, 80kDa);XRCC5
GeneID:7508	XPC	xeroderma pigmentosum, complementation group C;XPC
GeneID:6118	RPA2	replication protein A2, 32kDa;RPA2
GeneID:2547	G22P1	X-ray repair complementing defective repair in Chinese hamster cells 6 (Ku autoantigen, 70kDa);XRCC6
GeneID:2956	MSH6	mutS homolog 6 (E. coli);MSH6
GeneID:3066	HDAC2	histone deacetylase 2;HDAC2
GeneID:5887	RAD23B	RAD23 homolog B (S. cerevisiae);RAD23B
GeneID:4172	MCM3	MCM3 minichromosome maintenance deficient 3 (S. cerevisiae);MCM3
GeneID:3192	HNRPU	heterogeneous nuclear ribonucleoprotein U (scaffold attachment factor A);HNRPU
GeneID:3188	HNRPH2	heterogeneous nuclear ribonucleoprotein H2 (H');HNRPH2
GeneID:4904	NSEP1,YBX1	Y box binding protein 1;YBX1
GeneID:387893	SET8	PR/SET domain containing protein 8;SET8
GeneID:343069	HNRPCL1	heterogeneous nuclear ribonucleoprotein C-like 1;HNRPCL1
GeneID:9188	DDX21	DEAD (Asp-Glu-Ala-Asp) box polypeptide 21;DDX21
GeneID:9513	FXR2	fragile X mental retardation, autosomal homolog 2;FXR2
GeneID:142	PARP1	poly (ADP-ribose) polymerase family, member 1;PARP1
GeneID:8761	PABPC4	poly(A) binding protein, cytoplasmic 4 (inducible form);PABPC4
GeneID:51087	YBX2	Y box binding protein 2;YBX2
GeneID:10521	DDX17	DEAD (Asp-Glu-Ala-Asp) box polypeptide 17;DDX17
GeneID:25885	RPA1	polymerase (RNA) I polypeptide A, 194kDa;POLR1A
GeneID:8087	FXR1	fragile X mental retardation, autosomal homolog 1;FXR1
GeneID:80314	EPC1	enhancer of polycomb homolog 1 (Drosophila);EPC1
GeneID:57459	GATAD2B	GATA zinc finger domain containing 2B;GATAD2B
GeneID:6596	SMARCA3	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 3;SMARCA3
GeneID:6749	SSRP1	structure specific recognition protein 1;SSRP1
GeneID:5591	PRKDC	protein kinase, DNA-activated, catalytic polypeptide;PRKDC
GeneID:7150	TOP1	topoisomerase (DNA) I;TOP1
GeneID:23049	SMG1	PI-3-kinase-related kinase SMG-1;SMG1
GeneID:4678	NASP	nuclear autoantigenic sperm protein (histone-binding);NASP
GeneID:8531	CSDA	cold shock domain protein A;CSDA
GeneID:4204	MECP2	methyl CpG binding protein 2 (Rett syndrome);MECP2
GeneID:10038	PARP2	poly (ADP-ribose) polymerase family, member 2;PARP2

GeneID:2332	FMR1	fragile X mental retardation 1;FMR1
GeneID:6117	RPA1	replication protein A1, 70kDa;RPA1
GeneID:79009	DDX50	DEAD (Asp-Glu-Ala-Asp) box polypeptide 50;DDX50
GeneID:8350	HIST1H3A	histone 1, H3a;HIST1H3A
GeneID:8290	HIST3H3	histone 3, H3;HIST3H3
GeneID:54606	DDX21	DEAD (Asp-Glu-Ala-Asp) box polypeptide 56;DDX56
GeneID:7507	XPA	xeroderma pigmentosum, complementation group A;XPA
GeneID:3008	HIST1H1E	histone 1, H1e;HIST1H1E
GeneID:3009	HIST1H1B	histone 1, H1b;HIST1H1B
GeneID:3006	HIST1H1C	histone 1, H1c;HIST1H1C
GeneID:3024	HIST1H1A	histone 1, H1a;HIST1H1A
GeneID:8337	HIST2H2AA	histone 2, H2aa;HIST2H2AA
GeneID:8334	HIST1H2AC	histone 1, H2ac;HIST1H2AC
GeneID:3013	HIST1H2AD	histone 1, H2ad;HIST1H2AD
GeneID:317772	HIST2H2AB	histone 2, H2ab;HIST2H2AB
GeneID:8655	DNCL1	dynein, light chain, LC8-type 1;DYNLL1
GeneID:8091	HMGA2	high mobility group AT-hook 2;HMGA2
GeneID:171017	ZNF384	zinc finger protein 384;ZNF384
GeneID:3159	HMGA1	high mobility group AT-hook 1;HMGA1
GeneID:154790	CLEC2L	C-type lectin domain family 2, member L;CLEC2L

Protein metabolism and modification (BP00060) 54		
GeneID:6125	RPL5	ribosomal protein L5;RPL5
GeneID:6132	RPL8	ribosomal protein L8;RPL8
GeneID:207	AKT1	v-akt murine thymoma viral oncogene homolog 1;AKT1
GeneID:2288	FKBP4	FK506 binding protein 4, 59kDa;FKBP4
GeneID:10000	AKT3	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma);AKT3
GeneID:7314	UBB	ubiquitin B;UBB
GeneID:10277	UBE4B	ubiquitination factor E4B (UFD2 homolog, yeast);UBE4B
GeneID:1915	EEF1A1	eukaryotic translation elongation factor 1 alpha 1;EEF1A1
GeneID:3190	HNRPK	heterogeneous nuclear ribonucleoprotein K;HNRPK
GeneID:3303	HSPA1A	heat shock 70kDa protein 1A;HSPA1A
GeneID:51773	HBXAP	hepatitis B virus x associated protein;HBXAP
GeneID:8672	EIF4G3	eukaryotic translation initiation factor 4 gamma, 3;EIF4G3
GeneID:9531	BAG3	BCL2-associated athanogene 3;BAG3
GeneID:6124	RPL4	ribosomal protein L4;RPL4
GeneID:3326	HSPCB	heat shock 90kDa protein 1, beta;HSPCB

GeneID:1917	EEF1A2	eukaryotic translation elongation factor 1 alpha 2;EEF1A2
GeneID:6203	RPS9	ribosomal protein S9;RPS9
GeneID:6128	RPL6	ribosomal protein L6;RPL6
GeneID:6202	RPS8	ribosomal protein S8;RPS8
GeneID:6138	RPL15	ribosomal protein L15;RPL15
GeneID:6129	RPL7	ribosomal protein L7;RPL7
GeneID:208	AKT2	v-akt murine thymoma viral oncogene homolog 2;AKT2
GeneID:6139	RPL17	ribosomal protein L17;RPL17
GeneID:142	PARP1	poly (ADP-ribose) polymerase family, member 1;PARP1
GeneID:826	CAPNS1	calpain, small subunit 1;CAPNS1
GeneID:3320	HSPCA	heat shock 90kDa protein 1, alpha;HSPCA
GeneID:3312	HSPA8	heat shock 70kDa protein 8;HSPA8
GeneID:5591	PRKDC	protein kinase, DNA-activated, catalytic polypeptide;PRKDC
GeneID:4293	MAP3K9	mitogen-activated protein kinase kinase kinase 9;MAP3K9
GeneID:23049	SMG1	PI-3-kinase-related kinase SMG-1;SMG1
GeneID:6187	RPS2	ribosomal protein S2;RPS2
GeneID:3324	HSPCA	heat shock 90kDa protein 1, alpha-like 3;HSPCAL3
GeneID:10038	PARP2	poly (ADP-ribose) polymerase family, member 2;PARP2
GeneID:6141	RPL18	ribosomal protein L18;RPL18
GeneID:6136	RPL12	ribosomal protein L12;RPL12
GeneID:6135	RPL11	ribosomal protein L11;RPL11
GeneID:6143	RPL19	ribosomal protein L19;RPL19
GeneID:51187	RPL24	chromosome 15 open reading frame 15;C15orf15
GeneID:9349	RPL17	ribosomal protein L23;RPL23
GeneID:6154	RPL26	ribosomal protein L26;RPL26
GeneID:6208	RPS14	ribosomal protein S14;RPS14
GeneID:6205	RPS11	ribosomal protein S11;RPS11
GeneID:6204	RPS10	ribosomal protein S10;RPS10
GeneID:6223	RPS19	ribosomal protein S19;RPS19
GeneID:6229	RPS24	ribosomal protein S24;RPS24
GeneID:6152	RPL24	ribosomal protein L24;RPL24
GeneID:6155	RPL27	ribosomal protein L27;RPL27
GeneID:6159	RPL29	ribosomal protein L29;RPL29
GeneID:6160	RPL31	ribosomal protein L31;RPL31
GeneID:6181	RPLP2	ribosomal protein, large, P2;RPLP2
GeneID:6156	RPL30	ribosomal protein L30;RPL30
GeneID:6176	RPLP1	ribosomal protein, large, P1;RPLP1

GeneID:6228	RPS23	ribosomal protein S23;RPS23
GeneID:6217	RPS16	ribosomal protein S16;RPS16

Cell cycle (BP00203) 29		
GeneID:207	AKT1	v-akt murine thymoma viral oncogene homolog 1;AKT1
GeneID:2288	FKBP4	FK506 binding protein 4, 59kDa;FKBP4
GeneID:10000	AKT3	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma);AKT3
GeneID:7155	TOP2B	topoisomerase (DNA) II beta 180kDa;TOP2B
GeneID:3065	HDAC1	histone deacetylase 1;HDAC1
GeneID:8243	SMC1L1	SMC1 structural maintenance of chromosomes 1-like 1 (yeast);SMC1L1
GeneID:51773	HBXAP	hepatitis B virus x associated protein;HBXAP
GeneID:7486	WRN	Werner syndrome;WRN
GeneID:7278	TUBA2	tubulin, alpha 2;TUBA2
GeneID:4673	NAP1L1	nucleosome assembly protein 1-like 1;NAP1L1
GeneID:7153	TOP2A	topoisomerase (DNA) II alpha 170kDa;TOP2A
GeneID:51807	TUBA8	tubulin, alpha 8;TUBA8
GeneID:6118	RPA2	replication protein A2, 32kDa;RPA2
GeneID:3066	HDAC2	histone deacetylase 2;HDAC2
GeneID:4172	MCM3	MCM3 minichromosome maintenance deficient 3 (S. cerevisiae);MCM3
GeneID:208	AKT2	v-akt murine thymoma viral oncogene homolog 2;AKT2
GeneID:253959	GARNL1	GTPase activating Rap/RanGAP domain-like 1;GARNL1
GeneID:59	ACTA2	actin, alpha 2, smooth muscle, aorta;ACTA2
GeneID:10376	K-ALPHA-1	tubulin, alpha, ubiquitous;K-ALPHA-1
GeneID:84790	TUBA6	tubulin, alpha 6;TUBA6
GeneID:5591	PRKDC	protein kinase, DNA-activated, catalytic polypeptide;PRKDC
GeneID:7150	TOP1	topoisomerase (DNA) I;TOP1
GeneID:23049	SMG1	PI-3-kinase-related kinase SMG-1;SMG1
GeneID:70	ACTC	actin, alpha, cardiac muscle;ACTC
GeneID:6117	RPA1	replication protein A1, 70kDa;RPA1
GeneID:1069	CETN2	centrin, EF-hand protein, 2;CETN2
GeneID:8655	DNCL1	dynein, light chain, LC8-type 1;DYNLL1
GeneID:8091	HMGA2	high mobility group AT-hook 2;HMGA2
GeneID:3159	HMGA1	high mobility group AT-hook 1;HMGA1

Apoptosis (BP00179) 17

GeneID:57099	AVEN	apoptosis, caspase activation inhibitor;AVEN
GeneID:207	AKT1	v-akt murine thymoma viral oncogene homolog 1;AKT1
GeneID:10277	UBE4B	ubiquitination factor E4B (UFD2 homolog, yeast);UBE4B
GeneID:3065	HDAC1	histone deacetylase 1;HDAC1
GeneID:9531	BAG3	BCL2-associated athanogene 3;BAG3
GeneID:23049	SMG1	PI-3-kinase-related kinase SMG-1;SMG1
GeneID:4673	NAP1L1	nucleosome assembly protein 1-like 1;NAP1L1
GeneID:3190	HNRPK	heterogeneous nuclear ribonucleoprotein K;HNRPK
GeneID:208	AKT2	v-akt murine thymoma viral oncogene homolog 2;AKT2
GeneID:3066	HDAC2	histone deacetylase 2;HDAC2
GeneID:6949	TCOF1	Treacher Collins-Franceschetti syndrome 1;TCOF1
GeneID:9513	FXR2	fragile X mental retardation, autosomal homolog 2;FXR2
GeneID:8087	FXR1	fragile X mental retardation, autosomal homolog 1;FXR1
GeneID:10000	AKT3	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma);AKT3
GeneID:2332	FMR1	fragile X mental retardation 1;FMR1
GeneID:9131	PDCD8	programmed cell death 8 (apoptosis-inducing factor);PDCD8
GeneID:3956	LGALS1	lectin, galactoside-binding, soluble, 1 (galectin 1);LGALS1

Signal transduction (BP00102) 14		
GeneID:2288	FKBP4	FK506 binding protein 4, 59kDa;FKBP4
GeneID:1827	DSC1	Down syndrome critical region gene 1;DSCR1
GeneID:1823	DSC1	desmocollin 1;DSC1
GeneID:3190	HNRPK	heterogeneous nuclear ribonucleoprotein K;HNRPK
GeneID:301	ANXA1	annexin A1;ANXA1
GeneID:5144	PDE4D	phosphodiesterase 4D, cAMP-specific (phosphodiesterase E3 dunce homolog, Drosophila);PDE4D
GeneID:813	CALU	calumenin;CALU
GeneID:116984	CENTD1	centaurin, delta 1;CENTD1
GeneID:396	ARHGDIA	Rho GDP dissociation inhibitor (GDI) alpha;ARHGDIA
GeneID:55114	ARHGAP17	Rho GTPase activating protein 17;ARHGAP17
GeneID:826	CAPNS1	calpain, small subunit 1;CAPNS1
GeneID:1825	DSC1	desmocollin 3;DSC3
GeneID:4293	MAP3K9	mitogen-activated protein kinase kinase kinase 9;MAP3K9
GeneID:1069	CETN2	centrin, EF-hand protein, 2;CETN2

Cell structure and motility (BP00285) 13		
GeneID:301	ANXA1	annexin A1;ANXA1

GeneID:7278	TUBA2	tubulin, alpha 2;TUBA2
GeneID:51807	TUBA8	tubulin, alpha 8;TUBA8
GeneID:116984	CENTD1	centaurin, delta 1;CENTD1
GeneID:59	ACTA2	actin, alpha 2, smooth muscle, aorta;ACTA2
GeneID:7431	VIM	vimentin;VIM
GeneID:10376	K-ALPHA-1	tubulin, alpha, ubiquitous;K-ALPHA-1
GeneID:55114	ARHGAP17	Rho GTPase activating protein 17;ARHGAP17
GeneID:84790	TUBA6	tubulin, alpha 6;TUBA6
GeneID:4293	MAP3K9	mitogen-activated protein kinase kinase kinase 9;MAP3K9
GeneID:70	ACTC	actin, alpha, cardiac muscle;ACTC
GeneID:2670	GFAP	glial fibrillary acidic protein;GFAP
GeneID:8655	DNCL1	dynein, light chain, LC8-type 1;DYNLL1

Developmental processes (BP00193) 13		
GeneID:2626	GATA4	GATA binding protein 4;GATA4
GeneID:3980	LIG3	ligase III, DNA, ATP-dependent;LIG3
GeneID:207	AKT1	v-akt murine thymoma viral oncogene homolog 1;AKT1
GeneID:1827	DSC1	Down syndrome critical region gene 1;DSCR1
GeneID:8243	SMC1L1	SMC1 structural maintenance of chromosomes 1-like 1 (yeast);SMC1L1
GeneID:51773	HBXAP	hepatitis B virus x associated protein;HBXAP
GeneID:7486	WRN	Werner syndrome;WRN
GeneID:208	AKT2	v-akt murine thymoma viral oncogene homolog 2;AKT2
GeneID:2956	MSH6	mutS homolog 6 (E. coli);MSH6
GeneID:6949	TCOF1	Treacher Collins-Franceschetti syndrome 1;TCOF1
GeneID:7431	VIM	vimentin;VIM
GeneID:10000	AKT3	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma);AKT3
GeneID:8655	DNCL1	dynein, light chain, LC8-type 1;DYNLL1

Immunity and defense (BP00148) 13		
GeneID:2288	FKBP4	FK506 binding protein 4, 59kDa;FKBP4
GeneID:4293	MAP3K9	mitogen-activated protein kinase kinase kinase 9;MAP3K9
GeneID:23049	SMG1	PI-3-kinase-related kinase SMG-1;SMG1
GeneID:3303	HSPA1A	heat shock 70kDa protein 1A;HSPA1A
GeneID:7520	XRCC5	X-ray repair complementing defective repair in Chinese hamster cells 5 (double-strand-break rejoining; Ku autoantigen, 80kDa);XRCC5
GeneID:2547	G22P1	X-ray repair complementing defective repair in Chinese

		hamster cells 6 (Ku autoantigen, 70kDa);XRCC6
GeneID:3326	HSPCB	heat shock 90kDa protein 1, beta;HSPCB
GeneID:142	PARP1	poly (ADP-ribose) polymerase family, member 1;PARP1
GeneID:3320	HSPCA	heat shock 90kDa protein 1, alpha;HSPCA
GeneID:3312	HSPA8	heat shock 70kDa protein 8;HSPA8
GeneID:3324	HSPCA	heat shock 90kDa protein 1, alpha-like 3;HSPCAL3
GeneID:10038	PARP2	poly (ADP-ribose) polymerase family, member 2;PARP2
GeneID:3956	LGALS1	lectin, galactoside-binding, soluble, 1 (galectin 1);LGALS1
Intracellular protein traffic (BP00125) 10		
GeneID:3190	HNRPK	heterogeneous nuclear ribonucleoprotein K;HNRPK
GeneID:7278	TUBA2	tubulin, alpha 2;TUBA2
GeneID:51807	TUBA8	tubulin, alpha 8;TUBA8
GeneID:253959	GARNL1	GTPase activating Rap/RanGAP domain-like 1;GARNL1
GeneID:59	ACTA2	actin, alpha 2, smooth muscle, aorta;ACTA2
GeneID:10376	K-ALPHA-1	tubulin, alpha, ubiquitous;K-ALPHA-1
GeneID:84790	TUBA6	tubulin, alpha 6;TUBA6
GeneID:70	ACTC	actin, alpha, cardiac muscle;ACTC
GeneID:4678	NASP	nuclear autoantigenic sperm protein (histone-binding);NASP
GeneID:8655	DNCL1	dynein, light chain, LC8-type 1;DYNLL1
Cell proliferation and differentiation (BP00224) 5		
GeneID:207	AKT1	v-akt murine thymoma viral oncogene homolog 1;AKT1
GeneID:10000	AKT3	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma);AKT3
GeneID:51773	HBXAP	hepatitis B virus x associated protein;HBXAP
GeneID:208	AKT2	v-akt murine thymoma viral oncogene homolog 2;AKT2
GeneID:1069	CETN2	centrin, EF-hand protein, 2;CETN2
Transport (BP00141) 5		
GeneID:59	ACTA2	actin, alpha 2, smooth muscle, aorta;ACTA2
GeneID:9513	FXR2	fragile X mental retardation, autosomal homolog 2;FXR2
GeneID:8087	FXR1	fragile X mental retardation, autosomal homolog 1;FXR1
GeneID:70	ACTC	actin, alpha, cardiac muscle;ACTC
GeneID:2332	FMR1	fragile X mental retardation 1;FMR1
Nitrogen metabolism (BP00090) 4		
GeneID:207	AKT1	v-akt murine thymoma viral oncogene homolog 1;AKT1

GeneID:208	AKT2	v-akt murine thymoma viral oncogene homolog 2;AKT2
GeneID:10000	AKT3	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma);AKT3
GeneID:8655	DNCL1	dynein, light chain, LC8-type 1;DYNLL1

Cell adhesion (BP00124) 4		
GeneID:1823	DSC1	desmocollin 1;DSC1
GeneID:116984	CENTD1	centaurin, delta 1;CENTD1
GeneID:1825	DSC1	desmocollin 3;DSC3
GeneID:3956	LGALS1	lectin, galactoside-binding, soluble, 1 (galectin 1);LGALS1

Biological process unclassified (BP00216) 3		
GeneID:9774	BCLAF1	BCL2-associated transcription factor 1;BCLAF1
GeneID:128312	HIST3H2BB	histone 3, H2bb;HIST3H2BB
GeneID:6119	RPA3	replication protein A3, 14kDa;RPA3

Neuronal activities (BP00166) 2		
GeneID:2288	FKBP4	FK506 binding protein 4, 59kDa;FKBP4
GeneID:2521	FUS	fusion (involved in t(12;16) in malignant liposarcoma);FUS

Oncogenesis (BP00281) 3		
GeneID:2521	FUS	fusion (involved in t(12;16) in malignant liposarcoma);FUS
GeneID:286514	MAGEB18	melanoma antigen family B, 18;MAGEB18

Muscle contraction (BP00173) 1		
GeneID:1152	CKB	creatine kinase, brain;CKB

Lipid, fatty acid and steroid metabolism (BP00019) 1		
GeneID:301	ANXA1	annexin A1;ANXA1

Sensory perception (BP00182) 1		
GeneID:6949	TCOF1	Treacher Collins-Franceschetti syndrome 1;TCOF1

Table 3.7 Listing of all complex proteins in each of the 14 pathways

p53 pathway (P00059) 9		
GeneID:207	AKT1	v-akt murine thymoma viral oncogene homolog 1;AKT1
GeneID:8850	PCAF	p300/CBP-associated factor;PCAF

GeneID:3065	HDAC1	histone deacetylase 1;HDAC1
GeneID:7486	WRN	Werner syndrome;WRN
GeneID:23049	SMG1	PI-3-kinase-related kinase SMG-1;SMG1
GeneID:208	AKT2	v-akt murine thymoma viral oncogene homolog 2;AKT2
GeneID:3066	HDAC2	histone deacetylase 2;HDAC2
GeneID:5591	PRKDC	protein kinase, DNA-activated, catalytic polypeptide;PRKDC
GeneID:10000	AKT3	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma);AKT3

Apoptosis signaling pathway (P00006) 7		
GeneID:207	AKT1	v-akt murine thymoma viral oncogene homolog 1;AKT1
GeneID:10000	AKT3	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma);AKT3
GeneID:3303	HSPA1A	heat shock 70kDa protein 1A;HSPA1A
GeneID:9531	BAG3	BCL2-associated athanogene 3;BAG3
GeneID:208	AKT2	v-akt murine thymoma viral oncogene homolog 2;AKT2
GeneID:3312	HSPA8	heat shock 70kDa protein 8;HSPA8
GeneID:9131	PDCD8	programmed cell death 8 (apoptosis-inducing factor);PDCD8

p53 pathway feedback loops 2 (P04398) 5		
GeneID:207	AKT1	v-akt murine thymoma viral oncogene homolog 1;AKT1
GeneID:10000	AKT3	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma);AKT3
GeneID:208	AKT2	v-akt murine thymoma viral oncogene homolog 2;AKT2
GeneID:5591	PRKDC	protein kinase, DNA-activated, catalytic polypeptide;PRKDC
GeneID:23049	SMG1	PI-3-kinase-related kinase SMG-1;SMG1

Huntington disease (P00029) 5		
GeneID:207	AKT1	v-akt murine thymoma viral oncogene homolog 1;AKT1
GeneID:10000	AKT3	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma);AKT3
GeneID:208	AKT2	v-akt murine thymoma viral oncogene homolog 2;AKT2
GeneID:826	CAPNS1	calpain, small subunit 1;CAPNS1
GeneID:8655	DNCL1	dynein, light chain, LC8-type 1;DYNLL1

Wnt signaling pathway (P00057) 5		
GeneID:3065	HDAC1	histone deacetylase 1;HDAC1
GeneID:6597	SMARCA4	SWI/SNF related, matrix associated, actin dependent

		regulator of chromatin, subfamily a, member 4;SMARCA4
GeneID:8467	SMARCA5	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 5;SMARCA5
GeneID:3066	HDAC2	histone deacetylase 2;HDAC2
GeneID:6596	SMARCA3	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 3;SMARCA3

FAS signaling pathway (P00020) 5		
GeneID:207	AKT1	v-akt murine thymoma viral oncogene homolog 1;AKT1
GeneID:10000	AKT3	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma);AKT3
GeneID:208	AKT2	v-akt murine thymoma viral oncogene homolog 2;AKT2
GeneID:142	PARP1	poly (ADP-ribose) polymerase family, member 1;PARP1
GeneID:10038	PARP2	poly (ADP-ribose) polymerase family, member 2;PARP2

Insulin/IGF pathway-protein kinase B signaling cascade (P00033) 4		
GeneID:207	AKT1	v-akt murine thymoma viral oncogene homolog 1;AKT1
GeneID:10000	AKT3	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma);AKT3
GeneID:208	AKT2	v-akt murine thymoma viral oncogene homolog 2;AKT2
GeneID:253959	GARNL1	GTPase activating Rap/RanGAP domain-like 1;GARNL1

EGF receptor signaling pathway (P00018) 4		
GeneID:207	AKT1	v-akt murine thymoma viral oncogene homolog 1;AKT1
GeneID:10000	AKT3	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma);AKT3
GeneID:208	AKT2	v-akt murine thymoma viral oncogene homolog 2;AKT2
GeneID:23049	SMG1	PI-3-kinase-related kinase SMG-1;SMG1

DNA replication (P00017) 4		
GeneID:7155	TOP2B	topoisomerase (DNA) II beta 180kDa;TOP2B
GeneID:7153	TOP2A	topoisomerase (DNA) II alpha 170kDa;TOP2A
GeneID:6118	RPA2	replication protein A2, 32kDa;RPA2
GeneID:7150	TOP1	topoisomerase (DNA) I;TOP1

Angiogenesis (P00005) 3		
GeneID:207	AKT1	v-akt murine thymoma viral oncogene homolog 1;AKT1
GeneID:10000	AKT3	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma);AKT3
GeneID:208	AKT2	v-akt murine thymoma viral oncogene homolog 2;AKT2

Interleukin signaling pathway (P00036) 3		
GeneID:207	AKT1	v-akt murine thymoma viral oncogene homolog 1;AKT1
GeneID:10000	AKT3	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma);AKT3
GeneID:208	AKT2	v-akt murine thymoma viral oncogene homolog 2;AKT2

Inflammation mediated by chemokine and cytokine signaling pathway (P00031) 3		
GeneID:207	AKT1	v-akt murine thymoma viral oncogene homolog 1;AKT1
GeneID:10000	AKT3	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma);AKT3
GeneID:208	AKT2	v-akt murine thymoma viral oncogene homolog 2;AKT2

Hypoxia response via HIF activation (P00030) 3		
GeneID:207	AKT1	v-akt murine thymoma viral oncogene homolog 1;AKT1
GeneID:10000	AKT3	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma);AKT3
GeneID:208	AKT2	v-akt murine thymoma viral oncogene homolog 2;AKT2

Endothelin signaling pathway (P00019) 3		
GeneID:207	AKT1	v-akt murine thymoma viral oncogene homolog 1;AKT1
GeneID:10000	AKT3	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma);AKT3
GeneID:208	AKT2	v-akt murine thymoma viral oncogene homolog 2;AKT2

PI3 kinase pathway (P00048) 3		
GeneID:207	AKT1	v-akt murine thymoma viral oncogene homolog 1;AKT1
GeneID:10000	AKT3	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma);AKT3
GeneID:208	AKT2	v-akt murine thymoma viral oncogene homolog 2;AKT2

PDGF signaling pathway (P00047) 3		
GeneID:207	AKT1	v-akt murine thymoma viral oncogene homolog 1;AKT1
GeneID:10000	AKT3	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma);AKT3
GeneID:208	AKT2	v-akt murine thymoma viral oncogene homolog 2;AKT2

Ras Pathway (P04393) 3		
GeneID:10000	AKT3	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma);AKT3

GeneID:207	AKT1	v-akt murine thymoma viral oncogene homolog 1;AKT1
GeneID:208	AKT2	v-akt murine thymoma viral oncogene homolog 2;AKT2

VEGF signaling pathway (P00056) 3		
GeneID:10000	AKT3	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma);AKT3
GeneID:207	AKT1	v-akt murine thymoma viral oncogene homolog 1;AKT1
GeneID:208	AKT2	v-akt murine thymoma viral oncogene homolog 2;AKT2

T cell activation (P00053) 3		
GeneID:10000	AKT3	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma);AKT3
GeneID:207	AKT1	v-akt murine thymoma viral oncogene homolog 1;AKT1
GeneID:208	AKT2	v-akt murine thymoma viral oncogene homolog 2;AKT2

FGF signaling pathway (P00021) 3		
GeneID:10000	AKT3	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma);AKT3
GeneID:207	AKT1	v-akt murine thymoma viral oncogene homolog 1;AKT1
GeneID:208	AKT2	v-akt murine thymoma viral oncogene homolog 2;AKT2

Parkinson disease (P00049) 2		
GeneID:3303	HSPA1A	heat shock 70kDa protein 1A;HSPA1A
GeneID:3312	HSPA8	heat shock 70kDa protein 8;HSPA8

Oxidative stress response (P00046) 1		
GeneID:2521	FUS	fusion (involved in t(12;16) in malignant liposarcoma);FUS

Notch signaling pathway (P00045) 1		
GeneID:8850	PCAF	p300/CBP-associated factor;PCAF

Table 3.8 Complex proteins involving in mRNA transcription, DNA repair and DNA replication

mRNA transcription (BP00040) 30		
GeneID:84717	HDGF2	hepatoma-derived growth factor-related protein 2;HDGF2
GeneID:1107	CHD3	chromodomain helicase DNA binding protein 3;CHD3
GeneID:50810	HDGF2	hepatoma-derived growth factor, related protein 3;HDGFRP3
GeneID:3065	HDAC1	histone deacetylase 1;HDAC1
GeneID:7155	TOP2B	topoisomerase (DNA) II beta 180kDa;TOP2B
GeneID:8850	PCAF	p300/CBP-associated factor;PCAF
GeneID:51773	HBXAP	hepatitis B virus x associated protein;HBXAP
GeneID:3190	HNRPK	heterogeneous nuclear ribonucleoprotein K;HNRPK
GeneID:8467	SMARCA5	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 5;SMARCA5
GeneID:1106	CHD2	chromodomain helicase DNA binding protein 2;CHD2
GeneID:6664	SOX11	SRY (sex determining region Y)-box 11;SOX11
GeneID:6597	SMARCA4	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4;SMARCA4
GeneID:7153	TOP2A	topoisomerase (DNA) II alpha 170kDa;TOP2A
GeneID:23429	RYBP	RING1 and YY1 binding protein;RYBP
GeneID:3066	HDAC2	histone deacetylase 2;HDAC2
GeneID:51087	YBX2	Y box binding protein 2;YBX2
GeneID:4204	MECP2	methyl CpG binding protein 2 (Rett syndrome);MECP2
GeneID:8531	CSDA	cold shock domain protein A;CSDA
GeneID:7150	TOP1	topoisomerase (DNA) I;TOP1
GeneID:6749	SSRP1	structure specific recognition protein 1;SSRP1
GeneID:6596	SMARCA3	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 3;SMARCA3
GeneID:57459	GATAD2B	GATA zinc finger domain containing 2B;GATAD2B
GeneID:80314	EPC1	enhancer of polycomb homolog 1 (Drosophila);EPC1
GeneID:4904	NSEP1,YBX1	Y box binding protein 1;YBX1
GeneID:2626	GATA4	GATA binding protein 4;GATA4
GeneID:9406	ZNF265	zinc finger protein 265;ZNF265
GeneID:8091	HMGA2	high mobility group AT-hook 2;HMGA2
GeneID:171017	ZNF384	zinc finger protein 384;ZNF384

GeneID:3159	HMGA1	high mobility group AT-hook 1;HMGA1
GeneID:154790	CLEC2L	C-type lectin domain family 2, member L;CLEC2L

DNA repair (BP00036) 16		
GeneID:27301	APEX2	APEX nuclease (apurinic/apyrimidinic endonuclease) 2;APEX2
GeneID:5886	RAD23A	RAD23 homolog A (S. cerevisiae);RAD23A
GeneID:8243	SMC1L1	SMC1 structural maintenance of chromosomes 1-like 1 (yeast);SMC1L1
GeneID:6118	RPA2	replication protein A2, 32kDa;RPA2
GeneID:7508	XPC	xeroderma pigmentosum, complementation group C;XPC
GeneID:7520	XRCC5	X-ray repair complementing defective repair in Chinese hamster cells 5 (double-strand-break rejoining; Ku autoantigen, 80kDa);XRCC5
GeneID:5887	RAD23B	RAD23 homolog B (S. cerevisiae);RAD23B
GeneID:2956	MSH6	mutS homolog 6 (E. coli);MSH6
GeneID:2547	G22P1	X-ray repair complementing defective repair in Chinese hamster cells 6 (Ku autoantigen, 70kDa);XRCC6
GeneID:142	PARP1	poly (ADP-ribose) polymerase family, member 1;PARP1
GeneID:10038	PARP2	poly (ADP-ribose) polymerase family, member 2;PARP2
GeneID:23049	SMG1	PI-3-kinase-related kinase SMG-1;SMG1
GeneID:5591	PRKDC	protein kinase, DNA-activated, catalytic polypeptide;PRKDC
GeneID:3980	LIG3	ligase III, DNA, ATP-dependent;LIG3
GeneID:7507	XPA	xeroderma pigmentosum, complementation group A;XPA
GeneID:1069	CETN2	centrin, EF-hand protein, 2;CETN2

DNA replication (BP00035) 8		
GeneID:7155	TOP2B	topoisomerase (DNA) II beta 180kDa;TOP2B
GeneID:7486	WRN	Werner syndrome;WRN
GeneID:4673	NAP1L1	nucleosome assembly protein 1-like 1;NAP1L1
GeneID:7153	TOP2A	topoisomerase (DNA) II alpha 170kDa;TOP2A
GeneID:6118	RPA2	replication protein A2, 32kDa;RPA2
GeneID:4172	MCM3	MCM3 minichromosome maintenance deficient 3 (S. cerevisiae);MCM3
GeneID:7150	TOP1	topoisomerase (DNA) I;TOP1
GeneID:6117	RPA1	replication protein A1, 70kDa;RPA1

Chapter 4

In Vitro Acetylation Analysis of Ku70

4.1 Introduction

Post-translational modification (PTM) constitutes an important mechanism for modulating protein function. There are more than 200 described protein PTMs that vary from well-studied (e.g. phosphorylation) to those which are poorly understood (Krishna and Wold, 1993). More and more modifications have been discovered that they can impact the stability, interactions, biochemical activity and localization of proteins.

Lysine acetylation, or the transfer of an acetyl group from acetyl coenzyme A to the ε-amino group of a lysine residue, is a reversible and highly regulated posttranslational modification. It was initially discovered on histones about four decades ago (Vidali et al., 1968). It has been hypothesized that histone acetylation forms the histone code, a complex modification language that is fundamental to the regulation of chromatin structure and function (Strahl and Allis, 2000; Turner, 2000). It also regulates other nuclear processes such as DNA replication, recombination and repair.

However, it took another 30 years to identify the first nonhistone protein, p53, to be lysine acetylated (Gu, W. and Roeder, 1997) in 1997. Intensive research during the past decade has shown that acetylation is a common posttranslational modification and plays important roles in regulating the functions of eukaryotic and even viral and bacterial proteins (Kouzarides, 2000; Sterner and Berger, 2000). So far, about 40 transcription factors and over 30 other proteins have been described that can be acetylated by histone acetyltransferases (HATs) (Polevoda and Sherman, 2002; Yang, 2004;

Drummond et al., 2005). A more recent research applied a proteomic survey and identified 195 lysine-acetylated proteins derived from HeLa cells and mouse liver mitochondria (Kim, S. C. et al., 2006). The modification regulates diverse protein properties including chromatin dynamics and transcription (Kuo and Allis, 1998; Knoepfler and Eisenman, 1999; Kouzarides, 1999; Wolffe and Guschin, 2000), gene silencing (Bestor, 1998; Razin, 1998), cell cycle progression (Marzio et al., 2000; Morris et al., 2000; Takahashi et al., 2000; Zhang et al., 2000), apoptosis (Fu et al., 2004; Iyer et al., 2004; Roh et al., 2004), differentiation (Koipally et al., 1999; Wang, A. H. et al., 1999; Zhou et al., 2000; Miska et al., 2001; Sartorelli and Puri, 2001), DNA replication (Iizuka and Stillman, 1999; Aggarwal and Calvi, 2004), DNA repair (Gu, W. and Roeder, 1997; Sakaguchi et al., 1998; Liu et al., 1999; Abraham et al., 2000; Rubbi and Milner, 2003; Bernardi et al., 2004), nuclear import (Blander et al., 2002; Madison et al., 2002; Gay et al., 2003), neuronal repression (Huang et al., 1999; Roopra et al., 2000; Boutillier et al., 2003) and protein stability (Caron et al., 2005). In addition to its important roles in fundamental biology, lysine acetylation is also intimately linked to aging and several major diseases such as cancer, neurodegenerative disorders, and cardiovascular diseases (Carrozza et al., 2003; Blander and Guarente, 2004; McKinsey and Olson, 2004; Yang, 2004).

Acetylation is reversible *in vivo*. The reactions are catalysed by histone acetyltransferases (HATs) and histone deacetylases (HDACs). HATs transfer acetyl groups from acetyl-coenzyme A (AcCoA) to the ε-amino group

of specific lysines of substrate proteins while HDACs remove acetyl groups from specific acetyl-lysine residues in their substrates. HDACs can be suppressed by HDAC inhibitors through which HDACs inhibitors can result in protein acetylation indirectly (Johnstone and Licht, 2003). More than 20 acetyltransferases and 18 deacetylases have been identified so far (Marmorstein and Roth, 2001; Thiagalingam et al., 2003), but the mechanism of substrate selection and site specificity of these enzymes are still unknown. Interestingly, one substrate can interact and be acetylated by multiple acetyltransferases. How these HATs and HDACs interact with transcription factors and other non-histone proteins will be continuously proceeding in the coming few years. Because of the critical role of acetylation in the regulation of oncogenic processes, HDAC inhibitors have been developed as therapeutic agents for both hematologic malignancies and solid tumors (Egger et al., 2004).

Many techniques have been described to determine the presence of PTMs. Most of the known acetylated proteins have been identified by radioactive analysis or immuno detection. However, both radioactive analysis and immuno detection only provide evidence for acetylation but will not provide information about the number and location of acetylation sites in a protein. In recent years, the analysis of protein acetylation by mass spectrometry (MS) has become more and more popular since MALDI TOF MS and ESI MS provide fast and sensitive methods for the characterization of posttranslational modifications (Wilkins et al., 1999; Mann and Jensen, 2003;

McDonald and Yates, 2003). Mass spectrometric measurement of the molecular weight of *in vivo* or *in vitro* acetylated proteins can reveal the number and locations of acetyl groups attached to the protein besides the normal characterization. Since the signal of an acetylated protein or peptide is shifted by +42 Da per covalently bound acetyl group as compared to the signal of the unmodified protein or peptide, the number of acetylation sites in a protein can easily be determined by MS.

Ku70 is one of the proteins that can be acetylated by HAT. Five lysine residues in C terminal (539, 542, 544, 553, and 556) were reported to be acetylated by CBP and PCAF. Both Class I/II HDACs and Class III HDACs target Ku70 in vivo. A model was proposed that the acetylation of C terminal of Ku70 will release Bax from Ku70 thus activates the apoptosis process. Moreover, inhibition of HDAC activity by using TSA (class I and II HDACs inhibitor) and nicotinamide (class III HDACs inhibitor), results in increased acetylation of these residues , suggesting that the acetylation status of these same residues is also determined by HDACs . The similar result was demonstrated in NB cells.

Here we incubated Ku70 with two HATs, PCAF and p300, to induce the *in vitro* acetylation of Ku70 and employed MS to detect the acetylation site of Ku70 as well as the *in vivo* acetylation.

4.2 Materials and Methods

4.2.1 Plasmid Constructions

The Ku70 cDNA was amplified using the following oligonucleotides as primers for the PCR: forward primer 5' CTGTAGACATATGATGTCAGG GTGGGAGTC 3' and reverse primer 5' GATCGCTCACGCGGCCGCGTCC TGGAAGTGCT 3' which introduce a NdeI cutting site into the 5' terminal of Ku70 cDNA and a NotI cutting site into the 3'terminal of Ku70 cDNA. PCR product was purified from the gel (QIAquick Gel Extraction Kit, Qiagen, Germany) and double digested by NdeI and NotI. The digested PCR product was purified (QIAquick PCR Purification Kit, Qiagen, Germany) and cloned into the NdeI and NotI sites of pET42b (Novagen, Germany) in frame to generate pET42b-Chis-Ku70 which will produce Ku70 with a 6×His tag in its C terminal.

4.2.2 Protein Expression

The plasmid was transformed and expressed in BL21. The Correct integration of constructs was verified by colony PCR.

4.2.3 Time Course Analysis of Protein Expression

The correct colony was picked to inoculate a small amount of culture medium containing Kanamycin (50ug/ml) in a 15 ml tube and grew at 37°C

overnight. The overnight cultures were used to inoculate 100 ml media (with antibiotics) and grew at 37°C with vigorous shaking until an OD_{600} of 0.6 was reached. IPTG was added to induce expression to a final concentration of 1mM. The cultures were incubated for an additional 5 hours. Small amount of sample were collected at time points 0h, 1h, 2h, 3h, 4h,5h as well as a control which was incubated without induction by IPTG. Samples were spinned down and dissolved in 20ul 2×SDS loading dye and ready for electrophoresis.

4.2.4 Determination of Target Protein Solubility

E.Coli cell pellet was resuspended in 5ml of lysis buffer for native purification (50 mM NaH2PO4, 300 mM NaCl, 10 mM imidazole, 10 mM β-mecapitethanol, 0.5mM PMSF, pH8.0). Lysozyme was added to 1mg/ml and incubated on ice for 30 min. The lysates were then sonicated 6×10s with 10s pauses at 200-300 W and centrifuged at 10000g at 4°C for 20-30 mins. The supernatant and pellet were separated for electrophoresis.

4.2.5 Protein Purification

A Ni-NTA resin was added into a Poly-Prep column (Bio-Rad) and equilibrated by washing with 2×5 volumes of lysis buffer. Cleared cell lysate was added to Ni-NTA resin and mixed gently by shaking (200 rpm on a rotary shaker) at 4°C for 2h. The resin was washed by 2×5 volumes of wash buffer 1 (50 mM NaH2PO4, 300 mM NaCl, 40 mM imidazole, 0.5%TritonX-100, 10%

glycerol, 10 mM β-mecapitethanol, complete EDTA-free protease, 0.5 mM PMSF, pH 8.0) followed by 3×5 volume of wash buffer 2 (50 mM NaH2PO4, 1M NaCl, 40 mM imidazole, 0.5%TritonX-100, 10% glycerol, 10 mM β-mecapitethanol, complete EDTA-free protease inhibitor, 0.5 mM PMSF, pH 8.0) and 2×5 volume of wash buffer 1. The protein was eluted three times with 3 volumes of elution buffer (50 mM NaH2PO4, 300 mM NaCl, 350 mM imidazole, pH 8.0) for each time and incubated for 10 minutes. The purified Ku70 was quantified with Bradford method.

4.2.6 *In Vitro* Acetylation

4.2.6.1 PCAF Induced *In Vitro* Acetylation

10μg purified Ku70, 500ng PCAF, active (Upstate), 1nM Acetyl CoA and 5X HAT assay buffer (250mM Tris-HCl, pH 8.0, 0.5mM EDTA, 50% glycerol) were added into a tube. The whole system was top up to 50μl with sterile water and incubated in a 30°C shaking incubator for 60 minutes.

4.2.6.2 P300 Induced *In Vitro* Acetylation

10μg purified Ku70, 2μg p300 HAT domain (Upstate), 1nM Acetyl CoA and 5X HAT assay buffer (250mM Tris-HCl, pH 8.0, 0.5mM EDTA, 50% glycerol) were added into a tube. The whole system was top up to 30μl with sterile water and incubated in a 30°C shaking incubator for 60 minutes.

4.2.7 SDS-polyacrylamide Gel Electrophoresis (SDS-PAGE) and Western Blot

The purity of Ku70 was verified by SDS-polyacrylamide gel electrophoresis and the acetylation of Ku70 was verified by Western blotting. Anti-acetyl-lysine antibody (Cell Signaling) was used as primary antibody.

4.2.8 Simply Blue Staining

After electrophoresis, the gels were stained in Simply Blue solution (Invitrogen) by microwave method according to the manufacturer's instructions. The gel was then destained by MilliQ water until the background of the gel was no longer bluish.

4.2.9 Treatment of 293F Cells by HDAC Inhibitors

293F cells expressing TAP-tagged Ku70 were treated by 1μM TSA, 5 μM SAHA and 5mM nicotinamide respectively for 24h. Cells were then collected. Cell lysates were purified by TAP-tag purification.

4.2.10 Sample Preparation for Mass Spectrometry

The solution containing acetylated Ku70 was reduced by 10mM DTT ,

alkylated by 20mM IAA and digested by trypsin at the ratio of 1:50 (trypsin:protein) overnight (12-16h) at 37°C. After digestion, 0.1% TFA was added to terminate the reaction. The resulting peptides were sequenced by tandem mass spectrometry (LC-MS/MS).

4.2.11 Mass Spectrometry Analysis

For LC-MS/MS, samples were analyzed with a Finnigan Surveyor HPLC system coupled online to a LTQ ion trap mass spectrometer (Thermo Electron, San Jose, CA) equipped with nano-ESI source. A C18 nanocolumn (i.d. 75 μm, length 100 mm, tip 15 μm, New Objectives, Woburn, MA) was packed with Magic C18 particles (particle size 5 μm, pore size 300 Å, Michrom BioResources, Auburn, CA). The flow rate through the column was 200 nL/min. A water-acetonitrile gradient was employed with both mobile phases containing 0.1% formic acid. The gradient used was 5%-50% acetonitrile over 60 min. The injection volume was 100 μL. The electrospray mass spectra were acquired at a voltage of 1.8 kV, ion transfer tube temperature of 180 °C, collision gas pressure of 0.85 mTorr and normalized collision energy of 35% for MS2. Ion selection threshold was set to 500 counts for initiating MS2 while activation q was set to 0.25 and activation time to 30 ms. The MS scan range was 200 - 2000 m/z and mass spectra were collected in data-dependent MS2 mode in which five of the highest intensity peaks in each MS scan were chosen for CID, with an isolation width of 3 Da.

4.2.12 Data Analysis

MS/MS data were analyzed by MASCOT (Matrix Science, London) database search engine searching against IPI human database. Database search was performed with carbaminomethylation (Cys) set as fixed modification, oxidized methionine (+16 Da) and acetylation as viable modifications. Missed cleavages of trypsin were set to two. Peptide charges were confined to 1+, 2+ and 3+ and mass tolerance were set to 2.0 Da for peptide and 0.8 Da for fragment ions, respectively.

4.3 Results

4.3.1 Optimization of Protein Expression

BL21 cells were incubated for different time after induction by IPTG. Small amount of samples were collected at each time point for electrophoresis analysis. BL21 expressing Ku70 was lysed by lysis buffer for native purification. Supernatant and pellet were collected to determine the solubility of Ku70. As shown in Figure 4.1, Ku70 was induced to overexpress after the addition of IPTG. We found 5h is efficient for the expression of Ku70. Although most of Ku70 were still in the pellet, there are still soluble Ku70 in the supernatant. Exogenous Ku70 may form inclusion body in BL21 and are insoluble in lysis buffer for native purification.

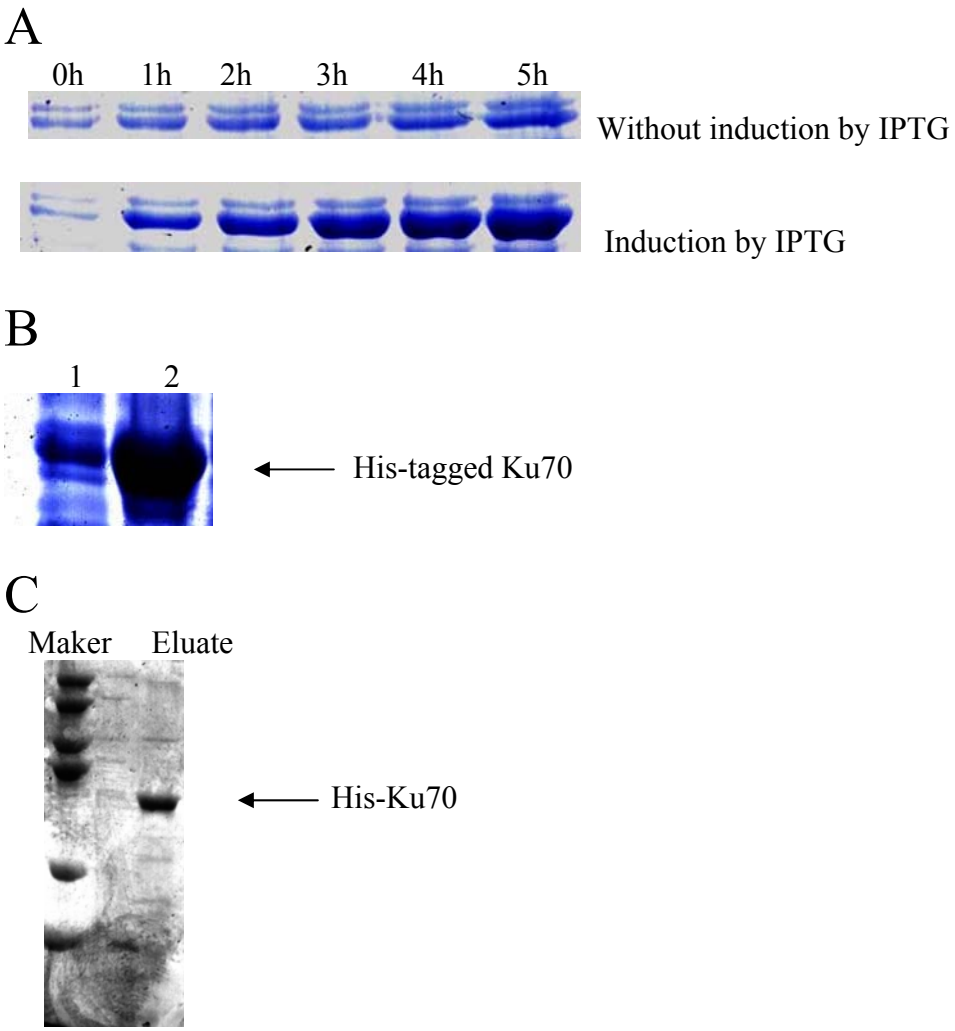


Figure 4.1

A. Time course analysis of protein expression
BL21 was induced by IPTG after an OD_{600} of 0.6 was reached. The cultures were incubated for an additional 5 hours. Small amount of sample were collected at timepoints 0h, 1h, 2h, 3h, 4h,5h as well as a control which was incubated without induction by IPTG.

B. Determination of Ku70 solubility in BL21
BL21 expressing his-tagged Ku70 was lysed by lysis buffer for native purification. The supernatant and pellet were separated for electrophoresis. 1, Supernatant for soluble proteins. 2, pellet for insoluble proteins.

C. Purification of His-tagged Ku70 on Ni-NTA matrices
BL21 was lysed by lysis buffer for native purification. The supernatant was purified on Ni-NTA matrices. The eluate was analyzed by SDS-PAGE.

4.3.2 Purification of His-tagged Ku70

The His-tagged Ku70 was purified on Ni-NTA matrices. But the purity was not high. We tried to optimize the condition by adding detergents, increasing salt concentration or changing concentration of imidazole in washing buffer. But the effect was not obvious (Figure 4.1). Although Ku70 was the predominant protein in eluate, there were still some other proteins.

4.3.3 *In Vitro* Acetylation of Ku70

The purified Ku70 was incubated with histone acetyltransferases PCAF and p300 respectively. PCAF and p300 are HATs that were used here to *in-vitro* acetylating Ku70. The aceytlation was detected by anti-acetyl-lysine antibody (Cell Signaling) (Figure 4.2, A). The acetylated Ku70 was then tryptic digested for mass spectrometry analysis. We repeated the experiments three times. Combined with all the data, we detected 21 acetylation sites by PCAF acetylation: K92, K182, K206, K299, K317, K351, K392, K443, K445, K461, K463, K468, K516, K539, K542, K565, K570, K575, K582, K591, K595 and 22 acetylation sites by p300 acetylation: K74, K94, K100, K129, K182, K189, K260, K265, K282, K287, K297, K299, K317, K445, K461, K463, K539, K542, K575, K582, K591, K595. Twelve lysines could be acetylated both by PCAF and p300. Nine lysines were acetylated only by PCAF and 10 lysines were acetylated only by p300.

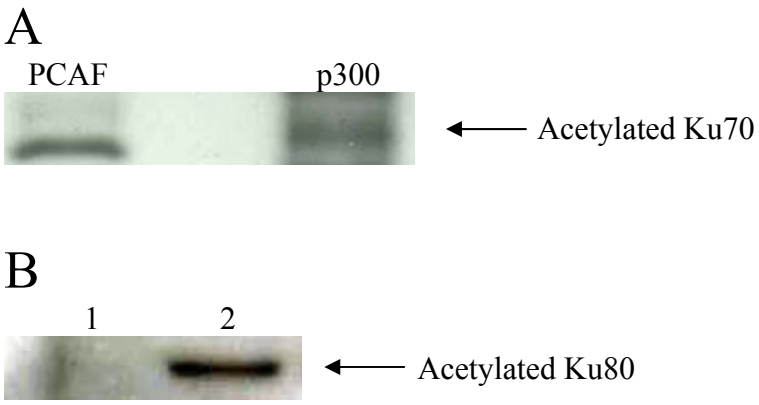


Figure 4.2 Acetylation of Ku70 and Ku80

A. His-tagged Ku70 was expressed in BL21and purified by Ni-NTA resin. The purified Ku70 was incubated with HATs- PCAF and p300. The acetylatioin was detected by anti-acetyl-lysine antibody.

B. The eluate from TAP tag purification of NTAP-tagged Ku70 was separated by SDS-PAGE and then transferred to PVDF membrane. Anti-acetyl-lysine antibody was used to detected acetylated proteins. Ku80 was found to be acetylated according to molecular weight.
1, Control-293F with empty vector
2, 293F-NTAP-Ku70

4.3.4 *In Vivo* Acetylation of Ku70

By including lysine acetylation as variable modification in Mascot search of TAP tag purified Ku70 from 293F cell, we found some acetylated lysines *in vivo* by endogenous acetyltransferase especially when cells were treated by HDAC inhibitors. Combined all the data, 19 acetylated lysine were found. They were K100, K245, K260, K282, K297, K299, K317, K451, K461, K468, K526, K539, K542, K544, K553, K556, K591, K595, and K596. Compared with the acetylated lysines in *in vitro* acetylation, 7 residues were novel. The comparison of *in vitro* and *in vivo* acetylation of Ku70 is shown in Table 4.1.

4.3.5 *In Vivo* Acetylation of Ku80

The acetylation of Ku80 was also detected by western blot analysis and mass spectrometry analysis. The eluate from TAP tag purification of NTAP-tagged Ku70 was separated by SDS-PAGE and then transferred to PVDF membrane. Anti-acetyl-lysine antibody was used to detect acetylated proteins. Ku80 was found to be acetylated according to its molecular weight (Figure 4.2, B). The peptides of Ku80 with acetylated lysines were also detected by mass spectrometry. We found 14 acetylated lysines: K60, K182, K197, K208, K248, K335, K338, K378, K391, K496, K587, K621, K718, and K755 in Ku80 protein.

In vitro	**PCAF**		92				182		206							299
	p300	74		94	100	129	182	189			260	265	282	287	297	299
In vivo					100					245	260		282		297	299
reported sites																

In vitro	**PCAF**	317			351	392	443	445		461	463	468	516		539	542
	p300	317						445		461	463				539	542
In vivo		317							451	461		468		526	539	542
reported sites		317	331	338											539	542

In vitro	**PCAF**				565	570	575	582	591	595	
	p300						575	582	591	595	
In vivo		544	553	556					591	595	596
reported sites		544	553	556							

Table 4.1 Comparision of acetylation sites of Ku70 under *in vitro, in vivo* acetylation and reported acetylation sit

4.4 Discussion

We first constructed a vector with His-tag at N-terminus of Ku70, but the purification was not good. Considering long recombinant proteins may be subject to premature termination and may produce a variety of truncated proteins, we constructed another vector placing His-tag at C-terminus of Ku70 so that only full-length proteins are purified. Functional Ku70 is essential to ensure the interaction between Ku70 and HATs. Incorrect folding will lead to the transfiguration of the interaction domain to prevent HATs from recognizing Ku70. However, eukaryotic proteins expressed intracellularly in *E.coli* are frequently sequestered into insoluble inclusion bodies that lack functional activity. So we first checked whether Ku70 would form inclusion body. The result showed that most of Ku70 were insoluble and lack of functionality. But some fraction of Ku70 still remained in a soluble and native form. They could be purified by Ni-NTA matrices under native condition.

The efficiency of His-tag purification was not ideal. Many other proteins were always co-purified by the nickel column. But purity was not an important element in our experiment. The enrichment of Ku70 would enable detection of more peptides and then improve the sequence coverage of the Ku70 protein by tandem mass spectrometry. Thus, more thorough acetylation sites could be discovered. The contamination by other proteins would not seriously affect the mass spectrometry analysis.

Eight lysines have been reported to be the target of HATs. They are K317, K331, K338, K539, K542, K544, K553 and K556. The acetylation of K539 and K542 are especially important for Bax-mediated apoptosis (Cohen et al., 2004). Except K331 and K338, all other lysines were detected by our *in vitro* and *in vivo* acetylation analysis. Although the peptide containing K331 and K338 was detected by MS, acetylation was not found. Thirty-two novel sites were discovered totally. K299, K461, K591 and K595 were both detected by *in vitro* and *in vivo* acetylation. There were slight differences between in vitro and in vivo acetylation indicating that although soluble Ku70 in *E.coli* was functional, it might be still a little different from native Ku70 in mammalian cells.

Ku80 was reported to be an acetylated protein (Cohen et al., 2004). But none of the acetylation sites were reported. Here we found 14 sites to be candidates of acetylated lysines for the first time. The sites need to be further validated by biochemical assays.

4.5 Conclusion

According to our *in vitro* and *in vivo* acetylation of Ku70 and Ku80, we found 38 acetylated lysines in Ku70, among which 32 were novel. K299, K461, K591 and K595 were found to be acetylated in both *in vivo* and *in vitro* samples. Fourteen acetylation sites were detected in Ku80 for the first time.

Chapter 5

Conclusion

Proteomics is still in its infancy stage. Innovative technology development is the key to advance the field to address more sophisticated biomedical questions. The protein sample preparation has direct and critical impact on the quality of mass spectrometric analysis, especially for low abundant protein. The current practical methods are not ideal and satisfactory yet. We have presented a novel sample preparation method for desalting and extraction of peptides from in-solution and in-gel digestion by using a small piece of Empore Disk. It can consistently detect down to 10 femtomol of BSA in in-solution digestion and 500 femtomol BSA in in-gel digestion. For in gel digestion, it is possible to extract and detect down to 10 femtomole of BSA; however, the results vary from experiment to experiment in this low level of proteins. It is not only suitable for sample preparation prior to MALDI-MS, but also LC-MS/MS analysis. The Empore Disk extraction minimizes the sample handling and transferring steps, thus reduces the losses of samples as well as offers a very effective and convenient tool for sample preparation. The procedure is simple, reproducible and economical. It is particularly suitable for parallel and batch samples preparation. These properties make it a useful tool for proteomics research.

Besides the importance of sample preparation prior to mass spectrometry analysis, the quality of protein complex purification from cell lysate plays a decisive role in the integrity of protein-protein interactions research. We have applied tandem affinity purification (TAP) method to pull down Ku70 complex in mammalian HEK293F cell line. This two-steps

protein purification technique isolated high quality Ku70 protein complex with minimal background. The protein complex was then identified by sensitive LTQ mass spectrometry. 151 proteins were found in the Ku70 complex. Most of them were involved in DNA repair, apoptosis, cell cycle regulation, gene transcription and signal transduction. Many Ku70 associated proteins participate in p53 pathway indicating the relationship between Ku70 and the development of cancer. PARP was validated to be a real Ku70 partner and interacted with Ku70 through Ku80. Our result provided an overview of the network of Ku70 in mammalian cells and clues for the roles of Ku70 in biological processes. We will validate and verify the roles of Ku70 in the apoptotic and signaling pathways by biochemical assays such as using RNAi technology to knock down the transcription level of *Ku70*.

The acetylation of Ku70 was proven to play an important role in modulating Ku70 biological functions. With our *in vitro* and *in vivo* Ku70 and Ku80 acetylation studies, 38 acetylation lysine sites in Ku70 were identified, among which 32 were novel. K299, K461, K591 and K595 were identified in both *in-vivo* and *in-vitro* analysis. Fourteen acetylation sites were detected in Ku80 for the first time.

Our studies indicated that Ku70 is an important protein with multiple functions. As it may be the key protein in linking the DNA damage to cellular apoptosis, it may be a potential target for cancer therapy. Our future research will focus on the detailed study on the roles of Ku70 in linking DNA damage and cellular apoptosis.

Reference

Abraham, J., J. Kelly, P. Thibault and S. Benchimol. (2000). Post-translational modification of p53 protein in response to ionizing radiation analyzed by mass spectrometry. *J. Mol. Biol.* **295**(4): 853-864.

Aebersold, R. and D. R. Goodlett. (2001). Mass spectrometry in proteomics. *Chem. Rev.* **101**(2): 269-295.

Aggarwal, B. D. and B. R. Calvi. (2004). Chromatin regulates origin activity in Drosophila follicle cells. *Nature*. **430**(6997): 372-376.

Anderson, C. W. (1993). DNA-Damage and the DNA-Activated Protein-Kinase. *Trends Biochem.Sci.* **18**(11): 433-437.

Aphasizhev, R., I. Aphasizheva, R. E. Nelson, G. H. Gao, A. M. Simpson, X. D. Kang, A. M. Falick, S. Sbicego and L. Simpson. (2003a). Isolation of a U-insertion/deletion editing complex from Leishmania tarentolae mitochondria. *Embo J.* **22**(4): 913-924.

Aphasizhev, R., I. Aphasizheva and L. Simpson. (2003b). A tale of two TUTases. *Proc. Natl. Acad. Sci. U. S. A.* **100**(19): 10617-10622.

Aravind, L. and E. V. Koonin. (2001). Prokaryotic homologs of the eukaryotic DNA-end-binding protein Ku, novel domains in the Ku protein and prediction of a prokaryotic double-strand break repair system. *Genome Res.* **11**(8): 1365-1374.

Bailey, S. M., J. Meyne, D. J. Chen, A. Kurimasa, G. C. Li, B. E. Lehnert and E. H. Goodwin. (1999). DNA double-strand break repair proteins are

required to cap the ends of mammalian chromosomes. *Proc. Natl. Acad. Sci. U. S. A.* **96**(26): 14899-14904.

Bernardi, R., P. P. Scaglioni, S. Bergmann, H. F. Horn, K. H. Vousden and P. P. Pandolfi. (2004). PML regulates p53 stability by sequestering Mdm2 to the nucleolus. *Nat. Cell Biol.* **6**(7): 665-672.

Bestor, T. H. (1998). Gene silencing - Methylation meets acetylation. *Nature*. **393**(6683): 311-312.

Bianchi, A. and T. de Lange. (1999). Ku binds telomeric DNA in vitro. *J. Biol. Chem.* **274**(30): 21223-21227.

Blander, G. and L. Guarente. (2004). The Sir2 family of protein deacetylases. *Annu. Rev. Biochem.* **73**(417-435.

Blander, G., N. Zalle, Y. Daniely, J. Taplick, M. D. Gray and M. Oren. (2002). DNA damage-induced translocation of the Werner helicase is regulated by acetylation. *J. Biol. Chem.* **277**(52): 50934-50940.

Boulton, S. J. and S. P. Jackson. (1996). Identification of a Saccharomyces cerevisiae Ku80 homologue: Roles in DNA double strand break rejoining and in telomeric maintenance. *Nucleic Acids Res.* **24**(23): 4639-4648.

Boutillier, A. L., E. Trinh and J. P. Loeffler. (2003). Selective E2F-dependent gene transcription is controlled by histone deacetylase activity during neuronal apoptosis. *J. Neurochem.* **84**(4): 814-828.

Bouwmeester, T., A. Bauch, H. Ruffner, P. O. Angrand, G. Bergamini, K. Croughton, C. Cruciat, D. Eberhard, J. Gagneur, S. Ghidelli, C. Hopf,

B. Huhse, R. Mangano, A. M. Michon, M. Schirle, J. Schlegl, M. Schwab, M. A. Stein, A. Bauer, G. Casari, G. Drewes, A. C. Gavin, D. B. Jackson, G. Joberty, G. Neubauer, J. Rick, B. Kuster and G. Superti-Furga. (2004). A physical and functional map of the human TNF-alpha NF-kappa B signal transduction pathway. *Nat. Cell Biol.* **6**(2): 97-+.

Brajenovic, M., G. Joberty, B. Kuster, T. Bouwmeester and G. Drewes. (2004). Comprehensive proteomic analysis of human par protein complexes reveals an interconnected protein network. *J. Biol. Chem.* **279**(13): 12804-12811.

Butland, G., J. M. Peregrin-Alvarez, J. Li, W. H. Yang, X. C. Yang, V. Canadien, A. Starostine, D. Richards, B. Beattie, N. Krogan, M. Davey, J. Parkinson, J. Greenblatt and A. Emili. (2005). Interaction network containing conserved and essential protein complexes in Escherichia coli. *Nature*. **433**(7025): 531-537.

Caron, C., C. Boyault and S. Khochbin. (2005). Regulatory cross-talk between lysine acetylation and ubiquitination: role in the control of protein stability. *Bioessays*. **27**(4): 408-415.

Carrozza, M. J., R. T. Utley, J. L. Workman and J. Cote. (2003). The diverse functions of histone acetyltransferase complexes. *Trends Genet.* **19**(6): 321-329.

Chechlacz, M., M. C. Vemuri and J. R. Naegele. (2001). Role of DNA-dependent protein kinase in neuronal survival. *J. Neurochem.* **78**(1): 141-154.

Cheeseman, I. M., C. Brew, M. Wolyniak, A. Desai, S. Anderson, N. Muster, J. R. Yates, T. C. Huffaker, D. G. Drubin and G. Barnes. (2001). Implication of a novel multiprotein Dam1p complex in outer kinetochore function. *J. Cell Biol.* **155**(7): 1137-1145.

Cohen, H. Y., S. Lavu, K. J. Bitterman, B. Hekking, T. A. Imahiyerobo, C. Miller, R. Frye, H. Ploegh, B. M. Kessler and D. A. Sinclair. (2004). Acetylation of the C terminus of Ku70 by CBP and PCAF controls Bax-mediated apoptosis. *Mol. Cell*. **13**(5): 627-638.

Critchlow, S. E., R. P. Bowater and S. P. Jackson. (1997). Mammalian DNA double-strand break repair protein XRCC4 interacts with DNA ligase IV. *Curr. Biol.* **7**(8): 588-598.

Dalziel, R. G., S. C. Mendelson and J. P. Quinn. (1992). The Nuclear Autoimmune Antigen Ku Is Also Present on the Cell-Surface. *Autoimmunity*. **13**(4): 265-267.

Delahunty, C. and J. R. Yates. (2005). Protein identification using 2D-LC-MS/MS. *Methods*. **35**(3): 248-255.

Denison, C., A. D. Rudner, S. A. Gerber, C. E. Bakalarski, D. Moazed and S. P. Gygi. (2005). A proteomic strategy for gaining insights into protein sumoylation in yeast. *Mol. Cell. Proteomics*. **4**(3): 246-254.

Devries, E., W. Vandriel, W. G. Bergsma, A. C. Arnberg and P. C. Vandervliet. (1989). Hela Nuclear-Protein Recognizing DNA Termini and Translocating on DNA Forming a Regular DNA Multimeric Protein Complex. *J. Mol. Biol.* **208**(1): 65-78.

Drummond, D. C., C. O. Noble, D. B. Kirpotin, Z. X. Guo, G. K. Scott and C. C. Benz. (2005). Clinical development of histone deacetylase inhibitors as anticancer agents. *Annu. Rev. Pharmacol. Toxicol.* **45**(495-528.

Dynan, W. S. and S. Yoo. (1998). Interaction of Ku protein and DNA-dependent protein kinase catalytic subunit with nucleic acids. *Nucleic Acids Res.* **26**(7): 1551-1559.

Egger, G., G. N. Liang, A. Aparicio and P. A. Jones. (2004). Epigenetics in human disease and prospects for epigenetic therapy. *Nature*. **429**(6990): 457-463.

Falzon, M., J. W. Fewell and E. L. Kuff. (1993). Ebp-80, a Transcription Factor Closely Resembling the Human Autoantigen Ku, Recognizes Single-Strand to Double-Strand Transitions in DNA. *J. Biol. Chem.* **268**(14): 10546-10552.

Featherstone, C. and S. P. Jackson. (1999). Ku, a DNA repair protein with multiple cellular functions? *Mutat. Res.-DNA Repair*. **434**(1): 3-15.

Feldmann, H. and E. L. Winnacker. (1993). A Putative Homolog of the Human Autoantigen Ku from Saccharomyces-Cerevisiae. *J. Biol. Chem.* **268**(17): 12895-12900.

Fenn, J. B., M. Mann, C. K. Meng, S. F. Wong and C. M. Whitehouse. (1989). Electrospray ionization for mass spectrometry of large biomolecules. **246**(4926): 64-71.

Fewell, J. W. and E. L. Kuff. (1996). Intracellular redistribution of Ku immunoreactivity in response to cell-cell contact and growth modulating components in the medium. *J. Cell Sci.* **109**(1937-1946.

Forler, D., T. Kocher, M. Rode, M. Gentzel, E. Izaurralde and M. Wilm. (2003). An efficient protein complex purification method for functional proteomics in higher eukaryotes. *Nat. Biotechnol.* **21**(1): 89-92.

Frit, P., P. Calsou, D. J. Chen and B. Salles. (1998). Ku70/Ku80 protein complex inhibits the binding of nucleotide excision repair proteins on linear DNA in vitro. *J. Mol. Biol.* **284**(4): 963-973.

FromontRacine, M., J. C. Rain and P. Legrain. (1997). Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens. *Nature Genet.* **16**(3): 277-282.

Fu, M. F., C. G. Wang, X. P. Zhang and R. G. Pestell. (2004). Acetylation of nuclear receptors in cellular growth and apoptosis. *Biochem. Pharmacol.* **68**(6): 1199-1208.

Gao, Y. J., J. Chaudhuri, C. M. Zhu, L. Davidson, D. T. Weaver and F. W. Alt. (1998). A targeted DNA-PKcs-null mutation reveals DNA-PK-independent functions for KU in V(D)J recombination. *Immunity*. **9**(3): 367-376.

Gay, F., D. Calvo, M. C. Lo, J. Ceron, M. Maduro, R. L. Lin and Y. Shi. (2003). Acetylation regulates subcellular localization of the Wnt signaling nuclear effector POP-1. *Genes Dev.* **17**(6): 717-722.

Giffin, W., H. Torrance, D. J. Rodda, G. G. Prefontaine, L. Pope and R. J. G. Hache. (1996). Sequence-specific DNA binding by Ku autoantigen and its effects on transcription. *Nature*. **380**(6571): 265-268.

Goodlett, D. R. and E. C. Yi. (2002). Proteomics without polyacrylamide: qualitative and quantitative uses of tandem mass spectrometry in proteome analysis. **2**(4-5): 138-153.

Gottlieb, T. M. and S. P. Jackson. (1993). The DNA-Dependent Protein-Kinase - Requirement for DNA Ends and Association with Ku Antigen. *Cell*. **72**(1): 131-142.

Graumann, J., L. A. Dunipace, J. H. Seol, W. H. McDonald, J. R. Yates, B. J. Wold and R. J. Deshaies. (2004). Applicability of tandem affinity purification MudPIT to pathway proteomics in yeast. *Mol. Cell. Proteomics*. **3**(3): 226-237.

Grawunder, U., N. Finnie, S. P. Jackson, B. Riwar and R. Jessberger. (1996). Expression of DNA-dependent protein kinase holoenzyme upon induction of lymphocyte differentiation and V(D)J recombination. *Eur. J. Biochem.* **241**(3): 931-940.

Gu, W. and R. G. Roeder. (1997). Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. *Cell*. **90**(4): 595-606.

Gu, Y. S., K. J. Seidl, G. A. Rathbun, C. M. Zhu, J. P. Manis, N. vanderStoep, L. Davidson, H. L. Cheng, J. M. Sekiguchi, K. Frank, P. StanhopeBaker, M. S. Schlissel, D. B. Roth and F. W. Alt. (1997).

Growth retardation and leaky SCID phenotype of Ku70-deficient mice. *Immunity*. **7**(5): 653-665.

Gu, Y. S., J. Sekiguchi, Y. J. Gao, P. Dikkes, K. Frank, D. Ferguson, P. Hasty, J. Chun and F. W. Alt. (2000). Defective embryonic neurogenesis in Ku-deficient but not DNA-dependent protein kinase catalytic subunit-deficient mice. *Proc. Natl. Acad. Sci. U. S. A.* **97**(6): 2668-2673.

Gullo, C., M. Au, G. Feng and G. Teoh. (2006). The biology of Ku and its potential oncogenic role in cancer. *Biochim. Biophys. Acta-Rev. Cancer*. **1765**(2): 223-234.

Gully, D., D. Moinier, L. Loiseau and E. Bouveret. (2003). New partners of acyl carrier protein detected in Escherichia coli by tandem affinity purification. *FEBS Lett.* **548**(1-3): 90-96.

Hannich, J. T., A. Lewis, M. B. Kroetz, S. J. Li, H. Heide, A. Emili and M. Hochstrasser. (2005). Defining the SUMO-modified proteome by multiple approaches in Saccharomyces cerevisiae. *J. Biol. Chem.* **280**(6): 4102-4110.

Hartley, K. O., D. Gell, G. C. M. Smith, H. Zhang, N. Divecha, M. A. Connelly, A. Admon, S. P. Leesmiller, C. W. Anderson and S. P. Jackson. (1995). DNA-Dependent Protein-Kinase Catalytic Subunit - a Relative of Phosphatidylinositol 3-Kinase and the Ataxia-Telangiectasia Gene-Product. *Cell*. **82**(5): 849-856.

Herceg, Z. and Z. Q. Wang. (2001). Functions of poly(ADP-ribose) polymerase (PARP) in DNA repair, genomic integrity and cell death. *Mutat. Res.-Fundam. Mol. Mech. Mutagen.* **477**(1-2): 97-110.

Hillenkamp, F. and M. Karas. (1990). Mass spectrometry of peptides and proteins by matrix-assisted ultraviolet laser desorption/ionization. **193**(280-295.

Honey, S., B. L. Schneider, D. M. Schieltz, J. R. Yates and B. Futcher. (2001). A novel multiple affinity purification tag and its use in identification of proteins associated with a cyclin-CDK complex. **29**(4): E24.

Horn, D. M., Y. Ge and F. W. McLafferty. (2000). Activated ion electron capture dissociation for mass spectral sequencing of larger (42 kDa) proteins. *Anal. Chem*. **72**(20): 4778-4784.

Hsieh, H. C., Y. H. Hsieh, Y. H. Huang, F. C. Shen, H. N. Tsai, J. H. Tsai, Y. T. Lai, Y. T. Wang, W. J. Chuang and W. Huang. (2005). HHR23A, a human homolog of Saccharomyces cerevisiae Rad23, regulates xeroderma pigmentosum C protein and is required for nucleotide excision repair. *Biochem. Biophys. Res. Commun.* **335**(1): 181-187.

Hsu, H. L., D. Gilley, E. H. Blackburn and D. J. Chen. (1999). Ku is associated with the telomere in mammals. *Proc. Natl. Acad. Sci. U. S. A.* **96**(22): 12454-12458.

Hsu, H. L., D. Gilley, S. A. Galande, M. P. Hande, B. Allen, S. H. Kim, G. C. Li, J. Campisi, T. Kohwi-Shigematsu and D. J. Chen. (2000). Ku acts

in a unique way at the mammalian telomere to prevent end joining. *Genes Dev.* **14**(22): 2807-2812.

Huang, Y. F., S. J. Myers and R. Dingledine. (1999). Transcriptional repression by REST: recruitment of Sin3A and histone deacetylase to neuronal genes. *Nat. Neurosci.* **2**(10): 867-872.

Iizuka, M. and B. Stillman. (1999). Histone acetyltransferase HBO1 interacts with the ORC1 subunit of the human initiator protein. *J. Biol. Chem.* **274**(33): 23027-23034.

Ishihama, Y., J. Rappsilber and M. Mann. (2006). Modular stop and go extraction tips with stacked disks for parallel and multidimensional peptide fractionation in proteomics. *J. Proteome Res.* **5**(4): 988-994.

Ismail, S. M., S. Prithivirajsingh, Y. Nimura and C. W. Stevens. (2004). Identification of proteins in the hamster DNA end-binding complex. *Int. J. Radiat. Biol.* **80**(4): 261-268.

Issaq, H. J., T. P. Conrads, G. M. Janini and T. D. Veenstra. (2002). Methods for fractionation, separation and profiling of proteins and peptides. *Electrophoresis*. **23**(17): 3048-3061.

Ito, T., K. Tashiro, S. Muta, R. Ozawa, T. Chiba, M. Nishizawa, K. Yamamoto, S. Kuhara and Y. Sakaki. (2000). Toward a protein-protein interaction map of the budding yeast: A comprehensive system to examine two-hybrid interactions in all possible combinations between the yeast proteins. *Proc. Natl. Acad. Sci. U. S. A.* **97**(3): 1143-1147.

Iyer, N. G., S. F. Chin, H. Ozdag, Y. Daigo, D. E. Hu, M. Cariati, K. Brindle, S. Aparicio and C. Caldas. (2004). P300 regulates p53-dependent apoptosis after DNA damage in colorectal cancer cells by modulation of PUMA/p21 levels. *Proc. Natl. Acad. Sci. U. S. A.* **101**(19): 7386-7391.

Jin, S. F. and D. T. Weaver. (1997). Double-strand break repair by Ku70 requires heterodimerization with Ku80 and DNA binding functions. *Embo J.* **16**(22): 6874-6885.

Johnstone, R. W. and J. D. Licht. (2003). Histone deacetylase inhibitors in cancer therapy: Is transcription the primary target? *Cancer Cell*. **4**(1): 13-18.

Karas, M. and F. Hillenkamp. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. **60**(20): 2299-2301.

Khanna, K. K. and S. P. Jackson. (2001). DNA double-strand breaks: signaling, repair and the cancer connection. *Nature Genet.* **27**(3): 247-254.

Kim, S. C., R. Sprung, Y. Chen, Y. D. Xu, H. Ball, J. M. Pei, T. L. Cheng, Y. Kho, H. Xiao, L. Xiao, N. V. Grishin, M. White, X. J. Yang and Y. M. Zhao. (2006). Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. *Mol. Cell*. **23**(4): 607-618.

Kim, S. H., D. Kim, J. S. Han, C. S. Jeong, B. S. Chung, C. D. Kang and G. C. Li. (1999). Ku autoantigen affects the susceptibility to anticancer drugs. *Cancer Res.* **59**(16): 4012-4017.

Knoepfler, P. S. and R. N. Eisenman. (1999). Sin meets NuRD and other tails of repression. *Cell*. **99**(5): 447-450.

Knuesel, M., Y. Wan, Z. Xiao, E. Holinger, N. Lowe, W. Wang and X. D. Liu. (2003). Identification of novel protein-protein interactions using a versatile mammalian tandem affinity purification expression system. *Mol. Cell. Proteomics*. **2**(11): 1225-1233.

Knuth, M. W., S. I. Gunderson, N. E. Thompson, L. A. Strasheim and R. R. Burgess. (1990). Purification and Characterization of Proximal Sequence Element-Binding Protein-1, a Transcription Activating Protein Related to Ku and Tref That Binds the Proximal Sequence Element of the Human U1 Promoter. *J. Biol. Chem.* **265**(29): 17911-17920.

Koike, M. (2002). Dimerization, translocation and localization of Ku70 and Ku80 proteins. *J. Radiat. Res.* **43**(3): 223-236.

Koike, M., T. Awaji, M. Kataoka, G. Tsujimoto, T. Kartasova, A. Koike and T. Shiomi. (1999). Differential subcellular localization of DNA-dependent protein kinase components Ku and DNA-PKcs during mitosis. *J. Cell Sci.* **112**(22): 4031-4039.

Koipally, J., A. Renold, J. Kim and K. Georgopoulos. (1999). Repression by Ikaros and Aiolos is mediated through histone deacetylase complexes. *Embo J.* **18**(11): 3090-3100.

Komuro, Y., T. Watanabe, Y. Hosoi, Y. Matsumoto, K. Nakagawa, N. Suzuki and H. Nagawa. (2005). Prognostic significance of Ku70 protein

expression in patients with advanced colorectal cancer. *Hepato-Gastroenterol.* **52**(64): 995-998.

Korabiowska, M., C. Cordon-Cardo, M. Schinagl, M. Karaus, J. Stachura, H. Schutz and U. Fischer. (2006a). Loss of Ku70/Ku80 expression occurs more frequently in hereditary than in sporadic colorectal tumors. Tissue microarray study. *Hum. Pathol.* **37**(4): 448-452.

Korabiowska, M., J. Voltmann, J. F. Honig, P. Bortkiewicz, F. Konig, C. Cordon-Cardo, F. Jenckel, P. Ambrosch and G. Fischer. (2006b). Altered expression of DNA double-strand repair genes Ku70 and Ku80 in carcinomas of the oral cavity. *Anticancer Res.* **26**(3A): 2101-2105.

Kouzarides, T. (1999). Histone acetylases and deacetylases in cell proliferation. *Curr. Opin. Genet. Dev.* **9**(1): 40-48.

Kouzarides, T. (2000). Acetylation: a regulatory modification to rival phosphorylation? *Embo J.* **19**(6): 1176-1179.

Krishna, R. G. and F. Wold. (1993). Post-translational modification of proteins. **67**(265-298.

Kuhn, A., V. Stefanovsky and I. Grummt. (1993). The Nucleolar Transcription Activator Ubf Relieves Ku Antigen-Mediated Repression of Mouse Ribosomal Gene-Transcription. *Nucleic Acids Res.* **21**(9): 2057-2063.

Kumar, J. K., S. Tabor and C. C. Richardson. (2004). Proteomic analysis of thioredoxin-targeted proteins in Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **101**(11): 3759-3764.

Kuo, M. H. and C. D. Allis. (1998). Roles of histone acetyltransferases and deacetylases in gene regulation. *Bioessays*. **20**(8): 615-626.

Larsen, M. R., S. J. Cordwell and P. Roepstorff. (2002). Graphite powder as an alternative or supplement to reversed-phase material for desalting and concentration of peptide mixtures prior to matrix-assisted laser desorption/ionization-mass spectrometry. *Proteomics*. **2**(9): 1277-1287.

Lee, S. W., K. J. Cho, J. H. Park, S. Y. Kim, S. Y. Nam, B. J. Lee, S. B. Kim, S. H. Choi, J. H. Kim, S. Do Ahn, S. S. Shin, E. K. Choi and E. Yu. (2005). Expressions of Ku70 and DNA-PKCS as prognostic indicators of local control in nasopharyngeal carcinoma. *Int. J. Radiat. Oncol. Biol. Phys.* **62**(5): 1451-1457.

Leesmiller, S. P., Y. R. Chen and C. W. Anderson. (1990). Human-Cells Contain a DNA-Activated Protein-Kinase That Phosphorylates Simian Virus-40 T-Antigen, Mouse P53, and the Human Ku-Autoantigen. *Mol. Cell. Biol.* **10**(12): 6472-6481.

Li, G. C., H. H. Ouyang, X. L. Li, H. Nagasawa, J. B. Little, D. J. Chen, C. C. Ling, Z. Fuks and C. Cordon-Cardo. (1998). Ku70: A candidate tumor suppressor gene for murine T cell lymphoma. *Mol. Cell.* **2**(1): 1-8.

Lindahl, T., M. S. Satoh, G. G. Poirier and A. Klungland. (1995). Posttranslational Modification of Poly(Adp-Ribose) Polymerase Induced by DNA Strand Breaks. *Trends Biochem.Sci.* **20**(10): 405-411.

Liu, L., D. M. Scolnick, R. C. Trievel, H. B. Zhang, R. Marmorstein, T. D. Halazonetis and S. L. Berger. (1999). p53 sites acetylated in vitro by

PCAF and p300 are acetylated in vivo in response to DNA damage. *Mol. Cell. Biol.* **19**(2): 1202-1209.

Madison, D. L., P. Yaciuk, R. P. S. 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**(41): 38755-38763.

Mann, M. and O. N. Jensen. (2003). Proteomic analysis of post-translational modifications. *Nat. Biotechnol.* **21**(3): 255-261.

Marmorstein, R. and S. Y. Roth. (2001). Histone acetyltransferases: function, structure, and catalysis. *Curr. Opin. Genet. Dev.* **11**(2): 155-161.

Marzio, G., C. Wagener, M. I. Gutierrez, P. Cartwright, K. Helin and M. Giacca. (2000). E2F family members are differentially regulated by reversible acetylation. *J. Biol. Chem.* **275**(15): 10887-10892.

McCormack, A. L., D. M. Schieltz, B. Goode, S. Yang, G. Barnes, D. Drubin and J. R. Yates, 3rd. (1997). Direct analysis and identification of proteins in mixtures by LC/MS/MS and database searching at the low-femtomole level. **69**(4): 767-776.

McDonald, W. H. and J. R. Yates. (2003). Shotgun proteomics: Integrating technologies to answer biological questions. *Curr. Opin. Mol. Ther.* **5**(3): 302-309.

McKinsey, T. A. and E. N. Olson. (2004). Cardiac histone acetylation - therapeutic opportunities abound. *Trends Genet.* **20**(4): 206-213.

Mimori, T., J. A. Hardin and J. A. Steitz. (1986). Characterization of the DNA-Binding Protein Antigen-Ku Recognized by Autoantibodies from Patients with Rheumatic Disorders. *J. Biol. Chem.* **261**(5): 2274-2278.

Miska, E. A., E. Langley, D. Wolf, C. Karlsson, J. Pines and T. Kouzarides. (2001). Differential localization of HDAC4 orchestrates muscle differentiation. *Nucleic Acids Res.* **29**(16): 3439-3447.

Morris, L., K. E. Allen and N. B. La Thangue. (2000). Regulation of E2F transcription by cyclinE-Cdk2 kinase mediated through p300/CBP co-activators. *Nat. Cell Biol.* **2**(4): 232-239.

Naldrett, M. J., R. Zeidler, K. E. Wilson and A. Kocourek. (2005). Concentration and desalting of peptide and protein samples with a newly developed C18 membrane in a microspin column format. *J Biomol Tech*. **16**(4): 423-428.

Nussenzweig, A., C. H. Chen, V. D. Soares, M. Sanchez, K. Sokol, M. C. Nussenzweig and G. C. Li. (1996). Requirement for Ku80 in growth and immunoglobulin V(D)J recombination. *Nature*. **382**(6591): 551-555.

Ouyang, H., A. Nussenzweig, A. Kurimasa, V. D. Soares, X. L. Li, C. CordonCardo, W. H. Li, N. Cheong, M. Nussenzweig, G. Iliakis, D. J. Chen and G. C. Li. (1997). Ku70 is required for DNA repair but not for T cell antigen receptor gene recombination in vivo. *J. Exp. Med.* **186**(6): 921-929.

Paillard, S. and F. Strauss. (1991). Analysis of the Mechanism of Interaction of Simian Ku Protein with DNA. *Nucleic Acids Res.* **19**(20): 5619-5624.

Panigrahi, A. K., A. Schnaufer, N. L. Ernst, B. B. Wang, N. Carmean, R. Salavati and K. Stuart. (2003). Identification of novel components of Trypanosoma brucei editosomes. *RNA-Publ. RNA Soc.* **9**(4): 484-492.

Pleschke, J. M., H. E. Kleczkowska, M. Strohm and F. E. Althaus. (2000). Poly(ADP-ribose) binds to specific domains in DNA damage checkpoint proteins. *J. Biol. Chem.* **275**(52): 40974-40980.

Polevoda, B. and F. Sherman. (2002). The diversity of acetylated proteins. **3**(5): reviews0006.

Prabhakar, B. S., G. P. Allaway, J. Srinivasappa and A. L. Notkins. (1990). Cell-Surface Expression of the 70-Kd Component of Ku, a DNA-Binding Nuclear Autoantigen. *J. Clin. Invest.* **86**(4): 1301-1305.

Puig, O., F. Caspary, G. Rigaut, B. Rutz, E. Bouveret, E. Bragado-Nilsson, M. Wilm and B. Seraphin. (2001). The tandem affinity purification (TAP) method: A general procedure of protein complex purification. *Methods*. **24**(3): 218-229.

Rappsilber, J., Y. Ishihama and M. Mann. (2003). Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal. Chem.* **75**(3): 663-670.

Rashmi, R., S. Kumar and D. Karunagaran. (2004). Ectopic expression of Bcl-XL or Ku70 protects human colon cancer cells (SW480) against curcumin-induced apoptosis while their down-regulation potentiates it. *Carcinogenesis*. **25**(10): 1867-1877.

Rathmell, W. K. and G. Chu. (1994). Involvement of the Ku Autoantigen in the Cellular-Response to DNA Double-Strand Breaks. *Proc. Natl. Acad. Sci. U. S. A.* **91**(16): 7623-7627.

Razin, A. (1998). CpG methylation, chromatin structure and gene silencing - a three-way connection. *Embo J.* **17**(17): 4905-4908.

Reeves, W. H. (1992). Antibodies to the P70/P80 (Ku) Antigens in Systemic Lupus-Erythematosus. *Rheum. Dis. Clin. North Am.* **18**(2): 391-414.

Rigaut, G., A. Shevchenko, B. Rutz, M. Wilm, M. Mann and B. Seraphin. (1999). A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* **17**(10): 1030-1032.

Roberts, M. R., W. K. Miskimins and F. H. Ruddle. (1989). Nuclear Proteins Tref1 and Tref2 Bind to the Transcriptional Control Element of the Transferrin Receptor Gene and Appear to Be Associated as a Heterodimer. **1**(1): 151-164.

Rodgers, W., S. J. Jordan and J. D. Capra. (2002a). Transient association of Ku with nuclear substrates characterized using fluorescence photobleaching. **168**(5): 2348-2355.

Rodgers, W., S. J. Jordan and J. D. Capra. (2002b). Transient association of Ku with nuclear substrates characterized using fluorescence photobleaching. *J. Immunol.* **168**(5): 2348-2355.

Rodriguez, K., J. Talamantez, W. Huang, S. H. Reed, Z. Wang, L. Chen, W. J. Feaver, E. C. Friedberg and A. E. Tomkinson. (1998). Affinity purification and partial characterization of a yeast multiprotein complex for nucleotide excision repair using histidine-tagged Rad14 protein. *J. Biol. Chem.* **273**(51): 34180-34189.

Roh, M. S., C. W. Kim, B. S. Park, G. C. Kim, J. H. Jeong, H. C. Kwon, D. J. Suh, K. H. Cho, S. B. Yee and Y. H. Yoo. (2004). Mechanism of histone deacetylase inhibitor Trichostatin A induced apoptosis in human osteosarcoma cells. *Apoptosis*. **9**(5): 583-589.

Rohila, J. S., M. Chen, R. Cerny and M. E. Fromm. (2004). Improved tandem affinity purification tag and methods for isolation of protein heterocomplexes from plants. *Plant J.* **38**(1): 172-181.

Roopra, A., L. Sharling, I. C. Wood, T. Briggs, U. Bachfischer, A. J. Paquette and N. J. Buckley. (2000). Transcriptional repression by neuron-restrictive silencer factor is mediated via the SIN3-histone deacetylase complex. *Mol. Cell. Biol.* **20**(6): 2147-2157.

Rubbi, C. P. and J. Milner. (2003). p53 is a chromatin accessibility factor for nucleotide excision repair of DNA damage. *Embo J.* **22**(4): 975-986.

Sadri-Vakili, G. and J. H. Cha. (2006). Histone deacetylase inhibitors: a novel therapeutic approach to Huntington's disease (complex mechanism of neuronal death). **3**(4): 403-408.

Sakaguchi, K., J. E. Herrera, S. Saito, T. Miki, M. Bustin, A. Vassilev, C. W. Anderson and E. Appella. (1998). DNA damage activates p53 through a phosphorylation-acetylation cascade. *Genes Dev.* **12**(18): 2831-2841.

Samper, E., F. A. Goytisolo, P. Slijepcevic, P. P. W. van Buul and M. A. Blasco. (2000). Mammalian Ku86 protein prevents telomeric fusions independently of the length of TTAGGG repeats and the G-strand overhang. *EMBO Rep.* **1**(3): 244-252.

Sartorelli, V. and P. L. Puri. (2001). The link between chromatin structure, protein acetylation and cellular differentiation. *Front. Biosci.* **6**(D1024-D1047.

Sawada, M., P. Hayes and S. Matsuyama. (2003a). Cytoprotective membrane-permeable peptides designed from the Bax-binding domain of Ku70. *Nat. Cell Biol.* **5**(4): 352-357.

Sawada, M., W. Y. Sun, P. Hayes, K. Leskov, D. A. Boothman and S. Matsuyama. (2003b). Ku70 suppresses the apoptotic translocation of Bax to mitochondria. *Nat. Cell Biol.* **5**(4): 320-329.

Schar, P., M. Fasi and R. Jessberger. (2004). SMC1 coordinates DNA double-strand break repair pathways. *Nucleic Acids Res.* **32**(13): 3921-3929.

Schild-Poulter, C., A. Shih, N. C. Yarymowich and R. J. G. Hache. (2003). Down-regulation of histone H2B by DNA-dependent protein kinase in

response to DNA damage through modulation of octamer transcription factor 1. *Cancer Res.* **63**(21): 7197-7205.

Schweitzer, J. K. and C. D'Souza-Schorey. (2005). A requirement for ARF6 during the completion of cytokinesis. *Exp. Cell Res.* **311**(1): 74-83.

Shen, Y. F. and R. D. Smith. (2002). Proteomics based on high-efficiency capillary separations. *Electrophoresis*. **23**(18): 3106-3124.

Shukla, A. K. and J. H. Futrell. (2000). Tandem mass spectrometry: dissociation of ions by collisional activation. *J. Mass Spectrom.* **35**(9): 1069-1090.

Sleno, L. and D. A. Volmer. (2004). Ion activation methods for tandem mass spectrometry. *J. Mass Spectrom.* **39**(10): 1091-1112.

Spagnolo, L., A. Rivera-Calzada, L. H. Pearl and O. Llorca. (2006). Three-dimensional structure of the human DNA-PKcs/Ku70/Ku80 complex assembled on DNA and its implications for DNA DSB repair. *Mol. Cell*. **22**(4): 511-519.

Sterner, D. E. and S. L. Berger. (2000). Acetylation of histones and transcription-related factors. *Microbiol. Mol. Biol. Rev.* **64**(2): 435-+.

Stewart, II, T. Thomson and D. Figeys. (2001). O-18 Labeling: a tool for proteomics. *Rapid Commun. Mass Spectrom.* **15**(24): 2456-2465.

Strahl, B. D. and C. D. Allis. (2000). The language of covalent histone modifications. *Nature*. **403**(6765): 41-45.

Stuiver, M. H., F. E. J. Coenjaerts and P. C. Vandervliet. (1990). The Autoantigen Ku Is Indistinguishable from Nf Iv, a Protein Forming Multimeric Protein-DNA Complexes. *J. Exp. Med.* **172**(4): 1049-1054.

Subramanian, C., A. W. Opipari, X. Bian, V. P. Castle and R. P. S. Kwok. (2005). Ku70 acetylation mediates neuroblastoma cell death induced by histone deacetylase inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **102**(13): 4842-4847.

Sucharov, C. C., S. M. Helmke, S. J. Langer, M. B. Perryman, M. Bristow and L. Leinwand. (2004). The Ku protein complex interacts with YY1, is up-regulated in human heart failure, and represses alpha myosin heavy-chain gene expression. *Mol. Cell. Biol.* **24**(19): 8705-8715.

Tagwerker, C., K. Flick, M. Cui, C. Guerrero, Y. M. Dou, B. Auer, P. Baldi, L. Huang and P. Kaiser. (2006). A tandem affinity tag for two-step purification under fully denaturing conditions - Application in ubiquitin profiling and protein complex identification combined with in vivo cross-linking. *Mol. Cell. Proteomics*. **5**(4): 737-748.

Takahashi, Y., J. B. Rayman and B. D. Dynlacht. (2000). Analysis of promoter binding by the E2F and pRB families in vivo: distinct E2F proteins mediate activation and repression. *Genes Dev.* **14**(7): 804-816.

Thiagalingam, S., K. H. Cheng, H. J. Lee, N. Mineva, A. Thiagalingam and J. F. Ponte (2003). Histone deacetylases: Unique players in shaping the epigenetic histone code. Epigenetics in Cancer Prevention: Early

Detection and Risk Assessment. New York, New York Acad Sciences. **983:** 84-100.

Tong, W. M., U. Cortes, M. P. Hande, H. Ohgaki, L. R. Cavalli, P. M. Lansdorp, B. R. Haddad and Z. Q. Wang. (2002). Synergistic role of Ku80 and poly(ADP-ribose) polymerase in suppressing chromosomal aberrations and liver cancer formation. *Cancer Res.* **62**(23): 6990-6996.

Toth, E. C., L. Marusic, A. Ochem, A. Patthy, S. Pongor, M. Giacca and A. Falaschi. (1993). Interactions of Usf and Ku Antigen with a Human DNA Region Containing a Replication Origin. *Nucleic Acids Res.* **21**(14): 3257-3263.

Turner, B. M. (2000). Histone acetylation and an epigenetic code. *Bioessays*. **22**(9): 836-845.

Tuteja, N., R. Tuteja, A. Ochem, P. Taneja, N. W. Huang, A. Simoncsits, S. Susic, K. Rahman, L. Marusic, J. Q. Chen, J. W. Zhang, S. G. Wang, S. Pongor and A. Falaschi. (1994). Human DNA Helicase-Ii - a Novel DNA Unwinding Enzyme Identified as the Ku Autoantigen. *Embo J.* **13**(20): 4991-5001.

Tuteja, R. and N. Tuteja. (2000). Ku autoantigen: A multifunctional DNA-binding protein. *Crit. Rev. Biochem. Mol. Biol.* **35**(1): 1-33.

Uetz, P., L. Giot, G. Cagney, T. A. Mansfield, R. S. Judson, J. R. Knight, D. Lockshon, V. Narayan, M. Srinivasan, P. Pochart, A. Qureshi-Emili, Y. Li, B. Godwin, D. Conover, T. Kalbfleisch, G. Vijayadamodar, M. J. Yang, M. Johnston, S. Fields and J. M. Rothberg. (2000). A

comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. *Nature*. **403**(6770): 623-627.

Vidali, G., E. L. Gershey and V. G. Allfrey. (1968). Chemical studies of histone acetylation. The distribution of epsilon-N-acetyllysine in calf thymus histones. **243**(24): 6361-6366.

Walker, J. R., R. A. Corpina and J. Goldberg. (2001). Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature*. **412**(6847): 607-614.

Wang, A. H., N. R. Bertos, M. Vezmar, N. Pelletier, M. Crosato, H. H. Heng, J. Th'ng, J. H. Han and X. J. Yang. (1999). HDAC4, a human histone deacetylase related to yeast HDA1, is a transcriptional corepressor. *Mol. Cell. Biol.* **19**(11): 7816-7827.

Wang, J. S., X. W. Dong, K. J. Myung, E. A. Hendrickson and W. H. Reeves. (1998). Identification of two domains of the p70 Ku protein mediating dimerization with p80 and DNA binding. *J. Biol. Chem.* **273**(2): 842-848.

Weaver, D., N. Boubnov, Z. Wills, K. Hall and J. Staunton (1995). V(D)J recombination: Double-strand break repair gene products used in the joining mechanism. Immunoglobulin Gene Expression in Development and Disease. New York, New York Acad Sciences. **764:** 99-111.

Weinfeld, M., M. A. Chaudhry, D. Damours, J. D. Pelletier, G. G. Poirier, L. F. Povirk and S. P. LeesMiller. (1997). Interaction of DNA-dependent

protein kinase and poly(ADP-ribose) polymerase with radiation-induced DNA strand breaks. *Radiat. Res.* **148**(1): 22-28.

Westermarck, J., C. Weiss, R. Saffrich, J. Kast, A. M. Musti, M. Wessely, W. Ansorge, B. Seraphin, M. Wilm, B. C. Valdez and D. Bohmann. (2002). The DEXD/H-box RNA helicase RHII/Gu is a co-factor for c-Jun-activated transcription. *Embo J.* **21**(3): 451-460.

Wilkins, M. R., E. Gasteiger, A. A. Gooley, B. R. Herbert, M. P. Molloy, P. A. Binz, K. L. Ou, J. C. Sanchez, A. Bairoch, K. L. Williams and D. F. Hochstrasser. (1999). High-throughput mass spectrometric discovery of protein post-translational modifications. *J. Mol. Biol.* **289**(3): 645-657.

Wolffe, A. P. and D. Guschin. (2000). Chromatin structural features and targets that regulate transcription. *J. Struct. Biol.* **129**(2-3): 102-122.

Yang, X. J. (2004). The diverse superfamily of lysine acetyltransferases and their roles in leukemia and other diseases. *Nucleic Acids Res.* **32**(3): 959-976.

Yu, E., K. Song, H. Moon, G. G. Maul and I. Lee. (1998). Characteristic immunolocalization of Ku protein as nuclear matrix. **17**(5): 413-420.

Zhang, Q. H., H. Yao, N. Vo and R. H. Goodman. (2000). Acetylation of adenovirus E1A regulates binding of the transcriptional corepressor CtBP. *Proc. Natl. Acad. Sci. U. S. A.* **97**(26): 14323-14328.

Zhou, M. S., M. A. Halanski, M. F. Radonovich, F. Kashanchi, J. M. Peng, D. H. Price and J. N. Brady. (2000). Tat modifies the activity of CDK9 to

phosphorylate serine 5 of the RNA polymerase II carboxyl-terminal domain during human immunodeficiency virus type 1 transcription. *Mol. Cell. Biol.* **20**(14): 5077-5086.

Zhu, C. M., M. A. Bogue, D. S. Lim, P. Hasty and D. B. Roth. (1996). Ku86-deficient mice exhibit severe combined immunodeficiency and defective processing of V(D)J recombination intermediates. *Cell*. **86**(3): 379-389.

Zhu, P. C., D. Zhang, D. Chowdhury, D. Martinvalet, D. Keefe, L. F. Shi and J. Lieberman. (2006). Granzyme A, which causes single-stranded DNA damage, targets the double-strand break repair protein Ku70. *EMBO Rep.* **7**(4): 431-437.

Zubarev, R. A., D. M. Horn, E. K. Fridriksson, N. L. Kelleher, N. A. Kruger, M. A. Lewis, B. K. Carpenter and F. W. McLafferty. (2000). Electron capture dissociation for structural characterization of multiply charged protein cations. *Anal. Chem.* **72**(3): 563-573.