424 R&R and PHS-398 Specific
Table Of Contents

OMB Number: 4040-0001
Expiration Date: 06/30/2016

RESEARCH & RELATED Other Project Information

1. Are Human Subjects Involved?* ❍ **Yes** ● **No**

1.a. If YES to Human Subjects

Is the Project Exempt from Federal regulations? ❍ Yes ❍ No

If YES, check appropriate exemption number: _ 1 _ 2 _ 3 _ 4 _ 5 _ 6

If NO, is the IRB review Pending? ❍ **Yes** ❍ **No**

IRB Approval Date:

Human Subject Assurance Number

2. Are Vertebrate Animals Used?* ❍ **Yes** ● **No**

2.a. If YES to Vertebrate Animals

Is the IACUC review Pending? ❍ Yes ❍ No

IACUC Approval Date:

Animal Welfare Assurance Number

Funding Opportunity Number: PA-13-303. Received Date: 2014-06-12T16:27:49.000-04:00

Facilities and other resources

1) Tackett Laboratory: Dr. Tackett has 3 laboratories (1250 total sq feet), which are adjacent and on the fourth floor of Biomedical Research Building 1. In total, the laboratories contain 1 tissue culture hood, 2 fume hoods, 6 desk areas for laboratory personnel, and 8 workbench areas.

2) Tackett Computers: Each laboratory desk is equipped with a Dell computer for daily operations. For mass spectrometric studies, we have two user licenses for Rosetta Elucidator client software, powered by an Oracle 10G database, designated for laboratory workflow creation and data analysis. A dedicated eight-core server with 64 GB of memory and 380 GB of primary storage capacity will be used for primary mass spectrometry peak statistics. Raw proteomics data will be stored in a separate RAID (Redundant Array of Independent Disks) array storage scheme with a total capacity of 8 GB, with backups provided by our institutions enerprise storage system. Additional data analysis resources include: a Matrix Science 4-cpu license for Mascot protein identification software and Mascot Distiller, four quad-core PCs for data analysis running Scaffold (Proteome Software), Bioworks (Thermo) and MaxQuant. Tackett laboratory personnel have considerable experience applying these programs to analyze proteomic data.

3) Tackett office space: 150 sq feet of office space, which is located on the same floor as my laboratories.

4) Support: A machine shop is located in the same building as our laboratories. This shop has the equipment needed to maintain and modify Dr. Tackett's custom MALDI sample stages. UAMS maintains a electronic subscriptions to all necessary scientific journals. Fischer scientific and Invitrogen maintain freezers for immediate access to molecular reagents. The Department provides secretarial support for assistance with manuscript preparation and grant administration. Core facilities in Proteomics, DNA sequencing, cell sorting, fluorescence microscopy, and electron microscopy are available. X-ray crystallographic core facilities are available including robotic set up of crystallographic treys. Videoconferencing through the internet is made routine by using the Access Grid studio in any of six different locations including one in the nearby UAMS library.

5) Scientific environment: The scientific environment at the University of Arkansas for Medical Sciences is extremely strong. We have all the necessary laboratory space to perform the work and unlimited access to the mass spectrometers and software necessary. The UAMS Proteomics Facility (directed by Dr. Tackett) and Dr. Tackett's laboratory are well versed in proteomic studies.

Dr. Tackett has 10% effort as the director of the UAMS Proteomics Facility. As stated in the applicant's terms of hire, he has exclusive access to all instrumentation in the facility. Furthermore, the applicant's samples take priority over submitted samples for analysis. All mass spectrometric analyses for the applicant are performed by his research personnel, while facility samples are handled by a separate technician. The facility exists on the same floor as the applicant's laboratory and office.

Equipment

Mass spectrometry: The following mass spectrometers are available for this proposal: a Thermo Orbitrap Velos ETD, Thermo Orbitrap, Thermo LTQ XL ETD, a PerkinElmerSciex MALDI-prOTOF, a Waters Q-TOF Micro, and a vMALDI Thermo LTQ. Mass spectrometers are equipped with Waters UPLC systems or Eksigent nanoLC-2D systems. Additional gel band analysis equipment includes a ProPic imaging and spot-picking robot from Genomic Solutions, a ProGest in-gel enzymatic digestion robot from Genomic Solutions, an Eksigent Ekspot and a Bio-Rad 2D gel electrophoresis system. For mass spectrometric studies, we have two user licenses for Rosetta Elucidator client software, powered by an Oracle 10G database, designated for laboratory workflow creation and data analysis. A dedicated eight-core server with 64 GB of memory and 380 GB of primary storage capacity will be used for primary mass spectrometry peak statistics. Raw data will be stored in a separate RAID (Redundant Array of Independent Disks) array storage scheme with a total capacity of 8 GB, with backups provided by our institutions enerprise storage system. Additional data analysis resources include: a Matrix Science 4-cpu license for Mascot protein identification software and Mascot Distiller, four quad-core PCs for data analysis running Scaffold (Proteome Software), Bioworks (Thermo) and MaxQuant. Tackett laboratory personnel have considerable experience applying these programs to analyze proteomic data.

All mass spectrometers and proteomic equipment are part of the University of Arkansas for Medical Sciences Proteomics Facility. Dr. Tackett spends 10% of his time as the director of this facility. The remaining time is devoted to the applicant's independent research. As stated in the applicant's terms of hire, the applicant has exclusive access to all equipment in the facility, and the applicant's mass spectrometric samples take priority over submitted samples. The applicant's laboratory personnel analyze their own samples, while a separate technician handles facility samples.

OMB Number: 4040-0001
Expiration Date: 06/30/2016

RESEARCH & RELATED Senior/Key Person Profile (Expanded)

PROFILE - Project Director/Principal Investigator

Prefix: First Name*: ALAN Middle Name J Last Name*: TACKETT Suffix:

Position/Title*: Professor
Organization Name*: University of Arkansas for Medical Sciences
Department: Biochemistry and Molecular Bio
Division: College of Medicine
Street1*: 4301 W Markham
Street2: Slot 516
City*: Little Rock
County: Pulaski
State*: AR: Arkansas
Province:
Country*: USA: UNITED STATES
Zip / Postal Code*: 722057101

Phone Number*: (501) 686-8152 Fax Number: E-Mail*: AJTACKETT@UAMS.EDU

Credential, e.g., agency login: AJTACKETT

Project Role*: PD/PI Other Project Role Category:

Degree Type: PhD Degree Year: 2002

	File Name
Attach Biographical Sketch*:	Tackett_biosketch.pdf
Attach Current & Pending Support:	

Tracking Number: GRANT11674906
Funding Opportunity Number: PA-13-303 . Received Date: 2014-06-12T16:27:49.000-04:00

BIOGRAPHICAL SKETCH

Provide the following information for the key personnel and other significant contributors.
Follow this format for each person. **DO NOT EXCEED FOUR PAGES.**

NAME	POSITION TITLE
Tackett, Alan, Jackson	Professor of Biochemistry & Molecular Biology Professor of Pathology
eRA COMMONS USER NAME AJTACKETT	

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)*

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	YEAR(s)	FIELD OF STUDY
Hendrix College, Conway, AR	B.A.	1994-1998	Chemistry
University of Arkansas for Medical Sciences, Little Rock, AR	Ph.D.	1998-2002	Biochemistry and Molecular Biology
The Rockefeller University, New York, NY	Postdoctoral	2002-2005	Mass Spectrometry

A. Personal statement

My areas of expertise include proteomics, epigenetic technology development, and melanoma cancer biology. We currently pursue a series of proteomic studies including the most comprehensive to date proteomic analysis of FFPE patient melanoma. Additionally, I direct the UAMS Proteomics Facility and have extensive experience in collaborative research projects involving various areas of mass spectrometry. My group has expertise in using mass spectrometry to identify proteins in complex mixtures as well as quantifying protein abundance using various procedures (e.g., label-free, spectral counting, stable isotope tags & SILAC). We have a PhD level bioinformaticist in the lab that provides sophisticated data analysis. We also perform various quantitative analyses of posttranslational modifications.

For our technology development in epigenetics, we have been funded through the NIH Director's Office, the NIH Epigenomics Roadmap program, NCI and NIGMS to develop new proteomic tools for epigenesists. We have had a rich history of developing new tools for researchers including (1) iDIRT for identification of *in vivo* protein interactions, (2) transient iDIRT for identification of transient *in vivo* protein interactions, (3) MassSQUIRM for quantitative histone lysine methylation measurements, and (4) ChAP-MS which provides the first ever method to purify a single genomic locus for proteomic identification of bound proteins and histone posttranslational modifications. We also have ongoing biochemical work with a particular histone acetyltransferase, NuA3, in *S. cerevisiae*.

I am well-versed in every aspect of this proposal and will serve as PI of this grant. I have a long standing track record for developing and applying novel technological platforms in the field of epigenetics (e.g., iDIRT, MassSQUIRM, ChAP-MS). Also, I have 15 years of experience working genetically and biochemically with *S. cerevisiae* and have performed work on arsenic response in this model system (Tackett et al., 2005a; Tackett et al., 2005b; Taverna et al., 2006; Gradolatto et al., 2008; Smart et al., 2009; Gradolatto et al., 2009). Our cutting-edge ChAP-MS technology should help us and others in environmental epigenetics elucidate the epigenetic mechanisms behind environmental exposures.

B. Positions and Honors

2002-2005 Postdoctoral fellow, The Rockefeller University, Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, advisor Dr. Brian Chait (Camille and Henry Dreyfus Professor)
2002-2005 NIH F32 postdoctoral fellow (F32GM066496)
2005-2009 Assistant Professor, University of Arkansas for Medical Sciences, Dept. of Biochemistry and Molecular Biology
2007-present Director UAMS Proteomics Core Facility
2009-present Associate Professor with tenure, University of Arkansas for Medical Sciences, Dept. of Biochemistry and Molecular Biology
2012-present Member Winthrop P. Rockefeller Cancer Institute
2013-present Graduate Program Director UAMS Dept of Biochemistry and Molecular Biology
2014-present Full Professor with tenure UAMS Depts of Biochemistry & Molecular Biology and Pathology

Study Sections

1. NIH S10 Shared Instrumentation grant PAR-09-028, ad hoc reviewer, 11/10/09-11/11/09, SRO David Jollie, ZRG1 BCMB-D(30)
2. NIH/NCI Innovative Molecular Analysis Technologies (IMAT) program CA-09-004/005/006/007, ad hoc reviewer, 02/24/10-02/25/10, SRO Jeffrey DeClue
3. NIH/NIDA Exploring Epigenomic Processes and Non-coding RNAs in HIV/AIDS, DA-10-010 and DA-10-011, ad hoc reviewer, 03/04/2010, SRO Scott Chen
4. NIH/NCI, ad hoc reviewer, 12/13/10-12/15/10, Clinical Proteomic Technologies for Cancer Initiative (CPTC): Proteome Characterization Centers (PCCs) (U24 mechanism, RFA-CA-10-016), SRO Adriana Stoica
5. NIH/NCI Special Emphasis Panel on Emerging Technologies for Cancer Research CA-10-003/004/013, ad hoc reviewer, 03/29/11-03/30/11, SRO Slava Soldatenkov
6. NIH/NIMH Epigenomic Modifications in Neurodevelopment RFA-MH-11-030, ad hoc reviewer, 03/01/11-03/02/2011, SRO David Miller
7. NIH/Common Fund, NIH Director's Transformative Research Awards (R01) RFA-RM-11-006, ad hoc reviewer, 03/21/12, SRO David Jollie
8. NIH S10 Shared Instrumentation grant PAR-09-028, ad hoc reviewer, 9/6/12-9/7/12, SRO David Jollie
9. NIH Phase III COBRE, ad hoc reviewer, 10/24/12-10/25/12
10. NIH Phase II COBRE, ad hoc reviewer, SRO Bob Horowitz 3/14/2013
11. Chair, NIH/NIDDK special emphasis panel, 2013/05 ZDK1 GRB-R (M4), ad hoc reviewer, SRO Carol Robinson, 4/10/2013
12. NIH Phase III COBRE, ad hoc reviewer, 10/24/13-10/25/13, SRO Lisa Newman
13. NIH/NCI Special Emphasis Panel on Emerging Technologies for Cancer Research CA-10-004, ad hoc reviewer, 11/20/2013, SRO Jeffery DeClue
14. NIH/NCI Special Emphasis Panel on Emerging Technologies for Cancer Research CA-10-004, ad hoc reviewer, 02/14/2014, SRO Jeffery DeClue
15. NIH Phase II COBRE, ad hoc reviewer, 4/11/14, SRO Lisa Dunbar
16. NIH Phase I COBRE, ad hoc reviewer, 7/2/14, SRO Shinako Takada
17. NIH/NCI Special Emphasis Panel on Emerging Technologies for Cancer Research CA-10-004, ad hoc reviewer, 07/16/2014, SRO Jeffery DeClue
18. NIH Review of K99 Applications, ad hoc reviewer, 07/22/2014, SRO Robert Horowits
19. NIH Program Project: Center for Computational Mass Spectrometry, ad hoc, 06/27/2014, SRO Raymond Jacobson

C. Selected (15 of 55) peer-reviewed publications (in chronological order).

1. Tackett, A.J., Dilworth, D.J., Davey, M.J., O'Donnell, M.D., Aitchison, J.D., Rout, M.P. and Chait, B.T. (2005) Proteomic and genomic characterization of chromatin complexes at a boundary. *J. Cell Biol.*, 169, 35-47. PMCID: PMC2171912
2. Dou, Y., Milne, T.A., Tackett, A.J., Smith, E.R., Fukuda, A., Wysocka, J., Allis, C.D., Chait, B.T., Hess, J.L. and Roeder, R.G. (2005) Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. *Cell*, 121, 873-885. PMID: 15960975
3. Tackett, A.J., DeGrasse, J.A., Oeffinger, M., Rout, M.P. and Chait, B.T. (2005) I-DIRT, A General Method for Distinguishing Between Specific and Non-specific Protein Interactions. *J. Proteome Res.*, 4, 1752-6. PMID: 16212429
4. Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tackett, A.J., Chait, B.T., Badenhorst, P., Wu, C. and Allis, C.D. (2006) *Nature*, 442, 86-90. PMID: 16728976
5. Taverna, S.D., Ilin, S., Rogers, R.S., Tanny, J.C., Lavender, H., Li, H., Baker, L., Boyle, J., Blair, L.P., Chait, B.T., Patel, D.J., Aitchison, J.D., [Allis, C.D. and Tackett, A.J.] (2006) Yng1 PHD finger binding to histone H3 trimethylated at lysine 4 promotes NuA3 HAT activity at lysine 14 of H3 and transcription at a subset of targeted ORFs. *Mol. Cell*, 24, 785-796. PMID: 17157260
6. Taverna, S.D., Ueberheide, B.M., Liu, Y., Tackett, A.J., Diaz, R.L., Shabonowitz, J., Chait, B.T., Hunt, D.F. and Allis, C.D. (2007) Long-distance combinatorial linkage between methylation and acetylation on histone H3 N termini. *Proc Natl Acad Sci USA,* 104, 2086-91. PMCID: PMC1892956

7. Gradolatto, A., Rogers, R.S., Lavender, H., Taverna, S.D., Allis, C.D, Aitchison, J.D. and Tackett, A.J. (2008) *Saccharomyces cerevisiae* Yta7 regulates histone gene expression. *Genetics,* 179, 291-304. PMCID: PMC2390607
8. Upreti, M., Galitovskaya, E.N., Chu, R., Tackett, A.J., Granell, S. and Chambers, T.C. (2008) Direct biochemical identification of the major phosphorylation site in Bcl-xL induced by microtubule inhibitors and analysis of its functional significance. *J Biol Chem*, 283, 35517-35525. PMCID: PMC2602892
9. Blair, L.P., Raney, K.D. and Tackett, A.J. (2009) Development and evaluation of a structural model for SF1B helicases Dda. *Biochemistry*, 48, 2321-2329. PMCID: PMC2664552
10. Gradolatto, A., Smart, S.K., Byrum, S., Blair, L.P., Rogers, R.S., Kolar, E.A., Lavender, H., Larson, S.K., Aitchison, J.D., Taverna, S.D. and Tackett, A.J. (2009) A noncanonical bromodomain in the AAA ATPase protein Yta7 directs chromosomal positioning and barrier chromatin activity. *Mol Cell Biol*, 29, 4604-11. PMCID: PMC2725702
11. Smart, S.K., Mackintosh, S.G., Edmondson, R.D., Taverna, S.D. and Tackett, A.J. (2009) Mapping the local protein interactome of the NuA3 histone acetyltransferase. *Protein Science*, 18, 1987-1997. PMCID: PMC2777373
12. Byrum, S., Mackintosh, S.G., Edmondson, R.D., Cheung, W.L., Taverna, S.D. and Tackett, A.J. (2011) Analysis of Histone Exchange during Chromatin Purification. *JIOMICS*, 1, 61-65. PMC3119864
13. Byrum SD, Larson SK, Avaritt NL, Moreland LE, Mackintosh SG , Cheung W and Tackett AJ (2013) Quantitative Proteomics Identifies Activation of Hallmark Pathways of Cancer in Patient Melanoma. *J Proteomics Bioinform*, 6(3), 43-50. PMCID: PMC3748992
    *Commentary in *Genetic Engineering and Biotechnology News,* March 1, 2013
14. Byrum, S, Raman, A, Taverna, SD and Tackett, AJ (2012) ChAP-MS: A Method for Identification of Proteins and Histone Posttranslational Modifications at a Single Genomic Locus. *Cell Reports*, 2(1), 198-205. PMCID: PMC3408609
    *Commentary in *Nature*, 2012, 491, 143-147.
15. Byrum, S.D., Taverna, S.D. and Tackett, A.J. (2013) Purification of a specific native genomic locus for proteomic analysis. *Nucleic Acids Res*, 1-6, doi:10.1093/nar/gkt822 PMC3814360

D. Research Support

Ongoing Research Support

R01GM106024 Tackett (PI) 4/01/2013-3/31/2017
Using ChAP-MS to Study Macromolecular Chromatin Composition during Transcription
We have developed a new technology termed ChAP-MS that provides for the analysis of macromolecular protein interactions on chromatin at a single defined genomic position *in vivo*. Our planned extension of ChAP-MS to human cells and tissues will undoubtedly provide a major tool for epigeneticists to explore mammalian transcription regulation as well as the epigenetic disregulation associated with human diseases.
Role: PI

R33CA173264 Tackett (PI) 07/01/2013-06/30/2016
Development of MassSQUIRM to Quantitatively Measure Lysine Methylation
In this grant, we apply a proteomic technology we developed called MassSQUIRM to measure histone lysine methylation in cancer cells and patient tissues.
Role: PI

Completed Research Support (last 3 years)

R01 DA025755 Tackett (PI) 9/20/2008-09/19/2013
Development of technology for high resolution epigenetic profiling of chromatin
Role: PI

PHS 398 Modular Budget

OMB Number: 0925-0001

Budget Period: 1

Start Date: 04/01/2015 **End Date:** 03/31/2016

A. Direct Costs

	Funds Requested ($)
Direct Cost less Consortium F&A*	150,000.00
Consortium F&A	0.00
Total Direct Costs*	**150,000.00**

B. Indirect Costs

	Indirect Cost Type	Indirect Cost Rate (%)	Indirect Cost Base ($)	Funds Requested ($)
1.	49% -- On-Campus Organized Research (MTDC)	49.00	150,000.00	73,500.00
2.				
3.				
4.				

Cognizant Agency
(Agency Name, POC Name and Phone Number)
Indirect Cost Rate Agreement Dated: 8/24/2010 DHHS/ Division of Cost Allocation POC Name: Arif Karim POC Phone #: (214) 767-3261

Indirect Cost Rate Agreement Date 12/17/2013

Total Indirect Costs **73,500.00**

C. Total Direct and Indirect Costs (A + B)

Funds Requested ($) **223,500.00**

Funding Opportunity Number: PA-13-303. Received Date: 2014-06-12T16:27:49.000-04:00

Personnel Justification

1. Dr. Alan Tackett, principal investigator, 2.4 person months per 12 month calendar year. Dr. Tackett will supervise the project and review all data. He is well-versed in every aspect of this proposal. Dr. Tackett has a long standing track record for developing and applying novel technological platforms in the field of epigenetics (e.g., iDIRT, MassSQUIRM, ChAP-MS). He also has extensive experience working with *S. cerevisiae* and has performed work on arsenic response in this model system (Tackett et al., 2005a; Tackett et al., 2005b; Taverna et al., 2006; Gradolatto et al., 2008; Smart et al., 2009; Gradolatto et al., 2009).

2. Dr. Stephanie Byrum, Instructor, 12 person months per 12 month calendar year. Dr. Byrum developed the ChAP-MS approach and is therefore our lab expert on this technology (Byrum et al., 2011a; Byrum et al., 2011b; Byrum et al., 2012; Byrum et al., 2013). She performed her NIH-funded Ruth L. Kirschstein NRSA postdoctoral fellowship in Dr. Tackett's laboratory (2010-2012). She received her PhD in Bioinformatics and therefore will additionally provide all statistical support for our work. Dr. Tackett has personally trained her in yeast genetics, mass spectrometry and affinity isolation techniques. Dr. Byrum is well known for developing ChAP-MS and has presented this approach in three recent international meetings.

Identifying the Epigenetic Regulation of Arsenic Exposure

A. Specific Aims

***Overview**: Using a S. cerevisiae system, we have developed a technology called ChAP-MS that allows one to isolate a native 1 kb section of a chromosome for proteomic analysis (Byrum et al., 2012). This approach is truly innovative as it provides the first technology for comprehensive and unbiased measurement of protein and histone posttranslational modification content on a chromosome at high resolution. In this proposal, we outline how we will use ChAP-MS to comprehensively identify all combinatorial histone post-translational modifications and proteins regulating transcription at the arsenic response locus in S. cerevisiae. The results from this study will re-define how environmental epigenesists study epigenetic regulation as a response to exposure.*

One of the most compositionally diverse structures in a eukaryotic cell is a chromosome. A multitude of macromolecular protein interactions must properly occur on chromatin to drive functional aspects of chromosome biology like gene transcription, DNA replication, recombination, repair and sister chromatid segregation. Analyzing how proteins interact *in vivo* with chromatin to direct these activities remains a significant challenge due to the temporal and dynamic nature of their associations. To work towards overcoming these obstacles, our laboratory has developed a suite of novel tools to study protein-protein interactions of macromolecular complexes on chromatin (Tackett et al., 2005b; Taverna et al., 2006; Smart et al., 2009). Most relevant here, we have recently developed a technique termed Chromatin Affinity Purification with Mass Spectrometry or ChAP-MS (Byrum et al., 2012). ChAP-MS provides for the enrichment of a native 1 kb section of a chromosome for site-specific identification of protein interactions and associated histone posttranslational modifications (PTMs). ChAP-MS is the only available technology to unbiasedly identify proteins and histone PTMs on chromosomes at their native locus and in high resolution. Using this revolutionary approach, we were able to identify protein associations and combinatorial histone PTMs on single histone molecules in transcriptionally active and repressed states of chromatin at the *GAL1* locus in *S. cerevisiae*. In this proposal, we plan to use this cutting-edge approach to define the histone PTMs and proteins regulating transcription at the arsenic response locus in budding yeast. The *GAL1* studies we have published are very similar in principle to those we propose with arsenic. In the *GAL1* studies we identified changes in histone epigenetics and proteins bound to the locus when galactose was added to the media, and now we will use the same methodological platform to study changes when arsenic is added to the media. Accordingly, *we hypothesize that ChAP-MS will provide for a comprehensive and unbiased identification of all histone modifications and proteins regulating transcription at the arsenic locus in S. cerevisiae*. We will pursue the following Aim to test this hypothesis:

Specific Aim. Use ChAP-MS to define the histone post-translational modifications and proteins regulating transcription at the arsenic response locus in *S. cerevisiae*. We will use our recently described ChAP-MS approach to define the dynamics of histone epigenetics and protein associations regulating these epigenetic marks at the arsenic response locus in budding yeast (Byrum et al., 2012). The significance of the histone PTMs and proteins will be studied with *in vivo* assays. Furthermore, the epigenetic mechanism regulating the arsenic response locus and other genome-wide arsenic induced loci will be investigated.

Administrative Note: This grant application serves as a *transition* grant from Dr. Tackett's recently expired R01 from the NIH Roadmap Epigenomics Program under the non-renewable Technology Development RFA (R01DA025755, 14 publications, 3 patents, 4 year grant). We are utilizing ChAP-MS technology developed in yeast from the Roadmap Epigenomics grant (Byrum et al., 2012), and extending it to other loci to study mechanisms of environmental epigenetic regulation.

B. Significance

Epigenetic Regulation of Arsenic Exposure

Eukaryotic DNA is packaged with histone proteins to form nucleosomes, which in turn condense into higher order structures that constitute different functional forms of chromatin (Grewal and Moazed, 2003). Although epigenetic regulation is an evolving field, a wealth of data suggests that the governing processes are modulated, at least in part, through post-translational modifications (PTMs) on histones. The presence of histone PTMs is often dynamic in nature. Proteins such as histone acetyltransferases (HATs), histone methyltransferases and kinases can specifically modify histones (Kouzarides, 2007). Removal of these modifications occurs through the activity of protein assemblies like histone deacetylation complexes and histone demethylases, or via the active replacement of histones (Dion et al., 2007). Histone PTMs have been correlated with activities like the formation of euchromatin or heterochromatin, activation or silencing of gene transcription, and DNA damage repair (Berger, 2007; Kouzarides, 2007). Increasing evidence suggests histone PTMs are translated into a given cellular effect through recognition of the PTM by specific proteins termed “effector” proteins (Bannister and Kouzarides, 2004; Kouzarides, 2007; Strahl and Allis, 2000; Taverna et al., 2007a). Combined with their physiologically-relevant ties to biology, the placement of PTMs within genomes suggests a conserved “histone/epigenetic code” (Fischle et al., 2003; Grewal and Moazed, 2003).

For this proposal, we seek to comprehensively identify all the histone PTMs and proteins regulating these PTMs at the arsenic response locus in *S. cerevisiae*. Arsenic is a nonmutagenic carcinogen that promotes cancer via unknown mechanisms, thus understanding the epigenetic mechanisms responsible for responding to environmental exposures of arsenic will provide novel targets for maximizing treatment. Arsenic exists as either trivalent arsenite (As III, As_2O_3) or pentavalent arsenate (As V, As_2O_5), with arsenite being 60 times more toxic. Interestingly, arsenic has been hypothesized to elicit its toxic effects in part through inhibition of approximately 200 enzymes (via binding to cysteine thiol groups), some of which are key enzymes in DNA replication and DNA repair; thus, providing some evidence for a link to defects in mechanisms relevant to cancer initiation and progression. Arsenic toxicity is a global health problem affecting millions worldwide primarily through contaminated drinking water, consumption and inhalation (Ratinaike, 2003). For example, it is estimated that 35-77 million people (of the 125 million total) in Bangladesh have consumed water with toxic levels of arsenic (Smith et al., 2000). In the US, it was found that domestic well users accounted for 12% of the US population, but 23% of overall arsenic exposure from drinking water – which is predicted to increase annual fatalities from domestic well drinking water from 23% to 29% (Kumar et al., 2010). Arsenic exposure is a growing problem receiving increasing attention as can be seen in the September 2012 release from Consumer Reports showing high levels of arsenic in the rice we are consuming – with rice eaten just once a day driving arsenic levels up 44% in the body and consumption of rice twice a day increasing arsenic levels by 70%.

A review of the scientific literature for what is understood about the epigenetics of arsenic response provides a minimal understanding of the mechanisms regulating gene expression upon arsenic exposure. The bulk of studies have focused on how arsenic exposure affects DNA methylation - a mark of gene silencing and only one aspect of epigenetic regulation. A recent review on this topic indicates that the field does observe fluctuations in DNA methylation upon arsenic exposure (Reichard and Puga, 2010). The studies on DNA methylation have found both hyper- and hypomethylation at specific loci; thus, there does appear to be a targeted epigenetic mechanism. It is known that DNA methylation recruits protein complexes such as histone deacetylases, which then deacetylate histone lysine residues and promote a silent state of chromatin. Thus, the studies on DNA methylation have actually only begun to uncover the complexity of epigenetic regulation at the DNA *and* histone level. A literature search for what is known about histone modifications and arsenic uncovered only two studies reporting that cells treated in cell culture or peripheral blood mononuclear cells from individuals exposed to arsenic had *bulk* changes in certain histone methylations, but nothing was reported for *specific* genomic loci (Zhou et al., 2008; Chervona et al., 2012). Furthermore, nothing is known about combinatorial histone PTMs. Different combinations of histone PTMs on individual histone molecules code for transcriptional activation or silencing (Taverna et al., 2007b). In our proposed work, we will use a technology we recently developed called ChAP-MS to comprehensively identify all proteins bound, all histone PTMs and all combinatorial histone PTMs on single histone molecules at the arsenic response locus in budding yeast.

Arsenic Response Locus in *S. cerevisiae*

One of the beauties of arsenic detoxification systems is that all organisms have them and they appear to work in similar ways (Rosen, 2002). Arsenic detoxification systems all have core components that (1) uptake extracellular arsenate via phosphate transporters on the cell membrane or uptake extracellular arsenite via aquaglyceroporins, (2) use an arsenate reductase to convert intracellular arsenate to arsenite and (3) use

membrane transporters to move intracellular arsenite to outside the cell or sequester arsenite in a vacuole. The core components of the arsenic response locus in *S. cerevisiae* are illustrated in Figure 1 (Rosen, 2002). The membrane transporter of extracellular arsenate is Pho87 (and additionally Pho86, Pho84), while Fps1 is the extracellular arsenite transporter (Bun-ya et al., 1996; Wysocki et al., 2001). Once arsenate enters the cell, Arr2 (aka Acr2) arsenate reductase converts arsenate to arsenite (Mukhopadhyay et al., 2000). Arsenite can then be transported out of the cell through Arr3 (aka Acr3) or can be sequestered in the vacuole via Ycf1 (Wysocki et al., 1997; Maciaszczyk-Dziubinska et al., 2010). Two of the core components of arsenic detoxification, Arr2 and Arr3, are coded for genomically at the extreme end of chromosome 16 adjacent to the telomere (Fig. 1) (Bobrowicz et al., 1997). The *ARR2* and *ARR3* genes are controlled in part by a transcription factor called Arr1 (aka Acr1 or Yap8) that is also coded for in this arsenic response locus. Arr1 is constitutively expressed and always bound to the divergent promoter region for the *ARR2* and *ARR3* genes, and it has been proposed that the interaction of intracellular arsenic with cysteine sulfhydryls on Arr1 activates this protein to promote transcription of *ARR2* and *ARR3* (Wysocki et al., 2004). This arsenic response locus containing *ARR1*/*ARR2*/*ARR3* will be the focus of this proposal as these are key response proteins. The budding yeast arsenic detoxification system is remarkably similar to the human system (Rosen, 2002); thus, justifying our studies in yeast, which has defined genetics and phenotypic assays – providing for rapid functional studies.

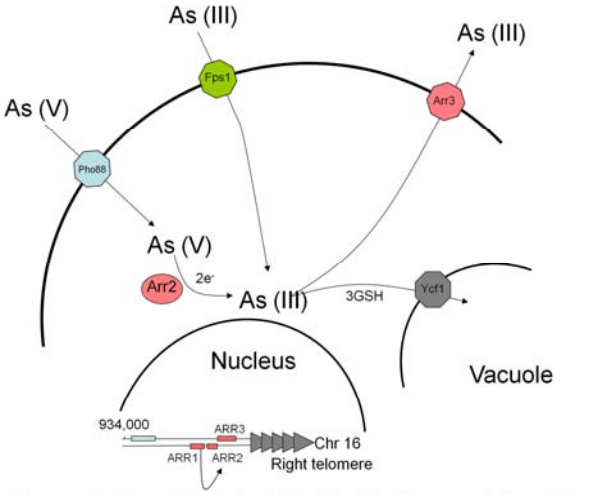

Figure 1. Arsenic detoxification in *S. cerevisiae*. The arsenic response genes are clustered towards the end of chromosome 16. Arr1 is a transcription factor that regulates the arsenic response locus. Arr2 is a reductase that converts arsenate to arsenite. Arr3 is a arsenite membrane transporter.

Available Technologies to Study Proteins / PTMs at a Specific Locus: Limitations and Obstacles

In order for one to begin to elucidate epigenetic mechanisms, an appropriate methodological approach has to be in place to define the proteins bound to the arsenic response locus and the histones PTMs – both in the presence and absence of arsenic. In general, technical challenges have precluded the ability to determine positioning of chromatin factors like proteins and histone PTMs along a chromosome. Chromatin immunoprecipitation (ChIP) assays have been used to better understand genome-wide distribution of known proteins and histone PTMs at the nucleosome level (Dedon et al., 1991; Ren et al., 2000; Pokholok et al., 2005; Robertson et al., 2007; Johnson et al., 2007; Barski et al., 2007; Mikkelsen et al., 2007). However, major drawbacks of current ChIP-based methods are their confinement to examining singular histone PTMs or proteins rather than simultaneous profiling of multiple targets, the inability to determine the co-occupancy of particular histone PTMs, and that ChIP is reliant on the previous identification of the molecular target. For example, how are we to know which protein and which histone PTM to look for at the arsenic response locus? And how do we identify combinatorial histone modifications on single histone molecules at the arsenic response locus? The best solution would be the biochemical isolation of a specific genomic locus for unbiased proteomic identification of proteins and histone PTMs. Approaches similar to this have been performed for structures like telomeres or engineered plasmids; however, nobody has accomplished the proteomic analysis of a native genomic region (Griesenbeck et al., 2003; Dejardin and Kingston, 2009; Unnikrishnan et al., 2010).

C. Innovation

ChAP-MS Technology

The ability to purify and characterize a single native locus for proteomic analysis has long been a goal for researchers in chromatin biology. We recently published a technology, called Chromatin Affinity Purification with Mass Spectrometry (ChAP-MS) that provides for the first time a method to identify all proteins and histone PTMs at a single, native genomic locus (Fig. 2) (Byrum et al., 2012). ChAP-MS provides for site-specific enrichment of a given ~1kb section of chromatin followed by identification of proteins and histone PTMs using high resolution mass spectrometry. Using ChAP-MS, we were able to purify the chromatin at the *S. cerevisiae GAL1* locus in transcriptionally silent and active states. We identified protein interactions and combinatorial histone PTMs unique to the *GAL1* gene in each of these functional states. This is the exact same methodological scenario of treating cells with or without arsenic and defining the proteins and histone PTMs at the arsenic response locus.

Purification and characterization of *GAL1* chromatin by ChAP-MS.

Figure 2A provides an overview of the ChAP-MS approach that was used to identify proteins and histone PTMs associated with the *GAL* locus. A LexA DNA binding site was engineered by homologous recombination upstream of the *GAL1* start codon in a *S. cerevisiae* strain constitutively expressing a LexA-Protein A (LexA-PrA) fusion protein from a plasmid. The LexA DNA binding site directs the localization of the LexA-PrA protein affinity "handle" to the *GAL1* chromatin. The positioning of the LexA DNA was designed so we could specifically enrich for chromatin-associated proteins and histone PTMs regulating gene expression near the 5'-end of *GAL1*. This strain was cultured in glucose to repress transcription, or galactose to activate transcription. Following *in vivo* chemical cross-linking to preserve the native chromatin, chromatin was sheared by sonication to ~1,000 base-pair sections. We have published studies showing that 1.25% formaldehyde cross-linking is optimal for preserving the *in vivo* state of the chromatin, while still providing for solubility during purification (Byrum et al., 2011a). The PrA moiety of the LexA-PrA was then used to affinity purify (on IgG-coated Dynabeads) the ~1,000 base-pair section of chromatin at the 5'-end of *GAL1* for mass spectrometric identification of proteins and histone PTMs.

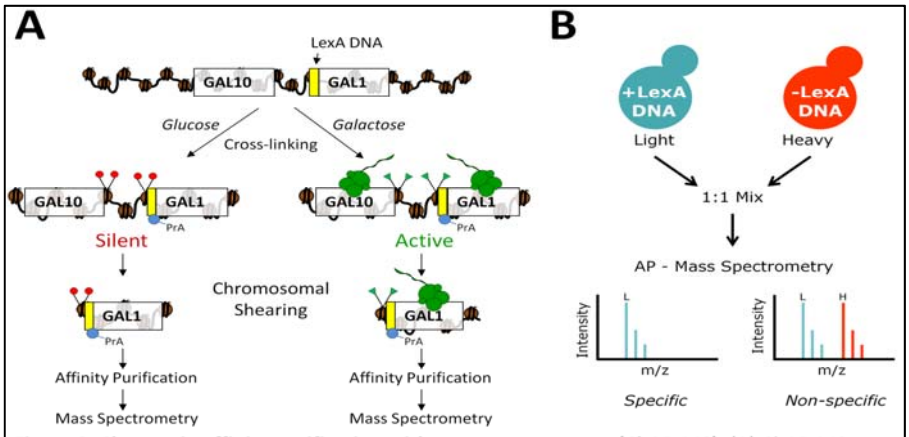

Figure 2. Chromatin affinity purification with mass spectrometry (ChAP-MS). (A) ChAP-MS provides for the specific enrichment of a small chromosome section and identification of specifically associated proteins and posttranslational modifications. (B) Isotope mixing control to identify non-specifically enriching proteins and PTMs. See text for details and Byrum et al 2012.

Due to the low abundance of the targeted chromatin region in cellular lysates, we fully anticipated that proteins non-specifically associating with *GAL1* chromatin would complicate our analysis of the resulting purified material. Co-purification of non-specifically associating proteins is one of the major complications of affinity purifications; however, isotopic labeling of media provides a means to gauge *in vivo* protein-protein interactions and quantitate differences in protein enrichment (Smart et al., 2009; Tackett et al., 2005b). We previously developed a variation of this labeling technique called iDIRT (isotopic differentiation of interactions as random or targeted) that provides a solution for determining which co-enriched proteins are specifically or non-specifically associated with a complex of proteins (Smart et al., 2009; Tackett et al., 2005b). The iDIRT technique was adapted to control for proteins non-specifically enriching with LexA-PrA and the resin (Fig. 2B). By using this adaptation of iDIRT separately on chromatin enriched from active and repressed chromatin states, the proteins non-specifically enriching with the isolated *GAL1* chromatin section were identified. A strain containing the LexA DNA binding site and LexA-PrA fusion protein was cultured in isotopically light media, while a strain lacking the LexA DNA binding site (but still containing the LexA-PrA fusion protein) was cultured in isotopically heavy media ($^{13}C_6{}^{15}N_2$-lysine). The light and heavy strains were mixed and co-lysed. The growth and mixing of light/heavy strains was performed separately under glucose and galactose conditions. The purification of the *GAL1* chromatin was performed from this mixture of light/heavy lysates. Proteins and histone PTMs specifically associated with the *GAL1* chromatin containing the LexA DNA binding site were isotopically light as they arose from the cells grown in light media. Proteins that were non-specific to the purification were a 1:1 mix of light and heavy as they were derived equally from the light and heavy lysates. Analysis of peptides from the enriched proteins with high resolution mass spectrometry was used to determine the level of isotopically light and heavy proteins – thereby

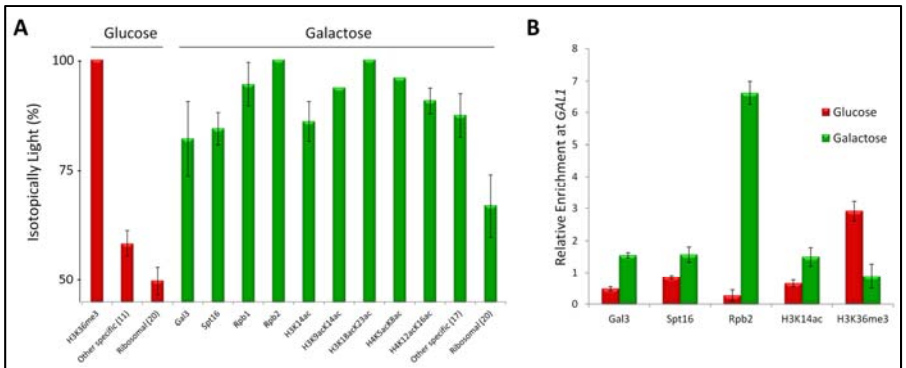

Figure 3. Identification and validation of proteins and histone posttranslational modifications enriched with *GAL1* chromatin. (A) Proteins and histone posttranslational modifications identified from the ChAP-MS analysis of *GAL1* chromatin in the transcriptionally active galactose and repressive glucose growth conditions are listed in accordance to their percent isotopically light. (B) qPCR-ChIP was used to validate proteins and histone posttranslational modifications on *GAL1* chromatin. (For further details see Byrum et al., 2012).

determining whether the detected protein was a specific *in vivo* constituent of *GAL1* chromatin or a non-specific contaminant. ChAP-MS analyses of *GAL1* chromatin revealed association of Gal3, Spt16, Rpb1, Rpb2, H3K14ac, H3K9acK14ac, H3K18acK23ac, H4K5acK8ac and H4K12acK16ac under transcriptionally active conditions; while repressive conditions showed H3K36me3 (Fig. 3A & B). The identification of the double acetylations is unique to the ChAP-MS approach as antibodies do not exist. The ChAP-MS approach identified the presence of RNA polymerase (Rpb1, Rpb2) and Spt16 (aids RNA pol) under active conditions.

D. Approach

Specific Aim. Use ChAP-MS to define the histone post-translational modifications and proteins regulating transcription at the arsenic response locus in *S. cerevisiae*.

Rationale. Using a grant from the NIH Roadmap Epigenomics Program under the Technology Development in Epigenetics initiative RFA (R01DA025755), we have developed a technology called ChAP-MS that provides for the site specific enrichment of a given 1kb section of a chromosome, unambiguous identification of the proteins associated with this chromosome section, and identification of the histone PTMs associated specifically with this enriched chromatin section (Fig. 2) (Byrum et al., 2012). In this Aim, we will utilize the ChAP-MS technology to perform a comprehensive analysis of the dynamics of histone PTMs and protein complexes recruited to the arsenic response locus chromatin in presence and absence of arsenic.

Experimental protocol, data analysis, and interpretation. ChAP-MS will be technically performed as detailed in Byrum et al., 2012. The specific approach for the arsenic response locus is detailed in Figure 4. We will create four different strains that contain separate engineered LexA DNA binding sites that tile the arsenic response locus at 1 kb resolution. LexA DNA binding sites will be engineered by targeted homologous recombination as described (Byrum et al., 2012). For ChAP-MS we will grow and lyse the following for both control and +arsenic growths: (1) strain containing one engineered LexA DNA binding site transformed with *pLexA-PrA* (plasmid for expression of LexA-PrA) in isotopically light media and (2) wild-type cells without an engineered LexA DNA binding site with *pLexA-PrA* in isotopically heavy $^{13}C_6$-Arg $^{13}C_6$-Lys media. We will separately test arsenite and arsenate treatment. We will explore varying levels and durations of arsenic treatment (chronic vs acute exposures), but will initially start with 2 mM arsenite and 10 mM arsenate treatment for 2 hours which we have previously used in the lab to ascertain yeast defects. Cells will be cross-linked with 1.25% formaldehyde to preserve chromatin during purification (Byrum et al., 2012; Byrum et al., 2011a; Byrum et al., 2011b). Prior to cryogenic lysis with a Retch mixer mill, we will mix equivalent amounts of isotopically light and isotopically heavy cells. The mixtures of these two strains will be from control growths or separately from growths with arsenic. Lysate from 10^11 mixed cells will be thawed into a routine purification buffer and chromatin will be sheared to ~1,000 bp sections with sonication (Byrum et al., 2012). Following shearing of chromatin, the PrA moiety on LexA will be enriched using IgG-coated Dynabeads as we have done extensively before (Byrum et al., 2011b; Smart et al., 2009; Tackett et al., 2005b). Purification of each of the four engineered LexA DNA binding sites is shown in Figure 4. Isolated proteins will be resolved by SDS-PAGE (heating in Laemmli loading buffer reverses protein cross-links) and visualized by Coomassie-staining. The entire gel lane will be excised as 2 mm sections and subjected to trypsin digestion of proteins. Peptides will be analyzed by high resolution MS and MS^2 with a Thermo Orbitrap Velos mass spectrometer. Proteins will be identified by database searching with Mascot and peptide intensities will be calculated with Mascot Distiller. We will search for histone PTMs such as acetylation, phosphorylation and methylation. The percentage of isotopically light peptide relative to heavy will be calculated from peak intensities, and proteins will be categorized as stable interactors or contaminants as detailed in Fig 2B. We will obtain the following data from this study for each of the 4 positions in transcriptionally active (+arsenic) and silent (control, –arsenic) states: (1) local proteome, (2) single and combinatorial histone PTMs, and (3) non-histone protein PTMs.

Validation and functional analysis of proteins and PTMs at ARR2/ARR3. We will use qPCR-ChIP to validate the proteins and PTMs we find localized to arsenic response locus chromatin. The protein and histone PTM ChIPs will be done in strains grown with or without arsenic for comparison and technically performed as in Fig 3B. For transcriptional studies, we will monitor transcription of *ARR2* and *ARR3* by real time rtPCR in strains (± arsenic) containing genomic deletions of the proteins found to associate at *ARR2/ARR3* chromatin or in strains with point mutations of histone amino acids containing PTMs identified at *ARR2/ARR3* chromatin. Protein deletion strains will be taken from our yeast ORF deletion set. Mutations for PTM-containing histone residues will be done with a histone “shuffle strain” (Ahn et al., 2005). We will combine deletions/mutations. We will also perform epistasis assays on plates ± arsenic with combinations of protein/histone PTM deletions (and *arr1Δ*, *arr2Δ* and *arr3Δ*). We will explore non-histone protein PTMs in similar ways. We will also use the transcription and epistasis approaches to further define the mechanism of transcriptional activation at the

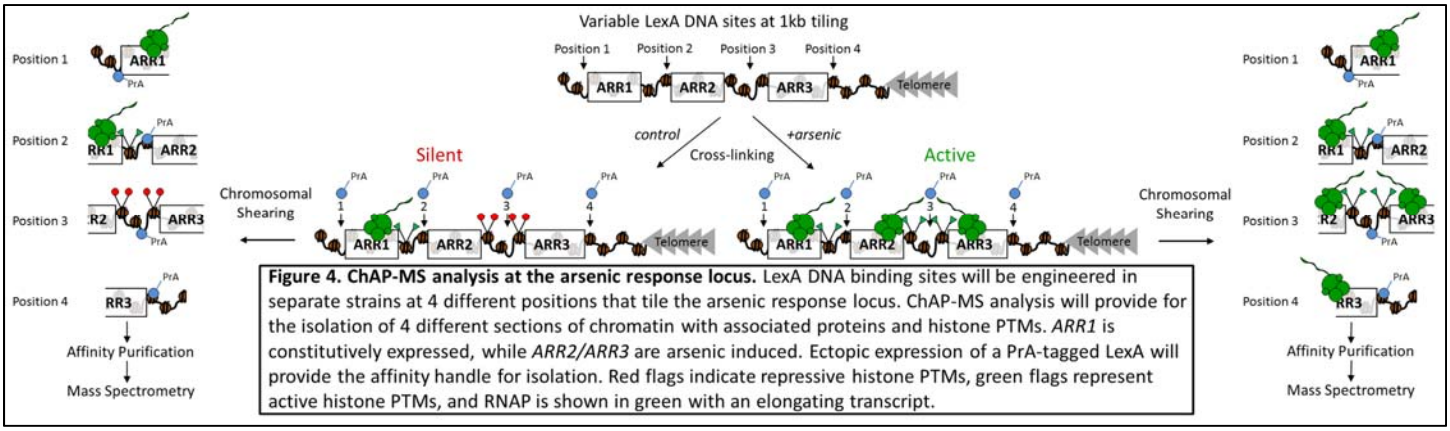

Figure 4. ChAP-MS analysis at the arsenic response locus. LexA DNA binding sites will be engineered in separate strains at 4 different positions that tile the arsenic response locus. ChAP-MS analysis will provide for the isolation of 4 different sections of chromatin with associated proteins and histone PTMs. *ARR1* is constitutively expressed, while *ARR2/ARR3* are arsenic induced. Ectopic expression of a PrA-tagged LexA will provide the affinity handle for isolation. Red flags indicate repressive histone PTMs, green flags represent active histone PTMs, and RNAP is shown in green with an elongating transcript.

arsenic locus. For a given histone PTM, we will delete the protein known to add or remove this mark and determine its relevance to response to arsenic exposure. For a given protein that we identify that has a known role in adding or removing a histone PTM, we will delete the amino acid modified on the histone and ascertain the importance for arsenic response. If we identify proteins with known histone PTM binding domains, we will delete these domains or mutate the amino acid they bind to and ascertain arsenic response.

A caveat to the transcriptional and epistasis studies is whether the observed effect at *ARR2/ARR3* is direct or indirect. We will pursue genome-wide transcription studies to determine whether the effects we observe are direct or indirect. For proteins/histone PTMs that are identified by ChAP-MS at *ARR2/ARR3* and demonstrate a functional phenotype in transcription/epistasis studies detailed in the paragraph above, we will perform genome-wide transcription microarray studies comparing wild-type (±As) and strains containing a deletion of the protein/histone PTM (±As). In Preliminary Studies, we have performed a triplicate transcription microarray study comparing wild-type yeast cells and wild-type cells treated with 2 mM arsenic. From these Preliminary Studies, we uncovered 91 genes that showed >3-fold ($p<0.05$) increase in transcription upon arsenic treatment. *ARR2* and *ARR3* were both represented in the 91 genes. In addition to *ARR2* and *ARR3*, other genes with a putative role in arsenic detoxification were identified - including *GRE1* & *GRE2* (induced by heavy metal exposure (Garay-Arroyo and Covarrubias, 1999)), and *GTT2* (glutathione transferase functioning in arsenic detoxification (Todorova et al., 2007)). This preliminary microarray study provides the panel of genes we expect to see transcriptionally up-regulated in the presence of arsenic; thus, providing a comparison for the deletion studies. If the deleted protein/histone PTM has a *direct* mechanism of action at *ARR2/ARR3*, we anticipate observing a specific lack of transcriptional activation of *ARR2/ARR3* in the presence of arsenic for the deletion strains. Additionally, we may observe defects in transcriptional activation of the other 89 genes identified in our preliminary microarray study if there is a conserved mechanism of epigenetic regulation between *ARR2/ARR3* and the other 89 arsenic induced genes. If the mechanism is *indirect* (i.e., the deletion has altered global transcription), then we would expect to observe a lack of transcriptional activation of *ARR2/ARR3* and many other non-arsenic induced genes (i.e., genes other than the 91 determined from our preliminary study). Results from epistasis studies (i.e., deletions of proteins/histone PTMs in combination with *arr1Δ*, *arr2Δ* or *arr3Δ*) detailed above will provide support for direct vs indirect mechanisms.

If we determine deletion of the protein/histone PTM causes global transcription defects or defects not specific to *ARR2/ARR3*, then we will perform targeted *in vivo* studies. For identified histone PTMs, we will use LexA-protein fusions of particular histone modifying enzymes (e.g., deacetylase, demethylase or phosphatase with specific activity toward the histone PTM) to localize the enzyme to *ARR2/ARR3* containing an engineered LexA DNA binding site, confirm localization of the LexA-enzyme fusion protein and removal of the given histone PTM by qPCR-ChIP, and measure transcriptional activation of *ARR2/ARR3* for wild-type and the genomically targeted strain (±As). Localizing a histone PTM modifying enzyme as a fusion protein has been shown to localize activity to a specific region of chromatin and remove (or add) specific histone PTMs in the region (Weerasinghe et al., 2010; Chiu et al., 2003; Taverna et al., 2002; Deng et al., 2008; Marcus et al., 1994; Candau et al., 1997; Kadosh and Struhl, 1997). These studies will effectively serve to erase the specific histone PTM from *ARR2/ARR3* rather than removing it globally. For identified proteins that bind a particular sequence of genomic DNA, we will genomically mutate the *in vivo* DNA binding site for the protein at *ARR2/ARR3* by homologous recombination, ensure the protein does not localize to the genomically mutated DNA by qPCR-ChIP, and measure transcriptional activation of *ARR2/ARR3* for wild-type and the genomic mutant (±As). These studies will effectively serve as deletion of the protein from only this specific locus. For proteins identified that do not bind to a specific DNA sequence, we will use LexA-fusions to localize them to *ARR2/ARR3* in the *absence* of arsenic and determine if they stimulate transcription by real time rtPCR.

Defining the epigenetic mechanism of regulation at the arsenic response locus and other genome-wide arsenic induced loci. We fully anticipate that an epigenetic program is activated upon arsenic exposure that alters the transcription of many genes including *ARR2* and *ARR3*. We hypothesize that the epigenetic mechanism of transcriptional regulation at *ARR2/ARR3* is conserved at other genes activated by arsenic exposure. In support of this hypothesis, the above-mentioned transcription microarray study identified 91 genes transcriptionally induced by arsenic. To test whether the same proteins and histone PTMs at *ARR2/ARR3* are localized to other arsenic induced genes, we will perform genome-wide ChIPseq (±As) to the proteins and histone PTMs **(1)** identified at *ARR2/ARR3* in the presence of arsenic with our ChAP-MS studies and **(2)** shown to have a functional role in the transcriptional activation of *ARR2/ARR3* (studies described in the above section). We believe that nucleosome resolution ChIPseq will be more efficient than targeted qPCR-ChIP since we do not know the precise region of the other arsenic-induced genes to target for PCR. Sequencing will be performed in the UAMS Pharmacogenomics core facility with an Illumina Cluster Station and Genome Analyzer IIx. We have performed numerous ChIPseq experiments targeting both proteins and histone PTMs using this platform. Protein and histone PTM sites of localization or 'peaks' will be identified using the model-based analysis for ChIPseq (MACS) computational tool to identify peaks with a false discovery rate <1% (Bardet et al., 2012). To compare samples, the ChIPseq read counts will be normalized by the total number of read counts and quantitative changes analyzed by calculating the $\log_2$ fold change. Significance testing for enrichment will be assessed using the student's t-test. Genome-wide profiles of protein and histone PTM locations from ChIPseq studies will be correlated to our preliminary transcription microarray studies from wild-type cells that showed 91 genes induced by arsenic. Correlations will be done by calculating Pearson correlation coefficients for genes found to be up-regulated 3-fold ($p<0.05$) by arsenic from triplicate transcription array data with the ChIPseq data showing proteins/histone PTMs significantly enriched upon arsenic exposure at $p<0.05$ by student's t-test and fold change > 2. These comparisons will tell us whether the protein/histone PTM initially identified and functionally characterized at *ARR2/ARR3* is also localized to other arsenic induced genes. Proteins and histone PTMs found located at genes transcriptionally activated by arsenic will be confirmed with standard qPCR-ChIP. Also, we will compare ChIPseq results with the transcription microarray studies (±As) described in the former section for strains containing deletions of proteins/histone PTMs identified at *ARR2/ARR3*, which will tell us whether the protein/histone PTM is required for direct or indirect transcriptional regulation of the given arsenic induced gene. As detailed above for indirect regulation, we will perform **(1)** epistasis studies and **(2)** targeted transcriptional studies at these other arsenic induced loci by localizing histone modifying enzymes to erase histone PTMs or abolishing the locus specific protein association and then measuring transcription (±As).

If we observe **(1)** localization to other arsenic induced genes of proteins and histone PTMs identified at *ARR2/ARR3* and **(2)** similar phenotypes in functional transcription studies, then we will conclude there are similar epigenetic mechanisms regulating arsenic induced genes. If we see similar epigenetic mechanisms regulating transcription at multiple arsenic induced genes, then we will have support for our hypothesis of a conserved mechanism of transcriptional activation; however, if we do not see a conserved mechanism at multiple arsenic induced loci, then we will perform ChAP-MS studies at other genes with roles in arsenic detoxification identified in our preliminary microarray study like *GRE1*/*GRE2*/*GTT2* to determine the different proteins/histone PTMs regulating the epigenetic state. Our approaches for defining the epigenetic regulation of genes **(1)** transcriptionally up-regulated by As exposure and **(2)** containing proteins and histone PTMs regulating As-induced transcription of *ARR2*/*ARR3* is superior to performing genome-wide studies on histone PTMs found to be differentially present in *bulk* histones because **(1)** locus specific alternations in histone PTMs will not necessarily be reflected in *bulk* levels and **(2)** protein targets cannot be identified this way.

Alternative approaches. In addition to the ChAP-MS approach using LexA-PrA binding to an engineered DNA site, we have recently developed another approach called TAL-ChAP-MS (Transcription Activator Like Chromatin Affinity Purification with Mass Spectrometry), which alleviates genomic engineering (Byrum et al., 2013). Key to this approach is the use of TALs (Transcription Activator Like effector proteins) to target DNA. TALs contain a central domain of 18 tandem repeats of 34 amino acids each, which direct sequence specific DNA binding (Doyle et al., 2012; Cermak et al., 2011). Binding to a given nucleobase in DNA is determined by two adjacent amino acids (12 and 13) within each of the 18 repeats (Scholze & Boch, 2010). Thus, by mutating these amino acids in each of the 18 tandem repeats, one can "program" binding to a given 18 nt region of DNA *in vivo*. The ectopically-expressed TALs used here would target the arsenic response locus. These TALs would be expressed as a fusion with PrA. Relative to LexA-PrA, the TAL-PrA approach has minor off target enrichment issues; thus, for the work in this proposal we will start with the LexA-PrA approach, but will have the ability to switch to the TAL-PrA system as we further develop it independently from this work.

References

Ahn, S. H., Cheung, W. L., Hsu, J. Y., Diaz, R. L., Smith, M. M., and Allis, C. D. (2005). Sterile 20 kinase phosphorylates histone H2B at serine 10 during hydrogen peroxide-induced apoptosis in S. cerevisiae. Cell *120*, 25-36.

Allis, C. D., Jenuwein, T., Reinberg, D., and Caparros, M. L. (2006). Epigenetics: Cold Spring Harbor Laboratory Press).

Akiyoshi, B., Nelson, C.R., Ranish, J.A. and biggins, S. (2009) Quantitative proteomic analysis of purified yeast kinetochores identified a PP1 regulatory subunit. Genes Dev *23*, 2887-2899.

Avaritt, N.L., Owens, R., Reynolds, M., Munkberg, J.M., Byrum, S., Smoller, B.R., Tackett, A.J. and Cheung, W.L. (2012) Misregulation of Rad50 expression in melanoma cells. *J Cutan Pathol*, 39, 680-4. PMC Journal – in progress

Bannister, A. J., and Kouzarides, T. (2004). Histone methylation: recognizing the methyl mark. Methods Enzymol *376*, 269-288.

Bardet AF, He, Q, Zeitlinger, J and Stark, A (2012) A computational pipeline for comparative ChIP-seq analyses. Nature Protocols 7, 45-61.

Barski, A.,Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I. and Zhao, K. (2007) High-resolution profiling of histone methylations in the human genome. Cell *129*, 823-37.

Bhattacharjee, H, Sheng, J, Ajees, AA, Mukhopadhyay, R and Rosen, BP (2010) Adventitious arsenate reductase activity of the catalytic domain of the human Cdc25B and Cdc25C phosphatases. Biochemistry 49, 802-9.

Berger, S. L. (2007). The complex language of chromatin regulation during transcription. Nature *447*, 407-412.

Bobrowicz, P, Wysocki, R, Owsianik, G, Goffeau, A, Ulaszewski, S (1997) Isolation of three contiguous genes, ACR1, ACR2 and ACR3, involved in resistance to arsenic compounds in the yeast Saccharomyces cerevisiae. Yeast 13, 819-28.

Bock, I., Dhayalan, A., Kudithipudi, S., Brandt, O., Rathert, P. and Jeltsch, A. (2011) Detailed specificity analysis of antibodies binding to modified histone tails with peptide arrays. Epigenetics *6*, 256-263.

Bun-ya, M, Shikata, K, Nakade, S, Yompakdee, C, Harashima, S and Oshima, Y (1996) Two new genes, PHO86 and PHO87, involved in inoraganic phosphate uptake in Saccharomyces cerevisiae. Curr Genet 29, 344-51

Byrum, S., Raman, A., Taverna, SD and Tackett, AJ (2012) ChAP-MS: A Method for Identification of Proteins and Histone Posttranslational Modifications at a Single Genomic Locus. *Cell Reports*, 2(1), 198-205. PMCID: PMC3408609.
http://download.cell.com/cell-reports/pdf/PIIS2211124712001933.pdf?intermediate=true

Byrum, S.D., Taverna, S.D. and Tackett, A.J. (2011a) Quantitative analysis of histone exchange for transcriptionally active chromatin. J Clin Bioinforma, *1*, 17. PMC3164610

Byrum, S., Mackintosh, S.G., Edmondson, R.D., Cheung, W.L., Taverna, S.D. and Tackett, A.J. (2011b). Analysis of Histone Exchange during Chromatin Purification. JIOMICS *1*, 61-65. PMC Journal – in progress

Byrum, S., Avaritt, N.L., Mackintosh, S.G., Munkberg, J.M., Badgwell, B.D., Cheung, W.L. and Tackett, A.J. (2011c) A Quantitative Proteomic Analysis of FFPE Melanoma. ***J Cutan Pathol***, 38, 933-6. PMC Journal – in progress

Byrum, S.D., Taverna, S.D. and Tackett, A.J. (2013) Purification of a specific native genomic locus for proteomic analysis. *Nucleic Acids Res*, 1-6, doi:10.1093/nar/gkt822

Candau R, Zhou JX, Allis CD, Berger SL 1997. Histone acetyltransferase activity and interaction with ADA2 are critical for GCN5 function in vivo. EMBO J 16: 555–565

Cermak, T., E. L. Doyle, et al. (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Research *39*(12): e82.

Chervona, Y, Hall, MN, Arita, A et al (2012) Association between arsenic exposure and global post-translational histone modifications amound adults in Bangladesh. Cancer Epidemiology, Biomarkers & Prevention, published online Oct 12, 2012

Chiu, YH, Yu, Q, Sandmeier, JJ and Bi, X (2003) A targeted histone acetyltransferase can create a sizeable region of hyperacetylated chromatin and counteract the propagation of transcriptionally silent chromatin. Genetics 165, 115-25.

Cohen, S. L., and Chait, B. T. (1997). Mass spectrometry of whole proteins eluted from sodium dodecyl sulfate-polyacrylamide gel electrophoresis gels. Anal Biochem *247*, 257-267.

Dedon, P.C., Soults, J.A., Allis, C.D. and Gorovsky, M.A. (1991) Formaldehyde cross-linking and immunoprecipitation demonstrate developmental changes in H1 association with transcriptionally active genes. Mol Cell Biol *11, 1729-33.*

Dejardin, J., and Kingston, R. E. (2009). Purification of proteins associated with specific genomic Loci. Cell *136*, 175-186.

Deng, H et al (2008) Ectopic histone H3S10 phosphorylation causes chromatin structure remodeling in Drosophilia. Development 135, 699-705.

Dilworth, D. J., Tackett, A. J., Rogers, R. S., Yi, E. C., Christmas, R. H., Smith, J. J., Siegel, A. F., Chait, B. T., Wozniak, R. W., and Aitchison, J. D. (2005). The mobile nucleoporin Nup2p and chromatin-bound Prp20p function in endogenous NPC-mediated transcriptional control. J Cell Biol *171*, 955-965. PMC2171315

Dion, M. F., Kaplan, T., Kim, M., Buratowski, S., Friedman, N., and Rando, O. J. (2007). Dynamics of replication-independent histone turnover in budding yeast. Science *315*, 1405-1408.

Doyle, E. L., N. J. Booher, et al. (2012) TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. Nucleic Acids Research *40*(1): 117-122.

Freeman, LEB, Dennis, LK, Lynch, CF, Thorne, PS, Just, CL (2004) Toenail arsenic content and cutaneous melanoma in Iowa. Am J Epidem 160, 679-87.

Garay-Arroyo, A and Covarrubias, AA (1999) Three genes whose expression is induced by stress in Saccharomyces cerevisiae. Yeast 15, 879-92.

Geibler et al (2011) Transcriptional activators of human genes with programmable DNA-specificity. PLoS ONE *6*, e19509.

Gradolatto, A., Rogers, R.S., Lavender, H., Taverna, S.D., Allis, C.D., Aitchison, J.D. and Tackett, A.J. (2008). Saccharomyces cerevisiae Yta7 regulates histone gene expression. Genetics 179, 291-304. PMC2390607

Gradolatto, A., Smart, S. K., Byrum, S., Blair, L. P., Rogers, R. S., Kolar, E. A., Lavender, H., Larson, S. K., Aitchison, J. D., Taverna, S. D., and Tackett, A. J. (2009). A noncanonical bromodomain in the AAA ATPase protein Yta7 directs chromosomal positioning and barrier chromatin activity. Mol Cell Biol *29*, 4604-4611. PMC2725702

Grewal, S. I., and Moazed, D. (2003). Heterochromatin and epigenetic control of gene expression. Science *301*, 798-802.

Griesenbeck, J., Boeger, H., Strattan, J.S. and Kornberg, R.D. (2003) Affinity Purification of Specific Chromatin Segments from Chromosomal Loci in Yeast. Mol Cell Biol *23*, 9275-9282.

Houseley, J., Rubbi, L., Grunstein, M., Tollervey, D., and Vogelauer, M. (2008). A ncRNA modulates histone modification and mRNA induction in the yeast GAL gene cluster. Mol Cell *32*, 685-695.

Hsueh YM, Cheng GS, Wu MM, et al. Multiple risk factors associated with arsenic-induced skin cancer: effects of chronic liver disease and malnutritional status. Br J Cancer 1995;71:109–14

Jenuwein, T. and Allis, C.D. (2001) Translating the histone code. Science 293, 1074-1080

Johnson, D.S., Mortazavi, A., Myers, R.M. and Wold, B. (2007) Genome-wide mapping of in vivo protein-DNA interactions. Science *316* 1497-502.

Kadosh D, Struhl K 1997. Repression by Ume6 involves recruitment of a complex containing Sin3 corepressor and Rpd3 histone deacetylase to target promoters. Cell 89: 365–371

Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693-705.

Kozul-Horvath, CD, Zandbergen, F, Jackson, BP, Enelow, RI, Hamilton, JW (2012) Effects of low-dose drinking water arsenic on mouse fetal and postnatal growth and development. PLoS ONE 7(5): e38249.

Kumar, A, Adak, P, Gurian, PL and Lockwood, JR (2010) Arsenic exposure in US public and domestic drinking water supplies: A comparative risk assessment. J Exp Sci and Environ Epidem 20, 245-254.

Maciaszczyk-Dziubinska, E, Migdal, I, Migocka, M, Bocer, T and Wysocki, R (2010) The yeast aquaglyceroporin Fps1p is a bidirectional arsenite channel. FEBS Lett 584, 726-32.

Marcus GA, Silverman N, Berger SL, Horiuchi J, Guarente L 1994. Functional similarity and physical association between GCN5 and ADA2: Putative transcriptional adaptors. EMBO J 13: 4807–4815.

Meneghini, MD, Wu, M and Madhani HD (2003) Conserved histone variant H2A.Z protects euchromatin from the ectopic spread of silent heterochromatin. Cell 112, 725-736

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., et al. (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553-60.

Miller et al (2011) A TALE nuclease architecture for efficient genome editing. Nat Biotech *29*, 143-148.

Mukhopadhyay, R, Shi, J and Rosen, BP (2000) Purification and characterization of ACR2p, the Saccharomyces cerevisiae arsenate reductase. J Biol Chem 275, 21149-57.

Pearce, D, Dowling, K and Sim, M (2011) Soil Arsenic and melanoma: is there a link? 4th International Conference on medical Geology. GeoMed2011 page 108.

Pesch B, Ranft U, Jakubis P, et al. Environmental arsenic exposure from a coal-burning power plant as a potential risk factor for nonmelanoma skin carcinoma: results from a case-control study in the district of Prievidza, Slovakia. Am J Epidemiol 2002;155:798–809

Phanstiel, D., Brumbaugh, J., Berggren, W. T., Conard, K., Feng, X., Levenstein, M. E., McAlister, G. C., Thomson, J. A., and Coon, J. J. (2008). Mass spectrometry identifies and quantifies 74 unique histone H4 isoforms in differentiating human embryonic stem cells. Proc Natl Acad Sci U S A *105*, 4093-4098.

Phelps, J (2005) Headliners: Skin Cancer: Linking Toenail Arsenic Content to Cutanewous Melanoma. Environ Health Perspect, 113, A377.

Pokholok, D. K., Harbison, C. T., Levine, S., Cole, M., Hannett, N. M., Lee, T. I., Bell, G. W., Walker, K., Rolfe, P. A., Herbolsheimer, E.*, et al.* (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. Cell *122*, 517-527.

Ratnaike, RN (2003) Acute and chronic arsenic toxicity. Postgrad Med J, 79, 391-6.

Reichard, JF and Puga, A (2010) Effects of arsenic exposure on DNA methylation and epigenetic gene regulation. Epigenomics 2, 87-104.

Ren, B., Robert, F., Wyrick, J.J., Aparicio, O., Jennings, E.G., et al (2000) Genome-wide location and function of DNA binding proteins. Science *290*, 2306-9.

Robertson, G., Hirst, M, Bainbridge, M., Bilenky, M., Zhao, Y. et al (2007) Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. Nat Methods *4*, 651-7.

Rosen, BP (2002) Biochemistry of arsenic detoxification. FEBS Letters 529, 86-92.

Rossman TG (2003) Mechanism of arsenic carcinogenesis: an integrated approach. Mutation Research 533, 37-65.

Scholze, H. and J. Boch (2010) TAL effector-DNA specificity. Virulence *1*(5): 428-432.

Scholze, H. and J. Boch (2011) TAL effectors are remote controls for gene activation. Current Opinion in Microbiology *14*(1): 47-53.

Shukla, A., Bajwa, P. and Bhaumik, S.R. (2006) SAGA-associated Sgf73p facilitates formation of the preinitiation complex assembly at the promoters either in a HAT-dependent or independent manner in vivo. Nucleic Acids Res *34*, 6225-32.

Smart, S. K., Mackintosh, S. G., Edmondson, R. D., Taverna, S. D., and Tackett, A. J. (2009). Mapping the local protein interactome of the NuA3 histone acetyltransferase. Protein Sci *18*, 1987-1997. PMC2777373

Smith, AH, Lingas, EO and Rahman, M (2000) Contamination of drinking-water by arsenic in Bangladesh: a public health emergency. Bulletin of the World Health Organization 78 (9).

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

Syka, J. E., Coon, J. J., Schroeder, M. J., Shabanowitz, J., Hunt, D. F. (2004). Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. Proc Natl Acad Sci U S A *101*, 9528-33.

Tackett, A.J., Dilworth, D.J., Davey, M.J., O'Donnell, M.D., Aitchison, J.D., Rout, M.P. and Chait, B.T. (2005a) Proteomic and genomic characterization of chromatin complexes at a boundary. *J. Cell Biol.*, 169, 35-47. PMCID: PMC2171912

Tackett, A. J., DeGrasse, J. A., Sekedat, M. D., Oeffinger, M., Rout, M. P., and Chait, B. T. (2005b). I-DIRT, a general method for distinguishing between specific and nonspecific protein interactions. J Proteome Res *4*, 1752-1756. PMC Journal – in progress

Tackett, A. J., Corey, D. R., and Raney, K. D. (2002) Non-Watson-Crick Interactions between PNA and DNA Inhibit the ATPase Activity of Bacteriophage T4 Dda Helicase. Nucleic Acids Research *30*, 950-957. PMC Journal – in progress

Tackett, AJ, Wei, L, Cameron CE and Raney KD (2001a) Unwinding of nucleic acids by HCV NS3 helicase is sensitive to the structure of the duplex. Nucleic Acids Res *29*, 565-72. PMC Journal – in progress

Tackett, A. J., Morris, P. D., Dennis, R., Goodwin, T. E., and Raney, K. D. (2001b) Unwinding of Unnatural Substrates by a DNA Helicase. Biochemistry *40*, 543-548. PMC Journal – in progress

Taverna, SD, Coyne, RS and Allis, CD (2002) Methylation of histone h3 at lysine 9 targets programmed DNA elimination in tetrahymena. Cell 110, 701-11.

Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D., and Patel, D. J. (2007a). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nat Struct Mol Biol *14*, 1025-1040.

Taverna, S. D., Ueberheide, B. M., Liu, Y., Tackett, A. J., Diaz, R. L., Shabanowitz, J., Chait, B. T., Hunt, D. F., and Allis, C. D. (2007b). Long-distance combinatorial linkage between methylation and acetylation on histone H3 N termini. Proc Natl Acad Sci U S A *104*, 2086-2091. PMC1892956

Taverna, S. D., Ilin, S., Rogers, R. S., Tanny, J. C., Lavender, H., Li, H., Baker, L., Boyle, J., Blair, L. P., Chait, B. T., Patel, D. J., Aitchison, J. D., Tackett, A. J. and Allis C. D. (2006). Yng1 PHD finger binding to histone H3 trimethylated at lysine 4 promotes NuA3 HAT activity at lysine 14 of H3 and transcription at a subset of targeted ORFs. Mol Cell *24*, 785-796. PMC Journal – in progress

Thompson, A., Schafer, J., Kuhn, K, Kienle, S, Schwarz, J, Schmidt, G, Neumann, T, Johnstone, R, Mohammed, AK and Hamon, C (2003) Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. Anal Chem *75*, 1895-904.

Todorova, T, Vuilleumier, S and Kujumdzieva, A (2007) Role of glutathione S-transferases and glutathione in arsenic and peroxide resistance in Saccharomyces cerevisiae: a reverse genetic analysis approach. Biotechnol. & Biotechnol. EQ. 21/2007/3, 348-352.

Unnikrishnan, A., Gafken, P. R., and Tsukiyama, T. (2010). Dynamic changes in histone acetylation regulate origins of DNA replication. Nat Struct Mol Biol *17*, 430-437.

Weerasinghe, SVW, Wambua, M and Pflum, MKH (2010) A histone deacetylase-dependent screen in yeast. Bioorg Med Chem 18, 7586-92.

Wysocki, R, Bobrowicz, P and Ulaszewski, S (1997) The Saccharomyces cerevisiae ACR3 gene encodes a putative membrane protein involved in arsenite transport. J Biol Chem 272, 30061-6.

Wysocki, R, Chery, CC, Wawrzycka, D, Van Hulle, M, Cornelis, R, Thevelein, JM, Tamas, MJ (2001) The glycerol channel Fps1p mediates the uptake of arsenite and antimonite in Saccharomyces cerevisiae. Mol Microbiol 40, 1391-401.

Wysocki, R, Fortier, PK, Maciaszczyk, E, Thorsen, M, Leduc, A, Odhagen, A, Owsianik, G, Ulaszewski, S, Ramator, D and Tamas, MJ (2004) Transcriptional activation of metalloid tolerance genes in Saccharomyces cerevisiae requires the AP-1-like proteins Yap1p and Yap8p. Mol Biol Cell 15, 2049-60.

Zhang, et al (2008) Epitope tagging of endogenous proteins for genome-wide ChIP-chip studies. Nat Methods, 5, 163-5.

Zhou, X, Sun, H, P.Ellen, T, Chen, H and Costa, M (2008) Arsenite alters global histone H3 methylation. Carcinogenesis 29, 1831-36.