Title

Quantitative Proteomic Analysis for Assessing the Mechanisms of Action of Anti-Cancer Drugs and Arsenite

Author

Zhang, Fan

Publication Date

2013

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Quantitative Proteomic Analysis for Assessing the Mechanisms of Action of Anti-Cancer Drugs and Arsenite

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Chemistry

by

Fan Zhang

December 2013

Dissertation Committee:
Dr. Yinsheng Wang, Chairperson
Dr. Cynthia Larive
Dr. Pingyun Feng

The Dissertation of Fan Zhang is approved:

Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENTS

This dissertation can never be completed without the help and support from many people. I owe my appreciation to all those who have made this dissertation possible and because of whom my graduate experience has been one that I will cherish forever.

First and foremost, I would like to give my deepest gratitude to my research adviser, Professor Yinsheng Wang, for his valuable guidance and consistent encouragement on all my research projects during my PhD study at UCR. Professor Wang encouraged me not only to grow as a solid chemist and experimentalist, but also as an independent thinker. His diligent and serious research attitude has impressive impact on the maturity of my personality. His patience and support helped me to overcome the obstacles and desperate situations during these past five years. I would say that I could never finish the PhD study without Professor Wang's great mentorship. For everything you have done for me, Professor Wang, thank you.

I would like to give my appreciation to my committees: Professor Cynthia Larive and Professor Pingyun Feng, for reading my dissertation and providing me their helpful comments; Professor Quan (Jason) Cheng, for teaching me the knowledge in his class; and Professor Jeff Bachant, for some wise advice he gave me on my research.

I would give my sincere thanks to some staff members in Chemistry Department and Center of Plant Cell Biology and Core Instrumentation Facility at UCR. I want to thank Dr. David Carter, for his kind assistance and advice on confocal microscopy; and

Ms. Holly Eckelhoefer for her training and help with flow cytometry. I also want to thank Dr. Dan Borchardt for the training on circular dichroism, Dr. Songqin Pan and Mr. Ronald New for their assistance on MALDI mass spectrometers.

I would express my thanks to all my friends in the Wang group. Thanks go to Dr. Haibo Qiu, Dr. Lei Xiong, and Dr. Xiaoli Dong for their generous help and patience when I started work in this group. I would give my gratitude to Dr. Bifeng Yuan who is always helpful and nice; my experimental techniques improved a lot thanks to all the assistance he offered. Thanks also go to Dr. Xiaoxia Dai, Qian Cai, Lijuan Fu, Pengcheng Wang and Yongsheng Xiao for their help in completing my research projects; and Dr. John Prins for his discussion on my research. Special thanks go to Candace Guerrero for her precious friendship, and the other members in the Wang group for their support. I would always appreciate and memorize the time in this laboratory.

Finally, and the most importantly, I would like to deeply thank my parents whose support, encouragement, patience and unwavering love are always the most valuable gifts at every stage of my personal and academic life. No matter what happened, they always stand behind me and offer their faith in me. In addition, I give special thanks to my boyfriend Dr. Qichi Yang whose understanding and accompany mean a lot to me. He is always the first person to encourage me in every possible way during the past four years. I would cherish the great time we had together and your continuous love and support.

COPYRIGHT ACKNOWLEDGEMENTS

The text and figures in Chapter 2, in part or in full, is a reprint of the material as it appears in ***Mol. Cell. Proteomics 2012***, 11, M111 016915. The co-author (Dr. Yinsheng Wang) in that publication directed and supervised the research which forms the basis for this chapter. The co-author (Xiaoxia Dai) listed in that publication performed siRNA knockdown and real-time PCR experiments.

ABSTRACT OF THE DISSERTATION

Quantitative Proteomic Analysis for Assessing the Mechanisms of Action of Anti-Cancer Drugs and Arsenite

by

Fan Zhang

Doctor of Philosophy, Graduate Program in Chemistry
University of California, Riverside, December 2013
Dr. Yinsheng Wang, Chairperson

Mass spectrometry is a powerful tool used in proteomic analysis. By combining with different isotope labeling methods, protein quantification at the entire proteome scale can be achieved. In this dissertation, I employed an LC-MS/MS coupled with stable isotope labeling by amino acid in cell culture (SILAC) strategy, to explore the molecular mechanisms of action of anti-cancer drugs and an environmental toxicant, arsenic.

In Chapters 2 and 3, I performed a quantitative assessment of global proteome changes upon treatment with two anti-cancer drugs, both are nucleoside analogues, i.e. 5-aza-2'-deoxytidine (5-Aza-CdR) and 6-thioguanine (^{S}G). We found, for the first time, that 5-Aza-CdR exerts its cytotoxic effects in leukemia and melanoma cells through epigenetic reactivation of *DPP4* gene and the resultant inhibition of cholesterol biosynthesis. In addition, proteome-wide analysis suggested that ^{S}G may exert its cytotoxic effect by inducing mitochondrial dysfunction and reactive oxygen species formation in acute lymphoblastic leukemia cells.

In Chapters 4 and 5, we utilized the same MS-based quantitative proteomic method, to study the mechanisms underlying trivalent arsenic-induced carcinogenic effect. Our results confirmed the previous findings and provided a more complete picture about the biological pathways that are altered upon arsenite treatment in human skin fibroblast cells. To further understand how arsenic affects RING-finger proteins that are known to be important in DNA damage repair, we performed *in vitro* and *in vivo* arsenite binding experiments. Our studies revealed, for the first time, that arsenite may exert its carcinogenic effect by targeting cysteine residues in the RING finger domains of the RNF20-RNF40 histone E3 ubiquitin ligase, thereby altering histone epigenetic mark and compromising DNA DSB repair.

TABLE OF CONTENTS

LIST OF FIGURES

Figure B1.. 159

Example mass spectra showing the incorporation of heavy labeled K and R. Shown are the ESI-MS of two trypsin-miscleaved peptides which carry two arginines or two lysines; IGEEEIQKPEEK and DILLRPELEELR are derived from splicing factor 3a (SAP114) and NADH-cytochrome b5 reductase (CYB5R3), respectively. The ions of m/z 751.4 and m/z 718.9 refer to the putative incompletely labeled peptides (i.e., with a single R or K being heavily labeled), which are present at similar abundance levels as the noise peaks.

Figure B2.. 160

Flow cytometry results for control CEM cells, and CEM cells that were treated for 24 hrs with 3 μM ^{S}G and 500 μM allopurinol, alone or in combination. In (a), no MitoTracker Deep Red was added, whereas MitoTracker Deep Red was added in (b).

Figure B3.. 161

Western blot for monitoring the expression levels of ectopically expressed Myc-Miro1 in HEK293T cells upon treatment with 3 μM ^{S}G (a). β-actin serves as the loading control. Shown in (b) are the quantification data, which represent the mean and S. D. of results from three independent drug treatment and Western blot experiments. "**", $p < 0.01$. The *p*-value was calculated using two-tailed, unpaired student t test.

Figure B4.. 162

Selected-ion chromatograms for monitoring the transitions of m/z 250→164 [(a), for unlabeled R-cdA] and m/z 255 → 169 [(b), for uniformly ^{15}N-labeled R-cdA].

RNF20 and RNF40 and detected with anti-RNF20 or anti-RNF40. β-actin was used as loading control in (c), (d) and (e).

Fluorescence microscopy results for cells transfected with GFP-RNF20 (top) or GFP-RNF40 (bottom), but without ReAsH treatment; no red fluorescence could be detected in these cells.

(a) Selected-ion chromatogram for monitoring the $[M +2H]^{2+}$ ion (m/z 549.8) of a tryptic peptide derived from histone H2B, AVTKGGVTSSK, with K120 being modified with a diglycine remnant. (b) Selected-ion chromatogram for monitoring the $[M +2H]^{2+}$ ion (m/z 408.7) of the reference tryptic peptide derived from histone H2B, EIQTAVR. Shown in the insets are the MS displaying the $[M+2H]^{2+}$ ions of the two peptides. (c) Product-ion spectrum of the $[M + 2H]^{2+}$ ion of H2B reference peptide EIQTAVR.

Quantitative real-time PCR results showing that the mRNA expression levels of five genes involved in DNA DSB repair are not reduced in HEK293T cells upon a 24-hr treatment with 5 μM $NaAsO_2$. The mRNA expression levels of these genes were normalized to that of *GAPDH*. The *p*-values were calculated by using unpaired two-tailed t-test. “N.S” indicates that the difference was not significant ($p > 0.05$)

Flow cytometry results showing the percentage of GFP+ cells in (a) U2OS-DRGFP and (b) U2OS-EJ5GFP cells transfected with I-SceI alone (I-Sce I) or together with a 24-hr treatment of 5 μM $NaAsO_2$ (I-Sce I + $NaAsO_2$), as well as those cells that were not transfected with I-Sce I (Control), or treated only with 5 μM $NaAsO_2$ for 24 hr (Control + $NaAsO_2$). (c) Quantitative results showed that the treatment with 5 μM $NaAsO_2$ does not affect the transfection efficiency in U2OS cells. "N.S" indicates that the difference was not significant ($p > 0.05$).

LIST OF TABLES

CHAPTER 1

General Overview

1.1 Introduction

Understanding the functions of living cells at the molecular level has become one of the most challenging endeavors in modern biological science. Although a complete human genome sequence has been obtained (1), the cellular function at the protein level in the post-genomic era is still demanded because the majority of the functional information of genes resides in the proteome (2). Proteomics, considered as a technology of "discovery science", is defined as the systematic study of various properties of proteins including the identities, quantities, structures, biochemical and cellular functions, and how they vary in space, time and physiological state (3).

Many technologies have been and are still being used in proteomics study including microarray and ChIP-Seq experiments. Microarray is a technology based on unbiased sample screening and data accumulation (4). With a small amount of material, microarray-based systems facilitate a large number of reactions to occur within a very limited area. Along this line, global gene expression analysis with microarray was used for assessing transcriptional responses of cancer cells towards anti-cancer agents (5). However, due to the regulatory processes following mRNA production including posttranscriptional, translational and protein degradation regulation which are crucial in controlling protein abundances, results from mRNA-based gene expression analysis cannot fully reflect protein expression data (6). In fact, global differentiation of protein abundance can also be measured by protein microarrays which mimic the chip layout of DNA microarray (7). However, the difficulties in obtaining a set of antibodies against the whole proteome prevented the application of this approach.

To fully understand transcriptional regulation, chromatin immunoprecipitation coupled with high-throughput sequencing (ChIP-Seq) has been developed recently for genome-wide mapping of DNA-protein interactions and histone modifications (8). ChIP-Seq is emerged as an alternative technique which produces profiles with higher resolution, dynamic range and genomic coverage than traditional microarray (8). Relatively small amount of sample is needed and the technique is not constrained by the species accessibility (9). However, the main drawbacks for ChIP-Seq are the cost and antibody availability which limited the its application (8).

Traditional two-dimensional gel electrophoresis (2-DE) is one of the technologies used in the proteomics study. 2-DE provides information about protein molecular weight, isoelectric point (*pI*), quantity and post-translational modifications (PTMs) by recording the different intensities of the stained spot derived from two populations of cells or tissues (10). However, due to the poor separation and low resolution for low-abundance proteins (11), this technique showed bias and limitations towards hydrophobic and membrane proteins.

Other than 2-DE, mass spectrometry (MS) has developed rapidly over the past 2-3 decades owing to its high sensitivity, wide dynamic range and good resolution to monitor the proteins of low abundance (12). It has been widely used in the study of protein-protein interactions (13), protein PTM (14), protein profiling (15) and intact protein characterization (16), etc. MS-based proteomic approach along with stable isotope labeling strategies has been widely used for large-scale quantitative analysis of proteins in complex samples (12). This technology, especially quantitative proteomics, has a fundamental importance in understanding the mechanisms of action of drugs, discovering for biomarkers, and assessing the toxicity of environmental toxicants.

In this chapter, I will review MS-based quantitative proteomic technique in combination with stable isotope labeling in the discovery of the mechanisms of action of anti-cancer drugs and arsenite.

1.2 MS-based quantitative proteomics

Quantitative proteomic analysis is highly desired given the reason that most changes induced in the biological system are only captured if some quantitative information is acquired (17). Prior to the development of high-throughput quantitative proteomic analysis, the study of cellular changes largely relied on the antibody-based immunostaining (i.e. Western blot and ELISA), which lacks efficiency and consistency and is limited by antibody availability (18, 19). Although the development of 2-DE facilities the protein quantification on a relatively large-scale, the low resolution, high variation in complex biological samples and the bias against low-abundance and membrane proteins have limited the application of this technique (11).

MS has established itself as a powerful tool and challenges the traditional assays for its ability to accurately identify and quantify a large-number of proteins in complex biological sample such as biofluids (20). MS-based quantitative analysis can be classified into two subgroups, the absolute quantification which determines the absolute amount of protein in the sample (21) and relative quantification which defines the amount in comparison to another experimental condition or biological sample (e.g. the protein expression level change upon drug treatment) (15). Absolute quantification of proteins is achieved by comparing the MS intensities between the endogenous peptides with a manually added, synthetic, isotope-labeled peptide with known concentration (22). It is a straightforward method which measures the exact amount of the sample with low detection limits, high dynamic range and low necessity of sample clean-up (23). With an appropriate selection of reference peptide on a triple-quadrupole mass spectrometer operating in selective-reaction monitoring (SRM) mode and an appropriate separation technique such as HPLC, this technique has been widely used for biomarker quantification and protein PTM studies (24, 25).

For the large-scale proteomics study, absolute quantification of the levels of all the identified proteins is, however, usually not feasible.

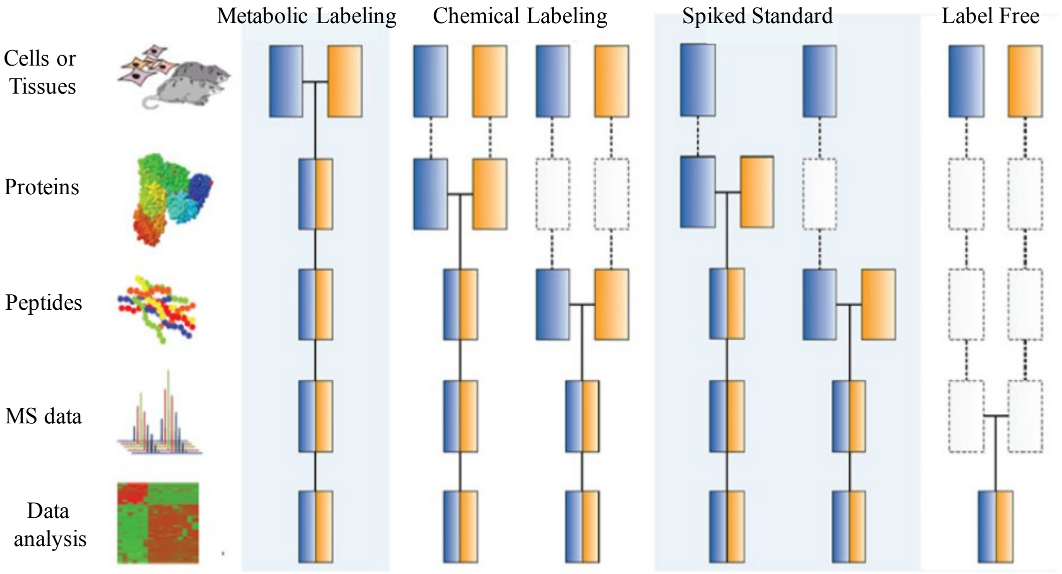


Figure 1.1 MS-based workflows for quantitative proteomics (Adapted from (17, 26)).

To this end, relative quantification, which compares the relative expression levels of cellular proteins by using stable isotope labeling strategies, is frequently employed. In this context, two populations of cells were labeled with light and heavy isotopes, respectively. By analyzing the mixture of light and heavy isotope-containing peptides, the difference in mass introduced by the weight of isotopic labeling can be easily distinguished by MS. The ratios of the signal intensities for the labeled and unlabeled peptides would reflect their relative quantities (27). Relative quantification can be achieved through (1) chemical labeling, (2) metabolic labeling, and (3) label-free quantifications (26) (Figure 1.1).

1.2.1 Chemical labeling strategies

Chemical labeling approaches label either intact proteins or proteolytic peptides with a stable isotope-coded chemical reagent *in vitro* (28). The labeled proteins or peptides introduced mass shifts are shown in MS or MS/MS spectra, and relative quantification is achieved by comparing the signal intensities of the peptide pairs.

1.2.1.1 Isotope-coded affinity tag (ICAT)

Developed by Aebersold and co-workers (29), ICAT was the first commercially available chemical labeling reagent. ICAT reagent consists of a cysteine-reactive iodoacetyl group which has specificity toward sulfhydryl group, a polyether linker where the heavy and light forms contain eight deuterons and eight hydrogens, respectively, and a biotin affinity tag as shown in Figure 1.2a. The standard ICAT workflow is displayed in Figure 1.2b. After labeling with light or heavy ICAT reagents, cell lysates from two cell states containing either heavy or light isotopic labeling are combined and fractionated. Proteolyzed peptide mixtures are then analyzed on a mass spectrometer. Pairs of ICAT-labeled peptides co-elute with 8 Da mass differences that can be observed in mass spectra. Relative quantification was determined by comparing the peak intensities between the light and heavy peptides. ICAT allows isolating the labeled peptide through biotin-avdin affinity chromatography under physiological pH, which significantly decreased the sample complexity by removing cysteine-free peptides.

Traditional ICAT labeling also has limitations. Because of the specificity of reagent reactivity toward thiol functionality, this method is limited to cysteine-containing proteins. In addition, the avidin affinity step suffers from considerable non-specific binding of Cys-free peptides (30). Moreover, the hydrophobic biotin moiety of the isotope-coded tag can influence the retention time of peptides (31). Owing to those disadvantages, modified ICAT reagents were developed including the solid-phase ICAT where an isotope-coded tag is immobilized on glass beads and it can be detached easily by UV irradiation; and a cleavable ICAT by replacing deuterium with ^{13}C and cleaving the linker under acidic conditions (31).

1.2.1.2 Isobaric tag for relative and absolute quantitation (iTRAQ^{TM})

iTRAQ is an alternative chemical labeling method widely used for quantitative proteomics study. iTRAQ labeling occurred on the primary amine groups of the proteolytic peptides by employing an N-hydroxysuccinimide (NHS) ester derivative, linking a mass balance group and a reporter group via an amide bond (32) (Figure 1.3a). Differentially labeled peptides result in a single peak in MS scans with the same *m/z*, and, by releasing the mass balancing carbonyl moiety as a neutral fragment, MS/MS analysis provides the quantitative information for the proteins. The availability of 8-plex labeling reagents that co-elute avoids increasing complexity at MS level. Because the labeled peptides are isobaric, the complexity of peptide in MS/MS spectra is not increased therefore theoretically increases sensitivity. Thus, iTRAQ labeling is suitable for multiplexed analysis of different biological samples (26). In fact, iTRAQ labeling already showed better peptide coverage than ICAT (27). However, due to the fact that the labeling occurs at peptide level after enzymatic digestion, an obvious drawback for this technique is the artificially introduced increase in sample complexity.

iTRAQ strategy for quantifying differential protein expression is shown in Figure 1.3b. Peptides from four different biological samples (can be applied up to eight samples) were labeled in parallel with four iTRAQ reagents. As depicted in MS/MS spectra, the relative intensities of the reporter ions which are associated with specific labeling reagents reflect the relative abundances of the peptides with the specific labelings in different samples.

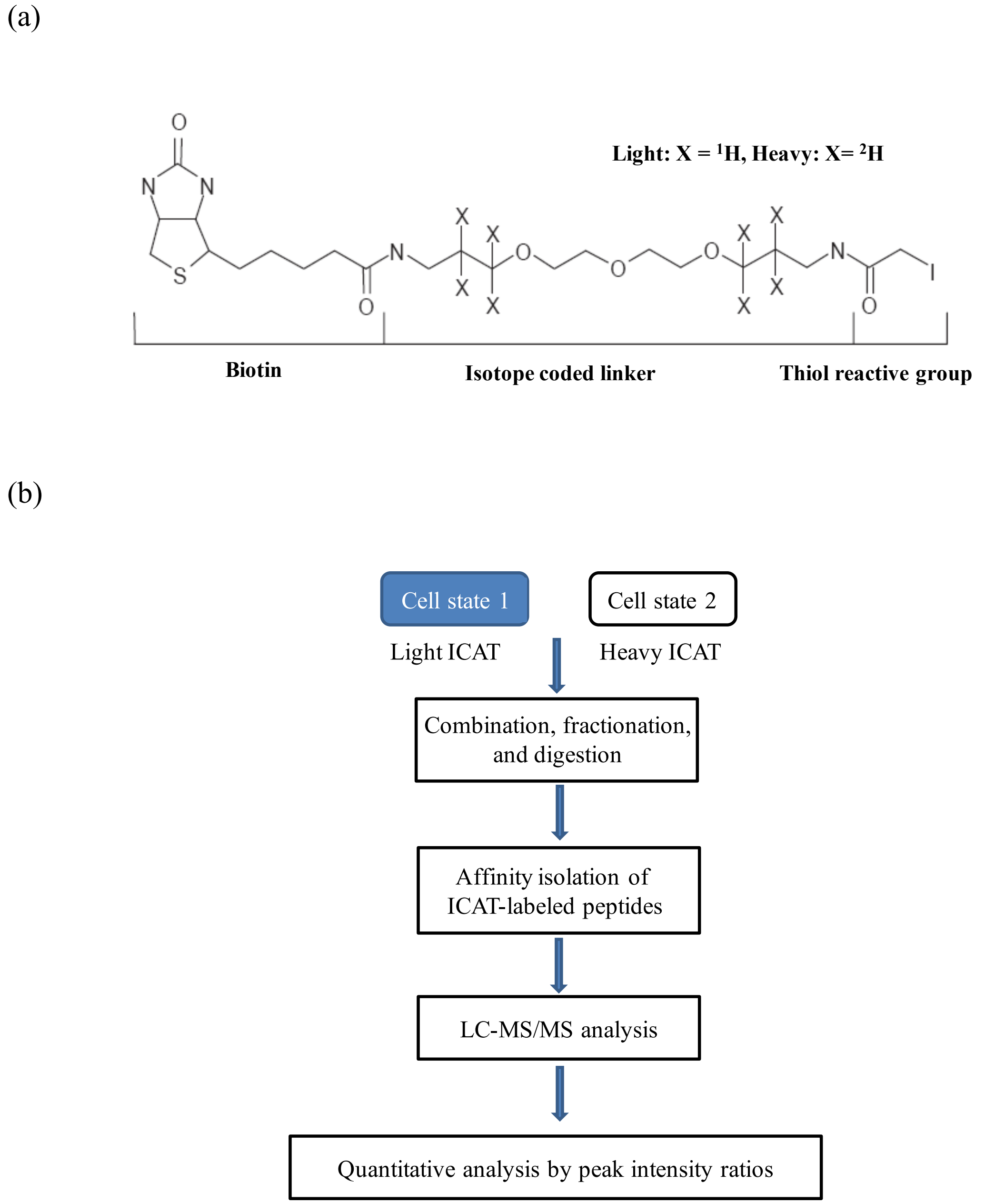


Figure 1.2. (a) The structure of ICAT reagents. It contains a biotin tag, a deuterium (or hydrogen)-labeled linker and a thiol reactive group towards cysteine side chain. (b). The workflow for using ICAT strategy in quantitative proteomic analysis.

1.2.1.3 Other chemical labeling strategies

Other than the widely used ICAT and iTRAQ techniques, tandem mass tags (TMTs), one of isobaric tags which label the amino group at the peptide level and achieves the quantification based on MS/MS (33), and methyl esterification which targets carboxylic acid on the side chains of aspartic and glutamic acids as well as peptide carboxyl terminus by converting them into the corresponding methyl esters are also broadly used in quantitative proteomics studies (34).

1.2.2 Metabolic labeling strategies

Metabolic labeling is another strategy which relies on the incorporation of stable isotopes into living cells, resulting in a replacement of the whole proteome with stable isotopes (35). Specific stable isotopes are introduced into culture media that, after sufficient cell doubling and protein turnover, facilitate the labeling of all the cellular proteins with stable isotopes. With proper sample mixing, enzymatic digestion, fractionation and clean-up processes, different fractions are subjected to LC-MS/MS for further analysis. Because of the complete labeling of the whole proteome, a mass shift between the unlabeled and labeled versions of all the peptides could be distinguished by MS. By comparing the relative peak intensities between the peptide pairs, protein quantification can be accomplished.

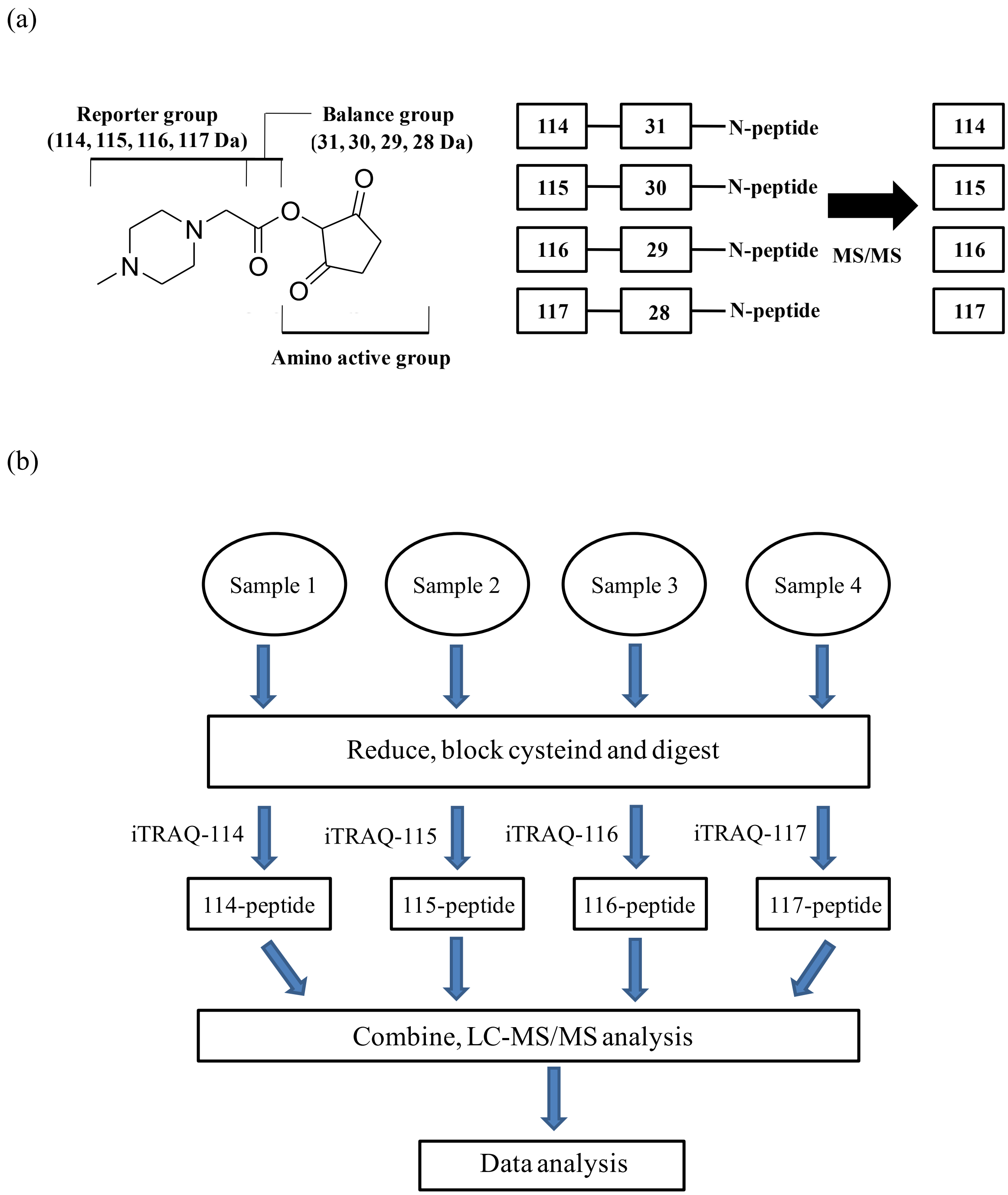


Figure 1.3. iTRAQ labeling and its workflow. (a) The structure of iTRAQ labeling reagents. (b) Workflow of a general iTRAQ experiment.

The most remarkable merit of metabolic labeling over chemical labeling is that the labeling is complete before subsequent fractionation, purification and enrichment steps, thereby minimizing artificial in errors. In addition, because no additional cleavage reaction occurs, the enzyme efficiency can be improved.

1.2.2.1 Stable isotope labeling by amino acids in cell culture (SILAC)

First reported by Mann and coworkers (36), SILAC has become the most widely used metabolic labeling method in the quantitative proteomics field. SILAC, as a successful metabolic labeling strategy, has many advantages: (1) It incorporates stable isotope(s) into the whole proteome by adding the labeled, essential amino acids into amino acid-depleted cell culture media; (2) it automatically labels cellular proteins without introducing chemical mass tags or affinity purification steps; (3) the incorporation is complete and it can capture subtle changes in protein expression; (4) it is a convenient, straightforward and affordable labeling strategy which is compatible with almost all cell types and culture conditions (36).

In a SILAC labeling experiment, two populations of cells, either of different types (e.g. normal cells and cancer cells) or under different treatment conditions, are cultured in parallel in “heavy” media, which contains isotope-labeled amino acids, and “light” media containing normal amino acids, respectively (Figure 1.4). Lysates from both heavy and light cells are extracted and combined at 1:1 ratio after at least five cell doublings.

Followed with appropriate separation using SDS-PAGE or SCX chromatography and enzymatic digestion, the peptide mixtures are analyzed by a mass spectrometer. Quantification is achieved by comparing the peak intensities between heavy and light peptides.

For a typical SILAC-based study on drug-induced proteome perturbation, at least one forward and one reverse labeling experiments should be performed. In the forward SILAC experiment, the light-labeled, drug-treated lysate and the heavy-labeled, control lysate are combined at 1:1 ratio (w/w); in the reverse SILAC experiment, the drug treatment and labeling are reversed (15). By performing both forward and reverse SILAC labelings, bias introduced by experimental conditions can be eliminated.

Recently, with the increasing needs on understating the proteome changes in tissues and even whole organisms, Mann et al. (37) developed a so-called “spike-in” SILAC cell line for tissue labeling. Labeling on the organisms can also be achieved by feeding mice with food containing heavy amino acids in place of normal ones for at least two generations (38).

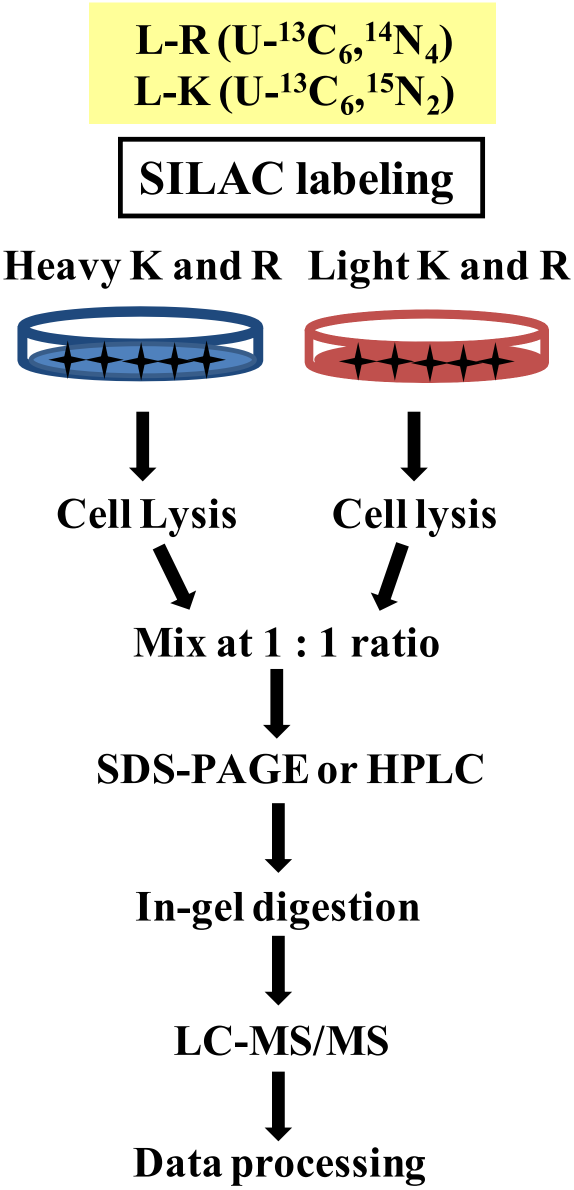


Figure 1.4. Workflow for standard SILAC labeling strategy.

1.2.2.2 ^{15}N labeling

^{15}N labeling has been used for the quantitative comparison of protein changes in bacteria and yeast (39) as well as mammalian cells (40) for more than a decade. Although ^{15}N labeling showed excellent promise in labeling microorganisms, it has difficulties in being extensively applied to mammalian systems. Firstly, the high cost of ^{15}N-substituted culture media prevents this method from being generally used for mammalian cell culture. Secondly, the labeling efficiency is not high enough because of the variable numbers of nitrogen atoms in different amino acids and the difficulties in eliminating the unlabeled nitrogen source in culture media, resulting in the complicated isotopic distributions in the mass spectra (17).

1.3 Label-free quantification

Label-free is a simple, fast quantification method by measuring and comparing either the peak intensity of the peptide precursor ions or the number of tandem mass spectra that are used to identify the peptides (41). Although label-free quantification is the least accurate method due to large variation(s) introduced by experimental steps, it is still frequently used for several reasons. First, label-free is not time-consuming because it avoids the labeling process, which renders the quantification very straightforward. Second, there is no limitation for the number of experiments or samples to compare and it is applicable to all kinds of biological samples. Third, being an economical method, it bears a very high dynamic range. However, label-free quantification has to be performed cautiously and the number of experimental steps should be as small as possible, particularly for spectral counting which has a relatively poor linearity and accuracy (42).

The beauty of quantitative proteomic strategy is to monitor, in the global proteome, the changes in protein expression levels at a time, which draws a complete picture of the whole proteome changes and provides implications for biological pathways and biomarker analysis. In this dissertation, I conducted several quantitative proteomics studies based on LC-MS/MS coupled with SILAC labeling strategy to understand the mechanisms of action of anti-cancer drugs and arsenite.

1.4 Nucleoside analogue-based anti-cancer drugs

Nucleoside analogues constitute a family of drugs which include antiviral agents, cytotoxic compounds and immnosuppressive molecules (43). Despite their extensive use in treating viral infection, nucleoside analogues have emerged as a class of chemotherapeutic agent. Anticancer nucleoside analogues are composed of purine and pyrimidine nucleosides and nucleobases. By mimicking the physiological characteristics of nucleosides in absorbance and metabolism, these analogues inhibit DNA synthesis through incorporation into newly synthesized DNA. Some of these drugs inhibit the activity of important enzymes involved in the endogenous nucleoside synthesis, giving rise to the cell death (44). Two primary purine analogues cladribine and fludarabine are known to be used in the treatment of low-grade malignant disorders of blood (45), whereas gemcitabine, which is widely prescribed in treating hematological malignancies and solid tumor, is a pyrimidine analogue (46).

The common mechanism of action of nucleoside analogues is shown in Figure 1.5. The hydrophilic nucleoside analogues are transported into cells by transporter proteins. After being activated into corresponding triphosphate derivatives, they can exert their cytotoxic effect through incorporation into DNA and RNA, inhibiting enzymatic activities and DNA synthesis (45).

1.4.1 Thiopurines

Thiopurines, including azathioprine, 6-mercaptopurine (6-MP) and 6-thioguanine (^{S}G) as shown in Figure 1.6a, are among the most successful chemotherapeutic agents and immunosuppressive agents that are used for treating various human diseases, including acute lymphoblastic leukemia (ALL) and chronic inflammation (47-49). Among them, 6-MP and ^{S}G are widely prescribed for ALL treatment.

The cytotoxic effects of thiopurines are thought to mainly depend on their metabolic activation and incorporation into DNA, resulting in the abnormal functions of various DNA metabolizing enzymes including DNA polymerases, ligases, and endonucleases (44). In addition, ^{S}G in DNA has been shown to be spontaneously methylated by *S*-adenosyl-L-methionine (SAM) to yield S^{6}-methylthioguanine (S^{6}mG). The misincorporation with thymine (T) opposite S^{6}mG during DNA replication triggers

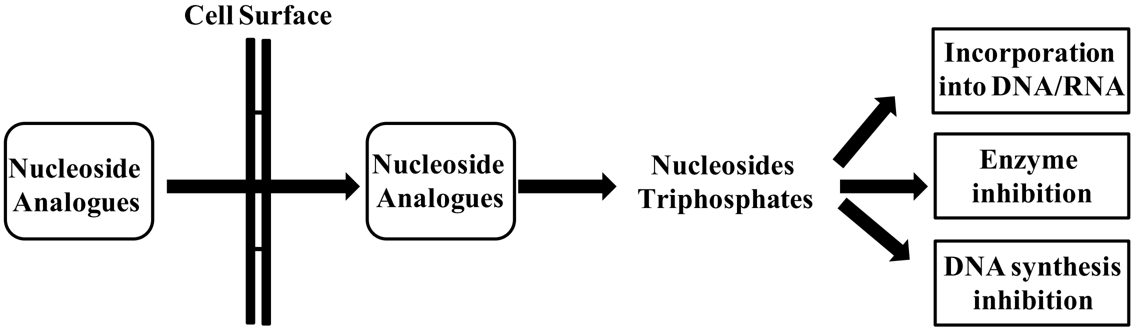


Figure 1.5. The common mechanism of action of nucleoside analogues as chemotherapeutic drugs (Adapted from (45)).

post-replicative mismatch repair (MMR) machinery (50), and the resultant futile cycles of repair synthesis may ultimately lead to cell death (51) (Figure 1.6b).

Despite the well-established MMR pathway, ${}^{S}G$ itself was found to be highly mutagenic and can direct the misincorporation of thymine during DNA replication at significant frequencies (*circa* 10%) in both *E. coli* and human cells (52, 53). In addition, some MMR-deficient cell lines are also sensitive towards thiopurine drugs (44). Thus, the activation of MMR may not be the sole pathway inducing the cytotoxic effects of thiopurine drugs.

1.4.2 5-aza-2'-deoxytidine and 5-azacytidine

DNA methylation, along with histone post-translational modification and microRNA pathway, are three major mechanisms constituting the epigenetic regulation, which plays significant roles in carcinogenesis and tumor progression (54). Among them, DNA methylation is the most extensively studied epigenetic modification in mammals. DNA is primarily methylated at 5-position of cytosine ring by DNA (cytosine-5) methyltransferases (DNMTs) (55). “CpG Island”, a CpG-enriched region which contains approximately 60% of human gene promoters and is responsible for maintaining transcriptional activities (56), remains free of methylation during development (57). Global DNA hypomethylation and CpG promoter DNA hypermethylation are known to occur in human tumors, where promoter cytosine methylation inhibits gene expression

(a)

[redacted] [redacted] [redacted]

6-thioguanine (^{S}G) **6-mercaptopurine** **azathioprine**

(b)

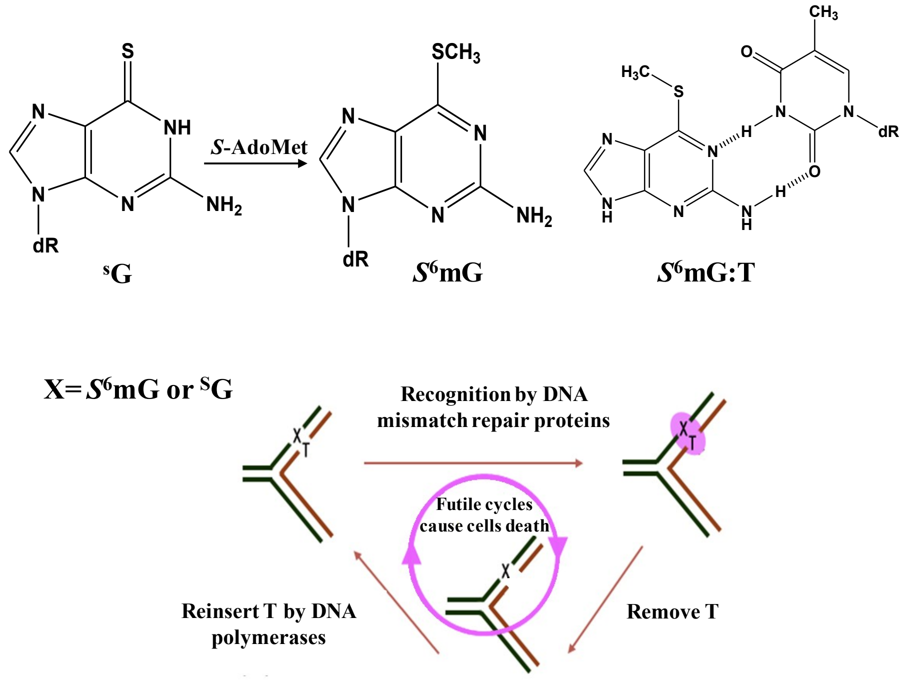


Figure 1.6. Thiopurine drugs and the triggering of mismatch repair pathway by their metabolite. (a) The structures of thiopurine drugs. (b) MMR-based mechanism of ^{S}G (Adapted from (58)).

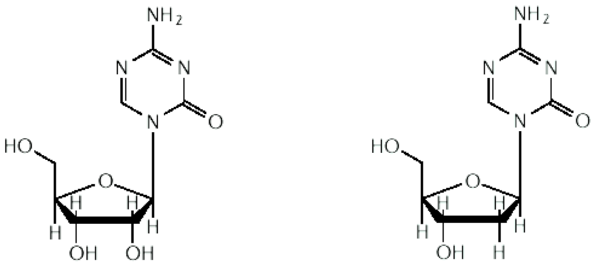
and results in long-term gene silencing (54). In this vein, the first epigenetic therapy was designed as DNA methylation inhibitors (55).

5-aza-2'-deoxycytidine (5-Aza-CdR) and 5-azacytidine (5-Aza-CR) have been reported to inhibit DNA methylation and induce cellular differentiation at micromolar concentrations (59). 5-Aza-CdR and 5-Aza-CR are two major cytidine nucleoside analogues which have been approved by FDA for the treatment of myelodysplastic syndrome, a hematological disease (60), and are widely used for treating hematological diseases including acute and chronic myeloid leukemia (AML and CML) (61) (Figure 1.7a). After being phosphorylated by deoxycytidine kinase and cytidine kinase into the corresponding mono-, di-, and triphosphates, 5-Aza-CdR and 5-Aza-CR are incorporated into DNA (62), a place where these nucleoside analogues can covalently bind to the cysteine residue at the active site of DNMTs and inhibit the transfer of methyl groups during DNA synthesis, which is believed to account for their cytotoxic effect (63) (Figure 1.7b). However, the detailed mechanisms underlying the cytotoxic effects of these drugs, particularly which target gene(s) becomes epigenetically reactivated, remain poorly defined.

1.5 Arsenite

Inorganic arsenic is a widespread environmental containment that has high abundance in soil, water, and airborne particles. As an ancient Chinese drug, arsenite trioxide (As_2O_3) has shown great therapeutic success in the treatment of acute promyelocytic leukemia (APL) (64). As_2O_3 exerts its therapeutic effect by promoting degradation of an oncogenic protein, i.e., the PML-RARα fusion protein that drives the growth of APL cells, where As(III) displaces Zn^{2+} ions in zinc fingers of the oncoprotein and triggers its proteasomal degradation (65). On the other hand, chronic exposure to arsenic, even at a very low dose, can induce cancers of skin, bladder, lung, kidney and liver (66). As(III), which is more toxic than As(V), may contribute to the increased cancer risks. Arsenite has been reported to stimulate the formation of oxyradicals (67), inhibit DNA damage repair (68), modulate DNA and histone methylation in mammalian cells (69), and induce apoptosis by introducing oxidative damage (70). A large body of evidence suggested that the mechanism of action of arsenite as a carcinogen is related to epigenetic regulation (71). It has been reported that the treatment with arsenite gave rise to a reduced level of *S*-adenosyl-L-methionine along with an elevated global DNA hypomethylation and decreased activity of DNA methyltransferase (72).

(a)



5-azacytidine (5-aza-CR) **5-aza-2’-deoxycytidine (5-aza-CdR)**

(b)

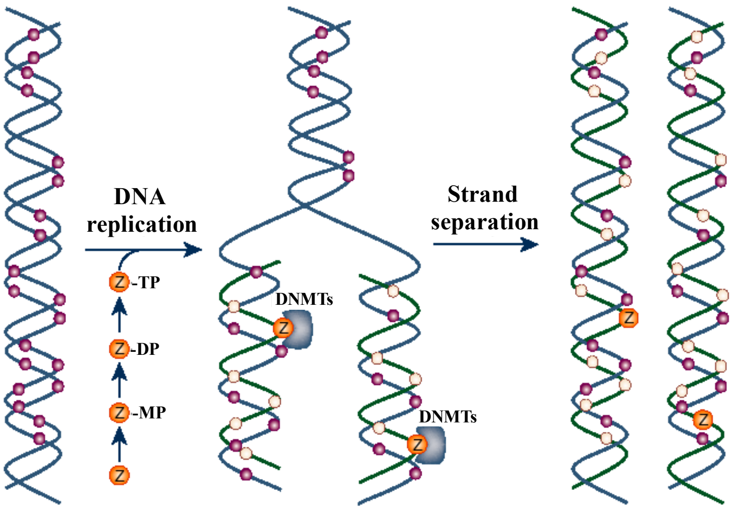


Figure 1.7. 5-Aza-CR and 5-Aza-CdR and their proposed mechanism of action.(a) The structures of 5-Aza-CR and 5-Aza-CdR. (b) The proposed mechanism of deoxynucleoside analogues (Adapted from (74)). ‘Z’ stands for the nucleoside analogues, dark and light circles designate methlyated and unmethylated cytosines, respectively.

Consistently, another study suggested that arsenite was associated with increased methylation of promoter regions in tumor suppressor genes, *RASSF1A* and *RPSS3* of bladder cancer (73).

In addition to DNA methylation, arsenite is also linked with hisone modification, another epigenetic mechanism. Arsenite was found to increase the level of dimethylated histone H3 at lysine 9 and trimethylated histone H3 at lysine 4 in A549 human lung carcinoma cells (71).

Although various possible mechanisms have been proposed, the mechanisms through which inorganic arsenite induces carcinogenesis remain incompletely understood. Quantitative proteomic analysis, as a mature method to study the alternation of the protein expression levels in the entire proteome scale, is capable of providing a more complete picture on understanding the fundamental mechanism of arsenite-induced cytotoxic effect.

1.6 Scope of the dissertation

The research covered in this dissertation focuses on the study, by employing a LC-MS/MS coupled with SILAC strategy, to understand the mechanisms of action of anti-cancer drugs and arsenite.

In Chapter two, we utilized a mass spectrometry-based quantitative proteomic method to analyze the 5-Aza-CdR-induced perturbation of protein expression in Jurkat-T cells at the global proteome scale. In particular, we found that drug treatment led to substantially reduced expression of farnesyl diphosphate synthase (FDPS) and farnesyl diphosphate farnesyltransferase (FDFT1), two important enzymes involved in *de novo* cholesterol synthesis. Exposure to 5-Aza-CdR also led to epigenetic reactivation of dipeptidyl peptidase 4 (*DPP4*) gene. Our study revealed, for the first time, that 5-Aza-CdR exerts its cytotoxic effects in leukemia and melanoma cells through epigenetic reactivation of *DPP4* gene and the resultant inhibition of cholesterol biosynthesis in these cells.

In Chapter three, we performed a quantitative assessment of global protein expression in control and ^{S}G-treated Jurkat-T cells by employing by SILAC and LC-MS/MS analysis. LC-MS/MS quantification results uncovered substantially decreased expression of a large number of proteins in mitochondrial respiratory chain complex, and the Ingenuity Pathway Analysis of the significantly altered proteins showed that ^{S}G treatment induced mitochondrial dysfunction. This was accompanied with diminished uptake of MitoTracker Deep Red and elevated formation of oxidatively induced DNA lesions, including 8,5’-cyclo-2’-deoxyadenosine and 8,5’-cyclo-2’-deoxyguanosine.

In Chapter four, we employed SILAC in conjunction with LC-MS/MS, to assess quantitatively the perturbation arsenite-induced protein expression in cultured human skin fibroblast cells.

By performing Ingenuity Pathway Analysis, we uncovered more than 10 biological pathways being perturbed by arsenite treatment. Particularly, we found that nine proteins involved in the Nrf2-mediated oxidative stress response pathway were significantly up-regulated. In addition, arsenite induced the altered expression of a family of selenoproteins and a number of metallothioneins.

In Chapter five, we further explored the molecular mechanisms underlying the carcinogenic effects of arsenic. In this study, we found that arsenite could bind directly to the RING finger domains of RNF20 and RNF40 *in vitro* and in cells, and treatment with arsenite resulted in significantly reduced H2B ubiquitination in multiple cell lines. Exposure to arsenite also diminished the recruitment of BRCA1 and RAD51 to laser-induced DNA double strand break (DSB) sites, reduced the rate of DNA DSB repair in human cells, and rendered cells sensitive toward a radiomimetic agent, neocarzinostatin.

1.7 REFERENCES

1. (2004) Finishing the euchromatic sequence of the human genome. *Nature* 431, 931-945.

2. Godovac-Zimmermann, J., and Brown, L. R. (2001) Perspectives for mass spectrometry and functional proteomics. *Mass Spectrom. Rev.* 20, 1-57.

3. Patterson, S. D., and Aebersold, R. H. (2003) Proteomics: the first decade and beyond. *Nat. Genet.* 33 311-323.

4. Hoheisel, J. D. (2006) Microarray technology: beyond transcript profiling and genotype analysis. *Nat. Rev. Genet.* 7, 200-210.

5. Cheok, M. H., Yang, W., Pui, C. H., Downing, J. R., Cheng, C., Naeve, C. W., Relling, M. V., and Evans, W. E. (2003) Treatment-specific changes in gene expression discriminate in vivo drug response in human leukemia cells. *Nat. Genet.* 34, 85-90.

6. Vogel, C., and Marcotte, E. M. (2012) Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat. Rev. Genet.* 13, 227-232.

7. Schweitzer, B., and Kingsmore, S. F. (2002) Measuring proteins on microarrays. *Curr. Opin. Biotechnol.* 13, 14-19.

8. Park, P. J. (2009) ChIP-seq: advantages and challenges of a maturing technology. *Nat. Rev. Genet.* 10, 669-680.

9. Ho, J. W., Bishop, E., Karchenko, P. V., Negre, N., White, K. P., and Park, P. J. (2011) ChIP-chip versus ChIP-seq: lessons for experimental design and data analysis. *BMC Genomics* 12, 134-145.

10. Klose, J., and Kobalz, U. (1995) Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome. *Electrophoresis* 16, 1034-1059.

11. Santoni, V., Molloy, M., and Rabilloud, T. (2000) Membrane proteins and proteomics: un amour impossible? *Electrophoresis* 21, 1054-1070.

12. Aebersold, R., and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* 422, 198-207.

13. Trinkle-Mulcahy, L., Boulon, S., Lam, Y. W., Urcia, R., Boisvert, F. M., Vandermoere, F., Morrice, N. A., Swift, S., Rothbauer, U., Leonhardt, H., and Lamond,

A. (2008) Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes. *J. Cell. Biol.* 183, 223-239.

14. Tan, M. J., Luo, H., Lee, S., Jin, F. L., Yang, J. S., Montellier, E., Buchou, T., Cheng, Z. Y., Rousseaux, S., Rajagopal, N., Lu, Z. K., Ye, Z., Zhu, Q., Wysocka, J., Ye, Y., Khochbin, S., Ren, B., and Zhao, Y. M. (2011) Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* 146, 1015-1027.

15. Zhang, F., Dai, X. X., and Wang, Y. S. (2012) 5-Aza-2 '-deoxycytidine induced growth inhibition of leukemia cells through modulating endogenous cholesterol biosynthesis. *Mol. Cell. Proteomics* 11, M111 016915.

16. Tran, J. C., Zamdborg, L., Ahlf, D. R., Lee, J. E., Catherman, A. D., Durbin, K. R., Tipton, J. D., Vellaichamy, A., Kellie, J. F., Li, M. X., Wu, C., Sweet, S. M. M., Early, B. P., Siuti, N., LeDuc, R. D., Compton, P. D., Thomas, P. M., and Kelleher, N. L. (2011) Mapping intact protein isoforms in discovery mode using top-down proteomics. *Nature* 480, 254-258.

17. Ong, S. E., and Mann, M. (2005) Mass spectrometry-based proteomics turns quantitative. *Nat. Chem. Biol.* 1, 252-262.

18. Towbin, H., Staehelin, T., and Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. 1979. *Proc. Natl. Acad. Sci. U.S.A.* 76, 4350-4354.

19. Ramos-Vara, J. A. (2005) Technical aspects of immunohistochemistry. *Vet. Pathol.* 42, 405-426.

20. Cravatt, B. F., Simon, G. M., and Yates, J. R., 3rd (2007) The biological impact of mass-spectrometry-based proteomics. *Nature* 450, 991-1000.

21. Kirkpatrick, D. S., Gerber, S. A., and Gygi, S. P. (2005) The absolute quantification strategy: a general procedure for the quantification of proteins and post-translational modifications. *Methods* 35, 265-273.

22. Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W., and Gygi, S. P. (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl. Acad. Sci. U.S.A.* 100, 6940-6945.

23. Bronstrup, M. (2004) Absolute quantification strategies in proteomics based on mass spectrometry. *Expert Rev. Proteomics* 1, 503-512.

24. Mayya, V., Rezual, K., Wu, L. F., Fong, M. B., and Han, D. K. (2006) Absolute quantification of multisite phosphorylation by selective reaction monitoring mass spectrometry - Determination of inhibitory phosphorylation status of cyclin-dependent kinases. *Mol. Cell. Proteomics* 5, 1146-1157.

25. Han, X. M., Aslanian, A., and Yates, J. R. (2008) Mass spectrometry for proteomics. *Curr. Opin. Chem. Biol.* 12, 483-490.

26. Bantscheff, M., Lemeer, S., Savitski, M. M., and Kuster, B. (2012) Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. *Anal. Bioanal. Chem.* 404, 939-965.

27. Nakamura, T., and Oda, Y. (2007) Mass spectrometry-based quantitative proteomics. *Biotechnol. Genet. Eng. Rev.* 24, 147-163.

28. Leitner, A., and Lindner, W. (2004) Current chemical tagging strategies for proteome analysis by mass spectrometry. *J. Chromatogr. B* 813, 1-26.

29. Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* 17, 994-999.

30. Moseley, M. A. (2001) Current trends in differential expression proteomics: isotopically coded tags. *Trends Biotechnol.* 19, 10-16.

31. Li, J., Steen, H., and Gygi, S. P. (2003) Protein profiling with cleavable isotope-coded affinity tag (cICAT) reagents: the yeast salinity stress response. *Mol. Cell. Proteomics* 2, 1198-1204.

32. Ross, P. L., Huang, Y. L. N., Marchese, J. N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., Purkayastha, S., Juhasz, P., Martin, S., Bartlet-Jones, M., He, F., Jacobson, A., and Pappin, D. J. (2004) Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* 3, 1154-1169.

33. Thompson, A., Schaefer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., Neumann, T., Johnstone, R. A. W., Mohammed, A. K. A., and Hamon, C. (2003) Tandem mass tags: A novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS *Anal. Chem.* 75, 1895-1904.

34. Goodlett, D. R., Keller, A., Watts, J. D., Newitt, R., Yi, E. C., Purvine, S., Eng, J. K., von Haller, P., Aebersold, R., and Kolker, E. (2001) Differential stable isotope labeling of peptides for quantitation and de novo sequence derivation. *Rapid Commun. Mass Spectrom.* 15, 1214-1221.

35. Mann, M. (2006) Functional and quantitative proteomics using SILAC. *Nat. Rev. Mol. Cell. Bio.* 7, 952-958.

36. Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* 1, 376-386.

37. Geiger, T., Cox, J., Ostasiewicz, P., Wisniewski, J. R., and Mann, M. (2010) Super-SILAC mix for quantitative proteomics of human tumor tissue. *Nat. Methods* 7, 383-385.

38. Kruger, M., Moser, M., Ussar, S., Thievessen, I., Luber, C. A., Forner, F., Schmidt, S., Zanivan, S., Fassler, R., and Mann, M. (2008) SILAC mouse for quantitative proteomics uncovers kindlin-3 as an essential factor for red blood cell function. *Cell* 134, 353-364.

39. Oda, Y., Huang, K., Cross, F. R., Cowburn, D., and Chait, B. T. (1999) Accurate quantitation of protein expression and site-specific phosphorylation. *Proc. Natl. Acad. Sci. U.S.A.* 96, 6591-6596.

40. Conrads, T. P., Alving, K., Veenstra, T. D., Belov, M. E., Anderson, G. A., Anderson, D. J., Lipton, M. S., Pasa-Tolic, L., Udseth, H. R., Chrisler, W. B., Thrall, B. D., and Smith, R. D. (2001) Quantitative analysis of bacterial and mammalian proteomes using a combination of cysteine affinity tags and N-15-Metabolic labeling. *Anal. Chem.* 73, 2132-2139.

41. Zhu, W., Smith, J. W., and Huang, C. M. (2010) Mass spectrometry-based label-free quantitative proteomics. *J. Biomed. Biotechnol.* 2010, 840518.

42. Old, W. M., Meyer-Arendt, K., Aveline-Wolf, L., Pierce, K. G., Mendoza, A., Sevinsky, J. R., Resing, K. A., and Ahn, N. G. (2005) Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Mol. Cell. Proteomics* 4, 1487-1502.

43. Sarafianos, S. G., Hughes, S. H., and Arnold, E. (2004) Designing anti-AIDS drugs targeting the major mechanism of HIV-1 RT resistance to nucleoside analog drugs. *Int. J. Biochem. Cell Biol.* 36, 1706-1715.

44. Galmarini, C. M., Mackey, J. R., and Dumontet, C. (2001) Nucleoside analogues: mechanisms of drug resistance and reversal strategies. *Leukemia* 15, 875-890.

45. Galmarini, C. M., Mackey, J. R., and Dumontet, C. (2002) Nucleoside analogues and nucleobases in cancer treatment. *Lancet Oncol.* 3, 415-424.

46. Cheson, B. D. (1992) New antimetabolites in the treatment of human malignancies. *Semin. Oncol.* 19, 695-706.

47. Elion, G. B. (1989) The purine path to chemotherapy. *Science* 244, 41-47.

48. Karran, P. (2006) Thiopurines, DNA damage, DNA repair and therapy-related cancer. *Br. Med. Bull.* 79-80, 153-170.

49. Pui, C. H., and Jeha, S. (2007) New therapeutic strategies for the treatment of acute lymphoblastic leukaemia. *Nat. Rev. Drug Discov.* 6, 149-165.

50. Swann, P. F., Waters, T. R., Moulton, D. C., Xu, Y. Z., Zheng, Q. G., Edwards, M., and Mace, R. (1996) Role of postreplicative DNA mismatch repair in the cytotoxic action of thioguanine. *Science* 273, 1109-1111.

51. Karran, P., and Attard, N. (2008) Thiopurines in current medical practice: molecular mechanisms and contributions to therapy-related cancer. *Nat. Rev. Cancer* 8, 24-36.

52. Yuan, B., O'Connor, T. R., and Wang, Y. (2010) 6-Thioguanine and S^6-methylthioguanine are mutagenic in human cells. *ACS Chem. Biol.* 5, 1021-1027.

53. Yuan, B., and Wang, Y. (2008) Mutagenic and cytotoxic properties of 6-thioguanine, S^6-methylthioguanine, and guanine-S^6-sulfonic acid. *J. Biol. Chem.* 283, 23665-23670.

54. Jones, P. A., and Baylin, S. B. (2002) The fundamental role of epigenetic events in cancer. *Nat. Rev. Genet.* 3, 415-428.

55. Yoo, C. B., and Jones, P. A. (2006) Epigenetic therapy of cancer: past, present and future. *Nat. Rev. Drug Discov.* 5, 37-50.

56. Wang, Y., and Leung, F. C. C. (2004) An evaluation of new criteria for CpG islands in the human genome as gene markers. *Bioinformatics* 20, 1170-1177.

57. Suzuki, M. M., and Bird, A. (2008) DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet.* 9, 465-476.

58. Swann, P. F., Waters, T. R., Moulton, D. C., Xu, Y. Z., Zheng, Q., Edwards, M., and Mace, R. (1996) Role of postreplicative DNA mismatch repair in the cytotoxic action of thioguanine. *Science* 273, 1109-1111.

59. Jones, P. A., and Taylor, S. M. (1980) Cellular-differentiation, cytidine analogs and DNA methylation. *Cell* 20, 85-93.

60. Kaminskas, E., Farrell, A., Abraham, S., Baird, A., Hsieh, L. S., Lee, S. L., Leighton, J. K., Patel, H., Rahman, A., Sridhara, R., Wang, Y. C., and Pazdur, R. (2005) Approval summary: azacitidine for treatment of myelodysplastic syndrome subtypes. *Clin. Cancer Res.* 11, 3604-3608.

61. Plimack, E. R., Kantarjian, H. M., and Issa, J. P. (2007) Decitabine and its role in the treatment of hematopoietic malignancies. *Leuk. Lymphoma* 48, 1472-1481.

62. Christman, J. K. (2002) 5-Azacytidine and 5-aza-2 '-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy. *Oncogene* 21, 5483-5495.

63. Ghoshal, K., Datta, J., Majumder, S., Bai, S. M., Kutay, H., Motiwala, T., and Jacob, S. T. (2005) 5-Aza-deoxycytidine induces selective degradation of DNA methyltransferase 1 by a proteasomal pathway that requires the KEN box, bromo-adjacent homology domain, and nuclear localization signal. *Mol. Cell. Biol.* 25, 4727-4741.

64. de The, H., and Chen, Z. (2010) Acute promyelocytic leukaemia: novel insights into the mechanisms of cure. *Nat. Rev. Cancer* 10, 775-783.

65. Zhang, X. W., Yan, X. J., Zhou, Z. R., Yang, F. F., Wu, Z. Y., Sun, H. B., Liang, W. X., Song, A. X., Lallemand-Breitenbach, V., Jeanne, M., Zhang, Q. Y., Yang, H. Y., Huang, Q. H., Zhou, G. B., Tong, J. H., Zhang, Y., Wu, J. H., Hu, H. Y., de The, H., Chen, S. J., and Chen, Z. (2010) Arsenic trioxide controls the fate of the PML-RARalpha oncoprotein by directly binding PML. *Science* 328, 240-243.

66. Weinberg, C. (2004) Arsenic and drinking water. *Epidemiology* 15, 255-256.

67. Liu, S. X., Athar, M., Lippai, I., Waldren, C., and Hei, T. K. (2001) Induction of oxyradicals by arsenic: implication for mechanism of genotoxicity. *Proc. Natl. Acad. Sci. U.S.A.* 98, 1643-1648.

68. Wang, F., Zhou, X., Liu, W., Sun, X., Chen, C., Hudson, L. G., and Liu, K. J. (2013) Arsenite-induced ROS/RNS generation causes zinc loss and inhibits the activity of poly (ADP-ribose) polymerase-1. *Free Radic. Biol. Med.* 61, 249-256.

69. Kitchin, K. T. (2001) Recent advances in arsenic carcinogenesis: modes of action, animal model systems, and methylated arsenic metabolites. *Toxicol. Appl. Pharmacol.* 172, 249-261.

70. Wang, T. S., Kuo, C. F., Jan, K. Y., and Huang, H. (1996) Arsenite induces apoptosis in Chinese hamster ovary cells by generation of reactive oxygen species. *J. Cell. Physiol.* 169, 256-268.

71. Zhou, X., Sun, H., Ellen, T. P., Chen, H., and Costa, M. (2008) Arsenite alters global histone H3 methylation. *Carcinogenesis* 29, 1831-1836.

72. Zhao, C. Q., Young, M. R., Diwan, B. A., Coogan, T. P., and Waalkes, M. P. (1997) Association of arsenic-induced malignant transformation with DNA hypomethylation and aberrant gene expression. *Proc. Natl. Acad. Sci. U.S.A.* 94, 10907-10912.

73. Marsit, C. J., Karagas, M. R., Danaee, H., Liu, M., Andrew, A., Schned, A., Nelson, H. H., and Kelsey, K. T. (2006) Carcinogen exposure and gene promoter hypermethylation in bladder cancer. *Carcinogenesis* 27, 112-116.

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

CHAPTER 2

5-Aza-2’-deoxycytidine Induced Growth Inhibition of Leukemia Cells through Modulating Endogenous Cholesterol Biosynthesis

INTRODUCTION

Epigenetic events, defined as mitotically and meiotically heritable changes in gene expression that are not due to alteration in primary DNA sequence (1), play important roles in carcinogenesis and tumor progression (2). DNA cytosine methylation, post-translational modifications of core histones, and microRNA pathway constitute three major mechanisms of epigenetic regulation. Global DNA hypomethylation and promoter DNA hypermethylation are known to occur in human tumors, where promoter cytosine methylation inhibits gene expression and results in long-term gene silencing (2).

Cytosine methylation pattern is maintained during cell division by DNA (cytosine-5)-methyltransferase 1 (DNMT1), and DNA methylation inhibitors were the first epigenetic drugs used for cancer treatment (3). 5-Azacytidine (5-Aza-CR) and 5-aza-2’-deoxycytidine (5-Aza-CdR) are among the many cytosine nucleoside analogs that can inhibit DNA methylation and induce cellular differentiation (4). These nucleoside analogs are incorporated into DNA of tumor cells during DNA replication and 5-azacytosine in DNA can bind to the cysteine residue at the active site of DNMTs (5).

This covalent and irreversible binding of the enzyme to drug-substituted DNA is believed to be the principal mechanism of cytotoxicity (5), though it was also found that 5-Aza-CdR treatment could lead to the proteasomal degradation of DNMT1 independent of its catalytic cysteine residue (6). 5-Aza-CR and 5-Aza-CdR have been approved by FDA for the treatment of myelodysplastic syndromes and are widely studied for the treatment of hematological diseases (7), including acute and chronic myeloid leukemia (AML and CML) (8). However, the detailed mechanisms underlying the cytotoxic effects of these drugs, particularly which target gene(s) becomes epigenetically reactivated and results in the growth inhibition of leukemic cells, remain poorly defined.

To exploit the molecular mechanisms contributing to the anticancer activity of 5-Aza-CdR in leukemia cells, we employed LC-MS/MS together with stable isotope labeling by amino acid in cell culture (SILAC) to assess, at the global proteome scale, the perturbation in protein expression of Jurkat-T human leukemia cells upon 5-Aza-CdR treatment. In this context, SILAC is a simple and efficient metabolic isotope-labeling method; when combined with LC-MS/MS analysis, the method can afford accurate quantification of subtle changes of protein abundance in the whole proteome (9). With this method, we quantified more than 2780 unique proteins, 188 of which were significantly altered upon 5-Aza-CdR treatment. Importantly, the quantitative proteomic experiment revealed the 5-Aza-CdR-induced down-regulation of two essential enzymes in cholesterol biosynthesis, namely, farnesyl diphosphate synthase (FDPS) and farnesyl diphosphate farnesyltransferase (FDFT1, a.k.a. squalene synthase). This finding, along with follow-up studies allowed us to discover, for the first time, that 5-Aza-CdR exerts its cytotoxic effect via modulating cholesterol biosynthesis in leukemia and melanoma cells, which involves epigenetic reactivation of *DPP4* gene.

MATERIALS AND METHODS

Cell Culture

All reagents unless otherwise stated were from Sigma, and all cell lines were obtained from ATCC (Manassas, VA). Jurkat-T, HL60 and K562 cells were cultured in Iscove's modified minimal essential medium (IMEM) supplemented with 10% fetal bovine serum (FBS, Invitrogen, Carlsbad, CA), 100 IU/mL penicillin and 100 μg/mL streptomycin in 75 cm^2 culture flasks. The WM-266-4 cells were cultured under the same conditions except that Eagle's minimum essential medium (EMEM) was used. Cells were maintained in a humidified atmosphere with 5% CO_2 at 37°C, with medium renewal of 2-3 times a week depending on cell density. For SILAC experiments, the IMEM medium without L-lysine or L-arginine was custom-prepared following ATCC formulation. The complete light and heavy IMEM media were prepared by adding light or heavy lysine ($[^{13}C_6, {}^{15}N_2]$-L-lysine) and arginine ($[^{13}C_6]$-L-arginine), along with dialyzed FBS (Invitrogen), to the lysine, arginine-depleted medium. The Jurkat-T cells were cultured in heavy IMEM medium for at least 10 days to achieve complete stable isotope incorporation.

5-Aza-CdR treatment and sample preparation

Jurkat-T cells, at a density of ~ 7×10^5 cells per mL in light or heavy IMEM medium, were treated with 5 μM 5-Aza-CdR for 24 hr. After treatment, the light and heavy cells were harvested by centrifugation at 300× g at 4°C for 5 min, and washed for three times with ice-cold PBS to remove culture medium and FBS. Cells were lysed with CelLytic^{TM} M lysis buffer supplemented with 1 mM PMSF and a protease inhibitor cocktail. The resulting cell lysate was centrifuged at 16,000× g at 4°C for 30 min and supernatant collected. The protein concentration in the cell lysate was measured using Quick Start^{TM} Bradford Protein Assay (Bio-Rad, Hercules, CA). In forward SILAC, the lysate of light labeled, drug-treated cells and that of the heavy labeled control cells were combined at 1:1 ratio (w/w), whereas the heavy labeled, drug-treated cell lysate was mixed equally with the light labeled, control lysate in the reverse SILAC experiment (Figure 2.1a).

SDS-PAGE separation and in-gel digestion

The above equi-mass mixture of light and heavy lysates was separated on a 12% SDS-PAGE with a 4% stacking gel and stained with Coomassie blue. The gel was cut into 20 slices, and the proteins in individual gel slices were separately reduced in-gel with dithiothreitol (DTT) and alkylated with iodoacetamide (IAM). The proteins were subsequently digested at 37°C with trypsin (Promega, Madison, WI) for overnight. Following digestion, peptides were extracted from gels with 5% acetic acid in H_2O and then with 5% acetic acid in CH_3CN/H_2O (1:1, v/v). The resulting peptide mixtures were dried and stored at –80°C until further analysis.

LC-MS/MS for protein identification and quantification

On-line LC-MS/MS analysis was performed on an LTQ-Orbitrap Velos mass spectrometer coupled with an EASY n-LCII HPLC system and a nanoelectrospray ionization source (Thermo, San Jose, CA). The sample injection, enrichment, desalting, and HPLC separation were conducted automatically on a homemade trapping column (150 μm×50 mm) and a separation column (75 μm×120 mm, packed with ReproSil-Pur C18-AQ resin, 5 μm in particle size and 300 Å in pore size, Dr. Maisch HPLC GmbH, Germany). The peptide mixture was first loaded onto the trapping column with a solvent mixture of 0.1% formic acid in CH_3CN/H_2O (2:98, v/v) at a flow rate of 3.0 μL/min. The peptides were then separated using a 120-min linear gradient of 2-40% acetonitrile in 0.1% formic acid at a flow rate of 300 nL/min.

The LTQ-Orbitrap Velos mass spectrometer was operated in the positive-ion mode, and the spray voltage was 1.8 kV. All MS/MS spectra were acquired in a data-dependent scan mode, where one full-MS scan was followed with twenty MS/MS scans. The full-scan MS spectra (from *m/z* 350 to 2000) were acquired with a resolution of 60,000 at *m/z* 400 after accumulation to a target value of 500,000. The twenty most abundant ions found in MS at a threshold above 500 counts were selected for fragmentation by collision-induced dissociation at a normalized collision energy of 35%.

Data processing

The LC-MS/MS data were employed for the identification and quantification of global proteome, which were conducted using Maxquant, Version 1.2.0.18 (10) against IPI database (11), version 3.68 with 87,061 entries to which contaminants and reverse sequences were added. The maximum number of miss-cleavages for trypsin was two per peptide. Cysteine carbamidomethylation and methionine oxidation were set as fixed and variable modifications, respectively. The tolerances in mass accuracy for MS and MS/MS were 25 ppm and 0.6 Da, respectively. Only those proteins with at least two distinct peptides being discovered from LC-MS/MS analyses were considered reliably identified. The protein expression ratio reported in the present study represented the normalized ratios determined by Maxquant, where the expression levels of the majority of proteins were assumed to be unchanged upon 5-Aza-CdR treatment and the median of log-transformed ratios of all quantified proteins was considered to be zero (12). The required false positive rate was set to 1% at the both peptide and protein levels, with the minimal required peptide length being set at 6 amino acids. The quantification was based on three independent SILAC and LC-MS/MS experiments, which included two forward and one reverse SILAC labelings. Only those proteins with alteration in expression levels being greater than 1.5 or less than 0.67 fold, and with quantification results from at least 2 sets (including both forward and reverse) of SILAC labeling experiments were considered significantly changed.

Exogenous cholesterol addition and cell viability assay

The cholesterol-BSA complex was prepared following a previously published method (13). Briefly, a 10-ml aliquot of 1% cholesterol in ethanol was mixed with an equal volume of doubly distilled water under continuous stirring at room temperature. The milk-like solution was then centrifuged at 2,000× g for 10 min. The supernatant was discarded, and the pellet was resuspended in a 10-mL solution containing 0.25 M sucrose and 1 mM EDTA (pH 7.3), followed by a gentle addition of 4 g BSA with continuous stirring at room temperature. Once the BSA was completely dissolved, the pH of the solution was adjusted to 7.3 with Tris, and the resulting solution was centrifuged at 12,000× g for 10 min at 4°C. The supernatant was collected and used for cholesterol addition experiments.

Jurkat-T, HL60, K562, and WM-266-4 cells were seeded in 6-well plates at a density of approximately 4×10^5 cells/mL and treated with 5-Aza-CdR at a final concentration of 5 μM. The cholesterol-BSA complex solution was added to the wells containing the control or 5-Aza-CdR-treated cells until the final cholesterol concentration reached 30 or 60 mg/L. After 12 or 24 hrs of treatment, cells were stained with trypan blue, and counted on a hemocytometer to measure cell viability.

Extraction and determination of the cellular cholesterol level

Cells were washed for 3 times with PBS and extracted with chloroform:methanol:water (2:1.1:0.9, v/v/v), following previously published procedures (14). The chloroform layer was washed 3 times with a methanol-water mixture (5:4, v/v), collected, and evaporated to dryness using a SpeedVac. The cholesterol level was measured by HPLC using methanol as mobile phase, and the effluent from the column was monitored with a UV detector at a wavelength of 210 nm. The cellular level of cholesterol was determined using a calibration curve constructed from HPLC analyses of different amounts of cholesterol, as described recently (15).

Quantitative real-time PCR

RNA was extracted using the RNeasy Mini Kit (QIAGEN). Approximately 1 μg RNA was reverse transcribed by employing M-MLV reverse transcriptase (Promega) with a poly(dT) primer. After incubating at 42°C for 60 min, the reverse transcriptase was deactivated by heating at 85°C for 5 min. Quantitative real-time PCR was performed using iQ SYBR Green Supermix kit (Bio-Rad) on a Bio-Rad iCycler system (Bio-Rad), and the running conditions were at 95°C for 3 min and 50 cycles at 95°C for 15 sec, 55°C for 30 sec, and 72°C for 45 sec. The comparative cycle threshold (Ct) method (ΔΔCt) was used for the relative quantification of gene expression, and *GAPDH* gene was used as the internal control (16). The mRNA level of each gene was normalized to that of the internal control.

siRNA treatment

Dharmacon (Lafayette, CO) ON-TARGET*plus SMART*pool siRNAs were employed to knockdown the expression of human *DPP4* genes in WM-266-4 cells, where Dharmacon siGENOME Non-Targeting siRNA was used as control.

Briefly, WM-266-4 cells were seeded in 6-well plates at 50-70% confluence level and transfected with approximately 1.5 μg DPP4 siRNA or control siRNA using

Lipofectamine 2000 (Invitrogen). After a 48-hr incubation, the cells were harvested by using trypsin-EDTA solution followed by centrifugation at 2,000 rpm at 4°C for 5 min and subsequently washed twice with PBS.

Western blot

Lysate of control and 5-Aza-CdR-treated Jurkat-T and K562 cells, along with that of DPP4 siRNA-treated WM-266-4 cells, were prepared following the above-described procedures. After SDS-PAGE separation, proteins were transferred to a nitrocellulose membrane using a solution containing 10 mM $NaHCO_3$, 3 mM Na_2CO_3, and 20% methanol. The membranes were blocked with 5% non-fat milk in PBS buffer containing 0.1% (v/v) Tween-20 (pH 7.5) for 7 hr and incubated overnight at 4°C with rabbit anti-FDPS antibody (1:400 dilution, Abgent, San Diego, CA) and rabbit anti-FDFT1 antibody (1:400 dilution) (Abgent). The membranes were washed with fresh PBS-T at room temperature for five times (10 min each). After washing, the membranes were incubated with HRP-conjugated secondary antibody (1:1000 dilution) at room temperature for 1 hr. The membranes were subsequently washed with PBS-T for five times. The secondary antibody was detected by using ECL Advance Western Blotting Detection Kit (GE Healthcare) and visualized with Hyblot CL autoradiography film (Denville Scientific, Inc., Metuchen, NJ). Intensities for immunoreactive bands were quantified using ImageJ (National Institutes of Health).

RESULTS AND DISCUSSION

5-Aza-CdR treatment and protein quantification

It is widely accepted that treatment of cells with 5-Aza-CdR could lead to the incorporation of the modified nucleoside into cellular DNA, which traps covalently the DNMTs thereby inducing DNA hypomethylation (5). However, the target gene(s), whose reactivation gives rise to the cytotoxic effect of the drug, remains to be identified. To explore the molecular targets of 5-Aza-CdR, we employed an unbiased quantitative proteomic approach to identify, at the entire proteome scale, the drug-induced perturbation of protein expression in Jurkat-T cells. To this end, we first established the optimal dose of 5-Aza-CdR by examining the survival rate of Jurkat-T cells upon treatment with different concentrations of 5-Aza-CdR. Based on trypan blue exclusion assay, a less than 5% cell death was observed after a 24-hr treatment with 5 μM 5-Aza-CdR; however, cell viability was significantly reduced (by ~20%) after a similar treatment with 10 μM 5-Aza-CdR. Thus, we chose 5 μM 5-Aza-CdR for the subsequent experiments to minimize the apoptosis-induced changes in protein expression. To obtain reliable quantification results, we conducted SILAC experiments in triplicate, including two forward and one reverse labelings (Figure 2.1a depicts the procedures for forward SILAC labeling).

LC-MS/MS allowed for the identification of a total of 3100 proteins in both forward and reverse SILAC experiments, and ~ 2780 of them were quantified (Figure 2.1b).

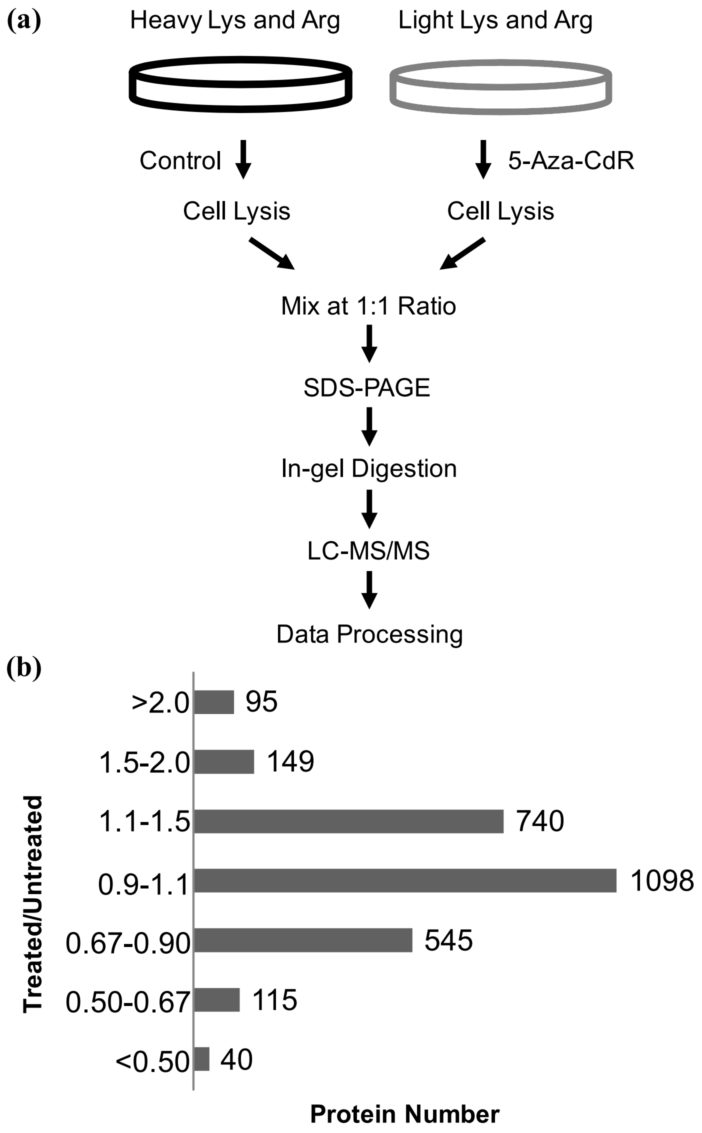


Figure 2.1. SILAC-based quantitative proteomic method for revealing the 5-Aza-CdR-induced perturbation of protein expression in the global proteome. (a) Flowcharts of forward SILAC combined with LC-MS/MS for the comparative analysis of protein expression in Jurkat-T cells upon 5-Aza-CdR treatment. (b) The distribution of expression ratios (treated/untreated) for the quantified proteins, including those quantified in only one set of SILAC labeling experiment.

The distribution of changes in protein expression levels arising from 5-Aza-CdR treatment is displayed in Figure 2.1b. By using a ratio of >1.5 or <0.67 as threshold for the significantly changed proteins, 188 exhibited significant changes with at least 2 peptides being identified in at least two sets of SILAC labeling experiments (Table A1). Figure 2.2 depicts the representative ESI-MS results for two tryptic peptides derived from farnesyl diphosphate synthase (FDPS), which reveals the 5-Aza-CdR-induced down-regulation of this protein.

5-Aza-CdR treatment led to down-regulation of FDPS and farnesyl diphosphate farnesyltransferase (FDFT1, a.k.a. squalene synthase**)**

We next performed protein interaction network and pathway analysis using the Ingenuity Pathway Analysis (IPA) software (17). Proteins exhibiting greater than a 1.5-fold change in expression upon the drug treatment were considered for the analysis. Networks represent a highly interconnected set of proteins derived from the input dataset. Biological functions and processes were assigned to networks by mapping the proteins in the network to functions in the Ingenuity ontology. Several pathways were found to be altered included steroid (cholesterol) biosynthesis, granzyme signaling, mitochondria dysfunction, etc. (Table A2).

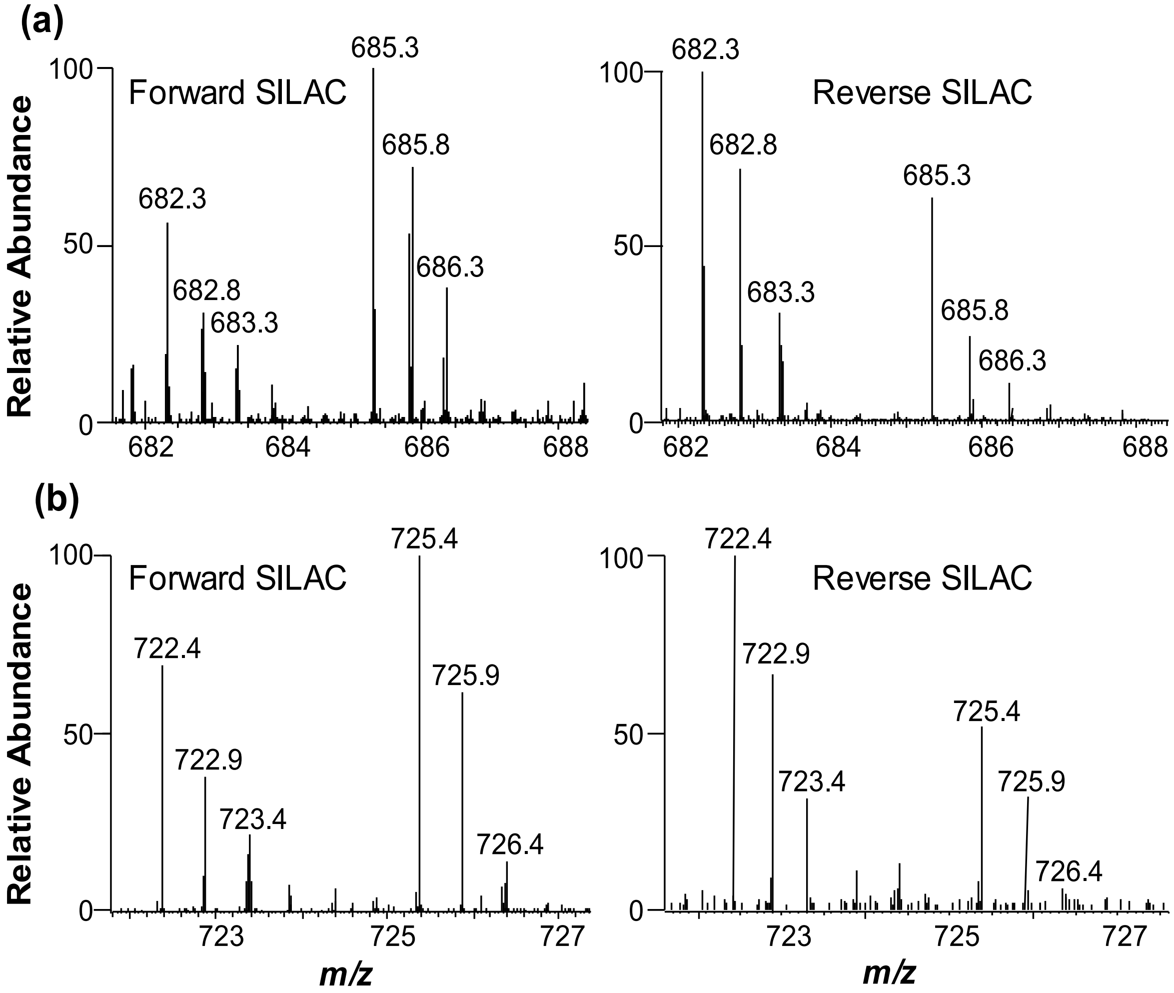


Figure 2.2. Representative ESI-MS results revealing the 5-Aza-CdR-induced down-regulation of FDPS. Shown are the MS for the $[M+2H]^{2+}$ ions of FDPS peptides EFWPQEVWSR and EFWPQEVWSR* (a), as well as TQNLPNCQLISR and TQNLPNCQLISR* (b) ('R*' designates the heavy arginine) from forward (left) and reverse (right) SILAC labeling experiments.

For the steroid biosynthesis pathway, results from our quantitative proteomic experiments showed that FDPS and FDFT1 were significantly reduced upon 5-Aza-CdR treatment. The expression ratios (treated/untreated) for FDPS and FDFT1 were determined to be 0.61 ± 0.03 and 0.48 ± 0.10, respectively, based on results from three independent SILAC labeling, drug treatment, and LC-MS/MS measurements. FDPS and FDFT1 catalyze the biosynthesis of FDP and the conversion of FDP to squalene, respectively (18). Viewing that both enzymes are essential for *de novo* cholesterol synthesis in human cells (14), we reasoned that 5-Aza-CdR treatment may give rise to diminished intracellular cholesterol levels. To test this, we extracted cholesterol from Jurkat-T cells and measured its level by HPLC analysis. It turned out that a 24-hr treatment with 5 μM 5-Aza-CdR led to a statistically significant decline of the cellular cholesterol content from 36 μg to 30 μg of cholesterol per 10^7 cells (Figure 2.3a). To examine if this observation is general, we measured cholesterol levels in two other leukemia cell lines (i.e. HL60 and K562) with and without 5-Aza-CdR treatment, and our results showed that the drug-induced decrease in cellular cholesterol content was even more pronounced in these two cell lines (Figure 2.3b&c). Therefore, the results confirmed the 5-Aza-CdR-induced down-regulation of FDPS and FDFT1, as observed from quantitative proteomic experiment. Along this line, the 5-Aza-CdR-induced down-regulation of FDPS and FDFT1 in Jurkat-T and K562 cells was further confirmed by Western blot analysis (Figure A1). Apart from the decreased expression of FDPS and

FDFT1, our quantitative proteomic results revealed the elevated expression of zyxin, Wiskott-Aldrich syndrome protein family member 2 (WASP2), as well as LIM and Src homology 3 domain protein 1 (LASP-1, Tables A1). These proteins co-localize with actin polymerization, which occurs in cholesterol-rich membrane microdomains (a.k.a. lipid rafts) (19-21). The overexpression of these proteins may reflect an enhanced membrane-cytoskeleton interaction in response to 5-Aza-CdR-induced cholesterol depletion (22, 23).

Diminished endogenous cholesterol biosynthesis in leukemic cells may contribute significantly to the cytotoxic effects of 5-Aza-CdR. In this vein, leukemia cells display enhanced rates of *de novo* cholesterol synthesis and lack of feedback inhibition of cholesterogenesis (24). Inhibition of endogenous cholesterol biosynthesis in leukemia cells suppresses their growth (25, 26). If this constitutes the major mechanism leading to the growth inhibition of leukemic cells, we expect that the 5-Aza-CdR-induced growth inhibition should be rescued by externally added cholesterol. It turned out that all three leukemia cell lines exhibited substantial growth inhibition upon 5-Aza-CdR treatment, which can indeed be abrogated by addition of cholesterol to the culture medium (Figure 2.4a&b displayed representative data for Jurkat-T and K562 cells). In this regard, the cholesterol content in 5-Aza-CdR-treated cells returned to control levels at 24 hr after the addition of exogenous cholesterol (Figure 2.3). These results demonstrated that the 5-Aza-CdR-induced growth inhibition of the three leukemia cell lines could be mainly attributed to the decreased levels of FDPS and FDFT1 and the resultant decline in endogenous cholesterol biosynthesis.

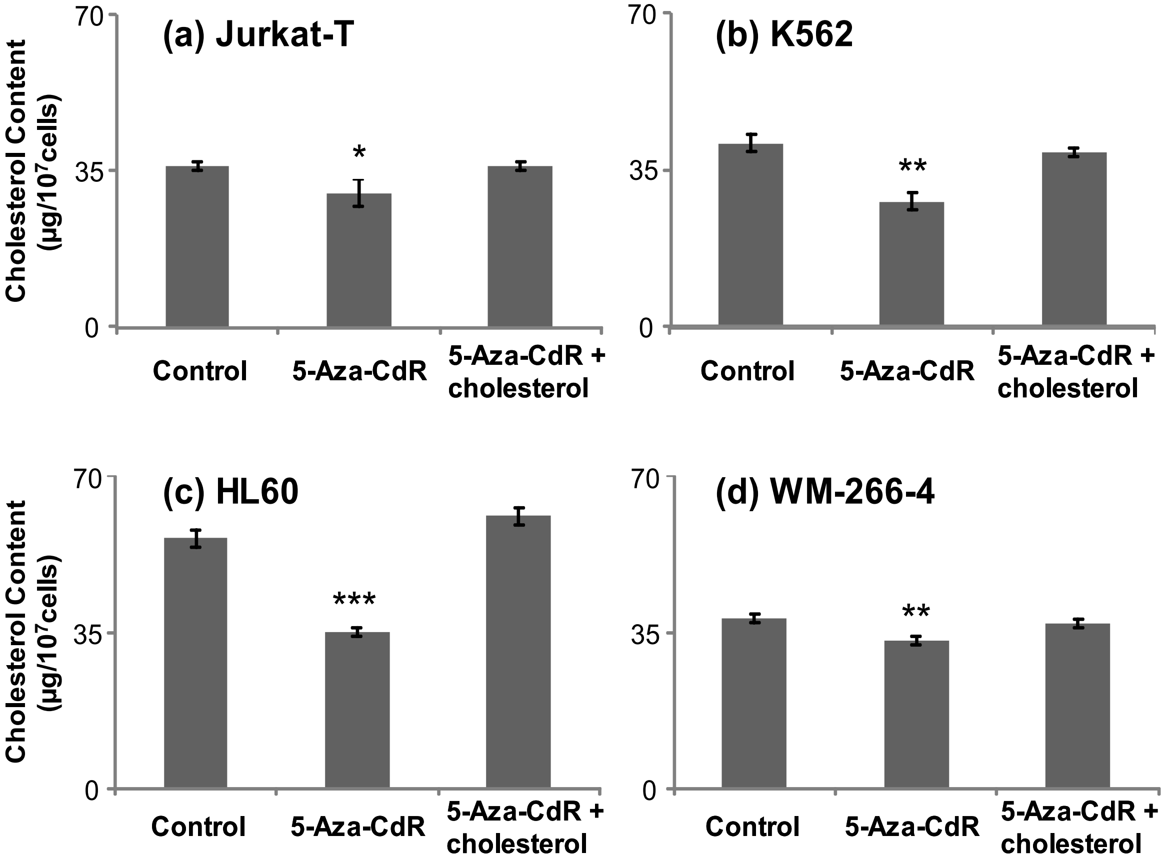


Figure 2.3. 5-Aza-CdR perturbed *de novo* cholesterol synthesis in leukemia cells. Shown are the histograms of cholesterol levels in Jurkat-T (a), K562 (b), HL60 (c) and WM-266-4 (d) cells that are untreated, treated with 5 μM 5-Aza-CdR treatment for 24 hrs alone or together with cholesterol. The values represent mean ± S.D. of results obtained from three independent experiments. '*', $p < 0.05$; '**', $p < 0.01$; '**', $p < 0.01$. The p -values were calculated by using unpaired two-tailed t-test.

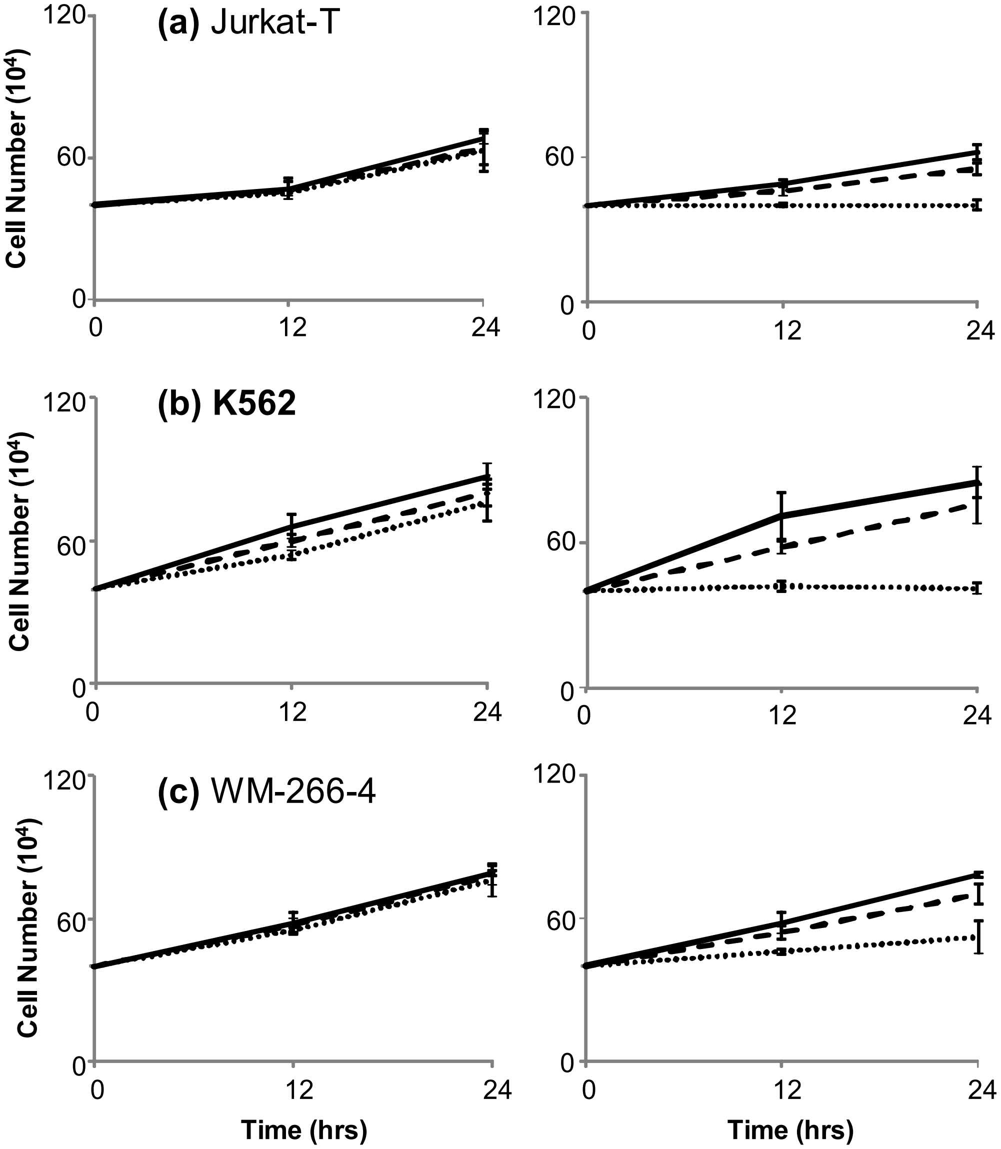


Figure 2.4. 5-Aza-CdR-induced growth inhibition of leukemia cells can be rescued by externally added cholesterol. The viability of Jurkat-T (a), K562 (b) and WM-266-4 (c) cells after 12 and 24 hrs of treatment with 0 (dotted line), 30 (dash line) or 60 mg/L (solid line) cholesterol alone (left), or together with 5 µM 5-Aza-CdR (right).

5-Aza-CdR-induced inhibition of cholesterol biosynthesis involves epigenetic reactivation of *DPP4* gene

Viewing that 5-Aza-CdR treatment can lead to cytosine demethylation and reactivation of epigenetically silenced genes, we surveyed literature for mammalian genes that are both epigenetically silenced in leukemia cells and involved in cholesterol biosynthesis. We found that dipeptidyl peptidase 4 (DPP4, a.k.a. CD26 in T cells or adenosine deaminase complexing protein) is the only known human gene satisfying both criteria. In this regard, *DPP4* was found to be epigenetically silenced in primary leukocytes isolated from adult T-lineage leukemic patients and in cultured human leukemia cells (27). In addition, promoter cytosine methylation of *DPP4* gene was observed to correlate positively with plasma HDL-cholesterol level in severely obese patients with metabolic syndrome (28). Although our proteomic experiment did not allow us to identify or quantify the DPP4 protein (likely due to its low level of expression in Jurkat-T cells), real-time PCR experiment showed that 5-Aza-CdR treatment led to the reactivation of *DPP4* gene in all three leukemia cell lines tested (Figure A2).

Apart from leukemia cells, *DPP4* is known to be epigenetically silenced in human melanoma cells (29). We found that 5-Aza-CdR treatment also resulted in elevated expression of *DPP4* (Figure A2) and diminished endogenous cholesterol biosynthesis (Figure 2.3d) in WM-266-4 human melanoma cells. Furthermore, the 5-Aza-CdR-induced growth inhibition of WM-266-4 cells can again be rescued by supplementing the culture medium with cholesterol (Figure 2.4c).

To examine whether epigenetic reactivation of *DPP4* gene plays a causative role in the decreased levels of FDPS and FDFT1 proteins, we employed siRNA to knockdown the expression of *DPP4* gene and assessed the expression levels of the two cholesterol biosynthesis enzymes as well as cellular cholesterol content. Owing to the difficulty in transfection of leukemia cells, we employed WM-266-4 cells for the experiment. Our results showed that the treatment of WM-266-4 cells with DPP4 siRNA led to a decrease in DPP4 mRNA expression by approximately 75%, as revealed by real-time PCR analysis (Figure 2.5a). More importantly, treatment of WM-266-4 cells with DPP4 siRNA, but not control non-targeting siRNA, could result in increased expression of FDPS and FDFT1 (Western blot results shown in Figure 2.5c&d) and elevated cellular cholesterol content (Figure 2.5b). This result demonstrated that DPP4 could regulate negatively the levels of FDPS and FDFT1 proteins. Together, the above results support that 5-Aza-CdR treatment induces epigenetic reactivation of *DPP4* gene, which leads to diminished expression of FDPS and FDFT1 proteins, thereby decreasing cholesterol biosynthesis and inhibiting the growth of leukemic cells (Figure 2.6).

CONCLUSIONS

Multiple epigenetic alterations have been observed in hematopoietic malignancies (30), and 5-Aza-CdR is an epigenetic drug used for the treatment of myelodysplastic syndromes and myelogenous leukemia (7, 8).

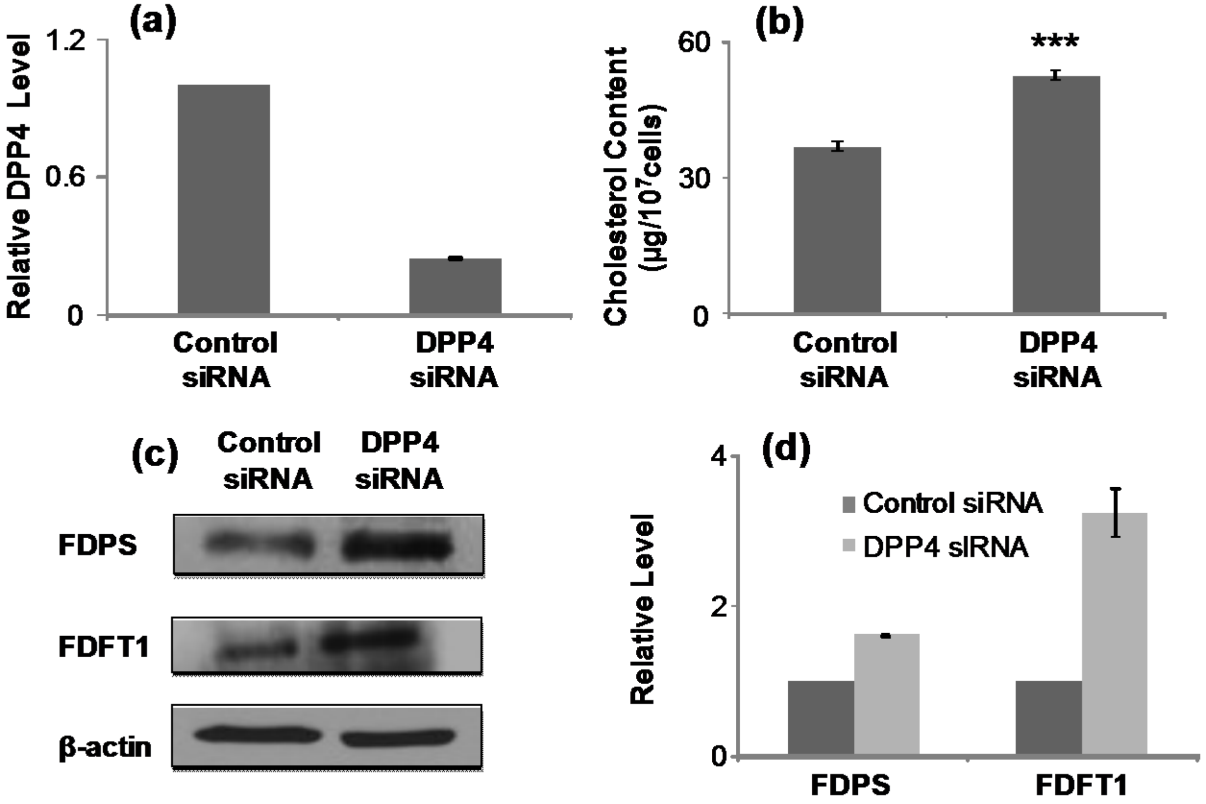


Figure 2.5. *DPP4* gene regulates negatively endogenous cholesterol biosynthesis in WM-266-4 cells. Displayed are histograms of *DPP4* expression level (a) and cholesterol content (b) in WM-266-4 cells after siRNA knockdown of *DPP4* gene. Shown in (c) and (d) are the Western blot image and quantification results for FDPS and FDFT1 protein levels in WM-266-4 cells after DPP4 siRNA knockdown.

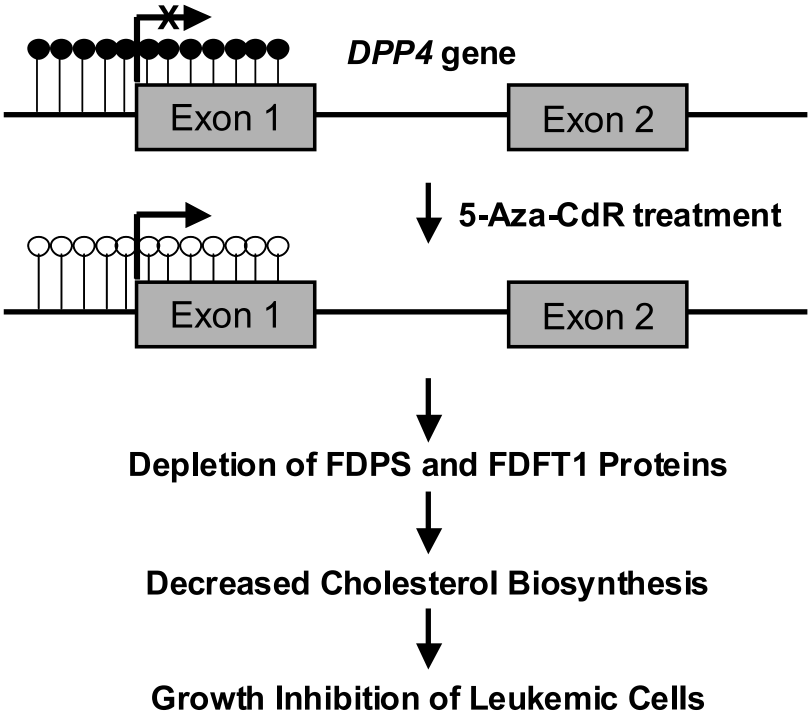


Figure 2.6. A mechanistic model underlying the anti-leukemic effect of 5-Aza-CdR.

Although the covalent binding of DNA 5-azacytosine with catalytic cysteine residue in DNMT proteins and the resulting diminished promoter cytosine methylation are considered the principal mechanism of toxicity (5), the target gene(s) underlying the therapeutic effect of this DNA hypomethylating agent remains ambiguous. We attempted to address this question by assessing the drug-induced alterations in protein expression at the whole proteome scale. Our results based on metabolic labeling using SILAC, together with LC-MS/MS analysis, revealed that the drug treatment led to significant changes in expression of 188 proteins. Among them, FDPS and FDFT1, two important enzymes involved in cholesterol biosynthesis, were decreased significantly upon 5-Aza-CdR treatment. In addition, we observed that 5-Aza-CdR-induced growth inhibition of Jurkat-T, HL60 and K562 cells could be abrogated by externally added cholesterol, supporting the conclusion that inhibition in endogenous cholesterol biosynthesis constitutes the major pathway leading to the growth inhibition of leukemia cells. The above observations, in conjunction with previous findings showing that *DPP4* gene is epigenetically silenced in leukemia cells (29) and promoter methylation of this gene is correlated positively with HDL-cholesterol level in severely obese patients with metabolic syndrome (28), led us to discover that 5-Aza-CdR-mediated epigenetic reactivation of *DPP4* gene accounts for the diminished expression of FDPS and FDFT1 at the protein level. Taken together, our quantitative proteomic experiment, together with follow-up studies, demonstrated that 5-Aza-CdR led to epigenetic reactivation of *DPP4* gene.

Elevated expression of *DPP4* gene gave rise to decreased levels of two important enzymes involved in cholesterol biosynthesis, and the resultant diminished cholesterol biosynthesis induced growth inhibition of leukemic cells (Figure 2.6). Future experiments will be needed for understanding how reactivation of *DPP4* gene leads to diminished levels of FDPS and FDFT1 proteins. In this vein, it is unlikely for these two enzymes to be direct substrates for DPP4 considering that the N-termini of neither proteins carry a DPP4 recognition site and DPP4 was found to be incapable of cleaving large protein molecules (31).

It is worth emphasizing the advantage of the quantitative proteomic method for discovering molecular mechanisms of action of anti-cancer drugs. The most notable outcome of the unbiased quantitative proteomic approach is perhaps the generation of novel hypotheses. Prior to this work, few would have thought that epigenetic reactivation of *DPP4* gene and the resulting inhibition in cholesterol biosynthesis would constitute an important molecular mechanism contributing to the anti-leukemic effect of 5-Aza-CdR. Now, this becomes evident on the basis of the validation experiments. The ability to discover molecular target of 5-Aza-CdR that lies outside of previous biological knowledge is a major motivation for the use of unbiased proteome-wide approaches for unraveling novel mechanisms of action of anti-neoplastic agents and is well supported by the results generated from the present study. It is also worth noting that 5-Aza-CdR treatment did not give rise to significant decreases in expression of FDPS and FDFT1 at the mRNA level (Figure 2.6); likewise, treatment with DPP4 siRNA did not lead to elevated mRNA expression of these two genes in WM-266-4 cells (Figure A3), revealing the lack of transcriptional regulation of these two genes by DPP4. Therefore, this result also underscored that the new pathway uncovered from quantitative proteomic strategy would not be discovered with the conventional microarray-based approach.

In conclusion, the current study uncovered a novel mechanism of 5-Aza-CdR-induced anticancer effect. In the future, it will be important to assess whether the findings made with cultured leukemia cells can be extended to leukemic patients administered with the drug. If the mechanism can be extended to *in vivo*, it will have a profound impact on the clinical use of 5-Aza-CdR in treating myeloid disorders. For instance, assessment of *DPP4* gene expression in patient leukocytes may provide an important basis for choosing the optimal therapeutic dose for the treatment, and leukocyte cholesterol level may serve as a biomarker for monitoring the clinical efficacy of the drug.

REFERENCES

1. Holliday, R. (1987) The inheritance of epigenetic defects. *Science* 238, 163-170.

2. Jones, P. A., and Baylin, S. B. (2002) The fundamental role of epigenetic events in cancer. *Nat. Rev. Genet.* 3, 415-428.

3. Yoo, C. B., and Jones, P. A. (2006) Epigenetic therapy of cancer: past, present and future. *Nat. Rev. Drug Discov.* 5, 37-50.

4. Jones, P. A., and Taylor, S. M. (1980) Cellular-differentiation, cytidine analogs and DNA methylation. *Cell* 20, 85-93.

5. Christman, J. K. (2002) 5-Azacytidine and 5-aza-2'-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy. *Oncogene* 21, 5483-5495.

6. Ghoshal, K., Datta, J., Majumder, S., Bai, S., Kutay, H., Motiwala, T., and Jacob, S. T. (2005) 5-Aza-deoxycytidine induces selective degradation of DNA methyltransferase 1 by a proteasomal pathway that requires the KEN box, bromo-adjacent homology domain, and nuclear localization signal. *Mol. Cell. Biol.* 25, 4727-4741.

7. Kaminskas, E., Farrell, A., Abraham, S., Baird, A., Hsieh, L. S., Lee, S. L., Leighton, J. K., Patel, H., Rahman, A., Sridhara, R., Wang, Y. C., and Pazdur, R. (2005) Approval summary: azacitidine for treatment of myelodysplastic syndrome subtypes. *Clin Cancer Res.* 11, 3604-3608.

8. Plimack, E. R., Kantarjian, H. M., and Issa, J. P. (2007) Decitabine and its role in the treatment of hematopoietic malignancies. *Leuk Lymphoma* 48, 1472-1481.

9. Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* 1, 376-386.

10. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26, 1367-1372.

11. Kersey, P. J., Duarte, J., Williams, A., Karavidopoulou, Y., Birney, E., and Apweiler, R. (2004) The International Protein Index: an integrated database for proteomics experiments. *Proteomics* 4, 1985-1988.

12. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol* 26, 1367-1372.

13. Martinez, F., Eschegoyen, S., Briones, R., and Cuellar, A. (1988) Cholesterol Increase in mitochondria - a new method of cholesterol incorporation. *J. Lipid. Res.* 29, 1005-1011.

14. Madden, E. A., Bishop, E. J., Fiskin, A. M., and Melnykovych, G. (1986) Possible role of cholesterol in the susceptibility of a human acute lymphoblastic leukemia cell line to dexamethasone. *Cancer Res.* 46, 617-622.

15. Dong, X., Xiong, L., Jiang, X., and Wang, Y. (2010) Quantitative proteomic analysis reveals the perturbation of multiple cellular pathways in Jurkat-T cells induced by doxorubicin. *J. Proteome Res.* 9, 5943-5951.

16. Livak Kj Fau - Schmittgen, T. D., and Schmittgen, T. D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-DDCt} method. *Methods* 25, 402-408.

17. Nilsson, C. L., Dillon, R., Devakumar, A., Shi, S. D. H., Greig, M., Rogers, J. C., Krastins, B., Rosenblatt, M., Kilmer, G., Major, M., Kaboord, B. J., Sarracino, D., Rezai, T., Prakash, A., Lopez, M., Ji, Y. J., Priebe, W., Lang, F. F., Colman, H., and Conrad, C. A. (2010) Quantitative Phosphoproteomic Analysis of the STAT3/IL-6/HIF1 alpha Signaling Network: An Initial Study in GSC11 Glioblastoma Stem Cells. *J. Proteome Res.* 9, 430-443.

18. Goldstein, J. L., and Brown, M. S. (1990) Regulation of the mevalonate pathway. *Nature* 343, 425-430.

19. Takenawa, T., and Suetsugu, S. (2007) The WASP-WAVE protein network: connecting the membrane to the cytoskeleton. *Nat. Rev. Mol. Cell Biol.* 8, 37-48.

20. Bernheim-Groswasser, A., Wiesner, S., Golsteyn, R. M., Carlier, M. F., and Sykes, C. (2002) The dynamics of actin-based motility depend on surface parameters. *Nature* 417, 308-311.

21. Lin, Y. H., Park, Z. Y., Lin, D., Brahmbhatt, A. A., Rio, M. C., Yates, J. R., 3rd, and Klemke, R. L. (2004) Regulation of cell migration and survival by focal adhesion targeting of Lasp-1. *J. Cell Biol.* 165, 421-432.

22. Dufour, S., Ramprasad, O. G., Srinivas, G., Rao, K. S., Joshi, P., Thiery, J. P., and Pande, G. (2007) Changes in cholesterol levels in the plasma membrane modulate cell

signaling and regulate cell adhesion and migration on fibronectin. *Cell. Motil. Cytoskel* 64, 199-216.

23. Sun, M., Northup, N., Marga, F., Huber, T., Byfield, F. J., Levitan, I., and Forgacs, G. (2007) The effect of cellular cholesterol on membrane-cytoskeleton adhesion. *J. Cell. Sci.* 120, 2223-2231.

24. Vitols, S., Norgren, S., Juliusson, G., Tatidis, L., and Luthman, H. (1994) Multilevel regulation of low-density lipoprotein receptor and 3-hydroxy-3-methylglutaryl coenzyme A reductase gene expression in normal and leukemic cells. *Blood* 84, 2689-2698.

25. Dimitroulakos, J., Nohynek, D., Backway, K. L., Hedley, D. W., Yeger, H., Freedman, M. H., Minden, M. D., and Penn, L. Z. (1999) Increased sensitivity of acute myeloid leukemias to lovastatin-induced apoptosis: A potential therapeutic approach. *Blood* 93, 1308-1318.

26. Li, H. Y., Appelbaum, F. R., Willman, C. L., Zager, R. A., and Banker, D. E. (2003) Cholesterol-modulating agents kill acute myeloid leukemia cells and sensitize them to therapeutics by blocking adaptive cholesterol responses. *Blood* 101, 3628-3634.

27. Tsuji, T., Sugahara, K., Tsuruda, K., Uemura, A., Harasawa, H., Hasegawa, H., Hamaguchi, Y., Tomonaga, M., Yamada, Y., and Kamihira, S. (2004) Clinical and oncologic implications in epigenetic down-regulation of CD26/dipeptidyl peptidase IV in adult T-cell leukemia cells. *Int. J. Hematol.* 80, 254-260.

28. Turcot, V., Bouchard, L., Faucher, G., Tchernof, A., Deshaies, Y., Perusse, L., Belisle, A., Marceau, S., Biron, S., Lescelleur, O., Biertho, L., and Vohl, M. C. (2011) DPP4 gene DNA methylation in the omentum is associated with its gene expression and plasma lipid profile in severe obesity. *Obesity* 19, 388-395.

29. Wesley, U. V., Albino, A. P., Tiwari, S., and Houghton, A. N. (1999) A role for dipeptidyl peptidase IV in suppressing the malignant phenotype of melanocytic cells. *J. Exp. Med.* 190, 311-322.

30. Claus, R., and Lubbert, M. (2003) Epigenetic targets in hematopoietic malignancies. *Oncogene* 22, 6489-6496.

31. De Meester, I., Korom, S., Van Damme, J., and Scharpe, S. (1999) CD26, let it cut or cut it down. *Immunol. Today* 20, 367-375.

CHAPTER 3

6-Thioguanine Induces Mitochondrial Dysfunction and Oxidative DNA Damage in Acute Lymphoblastic Leukemia Cells

INTRODUCTION

Thiopurines, including azathioprine, 6-mercaptopurine (6-MP) and 6-thioguanine (^{S}G), are widely used as cancer chemotherapeutic and immunosuppressive agents (1-3), where 6-MP and ^{S}G are commonly prescribed for the treatment of acute lymphoblastic leukemia (ALL) (3). A previous study showed that thiopurines may exert their cytotoxic effects via their metabolic activation and incorporation of ^{S}G into DNA, spontaneous methylation of DNA ^{S}G by *S*-adenosyl-L-methionine to yield S^6-methylthioguanine (S^6mG), misincorporation of thymine (T) opposite S^6mG during DNA replication, and triggering of post-replicative mismatch repair (MMR) by the ensuing S^6mG:T mispair (4).

Recent studies suggested that the activation of MMR might not be the sole pathway contributing to the cytotoxic effects of the drug. In this vein, very low levels (<0.02%) of DNA ^{S}G was found to be methylated to S^6mG in ^{S}G-treated leukemia cells (5) and ^{S}G itself could direct significant frequencies (*circa* 10%) of misincorporation of thymine during DNA replication in both *E. coli* and human cells (6, 7). These findings, in conjunction with the observation that the ^{S}G:T mispair could be recognized more readily by human MMR proteins than the S^6mG:T mispair (8, 9), suggest that DNA ^{S}G itself may trigger the MMR pathway without being converted to S^6mG.

Additionally, leukemic cells deficient in MMR are also sensitive to thiopurine drugs (10), indicating that thiopurines may also induce their cytotoxic effects through other pathways.

Global transcriptome and proteome analyses are very useful for revealing novel mechanisms of action of anti-cancer drugs (11). Along this line, global gene expression analysis with microarray was used for assessing transcriptional responses of cancer cells toward anticancer agents including thiopurines (12). However, results from mRNA-based gene expression analysis often exhibit poor correlation with protein expression data from proteomic measurements (13). This may be attributed to the fact that regulatory processes following mRNA production (i.e., posttranscriptional, translational, and protein degradation regulation) play substantial roles in controlling steady-state protein abundances (13). Mass spectrometry (MS)-based proteomics techniques, along with two-dimensional gel electrophoresis or stable-isotope labeling, have been widely used for large-scale quantitative analysis of proteins in complex samples (14, 15). Among the many isotope-labeling methods, stable isotope labeling by amino acids in cell culture (SILAC), a metabolic-labeling technique, has the advantages of being simple, efficient and capable of quantifying relatively small changes in protein expression (16).

Here we employed LC-MS/MS, in conjunction with SILAC, to assess, at the global proteome scale, the perturbation of protein expression in Jurkat-T human ALL cells upon treatment with a clinically relevant concentration of ^{S}G. The results from the quantitative proteomic analysis enabled us to conclude that the exposure to ^{S}G resulted in mitochondrial dysfunction in ALL cells, which was accompanied with drug-induced loss of active mitochondria and elevated generation of oxidatively induced DNA lesions.

MATERIALS AND METHODS

Cell culture

Jurkat-T, CEM and HEK293T cells were purchased from ATCC (Manassas, VA). COS7 cells were kindly provided by Prof. F. M. Sladek (University of California Riverside). Jurkat-T and CEM cells were cultured in RPMI-1640 medium (ATCC, Manassas, VA) supplemented with 10% fetal bovine serum (FBS, Invitrogen, Carlsbad, CA), 100 IU/mL penicillin. HEK293T and COS7 cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM, ATCC, Manassas, VA) under the same conditions without the addition of penicillin. Cells were maintained in a humidified atmosphere with 5% CO_2 at 37°C, with medium renewal of 2-3 times a week depending on cell density. For SILAC experiments, the RPMI-1640 medium without L-lysine or L-arginine was purchased from Cambridge Isotope Laboratories, Inc. (Andover, MA). The complete light and heavy RPMI-1640 media were prepared by the addition of light or heavy lysine ($[^{13}C_6, {}^{15}N_2]$-L-lysine) and arginine ($[^{13}C_6]$-L-arginine), along with dialyzed fetal bovine serum (FBS, Invitrogen, Carlsbad, CA) to the above lysine, arginine-depleted medium. We chose to label the proteome with lysine and arginine because all tryptic peptides except for the C-terminal peptides of some proteins carry a lysine or arginine, which provides better proteome coverage than labeling the proteome using other amino acids like leucine. The Jurkat-T cells were cultured in heavy RPMI-1640 medium for at least 10 days to achieve complete stable isotope incorporation, and Figure B1 displayed representative mass spectra illustrating the complete incorporation of the heavy labeled lysine and arginine.

^{S}G treatment and sample preparation

Jurkat-T cells, at a density of ~ 7×10^{5} cells/mL in light or heavy RPMI-1640 medium, were treated with 3 μM ^{S}G (Sigma, St. Louis, MO) for 24 hrs. In this vein, the mean peak concentrations of ^{S}G in plasma of ALL patients were (0.46 ± 0.68) and (2.7 ± 1.4) μM after oral ^{S}G administration at 60 mg/m^{2} and after 24-hrs of continuous intravenous infusion at 20 mg/m^{2}/hr, respectively (17). After treatment, the cells were harvested by centrifugation at 300 g at 4°C for 5 min and washed three times with ice-cold PBS to remove culture medium and FBS. To explore the role of reactive oxygen species in ^{S}G cytotoxicity, cells were also treated with 3 μM ^{S}G together with 500 μM allopurinol for 24 hrs. Cells were subsequently lysed with CelLytic^{TM} M lysis buffer (Sigma) that was supplemented with 1 mM PMSF and a protease inhibitor cocktail (Sigma). Cell lysates were centrifuged at 16,000 g at 4°C for 30 min, and the resulting supernatants were collected. The protein concentration in the cell lysate was measured using Quick Start Bradford Protein Assay (Bio-Rad, Hercules, CA). SILAC labeling experiments were performed in three biological replicates: In the two forward SILAC experiments, the light labeled, drug-treated lysate and the heavy labeled, control lysate were combined at 1:1 ratio (w/w) (Figure 3.1a); in the reverse SILAC experiment, the heavy labeled, drug-treated cell lysate was mixed equally with the light labeled, control lysate.

SDS-PAGE separation and in-gel digestion

The above equi-mass mixture of light and heavy lysates was separated using 12% SDS-PAGE with a 4% stacking gel, and the gel was stained with Coomassie blue. The gel was cut into 20 bands, and the proteins were reduced in-gel with dithiothreitol, alkylated with iodoacetamide, and digested with trypsin (Promega, Madison, WI) at 37°C overnight. Following the digestion, peptides were extracted into 5% acetic acid in H_2O and CH_3CN/H_2O (1:1, v/v) and dried in a Speed-vac. The resulting peptide mixtures were stored at –80°C until further analysis.

LC-MS/MS for protein identification and quantification

On-line LC-MS/MS analysis was performed on an LTQ-Orbitrap Velos mass spectrometer coupled with an EASY n-LCII HPLC system and a nanoelectrospray ionization source (Thermo, San Jose, CA). The sample injection, enrichment, desalting, and HPLC separation were conducted automatically on a homemade trapping column (150 μm×50 mm) and a separation column (75 μm×120 mm, packed with ReproSil-Pur C18-AQ resin, 5 μm in particle size and 300 Å in pore size, Dr. Maisch HPLC GmbH, Germany). The peptide mixture was first loaded onto the trapping column with a solvent mixture of 0.1% formic acid in CH_3CN/H_2O (2:98, v/v) at a flow rate of 3.0 μL/min. The peptides were then separated using a 120-min linear gradient of 2-40% acetonitrile in 0.1% formic acid at a flow rate of 300 nL/min.

The LTQ-Orbitrap Velos mass spectrometer was operated in the positive-ion mode, and the spray voltage was 1.8 kV. All MS/MS spectra were acquired in a data-dependent scan mode, where the 20 most abundant ions found in MS at a threshold above 500 counts were selected for fragmentation by collision-induced dissociation at a normalized collision energy of 35%. The full-scan mass spectra (from *m/z* 350 to 2000) were acquired with a resolution of 60,000 at *m/z* 400 after accumulation to a target value of 500,000.

Data processing

The LC-MS/MS data were employed for the identification and quantification of the global proteome, which were conducted using Maxquant, Version 1.2.0.18 (18) against the UniProt human database (with 538,585 sequence entries, release date: 2012.11.28) to which contaminants and reverse sequences were added. The maximum number of miss-cleavages for trypsin was two per peptide. Cysteine carbamidomethylation and methionine oxidation were set as fixed and variable modifications, respectively. The tolerances in mass accuracy for MS and MS/MS were 25 ppm and 0.6 Da, respectively. Only those proteins with at least two distinct peptides discovered from LC-MS/MS analyses were considered reliably identified. The required false positive discovery rate was set to be within 1% at both peptide and protein levels, with the minimal required peptide length being 6 amino acids. The quantification was based on three independent SILAC and LC-MS/MS experiments as noted above. To establish the threshold ratios for determining the significantly changed proteins, we calculated the standard scores (z-scores) for protein ratios as previously described (19, 20). Our analysis showed that, at 95% confidence level, those proteins with expression ratios that are greater than 1.22 or less than 0.73 could be considered significantly changed (21). To be stringent, we, however, only considered a protein to be significantly altered if its differential expression ratio (^{S}G treated/control) was greater than 1.5 or less than 0.67.

Ingenuity Pathway Analysis (IPA)

IPA (version 7.6, Ingenuity Systems Inc.) was employed to derive information of relationships, functions and pathways of the differentially regulated proteins with the use of the Ingenuity Knowledge Base (IPKB, Ingenuity Systems Inc.) (22, 23). The IPKB database consisted of gene regulatory and signaling pathways, which are integrated with other relevant databases such as NCBI Gene (http://www.ncbi.nlm.nih.gov/gene) and Gene Ontology (http://www.geneontology.org). The proteins that were significantly changed upon ^{S}G treatment, including those quantified only in one cycle of SILAC labeling experiment, were included for the pathway analysis. IPA determines the interaction by calculating a significance score using Fisher's exact test and exporting it as a p-value. For canonical pathways, a p-value of <0.05 is considered significant.

Plasmid transfection

HEK293T cells were seeded in 6-well plates at 70% confluence level ($\sim 3\times10^5$ cells/well), and ^{S}G was added to the culture medium at a final concentration of 3 μM. After a 24-hr incubation, the cells were transfected with 1.5 μg Myc-Miro1 expression plasmid, which was kindly provided by Prof. E. Soriano (24), using Lipofectamine 2000 (Invitrogen). The cells were subsequently harvested for Western analysis at 18-48 hrs after the transfection.

Western blot and flow cytometry

Mouse monoclonal anti-Myc (9E10, 1:5000 dilution, Santa Cruz Biotechnology) and mouse monoclonal anti-RHOT1 (1:10000 dilution, Sigma) were used for Western analysis of ectopically expressed and endogenous Miro1, respectively. MitoProfile Total OXPHOS Human WB Antibody Cocktail (1:10,000 dilution, ab110411, Abcam) was employed to assess the expression level of five mitochondrial respiratory chain proteins.

Cellular uptake of MitoTracker Deep Red FM (Invitrogen) was quantified using flow cytometry on a BD FACS Aria I instrument (BD Biosciences, San Jose, CA). CEM, Jurkat-T, and HEK293T cells, untreated or after a 24 hrs treatment with 3 μM ^{S}G and/or 500 μM allopurinol, were stained with MitoTracker Deep Red at a final concentration of 200 nM. After incubation for 15-20 min, the cells were washed thoroughly with pre-warmed media and PBS. Flow cytometry analysis (BD FACS Aria I, NJ) was performed immediately after resuspending cells in a sorting buffer containing 1× PBS, 25 mM HEPES, 1 mM EDTA, and 1% FBS (pH 7.0). Counting was based on 20,000 cells per sample.

Trypan blue exclusion assay

CEM, Jurkat-T and HEK293T cells were seeded in 6-well plates at a density of approximately 4×10^{5} cells/mL and treated with ^{S}G and/or allopurinol at final concentrations of 3 μM and 500 μM, respectively. At 24 or 48 hrs after treatment, the cells were stained with trypan blue, and counted on a hemocytometer to measure cell viability.

DNA extraction and LC-MS/MS measurement of oxidatively induced DNA lesions

DNA was isolated from control and ^{S}G-treated cells using a salt extraction method and digested to nucleosides with a cocktail of four enzymes (25). To the resulting mixture were subsequently added uniformly ^{15}N-labeled 8,5'-cyclo-2'-deoxyadenosine (cdA) and 8,5'-cyclo-2'-deoxyguanosine (cdG), and the cdA and cdG were enriched from the nucleoside mixtures using HPLC, as described previously (25). The enriched fractions were subjected to LC-MS/MS analysis on a TSQ Vantage triple quadrupole mass spectrometer (Thermo Fisher Scientific) operated in the multiple-reaction monitoring (MRM) mode. A 3.0×100 mm Hypersil GOLD column (particle size, 5 μm, Thermo Scientific) was used for the on-line analyses of the fractions containing cdA and cdG, and the flow rate was 100 μL/min. A solution of 0.1% (v/v) formic acid in water (solution A) and a solution of 0.1% (v/v) formic acid in methanol (solution B) were employed as mobile phases, and a gradient of 40 min 0-30% B followed by 2 min 30-80% B was used for the separation. The spray voltage and S-lens RF amplitude were 4.0 kV and 55 V, respectively, and the temperatures for the ion transport tube and vaporizer were set at 270°C and 226°C, respectively. The sheath and auxiliary gas flow rates were 15 and 10 arbitrary units, respectively. The respective MRM transitions for unlabeled and uniformly ^{15}N-labeled cdG were m/z 266→180 and m/z 271→185. The corresponding MRM transitions for unlabeled and uniformly ^{15}N-labeled cdA were m/z 250→164 and m/z 255→169. The collision energy was set at 17 V for cdG and 19 V for cdA, and the scan time was 300 ms.

RESULTS AND DISCUSSION

6-MP and ^{S}G are widely used for treating ALL and other human diseases (1-3). Although metabolic transformations of these prodrugs to ^{S}G nucleotides, their incorporation into nucleic acids, and triggering of the mismatch repair pathway are considered important (4), other mechanisms may also be at work. To exploit novel cellular pathways underlying the anti-neoplastic effects of the thiopurine drugs, we conducted a quantitative proteomic experiment, based on metabolic labeling using SILAC and LC-MS/MS, to exploit the ^{S}G-induced differential protein expression in human ALL cells.

^{S}G treatment, protein identification and quantification

We employed Jurkat-T human ALL cells for the quantitative proteomic experiment. To minimize systematic errors and to assess reliably the drug-induced alterations in protein expression, we performed the SILAC experiments in three biological replicates, including two forward and one reverse SILAC labelings (Figure 3.1a). The results from LC-MS/MS analyses of the three sets of SILAC samples facilitated us to quantify a total of 4049 proteins. While the majority of the quantified proteins displayed similar levels of expression in the ^{S}G-treated and control cells, 230 exhibited significant changes (by > 1.5-fold, see Materials and Methods) in expression levels after the drug treatment, with 166 and 64 being significantly down- and up-regulated, respectively (Figure 3.1b). Among these 230 proteins, 100 were quantified in both forward and reverse SILAC labeling experiments (Table B1). Figure 3.2 displays the representative LC-MS/MS results supporting the drug-induced down-regulation of Miro1, as reflected by lower signal intensity for the tryptic peptide, LPLILVGNK, in the protein lysate from the ^{S}G-treated cells.

^{S}G treatment led to mitochondrial dysfunction

To explore the cellular pathways that are altered by ^{S}G treatment, we next subjected the significantly changed proteins to Ingenuity Pathway Analysis (IPA). The IPA results revealed that the drug treatment led to the alteration of a number of canonical pathways (Table B2). Among them, mitochondrial dysfunction is the most significant, as manifested by the smallest p value ($-\log p = 7.41$). In this context, most down-regulated proteins in this pathway are part of the mitochondrial respiratory chain complex, where NDUFA2, NDUFA4, NDUFA10, and NDUFV1 are constituents of Complex I, UQCRC1, UQCRC2 and UQCRFS1 are components of Complex III, and MT-CO2, COX5B and COX6C belong to Complex IV (Table B1). Additionally, we observed the diminished expression of cytochrome b5 (Table 3.1), an outer mitochondrial membrane-bound hemoprotein functioning as an electron carrier for several membrane-bound oxygenases (26). Thus, treatment of Jurkat-T cells with ^{S}G led to the reduced expression of a large number of proteins involved in electron transport chain in mitochondria. It is important to emphasize that diminished expression is not a general phenomenon for mitochondrial proteins. For instance, the expression levels of mitochondrial ATP synthase α subunit (ATP5A), iron-sulfur subunit of complex II (SDH), and mitochondrial DNA-directed

Table 3.1. ^{S}G-induced alteration of expression of proteins indicating mitochondrial dysfunction. Fold changes represent mean ± S.D. of quantification results from multiple SILAC labeling experiments. “F1” and “F2” refer to the results obtained from two forward cycles of SILAC labeling experiments, whereas “R1” refers to the results from the reverse SILAC labeling experiment. Data without S.D. reflect that the proteins were quantified only in one set of SILAC labeling experiment, but at least 2 peptides were identified and quantified.

Gene Symbol	Gene Name	F 1	F 2	R 1	Fold Change (^{S}Gtreated/Control, Mean ± S.D.)
MT-CO2	Cytochrome c oxidase Subunit II	0.56		0.46	0.51 ± 0.07
COX5B	Cytochrome c oxidase Subunit Vb	0.39	0.51	0.42	0.45 ± 0.08
COX6C	Cytochrome c oxidase Subunit VIc		0.57	0.51	0.54 ± 0.04
CPT1	Carnitine palmitoyltransferase 1A	2.01	2.10	1.63	1.91 ± 0.25
CYB5A	Cytochrome b5	0.65			0.65
NDUFA2	NADH dehydrogenase alpha subcomplex 2	0.63			0.63
NDUFA4	NADH dehydrogenase alpha subcomplex 4	0.64			0.64
NDUFA10	NADH dehydrogenase alpha subcomplex 10	0.58		0.61	0.60 ± 0.02
NDUFV1	NADH dehydrogenase flavoprotein 1		0.48		0.48
UQCRC1	Ubiquinol-cytochrome c reductase core protein I	0.52	0.52	0.90	0.65 ± 0.22
UQCRC2	Ubiquinol-cytochrome c reductase core protein II	0.47	0.47	0.96	0.63 ± 0.28
UQCRFS1	Ubiquinol-cytochrome c reductase rieske iron-sulfur polypeptide1	0.61			0.61
RHOT1	Mitochondrial RhoGTPase 1 (MIRO1)		0.54	0.62	0.58 ± 0.06

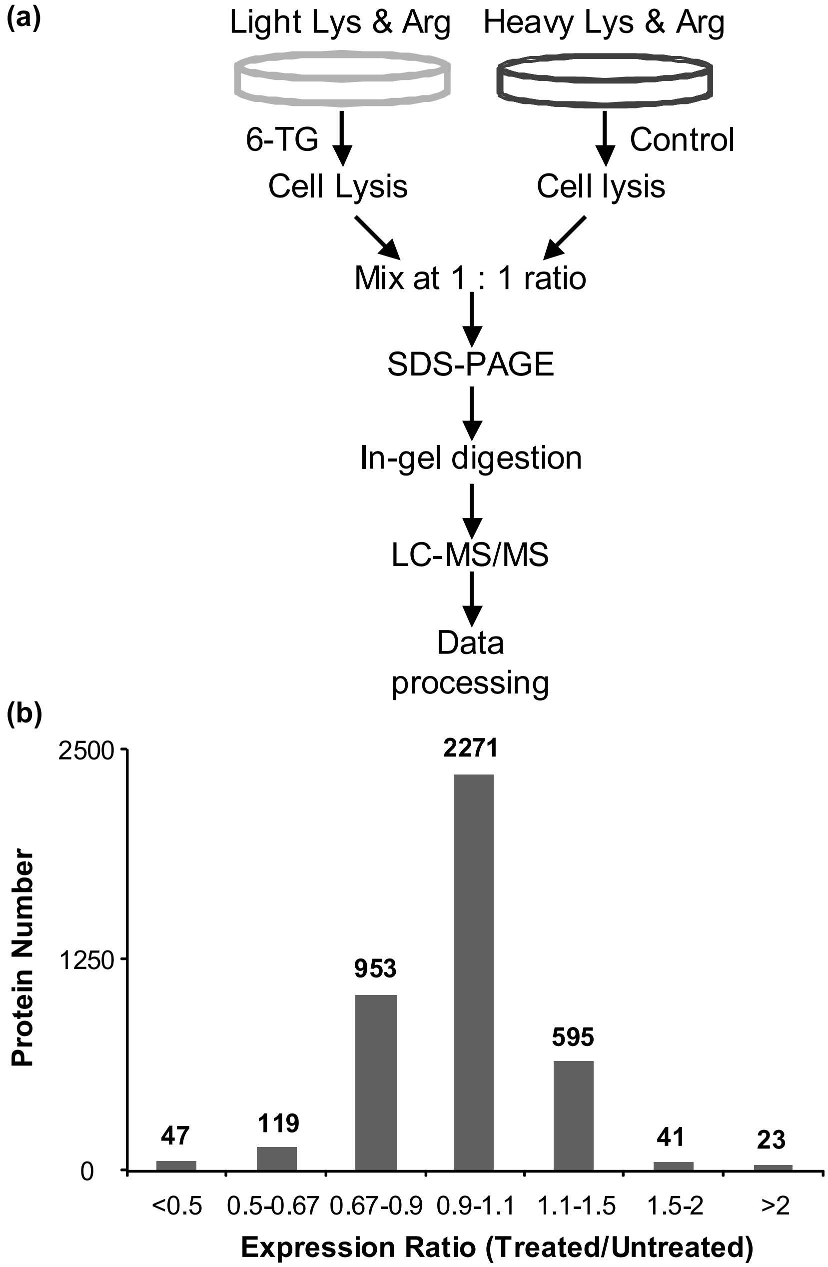


Figure 3.1. A flowchart of forward SILAC combined with LC-MS/MS for the comparative analysis of protein expression in Jurkat-T cells upon ^{S}G treatment (a). Shown in (b) is the distribution of expression ratios (treated/untreated) for the quantified proteins.

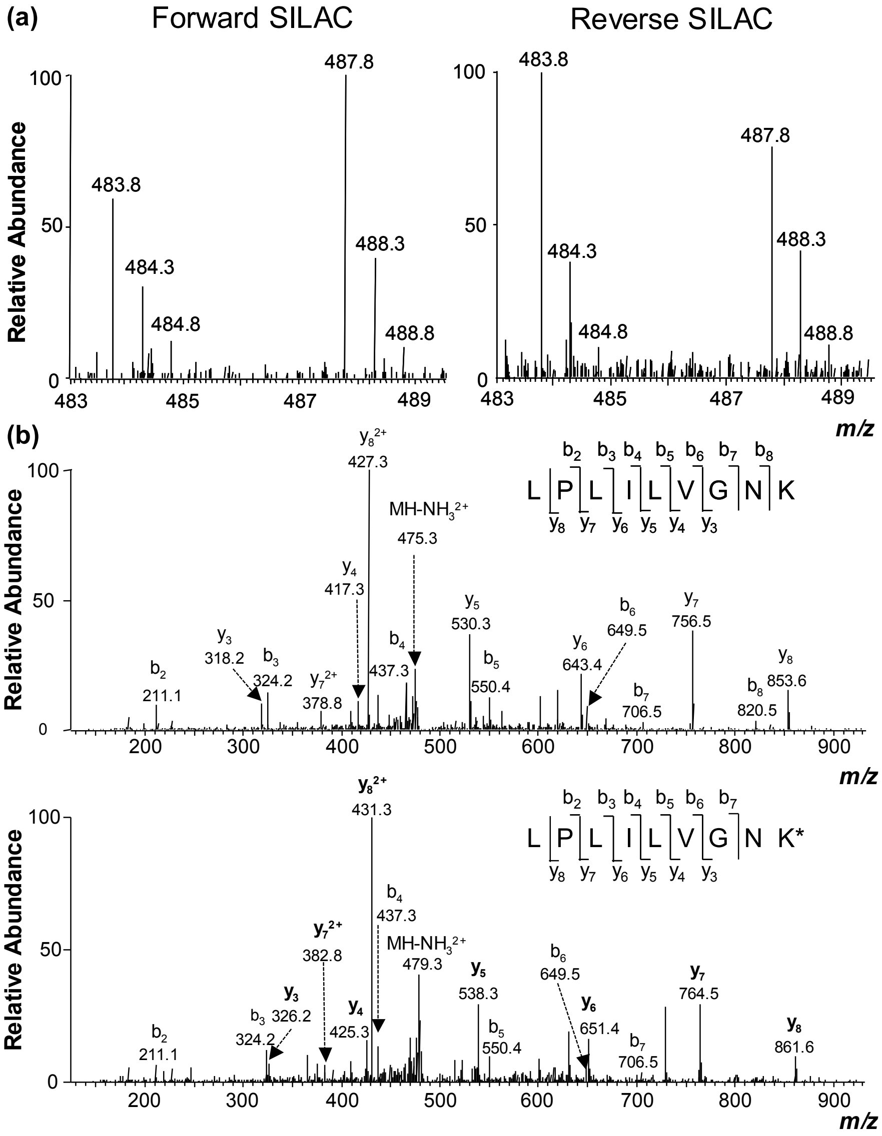


Figure 3.2. Example ESI-MS data revealing the ^{S}G-induced down-regulation of Miro1. Shown are the MS for the [M+2H]$^{2+}$ ions of Miro1 peptide LPLILVGNK and LPLILVGNK* ('K*'designates the heavy-labeled lysine) from the forward (right) and reverse (left) SILAC labeling experiments (a) and MS/MS for the [M+2H]$^{2+}$ ions of LPLILVGNK and LPLILVGNK* (b), where heavy-labeled lysine-containing y ions are labeled in bold.

RNA polymerase (POLRMT) did not change appreciably upon ^{S}G treatment.

In agreement with the LC-MS/MS quantification results, Western blot analysis revealed that the ^{S}G-induced decrease in Miro1 expression was dose-dependent and it occurred in multiple cell lines, including two ALL lines (Jurkat-T and CEM) and HEK293T human embryonic kidney epithelial cells (Figure 3.3a&b). Similar observation was made for ectopically expressed Myc-tagged Miro1 in HEK293T cells (Figure B3). Additionally, Western blot analysis with the use of a cocktail of antibodies recognizing five mitochondrial respiratory chain proteins (See Materials and Methods) revealed that the drug treatment resulted in a significant reduction in the expression level of UQCRC2, whereas no change was found for ATP5A or SDH (Figure 3.4a&b).

These results are again consistent with the LC-MS/MS quantification data. Western blot data did not reveal a significant alteration in expression of NDUFB8 or MT-CO2 (Figure 3.4a&b). The former protein was not quantified in LC-MS/MS experiments, whereas the latter displayed significant down-regulation based on both forward and reverse SILAC labeling experiments (Table 3.1). The discrepancy between Western blot and LC-MS/MS results for the expression level of MT-CO2 is likely attributed to inadequate specificity of the antibody toward this protein and/or co-migration of other protein(s) with MT-CO2 under our SDS-PAGE conditions.

The significant down-regulation of many proteins in mitochondrial respiratory chain complex prompted to ask whether ^{S}G treatment compromises mitochondrial integrity. Thus, we estimated the level of active mitochondria based on the cellular uptake of MitoTracker Deep Red. Our flow cytometry quantification results illustrated that the treatment with ^{S}G led to diminished uptake of MitoTracker Deep Red into Jurkat-T, CEM and HEK293T cells (Figure 3.5a), suggesting decreased levels of active mitochondria after ^{S}G treatment. Moreover, the diminution in the level of active mitochondria in these cell lines paralleled their relative sensitivity toward ^{S}G (Figure 3.5). In this context, we observed more pronounced loss in active mitochondria for CEM than Jurkat-T and HEK293T cells (Figure 3.5a), which mirrors the highest sensitivity of the CEM cells toward ^{S}G (Figure 3.5b-d).

^{S}G treatment led to elevated formation of oxidatively induced DNA lesions

We reason that the ^{S}G-induced down-regulation of many proteins in the mitochondrial respiratory chain complex may lead to decreased electron transport efficiency, a situation known to give rise to elevated formation of reactive oxygen species (ROS) (27). In this vein, increased ROS production was found in a series of nuclear-encoded genetic defects of Complex I in the mitochondrial respiratory chain (28). Therefore, we next asked whether the ^{S}G-induced mitochondrial dysfunction can also lead to increased formation of ROS and elevated generation of oxidatively induced DNA lesions in cells. In this respect, we found recently that the oxidatively generated 8,5'-cyclo-2'-deoxyadenosine (cdA) and 8,5'-cyclo- 2'-deoxyguanosine (cdG) can serve as robust biomarkers for oxidative stress (the mechanism for the formation of cdA is depicted in Figure 3.6a) (25, 29, 30).

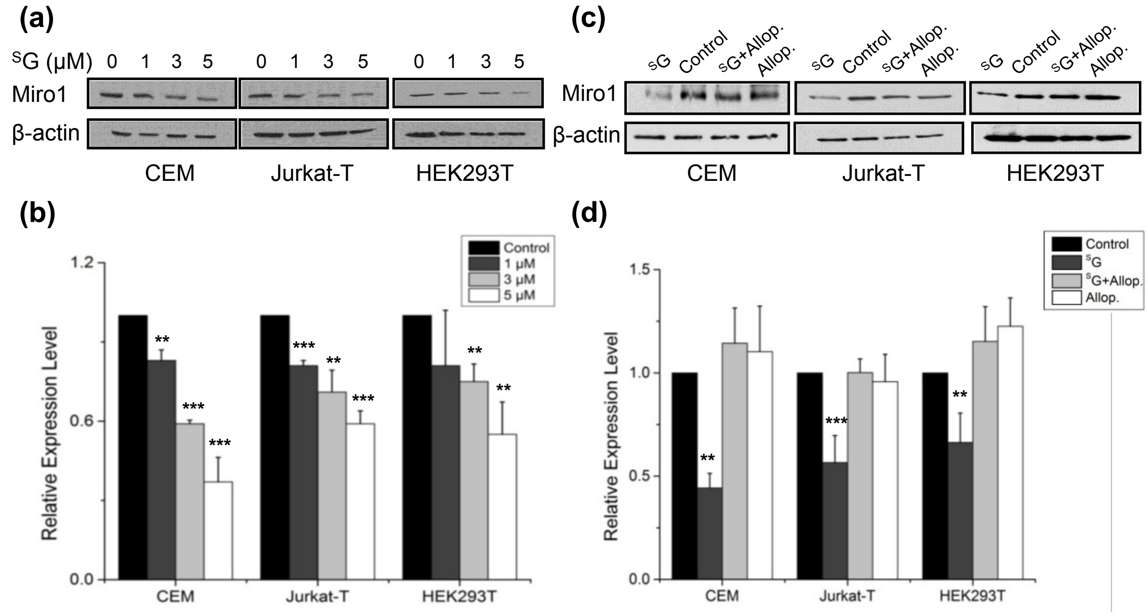


Fig. 3

Figure 3.3. Western blot for monitoring the expression levels of endogenous Miro1 in CEM, Jurkat-T and HEK293T cells upon treatment with increasing concentrations of ^{S}G (a) or treated for 24 hrs with ^{S}G and allopurinol, alone or in combination (c). β-actin served as the loading control. Shown in (b) and (d) are the quantification data representing the mean and S. D. of results from three independent drug treatment and Western blot experiments. "**", $p < 0.01$; "***", $p < 0.001$. The *p*-values were calculated using two-tailed, unpaired student *t* test.

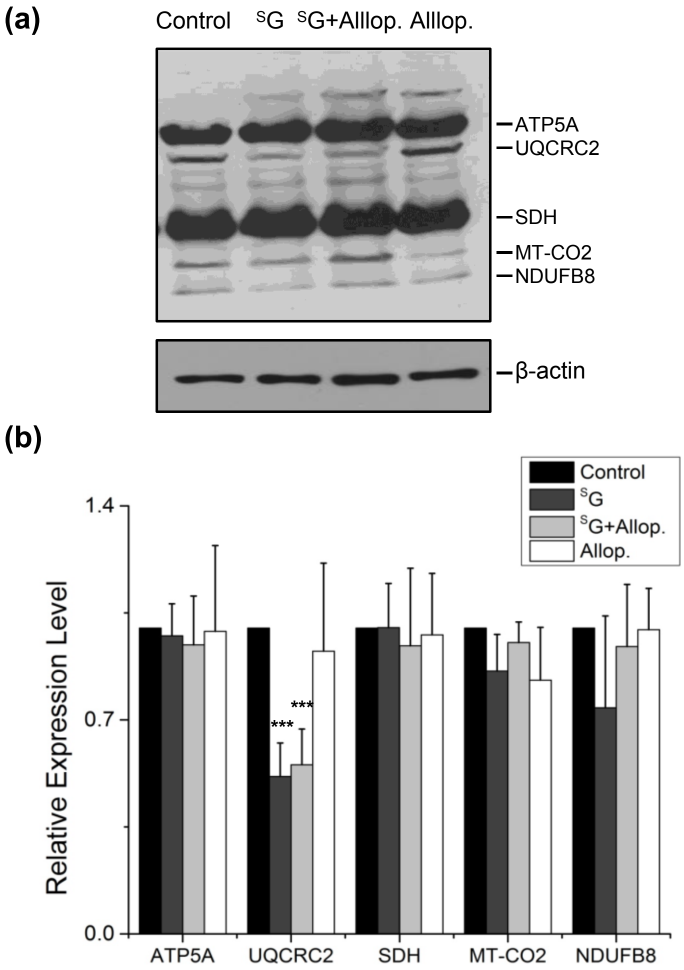


Figure 3.4. Western blot for monitoring the protein components of mitochondria respiratory complex. (a) Jurkat-T cells were treated with 3 μM ^{S}G, 3 μM ^{S}G + 500 μM allopurinol or 500 μM allopurinol alone for 24 hrs, analyzed on a 12% SDS-PAGE and probed for five representative proteins of the mitochondrial respiratory complex using an antibody cocktail (See Materials and Methods). β-actin served as the loading control. (b) Bar graph displaying the quantification results, which represent the mean and S. D. of results from three independent drug treatment and Western blot experiments. "***", $p < 0.001$. The p values were calculated using two-tailed, unpaired student t test.

We measured, by using LC-MS/MS, the levels of cdA and cdG in genomic DNA isolated from CEM, Jurkat-T and COS7 cells treated with ^{S}G (Representative LC-MS/MS data and calibrations curves are shown in Figures B4-B9). The LC-MS/MS results showed that the treatment with ^{S}G indeed stimulated the formation of the (5'*R*) diastereomer of cdA and cdG (Figure 3.6b-d), though we failed to observe significant increases in the levels of the (5'*S*) diastereomers of the two lesions (with the exception of *S*-cdG in CEM cells). This could be attributed to the less efficient formation of the (5'*S*) than (5'*R*) diastereomers of cdA and cdG, as found previously for the ionizing radiation-induced generation of these DNA lesions in isolated DNA (31).

In keeping with the elevated generation of oxidatively induced cdA and cdG lesions, we found that co-treatment of cells with allopurinol, an ROS scavenger, could abrogate the ^{S}G-induced reduction in expression level of Miro1 (Figure 3.3b). Nevertheless, co-treatment with allopurinol was not able to rescue the ^{S}G-induced diminished expression of UQCRC2 (Figure 3.4) or restore the ^{S}G-induced loss of active mitochondria, though allopurinol could rescue substantially the ^{S}G-induced growth inhibition of CEM, Jurkat-T and HEK293T cells (Figure 3.5a). These results suggest that decreased expression of proteins in the electron transport chain and mitochondrial dysfunction occur prior to ROS induction.

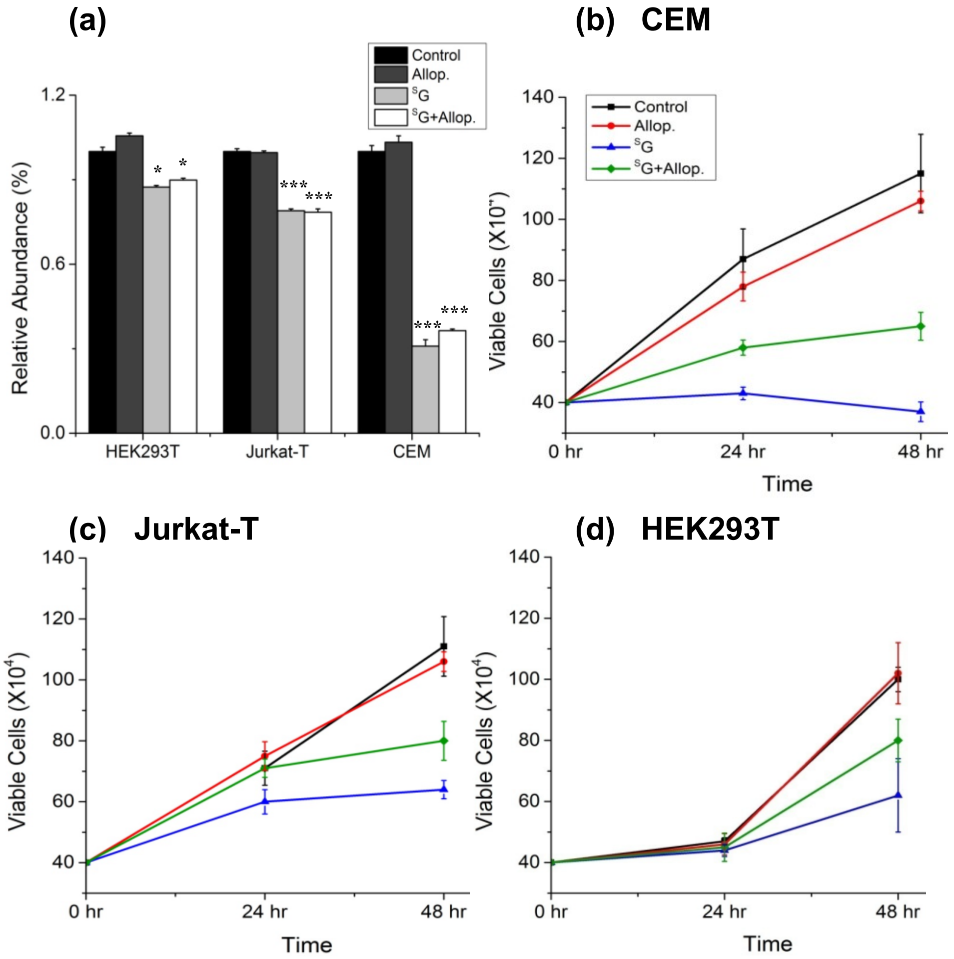


Figure 3.5. (a) Flow cytometry quantification data for the cellular uptake of MitoTracker Deep Red in CEM, Jurkat-T and HEK293T cells that were untreated (control), or treated with allopurinol (Allop.), ^{S}G (^{S}G), or both (^{S}G+Allop.). Representative plots for flow cytometry analysis are shown in Figure B2. "***", $p < 0.001$. The p-values were calculated using two-tailed, unpaired student t test. (b-d) Survival rates of CEM, Jurkat-T, and HEK293T cells that are untreated (control), or treated with allopurinol (Allop.), ^{S}G (^{S}G) or both (^{S}G+Allop.). The data represent the mean and S. D. of results from three independent experiments.

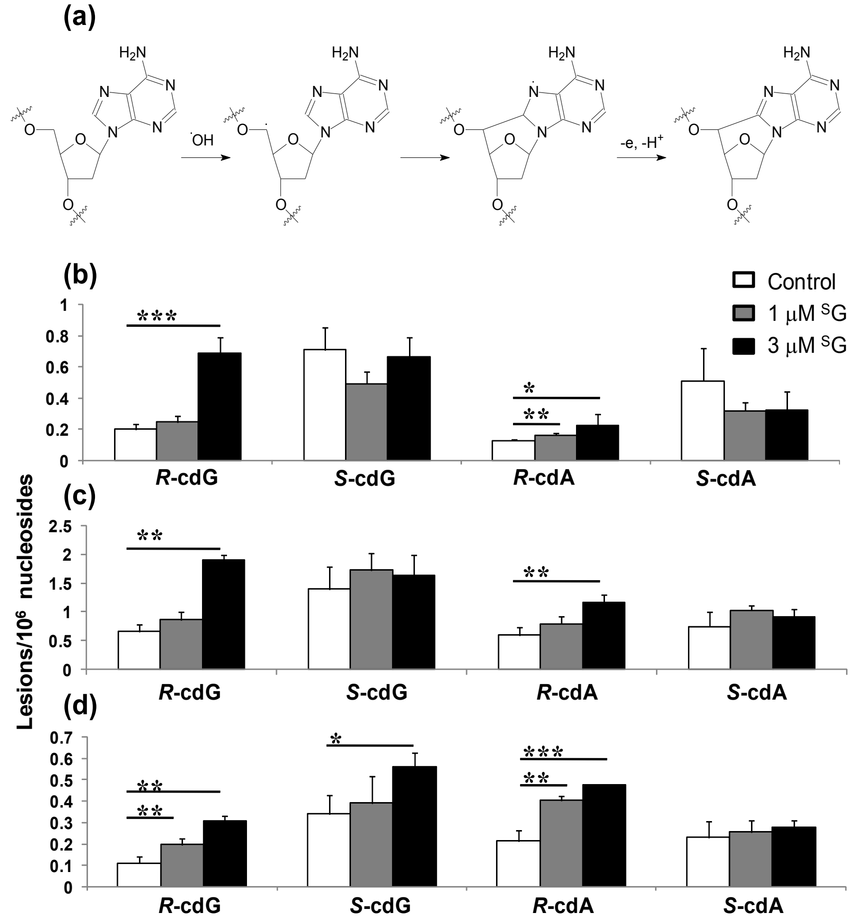


Figure 3.6. Proposed mechanism for the formation of cdA (a) and ^{S}G-induced formation of the (5’*R*) and (5’*S*) diastereomers of cdA and cdG in Jurkat-T (b), COS7 (c) and CEM (d) cells. The data represent the mean and S. D. of results from at least three independent drug treatments and LC-MS/MS measurements. “*”, $p < 0.05$; “**”, $p < 0.01$; “***”, $p < 0.001$. The *p* values were calculated using two-tailed, unpaired student *t* test.

CONCLUSIONS

Thiopurines are among the most successful chemotherapeutic agents for treating acute lymphoblastic leukemia and other human diseases (1-3). Although triggering of the MMR pathway by S^6mG in DNA was previously noted (4), it remains possible that other pathways may also contribute to the therapeutic efficacy of the drugs toward ALL (10). Thus, we set out to exploit novel mechanisms leading to the anti-leukemic effects of the thiopurine drugs by assessing the ^{S}G-induced alterations of global protein expression in Jurkat-T cells.

Our results revealed that the drug treatment led to diminished expression of a large number of proteins that are components of the electron transport chain of the mitochondrial respiratory complex (Table 3.1). In agreement with this observation, IPA results uncovered mitochondrial dysfunction as one of the major pathways altered by ^{S}G treatment (Table B2) and flow cytometry data showed the drug-induced loss of active mitochondria (Figure 3.5a). In addition, we observed elevated generation of oxidatively induced cdA and cdG lesions in ^{S}G-treated cells (Figure 3.6). Consistent with these findings, a recent study showed that ^{S}G treatment could give rise to enhanced formation of ROS, and the drug-induced ROS generation and cytotoxicity could be potentiated by glutathione depletion and rescued by allopurinol (32).

Mitochondrial dysfunction and the ensuing formation of oxidatively induced bulky cyclopurine DNA lesions may contribute to the cytotoxicity of the drug to ALL cells. In this vein, cdA and cdG are known to strongly block DNA replication and transcription in cells (36-38). Additionally, cdA and cdG can lead to mutations during DNA replication and transcription (36-38). Therefore, the ^{S}G-induced formation of cdA and cdG, if left unrepaired, may also contribute to the development of thiopurine therapy-related cancers (2, 39).

Our study also provides new knowledge about the sequence of events occurring underlying the ^{S}G-induced toxicity. Based on the fact that genetic defects of Complex I in the mitochondrial respiratory chain are known to result in augmented ROS production (28), we deduce that the reduced expression of a large number of proteins in the mitochondrial respiratory complex may constitute an upstream event leading to mitochondrial dysfunction and elevated generation of oxidatively induced DNA lesions. Additionally, we found that the reduced expression of mitochondrial electron transport chain protein UQCRC2 and loss of active mitochondria, as reflected by the cellular uptake of MitoTracker Deep Red, could not be rescued by co-treatment of cells with an ROS scavenger, allopurinol (Figures 3.4 and 3.5) (32). This result supported that the loss of active mitochondria did not arise from ROS induction. We also found that the ^{S}G-induced down-regulation of Miro1 could be fully abolished by co-treatment of the cells with allopurinol (Figure 3.3c&d), indicating that elevated ROS contributes to the decreased expression of Miro1. However, the molecular mechanisms underlying the ^{S}G-induced down-regulation of proteins in the mitochondrial respiratory chain complexes await further investigation.

Our findings also shed new light on the cellular effects of ^{S}G exposure observed in previous studies. In this vein, Kinsella and coworkers (33, 34) showed that ^{S}G treatment of cultured human cells could result in the activation of autophagy, a catabolic pathway for the lysosomal degradation of proteins and organelles. Considering the indispensable role of ROS, particularly those derived from mitochondria, in autophagy activation (35), our study also points to mitochondrial dysfunction as the potential origin for ^{S}G-induced autophagy.

Together, the results from the present study provided significant insights into the mechanisms of action of thiopurine drugs, namely, the drugs may exert their cytotoxic effect by triggering diminished expression of proteins in mitochondrial electron transport chain, inducing mitochondrial dysfunction and ROS generation, and resulting in augmented formation of oxidatively induced DNA lesions. Mitochondrial dysfunction may serve as a potential biomarker for monitoring the efficacy of the thiopurine therapy.

REFERENCES

1. Elion, G. B. (1989) The purine path to chemotherapy. *Science* 244, 41-47.

2. Karran, P. (2006) Thiopurines, DNA damage, DNA repair and therapy-related cancer. *Br. Med. Bull.* 79-80, 153-170.

3. Pui, C. H., and Jeha, S. (2007) New therapeutic strategies for the treatment of acute lymphoblastic leukaemia. *Nat. Rev. Drug Discov.* 6, 149-165.

4. Swann, P. F., Waters, T. R., Moulton, D. C., Xu, Y. Z., Zheng, Q. G., Edwards, M., and Mace, R. (1996) Role of postreplicative DNA mismatch repair in the cytotoxic action of thioguanine. *Science* 273, 1109-1111.

5. Wang, H., and Wang, Y. (2010) LC-MS/MS coupled with stable isotope dilution method for the quantification of 6-thioguanine and S^6-methylthioguanine in genomic DNA of human cancer cells treated with 6-thioguanine. *Anal. Chem.* 82, 5797-5803.

6. Yuan, B., O'Connor, T. R., and Wang, Y. (2010) 6-Thioguanine and S^6-methylthioguanine are mutagenic in human cells. *ACS Chem. Biol.* 5, 1021-1027.

7. Yuan, B., and Wang, Y. (2008) Mutagenic and cytotoxic properties of 6-thioguanine, S^6-methylthioguanine, and guanine-S^6-sulfonic acid. *J. Biol. Chem.* 283, 23665-23670.

8. Griffin, S., Branch, P., Xu, Y. Z., and Karran, P. (1994) DNA mismatch binding and incision at modified guanine bases by extracts of mammalian cells: implications for tolerance to DNA methylation damage. *Biochemistry* 33, 4787-4793.

9. Waters, T. R., and Swann, P. F. (1997) Cytotoxic mechanism of 6-thioguanine: hMutS$_a$, the human mismatch binding heterodimer, binds to DNA containing S^6-methylthioguanine. *Biochemistry* 36, 2501-2506.

10. Krynetski, E. Y., Krynetskaia, N. F., Gallo, A. E., Murti, K. G., and Evans, W. E. (2001) A novel protein complex distinct from mismatch repair binds thioguanylated DNA. *Mol. Pharmacol.* 59, 367-374.

11. Shankavaram, U. T., Reinhold, W. C., Nishizuka, S., Major, S., Morita, D., Chary, K. K., Reimers, M. A., Scherf, U., Kahn, A., Dolginow, D., Cossman, J., Kaldjian, E. P., Scudiero, D. A., Petricoin, E., Liotta, L., Lee, J. K., and Weinstein, J. N. (2007) Transcript and protein expression profiles of the NCI-60 cancer cell panel: an integromic microarray study. *Mol. Cancer Ther.* 6, 820-832.

12. Cheok, M. H., Yang, W., Pui, C. H., Downing, J. R., Cheng, C., Naeve, C. W., Relling, M. V., and Evans, W. E. (2003) Treatment-specific changes in gene expression discriminate in vivo drug response in human leukemia cells. *Nature Genet.* 34, 231-231.

13. Vogel, C., and Marcotte, E. M. (2012) Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat. Rev. Genet.* 13, 227-232.

14. Aebersold, R., and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* 422, 198-207.

15. Cravatt, B. F., Simon, G. M., and Yates, J. R. (2007) The biological impact of mass-spectrometry-based proteomics. *Nature* 450, 991-1000.

16. Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* 1, 376-386.

17. Lowe, E. S., Kitchen, B. J., Erdmann, G., Stork, L. C., Bostrom, B. C., Hutchinson, R., Holcenberg, J., Reaman, G. H., Woods, W., Franklin, J., Widemann, B. C., Balis, F. M., Murphy, R. F., and Adamson, P. C. (2001) Plasma pharmacokinetics and cerebrospinal fluid penetration of thioguanine in children with acute lymphoblastic leukemia: a collaborative Pediatric Oncology Branch, NCI, and Children's Cancer Group study. *Cancer Chemother. Pharmacol.* 47, 199-205.

18. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26, 1367-1372.

19. Graumann, J., Hubner, N. C., Kim, J. B., Ko, K., Moser, M., Kumar, C., Cox, J., Scholer, H., and Mann, M. (2008) Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC) and proteome quantitation of mouse embryonic stem cells to a depth of 5,111 proteins. *Mol. Cell. Proteomics* 7, 672-683.

20. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol* 26, 1367-1372.

21. Waanders, L. F., Hanke, S., and Mann, M. (2007) Top-down quantitation and characterization of SILAC-labeled proteins. *J. Am. Soc. Mass. Spectr.* 18, 2058-2064.

22. Muller, T., Schrotter, A., Loosse, C., Helling, S., Stephan, C., Ahrens, M., Uszkoreit, J., Eisenacher, M., Meyer, H. E., and Marcus, K. (2011) Sense and nonsense of pathway analysis software in proteomics. *J. Proteome. Res.* 10, 5398-5408.

23. Thomas, S., and Bonchev, D. (2010) A survey of current software for network analysis in molecular biology. *Hum. Genomics* 4, 353-360.

24. Lopez-Domenech, G., Serrat, R., Mirra, S., D'Aniello, S., Somorjai, I., Abad, A., Vitureira, N., Garcia-Arumi, E., Alonso, M. T., Rodriguez-Prados, M., Burgaya, F., Andreu, A. L., Garcia-Sancho, J., Trullas, R., Garcia-Fernandez, J., and Soriano, E. (2012) The Eutherian *Armcx* genes regulate mitochondrial trafficking in neurons and interact with Miro and Trak2. *Nat. Commun.* 3, 814.

25. Wang, J., Yuan, B., Guerrero, C., Bahde, R., Gupta, S., and Wang, Y. (2011) Quantification of oxidative DNA lesions in tissues of Long-Evans Cinnamon rats by capillary high-performance liquid chromatography-tandem mass spectrometry coupled with stable isotope-dilution method. *Anal. Chem.* 83, 2201-2209.

26. Schenkman, J. B., and Jansson, I. (2003) The many roles of cytochrome b5. *Pharmacol. Ther.* 97, 139-152.

27. Lenaz, G. (2001) The mitochondrial production of reactive oxygen species: mechanisms and implications in human pathology. *IUBMB Life* 52, 159-164.

28. Pitkanen, S., and Robinson, B. H. (1996) Mitochondrial complex I deficiency leads to increased production of superoxide radicals and induction of superoxide dismutase. *J. Clin. Invest.* 98, 345-351.

29. Mitra, D., Luo, X., Morgan, A., Wang, J., Hoang, M. P., Lo, J., Guerrero, C. R., Lennerz, J. K., Mihm, M. C., Wargo, J. A., Robinson, K. C., Devi, S. P., Vanover, J. C., D'Orazio, J. A., McMahon, M., Bosenberg, M. W., Haigis, K. M., Haber, D. A., Wang, Y., and Fisher, D. E. (2012) An ultraviolet-radiation-independent pathway to melanoma carcinogenesis in the red hair/fair skin background. *Nature* 491, 449-453.

30. Tilstra, J. S., Robinson, A. R., Wang, J., Gregg, S. Q., Clauson, C. L., Reay, D. P., Nasto, L. A., St Croix, C. M., Usas, A., Vo, N., Huard, J., Clemens, P. R., Stolz, D. B., Guttridge, D. C., Watkins, S. C., Garinis, G. A., Wang, Y., Niedernhofer, L. J., and Robbins, P. D. (2012) Inhibition of IKK/NF-κB delays the onset of senescence and aging-related degenerative diseases caused by DNA damage. *J. Clin. Invest.* 122, 2601-2612.

31. Belmadoui, N., Boussicault, F., Guerra, M., Ravanat, J. L., Chatgilialoglu, C., and Cadet, J. (2010) Radiation-induced formation of purine 5',8-cyclonucleosides in isolated and cellular DNA: high stereospecificity and modulating effect of oxygen. *Org. Biomol. Chem.* 8, 3211-3219.

32. Brem, R., and Karran, P. (2012) Oxidation-mediated DNA cross-linking contributes to the toxicity of 6-thioguanine in human cells. *Cancer Res.* 72, 4787-4795.

33. Zeng, X., and Kinsella, T. J. (2008) Mammalian target of rapamycin and S6 kinase 1 positively regulate 6-thioguanine-induced autophagy. *Cancer Res.* 68, 2384-2390.

34. Zeng, X., Yan, T., Schupp, J. E., Seo, Y., and Kinsella, T. J. (2007) DNA mismatch repair initiates 6-thioguanine--induced autophagy through p53 activation in human tumor cells. *Clin. Cancer Res.* 13, 1315-1321.

35. Scherz-Shouval, R., and Elazar, Z. (2011) Regulation of autophagy by ROS: physiology and pathology. *Trends Biochem. Sci.* 36, 30-38.

36. Yuan, B., Wang, J., Cao, H., Sun, R., and Wang, Y. (2011) High-throughput analysis of the mutagenic and cytotoxic properties of DNA lesions by next-generation sequencing. *Nucleic Acids Res.* 39, 5945-5954.

37. Jasti, V. P., Das, R. S., Hilton, B. A., Weerasooriya, S., Zou, Y., and Basu, A. K. (2011) (5'*S*)-8,5'-cyclo-2'-deoxyguanosine is a strong block to replication, a potent pol V-dependent mutagenic lesion, and is inefficiently repaired in *Escherichia coli*. *Biochemistry* 50, 3862-3865.

38. You, C., Dai, X., Yuan, B., Wang, J., Wang, J., Brooks, P. J., Niedernhofer, L. J., and Wang, Y. (2012) A quantitative assay for assessing the effects of DNA lesions on transcription. *Nat. Chem. Biol.* 8, 817-822.

39. Relling, M. V., Rubnitz, J. E., Rivera, G. K., Boyett, J. M., Hancock, M. L., Felix, C. A., Kun, L. E., Walter, A. W., Evans, W. E., and Pui, C. H. (1999) High incidence of secondary brain tumours after radiotherapy and antimetabolites. *Lancet* 354, 34-39.

CHAPTER 4

SILAC-based Quantitative Proteomic Analysis Unveiled the Arsenite-induced Perturbation of Multiple Pathways in Human Skin Fibroblast Cells

INTRODUCTION

Arsenic is one of the most toxic elements in the environment, and human exposure to arsenic in drinking water is a widespread public health concern (1). Epidemiological data from many regions around the world revealed the strong correlation between high levels of arsenic in drinking water and cancer risk in humans (2). Ingestion of high concentrations of arsenic causes cancers of the skin, lung, and urinary bladder and it is a suspected cause of kidney and other malignancies (3).

A number of studies have been conducted to explore the molecular mechanisms through which arsenic species exert their carcinogenic effect, and several modes of action have been proposed (4). In this vein, trivalent arsenic [As(III)] is transported into mammalian cells through aquaglyceroporins AQP7 and AQP9 (5). Once it is inside the cell, arsenite is believed to increase cancer risk, in part, by forming three-coordinate trigonal-pyramidal complexes with cysteine-containing proteins (6, 7). In addition, arsenic species may exert their carcinogenic effects via stimulating the formation of oxyradicals, inhibiting DNA damage repair, and modulating DNA and histone methylation (4, 8, 9). Nevertheless, the mechanisms through which inorganic arsenite induces carcinogenesis remain incompletely understood.

Global gene expression analysis with microarray has been used for exploiting the molecular mechanisms action of inorganic arsenite, and it was found that 133 genes were differentially expressed upon arsenite treatment (10). Although microarray is widely used to monitor global changes in gene expression, the method affords little information about protein expression changes (11). Mass spectrometry (MS)-based proteomic method allows for the identification and quantification of a large number of proteins in complex samples. In this context, LC-MS/MS together with various stable isotope-labeling methods allows for the quantitative assessment of protein expression at the entire proteome scale. Stable isotope labeling by amino acids in cell culture (SILAC) is a simple and efficient labeling method with minimal bias, which can afford accurate quantification of subtle changes of protein abundance in the whole proteome (12).

Herein, we investigated, by employing SILAC in conjunction with LC-MS/MS analysis, the perturbation of the entire proteome of GM00637 human skin fibroblast cells induced by $NaAsO_2$ treatment. We were able to quantify approximately 3880 unique proteins on an Orbitrap Velos mass spectrometer, 250 of which were significantly altered upon $NaAsO_2$ treatment. We further validated the quantification results of some select proteins by targeted analysis in multiple-reaction monitoring (MRM) mode on a triple quadrupole mass spectrometer. The results from our study provided a better understanding about the biological pathways perturbed by $NaAsO_2$ treatment and about the mechanisms contributing to carcinogenic effects of arsenic.

MATERIALS AND METHODS

Cell culture

GM00637 cells, which were kindly provided by Prof. Gerd P. Pfeifer (The City of Hope), were cultured in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum (FBS, Invitrogen, Carlsbad, CA), 100 IU/mL penicillin and 100 μg/mL streptomycin in 75 cm^2 culture flasks. Cells were maintained in a humidified atmosphere with 5% CO_2 at 37°C, with medium renewal of 2-3 times a week depending on cell density. For SILAC experiments, the complete light and heavy media were prepared by the addition of light or heavy lysine ($[^{13}C_6, {}^{15}N_2]$-L-lysine) and arginine ($[^{13}C_6]$-L-arginine), along with dialyzed FBS, to the DMEM medium without L-lysine or L-arginine (Cambridge Isotope Laboratories, Andover, MA). The GM00637 cells were cultured in heavy DMEM medium for at least 10 days to facilitate complete incorporation of the heavy lysine and arginine.

$NaAsO_2$ treatment and sample preparation

GM00637 cells, at a density of ~ 7×10^5 cells/mL in light or heavy DMEM medium, were treated with 5 μM $NaAsO_2$ (Sigma, St. Louis, MO) for 24 hr. The cells were subsequently harvested by centrifugation at 300 *g* at 4°C for 5 min, and washed for three times with ice-cold PBS. Cells were lysed with CelLytic™ M lysis buffer (Sigma) supplemented with 1 mM PMSF and a protease inhibitor cocktail (Sigma). The resulting cell lysate was centrifuged at 16,000 *g* at 4°C for 30 min and supernatant collected. The protein concentration in the cell lysate was measured using Quick Start™ Bradford

Protein Assay (Bio-Rad, Hercules, CA). In forward SILAC, the lysate of light labeled, arsenite-treated cells and that of the heavy labeled, control cells were combined at 1:1 ratio (w/w), and the labeling and arsenite treatment were reversed in the reverse SILAC experiment (Fig. 4.1a).

SDS-PAGE separation and in-gel digestion

The above equi-mass mixture of light and heavy lysates was separated on a 12% SDS-PAGE with a 4% stacking gel and stained with Coomassie blue. The gel was cut into 20 slices, and the proteins were reduced in-gel with dithiothreitol (DTT), alkylated with iodoacetamide (Sigma), and digested at 37°C overnight with trypsin (Promega, Madison, WI) at an enzyme/substrate ratio of 1:100. Following the digestion, peptides were extracted from the gels with 5% acetic acid in H_2O and then with 5% acetic acid in CH_3CN/H_2O (1:1, v/v). The resulting peptide mixtures were dried in a SpeedVac concentrator and stored at –80°C until further analysis.

LC-MS/MS for protein identification and quantification

On-line LC-MS/MS analysis was performed on an LTQ-Orbitrap Velos mass spectrometer coupled with an EASY n-LCII HPLC system and a nanoelectrospray ionization source (Thermo, San Jose, CA), as described previously (13). Briefly, the sample injection, enrichment, desalting, and HPLC separation were conducted automatically on a homemade trapping column (150 μm×40 mm) and a separation column (75 μm×200 mm, packed with ReproSil-Pur C18-AQ resin, 3 μm in particle size and 100 Å in pore size, Dr. Maisch HPLC GmbH, Germany). The peptide mixture was loaded onto the trapping column with a solvent mixture of 0.1% formic acid in CH_3CN/H_2O (2:98, v/v) at a flow rate of 3.0 µL/min, and subsequently separated with a 120-min linear gradient of 2-40% acetonitrile in 0.1% formic acid at a flow rate of 220 nL/min.

All tandem mass spectra were acquired in data-dependent scan mode. The full-scan mass spectra (from *m/z* 350 to 2000) were acquired with a resolution of 60,000 at m/z 400 after accumulation to a target value of 500,000. The twenty most abundant ions found in MS at a threshold above 500 counts were selected for fragmentation by low-energy collision-induced dissociation (CID) in the linear ion trap component of the instrument at a normalized collision energy of 35%. The same samples were also analyzed on the same instrument using higher energy collisional induced dissociation (HCD), which was performed by choosing the ten most abundant precursor ions for fragmentation in the HCD collision cell, where an activation time of 0.1 ms, an isolation window of 2.5 Da, and a normalized collision energy of 40% were used. The resolution for the HCD spectra was set to 7500 at m/z 400.

Data processing

The above LC-MS/MS data acquired from both low-energy CID and HCD were employed for the identification and quantification of the global proteome, which were conducted using Maxquant, Version 1.2.0.18 (14) against UniProt human database (with 538,585 sequence entries, release date: 2012.11.28) to which contaminants and reverse sequences were added. The maximum number of miss-cleavages for trypsin was two per peptide. Cysteine carbamidomethylation and methionine oxidation were set as fixed and variable modifications, respectively. The tolerances in mass accuracy for MS and MS/MS were 25 ppm and 0.6 Da, respectively. Only those proteins with at least two distinct peptides being discovered from LC-MS/MS analyses were considered reliably identified. Proteins with significant changes in SILAC experiments were determined by a combination of ratio and ratio significance calculated by MaxQuant normalization under the assumption that the median of all log-transformed ratio is zero (14) . The required false-positive discovery rate was set to 1% at both the peptide and protein levels, and the minimal required peptide length was 6 amino acids. SILAC experiments were conducted in three biological replicates, including two forward and one reverse SILAC labelings. To establish the threshold ratios for determining the significantly changed proteins, we calculated the standard scores (z-scores) for protein ratios as described by Mann et al. (14, 15). Our analysis showed that, at 95% confidence level, those proteins with expression ratios that are greater than 1.20 or less than 0.78 could be classified as significantly changed proteins. To be more stringent, we, however, only considered a protein to be significantly changed if its differential expression ratio ($NaAsO_2$ treated/control) was greater than 1.5 or less than 0.67.

Multiple-reaction monitoring (MRM) for targeted quantification of selected proteins

Raw data generated from LTQ-Orbitrap Velos were searched using Mascot 2.2 (Matrix Science, London, U.K.). The resulting DTA files were employed as input for processing in Skyline (16) to generate a list of fragment ions derived from targeted peptides. Four fragment ion pairs were selected for MRM transitions of each targeted peptide.

The same SILAC samples analyzed on LTQ-Orbitrap Velos were subjected to analysis on a TSQ-Vantage triple quadrupole mass spectrometer (Thermo Fisher Scientific) equipped with an Accela HPLC system with split nano-flow. The flow rate was 300 nL/min and a linear gradient of 5~35% acetonitrile in 0.1% formic acid was used. A spray voltage of 1.8 kV and a capillary temperature of 200°C were employed, and a resolution of 0.7 FWHM (full-width at half-maximum) was set for ion isolation in both Q1 and Q3. The collision gas pressure in Q2 was set at 1.2 mTorr, and a scan width of 0.002 m/z and a cycle time of 5 s were used for data acquisition. Collision energy was optimized using the default setting in Skyline. Targeted peptides were quantified using Skyline, where the ratio of the light and heavy versions of each peptide was calculated from the mean ratios of peak areas found in the chromatograms for individual pairs of MRM transitions for the peptide.

Ingenuity Pathway Analysis (IPA)

IPA (version 7.6, Ingenuity Systems Inc.) was employed to obtain information about relationships, functions and pathways of the differentially regulated proteins, and the Ingenuity Knowledge Base (IPKB, Ingenuity Systems Inc.) was used (17, 18). The IPKB database contained gene regulatory and signaling pathways, which are integrated with other relevant databases including Gene Ontology (http://www.geneontology.org) and NCBI Gene (http://www.ncbi.nlm.nih.gov/gene). The proteins that were significantly changed upon $NaAsO_2$ treatment, including those quantified only in one cycle of SILAC labeling experiment, were included for the pathway analysis. IPA determines the interaction by calculating a significance score with Fisher's exact test and exporting it as a p-value. A p-value of < 0.05 was considered significant for canonical pathways.

RESULTS AND DISCUSSION

To achieve a better understanding of the molecular mechanisms underlying the carcinogenic effects of arsenite, we set out to exploit the molecular targets and pathways affected by $NaAsO_2$. Toward this end, we employed an unbiased quantitative proteomic approach to assess the toxicant-induced alteration of the global proteome of GM00637 human skin fibroblast cells.

$NaAsO_2$ treatment, protein identification and quantification

We first identified the optimal dose of $NaAsO_2$ by examining the survival of GM00637 cells upon treatment with different concentrations of $NaAsO_2$. Trypan blue exclusion assay results revealed an approximately 5% cell death after a 24-hr treatment with 5 μM $NaAsO_2$; however, cell viability was significantly diminished (by ~15%) after a similar treatment using 10 μM $NaAsO_2$. Thus, 5 μM $NaAsO_2$ was chosen for the subsequent experiments to minimize the cell death-induced alterations of protein expression. In this vein, the concentration of arsenic in ground water can reach 5 μM in some heavily polluted areas in West Bengal, India (1). In addition, humans in contaminated areas are exposed, albeit at lower concentrations, to arsenic species across the entire lifespan; such long-term exposure, however, cannot be practically implemented in model studies using cultured cells in a laboratory setting.

To obtain reliable quantification results, we conducted the SILAC experiments in three biological replicates, including two forward and one reverse labelings (Fig. 4.1a). LC-MS/MS analysis of these SILAC samples enabled us to identify more than 5000 proteins, 3880 of which were quantified, with the expression ratio distribution shown in Fig. 4.1b. Among these proteins, 2719 were quantified in all three SILAC labeling experiments, and 2950 could be quantified in at least two SILAC labelings including the reverse cycle (Fig. 4.1c). By employing the criteria that: (1) the protein has to change by at least 1.5 fold ($NaAsO_2$ treated/control); (2) the protein has to be quantified in both forward and reverse SILAC labeling experiments, we found that 130 proteins were significantly altered upon $NaAsO_2$ treatment. The results for proteins with significant changes are summarized in Tables C1. Shown in Fig. 4.2a&b are the representative ESI-MS results for the quantification of a ferritin peptide, ALFQDIK, which demonstrated the arsenite-induced up-regulation of ferritin in both forward and reverse SILAC labeling experiments. In addition, MS/MS supported the identification of the light and heavy forms of this peptide (Fig. 4.2c&d).

We next validated the quantification results of 50 select proteins by employing targeted analysis in MRM mode on a triple quadrupole mass spectrometer, which was found to afford better sensitivity, reproducibility and accuracy in quantification than shotgun proteomics method based on data-dependent analysis (19). It turned out that ~ 90% of the proteins displayed consistent quantification results with those obtained from data-dependent analysis on the Orbitrap Velos (Table C2). Nevertheless, it is worth noting that, although both techniques showed the significant up-regulation of heme oxygenase 1, the expression ratio revealed by MRM-based targeted analysis (53.5 fold) was markedly higher than that determined by data-dependent analysis on the Orbitrap Velos (16.6 fold). This is likely attributed to the somewhat limited dynamic range of the Orbitrap Velos.

Ingenuity Pathway Analysis (IPA) of significantly changed proteins upon $NaAsO_2$ treatment

We next performed IPA analysis to acquire information with respect to the relationships, functions and pathways of the differentially expressed proteins induced by $NaAsO_2$ treatment (20). The IPA results showed the alterations of several pathways, including the Nrf2-mediated oxidative stress response, pancreatic adenocarcinoma signaling, chronic myeloid leukemia signaling and cell cycle regulation (Table C3). In addition, the IPA results revealed several biological functions perturbed by $NaAsO_2$ treatment (Tables C4).

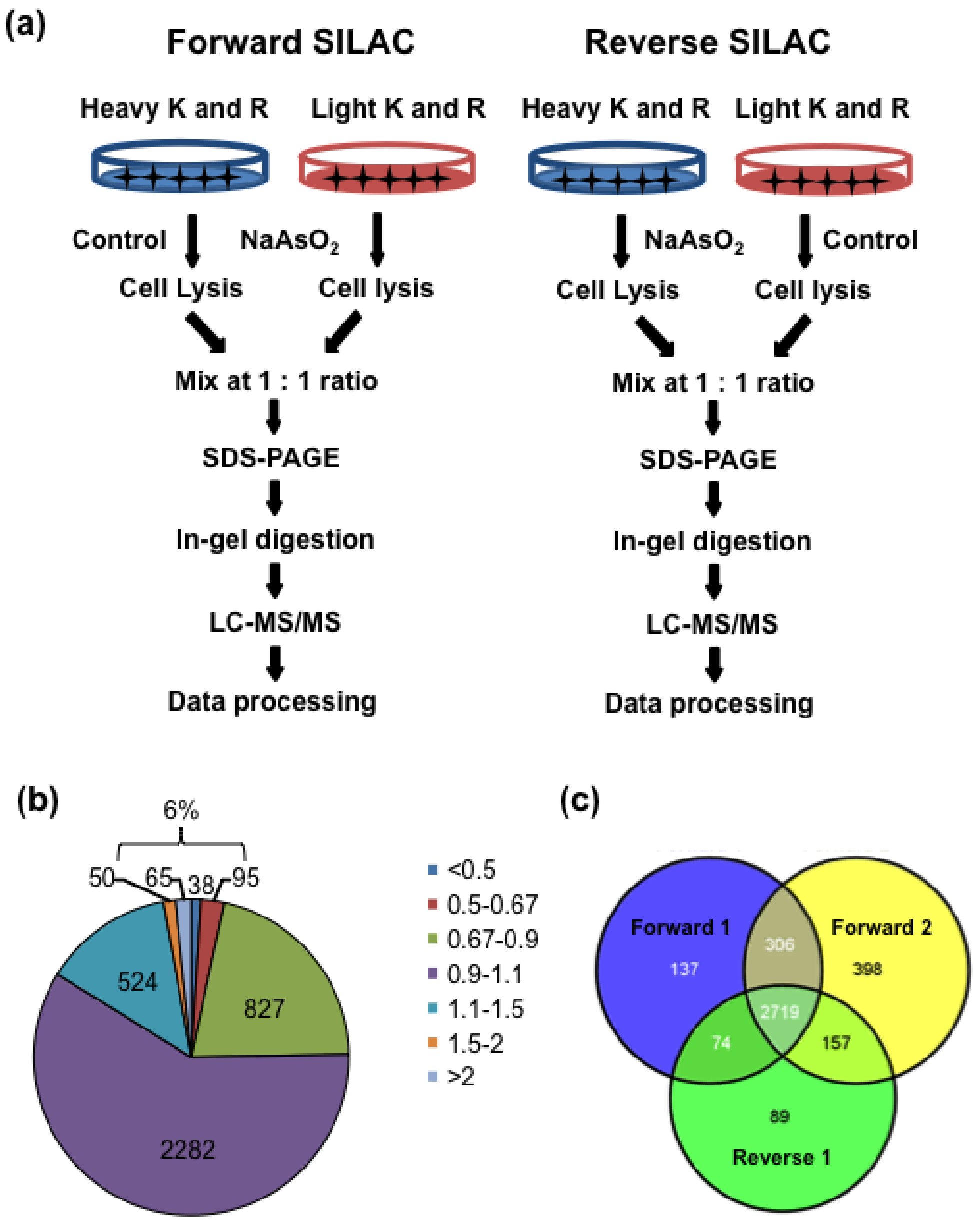


Figure 4.1. Forward- and reverse-SILAC combined with LC-MS/MS for the comparative analysis of protein expression in GM00637 cells upon arsenite treatment (A). Shown are the pie graph displaying the distribution of expression ratios (treated/untreated) for the quantified proteins (B) and Venn diagram revealing the number of quantified proteins (C) from three independent experiments.

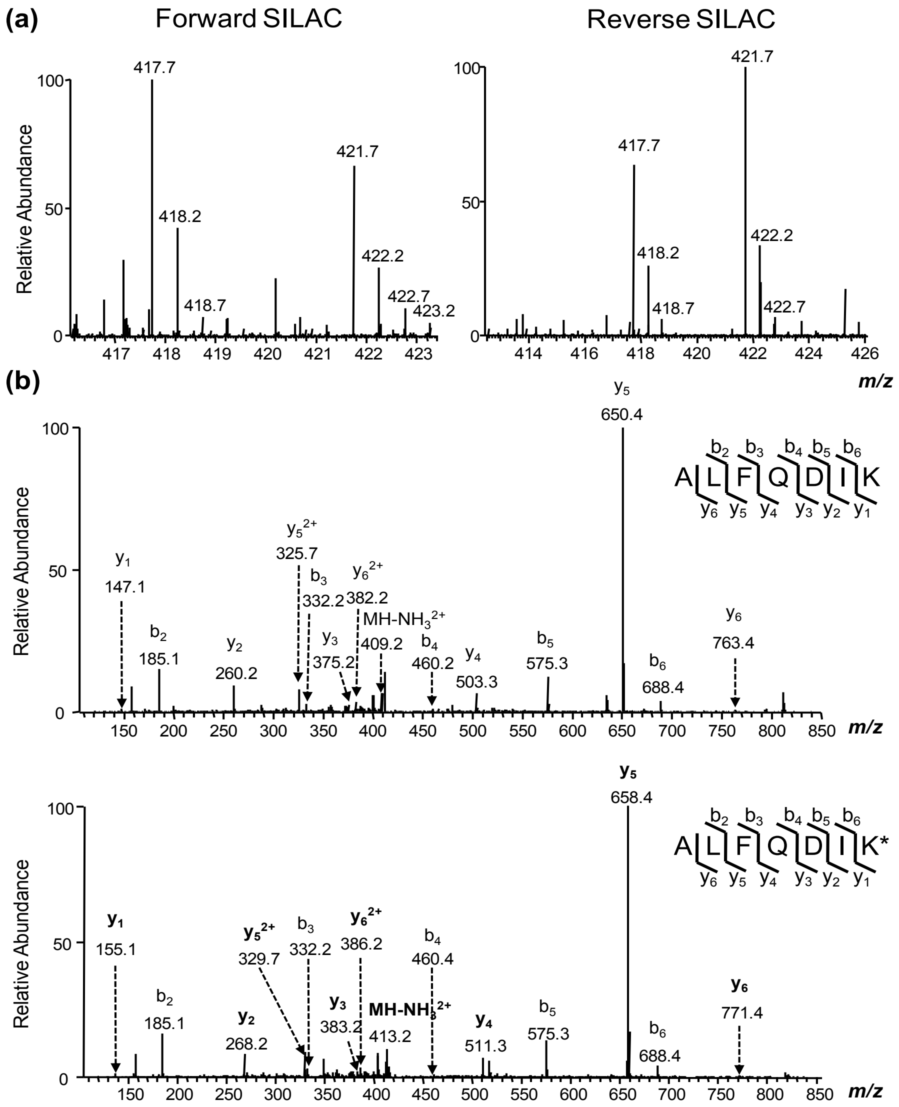


Figure 4.2. Example ESI-MS data revealing the arsenite-induced up-regulation of ferritin. Shown are the MS for the $[M+2H]^{2+}$ ions of ferritin peptide ALFQDIK (*m/z* 417.7) and ALFQDIK* (*m/z* 421.7)('K*'designates the heavy-labeled lysine) from the forward (left) and reverse (right) SILAC labeling experiments (a) and MS/MS for the $[M+2H]^{2+}$ ions of ALFQDIK and ALFQDIK* (b), where heavy-labeled lysine-containing y ions are labeled in bold.

Table 4.1. A select list of significantly changed proteins induced by arsenite treatment, with the UniProt ID, Protein Names, Expressions Ratios and S.D. being listed.

UniProt ID	Protein Name	Ratio (T/U) ± S.D.
	Nrf2-related proteins	
P09601	Heme oxygenase 1	16.56±2.58
P02794	Ferritin	1.52±0.02
O75190-1	DnaJ homolog subfamily B member 6	1.63±0.59
P50454	47 kDa heat shock protein	1.52±0.17
P25685	DnaJ homolog subfamily B member 1	1.53±0.17
Q9UDY4	DnaJ homolog subfamily B member 4	1.46±0.17
O75190-1	DnaJ homolog subfamily A member 1	1.39±0.13
P08107	Heat shock 70 kDa protein 1/2	2.23±0.47
B3KTT5	Highly similar to heat shock 70 kDa protein 1	2.30±0.56
P48507	Glutamate–cysteine ligase modifier subunits	2.18±0.29
P48506	Glutamate–cysteine ligase catalytic subunits	1.91
P41743	Atypical protein kinase C-lambda/iota	2.44±1.74
P27986-4	Phosphatidylinositol 3-kinase 85 kDa regulatory subunit alpha	0.63±0.34
	Selenoproteins	
Q9Y6D0	Selenoprotein K	0.49±0.08
Q8WWX9	Selenoprotein M	0.33±0.28
Q8IZQ5	Selenoprotein H	0.64±0.09
Q9BQE4	Selenoprotein S	0.37±0.08
P36969-1	Glutathione peroxidase 4	0.64±0.06
	Metallothionein proteins	
A8MWH4	Metallothionein-1F	2.43±0.27
P80294	Metallothionein-1H	2.97±0.03
P80297	Metallothionein-1X	6.57
P02795	Metallothionein-2	2.91±0.15
	Zinc finger proteins	
Q99942	E3 ubiquitin-protein ligase RNF5	0.66
A8K901	E3 ubiquitin-protein ligase UBR1	0.42
Q6ZT12	E3 ubiquitin-protein ligase UBR3	0.66
Q8N806	E3 ubiquitin-protein ligase UBR7	0.56±0.06
Q5BKZ1-1	Zinc finger protein 326	2.15±1.18
O95159	Zinc finger protein MCG4	1.64±0.10
	Cell cycle proteins	
Q13547	Histone deacetylase 1	0.67
Q96ST3	Histone deacetylase complex subunit Sin3a	0.48±0.38
O14519	Cyclin-dependent kinase 2-associated protein 1	1.99±0.39

$NaAsO_2$ treatment induced up-regulation of Nrf2-mediated oxidative stress response pathway

The above IPA results showed that Nrf2-mediated oxidative stress response is one of the most significantly enriched pathways perturbed by $NaAsO_2$ treatment (-log P = 3.34, Table S4), where nine proteins in this pathway were significantly altered upon $NaAsO_2$ treatment (Table 4.1).

Oxidative stress can occur when the production of reactive oxygen spices (ROS) in cells exceeds their detoxification capacity. Under normal physiological conditions, ROS form as byproducts of aerobic metabolism and play a crucial role in cell signaling and homeostasis (21). Upon exposure to some environmental agents (e.g., UV light and heavy metal ions), ROS levels, however, can increase drastically in cells and result in damage to cellular components (21). When cells are under oxidative stress, cytoplasmic Nrf2 protein is phosphorylated and translocated to the nucleus, where it can transactivate detoxifying and antioxidant enzymes, such as heme oxygenase 1 and superoxide dismutase (22). Here we found that $NaAsO_2$, when administered at 5 μM for 24 hr, can result in markedly elevated expression of a hallmark protein of the Nrf2 pathway, i.e., heme oxygenase 1 (22). Other downstream target proteins related to the Nrf2 pathway, including ferritin, glutamate-cysteine ligase, heat shock protein 40, heat shock protein 70 and atypical protein kinase C, also displayed elevated expression upon $NaAsO_2$ treatment (Table 4.1).

Activation of Nrf2 pathway is important in cellular defense against the deleterious effects of many environmental toxicants (22). Natural Nrf2 activators (e.g. sulforaphane and *tert*-butylhydroquinone) have been suggested for the dietary and therapeutic interventions against the adverse effects of arsenic (23), and these activators act through a Keap1-cysteine residue 151 (C151)-dependent mechanism (24). On the other hand, Nrf2 activation is also implicated in cancer promotion (25). In this vein, a previous study showed that, unlike the natural Nrf2 activators, arsenite may trigger Nrf2 activation through a mechanism independent of Keap1-C151 (26). In particular, arsenic may trigger Nrf2 activation via p62 accumulation, which induces autophagy dysregulation and Keap1 sequestration (27). The latter hampers Nrf2 ubiquitination by the Keap1-Cullin 3 E3 ubiquitin ligase and inhibits the degradation of Nrf2 by the proteasomal pathway (27). Thus, different from the protective effects of the natural Nrf2 activators, prolonged Nrf2 activation arising from arsenite exposure has the potential to elicit toxicity and carcinogenicity (27).Previous *in vitro* studies revealed that arsenite induces iron release from ferritin (32). Another study showed that arsenic exposure activated ferritin transcription via ARE along with elevation of methylation levels of histone H4 at arginine 3 and histone H3 at arginine 17 (33).

Glutamate-cysteine ligase (GCL), another Nrf2 target gene, was also markedly up-regulated in the $NaAsO_2$-treated cells. GCL acts as the rate-limiting enzyme catalyzing the synthesis of glutathione, which regulates the intracellular redox homeostasis and plays a predominant role in protecting cells from oxidative injury (34).

Sodium arsenite, when bound to GSH to form an $As(SG)_3$ complex (35), has been shown to reduce the intracellular level of GSH and induce cytotoxicity (36).

Other downstream target proteins related to Nrf2-induced oxidative stress pathway, including heat shock protein 40, heat shock protein 70 and atypical protein kinase C, also displayed elevated expression upon $NaAsO_2$ treatment. The elevated expression of the aforementioned Nrf2-target genes may reflect cells' protective response toward arsenite-induced ROS generation.

$NaAsO_2$ -induced down-regulation of selenoproteins

We observed significantly diminished expression of a group of selenoproteins, including glutathione peroxidase 4 (GPx4) and selenoproteins H, K, M, and S (Table 4.1). These proteins are known to play pivotal roles in antioxidant responses. For instance, glutathione peroxidases, which harbor a selenocysteine at their active sites, catalyze the reduction of harmful hydroperoxides (28). The importance of GPx4 in cellular antioxidant defense response is manifested by the observation that GPx4 KO mice displayed higher sensitivity to oxidative stress induced by ionizing radiation and H_2O_2 (29). Additionally, cells with reduced glutathione peroxidase level or activity exhibited higher sensitivity toward arsenite (30). Selenoprotein K (SelK) is also involved in protecting cells from the deleterious effects of ROS (31). In addition, selenoprotein H (SelH) was found to be important in regulating *de novo* glutathione synthesis and phase-II detoxification (32), whereas selenoprotein S (SelS) could modulate ER stress and was responsible for the control of inflammatory response (33).

Interestingly, all the above selenoproteins carry a selenocysteine that is capable of binding As(III) in cells (6). Moreover, arsenite is known to induce the decreased expression of GPx, a situation that diminishes the cell's ability to defend against ROS (34). Thus, the decreased expression of all the quantified selenoproteins indicates that arsenite may elicit its cytotoxic effect partly through perturbation of selenoprotein synthesis and/or covalently binding to selenocysteines in these proteins, both of which reduce the cells' capacity in defending against oxidative stress. In addition, the reduced expression of selenoproteins in GM00637 human skin fibroblast cells is consistent with previous findings made with human keratinocytes and mouse embryonic stem cells exposed with arsenite (34, 35). These results suggest that the alteration in expression of these selenoproteins might be a general effect of arsenite exposure.

$NaAsO_2$ induced up-regulation of metallothionein (MT) proteins

$NaAsO_2$ treatment also gave rise to considerable increases in the expression levels of MT proteins, including metallothioneins 1F, 1H, 1X and 2 (Table 1). MTs are a family of low-molecular weight polypeptides with 20-30% of amino acids being cysteines (36). Our results are in keeping with the previous finding that As(III) could induce the expression of MT in mice (37) and with the fact that these proteins are able to bind to a variety of heavy metals and metalloid including cadmium, copper, arsenic and zinc through their cysteine-rich domain (38, 39).

$NaAsO_2$ treatment induced alternation of zinc-finger proteins

Zinc-finger proteins are characterized by their capability in coordinating with one or more zinc ions to stabilize their structures (40). RING finger, a protein structural domain of the zinc-finger type with a $Cys_3HisCys_4$ amino acid motif, is also present in the majority of E3 ubiquitin ligases (41). In this context, As_2O_3 has been found to bind to the RING finger domain of PML in the oncogenic PML-RARα fusion protein and this binding ultimately results in the proteasomal degradation of the fusion protein (42). In addition, As(III) could inhibit oxidative DNA damage repair through binding to the zinc finger domain of poly(ADP-ribose) polymerase-1 (43, 44).

We found that a number of zinc-finger proteins were differentially expressed upon arsenite treatment. In particular, E3 ubiquitin-protein ligases UBR1, UBR3 and UBR7 were substantially down-regulated. UBR family proteins are also known as N-recognins due to the recognition of N-degrons or N-degron-like molecules in N-end rule pathway (45), which functions in the control of peptide import (46), chromosome segregation (47), apoptosis (48) and cardiovascular development (49). UBR1 recognizes type 1 and type 2 N-termini, whereas the recognition sites for UBR3 and UBR7 remain unclear (50). UBR3 was shown recently to polyubiquitinate APE1, a protein involved in DNA repair and transcriptional regulation (51). UBR7, a novel protein carrying a RING finger-like PHD domain, is known to be associated with transcriptional regulation (52). Owing to the important role of these RING-finger or RING-finger like proteins in cell cycle, DNA damage repair and transcriptional regulation (40), our quantification results indicate that $NaAsO_2$ may affect these biological processes and induce its carcinogenic effect through altering the expression of RING-finger domain-containing proteins.

CONCLUSIONS

Arsenic is an important human carcinogen, and human exposure to arsenicals in drinking water is known to be associated with the development of cancers of the skin, lung, urinary bladder, liver and kidney. Although accumulating evidence indicated that arsenic exposure could lead to chromosomal abnormalities, oxidative stress, altered DNA repair and cell cycle arrest (6), few studies have been carried out to examine comprehensively the alterations in expression of proteins involved in these processes (53).

In the present study, we assessed quantitatively the sodium arsenite-induced perturbation of the entire proteome of GM00637 human skin fibroblast cells. Our results showed that more than 250 proteins were significantly altered upon a 24-hr treatment with 5 μM $NaAsO_2$. IPA analysis of the significantly changed proteins revealed that arsenite exposure gave rise to the perturbation of more than ten pathways, including Nrf2-mediated oxidative stress response, pancreatic adenocarcinoma signaling, and cell cycle regulation, to name a few. Although many of these pathways were identified in previous studies, our unbiased quantitative proteomic approach led to the discovery of the differential expression of many proteins involved in each of these pathways. Thus, the current study painted a more complete picture for $NaAsO_2$-induced alterations of cellular pathways. It also demonstrated that the SILAC-based quantitative proteomic analysis is a powerful tool for the unbiased discovery of cellular pathways altered upon exposure to an environmental toxicant (i.e. arsenic).

REFERENCES

1. de The, H., and Chen, Z. (2010) Acute promyelocytic leukaemia: novel insights into the mechanisms of cure. *Nat. Rev. Cancer* 10, 775-783.

2. Zhang, X. W., Yan, X. J., Zhou, Z. R., Yang, F. F., Wu, Z. Y., Sun, H. B., Liang, W. X., Song, A. X., Lallemand-Breitenbach, V., Jeanne, M., Zhang, Q. Y., Yang, H. Y., Huang, Q. H., Zhou, G. B., Tong, J. H., Zhang, Y., Wu, J. H., Hu, H. Y., de The, H., Chen, S. J., and Chen, Z. (2010) Arsenic trioxide controls the fate of the PML-RARalpha oncoprotein by directly binding PML. *Science* 328, 240-243.

3. Das, D., Chatterjee, A., Samanta, G., Mandal, B., Chowdhury, T. R., Chowdhury, P. P., Chanda, C., Basu, G., Lodh, D., and et al. (1994) Arsenic contamination in groundwater in six districts of West Bengal, India: the biggest arsenic calamity in the world. *Analyst* 119, 168N-170N.

4. Cantor, K. P., and Lubin, J. H. (2007) Arsenic, internal cancers, and issues in inference from studies of low-level exposures in human populations. *Toxicol. Appl. Pharmacol.* 222, 252-257.

5. Hughes, M. F. (2002) Arsenic toxicity and potential mechanisms of action. *Toxicol Lett.* 133, 1-16.

6. Lubin, J. H., Pottern, L. M., Stone, B. J., and Fraumeni, J. F. (2000) Respiratory cancer in a cohort of copper smelter workers: Results from more than 50 years of follow-up. *Am. J. Epidemiol*. 151, 554-565.

7. Liu, Z., Shen, J., Carbrey, J. M., Mukhopadhyay, R., Agre, P., and Rosen, B. P. (2002) Arsenite transport by mammalian aquaglyceroporins AQP7 and AQP9. *Proc. Natl. Acad. Sci. U.S.A.* 99, 6053-6058.

8. Bertolero, F., Pozzi, G., Sabbioni, E., and Saffiotti, U. (1987) Cellular uptake and metabolic reduction of pentavalent to trivalent arsenic as determinants of cytotoxicity and morphological transformation. *Carcinogenesis* 8, 803-808.

9. Liu, S. X., Athar, M., Lippai, I., Waldren, C., and Hei, T. K. (2001) Induction of oxyradicals by arsenic: implication for mechanism of genotoxicity. *Proc. Natl. Acad. Sci. U.S.A.* 98, 1643-1648.

10. Wang, F., Zhou, X., Liu, W., Sun, X., Chen, C., Hudson, L. G., and Liu, K. J. (2013) Arsenite-induced ROS/RNS generation causes zinc loss and inhibits the activity of poly (ADP-ribose) polymerase-1. *Free Radic. Biol. Med.*

11. Kitchin, K. T. (2001) Recent advances in arsenic carcinogenesis: modes of action, animal model systems, and methylated arsenic metabolites. *Toxicol. Appl. Pharmacol.* 172, 249-261.

12. Wang, T. S., Kuo, C. F., Jan, K. Y., and Huang, H. (1996) Arsenite induces apoptosis in Chinese hamster ovary cells by generation of reactive oxygen species. *J. Cell. Physiol.* 169, 256-268.

13. Yih, L. H., Peck, K., and Lee, T. C. (2002) Changes in gene expression profiles of human fibroblasts in response to sodium arsenite treatment. *Carcinogenesis* 23, 867-876.

14. Russo, G., Zegar, C., and Giordano, A. (2003) Advantages and limitations of microarray technology in human cancer. *Oncogene* 22, 6497-6507.

15. Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* 1, 376-386.

16. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26, 1367-1372.

17. Graumann, J., Hubner, N. C., Kim, J. B., Ko, K., Moser, M., Kumar, C., Cox, J., Scholer, H., and Mann, M. (2008) Stable isotope labeling by amino acids in cell culture (SILAC) and proteome quantitation of mouse embryonic stem cells to a depth of 5,111 proteins. *Mol. Cell. Proteomics* 7, 672-683.

18. Waanders, L. F., Hanke, S., and Mann, M. (2007) Top-down quantitation and characterization of SILAC-labeled proteins. *J. Am. Soc. Mass. Spectrom.* 18, 2058-2064.

19. MacLean, B., Tomazela, D. M., Shulman, N., Chambers, M., Finney, G. L., Frewen, B., Kern, R., Tabb, D. L., Liebler, D. C., and MacCoss, M. J. (2010) Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* 26, 966-968.

20. Muller, T., Schrotter, A., Loosse, C., Helling, S., Stephan, C., Ahrens, M., Uszkoreit, J., Eisenacher, M., Meyer, H. E., and Marcus, K. (2011) Sense and nonsense of pathway analysis software in proteomics. *J. Proteome. Res.* 10, 5398-5408.

21. Thomas, S., and Bonchev, D. (2010) A survey of current software for network analysis in molecular biology. *Hum. Genomics* 4, 353-360.

22. Lange, V., Picotti, P., Domon, B., and Aebersold, R. (2008) Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol. Syst. Biol.* 4, 222.

23. Mayburd, A. L., Martinez, A., Sackett, D., Liu, H. T., Shih, J., Tauler, J., Avis, I., and Mulshine, J. L. (2006) Ingenuity network-assisted transcription profiling: Identification of a new pharmacologic mechanism for MK886. *Clin. Cancer Res.* 12, 1820-1827.

24. Devasagayam, T. P., Tilak, J. C., Boloor, K. K., Sane, K. S., Ghaskadbi, S. S., and Lele, R. D. (2004) Free radicals and antioxidants in human health: current status and future prospects. *J. Assoc. Physicians India* 52, 794-804.

25. Gopalakrishna, R., Chen, Z. H., and Gundimeda, U. (1994) Tobacco smoke tumor promoters, catechol and hydroquinone, induce oxidative regulation of protein kinase C and influence invasion and metastasis of lung carcinoma cells. *Proc. Natl. Acad. Sci. U.S.A*. 91, 12233-12237.

26. Liu, J., Kadiiska, M. B., Liu, Y., Lu, T., Qu, W., and Waalkes, M. P. (2001) Stress-related gene expression in mice treated with inorganic arsenicals. *Toxicol. Sci.* 61, 314-320.

27. Alam, J., Stewart, D., Touchard, C., Boinapally, S., Choi, A. M., and Cook, J. L. (1999) Nrf2, a Cap'n'Collar transcription factor, regulates induction of the heme oxygenase-1 gene. *J. Biol. Chem.* 274, 26071-26078.

28. Keyse, S. M., and Tyrrell, R. M. (1989) Heme oxygenase is the major 32-kda stress protein-induced in human-skin fibroblasts by uva radiation, hydrogen-peroxide, and sodium arsenite. *Proc. Natl. Acad. Sci. U.S.A.* 86, 99-103.

29. Arosio, P., and Levi, S. (2002) Ferritin, iron homeostasis, and oxidative damage. *Free Radic. Biol. Med*. 33, 457-463.

30. Torti, F. M., and Torti, S. V. (2002) Regulation of ferritin genes and protein. *Blood* 99, 3505-3516.

31. Rushworth, S. A., Chen, X. L., Mackman, N., Ogborne, R. M., and O'Connell, M. A. (2005) Lipopolysaccharide-induced heme oxygenase-1 expression in human monocytic cells is mediated via Nrf2 and protein kinase C. *J. Immunol.* 175, 4408-4415.

32. Ahmad, S., Kitchin, K. T., and Cullen, W. R. (2000) Arsenic species that cause release of iron from ferritin and generation of activated oxygen. *Arch. Biochem. Biophys.* 382, 195-202.

33. Huang, B. W., Ray, P. D., Iwasaki, K., and Tsuji, Y. (2013) Transcriptional regulation of the human ferritin gene by coordinated regulation of Nrf2 and protein arginine methyltransferases PRMT1 and PRMT4. *FASEB J.* 10.1096/fj.12-226043.

34. Meister, A., and Anderson, M. E. (1983) Glutathione. *Annu. Rev. Biochem.* 52, 711-760.

35. Scott, N., Hatlelid, K. M., Mackenzie, N. E., and Carter, D. E. (1993) Reactions of arsenic(Iii) and arsenic(V) Species with glutathione. *Chem. Res. Toxicol.* 6, 102-106.

36. Flora, S. J. (2011) Arsenic-induced oxidative stress and its reversibility. *Free Radic. Biol. Med.* 51, 257-281.

37. Imai, H., and Nakagawa, Y. (2003) Biological significance of phospholipid hydroperoxide glutathione peroxidase (PHGPx, GPx4) in mammalian cells. *Free Radic. Biol. Med.* 34, 145-169.

38. Imai, H., Hirao, F., Sakamoto, T., Sekine, K., Mizukura, Y., Saito, M., Kitamoto, T., Hayasaka, M., Hanaoka, K., and Nakagawa, Y. (2003) Early embryonic lethality caused by targeted disruption of the mouse PHGPx gene. *Biochem. Biophys. Res. Commun.* 305, 278-286.

39. Wang, T. S., Shu, Y. F., Liu, Y. C., Jan, K. Y., and Huang, H. (1997) Glutathione peroxidase and catalase modulate the genotoxicity of arsenite. *Toxicol.* 121, 229-237.

40. Lu, C., Qiu, F., Zhou, H., Peng, Y., Hao, W., Xu, J., Yuan, J., Wang, S., Qiang, B., Xu, C., and Peng, X. (2006) Identification and characterization of selenoprotein K: an antioxidant in cardiomyocytes. *FEBS Lett.* 580, 5189-5197.

41. Hwang, D. Y., Cho, J. S., Oh, J. H., Shim, S. B., Jee, S. W., Lee, S. H., Seo, S. J., Lee, S. K., and Kim, Y. K. (2005) Differentially expressed genes in transgenic mice carrying human mutant presenilin-2 (N141I): correlation of selenoprotein M with Alzheimer's disease. *Neurochem. Res.* 30, 1009-1019.

42. Panee, J., Stoytcheva, Z. R., Liu, W., and Berry, M. J. (2007) Selenoprotein H is redox-sensing high mobility group family DNA-binding protein that up-regulates genes involved in glutathione synthesis and phase II detoxification. *J. Biol. Chem.* 282, 23759-23765.

43. Gao, Y., Hannan, N. R. F., Wanyonyi, S., Konstantopolous, N., Pagnon, J., Feng, H. C., Jowett, J. B. M., Kim, K. H., Walder, K., and Collier, G. R. (2006) Activation of the selenoprotein SEPS 1 gene expression by pro-inflammatory cytokines in HepG2 cells. *Cytokine* 33, 246-251.

44. Kitchin, K. T., and Wallace, K. (2008) The role of protein binding of trivalent arsenicals in arsenic carcinogenesis and toxicity. *J. Inorg. Biochem.* 102, 532-539.

45. Ganyc, D., Talbot, S., Konate, F., Jackson, S., Schanen, B., Cullen, W., and Self, W. T. (2007) Impact of trivalent arsenicals on selenoprotein synthesis. *Environ. Health. Perspect.* 115, 346-353.

46. Margoshes, M., and Vallee, B. L. (1957) A Cadmium Protein from Equine Kidney Cortex. *J. Am. Chem. Soc.* 79, 4813-4814.

47. Peterson, C. W., Narula, S. S., and Armitage, I. M. (1996) 3D solution structure of copper and silver-substituted yeast metallothioneins. *FEBS Lett.* 379, 85-93.

48. Freisinger, E., and Vasak, M. (2013) Cadmium in metallothioneins. *Met. Ions Life Sci.* 11, 339-371.

49. Ngu, T. T., and Stillman, M. J. (2006) Arsenic binding to human metallothionein. *J. Am. Chem. Soc.* 128, 12473-12483.

50. Cherian, M. G., Jayasurya, A., and Bay, B. H. (2003) Metallothioneins in human tumors and potential roles in carcinogenesis. *Mutat. Res.* 533, 201-209.

51. Cherian, M. G., and Apostolova, M. D. (2000) Nuclear localization of metallothionein during cell proliferation and differentiation. *Cell Mol Biol* 46, 347-356.
52. Toyama, M., Yamashita, M., Hirayama, N., and Murooka, Y. (2002) Interactions of arsenic with human metallothionein-2. *J. Biochem.* 132, 217-221.

53. Albores, A., Koropatnick, J., Cherian, M. G., and Zelazowski, A. J. (1992) Arsenic induces and enhances rat hepatic metallothionein production in vivo. *Chem. Biol. Interact.* 85, 127-140.

54. Lipkowitz, S., and Weissman, A. M. (2011) RINGs of good and evil: RING finger ubiquitin ligases at the crossroads of tumour suppression and oncogenesis. *Nat. Rev. Cancer* 11, 629-643.

55. Lipkowitz, S., and Weissman, A. M. (2011) RINGs of good and evil: RING finger ubiquitin ligases at the crossroads of tumour suppression and oncogenesis. *Nat. Rev. Cancer* 11, 629-643.

56. Ding, W., Liu, W. L., Cooper, K. L., Qin, X. J., Bergo, P. L. D., Hudson, L. G., and Liu, K. J. (2009) Inhibition of poly(ADP-ribose) polymerase-1 by arsenite interferes with repair of oxidative DNA damage. *J. Biol. Chem.* 284, 6809-6817.

57. Zhou, X. X., Sun, X., Cooper, K. L., Wang, F., Liu, K. J., and Hudson, L. G. (2011) Arsenite interacts selectively with zinc finger proteins containing C3H1 or C4 motifs. *J. Biol. Chem.* 286, 22855-22863.

58. Bartel, B., Wunning, I., and Varshavsky, A. (1990) The recognition component of the N-end rule pathway. *Embo J.* 9, 3179-3189.

59. Du, F. Y., Navarro-Garcia, F., Xia, Z. X., Tasaki, T., and Varshavsky, A. (2002) Pairs of dipeptides synergistically activate the binding of substrate by ubiquitin ligase through dissociation of its autoinhibitory domain. *Proc. Natl. Acad. Sci. U.S.A.* 99, 14110-14115.

60. Rao, H., Uhlmann, F., Nasmyth, K., and Varshavsky, A. (2001) Degradation of a cohesin subunit by the N-end rule pathway is essential for chromosome stability. *Nature* 410, 955-959.

61. Ditzel, M., Wilson, R., Tenev, T., Zachariou, A., Paul, A., Deas, E., and Meier, P. (2003) Degradation of DIAP1 by the N-end rule pathway is essential for regulating apoptosis. *Nat. Cell. Biol.* 5, 467-473.

62. Kwon, Y. T., Kashina, A. S., Davydov, I. V., Hu, R. G., An, J. Y., Seo, J. W., Du, F., and Varshavsky, A. (2002) An essential role of N-terminal arginylation in cardiovascular development. *Science* 297, 96-99.

63. Tasaki, T., Mulder, L. C. F., Iwamatsu, A., Lee, M. J., Davydov, I. V., Varshavsky, A., Muesing, M., and Kwon, Y. T. (2005) A family of mammalian E3 ubiquitin ligases that contain the UBR box motif and recognize N-degrons. *Mol. Cell. Biol.* 25, 7120-7136.

64. Meisenberg, C., Tait, P. S., Dianova, I. I., Wright, K., Edelmann, M. J., Ternette, N., Tasaki, T., Kessler, B. M., Parsons, J. L., Kwon, Y. T., and Dianov, G. L. (2012) Ubiquitin ligase UBR3 regulates cellular levels of the essential DNA repair protein APE1 and is required for genome stability. *Nucleic Acids Res.* 40, 701-711.

65. Gozani, O., Karuman, P., Jones, D. R., Ivanov, D., Cha, J., Lugovskoy, A. A., Baird, C. L., Zhu, H., Field, S. J., Lessnick, S. L., Villasenor, J., Mehrotra, B., Chen, J., Rao, V. R., Brugge, J. S., Ferguson, C. G., Payrastre, B., Myszka, D. G., Cantley, L. C., Wagner, G., Divecha, N., Prestwich, G. D., and Yuan, J. Y. (2003) The PHD finger of the chromatin-associated protein ING2 functions as a nuclear phosphoinositide receptor. *Cell* 114, 99-111.

CHAPTER 5

Arsenite Binds to the RING Finger Domains of RNF20-RNF40 Histone E3 Ubiquitin Ligase and Inhibits DNA Double-strand Break Repair

INTRODUCTION

Owing to its natural abundance and industrial use, arsenic's environmental impact is felt by nearly 150 million people in more than 70 countries (1). Epidemiological studies revealed that human exposure to arsenic in drinking water is significantly associated with the prevalence of skin, lung, and bladder cancers (2-4). Inhibition of DNA repair has been suggested as a major mechanism in arsenic genotoxicity (5). In this respect, humans exposed to arsenite in drinking water were found to have increased frequencies of chromosome aberrations in peripheral lymphocytes and elevated frequencies of micronuclei in exfoliated cells (5). The inhibition of DNA repair might be attributable to the formation of three-coordinate trigonal-pyramidal complexes with cellular cysteine-containing proteins (6). Along this line, arsenite was previously found to displace Zn^{2+} in the zinc finger motifs of some DNA repair proteins including poly(ADP-ribose) polymerase 1 (PARP1) and XPA (7-10).

Aside from the carcinogenic effect of arsenic species, trivalent arsenic, in the form of arsenic trioxide, has demonstrated remarkable success in the clinical treatment of acute promyelocytic leukemia (11). In this regard, As(III) was found to bind to the RING finger domain of PML in the oncogenic PML-RARα fusion protein and this binding ultimately results in the proteasomal degradation of the fusion protein (12). RING finger, a protein structural domain of the zinc-finger type containing a $Cys_3HisCys_4$ amino acid motif, is also present in the majority of E3 ubiquitin ligases (13). Among them, the RNF20-RNF40 heterodimer targets histone H2B lysine 120 (K120) for monoubiquitination, and this ubiquitination event is important for decompacting the 30 nm chromatin fiber (14) and facilitating DNA double-strand break (DSB) repair (15, 16). We reasoned that arsenite may also bind to the RING finger motifs of RNF20 and RNF40, thereby perturbing H2B ubiquitination and suppressing DNA DSB repair. In the present study, we explored this possibility and found that As(III) can indeed interact directly with the RING finger motifs of RNF20 and RNF40 and result in reduced ubiquitination of histone H2B K120 in cells. As(III) treatment also gives rise to diminished recruitment of BRCA1 and RAD51 to laser-induced DNA DSB sites, leads to compromised DNA DSB repair via the homologous recombination (HR) and non-homologous end-joining (NHEJ) pathways, and sensitizes cells toward a radiomimetic agent. Thus, our study uncovers a novel mechanism underlying the carcinogenic effect of arsenite.

MATERIALS AND METHODS

Cell culture

HEK293T, HeLa and IMR90 cells were obtained from ATCC (Manassas, VA). U2OS cells harboring a chromosomally integrated copy of DR-GFP or EJ5-GFP reporter were provided by Prof. Jeremy M. Stark, (28) and GM00637 cells were obtained from Prof. Gerd P. Pfeifer (The City of Hope). GM00637, HeLa and HEK293T cells were cultured in Dulbecco's modified Eagle's medium (DMEM, ATCC). IMR90 cells were cultured in Eagle's minimal essential medium (EMEM, ATCC). U2OS cells were grown in DMEM with high glucose and L-glutamine, but without sodium pyruvate (Invitrogen). All culture media except those used for transfection were supplemented with 10% fetal bovine serum (FBS, Invitrogen, Carlsbad, CA) and 100 IU/mL penicillin. Cells were maintained in a humidified atmosphere with 5% CO_2 at 37ºC, with medium renewal of 2-3 times a week depending on cell density. For plasmid transfection, cells were switched to the same media as described above except that no penicillin was added.

***In vitro* arsenite binding assay**

The RING-finger peptides of RNF20 (with amino acid residues 921-964) and RNF40 (with residues 942-990) were obtained from Genemed Synthesis (San Antonio, TX), purified by HPLC and used for *in vitro* binding assays. Arsenite binding to the RING finger peptides was monitored by MALDI-TOF mass spectrometry on a Voyager DE STR instrument (Applied Biosystems, Framingham, MA) in linear, positive-ion mode.

Peptides were dissolved at a concentration of 1 mM in a buffer containing 20 mM Tris-HCl (pH 6.8) and 1 mM dithiothreitol. Aliquots of 100 μM peptides were incubated with 200 μM $NaAsO_2$ at room temperature for 1 hr. The resultant solution was diluted by 100 fold and mixed with an equal volume of 2,5-dihydroxybenzoic acid matrix solution before spotting onto a sample plate (10). The mass spectrometer was equipped with a pulsed nitrogen laser operating at 337 nm with a pulse duration of 3 ns. The acceleration voltage, grid voltage, and delayed extraction time were set at 20 kV, 65%, and 190 ns, respectively. Each mass spectrum was acquired from an average of signal from 100 laser shots.

The UV absorption spectra were recorded in the wavelength range of 240-400 nm on a Varian Cary 50 UV-visible spectrophotometer (Palo Alto, CA), where the RING finger peptides (100 μM) in 20 mM Tris-HCl (pH 6.8) were titrated with increasing amounts of $NaAsO_2$.

Extraction and enzymatic digestion of core histones

Core histones were isolated from cultured human cells following previously reported procedures (29). Briefly, GM00637, HEK293T and IMR90 cells, either untreated or after a 24-hr treatment with 5 μM arsenite, were harvested by centrifugation at 500g. The cell pellets were subsequently washed with a 5-mL lysis buffer containing 0.25 M sucrose, 10 mM $MgCl_2$, 0.5 mM PMSF, 50 mM Tris (pH 7.4) and 0.5% Triton X-100. The pellets were then resuspended in 5 mL of the same buffer and kept at 4ºC overnight. The histones were extracted from the cell lysis mixture with 0.4 M sulfuric acid by incubating at 4ºC for at least 4 hr with continuous vortexing, precipitated with cold acetone, centrifuged, dried and redissolved in water.

Core histone mixtures were digested with trypsin (Roche Applied Science, Indianapolis, IN) at a protein/enzyme ratio of 20:1 (w/w) in 100 mM NH_4HCO_3 (pH 8.0) at 37ºC overnight. The peptide mixtures were subjected to LC-MS/MS analysis.

LC-MS/MS for the identification and relative quantification of histone H2B K120 ubiquitination

On-line LC-MS/MS analysis was performed on an LTQ-Orbitrap Velos mass spectrometer coupled with an EASY n-LCII HPLC system and a nanoelectrospray ionization source (Thermo, San Jose, CA) for peptide sequencing and ubiquitination identification. The sample injection, enrichment, desalting, and HPLC separation were carried out automatically on a homemade trapping column (150 μm×50 mm) and a separation column (75 μm×120 mm, packed with ReproSil-Pur C18-AQ resin, 5 μm in particle size and 300 Å in pore size, Dr. Maisch HPLC GmbH, Germany). The peptide mixture was first loaded onto the trapping column with a solvent mixture of 0.1% formic acid in CH_3CN/H_2O (2:98, v/v) at a flow rate of 3.0 μL/min. The peptides were then separated using a 120-min linear gradient of 2-40% acetonitrile in 0.1% formic acid and at a flow rate of 300 nL/min.

The LTQ-Orbitrap Velos mass spectrometer was operated in the positive-ion mode with a spray voltage of 1.8 kV. All MS/MS data were acquired in a data-dependent scan mode where one full MS scan was followed with twenty MS/MS scans. To obtain high-quality MS/MS, the mass spectrometer was also set up in selected-ion monitoring (SIM) mode where the fragmentations of the $[M+2H]^{2+}$ ion of the H2B peptide containing diglycine-linked K120, which represents the ubiquitin remnant after tryptic digestion, was monitored. The quantification of ubiquitination level was conducted in SIM mode by comparing the relative abundances of ions corresponding to the target ubiquitinated peptide (AVTK$_{GG}$VTSSK) with respect to a reference H2B peptide EIQTAVR, which is not known to carry any post-translational modification, following previously published procedures (23).

Plasmid construction

The expression plasmids for GFP-RNF20 and Myc-RNF40 were provided by Dr. Yossi Shiloh (15). GFP-RNF40 vector was constructed by subcloning the RNF40 coding sequence from the Myc-RNF40 into the unique XhoI and BamH1 sites of pEGFP-C3 vector (provided by Dr. Alan R. Lehmann). The expression plasmids of GFP-RNF20 carrying the C922,924A or C957,960A mutations and GFP-RNF40 harboring the C948,950A or C983,986A mutations were obtained by site-directed mutagenesis, and the sequences for the mutated plasmids were verified by sequencing.

Fluorescence microscopy for monitoring the interaction between As(III) and RNF20 or RNF40

HEK293T cells, seeded in 6-well plates at a density of ~3×10^5 cells per well, were transfected with 1.5 μg wild-type or mutant GFP-RNF20 or GFP-RNF40 using Lipofectamine 2000 (Invitrogen). Cells were incubated at 37ºC for another 18-48 hr, transferred to cover glasses, and incubated with 5 μM ReAsH-EDT_2 (Invitrogen, Carlsbad, CA) at 37ºC for 1 hr in serum-free Opti-MEM (Invitrogen). After thoroughly washing with BAL buffer, cells were fixed with 4% paraformaldehyde and imaged with a Leica TCS SP5 confocal microscope (Leica Microsystems, Buffalo Grove, IL). Cells were also pretreated with 10 μM $NaAsO_2$ or PAPAO for 1 hr prior to transfection with the wild-type GFP-RNF20 or GFP-RNF40 and treatment with ReAsH-EDT_2.

Streptavidin agarose affinity assay and Western blot

Biotin-As was synthesized by conjugating *p*-aminophenylarsine oxide (PAPAO) to activated PFP-biotin (18, 19), and used for streptavidin agarose affinity assay following previously published procedures (12). Briefly, HEK293T cells were transfected with wild-type or mutant GFP-RNF20 or GFP-RNF40. At 24 hr after the transfection, cells were treated with 5 μM biotin-As for 2 hr and lysed in CelLytic^{TM} M lysis buffer supplemented with a protease inhibitor cocktail (Sigma-Aldrich). The cell lysates were incubated with streptavidin agarose at 4ºC for overnight. Streptavidin agarose beads were subsequently washed with 1×PBS and resuspended in SDS-PAGE loading buffer.

After SDS-PAGE separation, proteins were transferred to a nitrocellulose membrane using a solution containing 10 mM $NaHCO_3$, 3 mM Na_2CO_3, and 20% methanol. The membranes were blocked with 5% non-fat milk in PBS buffer containing 0.1% (v/v) Tween-20 (pH 7.5) for 7 hr and incubated overnight at 4°C with rabbit anti-GFP antibody (1:20000 dilution, Sigma-Aldrich). The membranes were washed with fresh PBS-T at room temperature for five times (10 min each). After washing, the membranes were incubated with HRP-conjugated secondary antibody at room temperature for 1 hr. The membranes were subsequently washed with PBS-T for five times. The secondary antibody was detected by using ECL Advance Western Blotting Detection Kit (GE Healthcare) and visualized with Hyblot CL autoradiography film (Denville Scientific, Inc., Metuchen, NJ). Similar experiments were also conducted by pretreating cells with 10 μM $NaAsO_2$, PAPAO, or Zn^{2+} for 1 hr prior to the biotin-As treatment.

Fluorescence microscopy for monitoring the recruitment of DNA repair proteins to laser-localized DNA DSBs

HeLa cells were seeded in a 35-mm glass bottom culture dish (MatTek) and incubated with 0, 5 or 20 μM sodium arsenite at 37°C for 8 hr prior to laser damage. Localized irradiation was performed using a Nikon Eclipse TE2000 confocal microscope equipped with an SRS NL100 nitrogen laser-pumped dye laser (Photonics Instruments, St. Charles, IL) that fires 3-ns pulses with a repetition rate of 10 Hz at 365 nm, with a power of 0.7 nW measured at the back aperture of the 60× objective. The laser was directed to a specified rectangular region of interest within the nucleus of a cell visualized with a Plan Fluor 60×/NA 1.25 oil objective. The laser beam was oriented by galvanometer-driven beam displacers and fired randomly throughout the region of interest until the entire region was exposed. Throughout an experiment, cells were maintained at 37°C, 5% CO_2, and 80% humidity using a Live Cell^{TM} environmental chamber. After the laser treatment, cells were incubated at 37°C for 10 min and fixed immediately with 4% formaldehyde in PBS at room temperature (RT) for 10 min or with 0.5% Triton X in PBS on ice for 10 min followed by 2% formaldehyde at RT for 10 min. Fixed cells were permeabilized with 0.5% Triton X-100, 1% BSA, 100 mM glycine and 0.2 mg/mL EDTA in PBS on ice for 10 min. The cells were subsequently digested with RNase A at 37ºC. For immunofluorescence staining, cells were incubated at 37ºC for 1.0 hr with primary antibodies for γ-H2AX (Millipore), BRCA1 (Abcam) and RAD51 (Santa Cruz). Cells were subsequently incubated with corresponding secondary antibodies (Alexa Fluor goat anti-mouse or Alexa Fluor goat anti-rabbit, Molecular Probes, Invitrogen) and mounted with ProLong Gold Antifade Reagent with DAPI (Molecular Probes, Invitrogen). Immunostained cells were visualized and imaged using Hamamatsu EM-CCD digital camera attached to the Nikon Eclipse TE2000 confocal microscope.

DSB repair assays

U2OS cells with a chromosomally integrated copy of DR-GFP or EJ5-GFP plasmid, either untreated or after a 24 hr treatment with $NaAsO_2$, were transfected with the I-SceI expression vector pCBASce (28). The U2OS-DR-GFP and U2OS-EJ5-GFP cells without I-SceI transfection were used as negative control. Three days after the transfection, cells were washed with PBS and stored in a sorting buffer (1×PBS, 1 mM EDTA, 25 mM HEPES, 1% FBS, pH 7.0) for flow cytometry analysis (BD FACS Aria I, NJ). The U2OS-DR-GFP cells, with or without $NaAsO_2$ treatment, were also transfected with pCAGGS-GFP to assess whether $NaAsO_2$ exposure affected the transfection efficiency (28).

Colony survival assay

HeLa cells either untreated or after a 24-hr treatment with 5 μM arsenite, were plated in 6-well plates in triplicate at densities of 150-8000 cells per well. The cells were subsequently exposed to various doses of neocarzinostatin (NCS) and the cells were immediately attached to the plates. A 'split-dose' protocol was applied (30), where each NCS dose was split into 15 ng/mL portions applied at 24 h intervals for up to 5 days. Cell colonies grown for 10-14 days were then fixed with 6% (v/v) glutaraldehyde and stained with 0.5% (w/v) crystal violet. Colonies containing at least 50 cells were subsequently counted under a microscope.

RESULTS AND DISCUSSION

In vitro binding between arsenite and RING finger peptides

To investigate whether $NaAsO_2$ can bind directly to the RING finger domain of RNF20 and RNF40, we performed *in vitro* binding assays using synthetic RING finger peptides of the two proteins. MALDI-TOF MS results revealed mass increases of 72 and 144 Da upon incubation of the peptides with arsenite (Figure 5.1a). The +72 Da mass shift reflects the binding of As(III) to the RING finger peptides with the release of three protons, suggesting the coordination of As(III) with three cysteines (10). The MS data, therefore, suggest that each molecule of the RING finger peptides can bind up to two As(III).

Thiol coordination to As(III) gives rise to new charge-transfer electronic transitions in the near UV range (250-320 nm) that can be monitored by UV absorbance (17). We found, from optical absorption experiments, that the two RING finger peptides displayed increased absorbance in this wavelength range upon addition with increasing amounts of $NaAsO_2$ (Figure 5.1b) (17). Together, these assays demonstrated that $NaAsO_2$ can bind to sulfhydryl group of cysteine residues in the RING finger domains of RNF20 and RNF40 *in vitro*.

Binding between arsenite and RING finger domains of RNF20 and RNF40 in cells

We next investigated the binding of arsenite with the RING domains of RNF20 and RNF40 in cells. To this end, we synthesized a biotin-As probe by conjugating *p*-aminophenylarsine oxide (PAPAO) with biotin (18, 19) and assessed arsenite binding by streptavidin agarose affinity assay. Western blot analysis revealed that the treatment of cells with biotin-As facilitated the pull-down of ectopically expressed RNF20 and RNF40, suggesting that RNF20 and RNF40 can bind to the biotin-As probe in cells (Figure 5.2a and Figure D1a). Moreover, we failed to pull down the two proteins when the cells were pretreated with $NaAsO_2$ or PAPAO, though pretreatment with Zn^{2+} still led to the successful pull-down of the two proteins, demonstrating that the interactions between biotin-As and RNF20 or RNF40 are specific (Figure 5.2a and Figure D1b). To determine if the cysteine residues in the RING finger domains are required for such interactions, we mutated the two cysteine residues in each of the two Zn^{2+}-binding sites in the RING finger motifs of RNF20 and RNF40 to alanines (i.e., the C922,924A and C957,960A mutants for RNF20 and the C948,950A and C983,986A mutants for RNF40) and performed the same pull-down assays. Our results showed that the mutations abolished the pull-down of the two proteins (Figure D1c), indicating that an intact RING finger is essential for the interaction between biotin-As and RNF20 or RNF40. Along this line, Western analysis revealed that the mutations did not lead to decreased expression of the two proteins (Figure D1d-e).

We further examined arsenite binding by performing fluorescence microscopy experiments with the use of ReAsH, which displays no red fluorescence until its arsenic moieties bind to nearby cysteine sulfhydryl groups in proteins (20). We found that the GFP-conjugated RNF20 or RNF40 can bind to As(III) in ReAsH as manifested by the emission of red fluorescence from the dye in ReAsH-treated cells (Figure 5.2 and Figure D2). Furthermore, ReAsH colocalized to nuclear foci with GFP-fused RNF20 and RNF40 in HEK293T cells. Such co-localization was abolished when the cells were pretreated with $NaAsO_2$ or PAPAO (but not with Zn^{2+}), or when two of the RING finger cysteine residues were mutated to alanines (Figure 5.2b&c). These results again demonstrate that the interactions between As(III) and RNF20 or RNF40 in human cells are specific and such interactions necessitate the intact RING finger motif.

$NaAsO_2$ perturbs histone H2B K120 ubiquitination and inhibits DNA DSB repair in human cells

The RNF20-RNF40 heterodimer induces in the ubiquitination of histone H2B at K120, which is important in transcription elongation (21, 22). Recently this ubiquitination was also found to be required for decompacting the 30 nm chromatin fiber (14) and play a crucial role in DNA DSB repair via the HR and NHEJ pathways (15, 16). In this context, cells lacking RNF20 exhibit reduced recruitment of HR and NHEJ factors to DSB sites and display markedly enhanced sensitivity to ionizing radiation and neocarzinostatin (NCS), a radiomimetic drug (15, 16). We reasoned that the binding of arsenite to the RING finger motifs of RNF20 and RNF40 might diminish the heterodimer's capability in inducing ubiquitination of histone H2B at K120 thereby suppressing DNA DSB repair. To test this hypothesis, we extracted core histones from the control and arsenite-treated cells, and monitored the levels of H2B K12

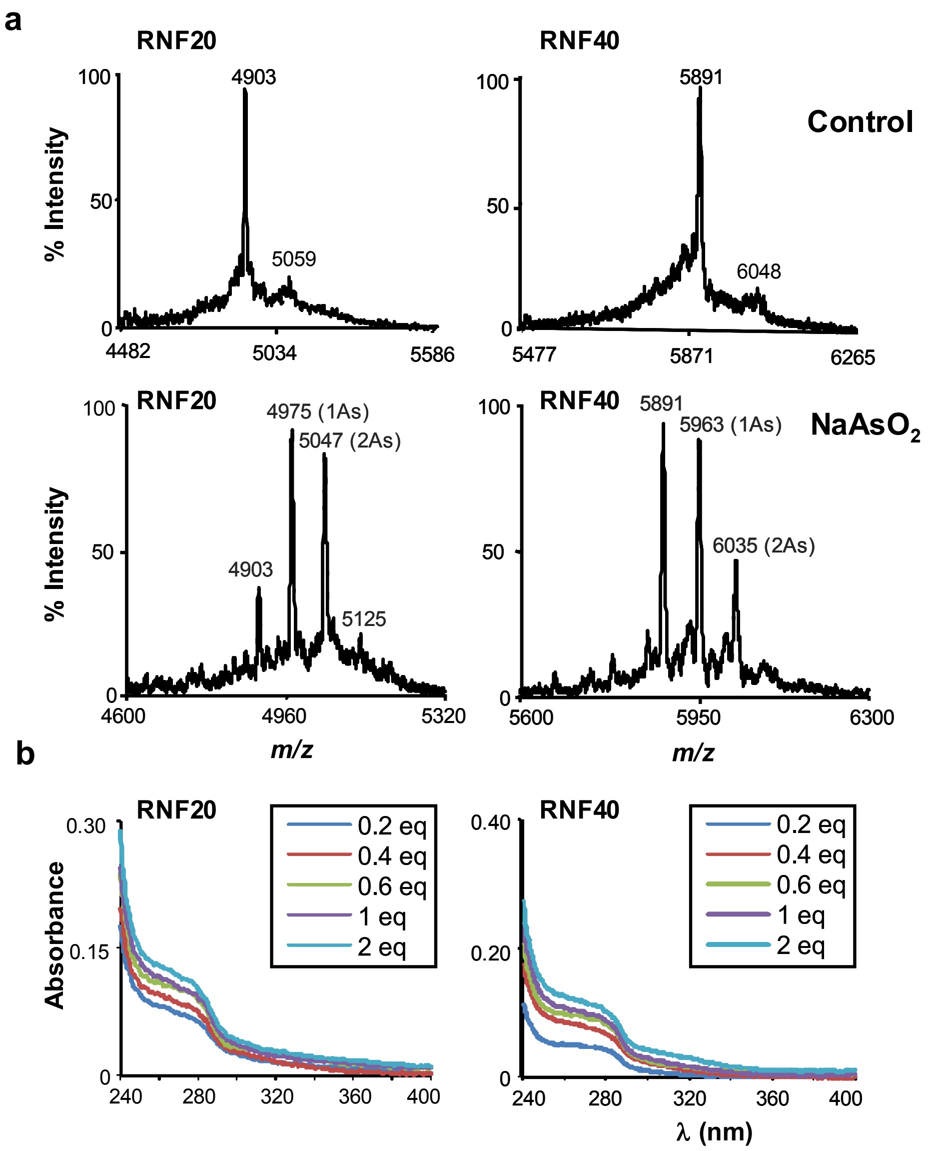


Figure 5.1. *In vitro* binding between $NaAsO_2$ and RING finger peptides of RNF20 and RNF40. (a) MALDI-TOF mass spectrometry for monitoring the interaction between arsenite and the RING finger peptides of RNF20 and RNF40. The molar ratios between the RING finger peptides and arsenite were 1:2. (b) UV absorption spectra of the RING finger peptides titrated with increasing amounts of $NaAsO_2$.

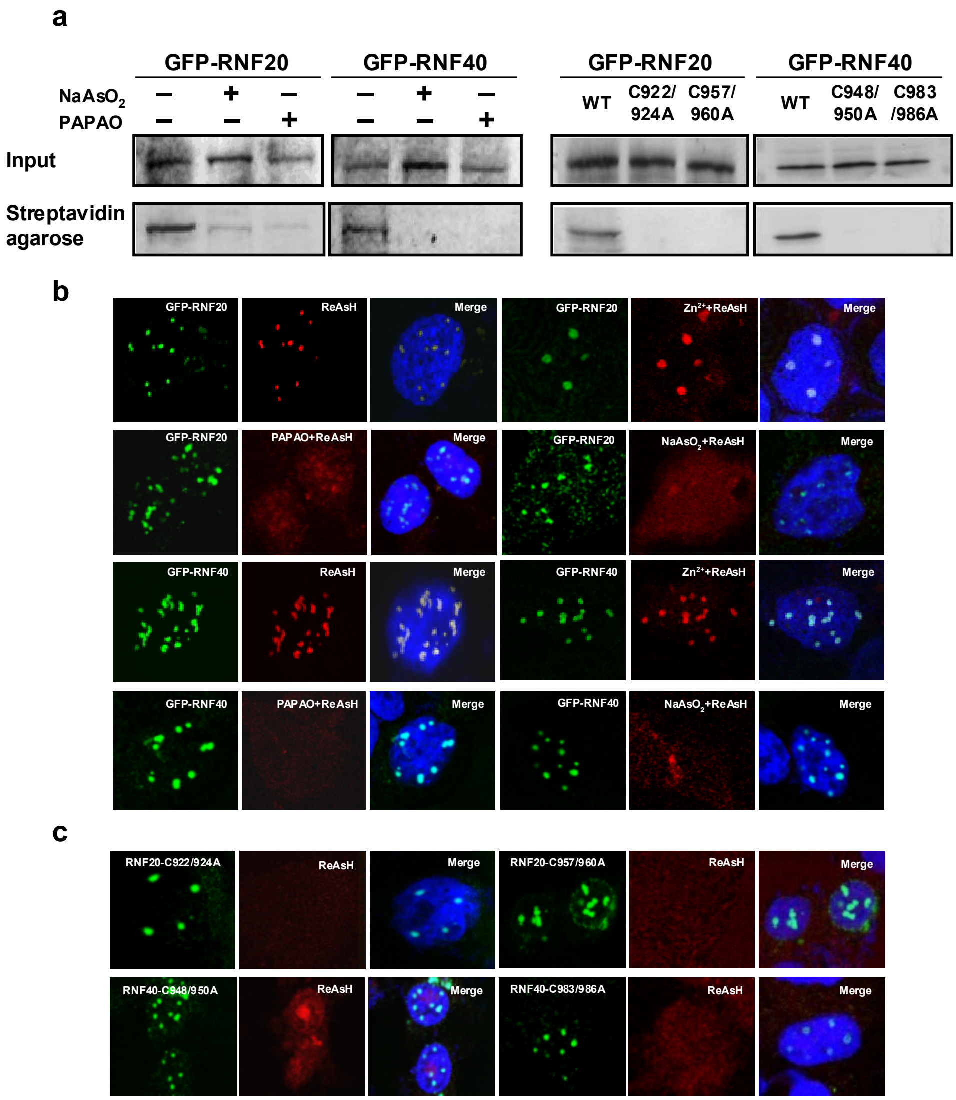


Figure 5.2. Binding between $NaAsO_2$ and the RING finger domains of RNF20 and RNF40 in cells. (a) Streptavidin agarose affinity pull-down assay using biotin-As as a probe to examine the binding between As(III) and RNF20 or RNF40 in cells. Pretreatment of cells with 10 μM $NaAsO_2$ or PAPAO for 1 hr attenuated the binding; (b) Mutations of RING finger cysteines to alanines abrogated the pull-down of RNF20 and RNF40 with the biotin-As probe. (c) Colocalization of As(III)-bearing ReAsH with GFP-RNF20 and GFP-RNF40 in HEK293T cells, and such colocalization is lost in cells pretreated with 10 μM $NaAsO_2$ or PAPAO, but not Zn^{2+}. (d) Mutations of RING finger cysteines to alanines abolished the colocalization of RNF20 or RNF40 with ReAsH.

ubiquitination by using LC-MS/MS (Figure 5.4a and Figure D3) (23). It turned out that a 24-hr treatment with 5 μM $NaAsO_2$ led to a significant diminution in the levels of H2B K120 ubiquitination in multiple cell lines (Figure 5.4a-b).

We next investigated whether arsenite-induced perturbation in H2B K120 ubiquitination confers diminished recruitment of DNA repair proteins to DNA DSB sites. To this end, we employed immunofluorescence microscopy to follow the recruitment of BRCA1 and RAD51 to laser-induced DNA DSB sites. Our results revealed that treatment with arsenite led to marked decreases in the recruitment of these two proteins to laser-induced DNA DSB sites (Figure 5.4). Additionally, the reduction in recruitment of these two proteins is not due to their decreased expression, as real-time PCR results showed that arsenite treatment did not lead to decreased expression of *BRCA1*, *RAD51*, or several other genes important in DNA DSB repair (i.e., *XRCC1*, *Ku70* and *Ku80*. Figure D4).

Diminished histone H2B ubiquitination and the resultant compromised recruitment of DNA repair factors to DNA DSB sites may perturb DSB repair. To test this, we utilized U2OS cells with a chromosomally integrated copy of DR-GFP and EJ5-GFP reporters to examine whether arsenite treatment leads to compromised DNA DSB repair via the HR and NHEJ pathways (24). DR-GFP contains the *SceGFP* cassette where the GFP coding sequence is interrupted by a single recognition site of the rare-cutting I-SceI nuclease, along with a 5' and 3' truncated fragment of GFP (*iGFP*) (24). HR-directed repair of the I-SceI-induced DSB, with the use of *iGFP* as the template, leads to the restoration of functional GFP gene (24). EJ5-GFP carries the pCAGGS promoter

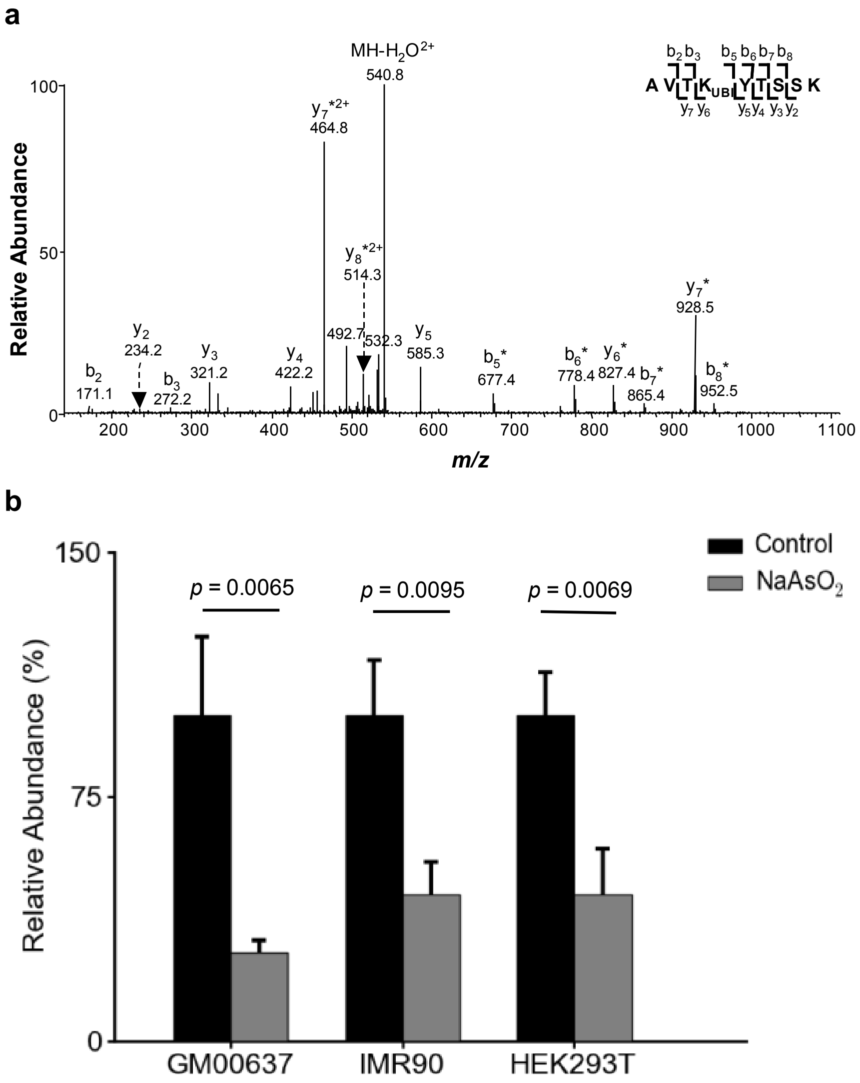


Figure 5.3. NaAsO$_2$ inhibits histone H2B K120 ubiquitination. (a) The product-ion spectrum (MS/MS) of the $[M + 2H]^{2+}$ ion of the tryptic peptide AVTK$_{GG}$VTSSK from histone H2B, where K120 is modified with a diglycine remnant. An asterisk (*) indicates those ions bearing a diglycine moiety. (b) Relative levels of H2B ubiquitination in GM00637, IMR90 and HEK293T cells without or with a 24-hr treatment of 5 μM NaAsO$_2$.

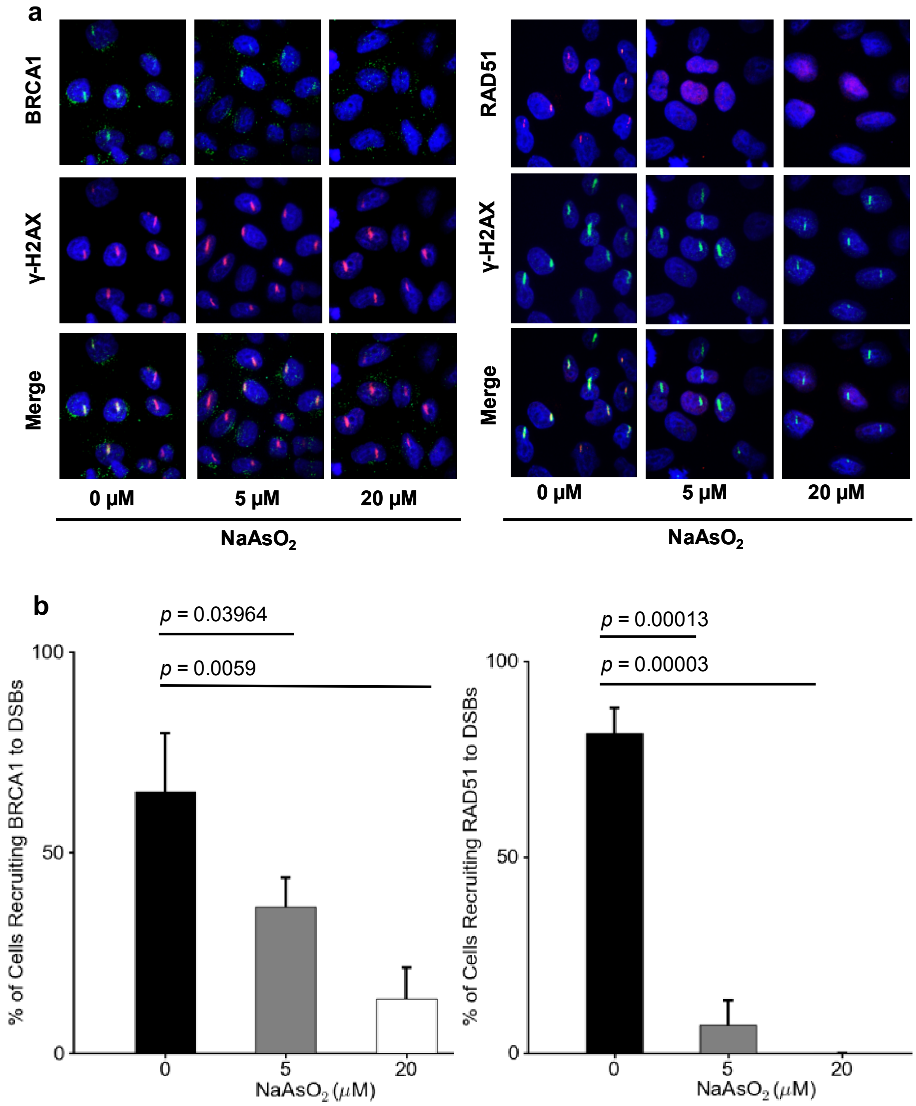


Figure 5.4. NaAsO$_2$ suppresses the recruitment of BRCA1 and RAD51 to DSB sites.(a) Immunofluorescence displaying that the treatment of HeLa cells with NaAsO$_2$ compromises the recruitment of BRCA1 and RAD51 to laser-induced DNA DSB sites. (b) Quantified percentages of cells recruiting BRCA1 and RAD51 to DSB sites upon treatment with 0, 5 and 20 μM of NaAsO$_2$. The *p*-values were calculated by using unpaired two-tailed *t*-test.

separated from the rest of a *GFP* expression cassette flanked by two I-SceI sites; NHEJ of I-SceI-produced distal DSB ends results in the restoration of the *GFP* cassette (24). Flow cytometry-based quantification of GFP-positive U2OS cells after I-SceI transfection facilitates us to quantify the levels of HR and NHEJ in control and arsenite-treated cells. Our results revealed significantly diminished DSB repair via HR (the GFP-positive cells decreased from 5.3% in control cells to 1.5% in arsenite-treated cells) and NHEJ (the GFP-positive cells were reduced from 2.2% to 0.6%) pathways (Figure 5.5a and Figure D5a). It is of note that treatment with $NaAsO_2$ does not alter the transfection efficiency, as assessed by using a GPF expression plasmid (Figure D5b). Additionally, we found that the percentage of GFP-positive cells in U2OS-DR-GFP and U2OS-EJ5-GFP cells pretreated with $NaAsO_2$ but without I-SceI transfection is similar to that of control cells with neither arsenite treatment nor I-SceI transfection (Figure 5.5a). Thus, $NaAsO_2$ does not interfere with the GFP-based reporter assays for monitoring HR and NHEJ activities. Moreover, consistent with the diminished DNA DSB repair, we found, from colony survival assay, that treatment with arsenite renders HeLa cells more sensitive toward treatment with NCS, a radiomimetic drug (Figure 5.5b).

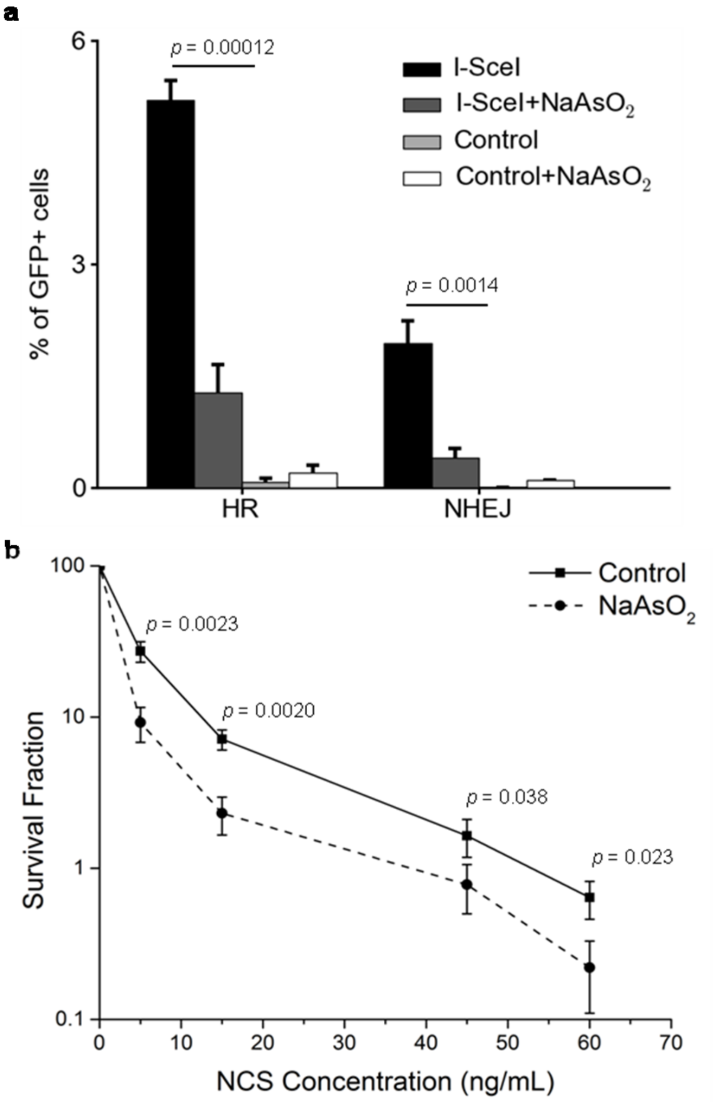


Figure 5.5. Effect of $NaAsO_2$ on DSB repair. (a) Quantitative results showing that $NaAsO_2$ treatment led to reduced DNA DSB repair via the HR and NHEJ pathways. (b) Survival curves based on clonogenic growth of HeLa cells with a treatment of 5 μM $NaAsO_2$ showed increased sensitivity to the radiomimetic drug neocarzinostatin (NCS). The values represent the mean ± S.D. of results obtained from three independent experiments. The *p*-values were calculated by using unpaired two-tailed *t*-test.

CONCLUSIONS

Trivalent arsenic is a known human carcinogen. Although arsenite has been shown to bind and inhibit the activity of zinc finger-harboring DNA repair proteins (i.e., XPA and PARP1) (7-10), this is the first report to demonstrate that arsenite can bind to the RING finger domains of a histone E3 ubiquitin ligase, which perturbs histone epigenetic mark and compromises DNA DSB repair. In this vein, DNA DSBs can emanate from collapse of replication fork, processing of DNA interstrand crosslinks, or following exposure to ionizing radiation (25, 26). DNA DSBs are among the most deleterious types of DNA lesions; failure to repair DSBs can lead to cell death or promote chromosomal rearrangements, which in turn can stimulate malignant transformation (27). Thus, arsenite may exert its carcinogenic effect partly through inhibiting DNA DSB repair. In keeping with its inhibitory effect on DNA repair, we also found that arsenite sensitizes cells toward a radiomimetic drug. Considering that ionizing radiation is widely used in cancer treatment, our results suggest the potential application of arsenite as a sensitizing agent for cancer radiation therapy.

Maintaining an open and biochemically active chromatin fiber conformation is also important for proteins involved in other DNA repair pathways to access damage sites in chromatin. Thus, compromised H2B K120 ubiquitination arising from arsenite exposure may also account for the diminished repair of other types of DNA lesions, including the bulky DNA lesions induced by benzo[*a*]pyrene diolepoxide and UV light.

Our findings that arsenite may promote cancer development through binding to the RING finger domains of a histone E3 ubiquitin ligase is reminiscent of the mechanism underlying arsenite’s role as a chemotherapeutic agent for the treatment of acute promyelocytic leukemia. In the latter case, arsenite binds directly to the RING finger motif of the PML-RARα fusion protein, which triggers eventually the proteasomal degradation of the oncoprotein (12). Viewing the crucial roles of the ubiquitin-proteasome system in physiology and pathophysiology, and the fact that the majority of the E3 ubiquitin-protein ligases harbor RING finger or RING finger-related motifs (13), our study suggests that arsenite may also perturb ubiquitination events mediated by other RING finger E3 ubiquitin ligases. In this context, the reduced efficiency in HR and NHEJ repair may also arise, in part, from the interaction between arsenite and other RING finger E3 ubiquitin ligases that are important in DNA damage response and repair. Along this line, we observed that As(III) can also bind to the RNF8 and RNF168 histone E3 ubiquitin ligases (Data not shown). Thus, arsenite may be considered as a general inhibitor for RING finger E3 ubiquitin ligases.

REFERENCES

1. Ravenscroft, P., Brammer, H., and Richards, K. (2009) *Arsenic pollution: a global synthesis*, Wiley-Blackwell, UK.

2. Chen, C. J., Chuang, Y. C., Lin, T. M., and Wu, H. Y. (1985) Malignant neoplasms among residents of a blackfoot disease-endemic area in Taiwan: high-arsenic artesian well water and cancers. *Cancer Res.* 45, 5895-5899.

3. Smith, A. H., Goycolea, M., Haque, R., and Biggs, M. L. (1998) Marked increase in bladder and lung cancer mortality in a region of Northern Chile due to arsenic in drinking water. *Am. J. Epidemiol.* 147, 660-669.

4. Council, N. R. (1999) Arsenic in the Drinking Water, *National Academy Press, Washington, DC.*

5. Kitchin, K. T. (2001) Recent advances in arsenic carcinogenesis: modes of action, animal model systems, and methylated arsenic metabolites. *Toxicol. Appl. Pharmacol.* 172, 249-261.

6. Kitchin, K. T., and Wallace, K. (2008) The role of protein binding of trivalent arsenicals in arsenic carcinogenesis and toxicity. *J. Inorg. Biochem.* 102, 532-539.

7. Ding, W., Liu, W., Cooper, K. L., Qin, X. J., de Souza Bergo, P. L., Hudson, L. G., and Liu, K. J. (2009) Inhibition of poly(ADP-ribose) polymerase-1 by arsenite interferes with repair of oxidative DNA damage. *J. Biol. Chem.* 284, 6809-6817.

8. Walter, I., Schwerdtle, T., Thuy, C., Parsons, J. L., Dianov, G. L., and Hartwig, A. (2007) Impact of arsenite and its methylated metabolites on PARP-1 activity, PARP-1 gene expression and poly(ADP-ribosyl)ation in cultured human cells. *DNA Repair* 6, 61-70.

9. Hartwig, A., Blessing, H., Schwerdtle, T., and Walter, I. (2003) Modulation of DNA repair processes by arsenic and selenium compounds. *Toxicology* 193, 161-169.

10. Zhou, X., Sun, X., Cooper, K. L., Wang, F., Liu, K. J., and Hudson, L. G. (2011) Arsenite interacts selectively with zinc finger proteins containing C3H1 or C4 motifs. *J. Biol. Chem.* 286, 22855-22863.

11. Shen, Z. X., Chen, G. Q., Ni, J. H., Li, X. S., Xiong, S. M., Qiu, Q. Y., Zhu, J., Tang, W., Sun, G. L., Yang, K. Q., Chen, Y., Zhou, L., Fang, Z. W., Wang, Y. T., Ma, J., Zhang, P., Zhang, T. D., Chen, S. J., Chen, Z., and Wang, Z. Y. (1997) Use of arsenic trioxide (As_2O_3) in the treatment of acute promyelocytic leukemia (APL): II. Clinical efficacy and pharmacokinetics in relapsed patients. *Blood* 89, 3354-3360.

12. Zhang, X. W., Yan, X. J., Zhou, Z. R., Yang, F. F., Wu, Z. Y., Sun, H. B., Liang, W. X., Song, A. X., Lallemand-Breitenbach, V., Jeanne, M., Zhang, Q. Y., Yang, H. Y., Huang, Q. H., Zhou, G. B., Tong, J. H., Zhang, Y., Wu, J. H., Hu, H. Y., de The, H., Chen, S. J., and Chen, Z. (2010) Arsenic trioxide controls the fate of the PML-RARa oncoprotein by directly binding PML. *Science* 328, 240-243.

13. Lipkowitz, S., and Weissman, A. M. (2011) RINGs of good and evil: RING finger ubiquitin ligases at the crossroads of tumour suppression and oncogenesis. *Nat. Rev. Cancer* 11, 629-643.

14. Fierz, B., Chatterjee, C., McGinty, R. K., Bar-Dagan, M., Raleigh, D. P., and Muir, T. W. (2011) Histone H2B ubiquitylation disrupts local and higher-order chromatin compaction. *Nat. Chem. Biol.* 7, 113-119.

15. Moyal, L., Lerenthal, Y., Gana-Weisz, M., Mass, G., So, S., Wang, S. Y., Eppink, B., Chung, Y. M., Shalev, G., Shema, E., Shkedy, D., Smorodinsky, N. I., van Vliet, N., Kuster, B., Mann, M., Ciechanover, A., Dahm-Daphi, J., Kanaar, R., Hu, M. C., Chen, D. J., Oren, M., and Shiloh, Y. (2011) Requirement of ATM-dependent monoubiquitylation of histone H2B for timely repair of DNA double-strand breaks. *Mol. Cell* 41, 529-542.

16. Mailand, N., Bekker-Jensen, S., Faustrup, H., Melander, F., Bartek, J., Lukas, C., and Lukas, J. (2007) RNF8 ubiquitylates histones at DNA double-strand breaks and promotes assembly of repair proteins. *Cell* 131, 887-900.

17. Spuches, A. M., Kruszyna, H. G., Rich, A. M., and Wilcox, D. E. (2005) Thermodynamics of the As(III)-thiol interaction: arsenite and monomethylarsenite complexes with glutathione, dihydrolipoic acid, and other thiol ligands. *Inorg. Chem.* 44, 2964-2972.

18. Zhang, X., Yang, F., Shim, J. Y., Kirk, K. L., Anderson, D. E., and Chen, X. (2007) Identification of arsenic-binding proteins in human breast cancer cells. *Cancer Lett.* 255, 95-106.

19. Kalef, E., and Gitler, C. (1994) Purification of vicinal dithiol-containing proteins by arsenical-based affinity chromatography. *Methods Enzymol.* 233, 395-403.

20. Martin, B. R., Giepmans, B. N., Adams, S. R., and Tsien, R. Y. (2005) Mammalian cell-based optimization of the biarsenical-binding tetracysteine motif for improved fluorescence and affinity. *Nat. Biotechnol.* 23, 1308-1314.

21. Kim, J., Hake, S. B., and Roeder, R. G. (2005) The human homolog of yeast BRE1 functions as a transcriptional coactivator through direct activator interactions. *Mol. Cell* 20, 759-770.

22. Zhu, B., Zheng, Y., Pham, A. D., Mandal, S. S., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2005) Monoubiquitination of human histone H2B: the factors involved and their roles in HOX gene regulation. *Mol. Cell* 20, 601-611.

23. Darwanto, A., Curtis, M. P., Schrag, M., Kirsch, W., Liu, P., Xu, G., Neidigh, J. W., and Zhang, K. (2010) A modified "cross-talk" between histone H2B Lys-120 ubiquitination and H3 Lys-79 methylation. *J. Biol. Chem.* 285, 21868-21876.

24. Pierce, A. J., and Jasin, M. (2005) Measuring recombination proficiency in mouse embryonic stem cells. *Methods Mol. Biol.* 291, 373-384.

25. Ciccia, A., and Elledge, S. J. (2010) The DNA damage response: making it safe to play with knives. *Mol. Cell* 40, 179-204.

26. Price, B. D., and D'Andrea, A. D. (2013) Chromatin remodeling at DNA double-strand breaks. *Cell* 152, 1344-1354.

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

28. Gunn, A., and Stark, J. M. (2012) I-SceI-based assays to examine distinct repair outcomes of mammalian chromosomal double strand breaks. *Methods Mol. Biol.* 920, 379-391.

29. Xiong, L., Ping, L. Y., Yuan, B. F., and Wang, Y. S. (2009) Methyl group migration during the fragmentation of singly charged Ions of trimethyllysine-containing peptides: Precaution of using MS/MS of singly charged ions for interrogating peptide methylation. *J. Am. Soc. Mass. Spectr.* 20, 1172-1181.

30. Ziv, Y., Bielopolski, D., Galanty, Y., Lukas, C., Taya, Y., Schultz, D. C., Lukas, J., Bekker-Jensen, S., Bartek, J., and Shiloh, Y. (2006) Chromatin relaxation in response to DNA double-strand breaks is modulated by a novel ATM- and KAP-1 dependent pathway. *Nat. Cell Biol.* 8, 870-876.

CHAPTER 6

Summary and Future directions

In this dissertation, I discussed several applications of mass spectrometry-based quantitative proteomic analysis. By employing SILAC labeling strategy, I assessed, at the global proteome scale, the alternation of protein expression levels in cells upon treatment with two anti-cancer nucleoside analogues, i.e. 5-Aza-CdR and ^{S}G. Also, I evaluated the proteome changes in human skin fibroblast cells upon $NaAsO_2$ treatment using the same method. Our studies provided some implications on the mechanisms of action of anti-cancer drugs and arsenite.

In Chapter two, we showed that the drug treatment led to significant alternation in expression of 188 proteins. Among them, FDPS and FDFT1, two important enzymes involved in cholesterol biosynthesis, were decreased significantly upon 5-Aza-CdR treatment. In addition, we observed that 5-Aza-CdR induced growth inhibition of Jurkat-T, HL60 and K562 cells, which could be abrogated by external cholesterol addition. Moreover, 5-Aza-CdR led to epigenetic reactivation of *DPP4* gene which is responsible for the decreased levels of FDPS and FDFT1, and the resultant diminished cholesterol biosynthesis induced growth inhibition of leukemic cells. In the future, it will be important to evaluate whether *DPP4* gene is reactivated after administering the drug in leukemic patients, which may provide an important basis for choosing the optimal therapeutic dose for the treatment and for monitoring the clinical efficacy of the drug.

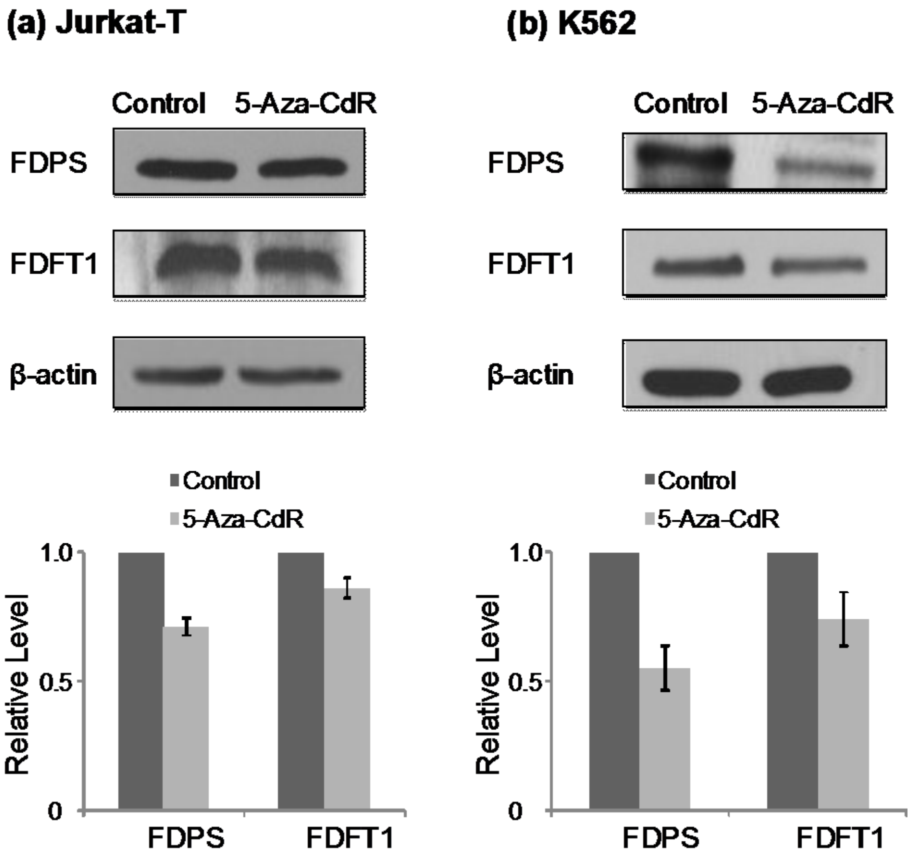
In Chapter three, we found that ^{S}G treatment led to diminished expression of a large number of proteins that are components of the electron transport chain of the mitochondrial respiratory complex and the perturbation of mitochondria dysfunction pathway. In agreement with these observations, flow cytometry data showed the drug-induced loss of active mitochondria in CEM and Jurkat-T human leukemia cells. In addition, we observed elevated generation of oxidatively induced cdA and cdG lesions in ^{S}G-treated cells. Our study suggested that mitochondrial dysfunction may serve as a potential biomarker for monitoring the efficacy of the thiopurine therapy. Future studies may be needed to investigate how these proteins in electron transport chain affect the ROS generation and how ^{S}G exposure led to the selective down-regulation of these proteins.

In Chapter four, we assessed the sodium arsenite-induced perturbation of the entire proteome of GM00637 human skin fibroblast cells. Our results revealed that $NaAsO_2$ treatment led to significant alternations of several canonical pathways including Nrf2-mediated oxidative stress response, pancreatic adenocarcinoma signaling, and cell cycle regulation. Our study painted a more complete picture for $NaAsO_2$-induced alterations of cellular pathways by using an unbiased quantitative proteomic approach.

In Chapter five, we further investigated the role of arsenite on the perturbation of RING finger E3 ubiquitin ligases. We demonstrated that arsenite can bind to the RING finger domains of the RNF20-RNF40 histone E3 ubiquitin ligase, which perturbs histone H2B K120 ubiquitination and leads to compromised DNA DSB repair. We also found that arsenite sensitizes cells toward a radiomimetic drug. Owing the fact that the majority of the E3 ubiquitin-protein ligases harbor RING finger or RING finger-related motifs, future studies can be focused on examining whether this inhibitory mechanism is general by examining arsenite binding with other RING finger proteins.

APPENDIX A. Supporting Information for Chapter 2

“5-Aza-2’-deoxycytidine Induced Growth Inhibition of Leukemia Cells through Modulating Endogenous Cholesterol Biosynthesis”



(a) Jurkat-T

(b) K562

Figure A1.Western blot showing decreased expression of FDPS and FDFT1 in Jurkat-T (a) and K562 (b) cells after a 24-hr treatment with 5 μM 5-Aza-CdR, and the quantification results are displayed in the histograms.

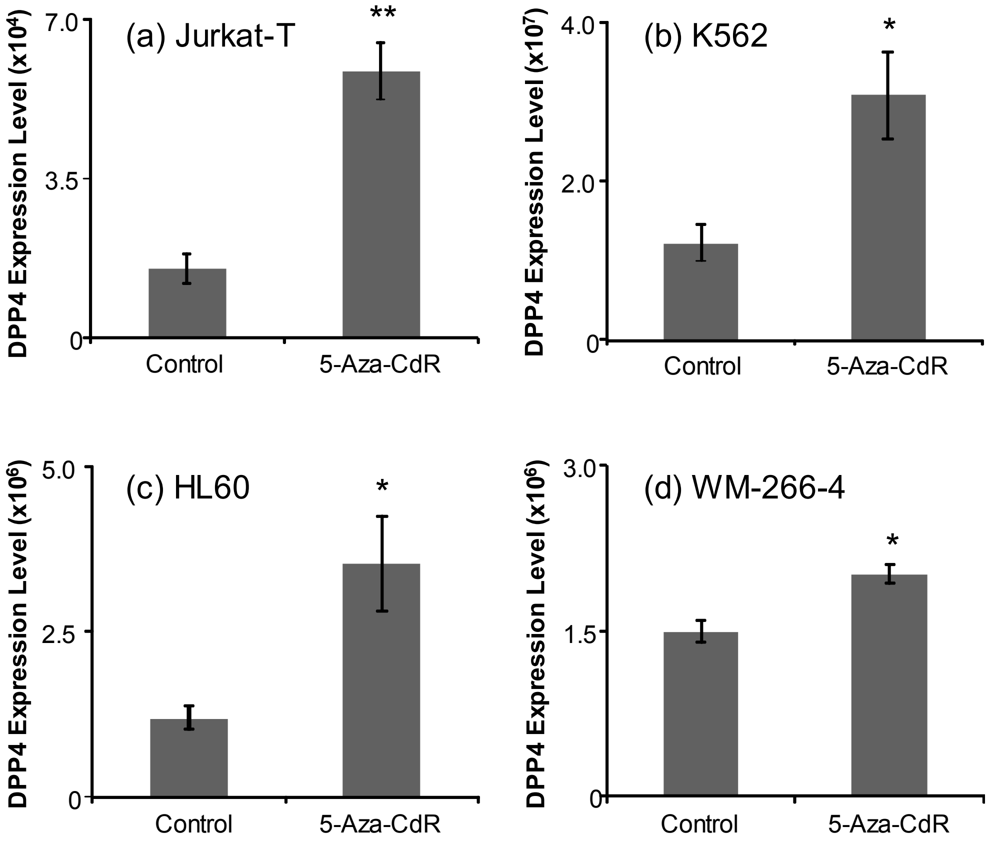


Figure A2. 5-Aza-CdR induced reactivation of DDP4 gene in leukemia and melanoma cells. Bar graphs of DPP4 expression level in Jurkat-T (a), K562 (b), HL60 (c) and WM-266-4 (d) cells after 24 hrs 5-Aza-CdR treatment. The results were obtained from real-time PCR analysis using *GAPDH* gene as reference. The values represent mean ± S.D. of results obtained from three independent experiments. ‘*’, $p < 0.05$; ‘**’, $p < 0.01$. The *p*-values were calculated by using unpaired two-tailed *t*-test.

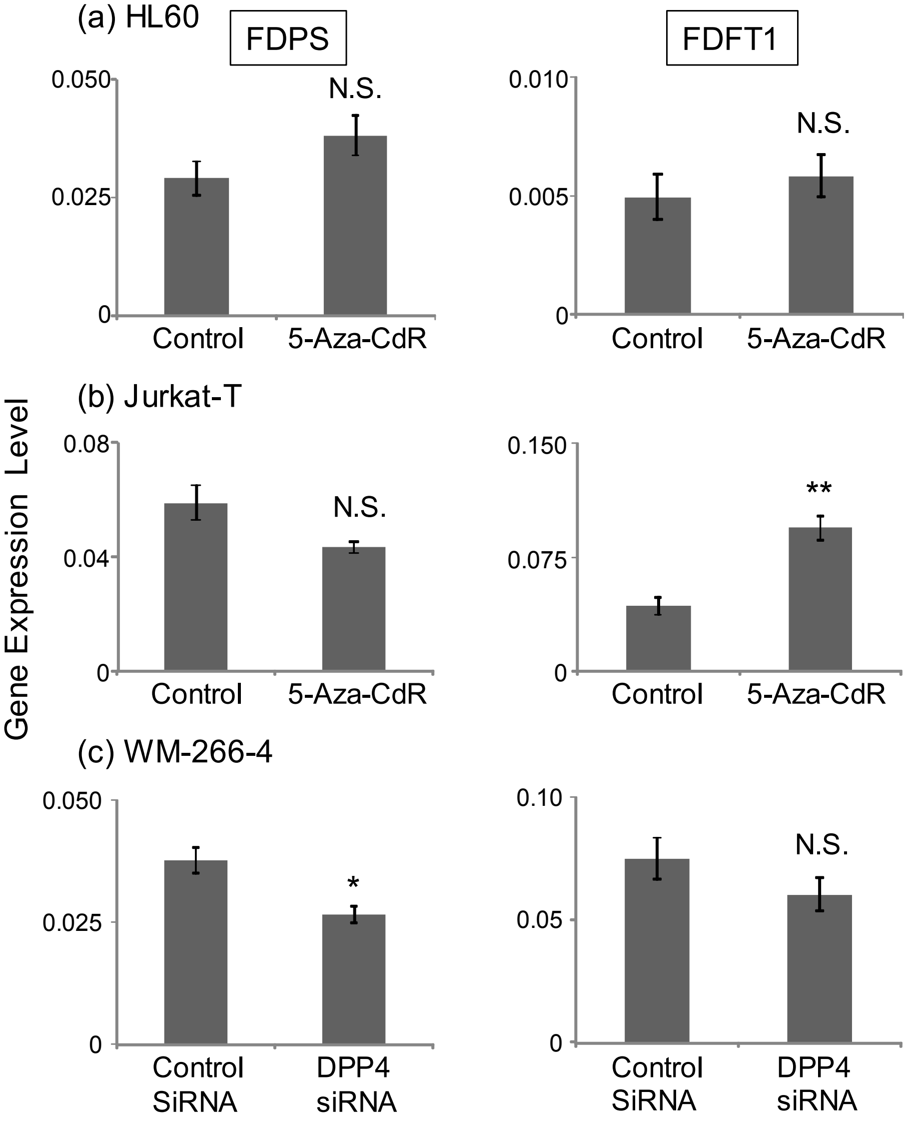


Figure A3. 5-Aza-CdR- and DPP4 siRNA-induced alteration in expression levels of FDPS and FDFT1 proteins occur through a post-transcriptional mechanism. Relative expression levels of *FDPS* (left) and *FDFT1* (right) in HL60 (a) and Jurkat-T (b) cells with and without 5-Aza-CdR (5 μM for 24 hrs) treatment, and in WM-266-4 cells treated with control or DPP4 siRNA (c). The results were obtained from real-time PCR analysis using *GAPDH* gene as reference. The values represent mean ± S.D. of results obtained from three independent experiments. '*', $p < 0.05$; '**', $p < 0.01$. The p-values were calculated by using unpaired two-tailed t-test. "N.S." indicates that the difference was statistically insignificant.

Table A1. A list of proteins which could be quantified in at least two sets of SILAC labeling experiments and whose expressions were changed significantly upon 5-Aza-CdR treatment ("S.D." represents standard deviation).

Protein_ID	**Protein_Name**	**Average Ratio (5-Aza-CdR/Control)**	**S.D.**
IPI00651738	1_2-dihydroxy-3-keto-5-methylthiopentene dioxygenase	1.88	0.56
IPI00010740	100 kDa DNA-pairing protein	2.14	0.88
IPI00413895	130 kDa cis-Golgi matrix protein	1.50	0.07
IPI00026625	155 kDa nucleoporin	0.61	0.07
IPI00074587	40 kDa SR-repressor protein	0.58	0.14
IPI00025329	60S ribosomal protein L19	0.63	0.10
IPI00029731	60S ribosomal protein L35a	0.58	0.05
IPI00414860	60S ribosomal protein L37a	0.50	0.01
IPI00646086	60S ribosomal protein L39	0.56	0.12
IPI00012772	60S ribosomal protein L8	0.67	0.05
IPI00294426	68 kDa TATA-binding protein-associated factor	0.63	0.09
IPI00643071	Abhydrolase domain-containing protein FAM108B1	1.54	0.12
IPI00010348	Acid DNase	1.95	0.56
IPI00784414	Acute-phase response factor	1.67	0.06
IPI00020530	Acyl-coenzyme A thioesterase 13	1.74	0.45
IPI00296441	Adenosine aminohydrolase	0.67	0.02
IPI00465256	Adenylate kinase 3	1.69	0.15
IPI00007426	ADP-ribosylation factor-like protein 6-interacting protein 5	1.51	0.51

IPI00291064	AN1-type zinc finger protein 1	1.61	0.16
IPI00384857	Androgen-regulated protein 2	2.57	0.38
IPI00939163	Antigen NY-CO-25	0.57	0.07
IPI00642971	Antigen NY-CO-4	0.48	0.16
IPI00007334	Apoptotic chromatin condensation inducer in the nucleus	0.65	0.02
IPI00433169	ARHGEF1 protein	1.80	0.35
IPI00007943	Associated molecule with the SH3 domain of STAM	1.56	0.13
IPI00296999	ATP synthase mitochondrial F1 complex assembly factor 2	1.63	0.02
IPI00099995	ATP:riboflavin 5'-phosphotransferase	2.00	0.70
IPI00465160	ATP-binding cassette sub-family F member 3	1.76	0.04
IPI00152535	ATP-dependent helicase CHD5	0.50	0.07
IPI00009328	ATP-dependent RNA helicase DDX48	0.67	0.01
IPI00022254	Autophagy-related protein 3	0.54	0.02
IPI00000643	BAG family molecular chaperone regulator 2	0.46	0.06
IPI00003269	Beta-actin-like protein 2	1.50	0.77
IPI00028955	Block of proliferation 1 protein	0.65	0.24
IPI00845479	Bruno-like protein 3	1.87	0.27
IPI00060715	BTB/POZ domain-containing protein KCTD12	0.33	0.10
IPI00007306	Bucentaur	0.21	0.07
IPI00027180	CAAX prenyl protease 1 homolog	2.69	0.11
IPI00383751	Calregulin	1.50	0.15
IPI00022810	Cathepsin C	1.68	0.14

IPI00011229	Cathepsin D	1.71	0.19
IPI00884105	CD107 antigen-like family member A	1.54	0.33
IPI00009803	CD49 antigen-like family member D	0.53	0.03
IPI00910064	cDNA FLJ50958_ highly similar to Homo sapiens GTPase_ IMAP family member 6 (GIMAP6)_ transcript variant 1_ mRNA	0.65	0.04
IPI00872556	cDNA FLJ52228_ highly similar to Mps one binder kinase activator-like 1A	0.67	0.02
IPI00916503	cDNA FLJ55789_ highly similar to Rattus norvegicus basic leucine zipper and W2 domains 1 (Bzw1)_ mRNA	1.62	0.16
IPI00018871	cDNA FLJ56285_ highly similar to ADP-ribosylation factor-like protein 8B	1.51	0.15
IPI00000875	cDNA FLJ56389_ highly similar to Elongation factor 1-gamma	0.61	0.09
IPI00394838	cDNA FLJ56442_ highly similar to ATP-citrate synthase (EC 2.3.3.8)	0.53	0.03
IPI00910581	cDNA FLJ57960	2.10	0.69
IPI00301579	cDNA FLJ59142_ highly similar to Epididymal secretory protein E1	1.79	0.46
IPI00301936	cDNA FLJ60076_ highly similar to ELAV-like protein 1	0.57	0.19
IPI00867739	cDNA FLJ60533_ highly similar to Kinesin-like protein KIF3A	1.91	0.35
IPI00027035	cDNA FLJ61157_ highly similar to XPA-binding protein 1	0.52	0.08
IPI00871890	cDNA FLJ75056_ highly similar to Homo sapiens phosphatidylinositol binding clathrin assembly protein (PICALM)_ mRNA	2.03	0.28

IPI00073110	cDNA FLJ77052_ highly similar to Homo sapiens FYN binding protein (FYB-120/130) (FYB)_ transcript variant 1_ mRNA	0.53	0.09
IPI00848090	cDNA_ FLJ94534_ highly similar to Homo sapiens capping protein (actin filament)_ gelsolin-like(CAPG)_ mRNA	1.93	0.20
IPI00023101	Cell differentiation protein RCD1 homolog	0.67	0.01
IPI00071189	Cell division cycle 2-like protein kinase 1	0.54	0.09
IPI00018236	Cerebroside sulfate activator protein	1.64	0.47
IPI00328753	CG-1 antigen	0.66	0.03
IPI00029665	Cob(I)alamin adenosyltransferase	1.58	0.39
IPI00177428	Coiled-coil-helix-coiled-coil-helix domain-containing protein 4	0.49	0.06
IPI00180954	Cold-inducible RNA-binding protein	3.46	0.27
IPI00419266	Complex I-B14	1.65	0.13
IPI00479905	Complex I-PDSW	1.75	0.30
IPI00163230	COP9 signalosome complex subunit 6	0.59	0.29
IPI00551062	CTG repeat protein 4a	0.64	0.03
IPI00646304	Cyclophilin B	1.92	0.07
IPI00026519	Cyclophilin F	2.39	0.13
IPI00465315	Cytochrome c	1.64	0.47
IPI00797738	Cytochrome c oxidase subunit 6B1	1.84	0.68
IPI00795769	Cytokine-induced protein of 29 kDa	0.67	0.04
IPI00940000	DEAD box polypeptide 17 isoform p82 variant	0.65	0.18
IPI00171856	Deoxyhypusine dioxygenase	2.31	0.23

IPI00026829	Deoxyhypusine synthase	0.58	0.05
IPI00477535	DNA excision repair protein ERCC-5	1.52	0.03
IPI00027705	DNA primase 58 kDa subunit	0.37	0.04
IPI00549205	DNA repair protein RAD50	1.69	0.23
IPI00296337	DNA-dependent protein kinase catalytic subunit	1.92	0.06
IPI00027808	DNA-directed RNA polymerase II 140 kDa polypeptide	1.51	0.40
IPI00006113	DNA-directed RNA polymerase II subunit I	1.69	0.37
IPI00024163	DNA-directed RNA polymerase III largest subunit	2.20	0.54
IPI00015947	DnaJ homolog subfamily B member 1	0.55	0.06
IPI00184477	Domain E	1.84	0.22
IPI00032401	Duncan disease SH2-protein	0.59	0.07
IPI00412497	Dynein_ light chain_ roadblock-type 1	0.65	0.02
IPI00017184	EH domain-containing protein 1	1.70	0.26
IPI00178440	Elongation factor 1-beta	0.52	0.01
IPI00218245	Ena/vasodilator-stimulated phosphoprotein-like	2.05	0.38
IPI00021570	Endothelial differentiation-related factor 1	0.52	0.20
IPI00299254	Eukaryotic translation initiation factor 5B	0.66	0.09
IPI00006408	eNOS-interacting protein	1.70	0.20
IPI00479786	Far upstream element-binding protein 2	1.98	0.41
IPI00914566	Farnesyl diphosphate synthase	0.61	0.03
IPI00020944	Farnesyl-diphosphate farnesyltransferase	0.48	0.10
IPI00397834	Fermitin family homolog 3	1.70	0.54

IPI00472523	Galectin-9B	0.64	0.07
IPI00061525	Glucosamine 6-phosphate N-acetyltransferase	1.55	0.10
IPI00304692	Glycoprotein p43	1.54	0.10
IPI00304925	Heat shock 70 kDa protein 1/2	0.51	0.00
IPI00215965	Helix-destabilizing protein	1.62	0.37
IPI00011913	Heterogeneous nuclear ribonucleoprotein A0	1.64	0.39
IPI00022228	High density lipoprotein-binding protein	0.63	0.10
IPI00419258	High mobility group protein 1	2.09	0.94
IPI00219097	High mobility group protein B2	2.48	0.50
IPI00217477	High mobility group protein B3	2.41	0.99
IPI00021924	Histone H1	5.48	0.37
IPI00216457	Histone H2A	5.90	1.40
IPI00646240	Histone H2B	4.93	0.13
IPI00465070	Histone H3	2.27	0.34
IPI00453473	Histone H4	3.74	1.34
IPI00915022	hTREX120	0.67	0.07
IPI00419844	Inactive ubiquitin-specific peptidase 39	0.58	0.03
IPI00376199	Interferon regulatory factor 2-binding protein 2	1.97	0.19
IPI00942945	Involucrin	0.44	0.11
IPI00216230	Lamina-associated polypeptide 2_ isoform alpha	2.63	0.53
IPI00000861	LIM and Src homology 3 domain protein 1	1.95	0.59
IPI00297169	Lymphocyte cytosolic protein 2	2.14	0.17
IPI00216172	Lysosomal-associated membrane protein 2_ isoform	1.79	0.22

	CRA_b		
IPI00019997	Mammalian lin-seven protein 3	0.60	0.05
IPI00219365	Membrane-organizing extension spike protein	1.50	0.47
IPI00151462	Microtubule-associated proteins 1A/1B light chain 3 beta 2	1.79	0.23
IPI00396171	Microtubule-associated protein 4	0.57	0.21
IPI00001589	Mitochondrial import inner membrane translocase subunit Tim13	1.57	0.05
IPI00787089	Modulator of non-genomic activity of estrogen receptor	0.62	0.02
IPI00926581	MYH14 variant protein	1.86	0.18
IPI00397730	Myosin phosphatase-targeting subunit 1	1.85	0.33
IPI00056432	NEDD8 carrier protein UBE2F	0.52	0.08
IPI00221172	N-recognin-7	0.55	0.03
IPI00005792	Nuclear poly(A)-binding protein 1	0.62	0.06
IPI00477040	Nucleoporin NUP188 homolog	1.61	0.11
IPI00742943	Nucleoporin Nup43	0.63	0.07
IPI00742682	Nucleoprotein TPR	1.94	0.11
IPI00291313	Nucleoside diphosphate-linked moiety X motif 8_ mitochondrial	0.63	0.01
IPI00031570	Nucleoside triphosphate phosphohydrolase	0.65	0.03
IPI00555902	OCIA domain-containing protein 2	2.27	0.44
IPI00032830	Oligoribonuclease_ mitochondrial	0.44	0.07
IPI00216308	Outer mitochondrial membrane protein porin 1	2.35	0.34
IPI00856098	p180/ribosome receptor	1.72	0.65

IPI00909552	Peptidyl-prolyl cis-trans isomerase	0.58	0.04
IPI00291608	PEST proteolytic signal-containing nuclear protein	0.62	0.03
IPI00449049	Poly [ADP-ribose] polymerase 1	4.43	1.55
IPI00641384	Protein transport protein Sec16A	1.51	0.12
IPI00297261	Protein-tyrosine phosphatase 1B	2.16	0.61
IPI00219825	Prosaposin	1.57	0.13
IPI00514856	Protein NICE-4	0.67	0.48
IPI00411356	Protein SKD2	1.89	0.09
IPI00873472	Protein transport protein Sec24A	1.60	0.10
IPI00472164	Protein WAVE-2	1.78	0.26
IPI00020515	Putative MAPK-activating protein PM26	0.60	0.04
IPI00927150	Putative uncharacterized protein CHCHD3	0.63	0.05
IPI00954006	Putative uncharacterized protein DKFZp667H197	1.69	0.31
IPI00413006	Putative uncharacterized protein ENSP00000368241	2.30	0.72
IPI00644386	Putative uncharacterized protein FUBP1	1.87	0.20
IPI00872817	Putative uncharacterized protein GLIPR2	0.57	0.11
IPI00011118	Putative uncharacterized protein RRM2	0.65	0.45
IPI00942450	Putative uncharacterized protein SFRS15	1.83	0.23
IPI00410162	Putative uncharacterized protein TCEB2	0.62	0.04
IPI00470610	Pyrroline-5-carboxylate reductase 2	0.62	0.09
IPI00016513	Ras-related protein Rab-10	0.55	0.02
IPI00024282	Ras-related protein Rab-8B	2.00	0.14
IPI00915437	ribonuclease H2_ subunit B isoform 2	0.57	0.09

IPI00013871	Ribonucleoside-diphosphate reductase large subunit	0.63	0.02
IPI00304232	Ribosome biogenesis protein WDR12	0.62	0.03
IPI00419473	RNA polymerase B transcription factor 3	2.03	0.48
IPI00024524	RNA-binding protein PNO1	0.59	0.06
IPI00719725	SAPS domain family member 3	1.87	0.67
IPI00218054	Selenoprotein H	1.74	0.23
IPI00017469	Sepiapterin reductase	0.56	0.03
IPI00025318	SH3 domain-binding glutamic acid-rich-like protein	1.50	0.17
IPI00555647	Splicing factor_ arginine/serine-rich 10 (Transformer 2 homolog_ Drosophila) variant	0.53	0.13
IPI00470883	Stromal antigen 2	2.23	0.17
IPI00008569	Synaptobrevin homolog YKT6	1.54	0.22
IPI00790248	Synaptobrevin-2	2.02	0.43
IPI00010438	Synaptosomal-associated protein 23	0.43	0.16
IPI00022934	T-cell receptor T3 delta chain	0.54	0.13
IPI00012923	T-cell surface antigen T3/Leu-4 epsilon chain	0.66	0.42
IPI00170855	Tetratricopeptide repeat protein 19	1.75	0.36
IPI00221108	Thymidylate synthase	0.56	0.06
IPI00719040	Uncharacterized protein C1orf77	0.55	0.04
IPI00017504	Uncharacterized protein C9orf40	0.62	0.21
IPI00166638	UPF0598 protein C8orf82	1.52	0.02
IPI00893918	Valyl-tRNA synthetase	0.54	0.01
IPI00418471	Vimentin	1.89	0.41

IPI00187011	Zinc finger CCCH domain-containing protein 4	1.55	0.14
IPI00062866	Zinc finger CCCH-type antiviral protein 1-like	0.51	0.02
IPI00926625	Zyxin	1.60	0.12

Table A2**.** Pathways perturbed by 5-Aza-CdR treatment, as identified by IPA. The expression ratios (5-Aza-CdR-treated/Untreated, expressed as mean ± S. D.) for individual proteins are listed in the parenthesis.

Pathways	**Protein Name**
Biosynthesis of steroids	Farnesyl-diphosphate farnesyltransferase (0.48 ± 0.10) Farnesyl diphosphate synthase (0.61 ± 0.03)
Granzyme A signaling	Histone H1 (5.48 ± 0.37) High-mobility group B2 (2.48 ± 0.50)
Granzyme B signaling	Cytochrome c (1.64 ± 0.47) Poly(ADP-ribose) polymerase (4.43 ± 1.55)
One-carbon pool by folate	Thymidylate synthease (0.56 ± 0.06) Serine hydroxymethyltransferase 2 (0.53)
Citrate cycle	cDNA FLJ56442_ highly similar to ATP-citrate synthase (0.53 ± 0.03) Isocitrate dehydrogenase [NAD] subunit alpha (0.69 ± 0.03)
Mitochondrial dysfunction	Cytochrome c (1.64±0.47) NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 4 (1.65±0.15) Cytochrome c oxidase subunit 6B1 (1.84±0.68)
Role of JAK2 in hormine-like cytokine signaling	Protein tyrosine phosphatase 1B (2.16±0.61) Signal transducer and activator of transcription 1-alpha/beta (1.22±0.25)

APPENDIX B. Supporting Information for Chapter 3

“6-Thioguanine Induces Mitochondrial Dysfunction and Oxidative DNA Damage in Acute Lymphoblastic Leukemia Cells”

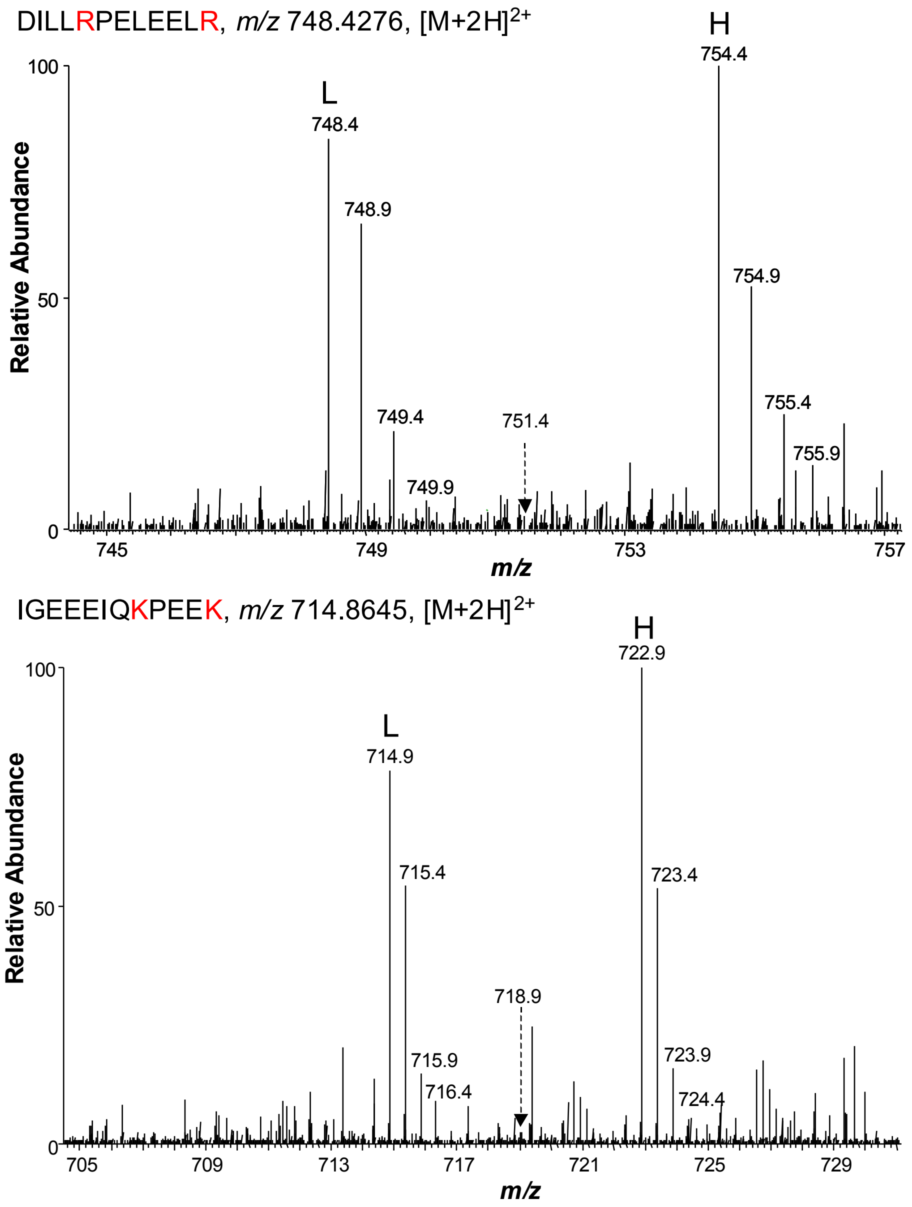


Figure B1. Example mass spectra showing the incorporation of heavy labeled K and R. Shown are the ESI-MS of two trypsin-miscleaved peptides which carry two arginines or two lysines; IGEEEIQKPEEK and DILLRPELEELR are derived from splicing factor 3a (SAP114) and NADH-cytochrome b5 reductase (CYB5R3), respectively. The ions of *m/z* 751.4 and *m/z* 718.9 refer to the putative incompletely labeled peptides (i.e., with a single R or K being heavily labeled), which are present at similar abundance levels as the noise peaks.

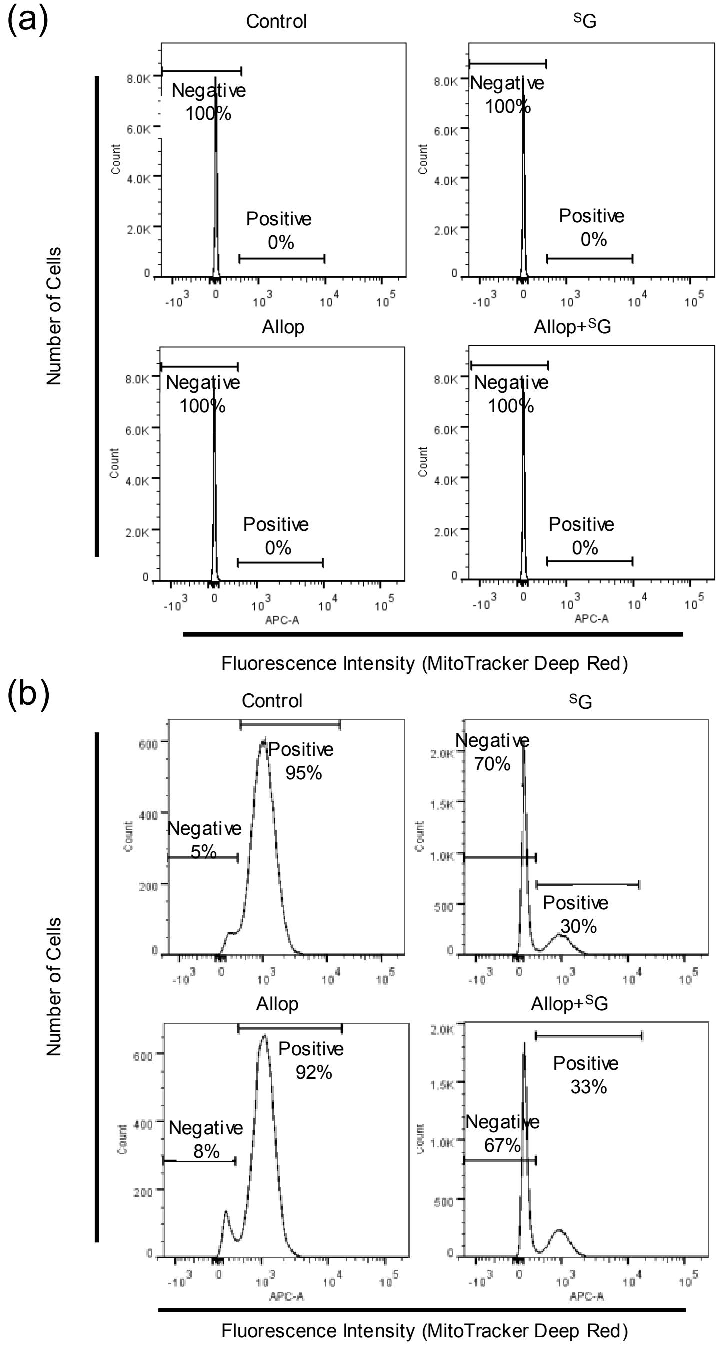


Figure B2. Flow cytometry results for control CEM cells, and CEM cells that were treated for 24 hrs with 3 μM ^{S}G and 500 μM allopurinol, alone or in combination. In (a), no MitoTracker Deep Red was added, whereas MitoTracker Deep Red was added in (b).

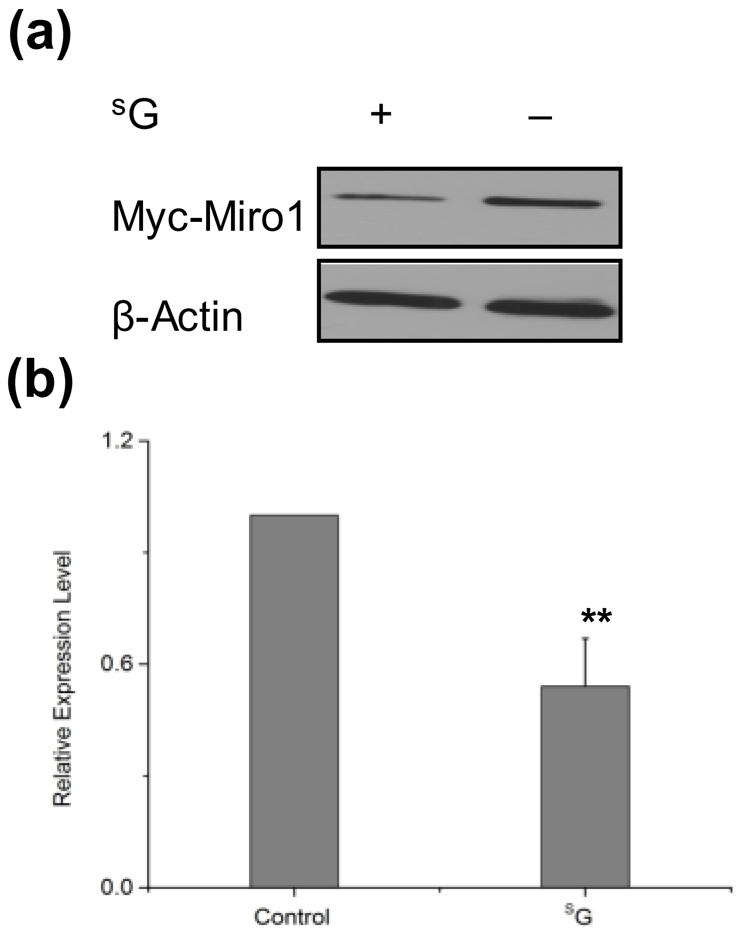


Figure B3. Western blot for monitoring the expression levels of ectopically expressed Myc-Miro1 in HEK293T cells upon treatment with 3 μM ^{S}G (a). β-actin serves as the loading control. Shown in (b) are the quantification data, which represent the mean and S. D. of results from three independent drug treatment and Western blot experiments. "**", $p < 0.01$. The p-value was calculated using two-tailed, unpaired student t test.

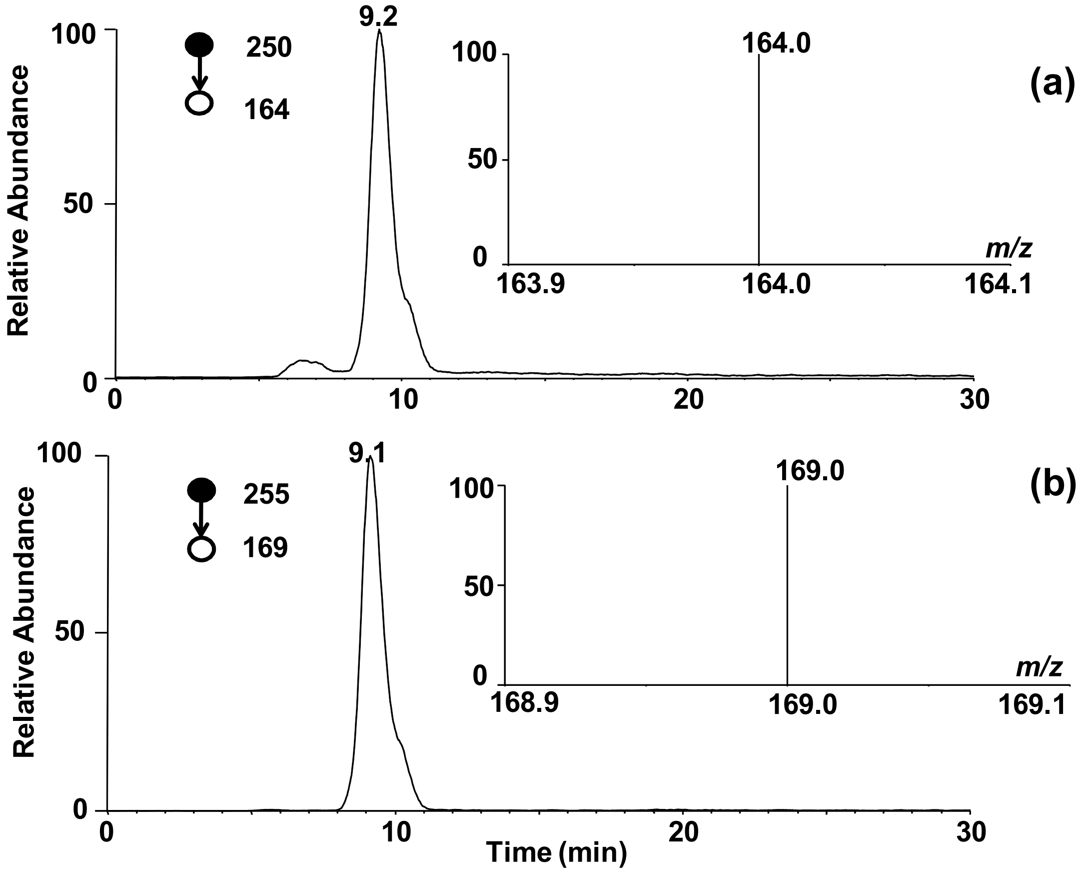


Figure B4. Selected-ion chromatograms for monitoring the transitions of *m/z* 250→164 [(a), for unlabeled *R*-cdA] and *m/z* 255 → 169 [(b), for uniformly ^{15}N-labeled *R*-cdA].

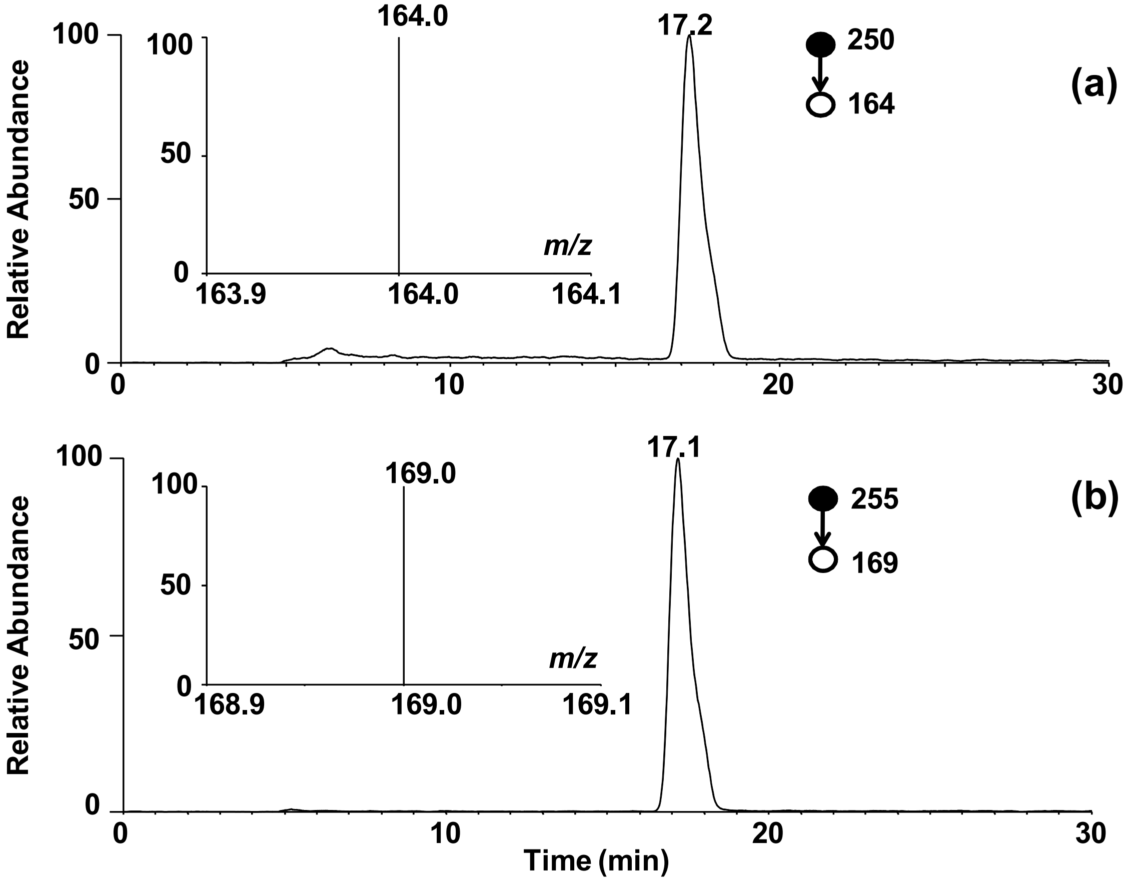


Figure B5. Representative selected-ion chromatograms for monitoring the transitions of *m/z* 250→164 [(a), for unlabeled *S*-cdA] and *m/z* 255 →169 [(b), for uniformly ^{15}N labeled *S*-cdA].

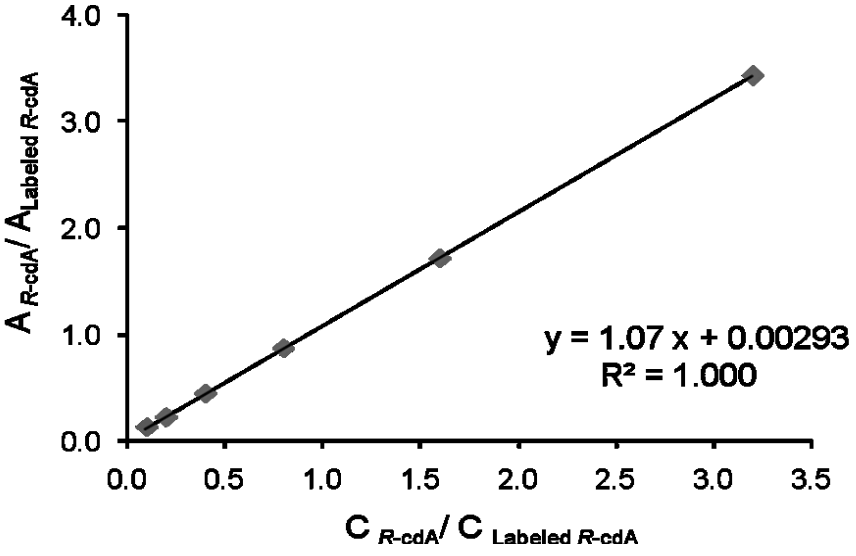


Figure B6. Calibration curve for the quantification of *R*-cdA, where 100 fmol of uniformly ^{15}N-labeled *R*-cdA was added.

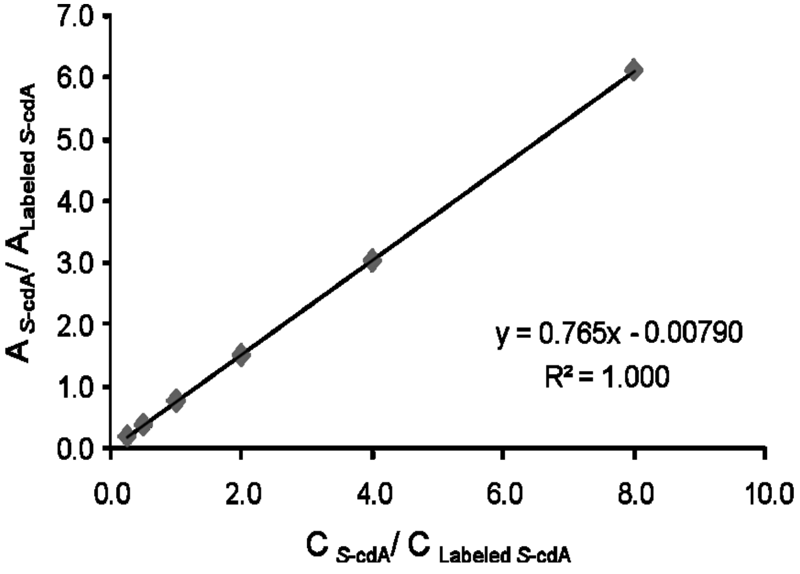


Figure B7. Calibration curve for the quantification of S-cdA, where 40 fmol uniformly ^{15}N-labeled *S*-cdA was added to each sample.

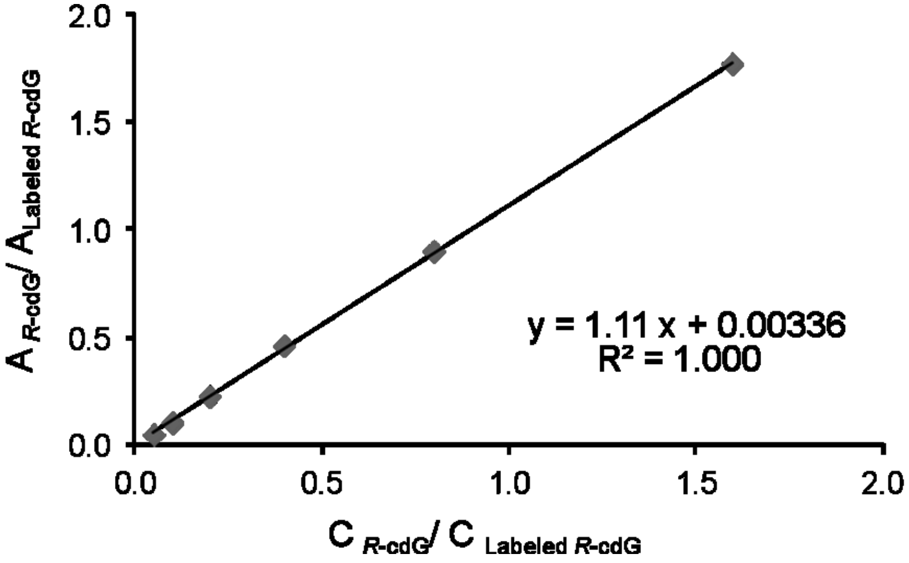


Figure B8. Calibration curve for the quantification of *R*-cdG, where 200 fmol of uniformly ^{15}N-labeled *R*-cdG was added to each sample.

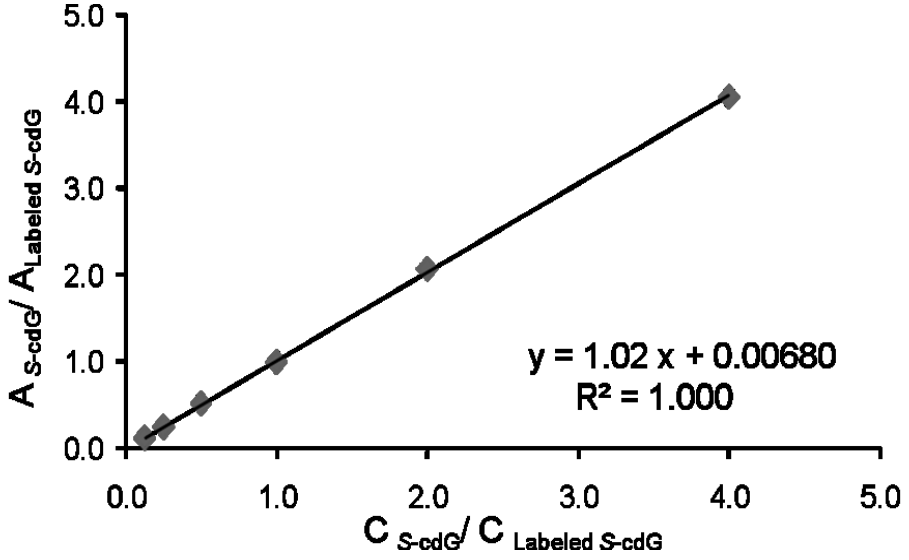


Figure B9. Calibration curve for the quantification of *S*-cdG, where 160 fmol uniformly ^{15}N-labeled *S*-cdG was added to each sample.

Table B1. A list of proteins which could be quantified in both forward and reverse SILAC labeling experiments and whose expressions were changed significantly upon ^{S}G treatment ("S.D." represents standard deviation).

Protein ID	Protein Name	Mean (T/U)	S.D.
P09382	14 kDa laminin-binding protein	0.31	0.01
Q6UXB8-1	Cysteine-rich secretory protein 9	0.36	0.08
Q8WXG9-1	G-protein coupled receptor 98	0.36	0.14
Q8NB37-2	Parkinson disease 7 domain-containing protein 1	0.43	0.11
Q99707	5-methyltetrahydrofolate--homocysteine methyltransferase	0.44	0.05
P10606	Cytochrome c oxidase polypeptide Vb	0.45	0.08
P63313	Thymosin beta-10	0.46	0.09
A6NMH8	Putative uncharacterized protein CD81	0.48	0.49
Q9H8V3-2	Epithelial cell-transforming sequence 2 oncogene	0.48	0.14
B7Z6C0	cDNA FLJ56946, highly similar to Cysteine-rich protein 2	0.49	0.24
O75794	Cell division cycle protein 123 homolog	0.51	0.19
P08590	Cardiac myosin light chain 1	0.51	0.01
P00403	Cytochrome c oxidase polypeptide II	0.51	0.07
O75879	Cytochrome oxidase assembly factor PET112 homolog	0.52	0.14
O95865	Dimethylargininase-2	0.52	0.11
P42575-1	Caspase-2	0.53	0.06
Q9NUV9	GTPase IMAP family member 4	0.53	0.02

Protein ID	Protein Name	Mean (T/U)	S.D.
Q01970	1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase beta-3	0.53	0.58
O75155-1	Cullin-associated and neddylation-dissociated protein 2	0.54	0.21
P09669	Cytochrome c oxidase polypeptide VIc	0.54	0.04
Q86UY6-1	N-acetyltransferase 11	0.55	0.37
Q15154-1	Pericentriolar material 1 protein	0.55	0.12
O95229	ZW10 interactor	0.55	0.06
O43683	Mitotic checkpoint serine/threonine-protein kinase BUB1	0.56	0.17
Q13190-1	Syntaxin-5	0.58	0.08
Q06945	Transcription factor SOX-4	0.58	0.13
P37268	Farnesyl-diphosphate farnesyltransferase	0.58	0.06
Q14191	DNA helicase, RecQ-like type 3	0.59	0.06
A6NED2	RCC1 domain-containing protein 1	0.59	0.08
P30281	G1/S-specific cyclin-D3	0.60	0.02
O95299	NADH dehydrogenase ubiquinone 1 alpha subcomplex	0.60	0.02
Q5TCU8	Tropomyosin 2 (Beta)	0.60	0.09
P63218	Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-5	0.60	0.07
Q6RFH5-1	NOP seven-associated protein 1	0.61	0.15
P84101	Gastric cancer-related protein VRG107	0.61	0.16
C9JG49	Protein tyrosine phosphatase, non-receptor type 6, isoform CRA_b	0.61	0.02

Protein ID	Protein Name	Mean (T/U)	S.D.
O60566-3	MAD3/BUB1-related protein kinase	0.61	0.07
Q9ULT8	E3 ligase for inhibin receptor	0.61	0.11
B4DGI9	cDNA FLJ52589, highly similar to Transcription factor 12	0.62	0.10
Q9BZL1	Ubiquitin-like protein 5	0.62	0.03
P51159-1	GTP-binding protein Ram	0.62	0.07
P56962	Syntaxin-17	0.62	0.03
Q5T7F6	Novel protein	0.63	0.02
P08670	Vimentin	0.63	0.15
P22695	Complex III subunit 2	0.63	0.28
Q96EB1-1	Elongator complex protein 4	0.64	0.08
Q53H12-1	Acylglycerol kinase, mitochondrial	0.64	0.08
Q8NDI1-1	EH domain-binding protein 1	0.64	0.13
Q96EL3	39S ribosomal protein L53, mitochondrial	0.65	0.08
Q9UIC8-2	Leucine carboxyl methyltransferase 1	0.65	0.11
Q9UPP1-1	Histone lysine demethylase PHF8	0.65	0.12
P31930	Complex III subunit 1	0.65	0.22
P46108-1	Adapter molecule crk	0.65	0.34
Q9NP61	ADP-ribosylation factor GTPase-activating protein 3	0.65	0.11
Q96RT1-1	Densin-180-like protein	0.65	0.00
B7Z6Z4	cDNA FLJ56329, highly similar to Myosin light polypeptide 6	0.65	0.03
O94830	DDHD domain-containing protein 2	0.65	0.12

Protein ID	Protein Name	Mean (T/U)	S.D.
Q14CJ1	Amino-terminal enhancer of split	0.65	0.14
Q9BQS8-1	FYVE and coiled-coil domain-containing protein 1	0.66	0.01
Q9Y448	Putative TRAF4-associated factor 1	0.66	0.02
Q9BXK1	Basic transcription element-binding protein 4	0.66	0.15
A2A305	Ubiquitin associated protein 2	0.66	0.15
Q71RC2-4	La ribonucleoprotein domain family member 4	0.66	0.17
P13051-1	Uracil-DNA glycosylase	0.67	0.03
O75146	Huntingtin-interacting protein 12	0.67	0.06
Q9Y3P9-1	GAP and centrosome-associated protein	0.67	0.07
P04818	Thymidylate synthase	0.67	0.04
Q9UJU2-1	Lymphoid enhancer-binding factor 1	0.67	0.05
O15294-1	O-GlcNAc transferase subunit p110	0.67	0.16
P18583-5	Bax antagonist selected in saccharomyces 1	1.50	0.14
Q13510-2	Acid ceramidase	1.54	0.37
O75153	Protein KIAA0664	1.54	0.28
P78330	L-3-phosphoserine phosphatase	1.56	0.16
Q9Y324	rRNA-processing protein FCF1 homolog	1.59	0.01
Q9UNK0	Syntaxin-8	1.60	0.83
Q5T440	Putative transferase C1orf69, mitochondrial	1.63	0.33
B4DHQ3	Phosphoserine aminotransferase	1.63	0.37

Protein ID	Protein Name	Mean (T/U)	S.D.
B1ANH1	Guanylate kinase 1	1.63	0.10
Q5T5C7	Seryl-tRNA synthetase	1.64	0.11
Q15758	Baboon M7 virus receptor	1.65	0.12
O00165-2	HCLS1-associated protein X-1	1.66	0.08
Q9BV57-1	1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase	1.67	0.10
Q9UM00-1	Transmembrane and coiled-coil domain-containing protein 1	1.72	0.44
Q96N66-1	1-acylglycerophosphatidylinositol O-acyltransferase	1.73	0.75
P02794	Cell proliferation-inducing gene 15 protein	1.75	0.30
O60610-1	Diaphanous-related formin-1	1.83	0.38
P50416-1	Carnitine O-palmitoyltransferase 1, liver isoform	1.91	0.25
Q5JTJ3-2	Uncharacterized protein C1orf31	1.92	0.44
Q14657	L antigen family member 3	1.92	0.27
B4DW73	cDNA FLJ50710, highly similar to Phosphoenolpyruvate carboxykinase (GTP), mitochondrial (EC 4.1.1.32)	1.94	0.54
Q01726	Melanocortin receptor 1	2.08	0.38
P53350	Polo-like kinase 1	2.22	0.78
P08243	Asparagine synthetase [glutamine-hydrolyzing]	2.27	0.35
P12236	Adenine nucleotide translocator 3	2.29	1.26
Q9NVP2	Anti-silencing function protein 1 homolog	2.33	0.76

Protein ID	Protein Name	Mean (T/U)	S.D.
	B		
Q9NQH7-1	Aminopeptidase P3	2.44	0.72
Q75MQ0	Putative uncharacterized protein EZH2	2.86	1.43
P05114	High mobility group nucleosome-binding domain-containing protein 1	4.75	1.59
A5JHP3	Dermcidin isoform 2	5.45	2.00
Q8TE01	DERP12 (Dermal papilla derived protein 12)	6.63	0.80

Table B2. Pathways perturbed by ${}^{S}G$ treatment, as identified by IPA. The most statistically significant canonical pathways and molecules identified are listed according to their *p*-value (-Log). The threshold of *p*-value is 0.05.

Ingenuity Canonical Pathways	**-log (p-value)**	**# of Molecules**	**Molecules**
Mitochondrial Dysfunction	7.41E00	13	NDUFA4 (0.64), NDUFV1 (0.48), CPT1 (1.91±0.25), MT-CO2 (0.51±0.07), NDUFA2 (0.63), UQCRC2 (0.63±0.28), UQCRFS1 (0.61), NDUFA10 (0.60±0.02), CYB5A (0.65), UQCRC1 (0.65±0.22), COX5B (0.45±0.08), COX6C (0.54±0.04), RHOT1 (0.58±0.06)
RhoGDI Signaling	4.45E00	8	ARHGAP9 (0.41), MYL6 (0.65±0.03), RHOT1 (0.58±0.06), ARH9 (0.66), GNG5 (0.60±0.07), ARHGEF11 (0.51), GNA14 (1.52), MYL3 (0.51±0.01)
CXCR4 Signaling	4.19E00	9	MYL6 (0.65±0.03), RHOT1 (0.58±0.06), ARH9 (0.66), PLCB3 (0.53±0.58), GNG5 (0.60±0.07), CRK (0.65±0.34), ARHGEF11 (0.51), GNA14 (1.52), MYL3 (0.51±0.01)
Signaling by Rho Family GTPases	3.97E00	9	SEPT8 (2.06), MYL6 (0.65±0.03), RHOT1 (0.58±0.06), ARH9 (0.66), VIM (0.63±0.15), GNG5 (0.60±0.07), ARHGEF11 (0.51), GNA14 (1.52), MYL3 (0.51±0.01)
RhoA Signaling	3.59E00	6	SEPT8 (2.06), ARHGAP9 (0.41), MYL6 (0.65±0.03), ANLN (0.53), ARHGEF11 (0.51), MYL3 (0.51±0.01)
ILK Signaling	3.48E00	8	MYL6 (0.65±0.03), RHOT1 (0.58±0.06), ARH9 (0.66),VIM (0.63±0.15), LEF1 (0.67±0.05), MYH7 (0.30), MYL3 (0.51±0.01), DSP (3.84)
Regulation of Actin-based Motility by Rho	2.97E00	4	MYL6 (0.65±0.03), RHOT1 (0.58±0.06), ARH9 (0.66), MYL3 (0.51±0.01)
Epithelial Adherens Junction Signaling	2.95E00	5	MYL6 (0.65±0.03), LEF1 (0.67±0.05), CRK (0.65±0.34), MYH7 (0.30), MYL3 (0.51±0.01)

Ingenuity Canonical Pathways	**-log (p-value)**	**# of Molecules**	**Molecules**
Thrombin Signaling	2.8E00	8	MYL6 (0.65±0.03), RHOT1 (0.58±0.06), ARH9 (0.66), PLCB3 (0.53±0.58), GNG5 (0.60±0.07), ARHGEF11 (0.51), GNA14 (1.52), MYL3 (0.51±0.01)
Phospholipase C Signaling	2.73E00	9	AHNAK (1.62), MYL6 (0.65±0.03), RHOT1 (0.58±0.06), ARH9 (0.66), CAGA (0.08), PLCB3 (0.53±0.58), GNG5 (0.60±0.07), ARHGEF11 (0.51), MYL3 (0.51±0.01)
Oxidative phosphorylation	2.51E00	11	MYL3 (0.51±0.01), COX5B (0.45±0.08), COX6C (0.54±0.04), NDUFA2 (0.63), NDUFA4 (0.64), NDUFA10 (0.60±0.02), NDUFV1 (0.48), UQCRC1 (0.65±0.22), UQCRC2 (0.63±0.28), ARHGEF11 (0.51), UQCRFS1 (0.61)
Actin Cytoskeleton Signaling	2.46E00	4	MYL6 (0.65±0.03), CRK (0.65±0.34), MYH7 (0.30), MYL3 (0.51±0.01)
Sphingosine-1-phosphate Signaling	2.33E00	4	RHOT1 (0.58±0.06), ARH9 (0.66), CASP2 (0.53±0.06), PLCB3 (0.53±0.58)
Actin Nucleation by ARP-WASP Complex	2.1E00	3	RHOT1 (0.58±0.06), ARH9 (0.66), FNRA (0.66)
Cardiac Hypertrophy Signaling	1.84E00	7	MYL6 (0.65±0.03), RHOT1 (0.58±0.06), ARH9 (0.66), PLCB3 (0.53±0.58), GNG5 (0.60±0.07), GNA14 (1.52), MYL3 (0.51±0.01)
Clathrin-mediated Endocytosis Signaling	1.79E00	4	LYZ (0.03), CAGA (0.08), CTTN (0.64), HIP12 (0.67±0.03)
Mitochondrial L-carnitine Shuttle Pathway	1.63E00	2	CPT1 (1.91±0.25), ACS4 (0.48)

Ingenuity Canonical Pathways	**-log (p-value)**	**# of Molecules**	**Molecules**
LXR/RXR Activation	1.54E00	4	FDFT1 (0.58±0.06), LYZ (0.03), CAGA (0.08), IRF3 (0.66)
GADD45 Signaling	1.50E00	2	CCND3 (0.60±0.02), CDK4 (2.27±0.04)
PAK Signaling	1.44E00	3	MYL6 (0.65±0.03), CAGA (0.08), MYL3 (0.51±0.01)
Cellular effects of sildenafil	1.44E00	4	MYL6 (0.65±0.03), PLCB3 (0.53±0.58), MYH7 (0.30), MYL3 (0.51±0.01)
Production of Nitric Oxide and Reactive Oxygen Species in Macrophages	1.38E00	5	LYZ (0.03), MPO (7.95), RHOT1 (0.58±0.06), ARH9 (0.66), CAGA (0.08)

APPENDIX C. Supporting Information for Chapter 4

“Quantitative Proteomic Analysis Revealed the Arsenite-induced Alteration of Cellular Pathways in Human Skin Fibroblast Cells”

Table C1. A list of proteins which could be quantified in at least two sets of SILAC labeling experiments and whose expressions were changed significantly upon arsenite treatment ("S.D." represents standard deviation).

UniProt	Protein Name	Mean (T/U)	S.D.
B2R960	cDNA, FLJ94230, highly similar to Homo sapiens thioredoxin-like 1 (TXNL1), mRNA	0.13	0.04
B4E0L0	cDNA FLJ54030, highly similar to Polymerase delta-interacting protein 3	0.16	0.08
P63218	Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-5	0.31	0.04
Q8WWX9	Selenoprotein M	0.33	0.28
Q9BQE4	Selenoprotein S	0.37	0.08
Q9UJ14-1	Gamma-glutamyltransferase 7	0.42	0.35
P02452	Alpha-1 type I collagen	0.45	0.05
O95810	Cavin-2	0.45	0.03
Q9Y263	Phospholipase A-2-activating protein	0.46	0.13
Q96ST3	Histone deacetylase complex subunit Sin3a	0.48	0.38
B7Z4S9	BUD31 homolog (Yeast), isoform CRA_b	0.48	0.05
Q9Y6D0	Selenoprotein K	0.49	0.08
Q16850-1	CYPLI	0.50	0.13
Q96CX6	Leucine-rich repeat-containing protein 58	0.50	0.04
Q9NUJ3	T-complex protein 11-like protein 1	0.51	0.03
Q9BW72	HIG1 domain family member 2A	0.54	0.23
Q9BW61	DET1- and DDB1-associated protein 1	0.55	0.09
Q8N806	E3 ubiquitin-protein ligase UBR7	0.56	0.06

Q9UBB4	Ataxin-10	0.57	0.05
Q7Z3T8-1	Endofin	0.58	0.58
P56556	Complex I-B14	0.58	0.01
Q15506	Cancer/testis antigen 22	0.58	0.03
Q16352	66 kDa neurofilament protein	0.59	0.74
Q8TF05-1	Serine/threonine-protein phosphatase 4 regulatory subunit 1	0.60	0.28
Q9P2B4	CTTNBP2 N-terminal-like protein	0.60	0.24
Q01628	Interferon-induced transmembrane protein 3	0.60	0.11
P98170	Baculoviral IAP repeat-containing protein 4	0.61	0.15
Q8IV08	Choline phosphatase 3	0.61	0.16
O15460-2	Procollagen-proline,2-oxoglutarate-4-dioxygenase subunit alpha-2	0.61	0.11
P62330	ADP-ribosylation factor 6	0.61	0.60
Q13907-2	Isopentenyl pyrophosphate isomerase 1	0.61	0.45
P63027	Synaptobrevin-2	0.61	0.03
Q96E14-1	BLM-associated protein of 18 kDa	0.62	0.09
O75886-1	Hrs-binding protein	0.62	0.17
Q9BY32	Inosine triphosphate pyrophosphatase	0.62	0.04
Q9GZU8	NEFA-interacting nuclear protein NIP30	0.62	0.31
P13674-2	Procollagen-proline,2-oxoglutarate-4-dioxygenase subunit alpha-1	0.63	0.05
Q9Y3D0	Protein FAM96B	0.63	0.06
Q8N128-2	Protein FAM177A1	0.63	0.02
Q96P48-6	Arf-GAP with Rho-GAP domain, ANK repeat and PH domain-containing protein 1	0.63	0.06

Q9NQE9	Histidine triad nucleotide-binding protein 3	0.63	0.05
P13489	Placental ribonuclease inhibitor	0.63	0.02
NDUFB1	Complex I-MNLL	0.63	0.14
C9IZA2	Putative uncharacterized protein TPM1	0.63	0.09
Q9BSG0	Protease-associated domain-containing protein of 21 kDa	0.63	0.04
P27986-4	Phosphatidylinositol 3-kinase 85 kDa regulatory subunit alpha	0.63	0.34
Q8IZQ5	Selenoprotein H	0.64	0.09
Q9H1E3-1	Nuclear ubiquitous casein and cyclin-dependent kinases substrate	0.64	0.22
B4DTY1	cDNA FLJ51091, highly similar to 65 kDa Yes-associated protein	0.64	0.07
O75787	ATPase H(+)-transporting lysosomal accessory protein 2	0.64	0.37
Q96BR5	Hcp beta-lactamase-like protein C1orf163	0.64	0.14
P12109	Collagen alpha-1(VI) chain	0.64	0.17
P36969-1	Glutathione peroxidase 4	0.64	0.06
Q5SSJ5-1	Heterochromatin protein 1-binding protein 3	0.65	0.12
P35658-5	214 kDa nucleoporin	0.65	0.24
Q9H089	Large subunit GTPase 1 homolog	0.65	0.47
O43181	Complex I-18 kDa	0.65	0.06
B4DJA0	NADH-ubiquinone oxidoreductase 75 kDa subunit	0.66	0.07
Q9NYJ1-2	Coiled-coil-helix-coiled-coil-helix domain-containing protein 8	0.66	0.14
A0PK02	PLXNB2 protein	0.66	0.08

O15355	Protein phosphatase 1C	0.66	0.03
Q15750	Mitogen-activated protein kinase kinase kinase 7-interacting protein 1	0.67	0.11
A3F719	MHC class I antigen	0.67	0.12
Q9UNW1-1	Inositol (1,3,4,5)-tetrakisphosphate 3-phosphatase	0.67	0.13
Q9Y608-1	Leucine-rich repeat flightless-interacting protein 2	1.50	0.63
O00622	CCN family member 1	1.51	0.06
P50454	47 kDa heat shock protein	1.52	0.17
P22626-1	Heterogeneous nuclear ribonucleoproteins A2/B1	1.52	0.24
P02794	Ferritin	1.52	0.02
Q9Y314	eNOS-interacting protein	1.52	0.46
P33981	Dual specificity protein kinase TTK	1.53	0.66
P25685	DnaJ homolog subfamily B member 1	1.53	0.17
B4DVB8	cDNA FLJ60076, highly similar to ELAV-like protein 1	1.59	0.26
Q9UKM9-1	Autoantigen p542	1.61	0.35
P51116	Fragile X mental retardation syndrome-related protein 2	1.61	0.36
B4E3T4	cDNA FLJ61582, highly similar to RNA-binding protein with multiple splicing	1.62	0.22
O75190-1	DnaJ homolog subfamily B member 6	1.63	0.59
O95159	Zinc finger protein MCG4	1.64	0.10
P38159	Glycoprotein p43	1.65	0.10
P52272-1	Heterogeneous nuclear ribonucleoprotein M	1.66	0.07
Q12929	Epidermal growth factor receptor kinase substrate 8	1.67	0.94
A8K4L9	cDNA FLJ75459	1.69	0.00

Q6ZSA3	cDNA FLJ45695 fis, clone FEBRA2013570, highly similar to 2-oxoisovalerate dehydrogenase alpha subunit, mitochondrial (EC 1.2.4.4)	1.71	0.03
Q9NS18-2	Glutaredoxin-2, mitochondrial	1.72	0.60
Q9BYN0	Sulfiredoxin-1	1.72	0.12
Q9NQW6-1	Actin-binding protein anillin	1.73	0.64
Q9P270	SLAIN motif-containing protein 2	1.77	0.07
O95164	Membrane-anchored ubiquitin-fold protein	1.80	0.12
B7ZM99	Methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1-like	1.82	0.49
A6NCE7	Microtubule-associated proteins 1A/1B light chain 3 beta 2	1.83	0.53
Q9UBU9	mRNA export factor TAP	1.84	1.16
Q96T37-1	One-twenty two protein 1	1.88	0.83
C9JJP5	Putative uncharacterized protein TFG	1.89	0.26
Q6PL18-1	AAA nuclear coregulator cancer-associated protein	1.94	0.01
P14866	Heterogeneous nuclear ribonucleoprotein L	1.96	0.21
O14519	Cyclin-dependent kinase 2-associated protein 1	1.99	0.39
P48681	Nestin	2.01	0.50
Q9NY12	H/ACA ribonucleoprotein complex subunit 1	2.03	0.01
Q969T9	WW domain-binding protein 2	2.04	0.34
Q6QNY1-1	Biogenesis of lysosome-related organelles complex 1 subunit 2	2.07	1.63
B3KY41	cDNA FLJ46786 fis, clone TRACH3028837, highly similar to Homo sapiens smoothelin (SMTN), transcript variant 2, mRNA	2.07	0.29
Q6ZRR7-1	Leucine-rich repeat-containing protein 9	2.07	0.63

A4D110	Putative uncharacterized protein	2.16	0.40
P48507	Glutamate–cysteine ligase modifier subunits	2.18	0.29
Q99541	Adipophilin	2.22	0.99
P08107	Heat shock 70 kDa protein 1/2	2.23	0.47
P05120	Monocyte Arg-serpin	2.24	0.31
Q9NZ23	Drug-sensitive protein 1	2.25	1.65
P0C0S8	Histone H2A type 1	2.26	1.63
B3KTT5	highly similar to heat shock 70 kDa protein 1	2.30	0.56
Q6FI13	Histone H2A type 2-A	2.38	1.67
A8MWH4	Metallothionein F	2.43	0.27
P41743	Atypical protein kinase C-lambda/iota	2.44	1.74
Q5BKZ1-1	Zinc finger protein 326	2.44	1.51
Q9UNI6	Dual specificity protein phosphatase 12	2.50	1.10
O95630	Associated molecule with the SH3 domain of STAM	2.60	2.37
Q9UBF6-1	CKII beta-binding protein 1	2.69	0.28
O60942-1	GTP--RNA guanylyltransferase	2.83	0.88
P62491	Ras-related protein Rab-11A	2.91	1.69
P02795	Metallothionein-2	2.91	0.15
P80294	Metallothionein-0	2.97	0.03
Q9NWS8-1	Required for meiotic nuclear division protein 1 homolog	3.36	0.36
Q9UDY8-1	MALT lymphoma-associated translocation	3.43	2.10
P36776	Lon protease homolog, mitochondrial	4.41	1.13
O94808	D-fructose-6-phosphate amidotransferase 2	4.77	3.18

Q6ZNL6	FYVE, RhoGEF and PH domain-containing protein 5	5.88	2.69
Q86YZ3	Hornerin	6.21	2.59
P12883	Myosin heavy chain 7	9.84	1.67
A5JHP3	Dermcidin isoform 2	10.80	5.95
P09601	Heme oxygenase 1	16.54	2.58

Table C2. A list of proteins quantified through SRM analysis on a triple quadrupole instrument. The protein expression ratios (Arsenite-treated/control, i.e., T/U) from both TSQ and Orbitrap-Velos were listed ("S.D." represents standard deviation).

ProteinName	**TSQ Ratio (T/U)**	**Orbitrap Ratio (T/U) ± S.D.**
Selenoprotein S	0.37	0.37±0.08
Selenoprotein M	0.55	0.33±0.28
E3 ubiquitin-protein ligase UBR3	0.62	0.66
Selenoprotein H	0.66	0.64±0.09
E3 ubiquitin-protein ligase UBR5	0.67	0.85
15 kDa selenoprotein	0.73	0.71±0.05
KRAB-associated protein 1	0.79	1.03±0.12
95 kDa retinoblastoma-associated protein	0.84	0.68±0.39
Zinc finger protein 185 variant 721	0.84	0.85±0.07
Fragile X mental retardation 1 protein	0.85	0.75±0.10
FYVE and coiled-coil domain-containing protein 1	0.85	0.88±0.02
Histone deacetylase 2	0.86	0.9±0.03
Histone deacetylase complex subunit Sin3a	0.88	0.48±0.38
Heme oxygenase 2	0.89	0.97±0.01
Zinc finger protein 593	0.94	1.70
RING finger protein 185	0.94	0.94±0.14
Histone acetyltransferase type B subunit 2	0.94	0.95±0.06
CKII beta-binding protein 1	0.95	2.69±0.28

Histone acetyltransferase type B catalytic subunit	0.95	1.02±0.08
Protein ZYP	0.96	0.99±0.12
RING finger protein 160	0.98	0.99±0.14
Selenoprotein Z	0.98	0.93±0.06
DnaJ homolog subfamily B member 6	1.00	1.62±0.59
Zinc finger protein 259	1.07	1.02±0.01
HBV pX-associated protein 8	1.10	1.88±0.61
M1G-methyltransferase	1.10	1.11±0.51
ADP-ribosylation factor domain-containing protein 1	1.12	1.05±0.08
E3 ubiquitin-protein ligase BRE1A	1.12	1.02±0.14
RING finger protein 114	1.12	1.03±0.09
E3 ubiquitin/ISG15 ligase TRIM25	1.13	0.95±0.18
E3 ubiquitin-protein ligase KCMF1	1.16	1.06±0.31
Zinc finger protein 207	1.19	1.07±0.12
BAT2 domain-containing protein 1	1.21	1.32±0.17
Fragile X mental retardation syndrome-related protein 2	1.23	1.6±0.36
Fanconi anemia group I protein	1.25	0.73±0.06
Cellular nucleic acid binding protein alpha variant 1	1.26	1.08±0.03
Fragile X mental retardation syndrome-related protein 1	1.31	1.3±0.36
DnaJ homolog subfamily B member 4	1.36	1.46±0.17
DnaJ homolog subfamily A member 1	1.41	1.38±0.13

DnaJ homolog subfamily B member 1	1.43	1.53±0.17
47 kDa heat shock protein	1.46	1.51±0.17
Cell proliferation-inducing gene 15 protein	1.51	1.52±0.02
Zinc finger protein MCG4	1.54	1.63±0.10
Metallothionein	1.72	2.42±0.27
Zinc finger protein 326	1.97	2.43±1.51
Gamma-ECS regulatory subunit	2.58	2.18±0.29
Metallothionein-1X	2.82	2.42±0.27
Metallothionein-0	2.96	2.96±0.03
HBV X protein up-regulated gene 4 protein	3.88	1.63
Heme oxygenase 1	53.48	16.5±2.58

Table C3. Pathways perturbed by arsenite treatment, as identified by IPA. The most statistically significant canonical pathways and molecules identified are listed according to their *p*-values. The threshold of *p*-value is 0.05.

Ingenuity Canonical Pathways	**- log(p-value)**	**# of Mol.**	**Molecules**
Pancreatic Adenocarcinoma Signaling	3.38E+00	7	HMOX1 (16.54±2.58), RAF1 (0.46), PLD3 (0.61±0.16), PIK3R1 (0.63±0.34), CDKN1A (0.43), ABL1 (5.57), SIN3A (0.48±0.38)
NRF2-mediated Oxidative Stress Response	3.34E+00	9	HMOX1 (16.54+2.58), PRKCI (2.44±1.74), PIK3R1 (0.63±0.34), DNAJB6 (1.63±0.59), DNAJB1 (1.53±0.17), GCLM (2.18±0.29), FTH1 (1.52±0.02), HSPA1 (2.23+0.47), GCLC (1.91)
Chronic Myeloid Leukemia Signaling	2.94E+00	6	RAF1 (0.46), PIK3R1 (0.63±0.34), CDKN1A (0.43), HDAC1 (0.67), ABL1 (5.57), SIN3A (0.48±0.38)
Fcγ Receptor-mediated Phagocytosis in Macrophages and Monocytes	2.91E+00	6	HMOX1 (16.54+2.58), ARF6 (0.61±0.60), PLD3 (0.61±0.16), PRKCI (2.44±1.74), PIK3R1 (0.63+0.34), RAB11A (2.91±1.69)
Glioma Signaling	2.89E+00	6	RAF1 (0.46), PRKCI (2.44±1.74), PIK3R1 (0.63±0.34), CDKN1A (0.43), ABL1 (5.57), SIN3A (0.48±0.38)
Gαq Signaling	2.58E+00	7	HMOX1 (16.54+2.58), RAF1 (0.46), PLD3 (0.61±0.16), PRKCI (2.44±1.74), RND3 (1.60), PIK3R1 (0.63±0.34), GNG5 (0.31±0.04)
Cyclins and Cell Cycle Regulation	2.52E+00	5	RAF1 (0.46), CDKN1A (0.43), HDAC1 (0.67), ABL1 (5.57), SIN3A (0.48±0.38)
Aldosterone Signaling in	2.49E+00	6	PLCD1 (0.63), RAF1 (0.46), PRKCI (2.44±1.74), PIK3R1 (0.63±0.34),

Epithelial Cells			DNAJB6 (1.63±0.59), DNAJB1 (1.53±0.17)
Prostate Cancer Signaling	2.45E+00	5	RAF1 (0.46), PIK3R1 (0.63±0.34), CDKN1A (0.43), ABL1 (5.57), SIN3A (0.48±0.38)
PI3K Signaling in B Lymphocytes	2.30E+00	6	ABL1 (5.57), MALT1 (3.43±2.10), PIK3R1 (0.63±0.34), PLCD1 (0.63), PRKCI (2.44±1.74), RAF1 (0.46)
Endothelin-1 Signaling	2.25E+00	7	PLCD1 (0.63), HMOX1 (16.54±2.58), RAF1 (0.46), PLD3 (0.61±0.16), PRKCI (2.44±1.74), PIK3R1 (0.63±0.34), CASP14 (0.34)
ErbB4 Signaling	2.19E+00	4	RAF1 (0.46), PRKCI (2.44±1.74), YAP1 (0.64±0.07), PIK3R1 (0.63±0.34)
Cell Cycle: G1/S Checkpoint Regulation	2.17E+00	4	CDKN1A (0.43), HDAC1 (0.67), ABL1 (5.57), SIN3A (0.48±0.38)
Telomerase Signaling	2.11E+00	5	RAF1 (0.46), PIK3R1 (0.63±0.34), CDKN1A (0.43), HDAC1 (0.67), ABL1 (5.57)
Clathrin-mediated Endocytosis Signaling	2.04E+00	7	LYZ (0.41), ARF6 (0.61±0.60), APOB (0.58), PIK3R1 (0.63±0.34), RAB11A (2.91±1.69), S100A8 (11.02), ARH (0.67)
IL-8 Signaling	2.04E+00	7	HMOX1 (16.54+2.58), RAF1 (0.46), PLD3 (0.61±0.16), PRKCI (2.44±1.74), RND3 (1.60), PIK3R1 (0.63±0.34), GNG5 (0.31±0.04)

Table C4. Biological functions perturbed by arsenite treatment, as identified by IPA. The most statistically significant biological functions are listed according to an FDR ≤ 0.05 (-log *P*-value = 1.33) to minimize false positives among the significantly-enriched functions.

TOP Biological Functions by IPA analysis		
Diseases and Disorders	**p-value**	**# Molecules**
Cancer	7.02E-05	55
Infectious Disease	2.39E-04	16
Respiratory Disease	2.39E-04	4
Cardiovascular Disease	3.96E-04	24
Developmetal Disorder	7.11E-04	31
Molecular and Cellular Functions	**p-value**	**# Molecules**
Cell Morphology	2.53E-05	20
Cellular Assembly and Organizaton	2.53E-05	41
Cellular Function and Maintenance	2.53E-05	30
Cell Cycle	9.38E-04	22
Cell Death and Survival	1.63E-03	27
Physiological System Development and Function	**p-value**	**# Molecules**
Cardiovascular System Develoment and Function	2.39E-04	15
Organismal Development	2.39E-04	10
Tissue Morphology	2.39E-04	11
Embryonic Development	7.30E-04	12
Organ Development	7.30E-04	7

APPENDIX D. Supporting Information for Chapter 5

“Arsenite Binds to the RING Finger Domains of RNF20-RNF40 Histone E3 Ubiquitin Ligase and Inhibits DNA Double-strand Break Repair”

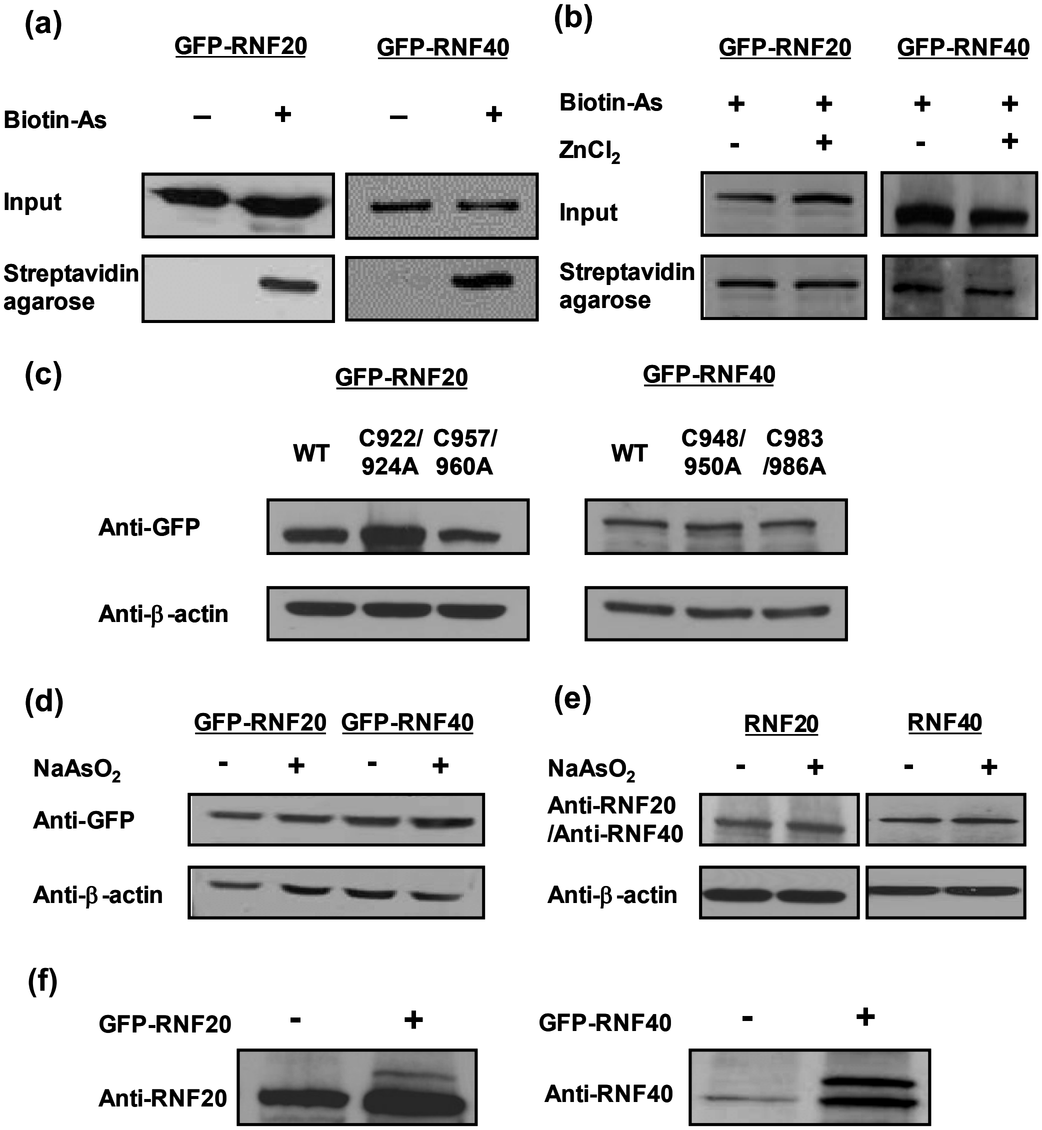


Figure D1. (a) Streptavidin agarose affinity pull-down assay without or with biotin-As probe to examine the binding between As(III) and RNF20 or RNF40 in cells. (b) Pretreatment of cells with 10 μM Zn^{2+} for 1 hr does not affect the binding between As(III) and RNF20 or RNF40 in cells. (c) Western blot results showing the expression levels of wild-type and mutant GFP-RNF20 and GFP-RNF40 in HEK293T cells. (d-e) Western blot results showing the expression levels of ectopically expressed and endogenous RNF20 and RNF40 with or without treatment with 5 μM $NaAsO_2$. (f) Cells without or with ectopically-expressed RNF20 and RNF40 and detected with anti-RNF20 or anti-RNF40. β-actin was used as loading control in (c), (d) and (e).

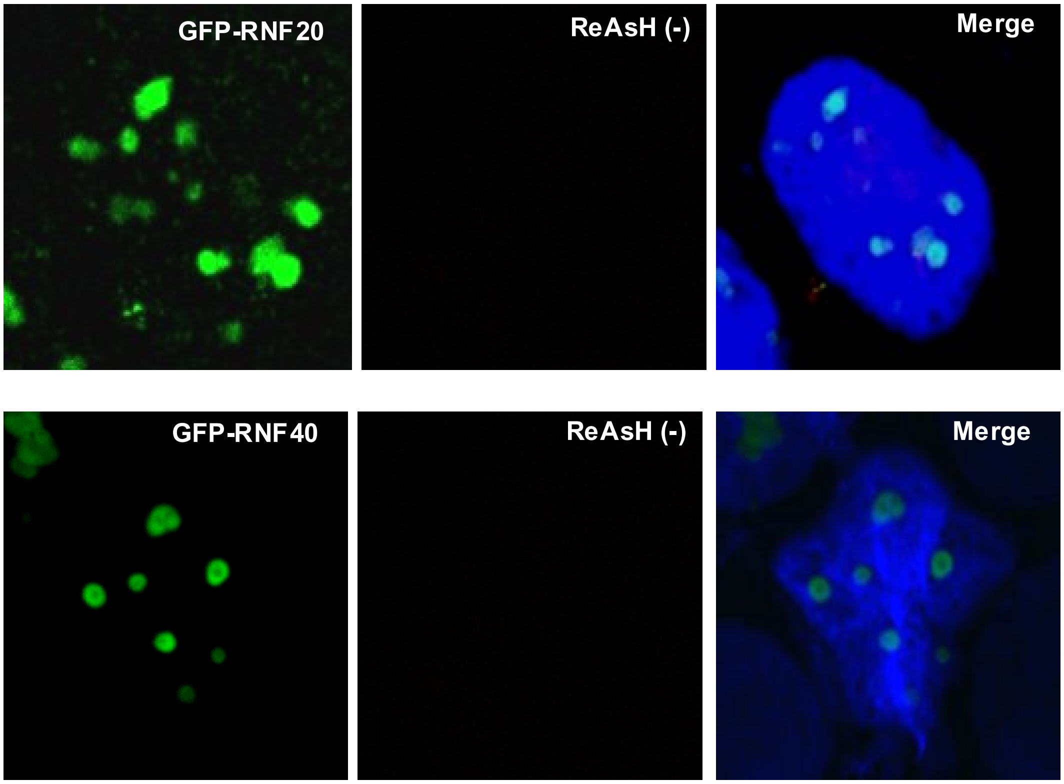


Figure D2. Fluorescence microscopy results for cells transfected with GFP-RNF20 (top) or GFP-RNF40 (bottom), but without ReAsH treatment; no red fluorescence could be detected in these cells.

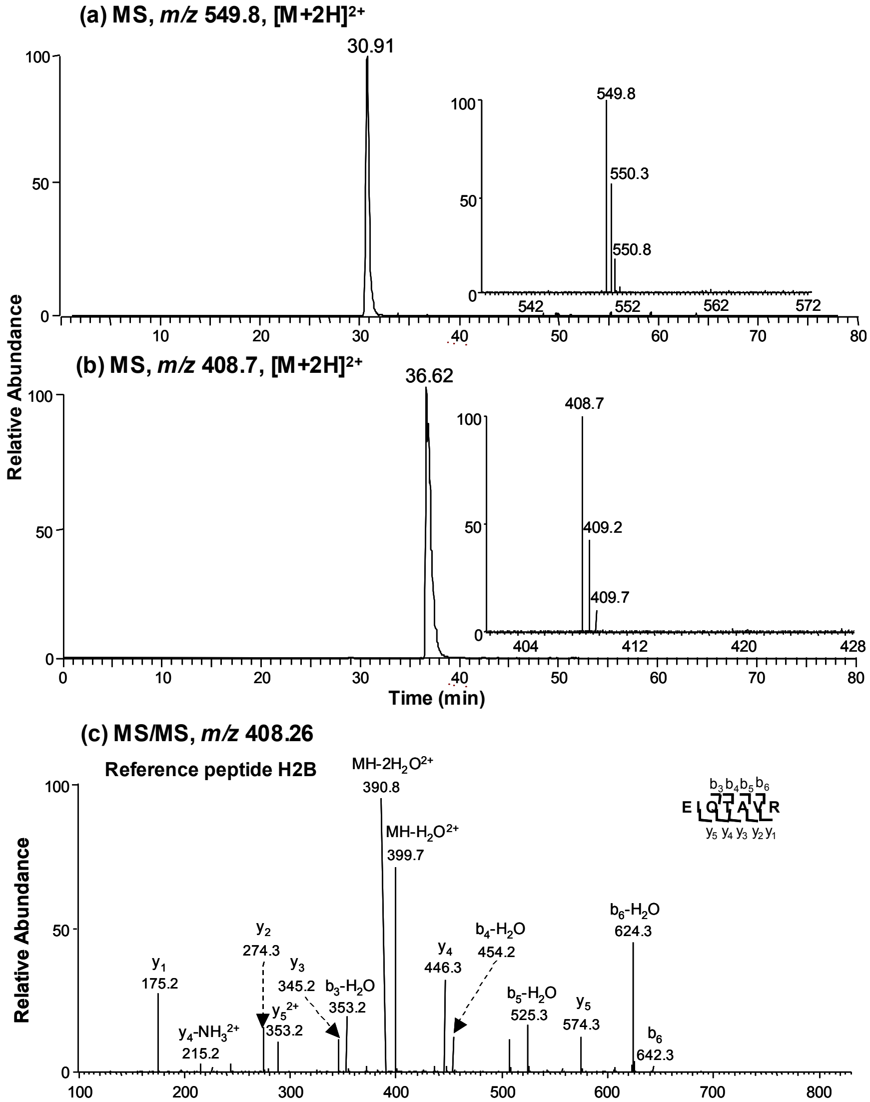


Figure D3. (a) Selected-ion chromatogram for monitoring the $[M+2H]^{2+}$ ion (*m/z* 549.8) of a tryptic peptide derived from histone H2B, AVTK$_{GG}$VTSSK, with K120 being modified with a diglycine remnant. (b) Selected-ion chromatogram for monitoring the $[M+2H]^{2+}$ ion (*m/z* 408.7) of the reference tryptic peptide derived from histone H2B, EIQTAVR. Shown in the insets are the MS displaying the $[M+2H]^{2+}$ ions of the two peptides. (c) Product-ion spectrum of the $[M + 2H]^{2+}$ ion of H2B reference peptide EIQTAVR.

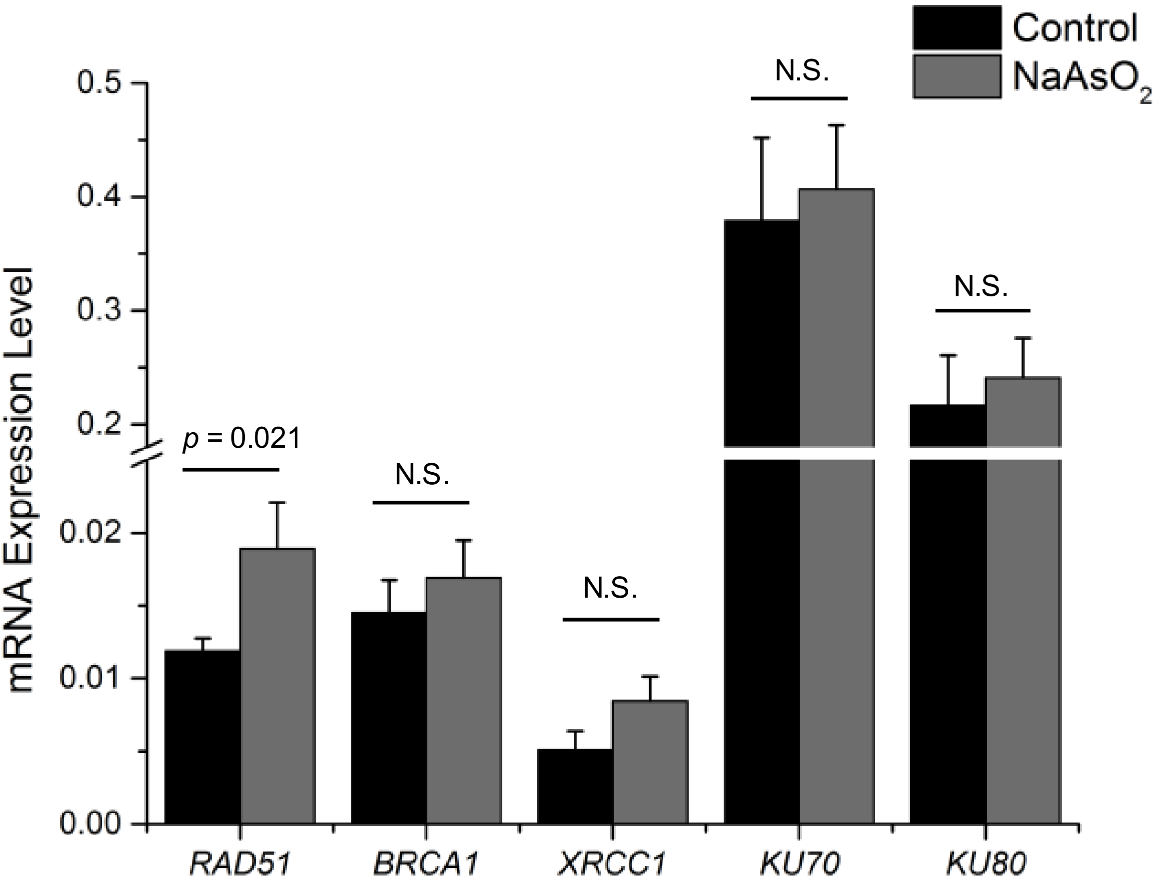


Figure D4. Quantitative real-time PCR results showing that the mRNA expression levels of five genes involved in DNA DSB repair are not reduced in HEK293T cells upon a 24-hr treatment with 5 μM NaAsO$_2$. The mRNA expression levels of these genes were normalized to that of *GAPDH*. The *p*-values were calculated by using unpaired two-tailed *t*-test. “N.S” indicates that the difference was not significant ($p > 0.05$).

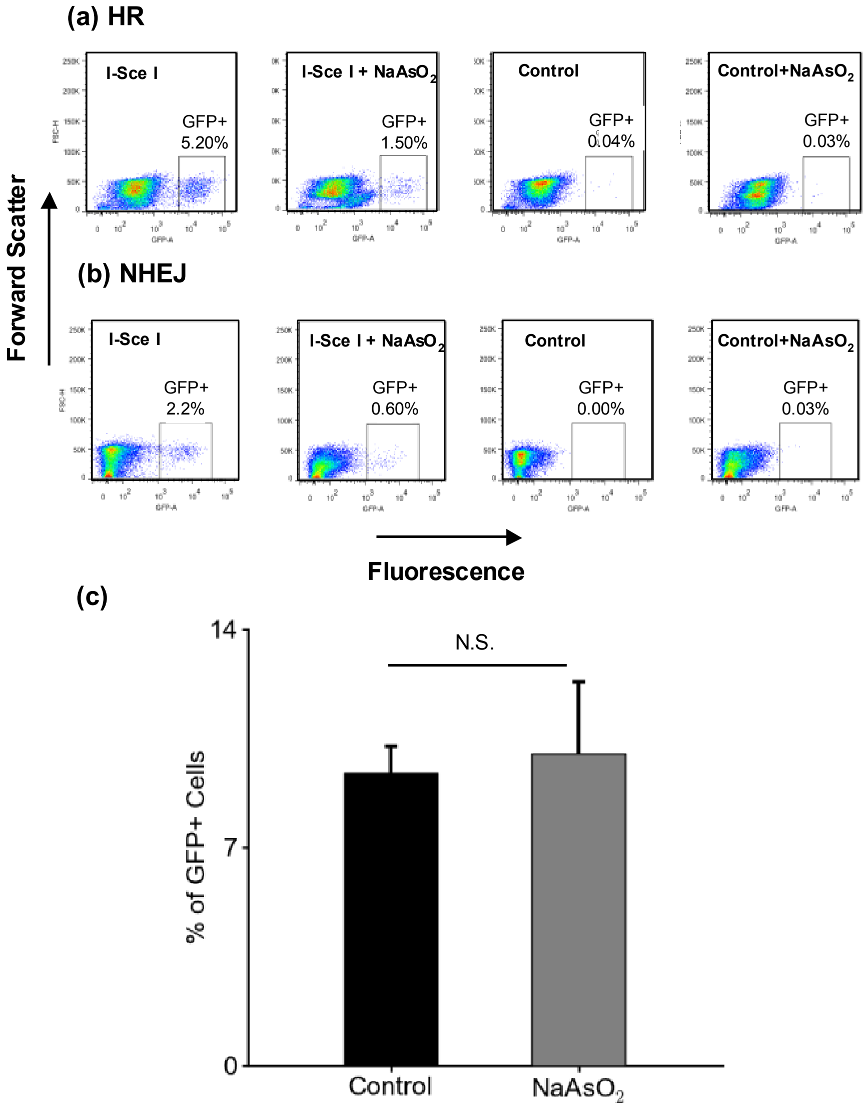


Figure D5. Flow cytometry results showing the percentage of GFP+ cells in (a) U2OS-DRGFP and (b) U2OS-EJ5GFP cells transfected with I-SceI alone (I-Sce I) or together with a 24-hr treatment of 5 µM $NaAsO_2$ (I-Sce I + $NaAsO_2$), as well as those cells that were not transfected with I-Sce I (Control), or treated only with 5 µM $NaAsO_2$ for 24 hr (Control + $NaAsO_2$). (c) Quantitative results showed that the treatment with 5 µM $NaAsO_2$ does not affect the transfection efficiency in U2OS cells. "N.S" indicates that the difference was not significant ($p > 0.05$).