*Electronic Supplementary Information*

# Arsenic trioxide targets Hsp60, triggering degradation of p53 and survivin

Xuqiao Hu,[‡a] Hongyan Li,[‡a] Tiffany Ka-Yan Ip,[a] Yam Fung Cheung,[a] Mohamad Koohi-Moghadam,[a,d] Haibo Wang,[a] Xinming Yang,[a] Daniel N Tritton,[a] Yuchuan Wang,[a] Yi Wang,[a] Runming Wang,[a] Kwan-Ming Ng,[a] Hua Naranmandura,[b] Eric Wai-Choi Tse[c] and Hongzhe Sun[a,*]

[a] *Department of Chemistry, The University of Hong Kong, Hong Kong, P.R. China;*
[b] *Department of Toxicology, School of Medicine and Public Health, Zhejiang University, Hangzhou, P.R. China*
[c] *Li Ka Shing Faculty of Medicine, The University of Hong Kong, Queen Mary Hospital, Hong Kong, P.R. China*
[d] *Division of Applied Oral Sciences and Community Dental Care, Faculty of Dentistry, University of Hong Kong, Hong Kong SAR, China*

‡These authors contributed equally to this work.
*Correspondence and request materials should be addressed to H.S. (e-mail: hsun@hku.hk).

## Supplementary Methods and Figures

**Synthesis of As-AC.** Coumarin azide (Compound **2**) was obtained through diazotization of 7-amino-4-methyl-3-coumarinylacetic acid (Compound **4**), **Supplementary Scheme 1**. 2-*p*-aminophenyl-1,3,2-dithiarsenolane (Compound **3)** was prepared from 4-aminophenylarsenoxide (compound **5**) by dithiol protection, while compound **5** was formed from *p*-arsanilic acid (compound **6**) and further protected by dithiol (Compound **5**), **Supplementary Scheme 2**. Coupling of compound **2** with compound **3** yields As-AC (Compound **1**), **Supplementary Scheme 3**.

Compound **2**: Compound **4** (48.8 mg, 0.21 mmol) was mixed with sodium nitrite (12.75 mg, 0.18 mmol) in water (1 mL) containing concentrated sulfuric acid in an ice-water bath. Subsequently, sodium azide (13.67 mg, 0.21 mmol) was added dropwise to the mixture within 5 min. After 2 hour vigorous stirring, full conversion of the material was achieved as indicated by TLC using 10% MeOH in DCM as an eluent. The precipitate was filtered, washed with ice-cold water and then dried under vacuum in dark to afford compound **2** as a white solid (46.3 mg, 0.179 mmol, 85%).

Compound **5**: *p*-Arsanilic acid (3 g, 13.8 mmol) was added to a solution containing methanol (9 mL), HCl (7.2 mL, 0.20 mmol) and potassium iodide (42.4 mg, 0.26 mmol). Sulfur dioxide was bubbled through the mixture for 30 minutes with the colour changes occurred from orange to pale yellow; 4-aminophenyldichloroarsine was immediately formed as a white precipitate and methanol (5 mL) was further added. The solution was cooled in an ice-water bath and the precipitate was filtered and washed with minimal amounts of diethyl ether. The filtered precipitate was added into 10% ammonium hydroxide (50 mL) for 1 hour, after which compound **5** was precipitated. The mixture was cooled in an ice-water bath and the precipitate was filtered and washed with diethyl ether, giving the residues as a white powder. The residues were transferred into a falcon tube and placed in a centrifuge, further washed with water once to remove ammonia, then washed with tetrahydrofuran (THF) once to remove water. The residue was dried under vacuum overnight to afford the product (1.08 g, 5.90 mmol, 43%). $^{1}$H NMR (300 MHz, DMSO-$d_6$) δ 7.29-7.32 (2H, d, $J$ = 9.0 Hz), 6.63-6.65 (2H, d, $J$ =6.0 Hz) [1].

Compound **3**: 4-Aminophenylarsenoxide (369.1 mg, 2.0 mmol) was dissolved in ethanol (6.6 mL) and 1,2-ethanedithiol (0.168 mL, 2.01 mmol) was added dropwise. After being refluxed for 30 minutes, the mixture was cooled to room temperature and then to 0-5 °C in an ice-water bath. A fluffy white crystal was precipitated. The white crystal was filtered and washed with diethyl ether (10 mL×2), then washed thoroughly with hexane (10 mL×3). The product can be used directly without further purification.

As-AC (Compound **1)**: Coumarin azide (129.2 mg, 0.50 mmol) and EDCI (363.7 mg, 1.90 mmol) were dissolved in dry DMF (5 mL), then 2-*p*-aminophenyl-1,3,2-dithiarsenolane (63.0 mg, 0.24 mmol) and DIPEA (0.3 mL, 1.72 mmol) were added to the above solution. The reaction mixture was stirred overnight under argon atmosphere, diluted with ethyl acetate (50 mL), and then washed with water (10 mL × 3) and

brine (10 mL×1). The organic phase was dried with $MgSO_4$. The organic solvent was evaporated under vacuum, and the residue was purified by silica gel chromatography (*n*-hexane/EtOAc = 2: 1 as an eluent). As-AC was obtained as white powder (19.0 mg, 15.6%). $^1$H NMR (400 MHz, $CDCl_3$) δ 8.46 (1H, s), δ 7.67-7.65 (1H, m), δ 7.58-7.55 (2H, m), 7.52-7.49 (2H, m), δ 7.02-7.00 (1H, m), δ 7.00 (1H, s), δ 3.73 (2H, s), δ 3.34-3.31 (2H, m), 3.14-3.10 (2H, m), δ 2.62(3H, s). $^{13}$C NMR (125 MHz $CDCl_3$) δ 167.60, δ 163.13, δ 153.25, δ 150.23, δ 144.04, δ 138.93, δ 138.81, δ 131.63, δ 126.79, δ 119.54, δ 118.70, δ 117.56, δ 116.08, δ 107.15, δ 41.89, δ 37.63. ESI-MS (m/z): $[M+Na]^+$ obsd. 523.00 (calcd. 522.99), HRMS (ESI) $[M+Na]^+$ found 522.99 (calcd. 522.99),. Melting point: 174-175 ºC.

**Protein expression and purification.** For the expression of BIR3 protein, pGST-his-BIR3 gene was cloned into pET-28a vector (Novogen), and expressed in *Escherichia coli* BL21 (DE 3). After 3 hour induction by 0.1 mM IPTG at 30 ºC, *E. coli* cells were then harvested by centrifugation. Bacterial cell lysate was obtained by ultrasound, and then centrifuged at 10,000 *g* for 30 min at 4 ºC. BIR3 protein was purified from the supernatant using affinity chromatography on Histrap HP column (GE Healthcare) according to the manufacturer's instructions. Proteins were buffer exchanged to PBS using Amicon Ultrafiltration device (Millipore). The purity of protein was verified by 15% SDS-PAGE to reach 95%, and the purified protein was stored at -80 ºC before further usage. Ubiquitin was expressed and purified similarly to BIR3 protein. SlyD, SlyDΔC and SlyD CmutA were overexpressed and purified according to the procedure we developed previously [2].

The GSTP and PPA1 genes were amplified by PCR using cDNA library of NB4 cell line and inserted into pRHS to generate the expression plasmid pRHS-GSTP and pRHS-PPA1. To express GSTP, a single clone of BL21(DE3) *E. coli* harboring the pRHS-GSTP vector was cultured in Luria Broth (LB) medium with 50 μg/ml Ampilicilin at 37℃ for overnight. The bacteria cultured overnight were diluted 1:100 into 1 liter of LB medium containing proper antibiotic for subculture at 37 37℃ for around 2h, until it was grown to an optical density of 0.8 at 600 nm. Expression of His-SUMO-GSTP fusion protein was induced by addition of 0.2 mM IPTG. The bacteria were further cultured at 25℃ for overnight (~16h).

The following steps were carried out at 4 ℃ unless stated otherwise. Bacteria were harvested by centrifugation (4,000 g, 30min) and resuspended in buffer A (20mM Hepes containing 500mM NaCl, pH 7.5) with 1 mM PMSF as an enzyme inhibitor. After cell lysis by sonication, supernatant was separated from pellets by centrifugation (16,000 g, 30 min). The supernatant was further filtered through Millex-HA filter (0.45 μm) and loaded onto a 5-ml HisTrap column (GE Healthcare) preloaded with nickel ion and pre-equilibrated by buffer A containing 50 mM imidazole. Five column volumes of buffer A with 50 mM imidazole were used to wash away the unspecific binding impurities. The His-SUMO-UreG protein was eluted by buffer A supplemented with 300 mM imidazole. The eluted protein fraction was changed buffer to

buffer B (20 mM Hepes, 100 mM NaCl, pH 7.5) to remove excess imidazole for SUMO protease cleavage. Fusion protein was incubated with 50 units of SUMO protease at 25 °C for 3 h with gentle shaking to cleave His-SUMO tag from free GSTP. His-SUMO tag and uncleaved fusion protein were removed by loading the digested protein sample to the 5-ml HisTrap column again. Fraction of flow-through was collected and incubated with 20 mM EDTA and 1 mM DTT at 4 °C overnight to obtain the apo-form GSTP protein, which was further purified by gel filtration using a HiLoad 16/60 Superdex 75 column (GE Healthcare) equilibrated with buffer C (20 mM Hepes, 300 mM NaCl, pH 7.2) containing 500 μM TCEP as a reducing agent. Elution fractions of purified protein were collected and concentrated to 2 ml. PPA1 was purified similarly. The purity of the proteins was confirmed by MALDI-TOF-MS.

For the expression of Hsp60 protein, pGST-his-HSPD1 gene was cloned into pET-28a vector (Novogen), and expressed in *Escherichia coli* BL21 (DE 3). After overnight induction by 0.1 mM IPTG at 25 ºC, *E. coli* cells were then harvested by centrifugation. Bacterial cell lysate was obtained by ultrasound, and then centrifuged at 15,000 *g* for 30 min at 4 ºC. Hsp60 protein was purified from the supernatant using affinity chromatography on Histrap HP column (GE Healthcare) according to the manufacturer's instructions. Proteins were buffer exchanged to PBS using Amicon Ultrafiltration device (Millipore). The purity of protein was verified by 15% SDS-PAGE to reach 95%, and the purified protein was stored at -80 ºC before further usage.

Binding of SlyD and BIR3 proteins to As(III) or As-AC. Either SlyD or BIR3 (30 μM each) was pre-incubated with 0 and 4 molar equivalents of As(III) (as ATO) at 4 ºC for 2 hrs, and then subjected to analysis by ABI4800 MALDI-TOF/TOF Analyzer. Similarly, the proteins (30 μM each) were pre-incubated with 0 and 4 molar equivalents of As-AC at 4 ºC overnight, and exposed to ultraviolet radiation at 365 nm for 20 min at room temperature, enabling the formation of covalent linkage, and then subjected to analysis by ABI4800 MALDI-TOF/TOF Analyzer.

The binding of As to GSTP or PPA1 was examined by MALDI-TOF-MS. Briefly, ATO was diluted to desired concentration (50μM, 100μM, 150μM, 200μM, 250μM) using Hepes buffer containing 5mM GSH, then GSTP or PPA1 protein was added slowly into ATO solution to a final concentration of 50μM. The mixture samples were incubated on ice for 3h or overnight. For MALDI-TOF-MS analysis, protein samples were cocrystallized with sinapinic acid (SA) as the matrix. Drop 1ul crystal on 1 spot of the MS sample plate. Following a short laser pulse, the analytes were protonated and desorbed into the gas phase, and their m/z values were determined in a TOF mass analyzer.

**Quantum yield and coefficient of As-AC.** Fluorescence spectra of As-AC were collected at room temperature (25 ºC) on a Hitachi F-7000 fluorescence spectrophotometer using a 1000 W xenon lamp. The excitation and emission slit width was set at 2.5 nm while the photomultiplier voltage was set at 700 V. A 1 cm×1 mm quartz cuvette with a sample volume of 0.6 mL (20 mM phosphate buffer, pH 7.2, containing 5‰

DMSO) was used for fluorescence measurement. The fluorescent probe As-AC was excited at near-UV wavelength ($\lambda_{ex}$= 355 nm) and emitted in the blue light region (emission maximum at $\lambda_{em}$= 448 nm), similar to Ni-*NTA-AC* [3]. The quantum yield of As-AC was determined to be 0.050 using quinine sulfate (Sigma) as the standard according to Williams' method[4]. The molar extinction coefficient (ε) at 355 nm of As-AC was calculated to be 1,841 $M^{-1} \cdot cm^{-1}$.

**Fluorescence turn-on response of As-AC.** The fluorescence response of As-AC towards a series of proteins or small ligands was monitored by fluorescence spectroscopy on a Hitachi F-7000 fluorescence spectrophotometer with 1000 W xenon lamp in a 1 cm × 1 mm quartz cuvette containing 600 μL samples. As-AC probe (1 μM) in 20 mM phosphate buffer, pH 7.2 was mixed with 10 molar equivalents of proteins including BIR3, Ubiquitin, SlyD, SlyDΔC or small ligands i.e., GSH, EDT and cysteine. The mixture was then exposed to long wavelength UV (365 nm) for different times (0, 1, 2, 3, 5, 10, 15, 30 min) prior to fluorescence recording.

**Examination of As-AC-protein binding *in vitro*.** Two groups of BIR3, ubiquitin, SlyD and SlyDΔC proteins were pre-incubated with equal molar equivalents of As-AC for 30 min at 37 ºC in PBS (pH 7.4). One group was then irradiated by UV light at 365 nm to allow the covalent linkage between the probe and the labeled proteins to be formed, while the other group was incubated under darkness for 20 min at room temperature for comparison. The samples were then separated by SDS-PAGE. The gel was imaged by $_{MY}$ECL Imager (Thermo Scientific) and further stained by Coomassie brilliant blue for protein visualization. The fluorescence intensity on the gel was quantified using ImageJ software.

**Fluorescent labeling of proteins in *E. coli* by As-AC.** *E. coli* KMI603 cells transformed with SlyD, SlyDΔC and SlyD CmutA were cultured separately in LB medium overnight at 37 ºC, with the supplementation of kanamycin (100 μg/mL) and ampicillin (100 μg/mL), and then sub-cultured by 1:100 dilutions. When $OD_{600}$ reached around 0.6, isopropyl β-D-thiogalactopyranoside (IPTG) (0.2 mM) was added to each culture to induce protein overexpression with further incubation at 30 ºC for 3 hrs. Bacterial cells were washed afterwards by PBS (pH 7.4) at 4 ºC for three times. Sterilized As-AC probe (10 μM) was then added into each group and further incubation for 30 min under darkness. The cells were then collected through centrifugation (3,000 rpm, 15 min) and washed by 20 mM phosphate, 100 mM NaCl, pH 7.2 at 4 ºC for three times, followed by 365 nm UV-light exposure for 20 min to allow photo-activation of the arylazide. Bacterial cell lysates were then separated by SDS-PAGE. Fluorescence gel images were then captured and the gels were stained by Coomassie brilliant blue for protein visualization.

**Identification of As-binding proteins in NB4 cells by As-AC.** NB4 cells were cultured in RPMI1640 medium containing penicillin (100 U/mL), streptomycin (100 μg/mL) and 10% FBS. Cells were incubated in 5% $CO_2$ humidified incubator at 37 ºC and sub-cultured in a ratio of 1:4 every two days. 10 mL of NB4 cells were then incubated with 10 μM As-AC under darkness for 30 min. Cells were collected and washed with PBS for three times. Cell pellets were re-suspended in 1 mL PBS buffer, and then exposed to UV (365 nm) for 20 min on ice. Cells were lysed through freeze-thaw cycles for three times (3 min × 3), and protein concentration was determined by BCA protein assay kit. About 120 μg proteins were applied to 2-DE separation. The samples were desalted using a 2D clean-up kit (GE Healthcare) and rehydrated. Isoelectric focusing (IEF) was performed using a precast Immobiline DryStrip (13 cm, pH 4-7 NL), followed by the second dimensional SDS-PAGE (13.5% arylamide gel) separation. Fluorescent spots on gels were imaged, and then subjected to silver staining for protein visualization. The fluorescence-labelled protein spots were then identified by peptide mass fingerprinting as described previously [5].

**Cytotoxicity of AS-AC.** HeLa and NB4 cells were seeded into 96-well plates (10,000 cells/well) and incubated in 100 μL of respective culture medium at 37 ºC with 5% $CO_2$ for 24 hrs. The cells were incubated with gradient amounts of As-AC probe (0-10 μM) in respective culture medium at 37 ºC incubator under darkness for 24 hrs. Cell viability was measured by XTT cell proliferation assay (Roche). Briefly, 50 μL XTT (sodium 3'-[1-(phenylaminocarbonyl)-3,4-tertrazolium]-bis(4-methoxy-6-nitro) benzene sulfonic acid hydrate) labeling mixture, which was prepared by mixing XTT labeling reagent (1 mg/mL in sterile phenol free RPMI 1640 medium) and electron coupling reagent (0.383 mg/mL in sterile phosphate buffered saline) at the ratio of 50: 1, was added to each well and incubated at 37 ºC with 5% $CO_2$ for 6 hrs. The absorbance at 450 nm was recorded using a microplate reader (Bio-rad, iMark™), and cell viability was reported relative to those of untreated cells.

**Confocal imaging of HeLa cells.** HeLa cells were transferred onto a 15 mm confocal dish for confocal imaging and allowed to stably adhere to the glass bottom for one day. Afterwards, the medium was removed from the culture plate and cells were washed with pre-warmed HBSS buffer. Cells were then incubated with 1 μM As-AC in HBSS buffer for 15 min, and washed by HBSS buffer for five times to remove unbound As-AC before confocal imaging. The fluorescent and phase contrast images were captured by Confocal Microscope-Zeiss LSM 710 using 40 × oil lens with excitation by 405 nm laser and the emission of 447-500 nm.

Competitive experiment was performed by pretreating HeLa cells with gradient concentrations (0, 2, 5, 10 and 20 μM) of As(III) for 30 min at 37 ºC, and then washed with HBSS buffer for three times to remove unbound As(III). Afterwards, the cells were treated with 10 μM As-AC in HBSS buffer for 20 min at 37 ºC.

Cells were washed with HBSS buffer again to remove unbound As-AC, and then subjected to confocal imaging. Fluorescent and phase contrast images were captured by a Carl Zeiss LSM780 Confocal Microscope (n= 5) using 405 nm laser under 40 × oil lens, while the measuring range of emission was fixed between 447-500 nm.

**Cellular thermal shift assay.** NB4 cells (50 mL) were treated with 1 μM ATO for 24 hrs, and then harvested (1.5 to 2× $10^6$ cells/mL) by centrifugation at 800 *g* for 4 min at 4 ºC. Cell pellets were resuspended in 1.5 mL ice-cold PBS, and 100 μL of cell suspension were transferred into 0.2 ml PCR tubes, followed by heating each tube in parallel in a PCR machine from 40 to 66 ºC for 6 min. The tubes were then incubated for 6 min at room temperature. Afterwards, the cells were snap-frozen in liquid nitrogen for 3 min and thawed shortly in a water bath at 25 ºC for 1 min. This freeze-thaw cycle was repeated four times. The cell lysate was then centrifuged at 14,000 *g* for 20 min at 4 ºC. The supernatant (30 μL) was transferred into a new tube, separated by SDS-PAGE, and subjected to western blot analysis.

**Western blot analysis.** After SDS-PAGE separation, proteins of interest were transferred to PVDF membranes using 110 V constant voltage for 2 hrs. Membranes were blocked with 5% (w/v) BSA in TBST buffer, and subsequently incubated with primary and secondary antibodies with optimized dilution ratios and incubation time. Western blot results were imaged and protein band densities were analyzed using ImageJ software. The total protein densities in each lane of the SDS-PAGE gel were quantified using loading control for protein normalization.

**Cell lysate protein preparation, digestion and TMT labeling.** Six plates of NB4 cells were seeded with a density of 5 × $10^5$ cells/well. Three plates were treated with 2 μM ATO and the other three were set as control. After ATO treatment for 24 hrs, the control and treated cells were collected, washed with ice-cold PBS for three times, and then resuspended in cell lysis buffer (8 M urea, 20 mM Tris·HCl, pH 8.0) with the supplementation of protease inhibitor cocktail. The cells were lysed through freeze-thaw cycles and protein concentration was determined by BCA assay. Cell lysate proteins (100 μg) were desalted by acetone precipitation. The protein pellet was resuspended in protein digestion buffer (8 M urea, 100 mM Hepes buffer, pH 8.5), and heated at 60 °C for 10 min for protein denaturation. Protein samples were then subjected to dithiothreitol (5 mM) and iodoacetamide (25 mM) treatment, and diluted with 100 mM Hepes buffer (pH 8.5) to lower urea concentration. The lysates were digested with 1 μg trypsin for 16 hrs at 37 °C, with 1: 100 mass ratio of trypsin to total protein amount, and a final concentration of 5% formic acid was used to quench the digestion. The digested samples were then desalted using pipet tip-based C18 StageTips according to published protocol [6].

The TMT labelling reaction was performed according to the instructions of the manufacturer. Briefly, the desalted and dried peptides were first reconstituted in 100 mM TEAB buffer at a concentration of 1 mg/mL. TMT-126 and TMT-127 reagents were equilibrated to room temperature. A volume of 41 μL acetonitrile was then added to each TMT vial to completely dissolve the reagent. The solutions were added to the peptide sample, and incubated for 1 hr at room temperature. The samples derived from ATO treated NB4 cells were labeled with TMT-126 agent whereas the control NB4 samples were labelled with TMT-127 agent. After that, the samples were quenched by 5% hydroxylamine solution for 15 min at room temperature. Each pair of TMT-126 labeled peptides and TMT-127 labeled peptides were combined in the same tube in 1: 1 ratio so that each tube contains 200 μg peptides, with 100 μg TMT-126 labelled and 100 μg TMT-127 labelled peptides.

**TMT-labeled peptide analysis and data acquisition.** A Surveyor MS Pump Plus HPLC system (Thermo Scientific) connected online with LTQ-Orbitrap Velos Mass Spectrometer was used for LC-MS/MS analysis of the TMT-labeled peptide samples. A reversed phase Glass PicoTip column was connected to a PicoTip ESI emitter (15 μm I.D., New Objective) that was packed in-house with YMC*GEL ODS-A HG 5-μm C18 resin (YMC Co. Ltd.). The peptide samples were dissolved in 0.1% formic acid at a concentration of 1 mg/mL and 10 μL samples containing 10 μg labeled peptides were injected into the HPLC system in each analysis. The peptides were eluted under a 135 min gradient, where the samples were eluted with different compositions of buffer A (0.1% formic acid in water) and buffer B (99.9% acetonitrile with 0.1% formic acid) at a flow rate of 250 nL/min to 570 nL/min.

MS data was acquired using data dependent acquisition (DDA) mode. The MS1 spectra were first acquired by cycling through a full scan (m/z 300–2,000) in the linear ion trap. MS2 mass spectra were acquired using the Orbitrap mass analyzer (resolution = 60,000, m/z 300–1,800) by selecting the 10 most abundant ions from the full scan and 10 CID MS/MS scans were conducted on the samples immediately after MS1 full scan. Precursor ions were isolated with a m/z window of 2, and were listed in a dynamic exclusion list for 60 seconds after they were selected for MS/MS. Ions that were singly-charged, or with unassigned charge were excluded from MS/MS.

**MS data processing and analysis.** The MS raw data were processed and analyzed using the quantitative proteomics software package MaxQuant [7] (version 1.5.5.1) built in Andromeda search engine with default parameters. The UniProt human proteome was used as the protein database for protein search and identification [8]. The protein and peptide identification false discovery rate (FDR) determined by target-decoy approach was set to < 1%. Quantification type was set as reporter ion MS2 with TMT-126 and TMT-127 set as isobaric label. Cysteine carbamido-methylation was set as a fixed modification; methionine

oxidation was set as a variable modification. Maximum number of missed cleavages was set as 2, minimum peptide length was set as 7 amino acids, maximum peptide mass was set as 4,600 Da. The first and main search peptide tolerance was set as 20 ppm and 4.5 ppm respectively.

After MaxQuant analysis, the MS/MS table that contains all the peptides identified through MS/MS and matches between runs was used for protein quantification. The leading razor protein entry of each peptide, representing the protein with the most identified peptides, was considered as the origin protein of the peptide. The peptides that were labelled as reverse decoy or potential contaminants were removed from analysis. The fold change of each peptide was then calculated, using the ratio of reporter ion intensity of each TMT tag. Reporter intensity 0 corresponds to TMT-126 ions originated from ATO-treated cells and reporter intensity 1 corresponds to TMT-127 ions originated from control cells, and the fold change of the peptide was calculated by the ratio of (Reporter intensity 0)/(Reporter intensity 1).

**Bioinformatics analysis.** The STRING webserver [9] was used to build the protein-protein interactions of As-associated proteins with medium confidence settings and "Homo Sapiens" selected as the target organism, which resulted in the construction of a PPI network with 260 nodes and 2488 edges. CytoScape v3.7.1[10] was used to visualize and process the network. CentiScape plugin was used to calculate the betweenness centrality (BC) score of each node and the top ten nodes with the highest BC scores were selected as the hub proteins for further analysis.

The Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.8[11] was used to identify the biological processes and molecular functions enriched by the As-associated proteins. The official gene names of each protein from Unitprot database were used as input of DAVID. The GO terms with *p*-value < 0.05 were extracted as significantly enriched categories. GOPlot package in R[12] was used to calculate a z-score for each GO term using the following equation:

$$zscore = \frac{up - down}{\sqrt{count}}$$

Here, up refers to the number of up-regulated genes with logFC > 0 and down is the number of down-regulated genes with logFC < 0. These z-score values were used to plot the GO terms. To build GO enrichment network, we used ClueGO plugin in CYtoScape.

**ATPase activity assay.** The human recombinant Hsp60 (2 μM) (BostonBiochem) was incubated for 30 min at 25 °C with or without various concentration of ATO in assay buffer (50 mM Tris, pH 7.6, 7 mM KCl and 7 mM $MgCl_2$), and then treated with ATP (0.1 mM). After incubation for 30 min at 37 °C, ATP content was determined by Malachite Green Phosphate Assay Kit (POMG-25H).

**Protein refolding assay.** Refolding of glow-fold protein was performed by Hsp60 and Hsp10 proteins using Hsp60/Hsp10 Protein Refolding Kit (Boston Biochem) according to the standard protocol. Briefly, 0.25 μM Glow-fold protein was unfolded in 25 mM Hepes buffer (pH 7.4) at 45 ºC for 7 min. The refolding was performed in 25 mM Hepes (pH 7.4) containing 20 mM KCl, 10 mM $MgCl_2$. About 0.36 μM of Hsp60 and Hsp10 were added into the refolding buffer in the presence or absence of graduate concentrations of As(III) (0.8, 1.6, 4.0, 8.0 μM). After 15 min incubation at room temperature, ATP was added to a final concentration of 1 mM. After 60 min refolding at 30 ºC, 4 μL of each sample mixture were mixed with 50 μL luciferin reagent in 96-well half area opaque white plate and luminescence was recorded immediately on a plate reader.

**Time-dependent expression levels of survivin and p53 upon ATO treatment.** NB4 cells were cultured in fresh RPMI 1640 medium supplemented with 10% FBS serum (Invitrogen) in a humidified atmosphere of 5% $CO_2$ and 95% air at 37°C. The cell density was adjusted to $1\times 10^5$ cells/mL, and 10 mL of the cells were seeded into 10 cm cell culture plates in triplicate. The plate was incubated for 24 hrs to obtain best-suited culture conditions for the cells. Cells were pretreated without or with 25 μM MG132 for 4 hrs, and then treated without or with ATO (0.8 μM) for 0, 6, 12 and 24 hrs. The cells were collected and washed by ice cooled PBS, and then lysed by lysis buffer (1 M Tris, pH 6.8, 2 M NaCl, 0.1 M EDTA, 0.05 M EGTA, 0.5% NP-40 and 1×EDTA-free protease inhibitor cocktail). Then, samples were vortexed briefly and incubated in lysis buffer for 1 h on ice and centrifuged at 15,000 *g* for 15 min. Samples of 20 μg of total protein were subjected to 15% SDS-PAGE separation, and transferred onto a nitrocellulose membrane. The expression levels of survivin and p53 were analyzed by western blotting, β-actin was used as the protein loading control.

**Immunoprecipitation assay.** For immunoprecipitation (Co-IP) studies, NB4 cells were lysed with immunoprecipitation buffer (50 mmol/L Tris buffer, pH 7.4, 150 mmol/L NaCl, 1 mmol/L EDTA, 1% NP-40, and 0.25% sodium deoxycholate) with the addition of 1× EDTA-free protease inhibitor cocktail. 500 μg of cell lysate were then immunopreciptated with 2 μg antibodies or nonimmune serum and 50 μL of Dynabeads™ protein G (Thermo Scientific) at 4 ºC overnight with constant rotation. After an immunoprecipitation assay buffer wash, the immunoprecipitated proteins were eluted by boiling with SDS sample buffer and analyzed by western blotting.

**Indirect immunofluorescence staining and imaging.** NB4 cells ($5\times 10^5$ cells/well) were seeded in 6-well plates. After overnight incubation, cells were treated with or without 1.2 μM ATO for 6 hrs, then collected and washed by ice-cold PBS three times, and fixed in -20 ºC pre-cooled methanol for 5 min. After the fixation step, cells were washed twice with PBS and softly suspended in fresh PBS. The cell suspension was placed on poly-L-lysine coated 8-well chamber plates (Thermo Scientific). Cells were dried on glass slides, blocked

for 2 hrs at room temperature in 2% BSA dissolved in PBS (0.1% Tween 20).

The blocked cells were incubated with primary antibodies (5 μg/mL) including rabbit anti-p53 monoclonal antibody (SP5) (Thermo Scientific), rabbit anti-survivin monoclonal antibody (J.33.5) (Thermo Scientific), mouse anti-HSP60 antibody (LK1) (Santa Cruz Biotechnology). The antibodies were paired by the following combination groups and incubated together: hsp60-p53 and hsp60-survivin. All cells were incubated with primary antibodies for 1 hr at room temperature. Afterwards, the cells were washed three times using PBS buffer with gentle shaking for 5 min, and incubated again with Alexa Fluor 488 dye goat anti-mouse antibody and Alexa Fluor 594 dye goat anti-rabbit antibody (Thermo Scientific) for 45 min at room temperature under darkness. The cells were then washed three times with PBS buffer with gentle shaking for 5 min. Negative control groups were set by adding just secondary antibodies or no antibodies. The excess liquid from the slides was removed and 10 μL of antifade mountant with DAPI medium were then added to each well. All images were captured using a Zeiss LSM 710 confocal laser scanning microscope (Carl Zeiss, Oberkochen, Germany) with a 63× oil lens objective. Images for AF488 (in green) and AF594 (in red) were obtained using multitrack channel mode settings for FITC channel at 488 nm and detected emission range of 493-580 nm, excitation for PI channel at 543 nm, and detected emission range of 580-679 nm. DAPI signals were additionally detected in a single track with an excitation at 405 nm and detected emission range of 408−478 nm.

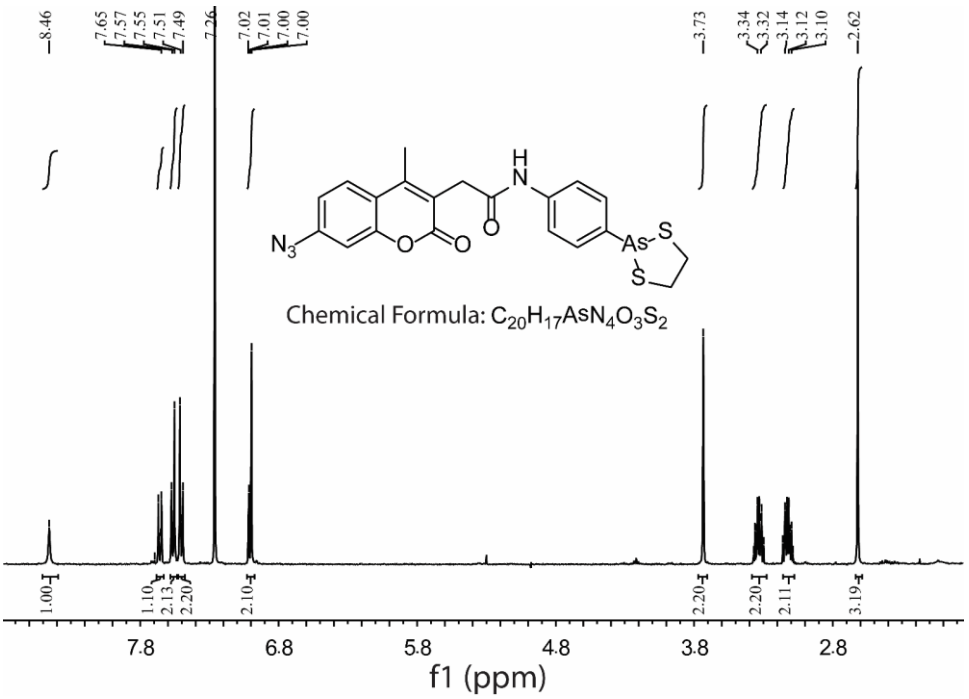



**Fig. S1** $^{1}$H NMR spectrum of As-AC.

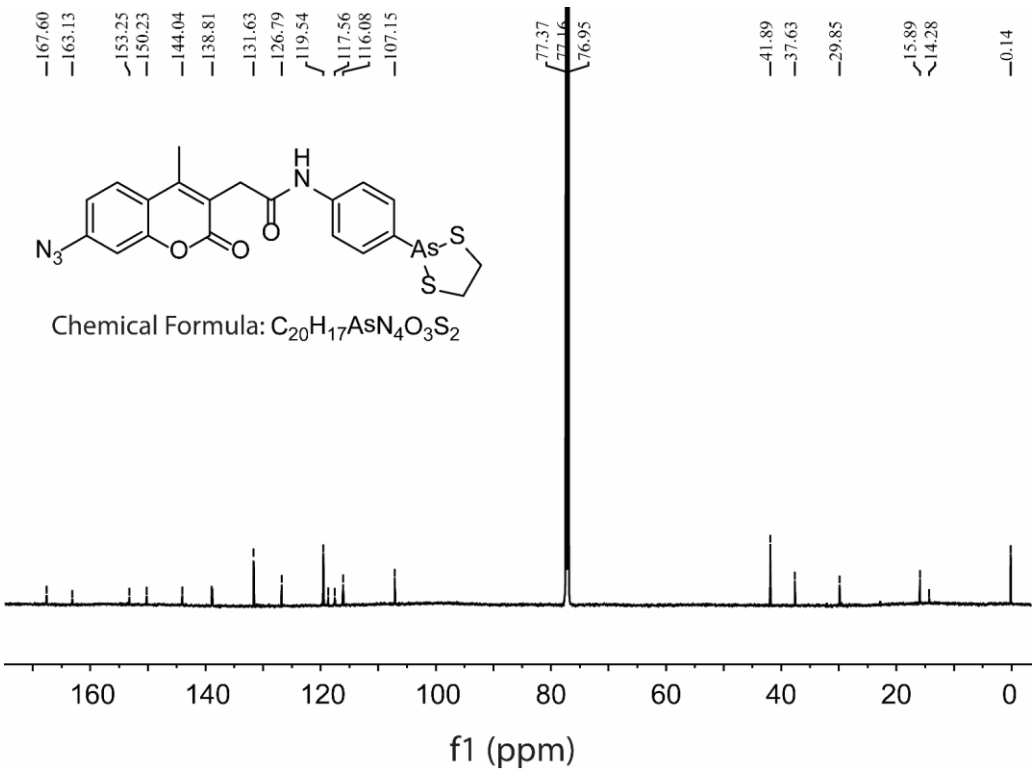



**Fig. S2** $^{13}$C NMR spectrum of As-AC.

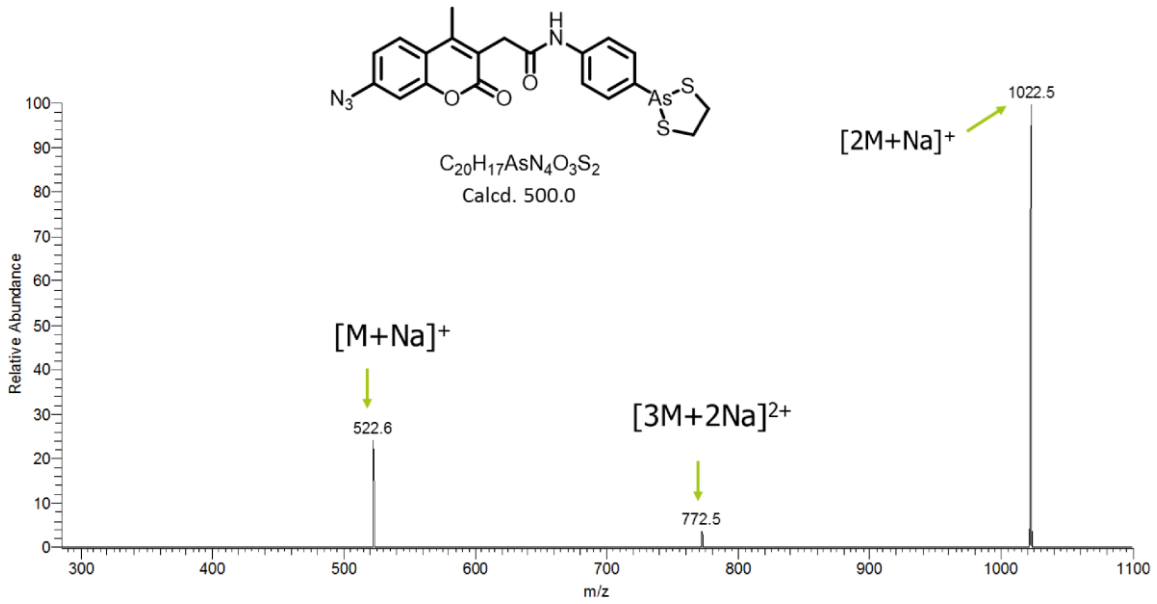



**Fig. S3** ESI-MS spectrum of As-AC.

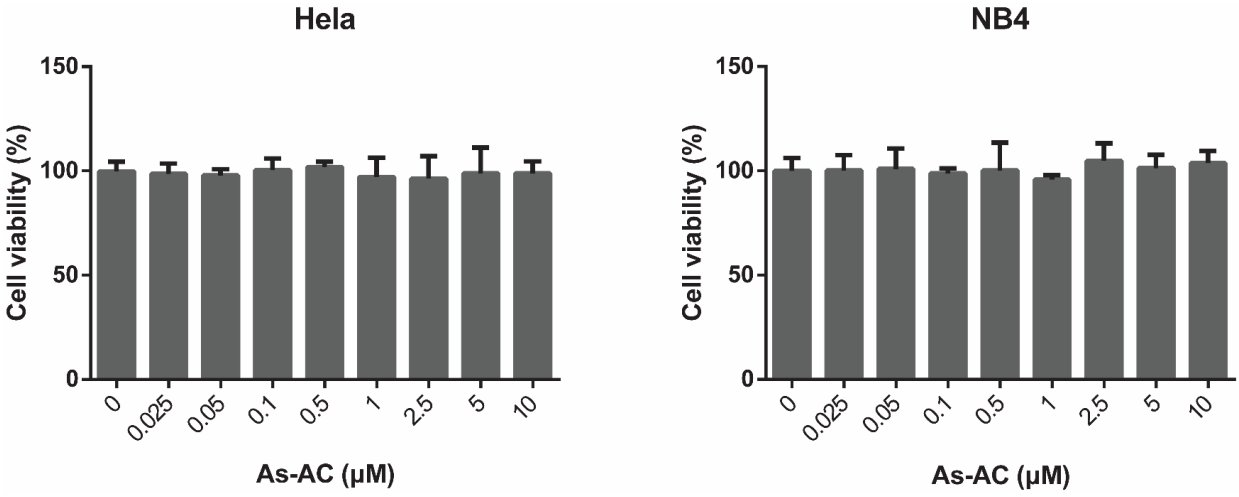



**Fig. S4** Cytotoxicity of As-AC at different concentration in HeLa and NB4 cells by XTT assay. The data are the average of quadruplicate measurements. A gradient amount of As-AC probe (0-10 μM) in respective culture medium at 37 ºC was incubated under darkness for 24 hrs. Cell viability was measured by XTT cell proliferation assay (Roche).

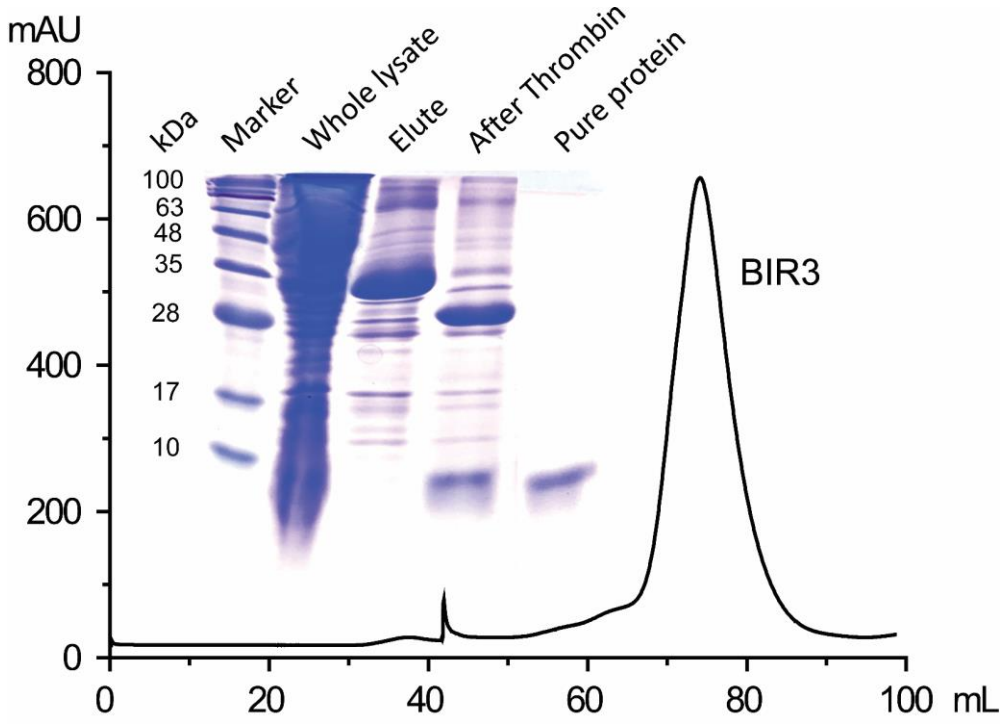



**Fig. S5** Gel filtration profile of BIR3 protein on a Hiload 16/60 Superdex200 column eluted by 20 mM HEPES containing 300 mM NaCl, pH 7.4. Insert: SDS-PAGE analysis of BIR3 protein.

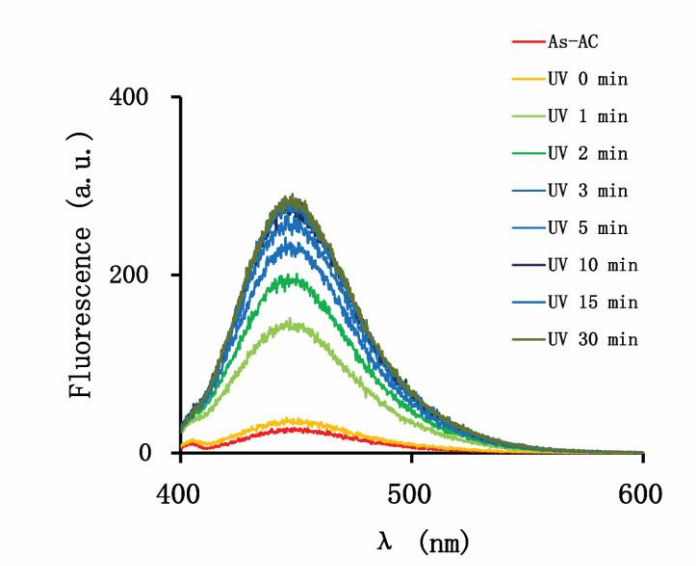



**Fig. S6** Fluorescence spectra ($\lambda_{em}$ = 448 nm) of As-AC (1 μM) in 20 mM phosphate buffer, pH 7.2 upon mixing with SlyD protein (10 molar equivalents) irradiated by UV light (365 nm) at different times (0, 1, 2, 3, 5, 10, 15, 30 min) prior to fluorescence recording.

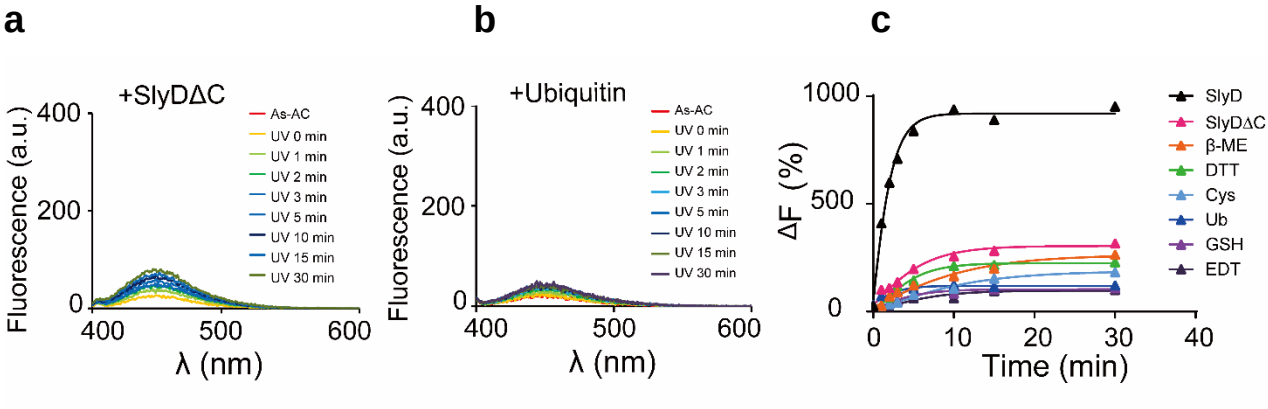



**Fig. S7** Fluorescence spectra ($\lambda_{em}$ = 448 nm) of As-AC (1 μM) in 20 mM phosphate buffer, pH 7.2 upon mixing with 10 molar equivalents of (a) SlyDΔC, (b) Ubiquitin protein for different times (0, 1, 2, 3, 5, 10, 15, 30 min). (c) Time-dependent fluorescence response of As-AC (1 μM) to 10 molar equivalents of different proteins or small ligands. Ub: Ubiquitin, DTT: 1,4-Dithiothreitol, Cys: Cysteine, EDT: 1,2-Ethanedithiol, GSH: glutathione, β-ME: 2-mercaptoethanol.

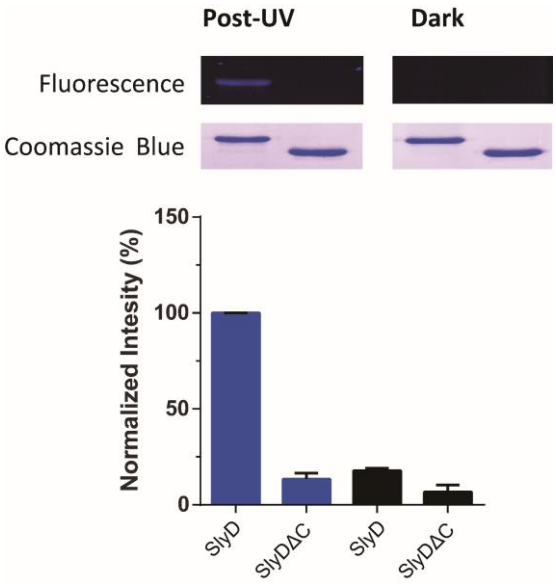



**Fig. S8** *In vitro* selectivity of As-AC towards SlyD and SlyDΔC post-UV activation (blue) and in darkness (black). Two groups of SlyD and SlyDΔC proteins were pre-incubated with equal molar equivalents of As-AC for 30 min at 37 °C in PBS (pH 7.4). One group was then irradiated by UV light at 365 nm to allow the covalent linkage between the probe and the labelled proteins to be formed; while the other group was incubated under darkness for 20 min at room temperature for comparison. The samples were then separated by SDS-PAGE. The gel was imaged by MYECL Imager (Thermo Scientific) and further stained by Coomassie brilliant blue for protein visualization.

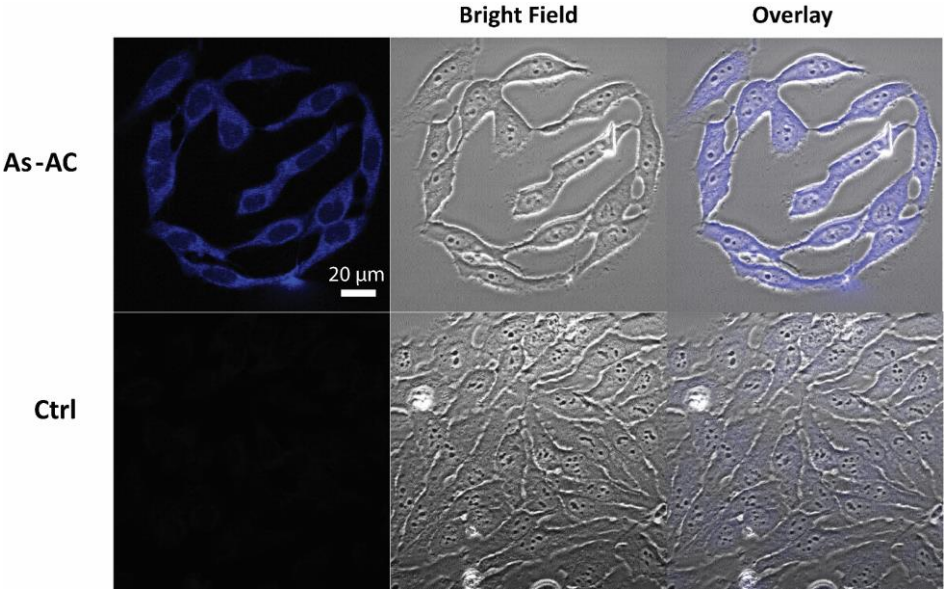



**Fig. S9** The confocal images of HeLa cells with As-AC (1 μM) treatment at 37 $^{o}$C for 15 min (n= 3). The images were collected at Confocal Microscope-Zeiss LSM 710. Scale bar: 20 μm.

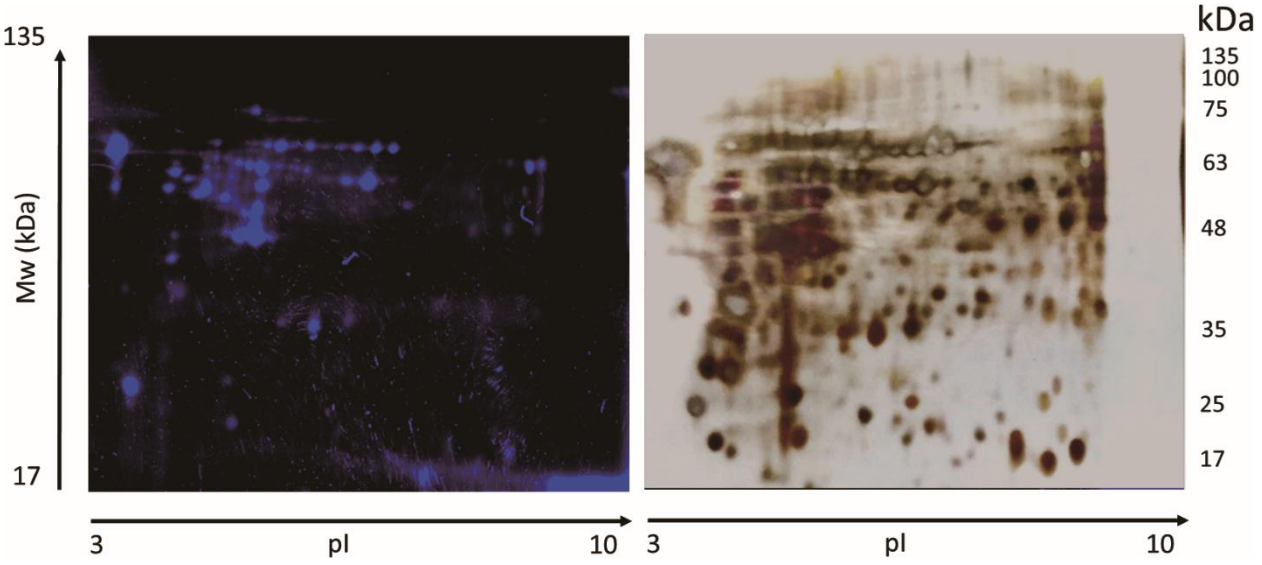



**Fig. S10** Two-dimensional electrophoresis (2-DE) gel of NB4 cells pretreated with As-AC (10 μM) for 30 min and exposed to UV (365 nm) for 20 min. The probe labelled proteins in NB4 cells were identified in 2-DE by fluorescence (left) and silver staining (right). The protein spots on the gels selected corresponding to the blue spots were exercised and subjected to peptide mass fingerprinting analysis for comparison.

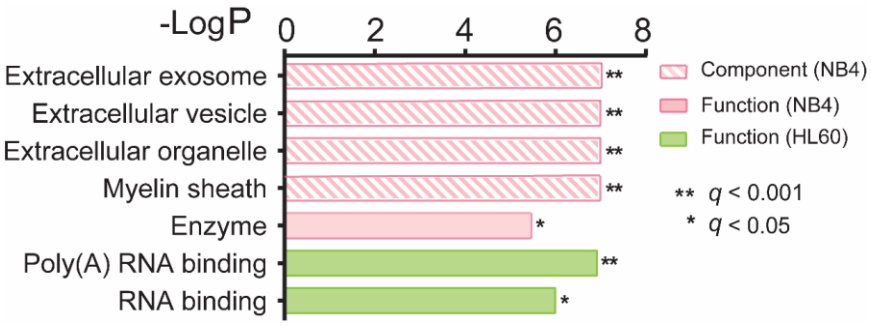



**Fig. S11** GO enrichment analysis illustrating the functional and component categories of non-mutual proteins from NB4 and HL60 cells by GOrilla web server. The p-value was set as '< 0.001', q-value '< 0.05' in functional categories and '< 0.001' in component categories are listed.

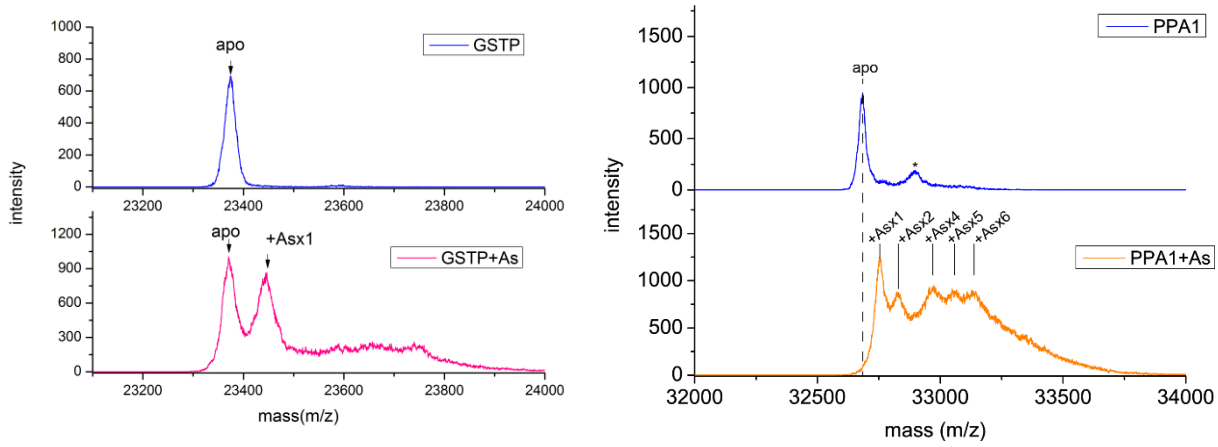



**Fig. S12** The binding of GSTP1 and PPA to $As^{III}$ (as ATO) by MALDI-TOF mass spectrometry. For GSTP, the peaks at m/z of 23370.534 and 23446.643 are assignable to apo- and As-GSTP. For PPA1, the peaks at m/z of 32680.489 is assigned as apo-PPA1, while the peaks at 32752.116, 32824.414, 32971.627, 33048.331 and 33128.103 are assigned as 1*As-, 2*As-, 3*As-, 5*As- and 6*As-bound PPA1.

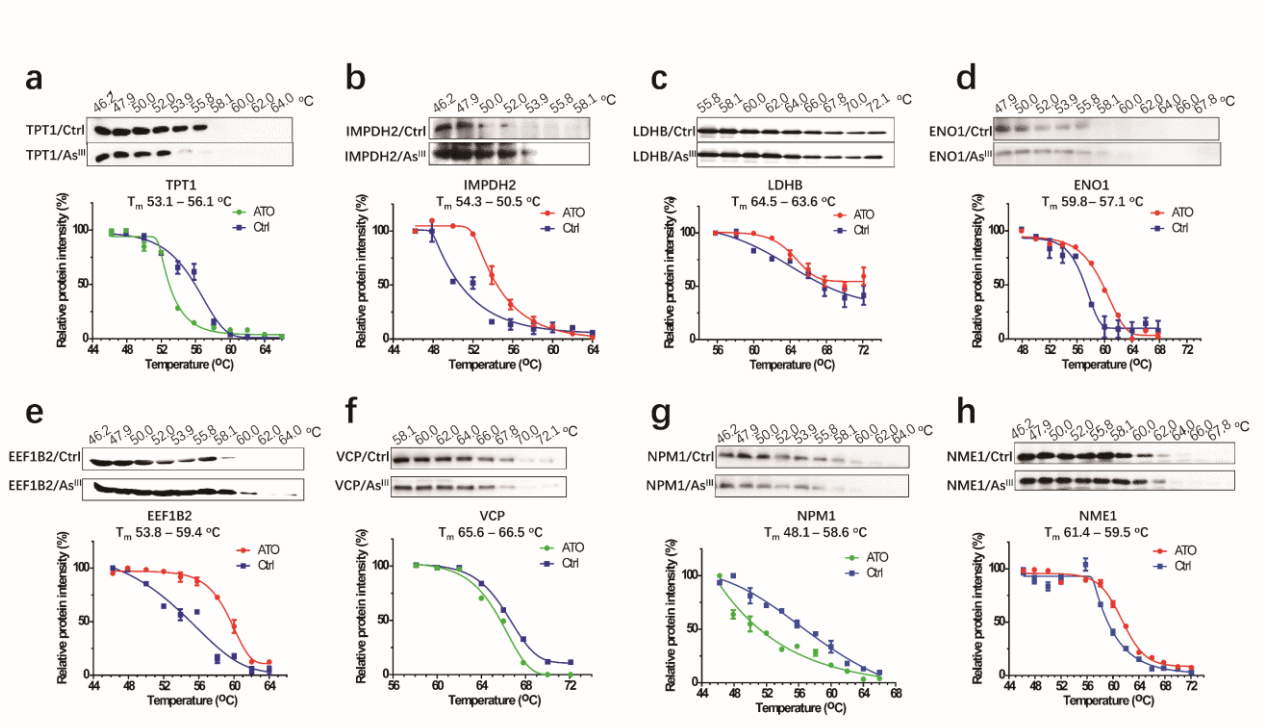



**Fig. S13** Protein thermal melting curves of different As-binding proteins in leukemia intact cells, treated with or without arsenic trioxide. Data are presented as the mean SEM from at least three independent experiments.

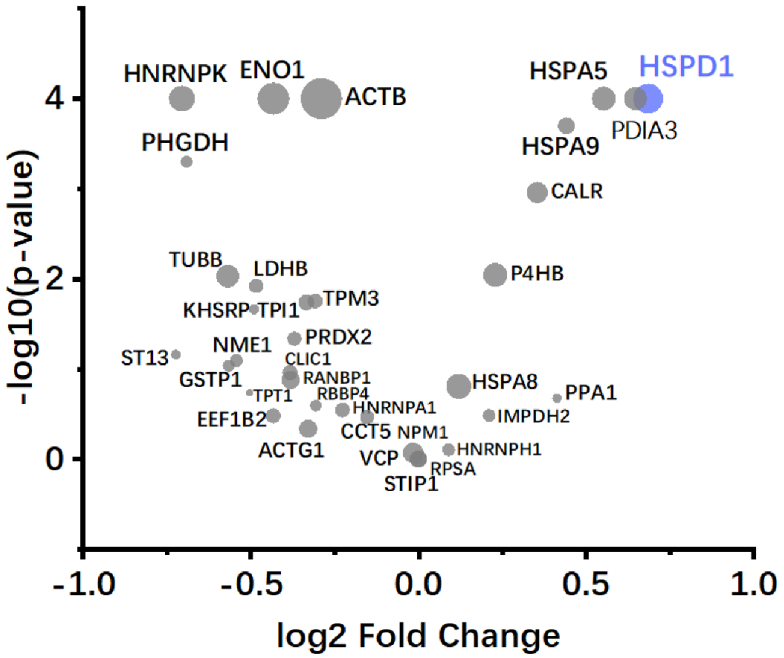



**Fig. S14** Volcano plot of log2 fold-change (x-axis) versus -log10 p-value (y-axis, representing the probability that the protein is differentially expressed). Proteins with log fold change above -log2(1.4) = 0.5 are upregulated, while proteins with log-fold change below log2(0.7) = -0.5 are considered as down regulated. Only As-binding proteins were included in this plot.

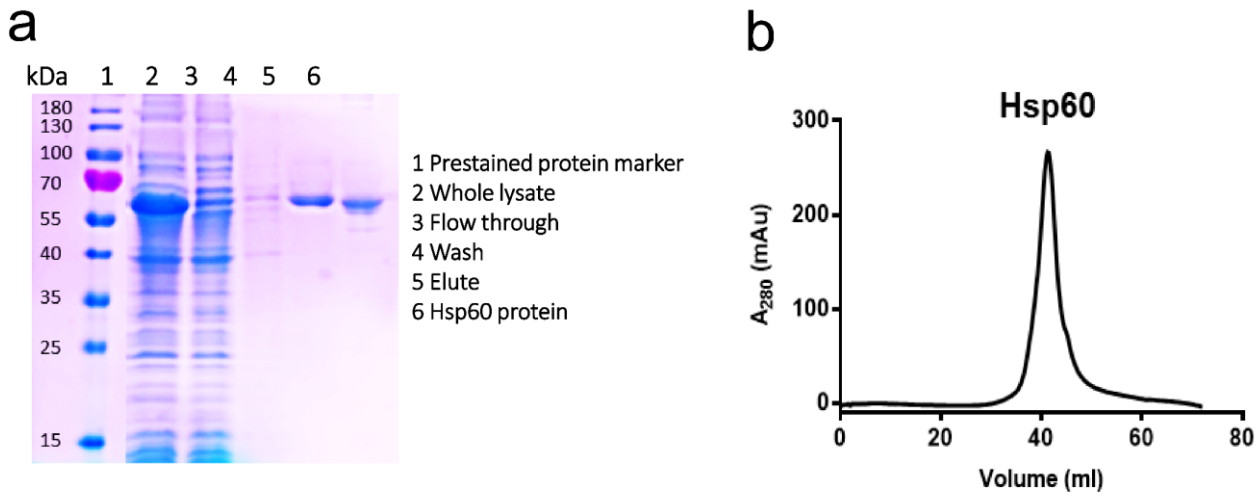



**Fig. S15** Overexpression and purification of Hsp60. (**a**) SDS-PAGE analysis showing the purity of Hsp60 protein. (**b**) Gel filtration profile of Hsp60 on a column Hiload 16/60 Superdex 200 (GE Healthcare) pre-equilibrated with 35 mM Tris-$HNO_3$, 100 mM $NaNO_3$, pH=7.4.

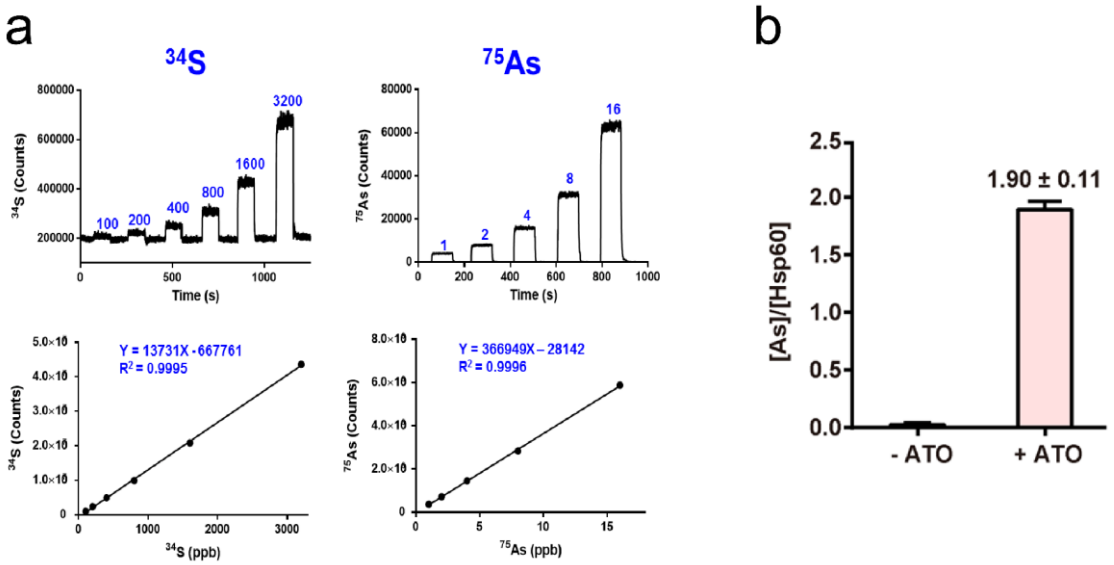



**Fig S16** (**a**) Standard time resolved ICP-MS profile of $^{34}S$ and $^{75}As$ in 1% $HNO_3$. (**b**) As-binding capability of Hsp60 measured by ICP-MS. About five molar equivalents of ATO were incubated with Hsp60 and bound arsenic was quantified by ICP-MS after removal of excess non-bound arsenic ions by size exclusion chromatography and Hsp60 concentration was calculated based on $^{34}S$ concentration.

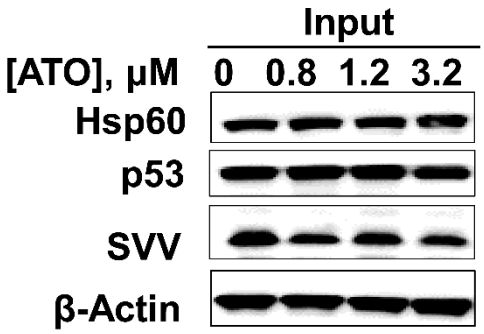
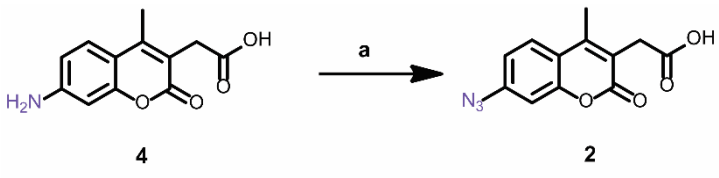
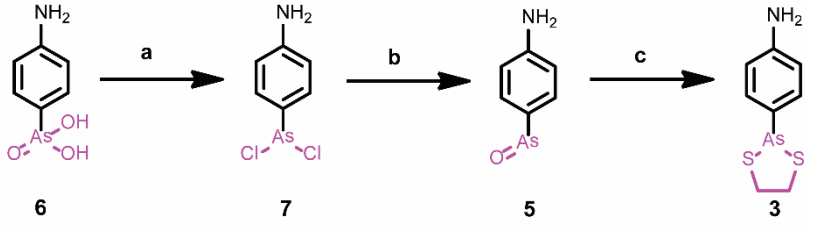


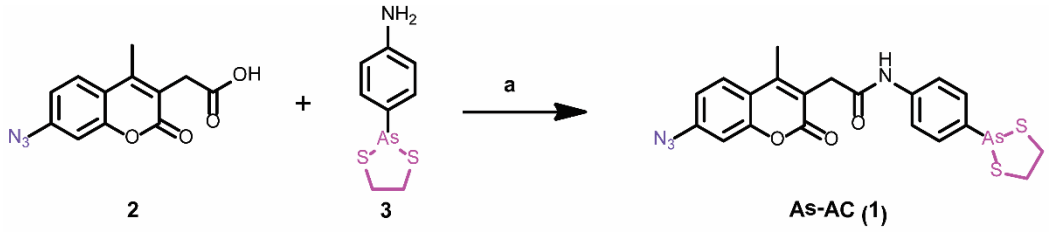

**Fig S17** Western blots of input groups for immunoprecipitation experiments. NB4 cell extracts were treated with 0, 0.8, 1.2 and 3.2 μM $As_2O_3$ for 6 hrs, and β-actin was used as a control.




**Scheme S1** Synthesis of coumarin azide (Compound 2). Condition: a) $H_2SO_4$, $H_2O$, $NaNO_2$, $NaN_3$, 0-5 $^oC$, 2 hrs, yield: 85%.




**Scheme S2** Synthesis of 2-*p*-aminophenyl-1,3,2-dithiarsenolane (Compound 3). Conditions: a)HCl, KI, $SO_2$, MeOH, r.t., 30 mins; b) 10% $NH_4OH$, r.t., 30 mins; c) 1,2-EDT, EtOH, r.t., 1h.




**Scheme S3** Synthesis of As-AC (Compound 1). Condition: a) EDCI, DIEA, dry DMF, r.t., overnight.

**Table S1** Amino acid sequences of purified proteins used for examination of protein labelling by As-AC *in vitro*.

| **Protein** | **Sequences** |
|---|---|
| BIR3 | ADSHMLPRNPSMADYEARIFTFGTWIYSVNKEQLARAGFYALGEGDKVK**C**FH**C**GGGLTDWKPSEDPWEQ HAKWYPG**C**KYLLEQKGQEYINNIHLTHSL |
| Ubiquitin | MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLR GG |
| SlyD | MQNHDLESIKQAALIEYEVREQGSSIVLDSNISKEPLEFIIGTNQIIAGLEKAVLKAQIGEWEEVVIAPEEAYGV YESSYLQEVPRDQFEGIELEKGMSVFGQTEDNQTIQAIIKDFSATHVMVDYNHPLAGKTLAFRFKVLGFREV SEEEILASHHGGGTG**CC**GGHGGHGGKKGGG**C**G**C**S**C**SHG |
| SlyDΔC | MQNHDLESIKQAALIEYEVREQGSSIVLDSNISKEPLEFIIGTNQIIAGLEKAVLKAQIGEWEEVVIAPEEAYGV YESSYLQEVPRDQFEGIELEKGMSVFGQTEDNQTIQAIIKDFSATHVMVDYNHPLAGKTLAFRFKVLGFREV SEEE |
| SlyD CmutA | MQNHDLESIKQAALIEYEVREQGSSIVLDSNISKEPLEFIIGTNQIIAGLEKAVLKAQIGEWEEVVIAPEEAYGV YESSYLQEVPRDQFEGIELEKGMSVFGQTEDNQTIQAIIKDFSATHVMVDYNHPLAGKTLAFRFKVLGFREV SEEEILASHHGGGTG**AA**GGHGGHGGKKGGG**A**G**A**S**A**SHG |

**Table S2** Summary of identified As(III)-binding proteins in NB4 cells by As-AC labelling.

| Gene name | Protein name | UniProt ID | MW (Da) | PI | The No. of Cys | Binding motif |
|---|---|---|---|---|---|---|
| EEF1B2 | Elongation factor 1-beta | P24534 | 24763.7 | 4.50 | 3 | $CX_nC$ |
| TPT1 | Translationally-controlled tumor protein | P13693 | 19595.3 | 4.84 | 2 | $CX_nC$ |
| NME1 | Nucleoside diphosphate kinase A | P15531 | 17148.7 | 5.81 | 3 | $CX_nC$ |
| TPI1 | Triosephosphate isomerase | P60174 | 30791.0 | 5.65 | 5 | $CX_nC$ |
| RANBP1 | Ran-specific GTPase-activating protein | P43487 | 23310.1 | 5.19 | 3 | $CX_nC$ |
| CLIC1 | Chloride intracellular channel protein 1 | O00299 | 26922.7 | 5.09 | 6 | $CX_nC$ |
| PPA1 | Inorganic pyrophosphatase | Q15181 | 32660.0 | 5.54 | 8 | $XCCX/CX_nC$ |
| LDHB | L-lactate dehydrogenase B chain | P07195 | 36638.5 | 5.71 | 5 | $CX_nC$ |
| RHOXF2 | Rhox homeobox family member 2 | Q9BQY4 | 31691.5 | 4.53 | 2 | $CX_nC$ |
| TPM3 | Tropomyosin alpha-3 chain | P06753 | 32950.0 | 4.68 | 1 | $CX_n$ |
| NPM1 | Nucleophosmin | P06748 | 32575.5 | 4.64 | 3 | $CX_nC$ |
| RPSA | 40S ribosomal protein SA | P08865 | 32854.1 | 4.79 | 2 | $CX_nC$ |
| CALR | Calreticulin | P27797 | 48141.6 | 4.29 | 3 | $CX_nC$ |
| RBBP4 | Histone-binding protein RBBP4 | Q09028 | 47655.7 | 4.74 | 5 | $CX_nC$ |
| P4HB | Protein disulfide-isomerase | P07237 | 57116.4 | 4.76 | 7 | $CXXC/CX_nC$ |
| HCLS1 | Hematopoietic lineage cell-specific protein | P14317 | 54014.0 | 4.73 | 2 | $CX_nC$ |
| HSPA5 | 78 kDa glucose-regulated protein | P11021 | 72333.0 | 5.07 | 2 | $CX_nC$ |
| TUBB | Tubulin beta chain | P07437 | 49670.8 | 4.78 | 8 | $CXCX/CX_nC$ |

| | | | | | | |
|---|---|---|---|---|---|---|
| VCP | Transitional endoplasmic reticulum ATPase | P55072 | 89321.8 | 5.14 | 12 | $CX_nC$ |
| ACTG1 | Actin, cytoplasmic 2 | P63261 | 41792.8 | 5.31 | 6 | $CX_nC$ |
| ST13 | Hsc70-interacting protein | P50502 | 41331.7 | 5.18 | 3 | $CX_nC$ |
| ACTB | Actin, cytoplasmic 1 | P60709 | 41736.7 | 5.29 | 6 | $CX_nC$ |
| HNRNPK | Heterogeneous nuclear ribonucleoprotein K | P61978 | 50976.3 | 5.39 | 5 | $CX_nC$/XCCX |
| HSPA8 | Heat shock cognate 71 kDa protein | P11142 | 70898.1 | 5.37 | 4 | $CX_nC$ |
| HSPA9 | Stress-70 protein, mitochondrial | P38646 | 73680.5 | 5.87 | 5 | $CX_nC$ |
| PDIA3 | Protein disulfide-isomerase A3 | P30101 | 56782.4 | 5.98 | 7 | CXXC/$CX_nC$ |
| HNRNPH1 | Heterogeneous nuclear ribonucleoprotein H | P31943 | 49229.5 | 5.89 | 5 | $CX_nC$ |
| STIP1 | Stress-induced-phosphoprotein 1 | P31948 | 62639.3 | 6.40 | 11 | $CX_nC$/CXXC |
| PHGDH | D-3-phosphoglycerate dehydrogenase 2 | O43175 | 56650.5 | 6.29 | 13 | XCCX/$CX_nC$ |
| IMPDH2 | Inosine-5’-monophosphate dehydrogenase 2 | P12268 | 55805.0 | 6.44 | 8 | $CX_nC$ |
| KHSRP | Far upstream element-binding protein 2 | Q92945 | 73115.3 | 6.85 | 5 | $CX_nC$ |
| ENO1 | Alpha-enolase | P06733 | 47169.0 | 7.01 | 6 | CXCX/$CX_nC$ |
| HNRNPA1 | Heterogeneous nuclear ribonucleoprotein A1 | P09651 | 38746.7 | 9.17 | 2 | $CX_nC$ |
| CCT5 | T-complex protein 1 subunit epsilon | P48643 | 59671.0 | 5.44 | 8 | $CX_nC$ |
| HSPD1 | 60 kDa heat shock protein | P10809 | 61054.6 | 5.70 | 3 | $CX_nC$ |
| PRDX2 | Peroxiredoxin-2 | P32119 | 21891.9 | 5.66 | 3 | $CX_nC$ |
| GSTP1 | Glutathione S-transferase P | P09211 | 23355.8 | 5.43 | 4 | $CX_nC$ |

**Table S3** Summary of identified As(III)-binding proteins in HL60 cells by As-AC labelling.

| Gene name | Protein name | UniProt ID | MW (Da) | PI | The No. of Cys | Binding motif |
|---|---|---|---|---|---|---|
| TPI1 | Triosephosphate isomerase | P60174 | 30791 | 5.65 | 5 | $CX_nC$ |
| NME2 | Nucleoside diphosphate kinase B | P22392 | 17298 | 8.52 | 2 | $CX_nC$ |
| PPIA | Peptidyl-prolyl cis-trans isomerase A | P62937 | 18012.5 | 7.68 | 4 | $CX_nC$ |
| NME1 | Nucleoside diphosphate kinase A | P15531 | 17148.73 | 5.81 | 3 | $CX_nC$ |
| EEF1B2 | Elongation factor 1-beta | P24534 | 24763.7 | 4.5 | 3 | $CX_nC$ |
| NPM1 | Nucleophosmin | P06748 | 32575.02 | 4.64 | 3 | $CX_nC$ |
| RPSA | 40S ribosomal protein SA | P08865 | 32854.08 | 4.79 | 2 | $CX_nC$ |
| CALR | Calreticulin | P27797 | 48141.56 | 4.29 | 3 | $CX_nC$ |
| P4HB | Protein disulfide-isomerase | P07237 | 57116.37 | 4.76 | 7 | $CXXC/CX_nC$ |
| PSMD4 | 26S proteasome non-ATPase regulatory subunit 4 | P55036 | 40736.7 | 4.68 | 4 | $CX_nC$ |
| ST13P4 | Putative protein FAM10A4 | Q8IZP2 | 27406.8 | 4.99 | 3 | $CX_nC$ |
| ACTG1 | Actin, cytoplasmic 2 | P63261 | 41792.8 | 5.31 | 6 | $CX_nC$ |
| PPA1 | Inorganic pyrophosphatase | Q15181 | 32660 | 5.54 | 8 | $CX_nC$ |
| HNRNPF | Heterogeneous nuclear ribonucleoprotein F | P52597 | 45671.9 | 5.37 | 6 | $CX_nC$ |
| PLIN3 | Perilipin-3 | O60664 | 47075 | 5.3 | 3 | $CX_nC$ |
| HSPA8 | Heat shock cognate 71 kDa protein | P11142 | 70898.09 | 5.37 | 4 | $CX_nC$ |

| | | | | | | |
|---|---|---|---|---|---|---|
| HSPA5 | 78 kDa glucose-regulated protein | P11021 | 72333 | 5.07 | 2 | $CX_nC$ |
| PDIA3 | Protein disulfide-isomerase A3 | P30101 | 56782.4 | 5.98 | 7 | $CXXC/CX_nC$ |
| HNRNPH1 | Heterogeneous nuclear ribonucleoprotein H | P31943 | 49229.5 | 5.89 | 5 | $CX_nC$ |
| CCT5 | T-complex protein 1 subunit epsilon | P48643 | 59671 | 5.44 | 8 | $CX_nC$ |
| HNRNPK | Heterogeneous nuclear ribonucleoprotein K | P61978 | 50976.3 | 5.39 | 5 | $CX_nC/XCCX$ |
| STIP1 | Stress-induced-phosphoprotein 1 | P31948 | 62639.3 | 6.4 | 11 | $CXXC/CX_nC$ |
| PHGDH | D-3-phosphoglycerate dehydrogenase | O43175 | 56650.5 | 6.29 | 13 | $XCCX/CX_nC$ |
| HNRNPA1 | Heterogeneous nuclear ribonucleoprotein A1 | P09651 | 38746.7 | 9.17 | 2 | $CX_nC$ |
| HNRNPM | Heterogeneous nuclear ribonucleoprotein M | P52272 | 77515.53 | 8.84 | 5 | $CX_nC$ |
| SERBP1 | Plasminogen activator inhibitor 1 RNA-binding protein | Q8NC51 | 44965.4 | 8.66 | 1 | $CX_n$ |
| PRDX1 | Peroxiredoxin-1 | Q06830 | 22110.4 | 8.27 | 4 | $CX_nC$ |
| HNRNPDL | Heterogeneous nuclear ribonucleoprotein D-like | O14979 | 46437.5 | 9.59 | 3 | $CX_nC$ |
| RANBP1 | Ran-specific GTPase-activating protein | P43487 | 23310.1 | 5.19 | 3 | $CX_nC$ |
| CLIC1 | Chloride intracellular channel protein 1 | O00299 | 26922.7 | 5.09 | 6 | $CX_nC$ |

| | | | | | | |
|---|---|---|---|---|---|---|
| GSTP1 | Glutathione S-transferase P | P09211 | 23355.8 | 5.43 | 4 | $CX_nC$ |

**Table S4** Biological process enrichment of genes differentially regulated in samples treated with ATO vs untreated ones. Top 10 GO terms with significant *p*-values were selected from DAVID. Z-score > 0 are those GO terms associated with genes up-regulated. X-score < 0 are those GO terms associated with genes down-regulated.

| GO ID | GO term name | Genes | Count | *p*-value | z-score |
|---|---|---|---|---|---|
| GO:0006413 | Translational initiation | RPL18, RPL17, RPL19, RPL15, RPL27A, RPL36, RPS3, LARP1, RPL30, EIF3B, RPL32, RPL7, RPL8, RPL11, RPL7A, RPL10A, RPL23A, RPL24, RPS6, RPS8, EIF4B, EIF4G2, RPS18, RPL23, EIF4H | 25 | 8.27E-20 | 2.60 |
| GO:0019083 | Viral transcription | RPL18, RPL17, RPL19, RPL15, RPL27A, RPL36, NUP85, RPL24, RPL23A, RPS6, RPS8, RPS3, RPL30, RPS18, RPL7, RPL32, RPL23, RPL8, RPL11, RANBP2, RPL7A, RPL10A | 22 | 4.15E-18 | 3.84 |
| GO:0006614 | ER translocation | RPL18, RPL17, RPL19, RPL15, RPL27A, RPL36, RPL24, RPL23A, RPS6, RPS8, RPS3, RPL30, RPS18, RPL7, RPL32, RPL23, RPL8, RPL11, RPL10A, RPL7A | 20 | 3.86E-17 | 3.58 |
| GO:0000184 | Nonsense-mediated mRNA decay | RPL18, RPL17, RPL19, MAGOH, RPL15, RPL27A, RPL36, RPL24, RPL23A, RPS6, RPS8, RPS3, RPL30, RPS18, RPL7, RPL32, RPL23, RPL8, RPL11, RPL7A, RPL10A | 21 | 2.49E-16 | 3.71 |
| GO:0006364 | rRNA processing | RPL18, RPL17, RPL19, EXOSC5, RPL15, RPL27A, RPL36, RPL24, RPL23A, RPS6, RPS8, RPS3, RPL30, RPS18, NOLC1, RPL7, RPL32, RPL23, DKC1, RPL8, DDX21, RPL11, RPL10A, RPL7A | 24 | 3.32E-14 | 3.26 |

| | | | | | |
|---|---|---|---|---|---|
| GO:0006457 | Protein folding | GRPEL1, PDIA3, FKBP5, ERP29, PDIA4, CANX, LRPAP1, AIP, ST13, TRAP1, HSP90B1, CSNK2A1, TXNDC5, BAG2, DNAJA1, HSPA4L, DNAJB2, HSPE1, RANBP2, HSPA9 | 20 | 8.49E-12 | 2.68 |
| GO:0006412 | Translation | RPL18, RPL17, RPL19, RPL15, RPL27A, RPL36, RPL24, RPL23A, RPS6, RPS8, RPS3, RPL30, RPS18, RPL7, RPL32, RPL23, RPL8, FARSB, RPL11, RPL7A, RPL10A | 21 | 4.52E-10 | 3.27 |
| GO:0002181 | Cytoplasmic translation | EIF4B, RPL7, EIF4H, RPL15, RPL8, RPL36 | 6 | 2.17E-05 | 0.82 |
| GO:0061077 | Chaperone-mediated protein folding | TRAP 1, CSNK2A1, PPIB, FKBP5, DNAJB2, PDIA4, CANX | 7 | 1.14E-05 | 1.89 |
| GO:0034975 | Protein folding in endoplasmic reticulum | HSP90B1, PDIA3, HSPA5, CANX | 4 | 6.98E-04 | 2.00 |
| GO:0051260 | Protein homooligomerization | BID, ST13, ANXA6, LONP1, CHMP4A, HMOX1, RBMX, ACAT1 | 8 | 0.012477 | -0.71 |
| GO:0006446 | Regulation of translational initiation | EIF4B, EIF4G2, EIF3B, EIF4H | 4 | 0.013753 | -1.00 |
| GO:0000387 | Spliceosomal snRNP assembly | PRMT5, GEMIN6, SNRPF, GEMIN5 | 4 | 0.006848 | -1.00 |
| GO:0010467 | Gene expression | HNRNPK, IGF2BP1, HNRNPR, RBMX, HNRNPA0 | 5 | 0.004492 | -0.45 |
| GO:1903507 | Negative regulation of nucleic acid-templated transcription | SRSF2, PRMT5, SF1, DDX54, APEX1 | 5 | 0.004163 | -0.45 |

| | | | | | |
|---|---|---|---|---|---|
| GO:0043488 | Regulation of mRNA stability | PSMB4, SET, PSMA3, EXOSC5, HSPA1B, PSMD6, APEX1 | 7 | 0.00326 | -0.38 |
| GO:0008380 | RNA splicing | SRSF2, PRPF8, SREK1, AKAP8L, SYNCRIP, ACIN1, SNRPF, SF3B2, PRPF6 | 9 | 0.002371 | -1.00 |
| GO:0016032 | Viral process | PSMB4, TLN1, HNRNPK, SET, IPO7, EIF4H, PSMA3, SYNCRIP, NUP85, RAB6A, HSPD1, RANBP2, KARS, SF3B2 | 14 | 3.14E-04 | -0.53 |
| GO:0000398 | mRNA splicing, via spliceosome | CRNKL1, MAGOH, SF1, SYNCRIP, RBMX, HNRNPR, HNRNPA0, YBX1, SF3B2, PRPF6, SRSF2, HNRNPK, PRPF8, GEMIN6, SNRPF, GEMIN5 | 16 | 5.63E-07 | -1.00 |
| GO:0098609 | Cell-cell adhesion | SEPT2, RPL15, SLC3A2, SNX1, HCFC1, EPS15L1, RPL24, RPL23A, LARP1, EIF4G2, ATXN2L, HNRNPK, EIF4H, USO1, HSPA5, TMPO, RPL7A, PAICS, AHNAK, ADD1 | 20 | 9.18E-09 | -1.34 |

**Table S5** Cellular component enrichment of genes differentially regulated in samples treated with or without ATO. Top 10 terms with significant *p*-values were selected from DAVID. Z-score > 0 are those GO terms associated with genes up-regulated. Z-score < 0 are those down-regulated.

| GO ID | GO term name | Genes | Count | *p*-value | z-score |
|---|---|---|---|---|---|
| GO:0022625 | Cytosolic large ribosomal subunit | RPL18, RPL17, RPL19, RPL15, RPL27A, RPL36, RPL24, RPL23A, RPL30, RPL7, RPL32, RPL23, RPL8, RPL11, RPL7A, RPL10A | 16 | 7.06E-15 | 3.00 |
| GO:0005840 | Ribosome | RPL18, RPL19, RPL15, SF1, RPL27A, RPL36, RPL24, RPL23A, RPS6, CANX, RPS8, RPS3, RPL30, RPS18, RPL7, RPL32, RPL23, RPL8, RPL11, RPL7A, RPL10A, APEX1 | 22 | 3.62E-15 | 3.41 |
| GO:0005925 | Focal adhesion | RPL18, TLN1, RPL19, PDIA3, HSPA1B, RPS3, ANXA6, RPL30, HNRNPK, RPL7, RPL8, RPL7A, RPL10A, SNAP23, CAP1, HSPA5, AHNAK, HSPA9, ACTN4, HMGA1, RPS8, CORO1C, HSP90B1, RPS18, PPIB, RPL23, ADD1 | 27 | 7.97E-12 | 2.89 |
| GO:0042470 | Melanosome | AHCY, PDIA3, ERP29, SLC3A2, NAP1L1, PDIA4, CANX, ANXA6, LAMP1, HSP90B1, PPIB, ANXA11, HSPA5, RAB27A | 14 | 5.85E-10 | 2.14 |
| GO:0043209 | Myelin sheath | SEPT2, UQCRC1, PDIA3, PHB, PRDX3, CANX, MIF, GOT2, NME1-NME2, DLD, PHGDH, HSPA5, HSPD1, TUBA1B, MDH2, MDH1, HSPA9 | 17 | 1.27E-10 | 1.21 |
| GO:0005739 | Mitochondrion | SRI, BCAT1, BID, GRPEL1, UQCRC1, ALDH18A1, GLUD1, HCFC1, ECHS1, HSPA1B, PRDX3, COX5B, ACAT1, | 44 | 1.84E-08 | 3.62 |

| | | | | | |
|---|---|---|---|---|---|
| | | HADHA, KARS, HADHB, GOT2, ANXA6, LONP1, SLC16A1, KRT5, PGAM5, DNAJA1, ACOT13, HSPE1, HSPA5, RANBP2, APEX1, HSPA9, SSBP1, PHB, QDPR, TIMM44, TRAP1, PHB2, DLD, MPO, ADSL, HSPD1, PMPCA, MDH2, GSTP1, MDH1, MYH10 | | | |
| GO:0005730 | Nucleolus | RPL18, XRCC5, CDV3, SEPT2, RPL19, MTDH, RPL36, RPS3, PPAN, DKC1, RPL7, HMOX1, RPL8, LEO1, DDX21, RPL11, RPL10A, RPL7A, ACIN1, APEX1, EWSR1, HSPA9, EXOSC5, RPL23A, RPS6, HNRNPR, NCL, RPL23, NOLC1, USO1, SMARCA5, CIRBP, DDX54 | 33 | 6.32E-08 | 1.91 |
| GO:0030529 | Intracellular ribonucleoprotein complex | HNRNPK, ACTN4, RPL7, IGF2BP1, SYNCRIP, RPS6, NCL, HNRNPR, RBMX, HNRNPA0, YBX1, RPS8, RPS3 | 13 | 1.98E-07 | 0.83 |
| GO:0005634 | Nucleus | XRCC5, RPL18, RPL17, GRPEL1, SEPT2, CHMP4A, RPL15, SYNCRIP, BTK,<br>CSNK2A1, CEP290, DDX21, ACIN1, ACTN4, MAGOH, STK4, RNASEH2B, RPS18,<br>PSMA3, SMARCA5, MDH2, ADD1, RAD23B, HMGB3, AHCY, ASS1, IGF2BP1,<br>RABGAP1L, KARS, PSMB4, PPAN, TUBB, HNRNPK, RPL7, HEXIM2, DNAJA1,<br>RPL10A, RPL7A, EWSR1, PHB, AKAP8L, RPL23A, RPS6, | 108 | 1.87E-07 | 0.19 |

| | | | | | |
|---|---|---|---|---|---|
| | | RBMX, HNRNPA0, RPS8, SRSF2, PPIB, POLD1, SUPT16H, DNMT1, CIRBP, DNAJB2, TMPO, DDX54, MYH10, PDIA3, NAP1L1, EPS15L1, PTMA, YBX1, RPS3, FUBP1, CASP3, LONP1, DKC1, PRMT5, HMOX1, ACOT13, AHNAK, HIST1H1C, SSBP1, EXOSC5, SLC3A2, SF1, NCL, HMGA1, MCM5, MCM6, PRPF6, PPM1G, GTF2I, UBE2K, GSTP1, CTSG, MTDH, HCFC1, RPL30, SET, KRT5, DDX19A, SAFB, PRPF8, KRT1, HSPA5, GEMIN6, APEX1, CHD5, GEMIN5, NUCKS1, SAMHD1, COTL1, HSP90B1, NME1-NME2, PHB2, HSPA4L, MPO, FAF1 | | | |
| GO:0005759 | Mitochondrial matrix | GRPEL1, SSBP1, GLUD1, ECHS1, PRDX3, TIMM44, ACAT1, KARS, RPS3, GOT2, TRAP1, LONP1, DLD, HSPE1, HSPD1, PMPCA, PCCA, MDH2 | 18 | 1.26E-06 | 3.77 |
| GO:0005654 | Nucleoplasm | XRCC5, SYNCRIP, PTMA, YBX1, RPS3, AIP, LARP1, FUBP1, CASP3, EIF4EBP1, LONP1, CSNK2A1, RAD21, DKC1, PRMT5, RPL11, DDX21, ACIN1, PSMD6, MAGOH, EXOSC5, SF1, POLR1C, NCL, HNRNPR, HMGA1, MCM5, MCM6, PRPF6, TRAP1, PPM1G, RPS18, BAZ1B, IPO7, GTF2I, PSMA3, DLD, SMARCA5, KPNA4, SNRPF, ADD3, MDH2, ADD1, SRI, RAD23B, FKBP5, RNH1, HCFC1, HSPA1B, SF3B2, MIF, PSMB4, EIF3B, HNRNPK, SET, HEXIM2, PRPF8, SAFB, LEO1, SNAP23, GEMIN6, APEX1, | 83 | 6.28E-14 | -1.21 |

| | | | | | |
|---|---|---|---|---|---|
| | | GEMIN5, PM20D2, SREK1, PHB, AKAP8L, SAMHD1, ACLY, RPS6, RBMX, RPS8, HNRNPA0, SRSF2, RPRD2, NOLC1, UBA2, POLD1, ANXA11, SUPT16H, DNMT1, CIRBP, TRMT112 | | | |
| GO:0005737 | Cytoplasm | RPL18, SEPT2, RPL19, CHMP4A, IQGAP2, TXLNA, MED22, AIP, BTK, EIF4EBP1, CEP290, RPL11, ACIN1, ACTN4, QDPR, STK4, ST13, IGBP1, PSMA3, STMN1, ADD3, MDH1, SRI, RAD23B, CDV3, HMGB3, AHCY, ASS1, GLUD1, SNX1, IGF2BP1, HSPA1B, KARS, PSMB4, TUBB, EIF3B, HNRNPK, HEXIM2, RPL7, PLCH1, RPL8, HSPE1, RPL10A, RPL7A, EWSR1, PHB, AKAP8L, RPL23A, OXSR1, RPS6, RPS8, PPA1, CORO1C, LAMP1, RABEP2, NOLC1, POLD1, PRKAR1A, CIRBP, HSPD1, TMPO, MYH10, TLN1, METAP2, CRNKL1, PRDX3, PTMA, YBX1, LARP1, RPS3, CASP3, LONP1, DKC1, PRMT5, AHNAK, KIF11, EXOSC5, SLC3A2, PPM1G, IPO7, UBE2K, GTF2I, DDT, FARSB, TMSB4X, GSTP1, BID, SNAP29, MTDH, ALDH18A1, RNH1, HCFC1, RPL36, MIF, RPL30, ATXN2L, SET, KRT5, DDX19A, CSDE1, SNAP23, GEMIN6, APEX1, GEMIN5, HSPA9, NUCKS1, ACLY, RPL24, COTL1, RPL23, NME1-NME2, PHB2, ANXA11, HSPA4L, PAICS | 115 | 7.88E-11 | -0.09 |
| GO:0005913 | Cell-cell adherents | TLN1, SEPT2, RPL15, SLC3A2, SNX1, HCFC1, EPS15L1, | 22 | 1.36E-09 | -1.28 |

| | | | | | |
|---|---|---|---|---|---|
| | junction | RPL24, RPL23A, HSPA1B, LARP1, EIF4G2, ATXN2L, HNRNPK, EIF4H, USO1, HSPA5, TMPO, RPL7A, PAICS, AHNAK, ADD1 | | | |
| GO:0005681 | Spliceosomal complex | SRSF2, CRNKL1, PRPF8, SREK1, SF1, SNRPF, HNRNPR, SF3B2, PRPF6 | 9 | 2.91E-05 | -0.33 |
| GO:0016607 | Nuclear speck | SRSF2, ATXN2L, CRNKL1, MAGOH, PRPF8, AKAP8L, ACIN1, APEX1, PRPF | 9 | 0.004696 | -1.00 |
| GO:0016281 | Eukaryotic translation initiation factor 4F complex | EIF4B, EIF4G2, EIF4H | 3 | 0.005615 | -0.58 |
| GO:0034719 | SMN-Sm protein complex | GEMIN6, SNRPF. GEMIN5 | 3 | 0.019824 | -1.73 |
| GO:0036464 | Cytoplasmic ribonucleoprotein granule | TUBB, RPS6, NCL | 3 | 0.04089 | -0.59 |
| GO:0005689 | U12-type spliceosomal complex | SNRPF, YBX1. SF3B2 | 3 | 0.043929 | -0.58 |

**Table S6** Molecular function enrichment of genes differentially regulated in samples treated with ATO vs untreated ones. Top 10 GO terms with significant *p*-values were selected from DAVID. Z-score > 0 are those GO terms associated with genes up-regulated. Z-score < 0 are those GO terms associated with genes down-regulated.

| GO ID | GO term name | Genes | Count | *p*-value | z-score |
|---|---|---|---|---|---|
| GO:0003723 | RNA binding | L18, RPL19, CRNKL1, RPL15, RPL27A, IGF2BP1, SYNCRIP, YBX1, RPS3, FUBP1, RPL30, HNRNPK, EIF3B, DKC1, RPL7, DDX19A, PRPF8, RPL8, RPL11, RPL7A, RPL10A, RANBP2, EWSR1, EXOSC5, SF1, SAMHD1, RPL24, RBMX, HNRNPR, NCL, HNRNPA0, PRPF6, EIF4B, EIF4G2, RPS18, HSP90B1, EIF4H, FARSB, CIRBP, DNMT1, SNRPF, DDX54 | 42 | 5.14E-19 | 1.54 |
| GO:0005515 | Protein binding | XRCC5, RPL17, RPL19, SEPT2, CHMP4A, RPL15, SYNCRIP, STOML2, MED22, TXLNA, COX5B, BTK, AIP, EIF4EBP1, CSNK2A1, RAD21, CEP290, RPL11, DDX21, ACIN1, DCTPP1, RAB27A, IRS2, ACTN4, MAGOH, HNRNPR, STK4, ST13, RPS18, BAZ1B, TXNDC5, IGBP1, PSMA3, USO1, SMARCA5, TUBA4A, STMN1, SNRPF, MDH1, SRI, RAD23B, AHCY, HMGB3, CCDC12, ASS1, DIAPH1, GLUD1, IGF2BP1, SNX1, HSPA1B, KARS, HADHA, HADHB, PEA15, PSMB4, TUBB, HNRNPK, EIF3B, HEXIM2, RPL7, DNAJA1, LEO1, HSPE1, RPL7A, RPL10A, EWSR1, PHB, SREK1, MET, AKAP8L, RPL23A, OXSR1, | 182 | 7.57E-16 | 0.59 |

| | | | | | |
|---|---|---|---|---|---|
| | | RPS6, RBMX, EIF4B, CORO1C, SRSF2, LAMP1, NOLC1, PPIB, EIF4H, POLD1, UBA2, PRKAR1A, SUPT16H, CIRBP, DNMT1, DNAJB2, HSPD1, MYH10, TLN1, PDIA3, CRNKL1, NAP1L1, EPS15L1, PRDX3, PDIA4, PTMA, CANX, YBX1, DSTN, RPS3, LARP1, FUBP1, LONP1, CASP3, DKC1, PRMT5, HMOX1, BAG2, RAB6A, RANBP2, SEC24C, PSMD6, TUBA1B, AHNAK, SSBP1, HIST1H1C, EXOSC5, SLC3A2, SF1, NUP85, POLR1C, TIMM44, NCL, HMGA1, MCM5, PRPF6, LRPAP1, MCM6, EIF4G2, TRAP1, PPM1G, IPO7, UBE2K, GTF2I, FARSB, TMSB4X, KPNA4, CTSG, GSTP1, BID, SNAP29, DCD, MTDH, ALDH18A1, FKBP5, RNH1, RPL27A, HCFC1, ECHS1, MIF, SF3B2, ANXA6, RPL30, SET, ATXN2L, KRT5, DDX19A, PRPF8, SAFB, KRT1, CSDE1, HSPA5, SNAP23, GEMIN6, APEX1, GEMIN5, HSPA9, PLP2, SAMHD1, RPL24, ACLY, COTL1, HSP90B1, SCFD1, RPL23, PHB2, ANXA11, FAF1, PAICS, TRMT112 | | | |
| GO:0051082 | Unfolded protein binding | GRPEL1, HSPA1B, CANX, SERPINH1, AIP, LRPAP1, ST13, TRAP1, HSP90B1, PPIB, DNAJA1, HSPE1, DNAJB2, HSPA5, HSPD1, HSPA9 | 16 | 3.13E-11 | 3.50 |
| GO:0003735 | Structural constituent of ribosome | RPL18, RPL17, RPL19, RPL15, RPL27A, RPL36, RPL24, RPL23A, RPS6, RPS8, RPS3, RPL30, RPS18, RPL7, RPL32, | 20 | 2.80E-10 | 3.58 |

| | | | | | |
|---|---|---|---|---|---|
| | | RPL23, RPL8, RPL11, RPL10A, RPL7A | | | |
| GO:0051087 | Chaperone binding | ST13, GRPEL1, ERP29, BAG2, DNAJA1, DNAJB2, HSPE1, HSPD1, HSPA5, TIMM44 | 10 | 1.78E-06 | 1.90 |
| GO:0019899 | Enzyme binding | PHB, HSPA1B, COTL1, ACAT1, HMGA1, RPS3, GOT2, LAMP1, HMOX1, POLD1, TUBA4A, ACIN1, HSPA5, PCCA | 14 | 7.91E-04 | 1.60 |
| GO:0019843 | rRNA binding | PPAN, RPS18, RPL8, DDX21, RPL11, RPL23A | 6 | 2.17E-04 | 1.63 |
| GO:0031625 | Ubiquitin protein ligase | BID, XRCC5, UQCRC1, HSPA1B, TUBB, UBE2K, PRKAR1A, DNAJA1, DNAJB2, HSPD1, FAF1, HSPA5, TUBA1B, HSPA9 | 14 | 1.92E-04 | 0.53 |
| GO:0005524 | ATP binding | XRCC5, ALDH18A1, ASS1, GLUD1, HSPA1B, KARS, BTK, LONP1, CSNK2A1, DDX19A, DNAJA1, HSPE1, DDX21, HSPA5, CHD5, HSPA9, KIF11, MET, ACLY, OXSR1, TIMM44, STK4, MCM5, MCM6, TRAP1, HSP90B1, BAZ1B, NME1-NME2, NOLC1, UBE2K, UBA2, FARSB, HSPA4L, SMARCA5, HSPD1, DDX54, PAICS, PCCA, MYH10 | 39 | 1.84E-04 | 1.44 |
| GO:0032403 | Protein complex binding | ST13, TUBB, CASP3, UQCRC1, KIF11, PPIB, PGAM5, PRMT5, IGBP1, RANBP2, APEX1, HADHA | 12 | 1.45E-04 | 1.15 |
| GO:0098641 | Cadherin binding involved in cell-cell adhesion | TLN1, SEPT2, RPL15, SLC3A2, SNX1, HCFC1, EPS15L1, RPL24, RPL23A, HSPA1B, LARP1, EIF4G2, ATXN2L, HNRNPK, EIF4H, USO1, HSPA5, TMPO, RPL7A, PAICS, AHNAK, ADD1 | 22 | 6.90E-10 | -1.28 |
| GO:0000166 | Nucleotide binding | SREK1, IGF2BP1, SYNCRIP, RPL23A, NCL, RBMX, | 17 | 2.96E-05 | -1.70 |

| | | | | | |
|---|---|---|---|---|---|
| | | HNRNPR, HNRNPA0, EIF4B, SRSF2, EIF3B, POLD1, EIF4H, SAFB, CIRBP, ACIN1, EWSR1 | | | |
| GO:0042802 | Identical protein binding | SREK1, IGF2BP1, SYNCRIP, RPL23A, NCL, RBMX, HNRNPR, HNRNPA0, EIF4B, SRSF2, EIF3B, POLD1, EIF4H, SAFB, CIRBP, ACIN1, EWSR1 | 24 | 3.09E-04 | -1.63 |
| GO:0051287 | NAD binding | AHCY, DLD, PHGDH, HADHA, MDH | 5 | 0.001835 | -0.45 |
| GO:0003684 | Damaged DNA binding | XRCC5, RAD22B, POLD1, APEX1, RPS3 | 5 | 0.011321 | -1.34 |
| GO:0005200 | Structural constituent of cytoskeleton | TUBB, TLN1, KRT5, TUBA4A, ADD3, TUBA1B | 6 | 0.018541 | -1.63 |
| GO:0034057 | RNA strand-exchange activity | EIF4B, EIF4H | 2 | 0.027532 | -1.41 |
| GO:0016597 | Amino acid binding | GOT2, ASS1, KARS | 3 | 0.03684 | -0.58 |
| GO:0004167 | Dopachrome isomerase activity | DDT, MIF | 2 | 0.041014 | -1.41 |

**Dataset S1 (separate file)** List of GO terms of biological process, cellular component and molecular function annotation lists of 250 As-regulated proteins in NB4 cells.

**Dataset S2 (separate file)** List of GO terms of biological process annotation list of 37 As-binding proteins in NB4 cells.

## References


1. Y. Liu, D. Duan, J. Yao, B. Zhang, S. Peng, H. Ma, Y. Song and J. Fang, *J. Med. Chem.*, 2014, **57**, 5203-5211.
2. T. Cheng, H. Li, W. Xia, H. Sun, *J. Biol. Inorg. Chem.*, 2012, **17**, 331-343.
3. Y. -T. Lai, Y. -Y. Chang, L. Hu, Y.Yang, A.Chao, Z. -Y. Du, J. A. Tanner, M. -L. Chye, C. Qian, K. -M. Ng, H. Li and H. Sun, *Proc. Natl. Acad. Sci. USA*, 2015, **112**, 2948-2953.
4. A. T. R. Williams, S. A. Winfield, J. N. Miller, *Analyst*, 1983, **108**, 1067-1071.
5. Y. -T. Lai, Y. Yang, L. Hu, T. Cheng, Y. -Y. Chang, M. Koohi-Moghadam, Y. Wang, J. Xia, J. Wang, H. Li and H. Sun, *Metallomics*, 2017, **9**, 38-47.
6. J. Rappsilber, M. Mann and Y. Ishihama, *Nat. Protoc.*, 2007, **2**, 1896-1906.
7. J. Cox and M. Mann, *Nat. Niotechnol.*, 2008, **26**, 1367-1372.
8. UniProt Consortium, *Nucleic Acid Res.*, 2016, **45**, D158-D169.
9. A. Franceschini, D. Szklarczyk, S. Franklid, M. Kuhn, M. Simonovic, A. Roth, J. Lin, P. Minguez, P. Bork, C. Von Mering and L. J. Jensen, *Nucleic Acids. Res.*, 2012, **41**, D808-D815.
10. P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski and T. Ideker, *Gemome Res.*, 2003, **13**, 2498-2504.
11. G. Dennis Jr., B. T. Sherman, D. A. Hosack, J. Yang, W. Gao, H. C. Lane and R. A. Lempicki, *Genome Biol.*, 2003, **4**, R60.
12. W. Walter, F. Sanchez-Cabo and M. Ricote, *Bioinformatics*, 2015, **31**, 2912-2914.