

Table S1. Arsenic binding affinity for single- and double-cysteine motifs.

| Type | Motif sequence | ΔG (kcal/mol) |
|-----------------|------------------------------|-----------------------|
| Single-cysteine | CEKKAAL ESK | -0.53 |
| | G C LAVERIEK | -1.07 |
| | EA C LEKKAAL | -1.21 |
| | SAD C LKTSAK | -1.06 |
| | QALR C LRQEE | -1.04 |
| | SSDRP C GLAV | -1.08 |
| | ASGQKS C KNG | -0.85 |
| | NRDQGLM C EK | -1.07 |
| | PSSRRTTL C G | -1.21 |
| | AKLAAAV P E C | -0.68 |
| Double-cysteine | KG CC MEIMSL | -3.81 |
| | VISN CV CNLS | -5.10 |
| | HQQ C WG C HEK | -4.53 |
| | KKI I CTIG C P | -3.78 |
| | QA C IPQA C AP | -1.22 |

Table S2. Arsenic-binding proteins with experimentally verified arsenic binding sites.

| No. | Protein Name | Abbr. | Accession | Ars binding site | Ref. |
|-----|--|-------|-----------|----------------------|------|
| 1 | Arsenical pump-driving ATPase | ArsA | P52145 | C113, C172, C422 | [1] |
| 2 | Arsenate reductase, ArsC | ArsC | P0AB96 | C12, C13, C18 | [2] |
| 3 | Arsenical resistance operon trans-acting repressor | ArsD | P46003 | C12, C13, C112, C113 | [3] |
| 4 | Arsenic methyltransferase | ArsM | C0JV69 | C72, C174, C224 | [4] |
| 5 | Arsenical resistance operon repressor | ArsR | P15905 | C32, C34, C37 | [1] |
| 6 | promyelocytic leukemia protein | PML | Q9BZY3 | C60, C77, C80 | [5] |
| 7 | Casitas B-lineage lymphoma | CBL | P22681 | C381, C384 | [6] |
| 8 | cAMP-specific 3',5'-cyclic phosphodiesterase 4B | PDE4B | Q07343 | C281, C348 | [7] |
| 9 | Thyroid hormone receptor beta | THRB | P10828 | C190, C236 | [8] |
| 10 | Acidic cytochrome c3 | | P94690 | | [9] |

Table S3. Proteins in the MALDI-MS data set.

| No. | Protein name | No. | Protein name |
|------------|---------------------|------------|---------------------|
| 1 | ACTB | 26 | KRT1 |
| 2 | ACTN4 | 27 | LDHB |
| 3 | AK1 | 28 | MCCC1 |
| 4 | ANXA2 | 29 | MCM6 |
| 5 | ARHGDI1 | 30 | MDH1 |
| 6 | ATP5F1A | 31 | NOP2 |
| 7 | CCT4 | 32 | NPEPPS |
| 8 | CCT8 | 33 | PC |
| 9 | CLTC | 34 | PDIA3 |
| 10 | EEF1D | 35 | PGD |
| 11 | EIF4A1 | 36 | PGK1 |
| 12 | ENO1 | 37 | PHGDH |
| 13 | ENO2 | 38 | PKM2 |
| 14 | FLNA | 39 | PPT1 |
| 15 | GAPDH | 40 | PRDX4 |
| 16 | GPB1 | 41 | PSMA7 |
| 17 | GRP94 | 42 | RNA |
| 18 | HNRNPA2 | 43 | RPN1 |
| 19 | HNRNPC | 44 | TCEA2 |
| 20 | HNRNPK | 45 | TKT |
| 21 | HSF1 | 46 | TPI1 |
| 22 | HSP90B | 47 | TUBB2 |
| 23 | HSPA4 | 48 | TUBB5 |
| 24 | HSPB1 | 49 | UBA1 |
| 25 | KPNB1 | 50 | VWHAQ |

Table S4. Proteins in the protein microarray data set.

| No. | Protein Name | No. | Protein Name | No. | Protein Name |
|-----|--------------|-----|--------------|-----|--------------|
| 1 | AAMP | 121 | FGL1 | 241 | PABPC4 |
| 2 | ABHD12B | 122 | FKBP5 | 242 | PACSN2 |
| 3 | ABHD5 | 123 | FLJ12716 | 243 | PAK2 |
| 4 | ACF | 124 | FLJ45256 | 244 | PCBP1 |
| 5 | ACOX1 | 125 | FLJ45983 | 245 | PCBP2 |
| 6 | ACSBG1 | 126 | FSCN1 | 246 | PCDH17 |
| 7 | ADH1B | 127 | FST | 247 | PCSK7 |
| 8 | ADH5 | 128 | FTH1 | 248 | PDCD2 |
| 9 | ADSL | 129 | G6PD | 249 | PFN2 |
| 10 | AES | 130 | GALE | 250 | PFN4 |
| 11 | AKR1A1 | 131 | GALK1 | 251 | PGD |
| 12 | AKR1C1 | 132 | GALK2 | 252 | PGK1 |
| 13 | AKR1C2 | 133 | GAPDH | 253 | PGM1 |
| 14 | AKR1C3 | 134 | GART | 254 | PGM3 |
| 15 | AKR1C4 | 135 | GAS2 | 255 | PHB |
| 16 | AKR7A3 | 136 | GBP5 | 256 | PHGDH |
| 17 | ALDH1A1 | 137 | GCK | 257 | PHOSPHO2 |
| 18 | ALDH1A3 | 138 | GDI1 | 258 | PHYHIPL |
| 19 | ALDH4A1 | 139 | GDI2 | 259 | PIGH |
| 20 | ALDOC | 140 | GGA1 | 260 | PKNOX2 |
| 21 | ALKBH4 | 141 | GGA3 | 261 | PLA2G6 |
| 22 | AMD1 | 142 | GKN1 | 262 | PLA2R1 |
| 23 | ANKHD1 | 143 | GLTP | 263 | PLCD1 |
| 24 | ANKRD39 | 144 | GLUL | 264 | PLEKHM2 |
| 25 | ANKRD40 | 145 | GMPS | 265 | PML |
| 26 | ANKS1A | 146 | GOLPH3 | 266 | POLD1 |
| 27 | ANXA11 | 147 | GOPC | 267 | PPARBP |
| 28 | ANXA9 | 148 | GPD1 | 268 | PPIL3 |
| 29 | AP2B1 | 149 | GPKOW | 269 | PPM1A |
| 30 | APEG1 | 150 | GPR119 | 270 | PRCC |
| 31 | APOE | 151 | GPR23 | 271 | PRMT5 |
| 32 | APTX | 152 | GSC | 272 | PRNP |
| 33 | ARFGAP3 | 153 | GSN | 273 | PRPS2 |
| 34 | ARL8B | 154 | GSTM5 | 274 | PSMB6 |
| 35 | ARMETL1 | 155 | GSTZ1 | 275 | PSMC3 |
| 36 | ARPC3 | 156 | HACL1 | 276 | PSMD3 |
| 37 | ASB11 | 157 | HAGH | 277 | PSMD7 |
| 38 | ASB9 | 158 | HARS | 278 | RAB17 |
| 39 | ASRGL1 | 159 | HARS2 | 279 | RAB2B |
| 40 | ATP6V1E1 | 160 | HK1 | 280 | RABL3 |

| | | | | | |
|----|-----------|-----|-----------|-----|----------|
| 41 | AURKC | 161 | HNMT | 281 | RAD23A |
| 42 | AXIIR | 162 | HSD3B2 | 282 | RAP1GDS1 |
| 43 | BCAR1 | 163 | HSPC142 | 283 | RBM12 |
| 44 | BCCIP | 164 | HSPC159 | 284 | RBP1 |
| 45 | BCL2L12 | 165 | IGFBP2 | 285 | RBP2 |
| 46 | BDH2 | 166 | IGHA1 | 286 | RNF144 |
| 47 | BIRC7 | 167 | IGHV1-69 | 287 | RNH1 |
| 48 | BXDC1 | 168 | IGKC | 288 | RNPEP |
| 49 | C10orf130 | 169 | IGKV1-5 | 289 | RP2 |
| 50 | C12orf5 | 170 | IGL@ | 290 | RPS6KB2 |
| 51 | C13orf1 | 171 | IGLV3-21 | 291 | RPSA |
| 52 | C16orf3 | 172 | IGLV3-25 | 292 | RQCD1 |
| 53 | C16orf75 | 173 | IL1F10 | 293 | RRM1 |
| 54 | C17orf57 | 174 | IL1F6 | 294 | RUVBL1 |
| 55 | C1orf123 | 175 | IL1RN | 295 | RUVBL2 |
| 56 | C1orf59 | 176 | IL3RA | 296 | SAT2 |
| 57 | C20orf195 | 177 | ILKAP | 297 | SBSN |
| 58 | C20orf52 | 178 | JOSD2 | 298 | SCARB2 |
| 59 | C8orf32 | 179 | KCTD5 | 299 | SCIN |
| 60 | C9orf103 | 180 | KCTD9 | 300 | SCLY |
| 61 | C9orf9 | 181 | KHK | 301 | SENP1 |
| 62 | CA3 | 182 | KIAA0828 | 302 | SGTB |
| 63 | CALB2 | 183 | KLHL14 | 303 | SH3D19 |
| 64 | CARS | 184 | KLHL32 | 304 | SH3GL2 |
| 65 | CASP4 | 185 | KRTAP19-1 | 305 | SIRT5 |
| 66 | CBLN4 | 186 | LASP1 | 306 | SLC45A4 |
| 67 | CCDC6 | 187 | LIMD1 | 307 | SMPDL3A |
| 68 | CCDC92 | 188 | LITAF | 308 | SMR3A |
| 69 | CCT7 | 189 | LOC113179 | 309 | SNX12 |
| 70 | CCT8 | 190 | LOC123688 | 310 | SPDYC |
| 71 | CENPP | 191 | LOC147804 | 311 | SPHAR |
| 72 | CHORDC1 | 192 | LOC197350 | 312 | SPINT1 |
| 73 | CLIC3 | 193 | LOC387693 | 313 | SPR |
| 74 | CNNM4 | 194 | LOC440337 | 314 | SPRR2G |
| 75 | CORO6 | 195 | LRRK1 | 315 | SPRR3 |
| 76 | COX6B2 | 196 | LSM12 | 316 | SRI |
| 77 | CPNE1 | 197 | LYPLA2 | 317 | STAT4 |
| 78 | CPNE6 | 198 | MAP2K2 | 318 | STIP1 |
| 79 | CRABP1 | 199 | MAPK11 | 319 | STOML2 |
| 80 | CRBN | 200 | MAT2B | 320 | STUB1 |
| 81 | CRIP1 | 201 | MCM2 | 321 | SYNJ2BP |
| 82 | CRLF3 | 202 | MCM7 | 322 | TAGLN |
| 83 | CRYGN | 203 | MDFI | 323 | TERF2IP |

| | | | | | |
|-----|----------------|-----|----------|-----|---------|
| 84 | CTBP1 | 204 | MDH1 | 324 | TES |
| 85 | CTF1 | 205 | MESDC2 | 325 | THOC3 |
| 86 | CTH | 206 | METAP1 | 326 | THOP1 |
| 87 | CTSL1 | 207 | METTL1 | 327 | TKT |
| 88 | CYGB | 208 | MGC26718 | 328 | TMEM128 |
| 89 | CYP2D6 | 209 | MIF | 329 | TMEM130 |
| 90 | CYP3A43 | 210 | MLZE | 330 | TMEM55B |
| 91 | DARS | 211 | MRPL37 | 331 | TNNT2 |
| 92 | DCTD | 212 | MSH2 | 332 | TOMM34 |
| 93 | DCUN1D3 | 213 | MT1G | 333 | TOX4 |
| 94 | DDI2 | 214 | MT1H | 334 | TPI1 |
| 95 | DDT | 215 | MT1JP | 335 | TPMT |
| 96 | DDX1 | 216 | MT1X | 336 | TPO |
| 97 | DFFA | 217 | MT2A | 337 | TRIM10 |
| 98 | DFFB | 218 | MT3 | 338 | TRIM27 |
| 99 | DHRS1 | 219 | MTAP | 339 | TRUB1 |
| 100 | DKFZp686K16132 | 220 | MTF2 | 340 | TSPAN31 |
| 101 | DNAJA4 | 221 | MYBPH | 341 | TSTA3 |
| 102 | DNAJB11 | 222 | N6AMT2 | 342 | TTC1 |
| 103 | DNAJC7 | 223 | NANP | 343 | TYW3 |
| 104 | DNMT3L | 224 | NANS | 344 | UBE1DC1 |
| 105 | DOK3 | 225 | NAPB | 345 | UBE2D2 |
| 106 | DPH2 | 226 | NAPG | 346 | UBE2Z |
| 107 | ECGF1 | 227 | NBPF4 | 347 | VPS37B |
| 108 | EEF1B2 | 228 | NDOR1 | 348 | WDR34 |
| 109 | EEF1D | 229 | NPHP4 | 349 | WIPF1 |
| 110 | EEF1E1 | 230 | NPLOC4 | 350 | YKT6 |
| 111 | EFHA2 | 231 | NSDHL | 351 | YPEL5 |
| 112 | EIF4B | 232 | NUBP1 | 352 | ZHX3 |
| 113 | ELA1 | 233 | NUDT10 | 353 | ZNF213 |
| 114 | EXOC5 | 234 | NUDT12 | 354 | ZNF259 |
| 115 | EXOSC5 | 235 | NUDT4 | 355 | ZNF428 |
| 116 | FAM104B | 236 | OBFC2A | 356 | ZNF576 |
| 117 | FAM46C | 237 | ODC1 | 357 | ZNF641 |
| 118 | FAM54B | 238 | OLIG2 | 358 | ZNF72 |
| 119 | FAM9C | 239 | OPTN | 359 | ZNF81 |
| 120 | FERD3L | 240 | OVOL2 | 360 | ZSCAN2 |

Table S5. Functional enrichment of the arsenic binding proteins in signaling pathways.

| Pathway | Count | Percentage | P value | Enriched gene |
|--|-------|------------|----------|---|
| hsa04919:Thyroid hormone signaling pathway | 30 | 1.06 | 2.28E-05 | HRAS, THRB, MED24, ITGB3, WNT4, ITGAV, GATA4, TBC1D4, PIK3CA, PLCD1, KAT2B, RXRB, RXRA, CREBBP, MED12, RXRG, ESR1, MED14, MED13, NOTCH3, NOTCH2, NCOA1, CCND1, NOTCH1, HIF1A, EP300, MED16, NOTCH4, TSC2, NCOR1 |
| hsa04310:Wnt signaling pathway | 34 | 1.20 | 2.54E-05 | PPARD, APC2, WNT3A, CXXC4, WNT1, WNT4, WNT6, WNT8A, WNT8B, WNT10A, CTBP1, CTBP2, ROCK2, CREBBP, PORCN, WNT2B, SFRP5, DKK2, WNT7B, CCND1, EP300, SOST, CCND3, PRICKLE1, SFRP2, CCND2, JUN, SFRP4, PRICKLE2, WNT9B, WIF1, WNT11, SIAH1, WNT9A |
| hsa04330:Notch signaling pathway | 17 | 0.60 | 5.03E-05 | DTX4, CTBP1, KAT2B, CTBP2, DTX1, CREBBP, JAG2, JAG1, RFNG, NOTCH3, NOTCH2, NOTCH1, EP300, DLL4, NOTCH4, ADAM17, RBPJL |
| hsa04390:Hippo signaling pathway | 34 | 1.20 | 0.0002 | PRKCZ, APC2, WNT3A, ITGB2, WTIP, LATS2, TGFB2, WNT1, WNT4, CRB1, CTGF, CRB2, DLG3, LIMD1, WNT6, DLG2, WNT8A, WNT8B, WNT10A, PRKCI, SNAI2, TP73, CTNNA3, STK3, WNT2B, AJUBA, CCND1, WNT7B, CCND3, CCND2, RASSF1, WNT9B, WNT11, WNT9A |
| hsa04151:PI3K-Akt signaling pathway | 63 | 2.23 | 0.0002 | HRAS, LPAR3, PIK3CA, PDGFC, COL11A2, INSR, |

| | | | | |
|---|----|------|--------|---|
| | | | | EGFR, SGK1, RXRA, BCL2L11, CCND1, CCND3, LAMC3, CCND2, RELN, LAMC2, LAMC1, NGFR, PHLPP1, PPP2R3A, TNC, ITGB4, ITGB5, ITGA10, ITGB3, ITGB1, CDC37, IGF1R, LAMB4, LAMB3, LAMB2, ITGB8, COL6A5, TNFR, ITGAV, ITGB7, COMP, TEK, ITGB6, COL6A1, TNN, LAMB1, THBS1, THBS2, FN1, THBS4, COL4A4, TNXB, COL4A1, FLT4, NR4A1, ITGA3, HGF, BRCA1, LAMA2, VWF, LAMA1, LAMA4, LAMA3, ITGA6, ITGA5, LAMA5, TSC2 |
| hsa04550:Signaling pathways regulating pluripotency of stem cells | 27 | 0.96 | 0.0085 | HRAS, APC2, WNT3A, REST, MEIS1, WNT1, IGF1R, WNT4, PCGF6, PIK3CA, WNT6, WNT8A, WNT8B, SETDB1, WNT10A, ESRRB, ISL1, WNT2B, ACVR2A, WNT7B, INHBE, INHBC, WNT9B, WNT11, WNT9A, ZFHX3, ACVR1 |
| hsa04350:TGF-beta signaling pathway | 17 | 0.60 | 0.028 | LTBP1, ROCK1, MINOS1-NBL1, CREBBP, FST, TGFB2, NBL1, ACVR2A, EP300, SP1, INHBE, ZFYVE9, ZFYVE16, INHBC, THBS1, CHR1, ACVR1 |
| hsa04115:p53 signaling pathway | 14 | 0.50 | 0.039 | BID, RFWD2, CCND1, CCND3, CCND2, ZMAT3, TSC2, SIAH1, RCHY1, THBS1, CCNG1, ADGRB1, CCNG2, TP73 |

REFERENCES

1. Ordóñez E, Thiyagarajan S, Cook JD, Stemmler TL, Gil JA, Mateos LM, Rosen BP. Evolution of metal(loid) binding sites in transcriptional regulators. *J. Biol. Chem.* 2008, 283(37):25706-25714.
2. Martin P, DeMel S, Shi J, Gladysheva T, Gatti DL, Rosen BP, Edwards BF. Insights into the structure, solvation, and mechanism of ArsC arsenate reductase, a novel arsenic detoxification enzyme. *Structure* 2001, 9(11): 1071-1081.
3. Yang J, Rawat S, Stemmler TL, Rosen BP. Arsenic binding and transfer by the ArsD As(III) metallochaperone. *Biochemistry* 2010, 49(17): 3658-3666.
4. Ajees AA, Marapakala K, Packianathan C, Sankaran B, Rosen BP. Structure of an As(III) S-adenosylmethionine methyltransferase: insights into the mechanism of arsenic biotransformation. *Biochemistry* 2012, 51(27): 5476-5485.
5. Zhang XW, Yan XJ, Zhou ZR, Yang FF, Wu ZY, et al. (2010) Arsenic trioxide controls the fate of the PML-RARalpha oncoprotein by directly binding PML. *Science* 328: 240-243.
6. Mao JH, Sun XY, Liu JX, Zhang QY, Liu P, et al. (2010) As4S4 targets RING-type E3 ligase c-CBL to induce degradation of BCR-ABL in chronic myelogenous leukemia. *Proc Natl Acad Sci U S A* 107: 21683-21688.
7. Xu RX, Hassell AM, Vanderwall D, Lambert MH, Holmes WD, Luther MA, Rocque WJ, Milburn MV, Zhao Y, Ke H, Nolte RT. Atomic structure of PDE4: insights into phosphodiesterase mechanism and specificity. *Science* 2000, 288(5472): 1822-1825.
8. Huber BR, Desclozeaux M, West BL, Cunha-Lima ST, Nguyen HT, Baxter JD, Ingraham HA, Fletterick RJ. Thyroid hormone receptor-beta mutations conferring hormone resistance and reduced corepressor release exhibit decreased stability in the N-terminal ligand-binding domain. *Mol. Endocrinol.* 2003, 17(1): 107-116.
9. Nørager S, Legrand P, Pieulle L, Hatchikian C, Roth M. Crystal structure of the oxidised and reduced acidic cytochrome c3 from *Desulfovibrio africanus*. *J. Mol. Biol.* 1999, 290(4): 881-902.