Online Supporting Information for

"Facile Discovery of Cell-surface Protein Targets of Cancer Cell Aptamers"

by

Tao Bing, Dihua Shangguan and Yinsheng Wang

Table S1. Specificity of Sgc-3b and Sgc-4e aptamers to different cell lines.

Cell lines	Origin	Sgc-3b	Sgc-4e
RAEC	rat aortic endothelial cells	-	-
MRC-5	human fetal lung fibroblast	-	++
A549	human alveolar basal epithelial cells	-	-
Hela	human cervical cancer	-	-
Huh-7	human hepatoma	-	-
T24	human bladder carcinoma	-	++
SK-Hep-1	human hepatocellular carcinoma	-	+++
MCF-7	human breast adenocarcinoma	-	-
MCF-7R	human breast adenocarcinoma	-	+++
SKOV-3	human ovarian carcinoma	-	+
K562	human myelogenous leukemia	-	-
MDA-MB-231	human breast cancer	-	-
A431	human epithelial carcinoma	-	-
Jurkat E6-1	human T-cell leukemia	+++	++++
HCT-8	human ileocecal adenocarcinoma	-	-
HEK-293	human embryonic kidney	-	++
PC-3	human prostate cancer	-	-
Hut-78	human T cell lymphoma	-	-
LoVo	human colon cancer	-	-

The binding capacity of the aptamer to the cells: -, <10%; +, 10-35%; ++, 35-60%; +++, 60-85%; ++++, >85%.

Table S2. The identification of endogenous biotinylated proteins with or without formaldehyde cross-linking using. The data represent the means and standard deviations of results obtained from three independent SILAC experiments.

Refseq Protein Accession	Gene Symbol	Protein Name	Unique Peptides	Sequence coverage [%]	PEP[a]	Protein Abundance Ratio (streptavidin beads/ biotin-saturated streptavidin beads)	
						With Cross-linking	No Cross-linking
NP_000911	PC	Pyruvate carboxylase	11	13.1	0	17.5±1.2	13.7±3.4
NP_942131	ACACA	Isoform 4 of Acetyl-CoA carboxylase 1	22	11.7	0	>20	17±2.4
NP_000273	PCCA	propionyl-coenzyme A carboxylase, α polypeptide isoform a precursor	4	6.9	2.48E-85	>20	18.7±2.6
NP_000523	PCCB	Propionyl coenzyme A carboxylase, β polypeptide, isoform CRA_c	9	22.4	2.9E-263	>20	>20
NP_064551	MCCC1	Methylcrotonoyl-CoA carboxylase subunit α	7	14.1	9.1E-178	>20	14±4.1

[a]PEP, the posterior error probability. This value essentially operates as a p-value, where smaller is better.

Fig. S1. A) Binding of Sgc-3b to Jurkat E6-1cells in the absence or presence of Mg^{2+} (5 mM), Ca^{2+} (5 mM) or EDTA (5 mM). B) Binding of Sgc-4e to Jurkat E6-1 cells in the absence or presence of Mg^{2+} (5 mM), Ca^{2+} (5 mM) or EDTA (5 mM). C) Fluorescence of Jurkat E6-1 cells cross-linked with Sgc-4e in the presence of Mg^{2+} or EDTA in the buffer for different periods of time.

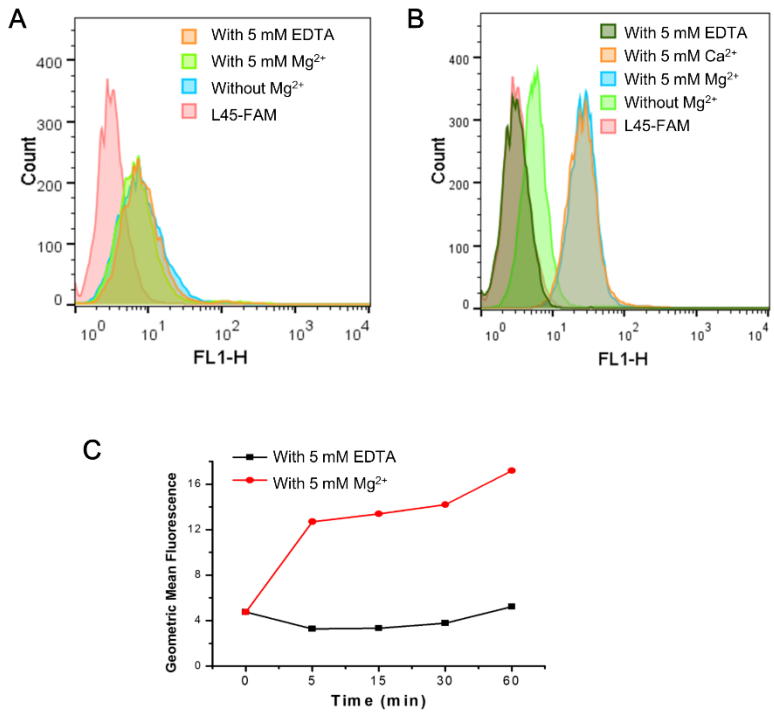

Fig. S2. A) Binding of Sgc-3b to Jurkat E6-1 cells after cross-linking (Control) and elution of aptamer-target complex using 1% Triton X-100. B) Binding of Sgc-4e to Jurkat E6-1 cells following cross-linking and elution of aptamer-target complex using 1% Triton X-100 or 2% Triton X-100 and 0.4% SDS.

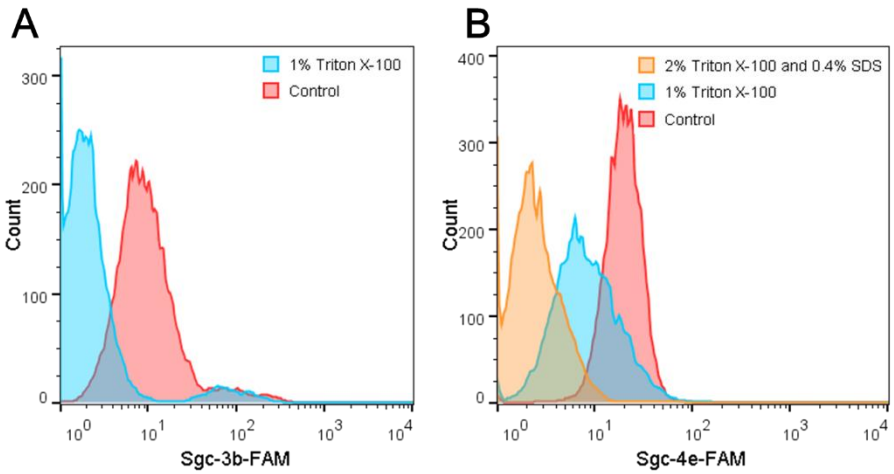

Fig. S3. Fluorescence images of streptavidin-coated beads (GE company) pulling down the aptamer-target complex from Jurkat E6-1 cells: A) Blank; B) aptamer Sgc-3b; C) aptamer Sgc-4e.

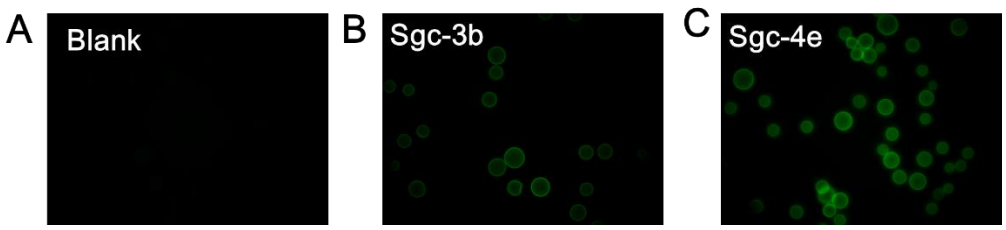

Fig. S4. A) Sliver-stained SDS-PAGE (10%) of proteins that were pulled down using streptavidin-coated beads (SA) or biotin-saturated streptavidin-coated beads (SA-biotin)**.** B) Sliver-stained SDS-PAGE (10%) of proteins that were pulled down using streptavidin beads coated with Sgc-3b or Sgc-4e.

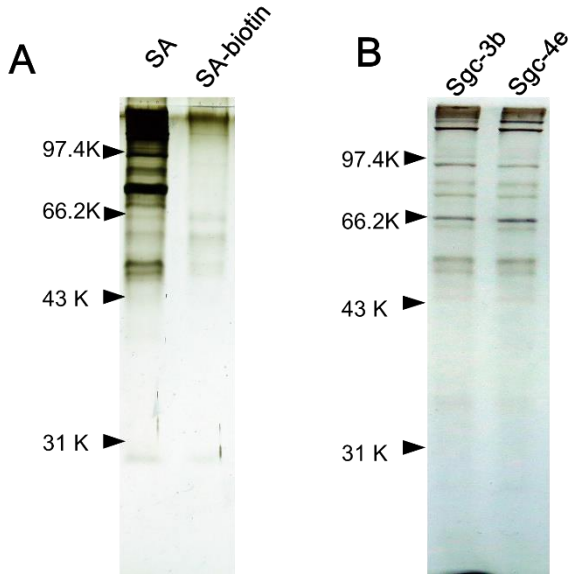

Fig. S5. Flow cytometry assay results showed the expression of CD62L, CD49d or CD29 on K562(A), HCT-8(B), LoVo(C), PC-3(D), MD-MB-231(E), MCF-7(F), MCF-7R(G), HEK-293(H), MRC-5(I) and SK-Hep-1(J) cell lines. IgG-PE, which do not bind to cells, was used as negative control.

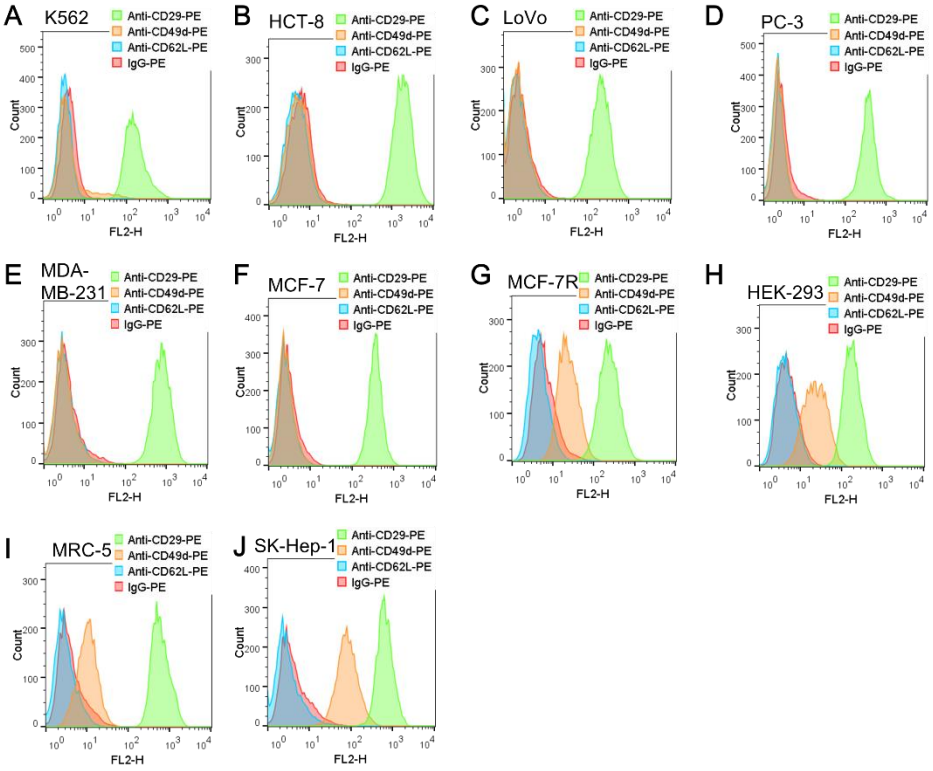

Fig. S6. Representative ESI-MS and MS/MS of a tryptic peptide from CD49d (the target protein for aptamer Sgc-4e). R* designates the heavy arginine. Shown in (A) and (B) are the ESI-MS for the light (*m/z* 543.3 for the monoisotopic peak of the $[M+2H]^{2+}$ ion) and heavy (*m/z* 546.3 for the monoisotopic peak of the $[M+2H]^{2+}$ ion) arginine-containing peptide observed in forward and reverse SILAC experiments. Displayed in (C) and (D) are the MS/MS for the $[M+2H]^{2+}$ ions of the heavy- and light-arginine-bearing peptide.

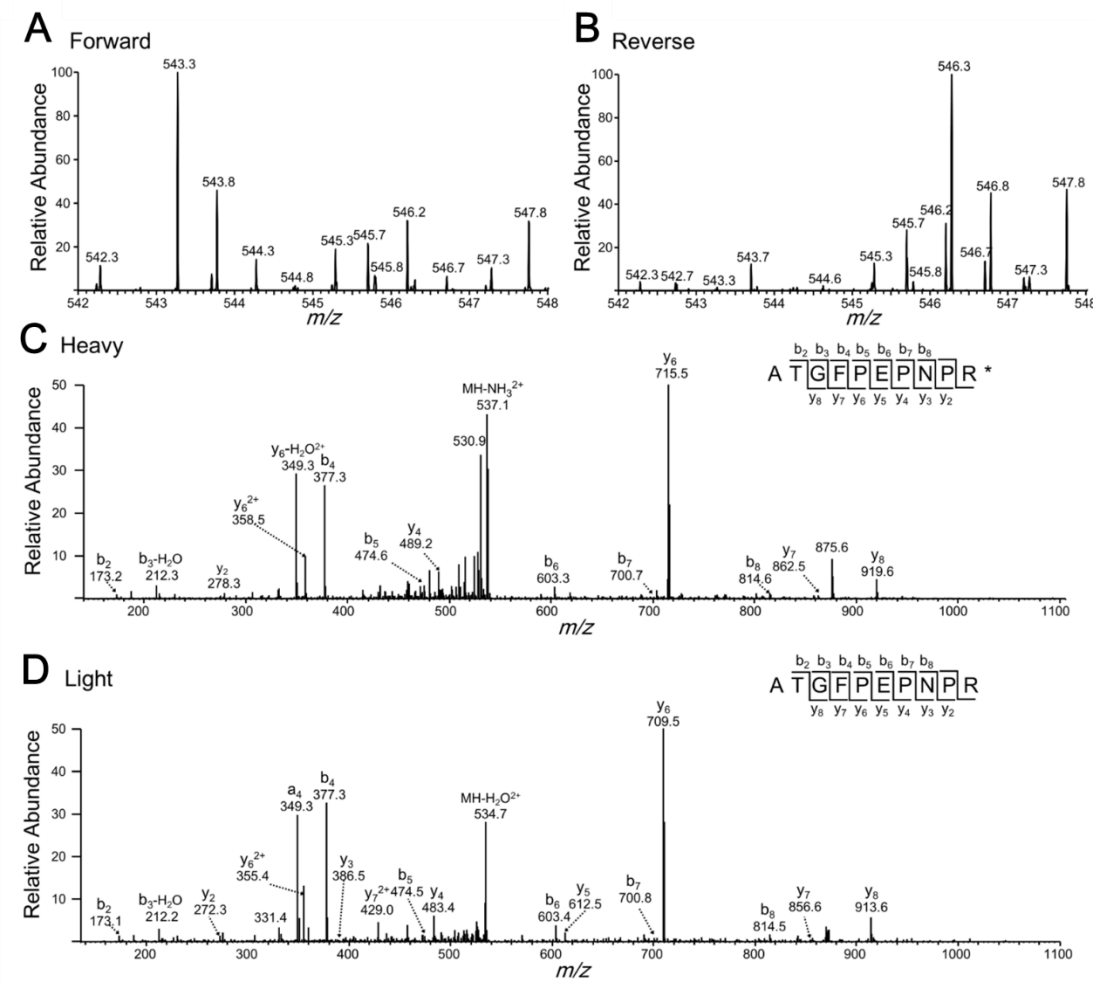

Fig. S7. Representative ESI-MS and MS/MS of a tryptic peptide from CD29 (the target protein for aptamer Sgc-4e). K* designates the heavy lysine. Shown in (A) and (B) are the ESI-MS for the light (*m/z* 508.3 for the monoisotopic peak of the $[M+2H]^{2+}$ ion) and heavy (*m/z* 512.3 for the monoisotopic peak of the $[M+2H]^{2+}$ ion) lysine-containing peptide observed in forward and reverse SILAC experiments. Displayed in (C) and (D) are the MS/MS for the $[M+2H]^{2+}$ ions of the heavy- and light-lysine-bearing peptide, respectively.

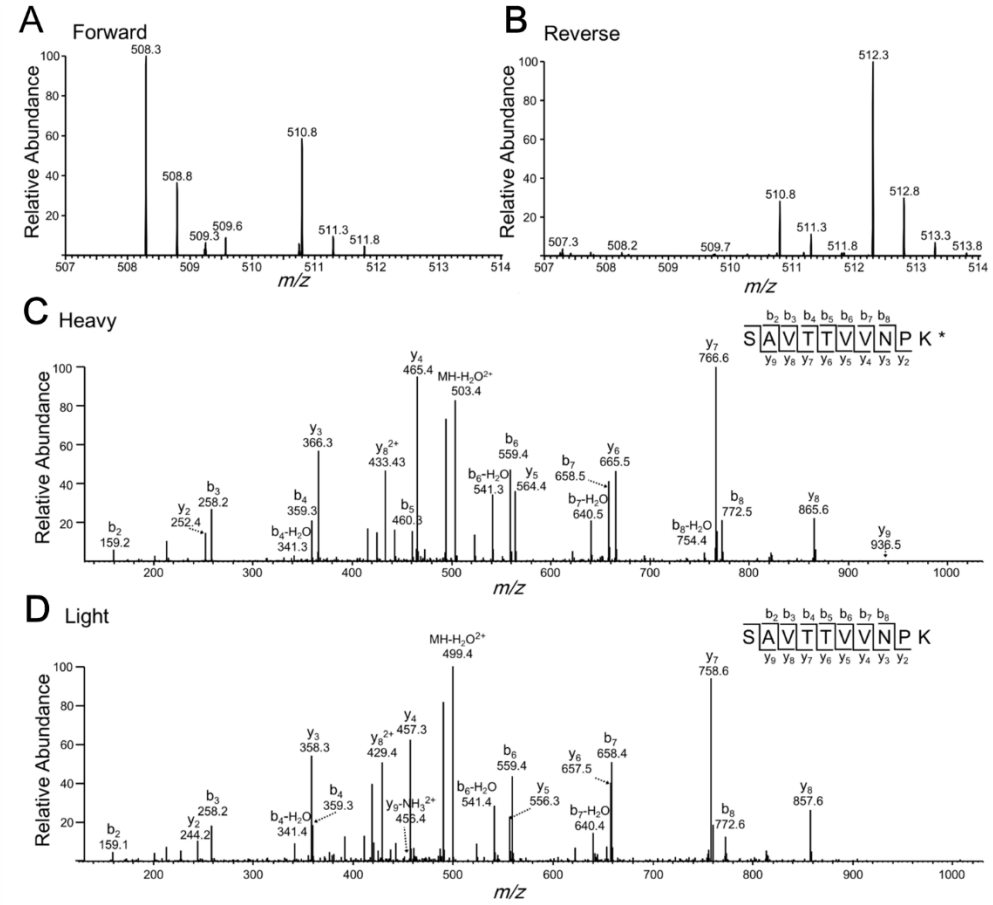

Fig. S8. Representative ESI-MS and MS/MS of a tryptic peptide from talin-1 (the target protein for aptamer Sgc-4e). R* designates the heavy arginine. Shown in (A) and (B) are the ESI-MS for the light (*m/z* 508.6 for the monoisotopic peak of the $[M+2H]^{3+}$ ion) and heavy (*m/z* 510.6 for the monoisotopic peak of the $[M+2H]^{3+}$ ion) arginine-containing peptide observed in forward and reverse SILAC experiments. Displayed in (C) and (D) are the MS/MS for the $[M+2H]^{3+}$ ions of the heavy- and light-arginine-bearing peptide.