Electronic Supplementary Material (ESI) for ChemComm.

Supplementary Information

A Peptide-based Synthetic Transcription Factor Selectively Down-regulates the Proto-oncogene CFOS in Tumour Cells and Inhibits Proliferation

Madhumita Chakraborty and Siddhartha Roy[1]*

[1]Department of Biophysics, Bose Institute, P-1/12 CIT Scheme VII M, Kolkata 700 054, India

MATERIALS AND METHODS

All reagents were purchased from Sigma Chemical Company (Bangalore, India) unless otherwise indicated. All protected amino acids were from Novabiochem (Switzerland). All oligonucleotides were purchased from TriLink (San Diego, CA). The different oligonucleotide sequences used are presented in Supporting Table S1.

Design and construction of peptides. Design of peptides was guided by three-dimensional structures of the protein-DNA complexes. The major groove binding helix of ELK1 and minor groove binding elements of the SRF were selected as building blocks from the model constructed, based on pdb1K6O and 1DUX by docking free ELK1 structure with SRF-SRE structure (without SAP1) in HADDOCK web server (*1*).Conformational constraint on the helix was imposed by substitution of α-amino isobutyric acid in place of residues that are not part of the protein-DNA interface (*2, 3*). For conjoined peptide mimics, the linker between the ELK1 helix and SRF arm (N-terminal) were chosen by calculating the actual distance between the two (at suitable points) in the original three-dimensional structures of the proteins. The mutant peptide was designed by replacing a few residues, that are in contact with DNA bases, by alanine. The sequences of the different peptides synthesized are given in Supporting Tables S2 and S3.

Peptide Synthesis and purification. All peptides were synthesized in the solid phase via Fmoc chemistry on Rink amide MBHA resin in a peptide synthesizer manufactured by Protein Technologies Inc (AZ, USA). After synthesis, the peptides were cleaved from the resin by using a cocktail containing 87.5% TFA, 5% H_2O, 2.5% Ethane dithiol and 5% methyl phenyl sulfide for 3 hrs at 25°C followed by precipitation with cold diethyl ether. The precipitate was then collected after centrifugation, dried and dissolved in water. The peptides were purified by Reverse Phase HPLC. The molecular weights of the peptides were verified

by MALDI Mass Spectrometry in ABI-4800 MALDI-TOF-TOF mass spectrometer (The HPLC profiles and the mass spectra of all the synthesized peptides are shown in Figure S16. two important ones are shown in Supporting Figures S17 and S18.). Sequences of the peptides synthesized and their respective molecular weights verified from mass spectrometry are listed in Supporting Tables S2 and S3.

Labeling of peptides. Peptides were labeled with fluorescent probes in the solid phase. N-terminal deprotected peptides, while linked with resins, were incubated with ten times molar excess of carboxy-fluorescein (Invitrogen, molecular probes, New Delhi), HOBT and DIPC in DMF. The labeling reaction was carried out for approximately 18 h at 25°C. After completion of the reaction, resins were washed several times with 20% piperidine followed by DMF. The labeled peptides were then cleaved from the resins using the cocktail as described above (*4*)**.**

Peptide-DNA binding assay by Fluorescence Anisotropy. All fluorescence studies were done in a Perkin-Elmer LS55 spectrofluorometer at 25°C.The experiments were carried out in a 0.5 cm path length quartz cuvette. The excitation wavelength was at 490 nm and the emission was set at 530 nm. The excitation and emission band passes were 5 nm each. The titrations were performed in 20 mM Tris-HCl buffer, pH 7.5 containing 150 mM NaCl. The measured anisotropy values were fitted to the following single site binding equation using Kyplot:

$$A_{obs} = A_0 + [(A_\infty - A_0)(K_d + X + [DNA]) - \sqrt{(K_d + X + [DNA])^2 - 4X[DNA]}] / 2[DNA]$$

Where A_{obs} is the observed anisotropy, A_0 is the initial anisotropy, A_∞ is the final limiting anisotropy, [DNA] is the DNA concentration, X is the total peptide concentration, and K_d is the dissociation constant. All concentrations were expressed in molar units.

CD spectroscopy

CD spectra of E3x2GS peptide (30μM) with and without the target DNA in 10 mM phosphate buffer, pH 7.5 containing 50 mM NaCl were recorded in the JASCO-810 spectropolarimeter using a rectangular quartz cell of path length of 0.1cm in the wavelength range of 190 nm to 250 nm. The peptide was incubated with equimolar DNA for 15 mins at room temperature for recording CD spectra of peptide-DNA complex. Similar conditions were used for the DNA CD experiments.

Expression and purification of ELK1 DNA binding domain

The recombinant ELK1-DBD (1-94 amino acid from PDB structure 1DUX) in pGEX-4T1 plasmid was purchased from GenScript the service provider of synthetic gene after codon optimization. The plasmid was transformed into E. coli BL21-DE3. Bacterial cultures were incubated at 37°C in LB medium containing ampicilin until an OD_{600} of 0.6 was reached. Recombinant protein expression was induced by the addition of IPTG to a final concentration of 1 mM. Following 3 h incubation at 37°C, bacteria were harvested by centrifugation and frozen at −20°C. Bacterial pellets were resuspended with lysis buffer (20 mM Tris-HCl buffer, pH 7.5 containing 150 mM NaCl, 5% glycerol and 10 mM PMSF) cells were lysed by pulsed sonication in ice and subsequent freeze thawing. The lysates were centrifuged at 14,000 g for 30 min at 4°C. Pre-equilibrated Glutathione sepharose 4B beads were added to the supernatant and kept on a shaker for 3 hrs at 4°C. It was then spun for 5 mins at 700Xg and the supernatant was removed and GST-fusion protein remained attached to the bead. After extensive washing, beads were incubated over night at 4°C with restriction grade thrombin from MP Biomedicals for in bead digestion of the protein. The free ELK1-DBD was collected and treated with benzamidine-sepharose 4 fast flow for 1 hr at 4°C to remove thrombin. The protein was then collected, concentrated and estimated.

Competitive titration of E3x2GS peptide and ELK1 DBD to the consensus binding site

The titration of E3x2GS peptide with SRE DNA was first performed under similar conditions as described before. The ELK1 DBD protein was added to the mixture before the saturation was achieved with peptide-DNA titration starting with a molar excess of 50 fold to 250 fold and anisotropy values were recorded.

Cell Culture

The cell line used in this study is the human lung carcinoma cell line (A549). It was purchased from American Type Culture Collection (ATCC, USA). A549 was grown in DMEM medium (Dulbecco's modified Eagle's medium from Gibco). To ensure growth and viability of the cells, the medium was supplemented with 10% FBS (Gibco) and incubated in a humidified atmosphere with 5% CO_2 at 37^{o}C.

Cell morphology and growth curve

To observe cell morphology with the treated peptides and to observe the growth curve for the control cell and treated cell 10^4 cells were treated with wild type and mutant peptides in different wells and kept for 24 hrs, 48 hrs, and 72 hrs, respectively with a peptide dose of 15 μM each. A control well was also kept with untreated cells under similar condition. The morphology of the cells was observed under optical microscope and captured at different time points. The total cells were counted in hemocytometer for control and treated wells (WT and mutant peptide treated) by trypan blue exclusion method after 24 hrs, 48 hrs, and 72 hrs, respectively.

Cell viability assay

Cell viability assay of A549 was performed with MTT (3-(4, 5-dimethylthiazol-2yl-)-2, 5-diphenyl tetrazolium bromide).For performing the experiment, cells were grown to 80-90%

confluence and then harvested with 0.025% trypsin in PBS containing 0.52mM EDTA. Approximately, 10^3 cells/well were seeded into 96-well tissue culture plates and 24 h after seeding, the medium was changed and cells were treated with various concentrations of the peptide and incubated for 96 hrs. After incubation in a CO_2 incubator, the cell viability was determined by the MTT assay. 20 µl of MTT solution (5 mg/ml in PBS) was added to each well. The plates were wrapped with aluminum foil and incubated for 4 h at 37°C. The purple formazan crystals formed at the bottom of the wells were dissolved with 0.1 N HCl in isopropanol and the absorbance at 550 nm was read on a microplate reader. Cell viability was expressed as the ratio of the optical density of the treated cells to that of the control (% control). The 50% growth inhibition values were obtained from dose response growth inhibitory curves.

Flow cytometry

For cell cycle assay $5X10^4$ cells/well were plated in 6 well plate and after 24 hrs of seeding cells were treated with 15µM of E3x2GS-d6NLS peptide for incubation of 48 hr, and 72 hrs followed by trypsinization. Untreated cells were also kept for control experiment. Cells were prepared and stained with BD Cycletest™ Plus DNA Reagent Kit according to manufacturer's protocol. Analysis was done on a Becton-Dickinson flow cytometer.

Confocal microscopy

A549 cells ($2 X 10^4$) cells were grown on cover slips in 500µl culture medium. Twenty-four hours after seeding, cells were treated with both 1 µM of E3x2GS-d6NLS and mE3x2GS-d6NLS peptides separately for 5hrs. Cells were then washed thrice with PBS and were fixed with 4% (w/v) formaldehyde for 10 min. Again after washing with PBS cells were incubated with Hoechst stain for 10 min and washed. Cover slips are then mounted on slides with a

drop of anti-fading agent. Images were independently captured by Leica TCS SP8 Confocal Laser Scanning microscope (Germany).

RNA preparation

For differential gene expression study 10^5 cells were treated with wild type and mutant peptides in different wells and kept for 24 h. Peptide concentration was either 15 μM or 10 μM. A control well consisting of untreated cells was also included in the experimental set. The total RNA was isolated using Qiagen RNeasy Mini Kit. Extracted amount of RNA was determined from absorbance at 260nm.

Reverse transcription

cDNA synthesis was performed using MuLV Reverse Transcriptase Plus from Bio Bharati (Kolkata, India). About 1 μg of total RNA was used along with a random primer and dNTP mix in a total volume of 12 μl. The mixture was heated to 65^{o}C for 5 min and quickly chilled on ice. To the contents, buffer, DTT and RNase inhibitors were added, mixed well and kept for 5 minutes. Then 1 μl of RT (200 U/μl) was added. The mixture was incubated at 42^{o}C for 50 min and finally the reaction was inactivated by heating at 70^{o}C for 15 min.

PCR

PCR amplifications were performed in buffer of 10 mM Tris-HCl buffer, pH 8.8, containing 50 mM KCl, 4 mM $MgCl_2$ and 10 pico mole of each forward and reverse oligonucleotide along with 1.25 units of Taq DNA polymerase (Fermentas) and 1 μl cDNA mixture in a final volume of 25 μl. After an initial denaturation step, PCR was carried out under the following conditions: 30 s denaturation at 95°C, 45 s annealing at 57°C (for CFOS, BAX, p21and GAPDH), 56°C (for noxa), 55°C (for EGR1, APAF1 and GNGT1), 59°C (ATF3), followed by 1 min extension at 72°C. RT-PCRs were performed by the use of specific primers of the

respective genes. The primers used again all the genes are listed in Table S6. Amplification products were run on a 1.5% agarose gel in equal volume. Gel quantification (densitometric analysis) was done using Image J. Expression levels were determined on the basis of three independent sets of RT-PCR data.

Real Time PCR

Real-time RT-PCR amplifications were performed in 7500 Fast Real-Time PCR System of Applied Biosystems with a standard reaction mixture containing the sample DNA, one set of primers and the SYBR Green I PCR master mix (Applied Biosystems) at a total volume of 10 µl. Samples were run in triplicates. PCR was performed under the following conditions: 20 sec at 50°C then 10 min at 95°C followed by 50 cycles of 15 sec at 95°C for denaturationand1min at 60°C for annealing and extension. Relative changes in gene expression were analyzed using $2^{-\Delta\Delta}C_T$ method (*5*).

si-RNA and transfection

si-RNA duplex against ELK1 was purchased from Biogene India (Dharmacon) corresponding to the sequence GGCAATGGCCACATCATCT obtained from (*6*). Scrambled si-RNA which is a non-targeting 20-25 nt si-RNA (SANTA CRUZ BIOTECHNOLOGY, INC) was used as negative control. A549 cells ($10X10^4$ cells/well) were grown and transfected with si-RNA (20nM) and Lipofectamine^{TM}RNAiMax (invitrogen) according to the protocol provided by the manufacturer. Si-RNA transfection was performed for 48 hrs where after 24 hrs optimem media was replaced with complete media. Further experiments of RNA isolation and semi quantitative-PCR were carried out as described previously.

Whole genome expression profiling

The effect of synthetic transcription factor on gene expression was observed by microarray experiment. WT and mutant peptides were added to ~10^7 cells in separate flasks at a concentration of 15 μM for 24 h. A control flask was also kept with untreated cells. Cells were then trypsinized, centrifuged, washed and counted. Equal number of cells ($2X10^6$) from each flask was collected separately and dissolved in a RNA-later solution. The microarray experiments were conducted by Genotypic, Inc. (Bangalore, India) on payment for service basis. The array chosen was Whole Genome-Human Gene Expression 8X60k and hybridization was done with single color. Three biological replicates for each of control and WT peptide treated samples were taken and the fold changes were calculated as log base 2. Normalization of the raw data was done using GeneSpring GX Software. The results were analyzed on PASTAA server (http://trap.molgen.mpg.de/cgi-bin/pastaa.cgi) for transcription factor associations (*7*). The same experiment was later repeated with an earlier time point of 8 hr with control and WT peptide treated cells having three independent replicate experiments of each sample. The functions and pathway analysis were done for the differentially regulated genes using DAVID database.

Table S1. Oligonucleotide sequences

Oligonucleotides	Sequences
SRE (F) **SRE (R)**	**5' CACAGGATGTCCATATTAGGACA 3'** **3' GTGTCCTACAGGTATAATCCTGT 5'**
SBM1(F) * **SBM1 (R)**	**5' CACAGTATGTCCATATTAGGACA 3'** **3' GTGTCATACAGGTATAATCCTGT 5'**
SBM2 (F) **SBM2 (R)**	**5'CACAGGATGTCTATATTAGGACA 3'** **3' GTGTCCTACAGATATAATCCTGT 5'**
DBM (F)** **DBM (R)**	**5' CACAGTATGTCTATATTAGGACA 3'** **3' GTGTCATACAGATATAATCCTGT 5'**

All oligos are 5'-C6 aminolinked

* SBM :-Single base pair mutant, 1:- mutation in Elk1 binding part, 2:- mutation in SRF binding part

** DBM:- Double base pair mutant

Table S2: List of peptides synthesized

Peptide Name	Peptide Description and Sequence (N to C terminal)	Expected Mass	Experimental Mass
ELK (helix)	15-mer peptide containing helix from Elk NYBKBSRABRYYYDK	1882	1882
SRF (arm)	11-mer peptide containing arm from SRF GKKTRGRVKIK	1270	1270
SRF (helix)	16-mer peptide from SRF containing major groove binding helix NBLRRBTTBSKRKBGI	1770	1770
SRF (arm +helix)	32-mer peptide from SRF containing minor and major groove binding loop and helix GKKTRGRVKIKMEFIDNBLRRBTTBSKRKBGI	3658	3655
EGGGS	35-mer peptide containing both Elk and SRF helix linked by GGG NYDKBSRABRYYYDKNGGGNBLRRBTTBSKRKBGI	3949	3942
EGPGPGS	37-mer peptide containing both Elk and SRF helix linked by GPGPG NYBKBSRABRYYYDKNGPGPGNBLRRYTTBSKRKBGI	4192	4192
E3xS (G=0)	29-mer peptide containing helix from Elk and loop from SRF linked by 3 Ahx residues NYBKBSRABRYYYDKAhxAhxAhxGKKTRGRVKIK	3473	3471
E3x1GS (G=1)	30-mer peptide containing helix from Elk and loop from SRF linked by 1 Glycine and 3 Ahx residues NYBKBSRABRYYYDKGAhxAhxAhxGKKTRGRVKIK	3531	3530
E3x3GS (G=3)	32-mer peptide containing helix from Elk and loop from SRF linked by 3 Glycine and 3 Ahx residues NYBKBSRABRYYYDKGAhxAhxAhxGKKTRGRVKIK	3646	3645

Table S3. List of different peptides synthesized

Peptide Name	Peptide Description and Sequence (N to C terminal)	Expected Mass	Experimental Mass
E3x2GS Elk-loop (G=2)	31-mer peptide containing helix from Elk and loop from SRF linked by 2 Glycine and 3 Ahx residues NYBKBSRABRYYYDKGGAhxAhxAhxGKKTRGR VKIK	3588	3588
E3x2GS-d6NLS Elk-loop (G=2) with NLS and D-Arg.	44-mer peptide containing helix from Elk and loop from SRF linked by 2 Glycine and 3 Ahx residues with nuclear localization signal and 6 D-Arg at N-terminal d6PKKKRKVNYBKBSRABRYYYDKGGAhxAhxA hxGKKTRGRVKIK d6= six D-Arg	5390	5392
mE3x2GS Elk-loop (G=2) Mutant peptide	31-mer peptide containing helix from Elk and loop from SRF linked by 2 Glycine and 3 Ahx residues and few DNA binding residues were replaced by Ala. NYBKBS**A**AB**AA**YYDKGGAhxAhxAhxGKK**A**RG **A**VKIK	3210	3211
mE3x2GS-d6NLS Elk-loop (G=2) Mutant peptide with NLS and D-Arg.	44-mer peptide containing helix from Elk and loop from SRF linked by 2 Glycine and 3 Ahx residues and few DNA binding residues were replaced by Ala. Also contains nuclear localization signal and 6 D-Arg at N-terminal d6PKKKRKVNYBKBS**A**AB**AA**YYDKGGAhxAhxA hxGKK**A**RG**A**VKIK	5013	5017

Red colored amino acids are substituted ones. B stands for Aib, d6 stands for six D-Arginines and x for ω-amino acids, Ahx.

Table S4: Dissociation constants of different synthesized constructs and DNA complexes

Peptide Name	K_d**M**				
	SRE(WT)	SBM1 Mutation in ELK1-binding site	SBM2 Mutation in SRF-binding site	DBM	c-myc
ELK (helix)	$6.79X10^{-5}$				
SRF (arm)	$3.3\ X10^{-4}$	–	–	–	
SRF (helix)	$3.0\ X10^{-4}$	–	–	–	
SRF (arm +helix)	$(2.1\pm0.4)\ X10^{-7}$	–	–	–	
EGPGPGS	$(4.9\pm1.5)\ X10^{-7}$	$(2.1\pm1.0)\ X10^{-6}$	$(5.3\pm3.5)\ X10^{-7}$	$(6.0\pm0.7)\ X10^{-6}$	
EGGGS	$9.6\ X10^{-6}$	–	–	–	
E3xS (G=0)	$(3.2\pm0.8)\ X10^{-7}$	–	–	–	
E3x1GS (G=1)	$(3.0\pm1.1)\ X10^{-7}$	–	–	–	
E3x2GS (G=2)	**$(4.7\pm0.31)\ X10^{-8}$**	$(2.0\pm0.2)\ X10^{-7}$	$(1.8\pm0.1)\ X10^{-7}$	–	$(1.9\pm0.03)\ X10^{-4}$
mE3x2GS (G=2) mutant	$1.0\ X10^{-4}$	-	-	-	-
E3x3GS (G=3)	$(2.34\pm1.1)\ X10^{-6}$	–	–	–	

Dissociation constants of different constructs synthesized with their DNA. The dash signifies corresponding experiments have not been performed. The values represent $K_d \pm$ SEM based on at least three independent measurements.

Table S5: List of genes involved in cell cycle and mitosis which are down regulated by E3x2GS peptide

GENE NAME	*FUNCTION*
CENPO	Centromere protein O, required for bipolar spindle assembly, chromosome segregation and checkpoint signaling during mitosis
GINS1	Plays an essential role in the initiation of DNA replication, and progression of DNA replication forks
CDK1	Cyclin-Dependent Kinase 1, essential for G1/S and G2/M phase transitions of eukaryotic cell cycle
GINS2	Plays an essential role in the initiation of DNA replication, and progression of DNA replication forks
HSP90AA1	Heat Shock Protein 90kDa Alpha Family Class A Member 1, Molecular chaperone that promotes the maturation, structural maintenance and proper regulation of specific target proteins involved for instance in cell cycle control and signal transduction.
SGOL2	Shugoshin-Like 2, that regulate sister chromatid cohesion during meiosis
CASC5	Cancer Susceptibility Candidate 5, essential for spindle-assembly checkpoint signaling and for correct chromosome alignment during mitosis
AURKAPS1	Aurora Kinase A Pseudogene 1, a Pseudogene
CCNE2	Cyclin E2, essential for the control of the cell cycle at the late G1 and early S phase.
CCNB1	Cyclin B1, Essential for the control of the cell cycle at the G2/M (mitosis) transition.
PLK4	Polo Like Kinase 4, regulates centriole duplication during the cell cycle, related pathways are Cell Cycle, Mitotic and RB in

	Cancer
CDCA8	Cell Division Cycle Associated 8, required for chromatin-induced microtubule stabilization and spindle formation, a key regulator of mitosis
POLE2	Polymerase (DNA) Epsilon 2, Accessory Subunit, Participates in DNA repair and in chromosomal DNA replication.
CENPA	Centromere Protein A, Required for recruitment and assembly of kinetochore proteins, mitotic progression and chromosome segregation
INCENP	Inner Centromere Protein, Component of the chromosomal passenger complex (CPC), a complex that acts as a key regulator of mitosis
RRM2	Ribonucleotide Reductase Regulatory Subunit M2, Provides the precursors necessary for DNA synthesis. Catalyzes the biosynthesis of deoxyribonucleotides from the corresponding ribonucleotides
ERCC6L	Excision Repair Cross-Complementation Group 6 Like, Contributes to the mitotic checkpoint by recruiting MAD2 to kinetochores

The functionality of the genes are mentioned from Gene Cards, Human Gene Database

Table S6: List of primers

Gene Name	Primer
CFOS	Forward 5'-GTGGGAATGAAGTTGGCACT -3' Reverse 5'-CTACCACTCACCCGCAGACT -3'
NOXA	Forward 5'-AAGTTTCTGCCGGAAGTTCA -3' Reverse 5'-GCAAGAACGCTCAACCGAG -3'
BAX	Forward 5'-AGCTTCTTGGTGGACGCAT -3' Reverse 5'-CAGAGGCGGGGTTTCATC -3'
P21	Forward 5'-AGTCAGTTCCTTGTGGAGCC-3' Reverse 5'-CATGGGTTCTGACGGACAT-3'
GAPDH	Forward 5'-ACCACAGTCCATGCCATCAC-3' Reverse 5'-TCCACCACCCTGTTGCTGTA-3'
EGR1	Forward 5'-GACCGCAGAGTCTTTTCCTG -3' Reverse 5'-AGCGGCCAGTATAGGTGATG -3'
GNGT1	Forward 5'-TGAGGACCTGACAGAAAAGGA -3' Reverse 5'-CTCGCCAGATCGTTCTTCA-3'
APAF1	Forward 5'-CCTCTCATTTGCTGATGTCG -3' Reverse 5'-TCACTGCAGATTTTCACCAGA-3'
ATF3	Forward 5'-AGGTTTGCCATCCAGAACAA -3' Reverse 5'-AGGCACTCCGTCTTCTCCTT -3'

Figure S1

A

SAP-1
YDKLSRALRYYYVKN

Elk-1
YDKLSRALRYYYDKN

B

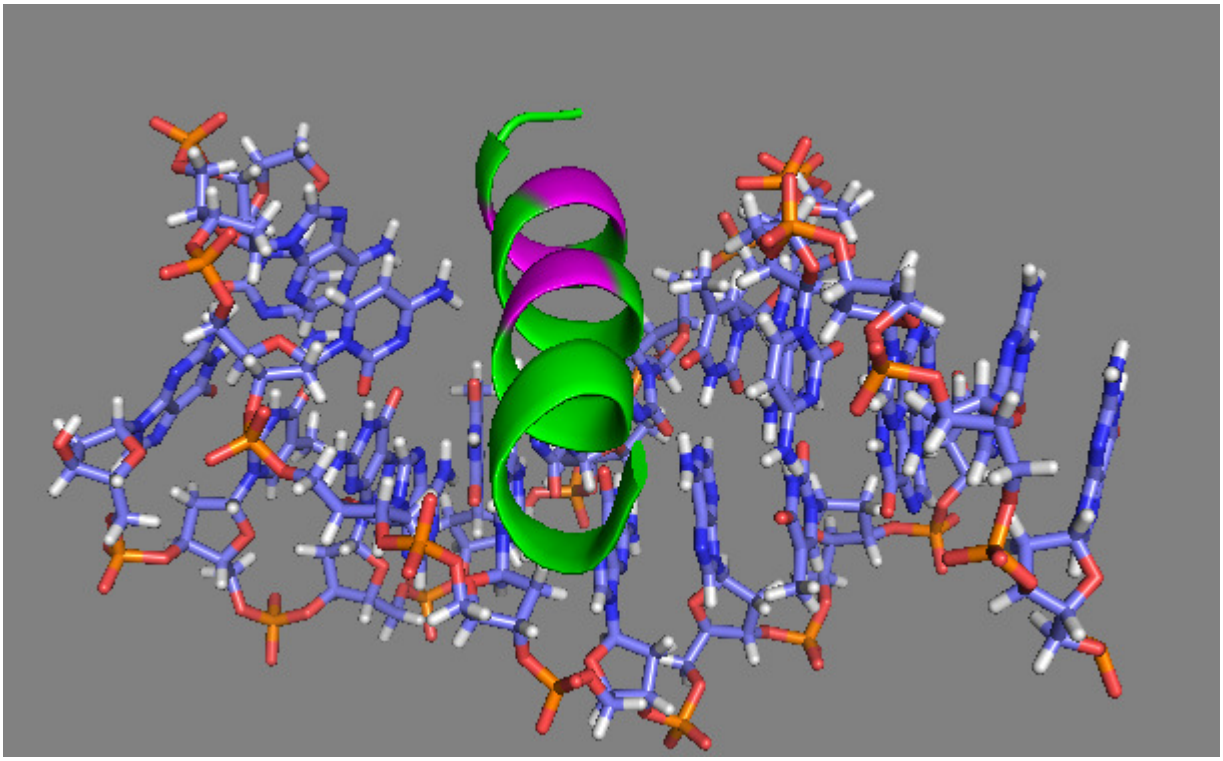


Figure S1. (A) Amino acid sequence of the recognition helix of ELK1 and SAP1. (B) A molecular model of the Elk1 recognition helix with the target DNA (derived from the model of Elk1-SRE complex). The magenta colored non-DNA-binding amino acids were substituted with aib.

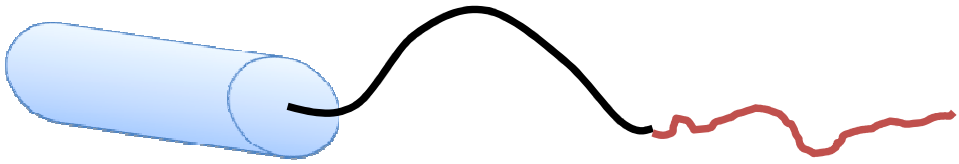
Figure S2

A

5’ CACAG G ATGTCCATAT TAGGACA
GTGT C CT A CAGGTATA ATCCT G T 5’

ETS motif CArg Box

B



E3xS: NYBKBSRABRYYYDK-Ahx-Ahx-Ahx-GKKTRGRVKIK

E3xGS: NYBKBSRABRYYYDK-G-Ahx-Ahx-Ahx-GKKTRGRVKIK

E3x2GS: NYBKBSRABRYYYDK-G-G-Ahx-Ahx-Ahx-GKKTRGRVKIK

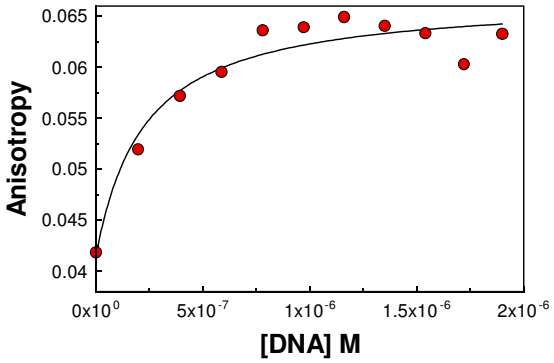
C

NH_2-CH_2-CH_2-CH_2-CH_2-CH_2-COOH

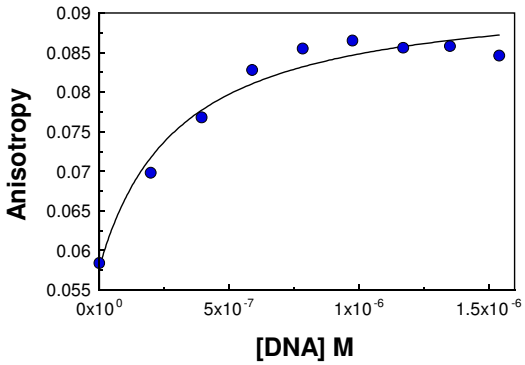
Figure S2. (A) oligonucleotide sequence of the target site. ELK1 binds to the ETS motif and SRF binds to the CArg-Box (B) Sequence of peptides used in this study having different linker lengths with 0, 1, 2, 3 glycine residues. Only E3x2GS showed the desired affinity and specificity. (C) Chemical structure of 6-amino hexanoic acid (Ahx).

Figure S3

A



B



C

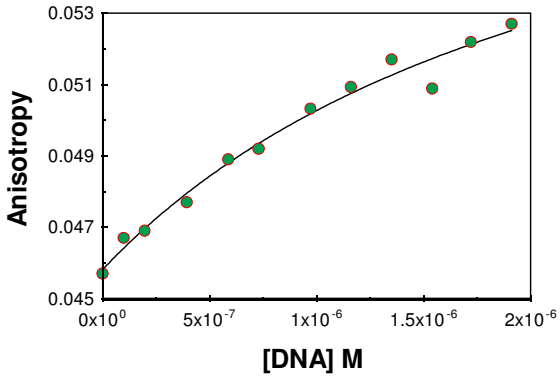


Figure S4. Binding isotherms of (A) E3xS, (B) E3xGS and (C) E3x3GS with the WT DNA.

Figure S4

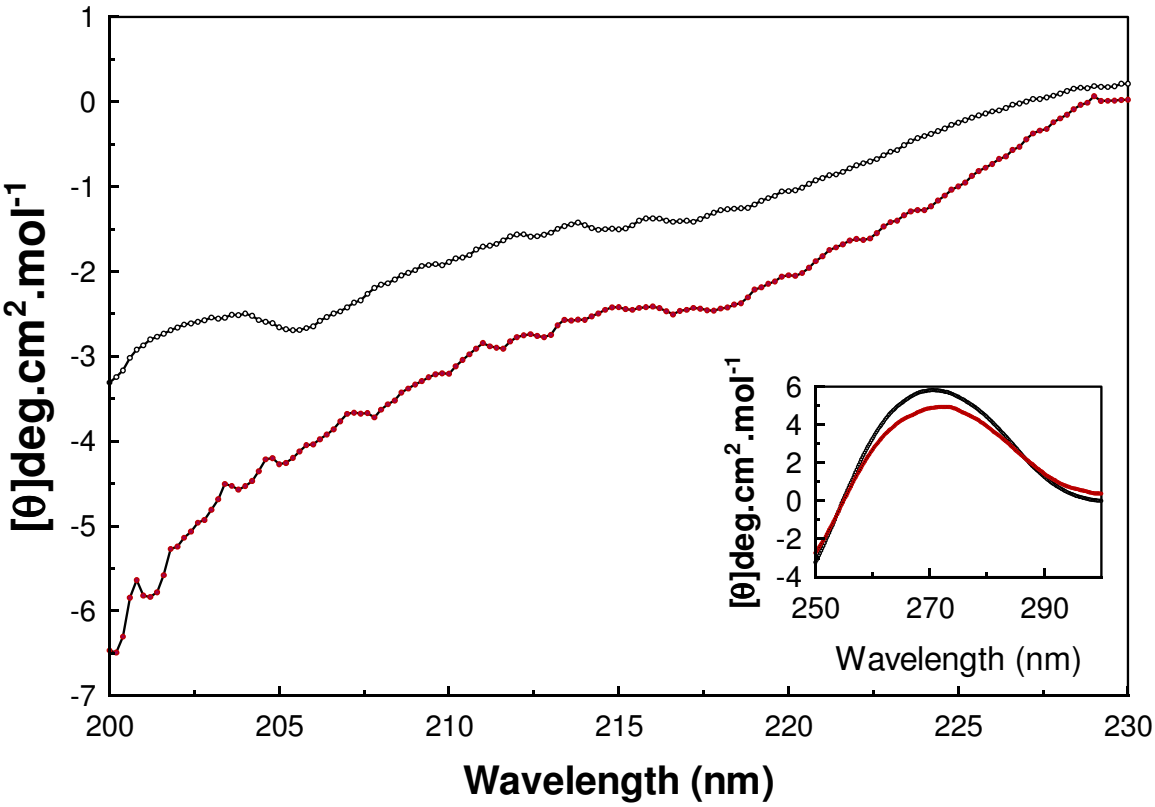
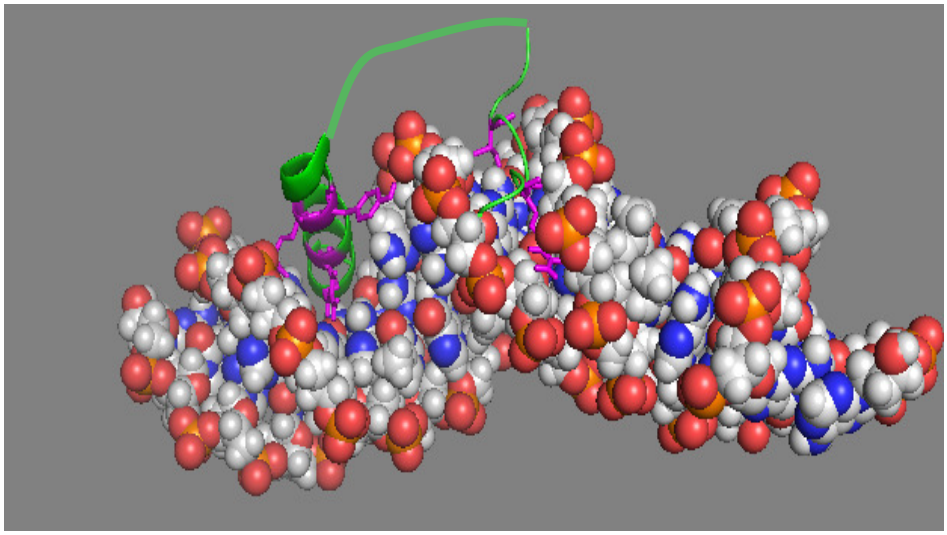


Figure S4. Far UV-CD spectra of free E3x2GS peptide (black) and in presence of DNA (red). The inset shows the CD spectra of SRE-DNA (black) and same DNA in presence of E3x2GS peptide (red)

Figure S5

A



B

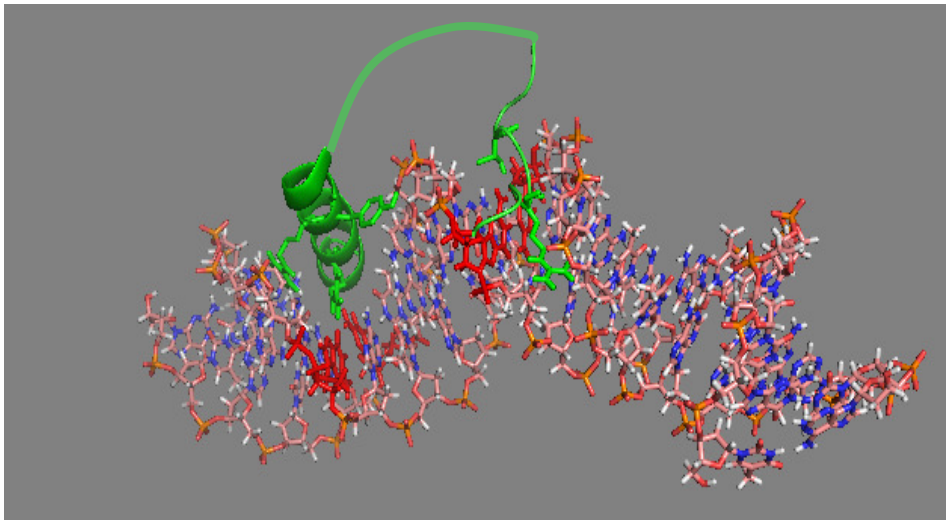


Figure S5. Molecular models of the target DNA with the Synthetic Transcription factor, constructed from the ELK1/SRF/DNA ternary complex model structure. (A) The sidechains are where the mutations were introduced in the peptide both from ELK1 and SRF parts to create the mutant peptide. (B) Red colored basepairs were mutated to create single base mutants SBM1 and SBM2 respectively of the target site.

Figure S6

Gene sequence of ELK1-DBD

ATGGACCCATCTGTGACGCTGTGGCAGTTTCTGCTGCAGCTGCTGAGAGAGCAAGGCAAT
GGCCACATCATCTCCTGGACTTCACGGGATGGTGGTGAATTCAAGCTGGTGGATGCAGAG
GAGGTGGCCCGGCTGTGGGGACTACGCAAGAACAAGACCAACATGAATTACGACAAGCT
CAGCCGGGCCTTGCGGTACTACTATGACAAGAACATCATCCGCAAGGTGAGCGGCCAGAA
GTTCGTCTACAAGTTTGTGTCCTACCCTGAGGTCGCAGGGTGCTAA

Protein sequence of ELK1-DBD

MDPSVTLWQFLLQLLREQGNGHIISWTSRDGGEFKLVDAEEVARLWGL

RKNKTNMNYDKLSRALRYYYDKNIIRKVSGQKFVYKFVSYPEVAGC

Figure S7

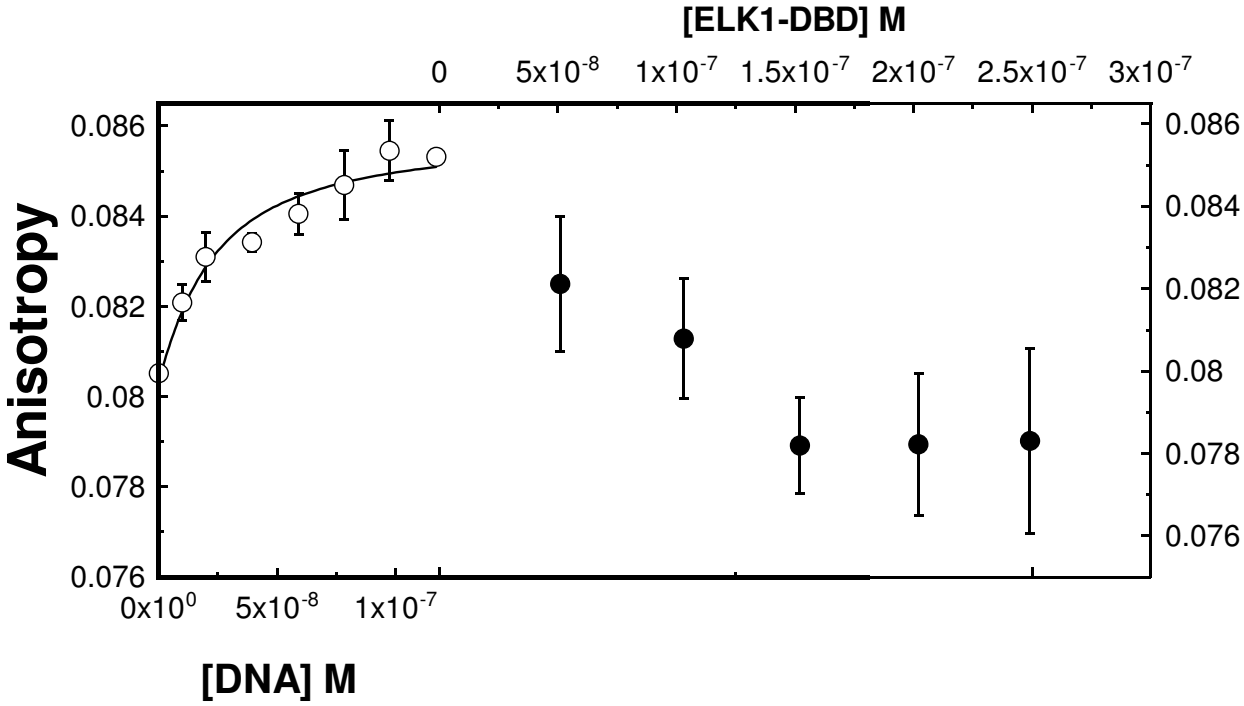


Figure S7. Binding isotherm showing competition of E3x2GS peptide and ELK1-DBD to the target DNA site. Binding isotherm with open circle shows binding of the peptide with the DNA with increase in anisotropy whereas the solid circle represents the decrease in anisotropy upon addition of ELK1-DBD, clearly suggesting the competitive binding by releasing the peptide fully from its target DNA. The error bars represent SEM based on at least three independent measurements.

Figure S8

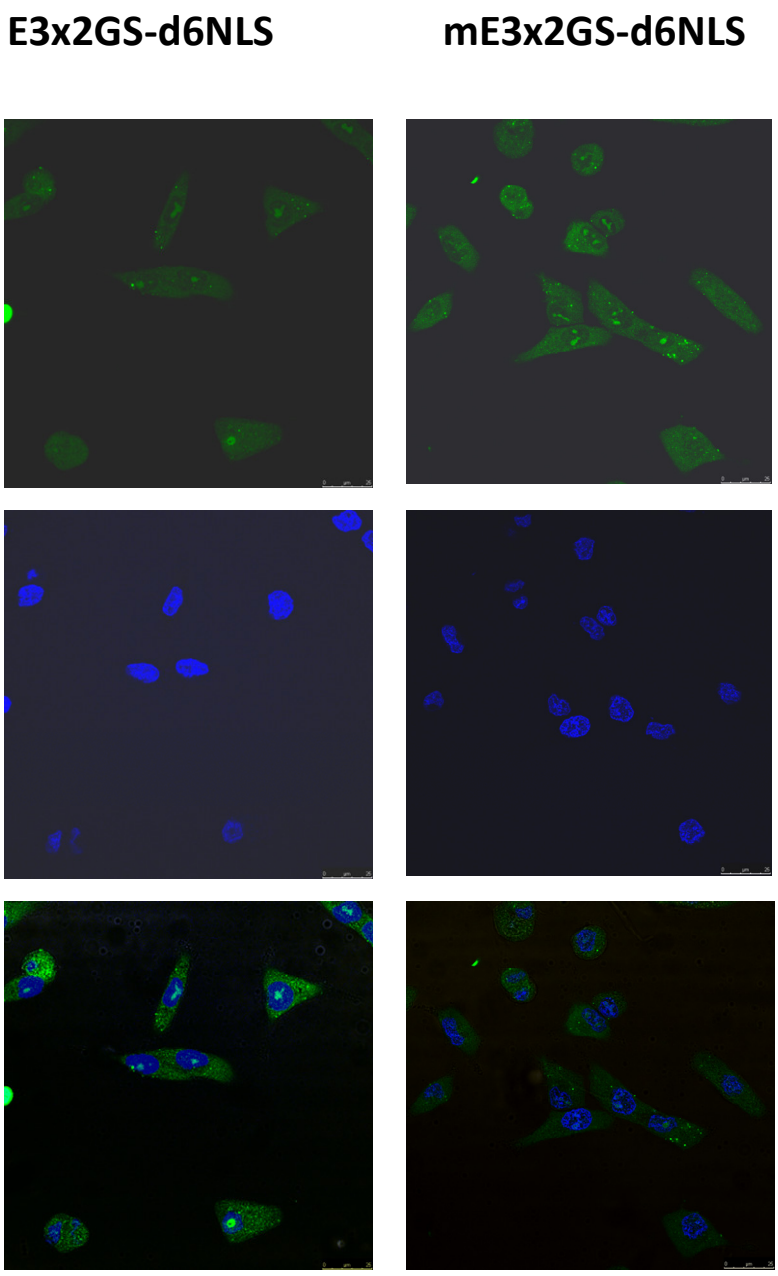


Figure S8. Entry of Fluorescein labeled peptides into A549 cells (after 5 hrs). Cells were treated with 1µM of the peptides for 5 hrs and imaged with Leica TCS SP8 Confocal Laser Scanning microscope. Brightness and contrast enhancements were made to increase visibility.

Figure S9

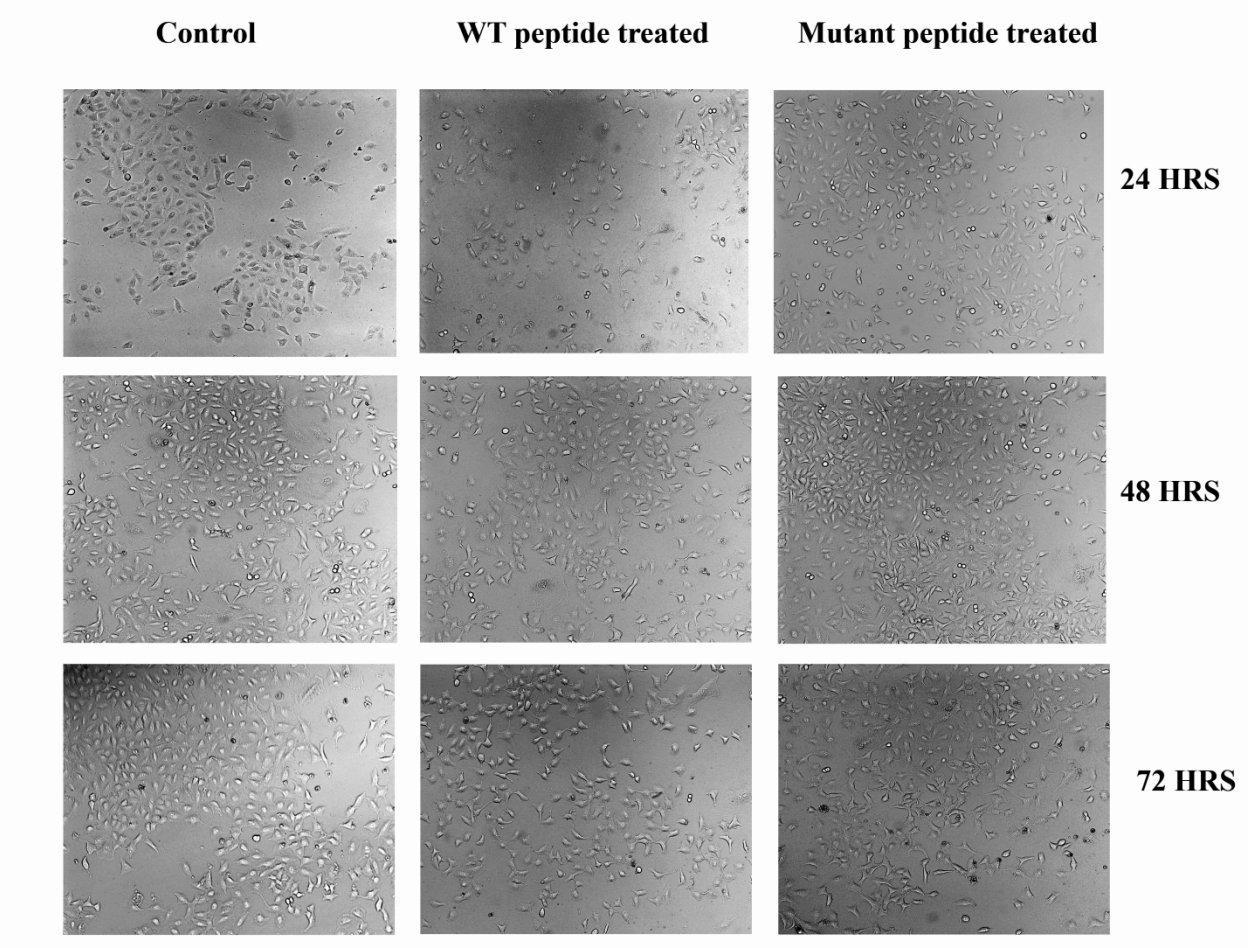


Figure S9. Morphology of A549 cells for untreated (left), E3x2GS-d6NLS peptide treated (middle), m E3x2GS-d6NLS peptide treated (right) after 24 hrs, 48 hrs, and 72 hrs

Figure S10

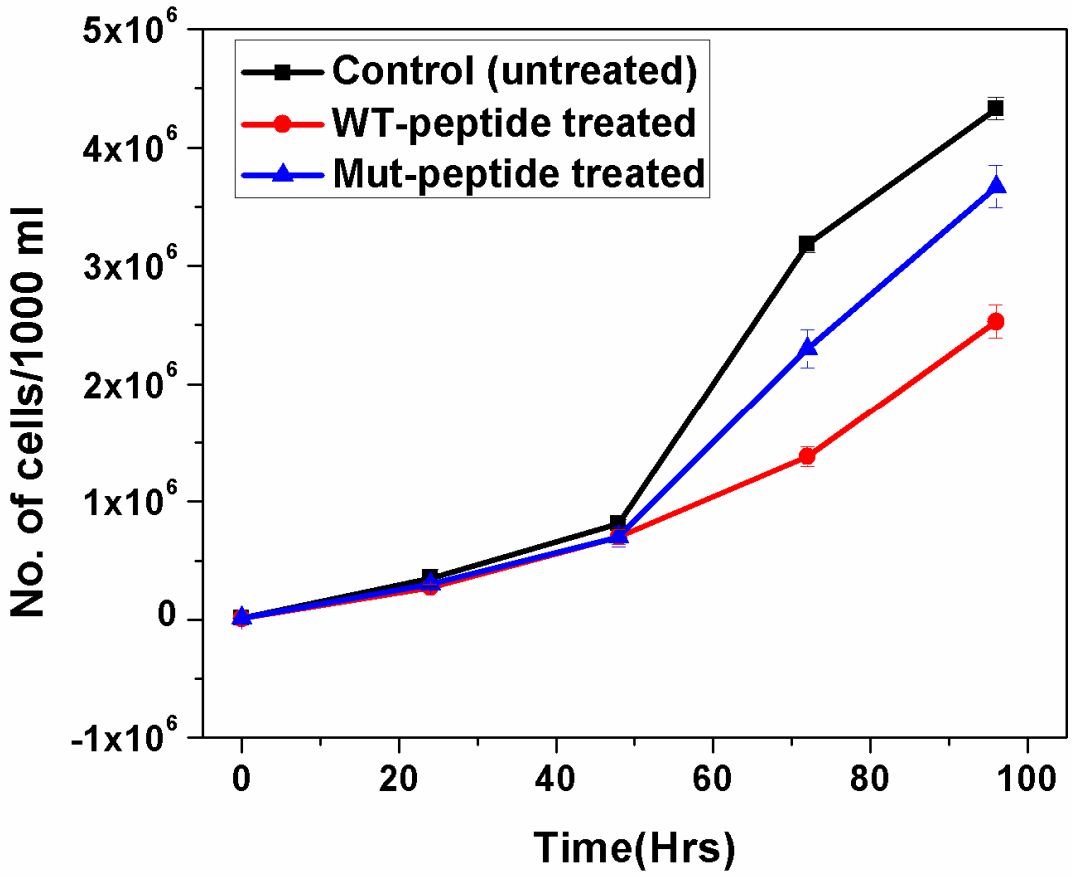
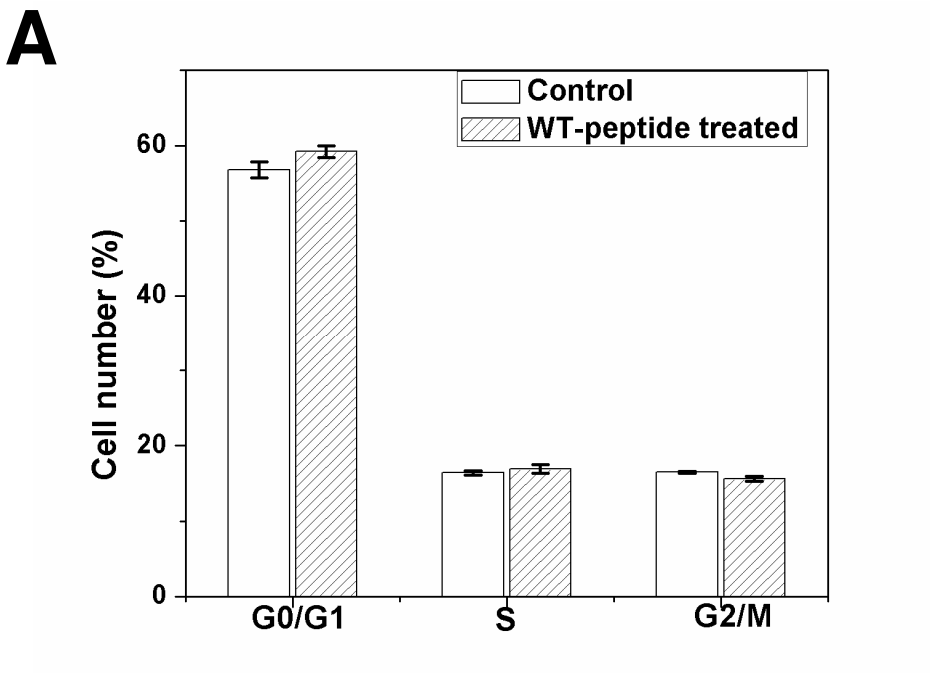


Figure S10. Comparative Growth curve for A549 cells either untreaed (black) or treated with either E3x2GS-d6NLS (red) or m E3x2GS-d6NLS (blue) peptides**.** The error bars represent SEM based on at least three independent measurements.

Figure S11



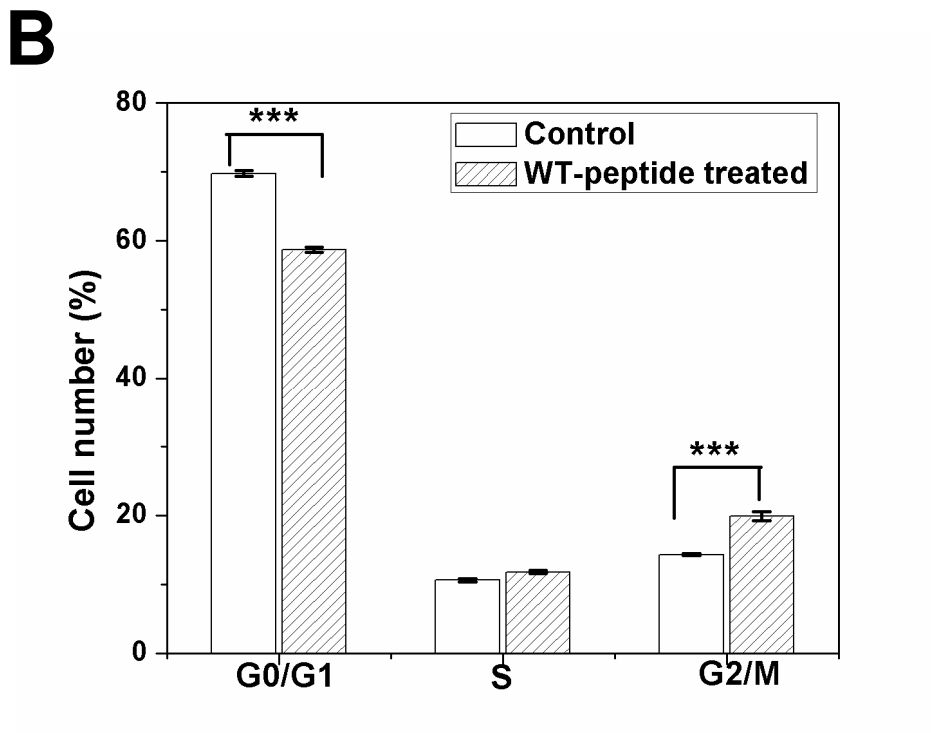


Figure S11. Cell cycle studies for (A) 48 and (B) 72 hrs. Histograms represent DNA contents according to the cell cycle phase G0/G1phase, S phase and G2/M phase for untreated and E3x2GS-d6NLS treated A549 cells. Data are expressed as the mean of percentage ± SEM. Histograms represent means of 3 independent experiments. ***: $p<0.001$.

Figure S12

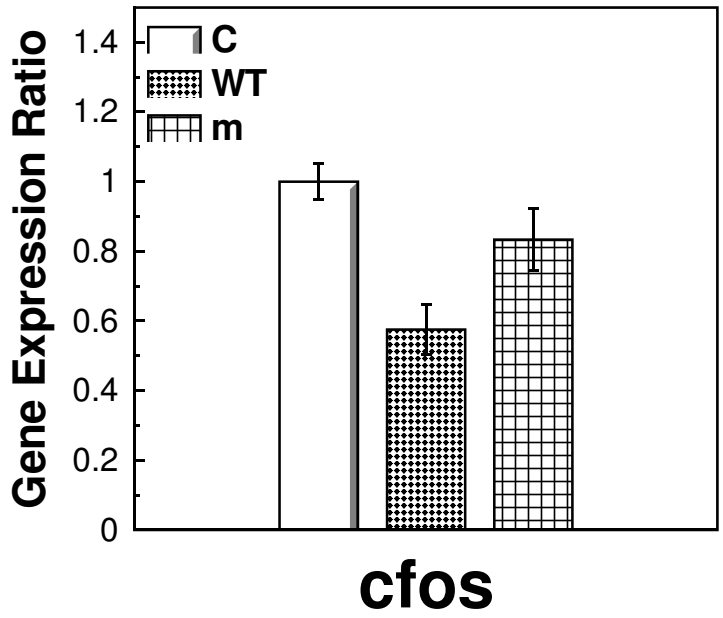


Figure S12. Real-time PCR validation of differentially expressed CFOS gene in presence of either WT or mutant peptides in comparison with untreated control. Data are expressed as the mean ± SEM of the fold-change in gene expression and based on at least three independent experiments.

Figure S13

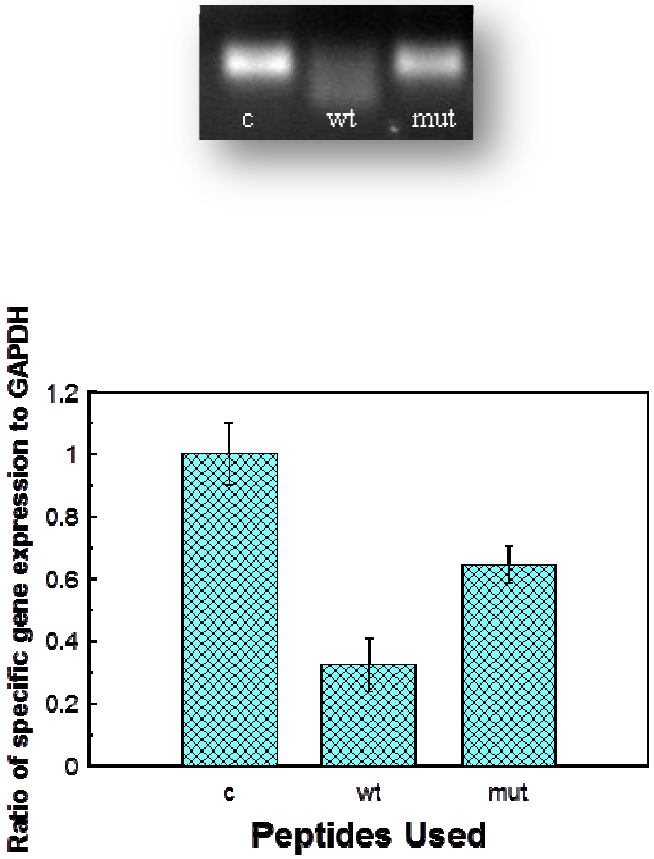


Figure S13. Top panel shows the band intensities obtained from a typical semi-quantitative PCR for *P21* gene expression. C, represents RNA isolated from untreated A549 cells; wt, represents RNA isolated from E3x2GS-d6NLS treated cells; mut, represents RNA isolated from mE3x2GS-d6NLS treated cells. The bottom panel is the quantitation of the semi-quantitative PCR results from several repeat experiments. Data are expressed as the mean of fold change ± SEM, based on at least three independent measurements.

Figure S14

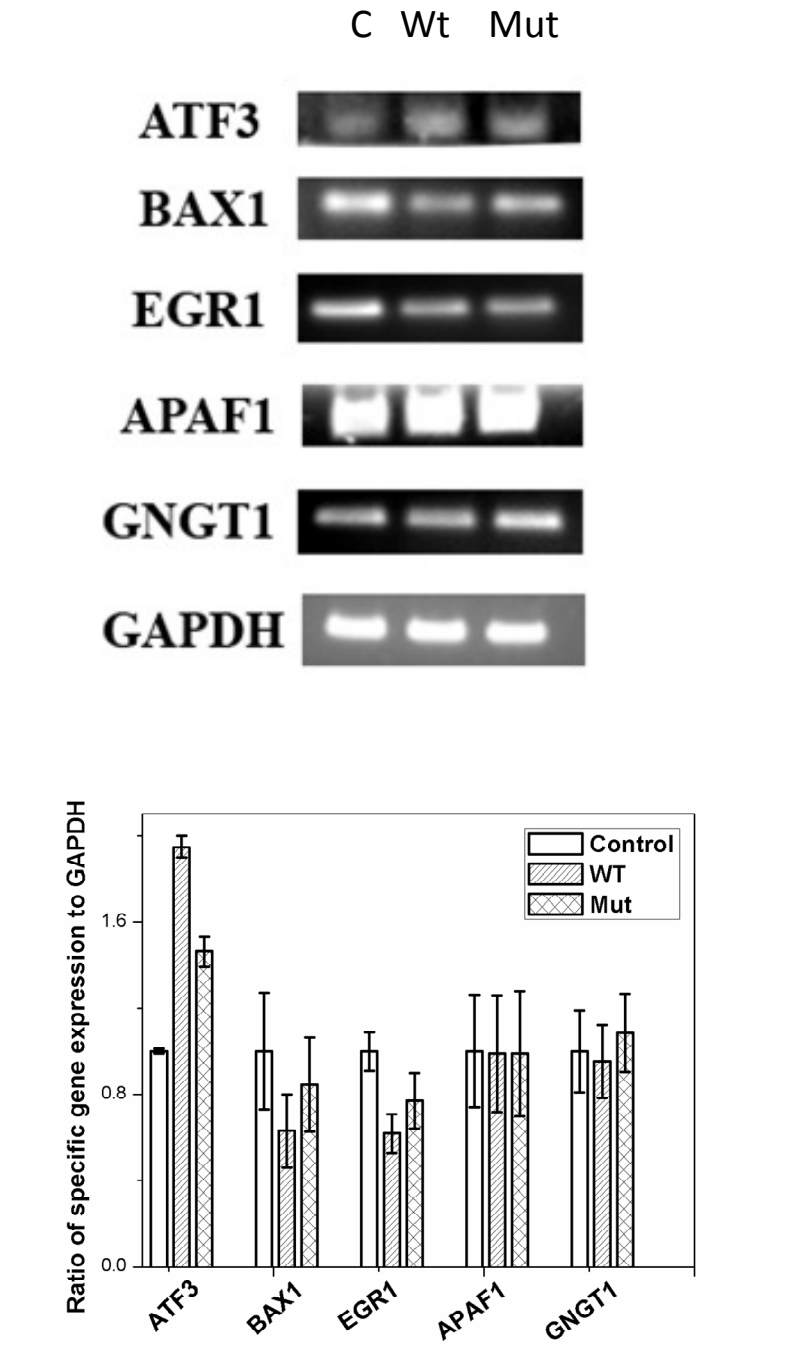


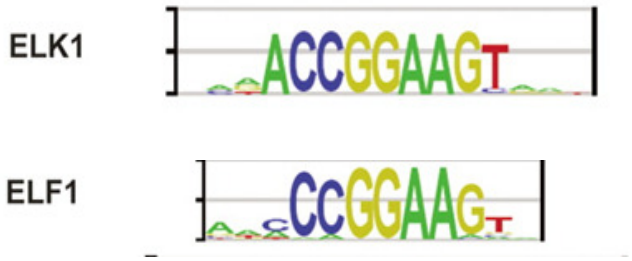
Figure S14. Semi-quantitative PCR profiles of three ELK1-target genes, ATF3, BAX1 and EGR1 and two off-target genes APAF1 and GNGT1. In the top panel, C refers to control with no added peptide, wt refers to E3x2GS-d6NLS and mut refers to mE3x2GS-d6NLS treated cells. The bar diagram shows ratio of gene expression of these genes, upon WT and mutant peptide treatment as determined by semi-quantitative PCR. Data are expressed as the mean of fold change ± SEM, based on at least three independent measurements.

Figure S15

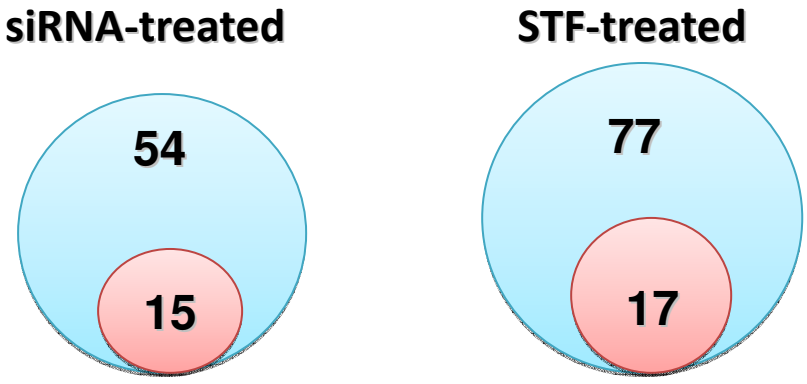
A

Run	Matrix	Transcription Factor	Show Target Genes	Association Score	P-Value
1	ETS_Q6	Elf-1 , Elfr	Targets	8.011	1.00e-06
2	ETS_Q6	Elf-1 , Elfr	Targets	4.716	1.22e-03

B



C



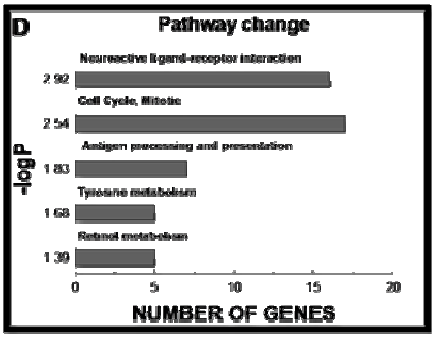
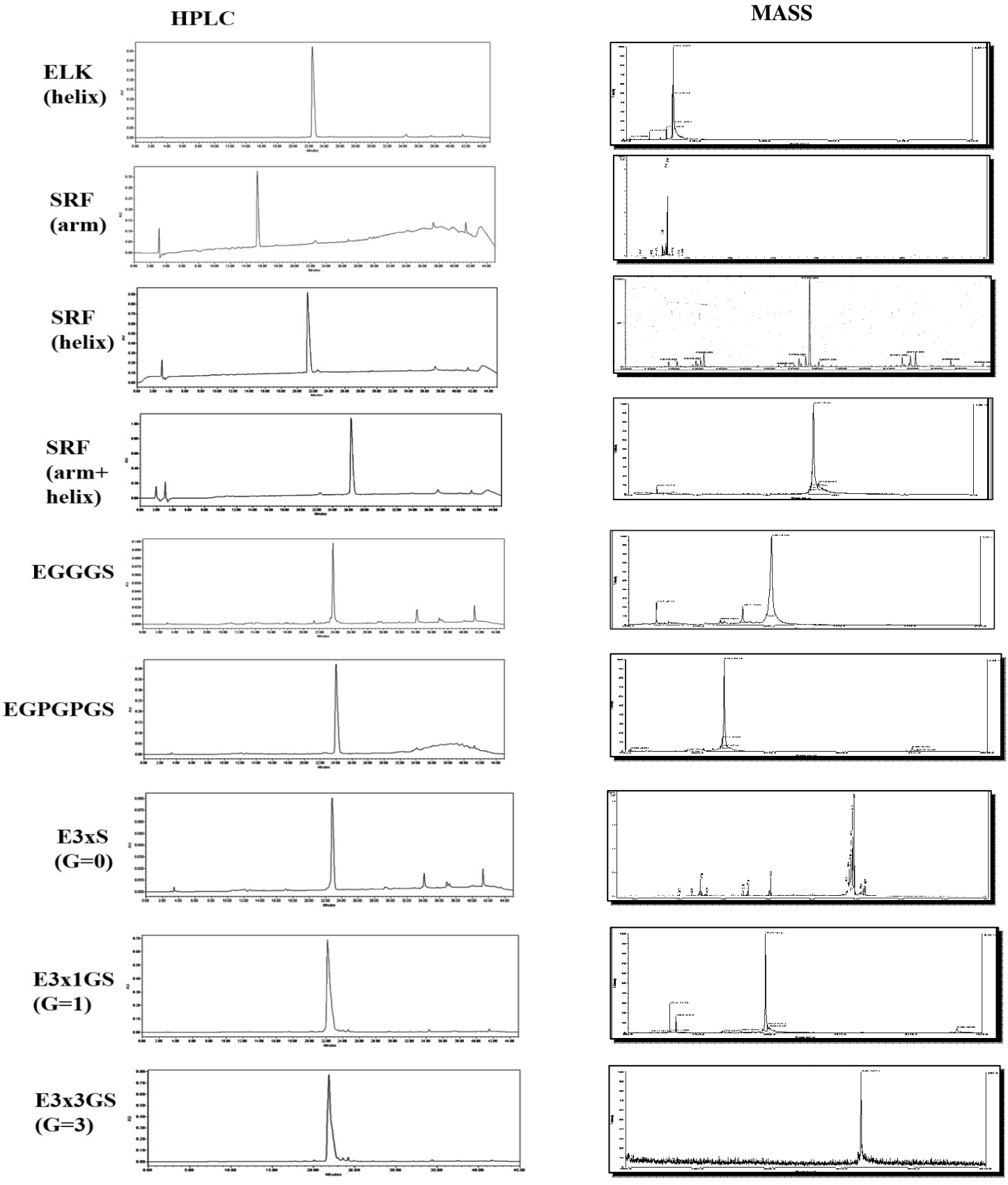


Figure S15. (A) Association of ETS family transcription factors with the up and down regulated genes. Up and down regulated gene set (combined) from two separate microarray experiments were run on PASTAA web server using standard parameter sets. (B) Logos of consensus binding sequences of ELK1 and ELF1. (C) The number of down-regulated genes in siRNA treated (left panel) as reported by Odrowaz *et al*. Synthetic Transcription Factor treated A549 cells (right panel). The red circles represent gene down-regulated four-fold or more and blue circles represent genes down-regulated three-fold or more. (D) The number of significantly down-regulated genes observed in different cellular pathways upon addition of E3x2GS-d6NLS peptide. Plot has been generated on the basis of microarray analysis at earlier time point (8 hr).

Figure S16


HPLC
MASS
ELK (helix)
SRF (arm)
SRF (helix)
SRF (arm+ helix)
EGGGS
EGPGPGS
E3xS (G=0)
E3x1GS (G=1)
E3x3GS (G=3)

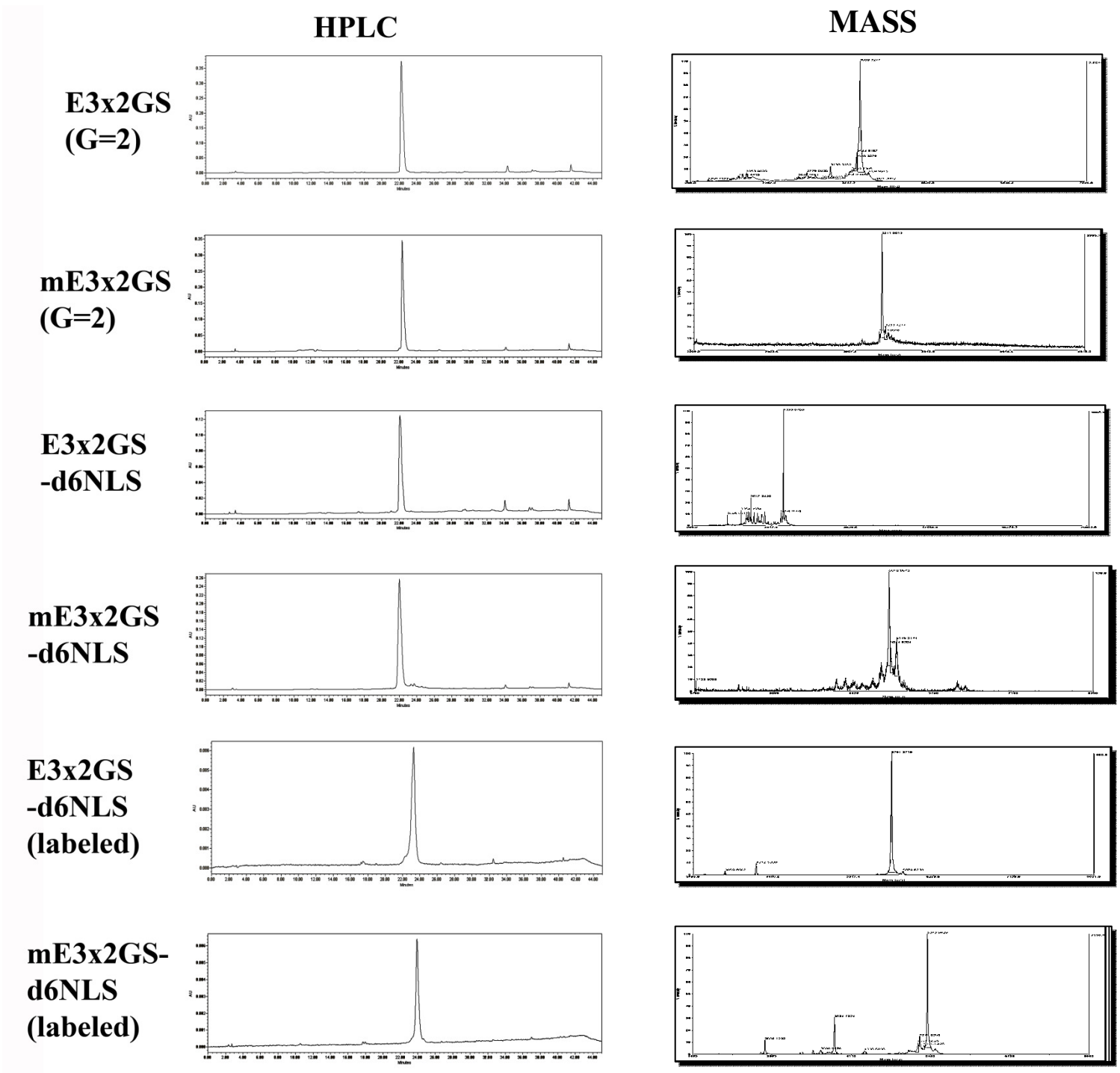


Figure S16. HPLC profile (left) and Mass spectra (right) of all synthesized peptides listed in Table S2 and S3. The last two rows are for labeled peptides used for nuclear localization experiments.

Figure S17

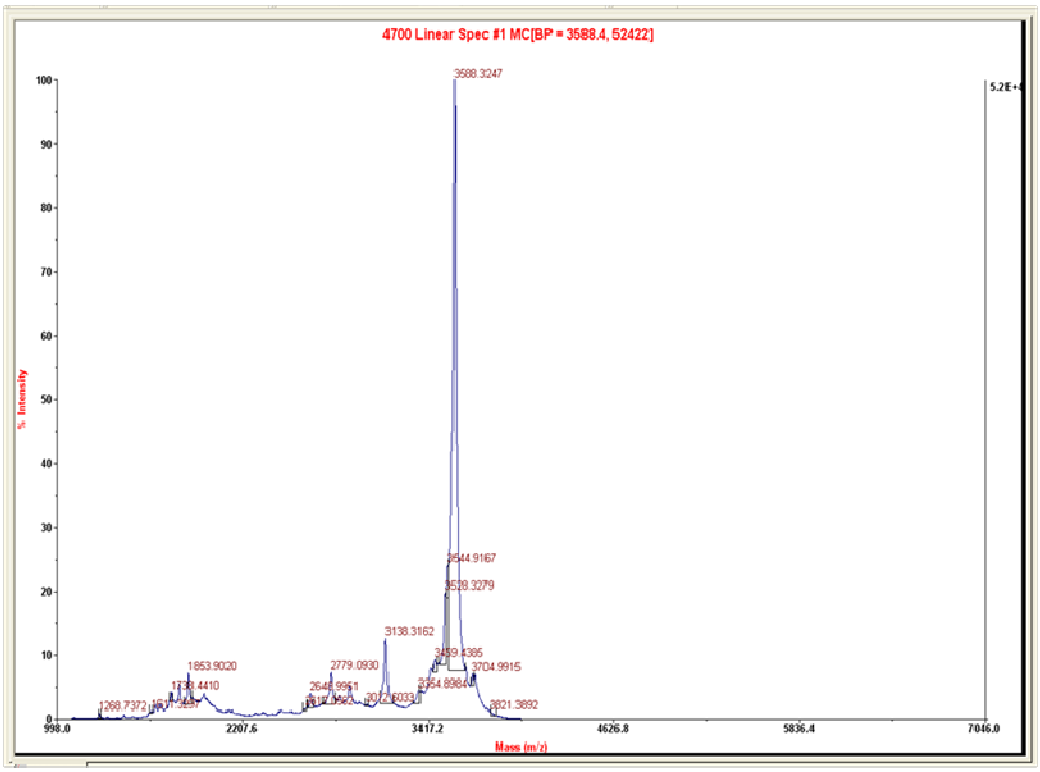


Figure S17. Maldi-Tof mass spectra of E3x2GS.

Figure S18

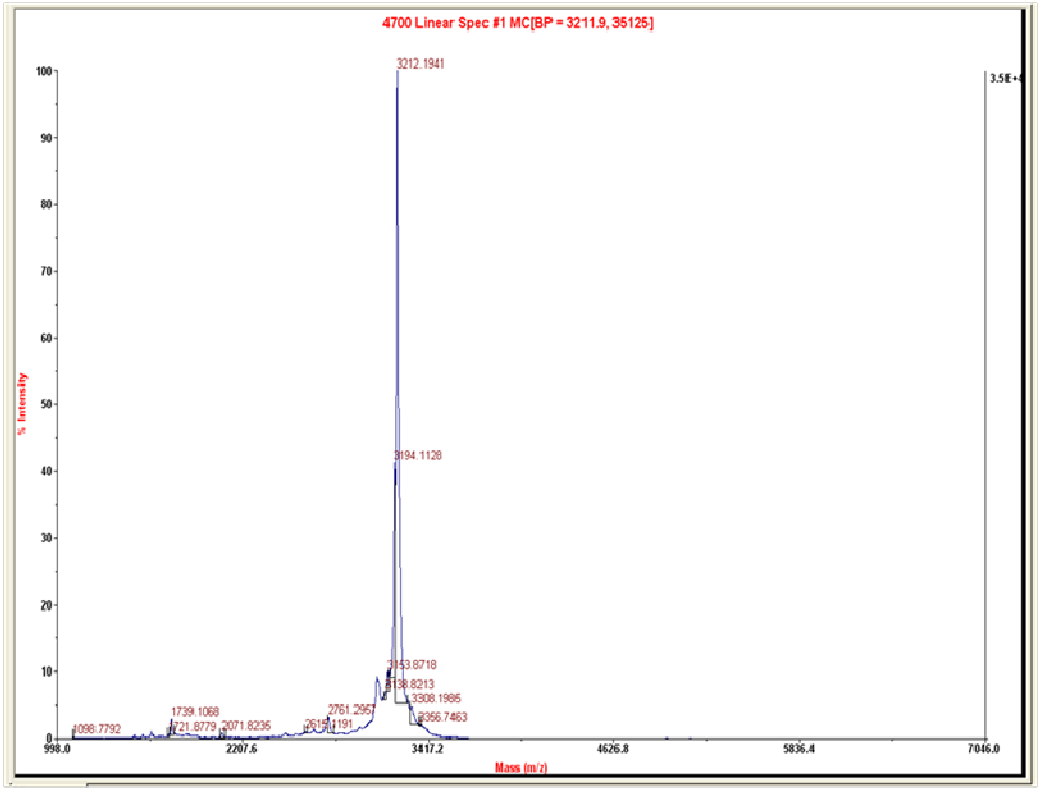


Figure S18. Maldi-Tof mass spectra of mE3x2GS.