Supporting Information

Post cross-linked ROS-responsive poly (β-amino ester)-plasmids polyplex NPs for gene therapy of EBV-associated nasopharyngeal carcinoma

Caiyan Yuan[a,c,e], Shuangyan Chang[b,e], Chong Zhang[a,e], Dirong Dong[b], Jiahui Ding[a], Ali Reza Mahdavian[d], Zheng Hu[b], Lili Sun[b,*] Songwei Tan[a,*],

a. School of Pharmacy, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China. E-mail:Tansw@hust.edu.cn, tansongwei@gmail.com

b. Department of Obstetrics and Gynecology, Zhongnan Hospital of Wuhan University, Donghu 169th Road, Wuchang District, Wuhan 430062, Hubei, China. E-mail: huzheng1998@163.com

c. The first hospital of Nanchang, Nanchang 330008, China

d. Polymer Science Department, Iran Polymer and Petrochemical Institute, Tehran 14967, Iran;

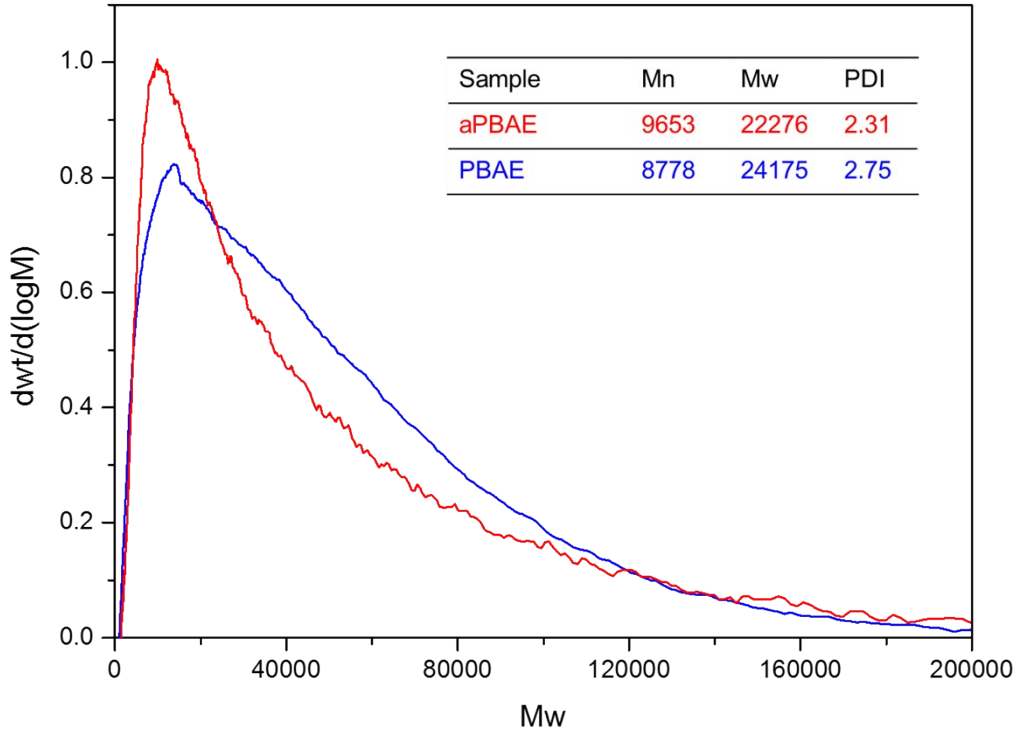

Fig S1. Gel Permeation Chromatography (GPC) results of aPBAE and PBAE.

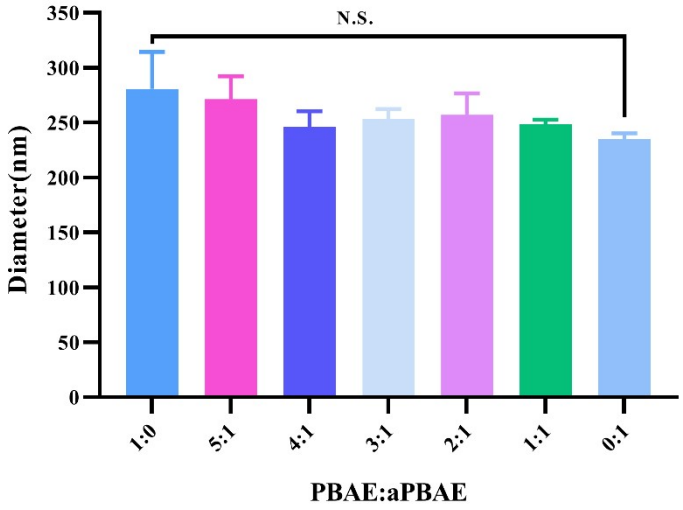

Fig. S2. Particles size of polyplex NPs at the different ratios between PBAE and aPBAE. N.S.: no significant difference.

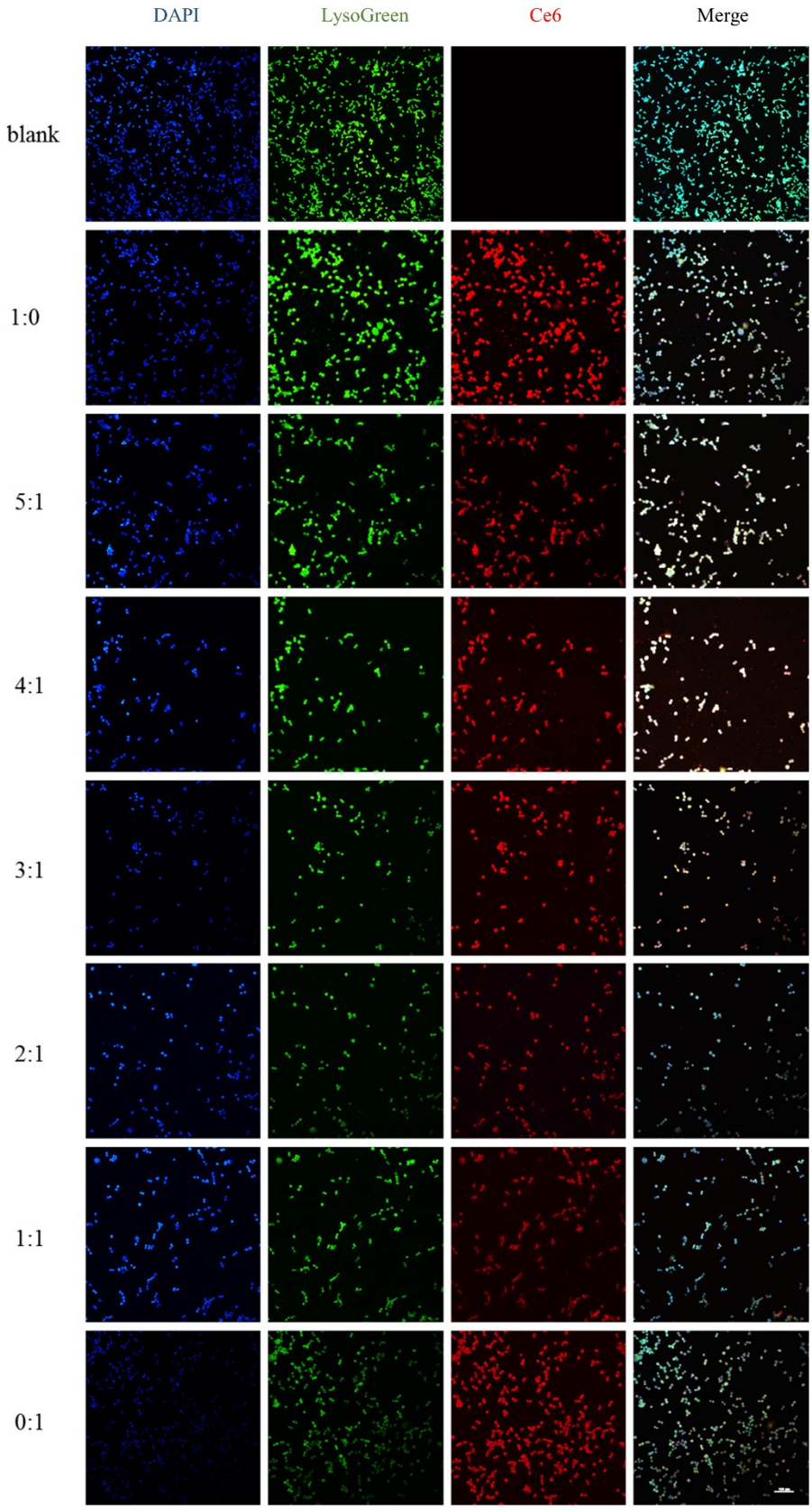

Fig. S3. Typical fluorescence microscopy images of intracellular uptake of C666-1 cells at the different ratios between PBAE and aPBAE at 2h (Scale bar, 100 μm).

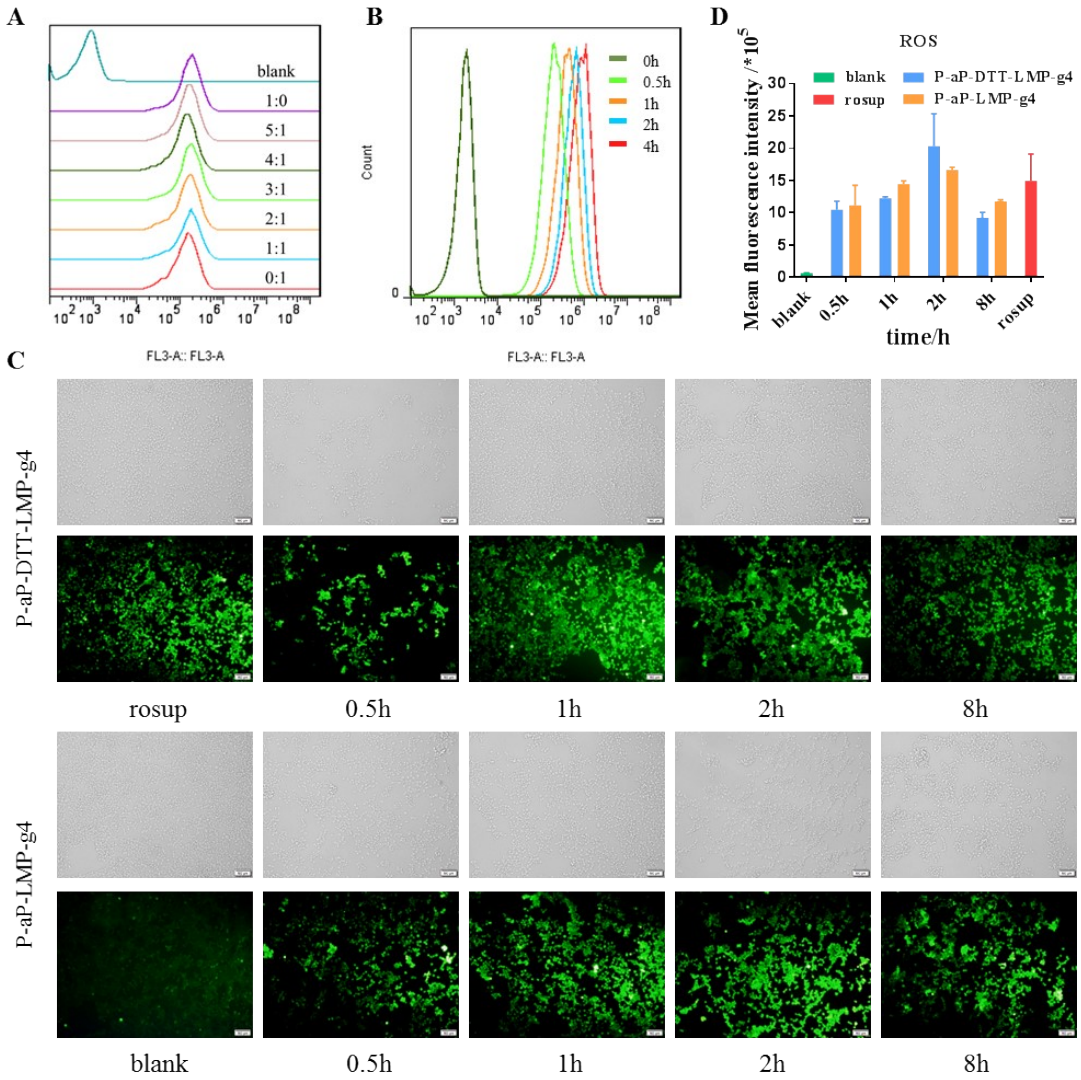

Fig. S4. Intracellular uptake of P-ap-DTT-plasmids polyplex NPs and ROS on 293T cells. (A) Uptake migration diagram at the different ratios between PBAE and aPBAE at 2h (plasmids: GFP) and (B) at different time points. (C) Typical fluorescence microscopy of intracellular ROS level of HEK 293T cells incubated with P-aP-DTT/P-aP-LMP-g4 polyplex NPs and was detected by DCFH-DA probe at different time points (Scale bar, 100 μm) and (D) corresponding mean fluorescence intensity statistics (n = 2).

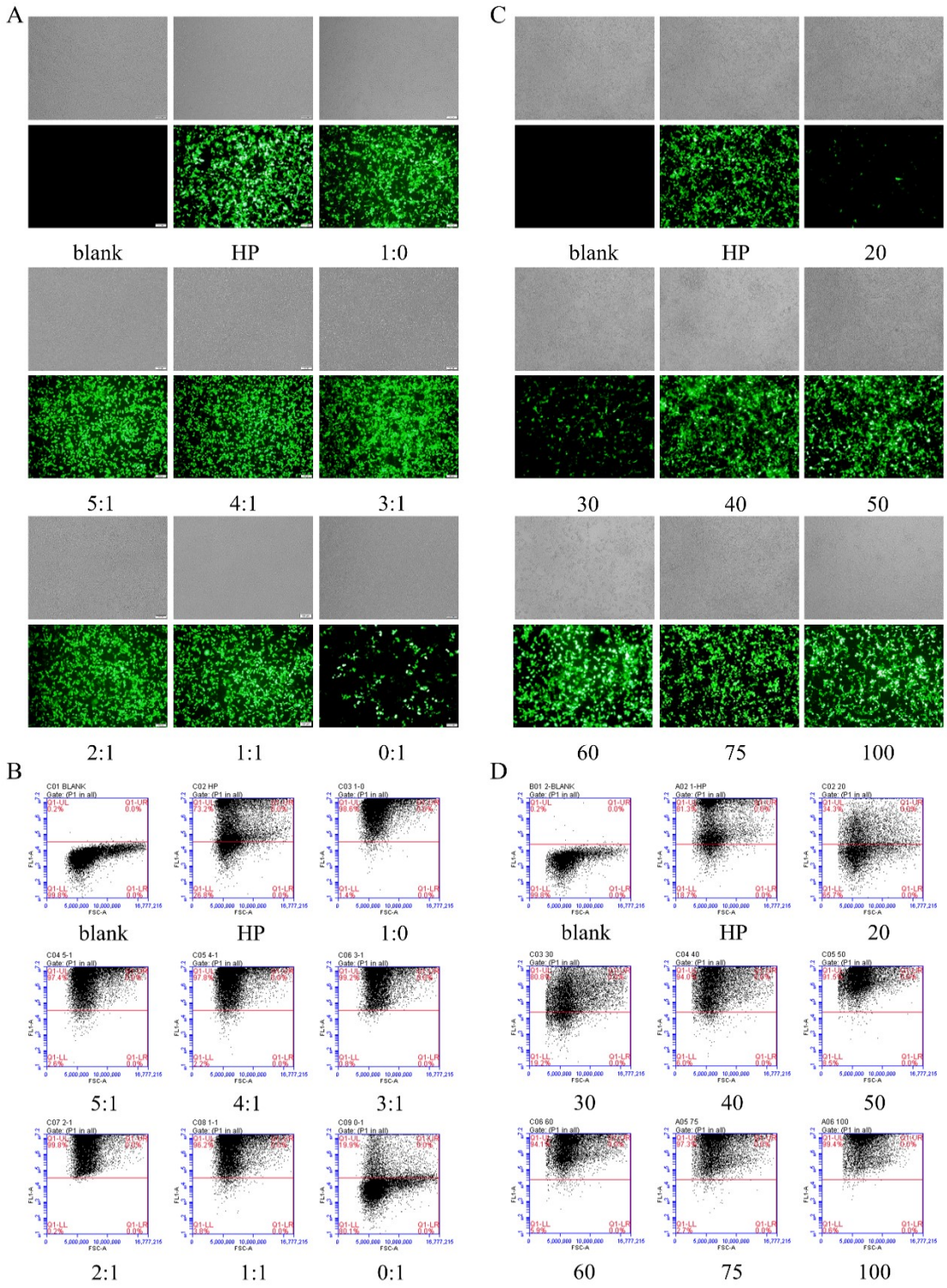

Fig. S5. Transfection efficiency (representative fluorescence microscopy images and flow cytometry results) of P-aP-DTT-GFP polyplex NPs in HEK 293T cells. (A) Typical fluorescence microscopy of transfection of HEK 293T cells at the different ratios between PBAE and aPBAE and (B) flow cytometry results. (C) Typic fluorescene microscopy of transfection of HEK 293T cells at the different mass ratios between polymer and GFP and D) flow cytometry results (Scale bar, 100 μm).

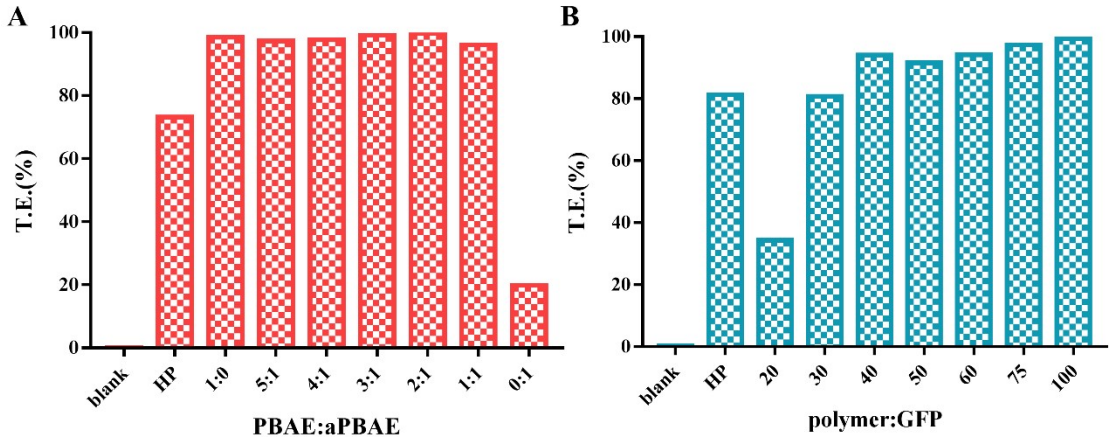

Fig. S6. Transfection efficiency (T.E.) of P-aP-DTT-GFP polyplex NPs in HEK 293T cells. (A) The quantitative results at the different ratios between PBAE and aPBAE and (B) at the different mass ratios between polymer and GFP.

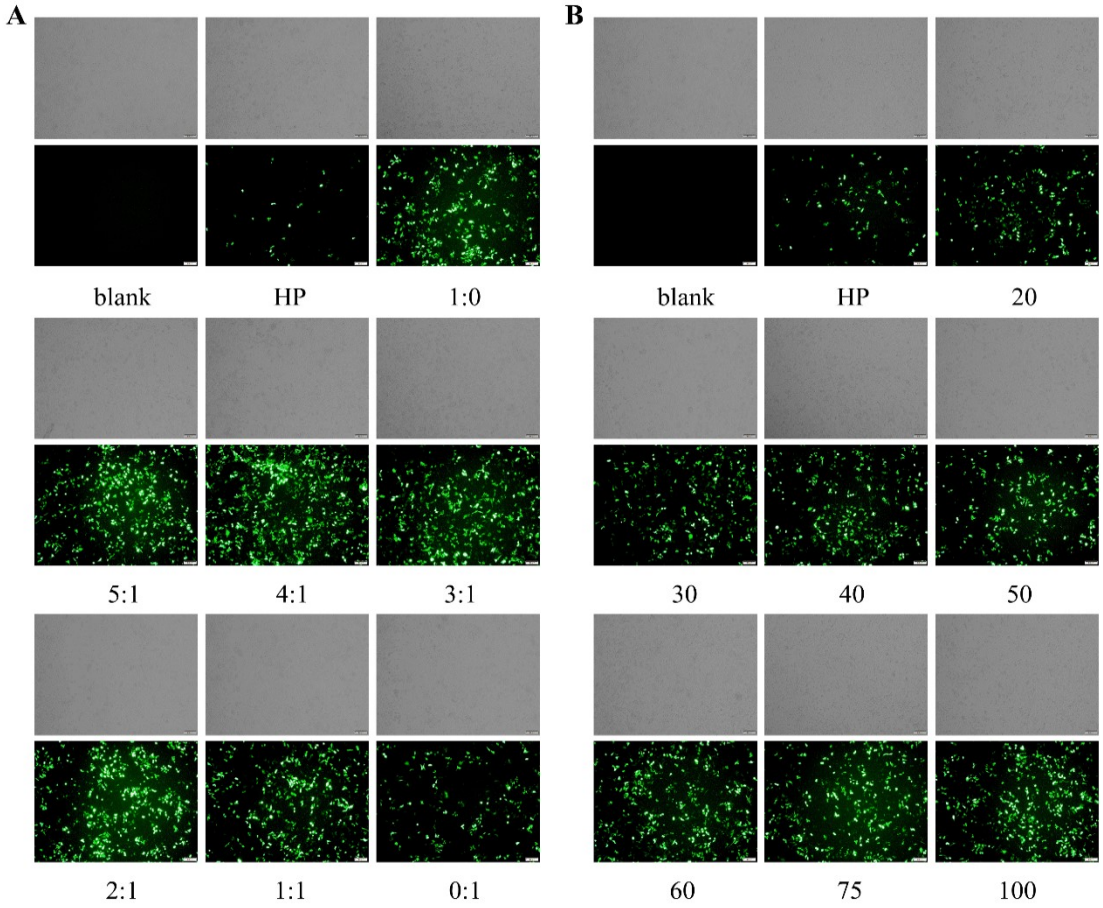

Fig. S7. Transfection efficiency (representative fluorescence microscopy images) of P-aP-DTT-GFP polyplex NPs in C666-1 cells (A) at the different ratios between PBAE and aPBAE and (B) at the different mass ratios between polymer and GFP (Scale bar, 100 μm).

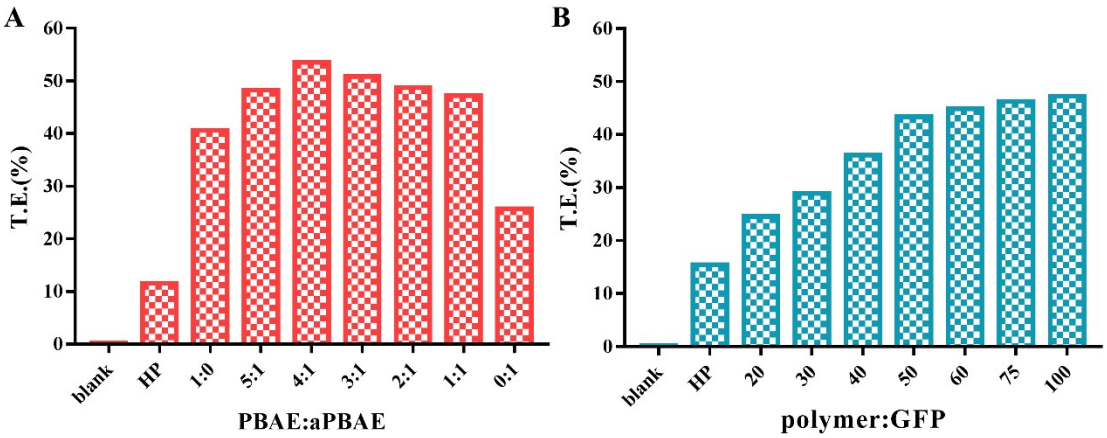

Fig. S8. Transfection efficiency (T.E.) of P-aP-DTT-GFP polyplex NPs in C666-1 cells. (A) The quantitative results at the different ratios between PBAE and aPBAE and (B) at the different mass ratios between polymer and GFP.

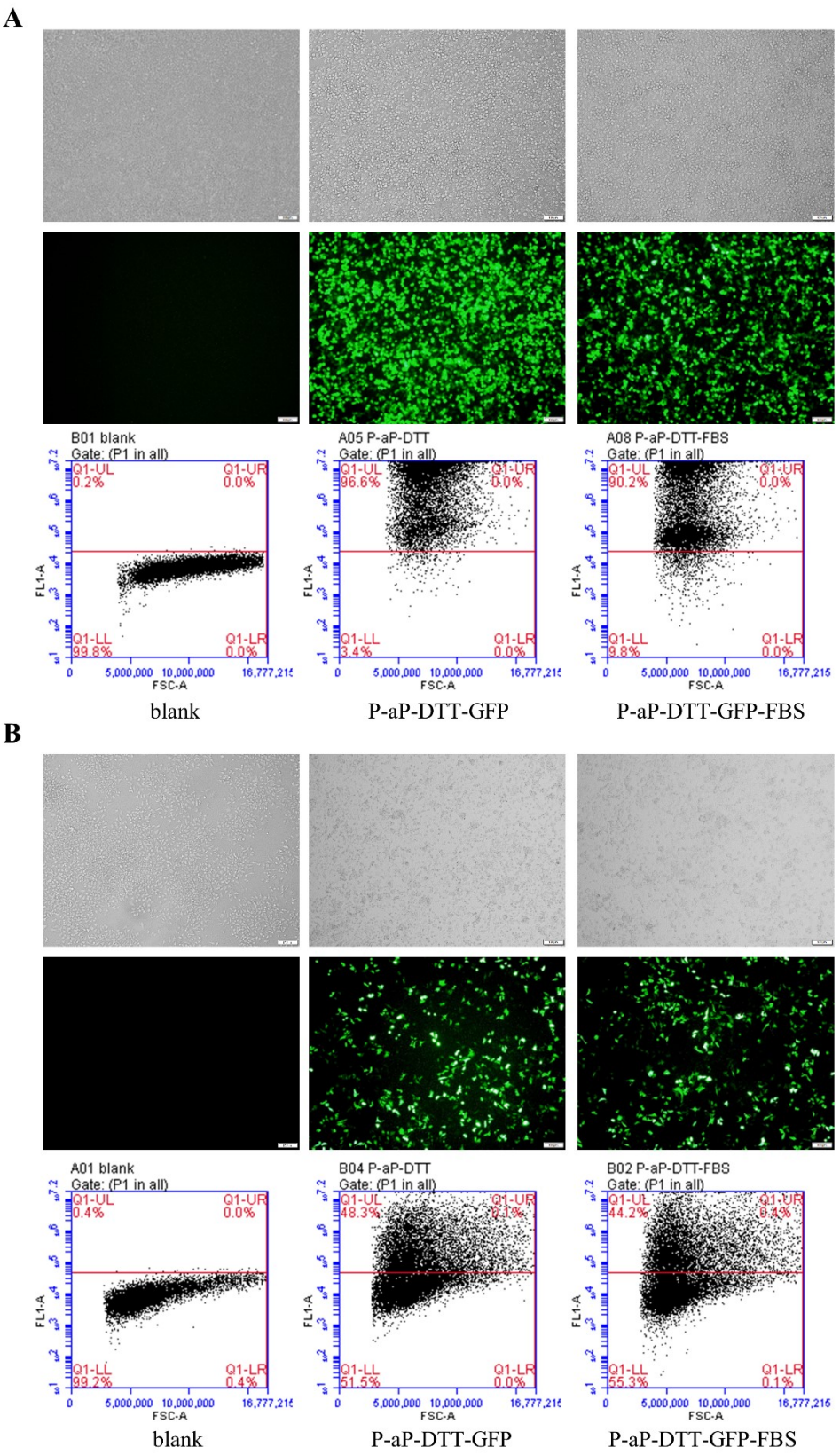

Fig. S9. Transfection efficiency (representative fluorescence microscopy images and flow cytometry results) of P-aP-DTT-GFP polyplex NPs with and without FBS in (A) HEK 293T cells or (B) C666-1 cells (Scale bar, 100 μm).

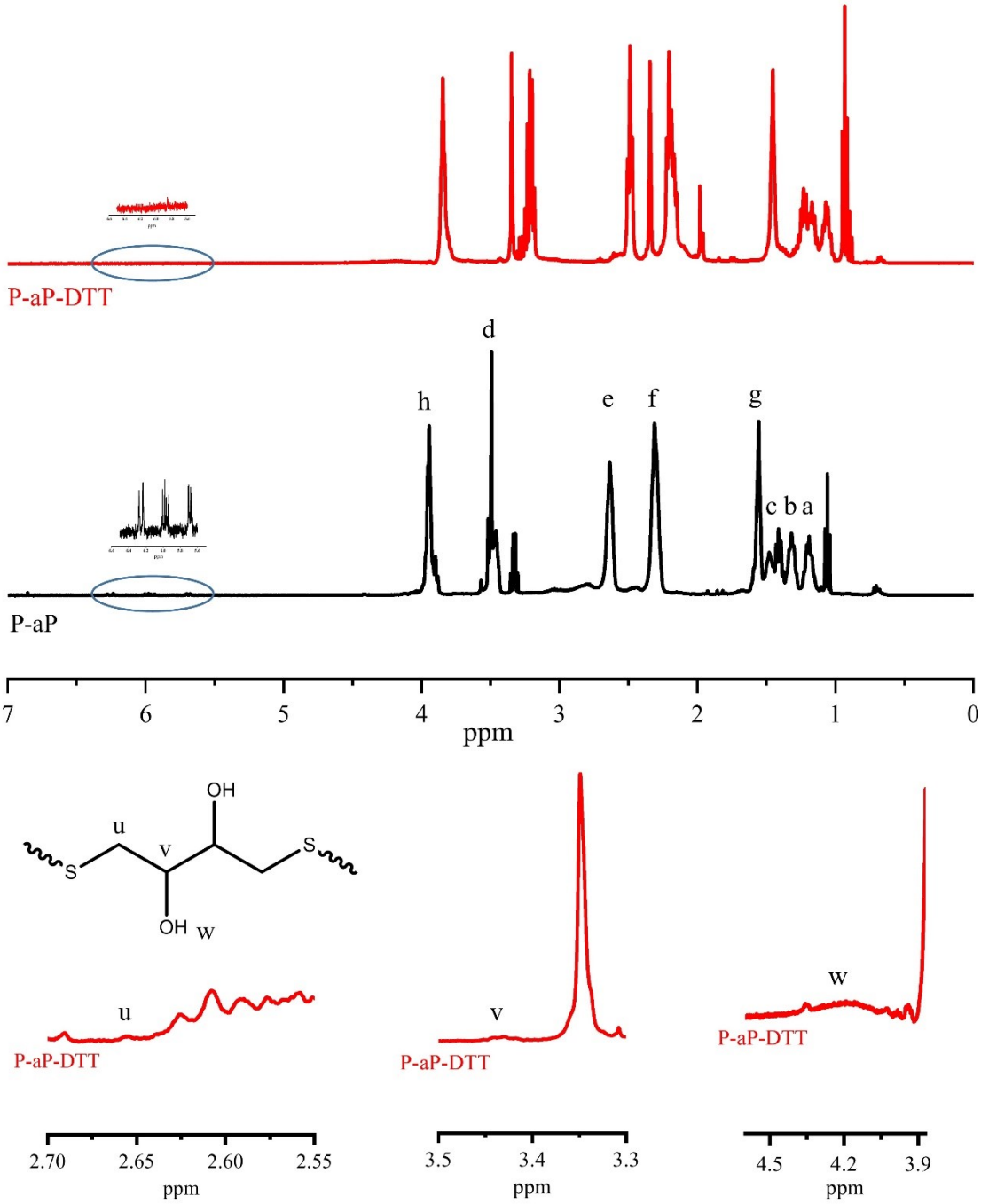

Fig. S10. ^{1}H-NMR spectra of of P-aP-DTT and P-aP.

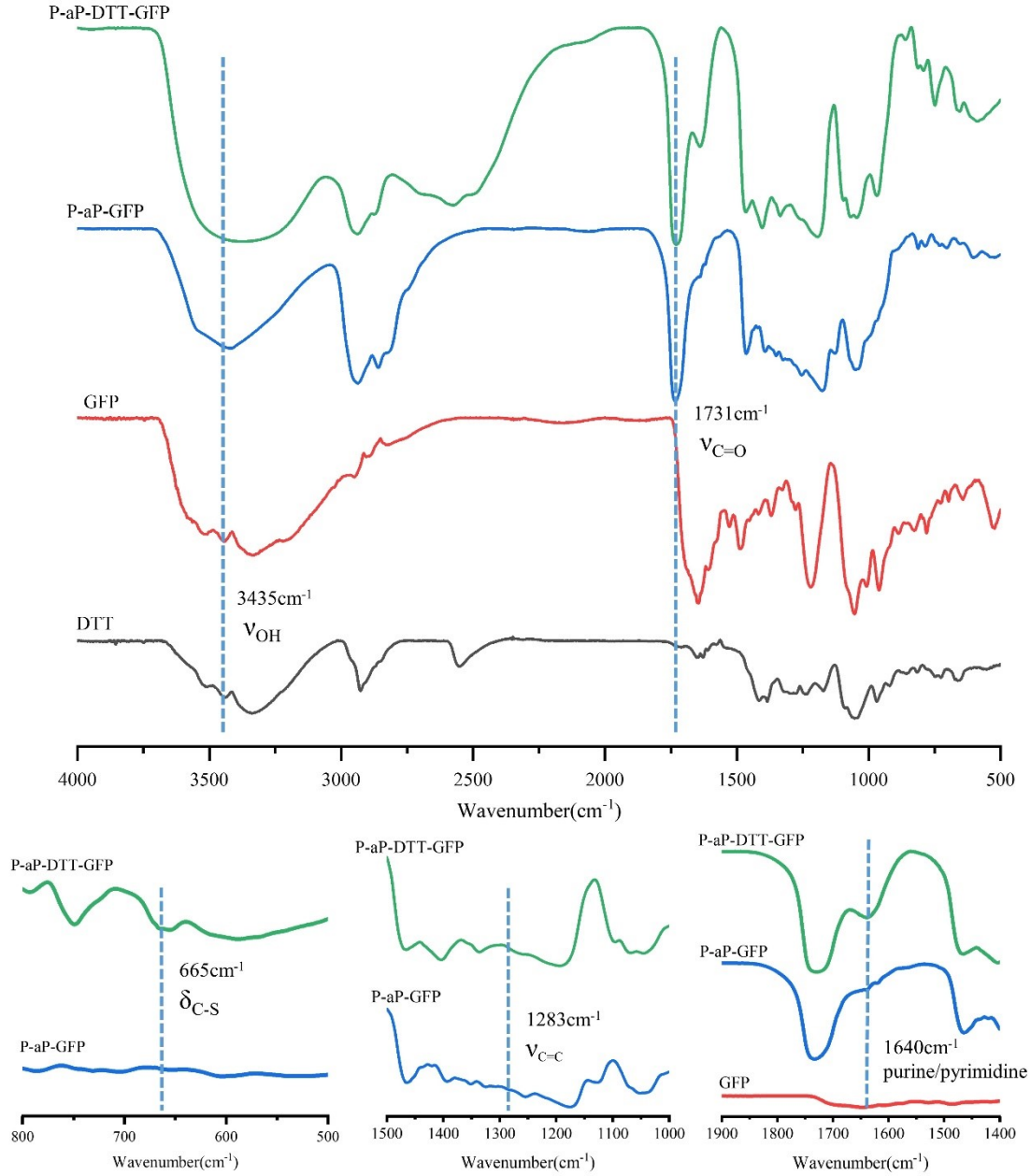

Fig. S11. FT-IR spectra of of P-aP-DTT-GFP polyplex NPs, P-aP-GFP polyplex NPs, GFP and DTT.

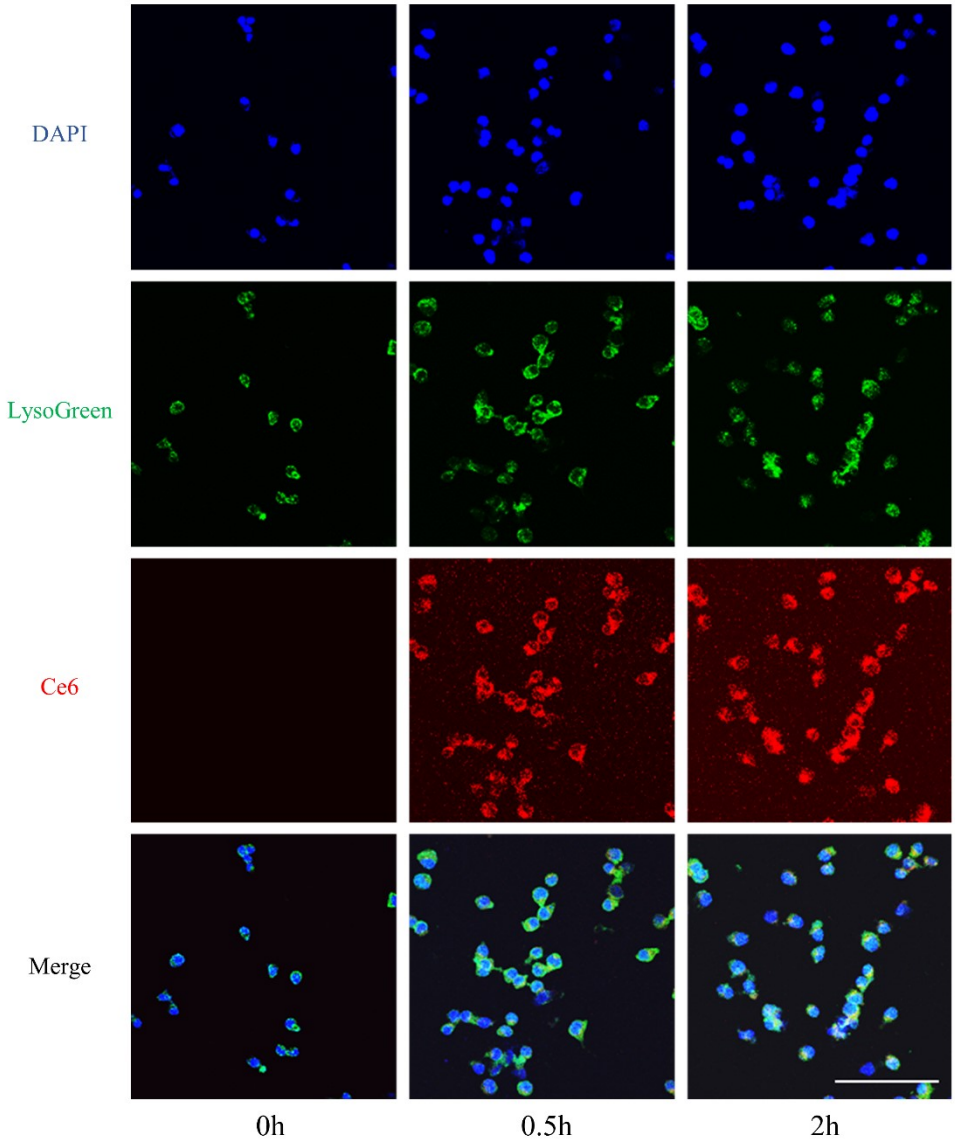

Fig. S12. Typical fluorescence microscopy images of intracellular uptake of C666-1 cells at different time points, blue: nucleus, green: lysosome, red: Ce6-P-aP-DTT-LMP-g4. (Scale bar, 100 μm)

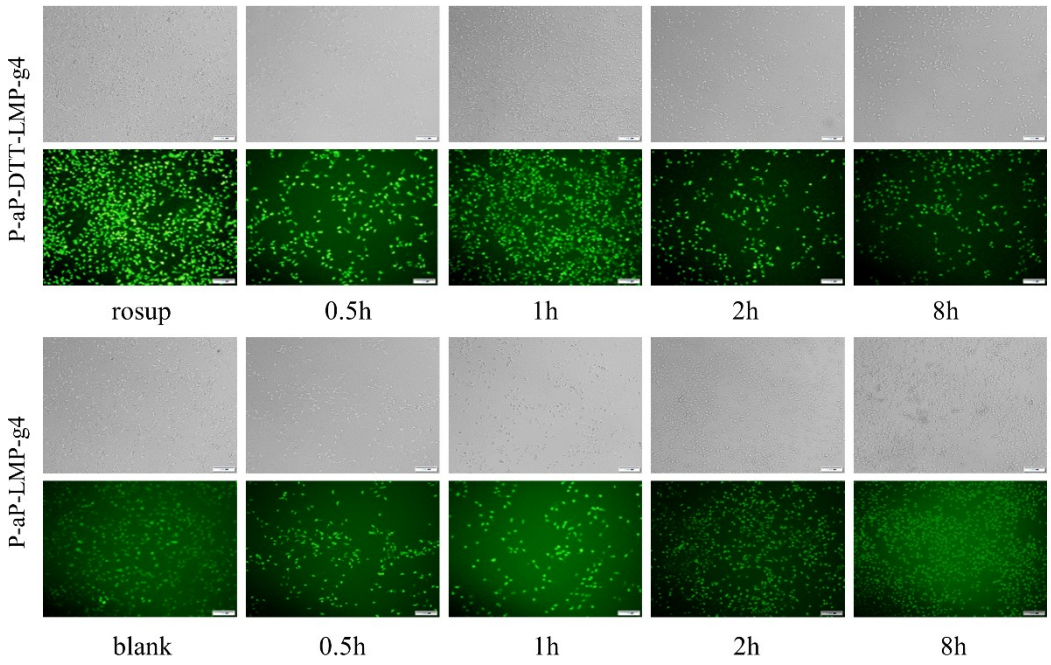

Fig. S13. Typical fluorescence microscopy of intracellular ROS level of C666-1 cells incubated with P-aP-DTT/P-aP-LMP-g4 polyplex NPs and was detected by DCFH-DA probe at different time points (Scale bar, 100 μm).

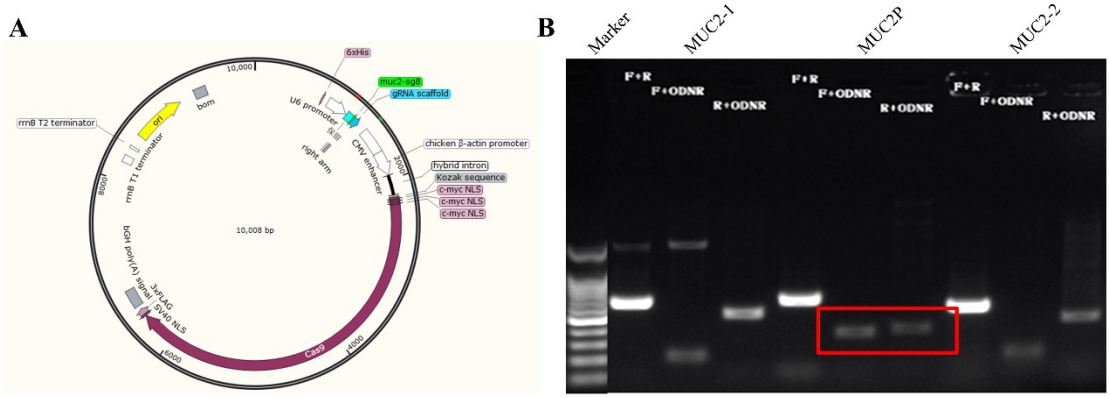

Fig. S14. Construction of *Muc2* targeting CRISPR/Cas9 plasmid (A) The schematic of plasmid. (B) Selection of *Muc2* targeting sgRNA before *in vivo* experiments.

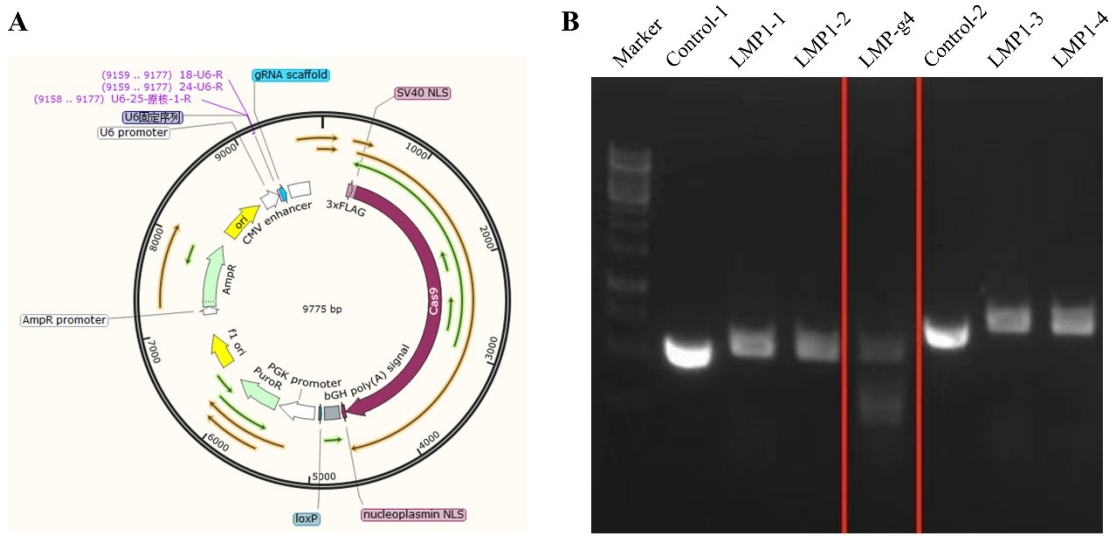

Fig. S15. Construction of *Lmp1* targeting CRISPR/Cas9 plasmid (A) The schematic of plasmid. (B) Selection of *Lmp1* targeting sgRNA before in vivo experiments.

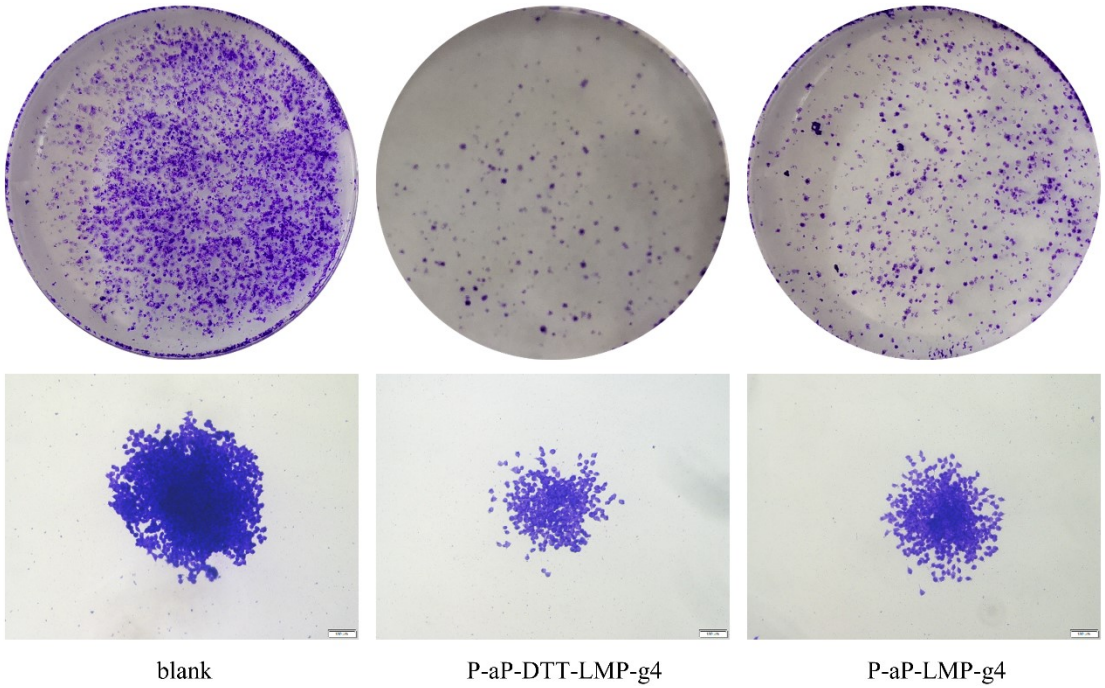

Fig. S16. C666-1 cells proliferation reflected by cell colony formation assay (Scale bar, 100 μm).

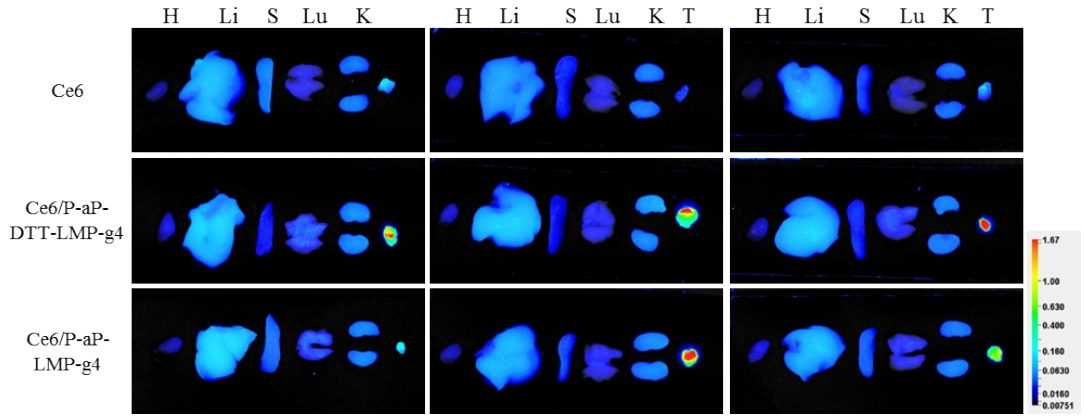

Fig. S17. Fluorescent images of hearts (H), livers (Li), spleens (S), lungs (Lu), kidneys (K) and tumors(T) in each group after administration for 24 h (n = 3).

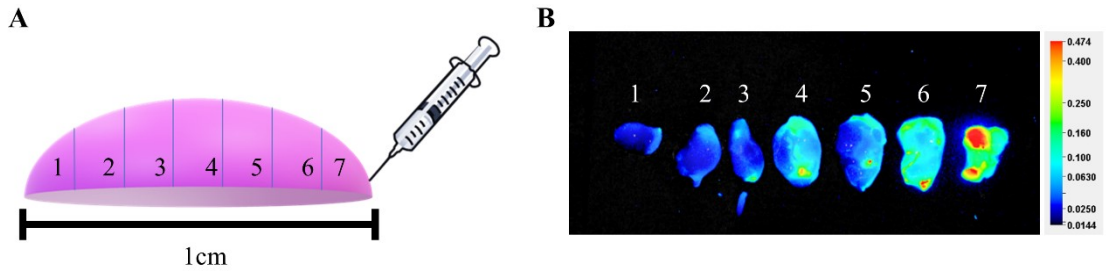

Fig. S18. Penetration of P-aP-DTT-GFP polyplex NPs after peritumoral administration for 24 h. (A) Schematic diagram of the direction of administration and tumor lateral resection. (B) Fluorescence images of tumor slices.

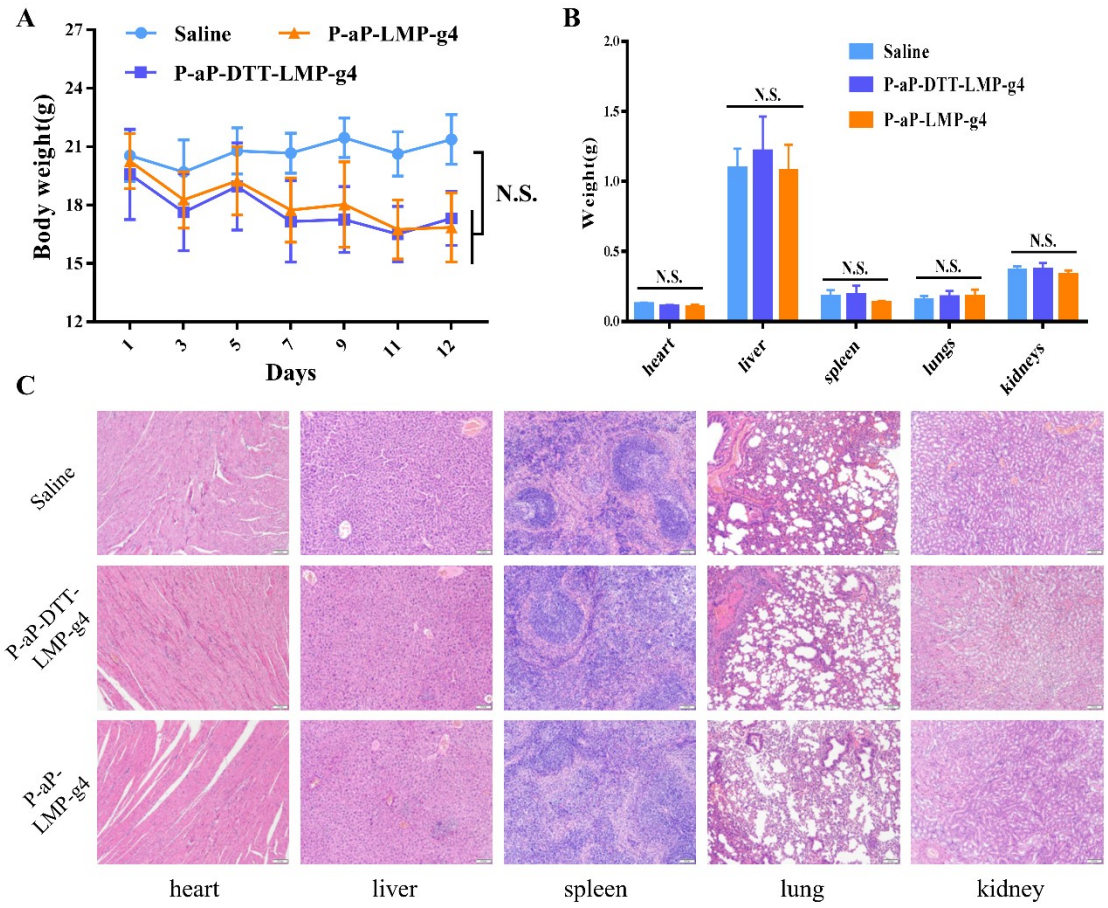

Fig. S19. *In vivo* C666-1 xenograft tumor growth treated by P-aP-DTT/P-aP-LMP-g4 polyplex NPs. (A) Body weight. (B) The weight of main organs (heart, liver, spleen, lungs and kidneys). (C) Representative images of H&E staining of major organs (heart, liver, spleen, lungs and kidneys) of mice that were treated with P-aP-DTT/P-aP-LMP-g4 polyplex NPs and Saline (Scale bar, 100 μm). One-way ANOVA was for statistical analysis, N.S.: no significant difference.

Table S1. sgRNA sequence and corresponding PAM sequences

Name	Forward	Reverse
MUC2P	CAGCACGTCATCCTGAAGGT	AAGACACCCTGGAGACACCT
LMP-g4	ACTCTGCTCTCAAAACCTAGGC	CATGTCCTCCTTTCCCCTTGTT
OND	TTGAGTTGTCATATGTTAATAACGGT	ACCGTTATTAACATATGACAACTCAA

Name	sgRNA sequence (5'-3')	PAM sequence (5'-3')	Target gene
LMP1-1	CAGCCCTTCACACACCACAC	AGG	LMP1
LMP1-2	AGTCAGGCAAGCCTATGACA	TGG	LMP1
LMP1-3	TGGTGTTCATCAGTGTGTCG	TGG	LMP1
LMP1-4	GATGAGTAGGAGGGTGACTG	GGG	LMP1
LMP-g4	GACCCGCCTTCGATGACAGA	CGG	LMP1
MUC2-1	ATGTCTGCATCGTGACCCTT	GGG	MUC2
MUC2-2	AGGTCACTGTCGCAGCACTG	GGG	MUC2
MUC2P	GGGGCACCTAGAGTGACCAG	AGG	MUC2

Table S2. Sequences of Sanger sequencing

The sequences of LMP-g4:

GGTGCGGCGGCAGCCAATCAGAGCGGCGCGCTCCGAAAGTTTCCTTTTAT
GGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAAGCGAAGCGCGCGGCGG
GCGGGAGTCGCTGCGCGCTGCCTTCGCCCCGTGCCCCGCTCCGCCGCCGC
CTCGCGCCGCCCGCCCCGGCTCTGACTGACCGCGTTACTCCCACAGGTGA
GCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCTGAGCAAGAGG
TAAGGGTTTAAGGGATGGTTGGTTGGTGGGGTATTAATGTTTAATTACCTG
GAGCACCTGCCTGAAATCACTTTTTTTCAGGTTGGACCGGTGCCACCATGG
ACTATAAGGACCACGACGGAGACTACAAGGATCATGATATTGATTACAAA
GACGATGACGATAAGATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCC

ACGGAGTCCCAGCAGCCGACAAGAAGTACAGCATCGGCCTGGACATCGG
CACCAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTACAAGGTGCCCA
GCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAGCATCAAGAA
GAACCTGATCGGAGCCCTGCTGTTCGACAGCGGCGAAACAGCCGAGGCCA
CCCGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAGAACCG
GATCTGCTATCTGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGGACG
ACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGGAAGAGGATAAG
AAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTA
CCACGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACA
GCACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATG
ATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAA
CAGCGACGTGGACAAGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGC
TGTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGGACGCCAAGGCCATC
CTGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATCGCCCA
GCTGCCCGGCGAGAAGAAGAATGGCCTGTTCGGAAACCTGATTGCCCTGA
GCCTGGGCCTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGAT
GCCAAACTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACC
TGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAG
AACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGA
GATCACCAAGGCCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGC
ACCACCAGGACCTGACCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCT
GAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACGCCG
GCTACATTGACGGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAG
CCCATCCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGA
ACAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCAT
CCCCCACCAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGG
AAGATTTTTACCCATTCCTGAAGGACAACCGGGAAAAGATCGAGAAGATC
CTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACAG
CAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGG
AACTTCGAGGAAGTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATCGA
GCGGATGACCAACTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCA
AGCACAGCCTGCTGTACGAGTACTTCACCGTGTATAACGAGCTGACCAAA
GTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTGAGCGGCG
AGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGT
GACCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCG
ACTCCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGC
ACATACCACGATCTGCTGAAAATTATCAAGGACAAGGACTTCCTGGACAA
TGAGGAAAACGAGGACATTCTGGAAGATATCGTGCTGACCCTGACACTGT
TTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCCCACCT
GTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGC
TGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGT
CCGGCAAGACAATCCTGGATTTCCTGAAGTCCGACGGCTTCGCCAACAGA
AACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACAT
CCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTG

CCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTG
AAGGTGGTGGACGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGAGA
ACATCGTGATCGAAATGGCCAGAGAGAACCAGACCACCCAGAAGGGACA
GAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAGAGGGCATCAAAGA
GCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTG
CAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCGGGATATGTA
CGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTACGATGTGGACC
ATATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAAGGTG
CTGACCAGAAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCCTCCG
AAGAGGTCGTGAAGAAGATGAAGAACTACTGGCGGCAGCTGCTGAACGC
CAAGCTGATTACCCAGAGAAAGTTCGACAATCTGACCAAGGCCGAGAGA
GGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAGCTGGT
GGAAACCCGGCAGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGG
ATGAACACTAAGTACGACGAGAATGACAAGCTGATCCGGGAAGTGAAAG
TGATCACCCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGT
TTTACAAAGTGCGCGAGATCAACAACTACCACCACGCCCACGACGCCTAC
CTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAGTACCCTAAGCTGGA
AAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGAAGATG
ATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTT
CTACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACG
GCGAGATCCGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACCGGGGA
GATCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGCGGAAAGTGCTGA
GCATGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTGCAGACAGGCGG
CTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGATAAGCTGATCG
CCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCC
CACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGT
CCAAGAAACTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCATGGA
AAGAAGCAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCT
ACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAGTACTCCCTG
TTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACT
GCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGT
ACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATGAG
CAGAAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATCAT
CGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATC
TGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGA
GAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGC
CCCTGCCGCCTTCAAGTACTTTGACACCACCATCGACCGGAAGAGGTACA
CCAGCACCAAAGAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACC
GGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGGAGGCGACAAAA
GGCCGGCGGCCACGAAAAAGGCCGGCCAGGCAAAAAAGAAAAAGTAAG
AATTCCTAGAGCTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCC
ATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCAC
TCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAG
TAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGG

AGGATTGGGAAGAGAATAGCAGGCATGCTGGGGAGCGTACAAGTAAGAA
TTCCGATCATATTCAATAACCCTTAATATAACTTCGTATAATGTATGCTAT
ACGAAGTTATTAGGTCTGAAGAGGAGTTTACGTCCAGCCAAGCTTAGGAT
CTCGACCTCGAAATTCTACCGGGTAGGGGAGGCGCTTTTCCCAAGGCAGT
CTGGAGCATGCGCTTTAGCAGCCCCGCTGGGCACTTGGCGCTACACAAGT
GGCCTCTGGCCTCGCACACATTCCACATCCACCGGTAGGCGCCAACCGGC
TCCGTTCTTTGGTGGCCCCTTCGCGCCACCTTCTACTCCTCCCCTAGTCAGG
AAGTTCCCCCCCGCCCCGCAGCTCGCGTCGTGCAGGACGTGACAAATGGA
AGTAGCACGTCTCACTAGTCTCGTGCAGATGGACAGCACCGCTGAGCAAT
GGAAGCGGGTAGGCCTTTGGGGCAGCGGCCAATAGCAGCTTTGCTCCTTC
GCTTTCTGGGCTCAGAGGCTGGGAAGGGGTGGGTCCGGGGGCGGGCTCAG
GGGCGGGCTCAGGGGCGGGGCGGGCGCCCGAAGGTCCTCCGGAGGCCCG
GCATTCTGCACGCTTCAAAAGCGCACGTCTGCCGCGCTGTTCTCCTCTTCC
TCATCTCCGGGCCTTTCGACCTGCATCCATCTAGATCTCGAGCAGCTGAAG
CTTACCATGACCGAGTACAAGCCCACGGTGCGCCTCGCCACCCGCGACGA
CGTCCCCAGGGCCGTACGCACCCTCGCCGCCGCGTTCGCCGACTACCCCG
CCACGCGCCACACCGTCGATCCGGACCGCCACATCGAGCGGGTCACCGAG
CTGCAAGAACTCTTCCTCACGCGCGTCGGGCTCGACATCGGCAAGGTGTG
GGTCGCGGACGACGGCGCCGCGGTGGCGGTCTGGACCACGCCGGAGAGC
GTCGAAGCGGGGGCGGTGTTCGCCGAGATCGGCCCGCGCATGGCCGAGTT
GAGCGGTTCCCGGCTGGCCGCGCAGCAACAGATGGAAGGCCTCCTGGCGC
CGCACCGGCCCAAGGAGCCCGCGTGGTTCCTGGCCACCGTCGGCGTCTCG
CCCGACCACCAGGGCAAGGGTCTGGGCAGCGCCGTCGTGCTCCCCGGAGT
GGAGGCGGCCGAGCGCGCCGGGGTGCCCGCCTTCCTGGAGACCTCCGCGC
CCCACAACCTCCCCTTCTACGAGCGGCTCGGCTTCACCGTCACCGCCGACG
TCGAGGTGCCCGAAGGACCGCGCACCTGGTGCATGACCCGCAAGCCCGGT
GCCTGACGCCCGCCCCACGACCCGCAGCGCCCGACCGAAAGGAGCGCAC
GACCCCATGCATCGATGATATCAGATCCCCGGGATGCAGAAATTGATGAT
CTATTAAACAATAAAGATGTCCACTAAAATGGAAGTTTTTCCTGTCATACT
TTGTTAAGAAGGGTGAGAACAGAGTACCTACATTTTGAATGGAAGGATTG
GAGCGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCA
CACCGCATACGTCAAAGCAACCATAGTACGCGCCCTGTAGCGGCGCATTA
AGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAG
CGCCTTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTC
GCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGA
TTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTTGGGTGATGGT
TCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTG
GAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTC
AACTCTATCTCGGGCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGG
TCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTA
ACAAAATATTAACGTTTACAATTTTATGGTGCACTCTCAGTACAATCTGCT
CTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGAC
GCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGT
GACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGA

AACGCGCGAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAAT
GTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAAT
GTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTAT
CCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAG
GAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGC
GGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAA
AGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATC
TCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCA
ATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATT
GACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGA
CTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGA
CAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCG
GCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTT
TTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGG
AGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTA
GCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCT
AGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAG
GACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAAT
CTGGAGCCGGTGAGCGTGGAAGCCGCGGTATCATTGCAGCACTGGGGCCA
GATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGC
AACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGA
TTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTG
ATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTG
ATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGT
CAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGC
GCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTT
TGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTC
AGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGG
CCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAAT
CCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTT
GGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACG
GGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACT
GAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGG
AGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGC
GCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTC
GGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGG
GGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCT
GGCCTTTTGCTGGCCTTTTGCTCACATGTGAGGGCCTATTTCCCATGATTC
CTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTGGAATTA
ATTTGACTGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAG
TAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAATGGAC
TATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATA
TCTTGTGGAAAGGACGAAACACCGACCCGCCTTCGATGACAGAGTTTTAG
AGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAA

AGTGGCACCGAGTCGGTGCTTTTTCTAGCGCGTGCGCCAATTCTGCAGAC
AAATGGCTCTAGAGGTACCCGTTACATAACTTACGGTAAATGGCCCGCCT
GGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAGTAACGCCAAT
AGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCC
ACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGAC
GTCAATGACGGTAAATGGCCCGCCTGGCATTGTGCCCAGTACATGACCTT
ATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTAC
CATGGTCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCC
CTCCCCACCCCCAATTTTGTATTTATTTATTTTTTAATTATTTTGTGCAGCG
ATGGGGGCGGGGGGGGGGGGGGGGCGGGGCGAGGGGCGGGGCGGGGCG
AGGCGGAG

The sequences of MUC2P:

TCTTCTCTCATCCGCCAAAACAGCCAAGCTGGAGACCGTTTGACATTACCC

TGTTATCCCTAGATACATTACCCTGTTATCCCAGATGACATACCCTGTTAT

CCCTAGATGACATTACCCTGTTATCCCAGATGACATTACCCTGTTATCCCT

AGATACATTACCCTGTTATCCCAGATGACATACCCTGTTATCCCTAGATGA

CATTACCCTGTTATCCCAGATGACATTACCCTGTTATCCCTAGATACATTA
CCCTGTTATCCCAGATGACATACCCTGTTATCCCTAGATGACATTACCCTG
TTATCCCAGATGACATTACCCTGTTATCCCTAGATACATTACCCTGTTATC
CCAGATGACATACCCTGTTATCCCTAGATGACATTACCCTGTTATCCCAGA
TGACATTACCCTGTTATCCCTAGATACATTACCCTGTTATCCCAGATGACA
TACCCTGTTATCCCTAGATGACATTACCCTGTTATCCCAGATGACATTACC
CTGTTATCCCTAGATACATTACCCTGTTATCCCAGATGACATACCCTGTTA
TCCCTAGATGACATTACCCTGTTATCCCAGATGACATTACCCTGTTATCCC
TAGATACATTACCCTGTTATCCCAGATGACATACCCTGTTATCCCTAGATG
ACATTACCCTGTTATCCCAGATGACATTACCCTGTTATCCCTAGATACATT
ACCCTGTTATCCCAGATGACATACCCTGTTATCCCTAGATGACATTACCCT
GTTATCCCAGATAAACTCAATGATGATGATGATGATGGTCGAGACTCAGC
GGCCGCGGTGCCAGGGCGTGCCCTTGGGCTCCCCGGGCGCGGAATTCCGA
GGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTGT
TAGAGAGATAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGTAC
AAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAA
AATTATGTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATT
TCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACcGAACAAA
GCATTGGTGGTTCAGTGGTAGAATTCTCGCCTGCCACGCGGGAGGCCCGG
GTTCGATTCCCGGCCAATGCAggggcacctagagtgaccagGTTTTAGAGCTAGAAA
TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACC
GAGTCGGTGCTTTTTTTTTTGTTTTAGAGCTAGAAATAGCAAGTTAAAATA
AGGCTAGTCCGTTTTTAGCGCGTGCGCCAATTCTGCAGACAAATGGCTCTA
GAGGTACCCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCC
CAACGACCCCCGCCCATTGACGTCAATAGTAACGCCAATAGGGACTTTCC
ATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTA
CATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGG
TAAATGGCCCGCCTGGCATTGTGCCCAGTACATGACCTTATGGGACTTTCC
TACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTCGAGGT
GAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCCACCCCC

AATTTTGTATTTATTTATTTTTTAATTATTTTGTGCAGCGATGGGGGCGGGG
GGGGGGGGGGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGG
GGCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGAGCGGCGCGC
TCCGAAAGTTTCCTTTTATGGCGAGGCGGCGGCGGCGGCGGCCCTATAAA
AAGCGAAGCGCGCGGCGGGCGGGAGTCGCTGCGACGCTGCCTTCGCCCCG
TGCCCCGCTCCGCCGCCGCCTCGCGCCGCCCGCCCCGGCTCTGACTGACCG
CGTTACTCCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGT
AATTAGCTGAGCAAGAGGTAAGGGTTTAAGGGATGGTTGGTTGGTGGGGT
ATTAATGTTTAATTACCTGGAGCACCTGCCTGAAATCACTTTTTTTCAGGT
TGGACCGGTGCCACCGCCACCATGGCGCCAGCGGCAAAGCGCGTAAAATT
GGACCCAGCAGCCAAACGAGTAAAACTCGATCCGGCTGCTAAACGAGTTA
AGCTGGACGACAAGAAGTACAGCATCGGCCTGGACATCGGCACCAACTCT
GTGGGCTGGGCCGTGATCACCGACGAGTACAAGGTGCCCAGCAAGAAATT
CAAGGTGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATC
GGAGCCCTGCTGTTCGACAGCGGCGAAACAGCCGAGGCCACCCGGCTGA
AGAGAACCGCCAGAAGAAGATACACCAGACGGAAGAACCGGATCTGCTA
TCTGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCT
TCCACAGACTGGAAGAGTCCTTCCTGGTGGAAGAGGATAAGAAGCACGA
GCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGA
AGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAGCACCGAC
AAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTT
CCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACG
TGGACAAGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAG
GAAAACCCCATCAACGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGC
CAGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATCGCCCAGCTGCCC
GGCGAGAAGAAGAATGGCCTGTTCGGAAACCTGATTGCCCTGAGCCTGGG
CCTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAAAC
TGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCC
CAGATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGTC
CGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAGATCACCA

AGGCCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAG
GACCTGACCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTA
CAAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATTG
ACGGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCCATCCTG
GAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGG
ACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAG
ATCCACCTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTA
CCCATTCCTGAAGGACAACCGGGAAAAGATCGAGAAGATCCTGACCTTCC
GCATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACAGCAGATTCGCC
TGGATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGGAACTTCGAGG
AAGTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATCGAGCGGATGACC
AACTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCT
GCTGTACGAGTACTTCACCGTGTATAACGAGCTGACCAAAGTGAAATACG
TGACCGAGGGAATGAGAAAGCCCGCCTTCCTGAGCGGCGAGCAGAAAAA
GGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGC
AGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAA
ATCTCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGA
TCTGCTGAAAATTATCAAGGACAAGGACTTCCTGGACAATGAGGAAAACG
AGGACATTCTGGAAGATATCGTGCTGACCCTGACACTGTTTGAGGACAGA
GAGATGATCGAGGAACGGCTGAAAACCTATGCCCACCTGTTCGACGACAA
AGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGCTG
AGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAA
TCCTGGATTTCCTGAAGTCCGACGGCTTCGCCAACAGAAACTTCATGCAG
CTGATCCACGACGACAGCCTGACCTTTAAAGAGGACATCCAGAAAGCCCA
GGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAATCTGGCCG
GCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGA
CGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATC
GAAATGGCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCC
GCGAGAGAATGAAGCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCA
GATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAACGAGAAG

CTGTACCTGTACTACCTGCAGAATGGGCGGGATATGTACGTGGACCAGGA
ACTGGACATCAACCGGCTGTCCGACTACGATGTGGACCATATCGTGCCTC
AGAGCTTTCTGAAGGACGACTCCATCGACAACAAGGTGCTGACCAGAAGC
GACAAGAACCGGGGCAAGAGCGACAACGTGCCCTCCGAAGAGGTCGTGA
AGAAGATGAAGAACTACTGGCGGCAGCTGCTGAACGCCAAGCTGATTACC
CAGAGAAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCG
AACTGGATAAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCA
GATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAACACTAAGT
ACGACGAGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCACCCTGAA
GTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCG
CGAGATCAACAACTACCACCACGCCCACGACGCCTACCTGAACGCCGTCG
TGGGAACCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTG
TACGGCGACTACAAGGTGTACGACGTGCGGAAGATGATCGCCAAGAGCG
AGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGCAACATC
ATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAA
GCGGCCTCTGATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGAT
AAGGGCCGGGATTTTGCCACCGTGCGGAAAGTGCTGAGCATGCCCCAAGT
GAATATCGTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAGAG
TCTATCCTGCCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGG
ACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTAT
TCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGA
AGAGTGTGAAAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTT
CGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTG
AAAAAGGACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTCGAGCTGGA
AAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACTGCAGAAGGGA
AACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCCAG
CCACTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACAG
CTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATCATCGAGCAGAT
CAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAAG
TGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCC

GAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCC
TTCAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAA
AGAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTACG
AGACACGGATCGACCTGTCTCAGCTGGGAGGCGACCCCAAGAAGAAGAG
GAAAGTCTCGAGCGACTACAAAGACCATGACGGTGATTATAAAGATCATG
ACATCGATTACAAGGATGACGATGACAAGGCTGCAGGATGAGAATTCCTA
GAGCTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTG
TTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTG
TCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTC
ATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTG
GGAAGAGAATAGCAGGCATGCTGGGGACATGGGGGCCCGCCCCAACTGG
GGTAACCTTTGAGTTCTCTCAGTTGGGGGTAATCAGCATCATGATGTGGTA
CCACATCATGATGCTGATTATAAGAATGCGGCCGCCACACTCTAGTGGAT
CTCGAGTTAATAATTCAGAAGAACTCGTCAAGAAGGCGATAGAAGGCGAT
GCGCTGCGAATCGGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGTCA
GCCCATTCGCCGCCAAGCTCTTCAGCAATATCACGGGTAGCCAACGCTAT
GTCCTGATAGCGGTCCGCCACACCCAGCCGGCCACAGTCGATGAATCCAG
AAAAGCGGCCATTTTCCACCATGATATTCGGCAAGCAGGCATCGCCATGG
GTCACGACGAGATCCTCGCCGTCGGGCATGCTCGCCTTGAGCCTGGCGAA
CAGTTCGGCTGGCGCGAGCCCCTGATGCTCTTCGTCCAGATCATCCTGATC
GACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCG
CTTGGTGGTCGAATGGGCAGGTAGCCGGATCAAGCGTATGCAGCCGCCGC
ATTGCATCAGCCATGATGGATACTTTCTCGGCAGGAGCAAGGTGTAGATG
ACATGGAGATCCTGCCCCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCC
CGCTTCAGTGACAACGTCGAGCACAGCTGCGCAAGGAACGCCCGTCGTGG
CCAGCCACGATAGCCGCGCTGCCTCGTCTTGCAGTTCATTCAGGGCACCG
GACAGGTCGGTCTTGACAAAAAGAACCGGGCGCCCCTGCGCTGACAGCCG
GAACACGGCGGCATCAGAGCAGCCGATTGTCTGTTGTGCCCAGTCATAGC
CGAATAGCCTCTCCACCCAAGCGGCCGGAGAACCTGCGTGCAATCCATCT
TGTTCAATCATGCGAAACGATCCTCATCCTGTCTCTTGATCAGAGCTTGAT

CCCCTGCGCCATCAGATCCTTGGCGGCGAGAAAGCCATCCAGTTTACTTTG
CAGGGCTTCCCAACCTTACCAGAGGGCGCCCCAGCTGGCAATTCCGGTTC
GCTTGCTGTCCATAAAACCGCCCAGTCTAGCTATCGCCATGTAAGCCCACT
GCAAGCTACCTGCTTTCTCTTTGCGCTTGCGTTTTCCCTTGTCCAGATAGCC
CAGTAGCTGACATTCATCCGGGGTCAGCACCGTTTCTGCGGACTGGCTTTC
TACGTGCTCGAGGGGGGCCAAACGGTCTCCAGCTTGGCTGTTTTGGCGGA
TGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGAACGCAGAAGCGG
TCTGATAAAACAGAATTTGCCTGGCGGCAGTAGCGCGGTGGTCCCACCTG
ACCCCATGCCGAACTCAGAAGTGAAACGCCGTAGCGCCGATGGTAGTGTG
GGGTCTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAACGA
AAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTG
AACGCTCTCCTGAGTAGGACAAATCCGCCGGGAGCGGATTTGAACGTTGC
GAAGCAACGGCCCGGAGGGTGGCGGGCAGGACGCCCGCCATAAACTGCC
AGGCATCAAATTAAGCAGAAGGCCATCCTGACGGATGGCCTTTTTGCGTT
TCTACAAACTCTTTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCT
CATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCC
CGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAAT
CTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGC
CGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGA
GCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCA
CTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTT
ACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACT
CAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGG
TTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGAT
ACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAA
GGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACG
AGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTT
TCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCG
GAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCT
TTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGT

GGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCC
GAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCT
GATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATG
GTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATAC
ACTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACACCCGCC
AACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTA
CAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCAC
CGTCATCACCGAAACGCGCGAGGCAGCAGATCAATTCGCGCGCGAAGGC
GAAGCGGCATGCATAATGTGCCTGTCAAATGGACGAAGCAGGGATTCTGC
AAACCCTATGCTACTCCGTCAAGCCGTCAATTGTCTGATTCGTTACCAATT
ATGACAACTTGACGGCTACATCATTCACTTTTTCTTCACAACCGGCACGGA
ACTCGCTCGGGCTGGCCCCGGTGCATTTTTTAAATACCCGCGAGAAATAG
AGTTGATCGTCAAAACCAACATTGCGACCGACGGTGGCGATAGGCATCCG
GGTGGTGCTCAAAAGCAGCTTCGCCTGGCTGATACGTTGGTCCTCGCGCC
AGCTTAAGACGCTAATCCCTAACTGCTGGCGGAAAAGATGTGACAGACGC
GACGGCGACAAGCAAACATGCTGTGCGACGCTGGCGAT