Electronic Supplementary Material (ESI) for Chemical Communications


**Supporting Information**

# Screening mutant libraries of T7 RNA polymerase for candidates with increased acceptance of 2'-modified nucleotides

**Vanessa Siegmund, Tobias Santner , Ronald Micura and Andreas Marx**

## Materials.

2'-O-Methyluridine was purchased from Molekula. The *E. coli* codon optimized gene for T7 RNA polymerase wildtype was ordered from Geneart. *E. coli* BL21 cells were from Amersham Biosciences. Oligonucleotides were obtained from Metabion and Biomers. Primers with randomized positions were obtained from Purimex. Regular ribonucleotides were purchased from Fermentas and Roche. Phusion High-Fidelity DNA polymerase was from Finnzymes. *Taq* DNA Polymerase was from Fermentas. Restriction endonucleases and T4 DNA Ligase were obtained from Fermentas or New England Biolabs. Pipetting steps during screening were performed with a Microlab Star automated liquid handling device from Hamilton. SYBR Green II RNA gel stain was obtained from Sigma Aldrich. Fluorescence readout was performed using a Polarstar Optima microplate reader from BMG Labtechnologies GmbH. [α-$^{32}$P]GTP was purchased from Hartmann Analytics. Phosphorimaging was done on a Molecular Imager FX (BioRad). Sequencing was done at GATC Biotech.

## Synthesis of 2'-O-Methyl-2'-deoxyuridine 5'-triphosphate 2'-Methylseleno-2'-deoxyuridine 5'-triphosphate

2'-O-Methyl-2'-deoxyuridine 5'-triphosphate was synthesized starting from the commercial available nucleoside according to a method reported by Kovács and Ötvös.[1] 2'-Methylseleno-2'-deoxyuridine 5'-triphosphate was synthesized as described before.[2]

## Construction of mutant library by saturated mutagenesis.

The T7 RNA polymerase wildtype gene was cloned into the expression vector pGDR11[3] which is a derivative of the pQE31 vector carrying the *lacIq* gene as described before.[2] The

library was constructed by randomizing the T7 RNA polymerase wildtype gene at positions Y639 and H784. DNA fragments containing each of the randomized regions were made by PCR using primers in which the codons encoding amino acids Y639 and H784 were randomized by NNK, where N stands for all four bases and K for T and G. PCR reactions (1× Phusion HF buffer, 500 nM of each primer, 200 µM dNTPs, 100 pM template, 0.02 units/µL Phusion DNA polymerase) were performed using Phusion DNA polymerase. The PCR products containing the randomized regions were purified by agarose gel electrophoresis and assembled by overlapping PCR (1 × Phusion HF buffer, 1 nM of each DNA fragment, 500 nM flanking primers, 200 µM dNTPs, 0.01 units/µL Phusion DNA polymerase). The flanking primers were designed to introduce *NdeI* and *HindIII* restriction sites. The assembled fragment was cloned into pGDR11-NdeI, a vector where the outer *NdeI* restriction site was removed by introducing a silent mutation using site directed mutagenesis. Electrocompetent *E. coli* BL21 cells were transformed with the resulting library and cells grown on LB-agar plates supplemented with 100 µg/mL carbenicillin overnight. 3200 single colonies were picked and separately grown overnight at 25 °C in 384 deepwell plates containing 150 µL/well LB medium (100 µg/mL carbenicillin). 100 µL/well glycerol was added to a final concentration of 20% (v/v) for storage at -80 °C.

**Construction of randomized mutant library by error-prone PCR.**

Mutant 'RGVG', E593G, V685A (Y639V, H784A, E593G, V685A) was generated as described before.[2] Random mutagenesis was done by error-prone PCR using the gene of T7 RNA polymerase mutant 'RGVG', E593G, V685A as template. The gene was amplified by PCR using 0.05 units/µL of *Taq* DNA polymerase, 1×reaction buffer (10 mM Tris-HCl pH 8.8, 50 mM KCl, 0.08% (v/v) Nonidet P40, 7 mM $MgCl_2$, 0.1 mM $MnCl_2$) 200 nM of each primer (FP: 5'-GGATCCGCATGCAATTGAAGGTCGCATGAATACC-3'; RP: 5'-CAGGAGTCCAAGCTCAGCTAATTAAGCTTTTA-3'), 1 mM dCTP/dTTP, 0.2 mM dATP/dGTP and 20 pM template in a total reaction volume of 50 µL according to the following program: 95 °C for 2 min followed by 15 cycles of 95 °C for 1 min and 72 °C for 4 min and a final extension step at 72 °C for 5 min. The obtained PCR product was gel purified by electrophoresis using 0.8% TAE agarose and cloned into the pGDR11 vector using *SphI* and *HindIII* restriction sites. Electrocompetent *E. coli* BL21 cells were transformed with the resulting library and cells grown on LB-agar plates supplemented with 100 µg/mL carbenicillin overnight. 1600 single colonies were picked and separately grown overnight at

25 °C in 384 deepwell plates containing 150 µL/well LB medium (100µg/mL carbenicillin). For storage at -80 °C, 100 µL/well glycerol was added to a final concentration of 20% (v/v).

**Library expression and screening.**

Expression and screening of the mutant libraries was established according to previously described procedures.[4] T7 RNA polymerase variants were expressed in parallel in 96-deepwell plates for 5 h at 37 °C. For this purpose, cells were grown in 1 mL LB-medium (100µg/mL carbenicillin) to an $OD_{600}$ of 0.6-0.7 and expression induced by the addition of IPTG to a final concentration of 1 mM. Cells were harvested by centrifugation (4.500 × g for 20 min, 4 °C) and lysis conducted by the addition of 300 µL/well lysis buffer (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 0.1% (v/v) Triton X-100, 0.1 mg/ml lysozyme) for 20 min at 25°C. Afterwards, lysates were diluted by the addition of 1.2 mL dilution buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 5 mM DTT) and cleared by centrifugation (4.500 × g for 45 min, 4 °C). Supernatants containing RNA polymerase mutants were directly used for screening. As control reactions, cells harboring the T7 RNA polymerase wildtype plasmid and cells containing the pGDR11 vector were present on all plates. For screening of the randomized mutant library, cells containing the plasmid of mutant 'RGVG', E593G, V685A (Y639V, H784A, E593G, V685A) were also used as control. Transcription reactions for screening were performed in an overall volume of 20 µL in 384 well plates. The reaction mixtures for library screening contained 5 µL of supernatant and 1× transcription buffer (40 mM Tris-HCl pH 8.0, 2 mM spermidine, 10 mM 1,4-dithiothreitol (DTT), 6 mM $MgCl_2$), 0.2 µM dsDNA template (5'-CGTTGGTCCTGAAGGAGGATAGGTTGATTTTCTGCAGTCCTCTGTCCAC GGCGGCGCGTGCTGCGCGACGGCACAGCTGACGGTCTCCCTATAGTGAGTCGTAT TAGGAGGC-3') and 1 mM NTPs. After an incubation time of 1.5 h at 37 °C, DNA templates were digested by the addition of 2 units of DNaseI in 5 µL buffer (50 mM Tris-HCl pH 7.5, 10 mM $CaCl_2$). Digestion was performed for 30 min at 37 °C and the reactions stopped by the addition of 45 µL/well SYBR Green II containing stop solution (50 mM Tris-HCl pH 8.0, 25 mM EDTA, 4 × SYBR Green). After 5 min of incubation, fluorescence intensities were determined by fluorescence readout using a microplate reader with excitation at 485 nm and emission at 520 nm.

**Purification of selected mutants.**

Mutant T7 RNA polymerases were expressed and purified as described before via the N-terminal polyhistidine (6 × His) tag by nickel ion affinity chromatography. Enzymes were >95% pure as determined by colloidal Coomassie blue stained SDS-PAGE gel (12%). Protein concentrations were determined by Bradford protein assay and adjusted to the same level. Wildtype T7 RNA polymerase and mutant 'RGVG', E593G, V685A were expressed and purified simultaneously for control reactions.

**Incorporation assays.**

Transcription assays using regular NTPs and 2'-modified NTPs were performed for 1-2 h at 37 °C in 1× transcription buffer (40 mM Tris-HCl pH 8.0, 2 mM spermidine, 10 mM 1,4-dithiothreitol (DTT), 6 mM $MgCl_2$) with 0.2 μM double-stranded template DNA (5'-CGTTGGTCCTGAAGGAGGATAGGTTGATTTTCTGCAGTCCTCTGTCCACGGCGGC GCGTGCTGCGCGACGGCACAGCTGACGGTCTCCCTATAGTGAGTCGTATTAGGAG GC-3'), 1 mM of each NTP and 250 nM RNA polymerase in a final reaction volume of 20 μL. Transcripts (89nt: 5'-GGGAGACCGUCAGCUGUGCCGUCGCGCAGCACGCGCC GCCGUGGACAGAGGACUGCAGAAAAUCAACCUAUCCUCCUUCAGGACCAACG-3') were body labeled by the inclusion of 0.75% (v/v) 3000 Ci/mmol [α-$^{32}$P]GTP. Reactions were quenched by adding an equal volume of 90% (v/v) formamide, 50 mM EDTA, 0.01% (w/v) bromophenol blue and 0.01% (w/v) xylene cyanol and transcripts resolved by electrophoresis on 10% denaturing polyacrylamide gels. Analysis was done by phosphorimaging.

**Time course experiment.**

Transcription reactions were done as described above. Aliquots (2 μL) were withdrawn from the reaction mixture at the corresponding time points and added to an equal volume of 90% (v/v) formamide, 50 mM EDTA, 0.01% (w/v) bromophenol blue and 0.01% (w/v) xylene cyanol. Transcripts were resolved by electrophoresis on 10% denaturing polyacrylamide gels and visualized by phosphorimaging.

**A**

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 28595 | 3248 | 1348 | 1256 | 3268 | 922 | 24896 | 1773 | 218 | 2859 | 2569 | 3745 | 31804 | 895 | 1195 | 688 | -900 | 791 | 41209 | 611 | -585 | 949 | 205 | 2536 |
| B | 31020 | -46 | 709 | 313 | -288 | 1543 | 18910 | 1194 | 2038 | 3226 | 4087 | 668 | 29966 | 922 | -777 | -274 | 353 | -313 | 32700 | 80 | -297 | 348 | 2195 | 997 |
| C | 19024 | 396 | 1398 | 652 | 668 | 1231 | 22285 | 6704 | 836 | 856 | 1338 | 1202 | 26400 | 531 | -1328 | -2 | -1363 | -532 | 31893 | -556 | -316 | -535 | -422 | 392 |
| D | 23312 | -233 | 135 | -183 | 220 | -690 | 20075 | 484 | 1283 | 535 | 2295 | 1323 | 27750 | 307 | -355 | -291 | -404 | -371 | 27449 | 179 | 821 | -80 | 351 | 389 |
| E | 21619 | 834 | 593 | 364 | 541 | 43 | 23086 | 1696 | 475 | 260 | 20414 | -131 | 24242 | -820 | -1221 | -1184 | -1312 | -751 | 23700 | -1021 | -612 | -693 | -878 | -447 |
| F | 23723 | -1011 | -545 | -922 | -429 | -706 | 16454 | -639 | 1208 | 92 | 1563 | 1383 | 28367 | 388 | -369 | 917 | 135 | 350 | 30452 | -348 | -283 | 182 | 166 | 290 |
| G | 22156 | 1405 | 20694 | 206 | -4 | 895 | 22908 | 607 | 788 | 682 | 344 | -346 | 26129 | -1073 | -1525 | -882 | 21552 | -829 | 23688 | -529 | -754 | -547 | -571 | -647 |
| H | 22650 | -562 | -646 | -541 | -790 | -177 | 16924 | -277 | -71 | 240 | 1323 | 1574 | 18633 | 391 | 192 | 485 | 240 | 183 | 27750 | 420 | -9 | -316 | -113 | 649 |
| I | 860 | 445 | 262 | 740 | 709 | -167 | -152 | -252 | 178 | 693 | 739 | 395 | -761 | -1053 | -1214 | -839 | -1553 | -1188 | -1218 | -752 | -987 | -705 | -680 | -437 |
| J | -546 | 167 | -29 | -450 | 2 | -982 | -127 | -399 | 315 | 170 | 1420 | 1638 | -19 | 188 | 509 | 756 | 656 | 320 | 578 | 563 | 163 | 262 | 1047 | 137 |
| K | 1068 | 586 | 96 | 19 | 307 | 871 | 734 | -29 | 324 | -173 | 480 | 141 | 525 | -1474 | -1106 | -1198 | -1104 | -1647 | -1008 | 9664 | -846 | -1129 | -809 | 171 |
| L | 161 | -258 | 70 | -497 | -681 | -685 | 341 | 10264 | -60 | 345 | 1240 | 1288 | 6 | 874 | 495 | 421 | 245 | 175 | 42 | 563 | 420 | -234 | 426 | 963 |
| M | 1297 | 421 | 5161 | 154 | 476 | -246 | -475 | -122 | -650 | -328 | -262 | -328 | -851 | 307 | -854 | -1357 | -875 | -1365 | -1038 | -1475 | -1616 | -1009 | -1174 | 131 |
| N | 91 | 73 | -484 | 2074 | 190 | -429 | 178 | -378 | -169 | -80 | 731 | 529 | 270 | 423 | 1770 | 1366 | 1197 | 24028 | 50 | 245 | -354 | 625 | 367 | 348 |
| O | 603 | 862 | 315 | 477 | -398 | 101 | 191 | -367 | 256 | -557 | 1504 | 463 | 890 | 290 | -1012 | -87 | -1361 | -1300 | -1497 | -526 | -308 | -403 | -1449 | -205 |
| P | -314 | 297 | 1150 | -564 | 453 | -559 | -253 | -33 | 570 | -98 | -279 | 58 | 157 | 55 | 734 | 1878 | 517 | 1423 | 224 | 331 | 217 | 107 | -191 | 258 |

**B**

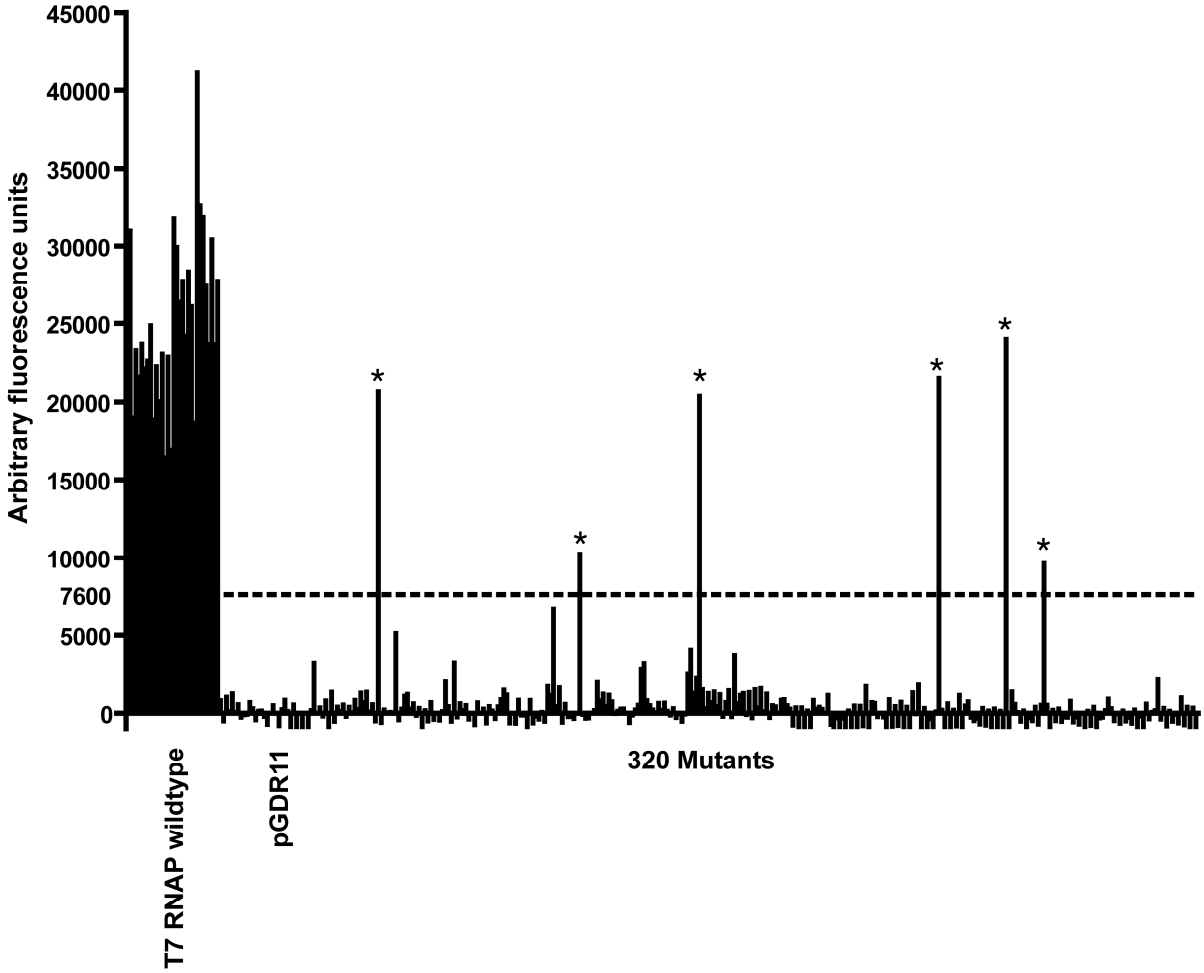
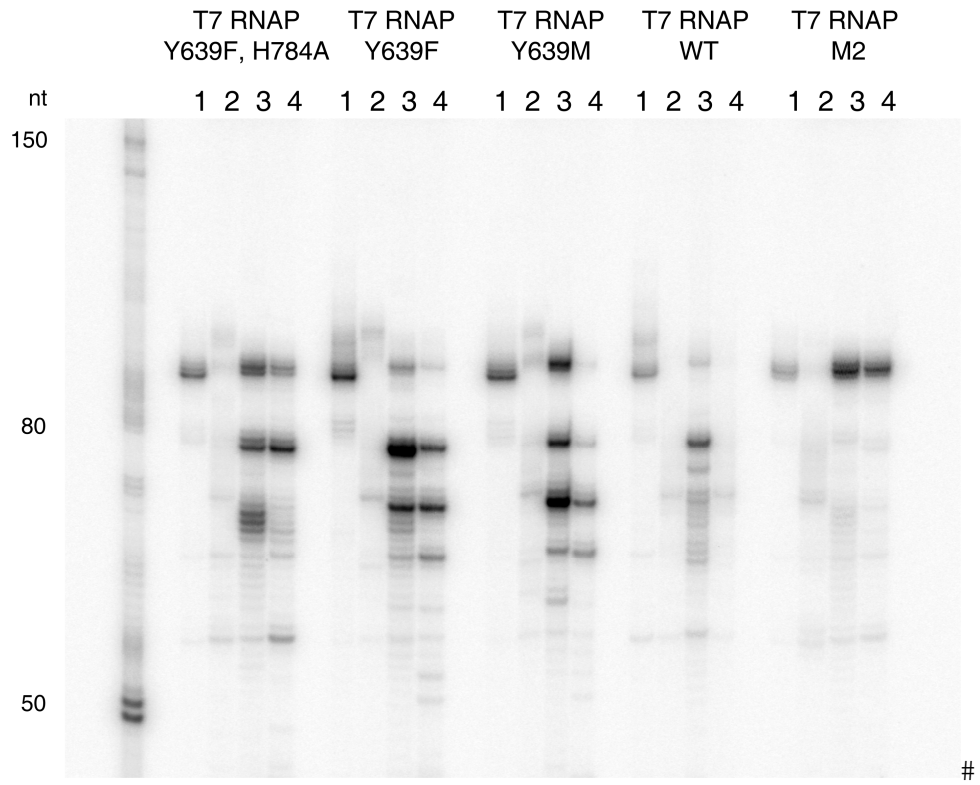



**Figure S1**: Fluorescence readings from a representative 384-well plate resulting from the screening of the library in which amino acid positions Y639 and H784 were randomized. Fluorescence intensities were determined by fluorescence readout using a microplate reader with excitation at 485 nm and emission at 520 nm. (A) Represented values show arbitrary fluorescence units of a 384-well plate that were calculated by subtracting the mean fluorescence of the assay blank (pGDR11) from all readings. Positive controls (T7 RNA polymerase wildtype) were present in columns 1, 7, 13 and 19 in rows A-H. Negative controls (pGDR11) were present in columns 1, 7, 13 and 19 in rows I-P. 320 mutants were present in the remaining wells of the 384-well plate. Positive hits that showed at least 30% activity relative to the mean value of the wildtype positive controls were selected and are highlighted in gray. (B) Bar chart depicting the results from the representative 384-well plate. Mutants labeled with an asterisk (*) showed at least 30% activity relative to the mean value of the wildtype positive controls and were selected.




#

**Figure S2**: Transcription assays with 2'-modified UTP using selected mutants from a library of T7 RNA polymerase where the positions Y639 and H784 were randomized. Activity of the mutants was compared to reactions with T7 RNA polymerase wildtype (WT) and mutant 'RGVG', E593G, V685A (M2). Transcription of an 89 nt long RNA sequence was done for 2 h at 37 °C. Transcripts were labeled by the inclusion of [$\alpha^{32}$P]GTP and analyzed by denaturing PAGE (10%). Lane 1: Transcription in the presence of ATP, CTP, GTP and UTP; Lane 2: Transcription in the presence of ATP, CTP and GTP, but no UTP; Lane 3: Transcription in the presence of ATP, CTP, GTP and 2'-OMe-UTP; Lane 4: Transcription in the presence of ATP, CTP, GTP and 2'-SeMe-UTP. Reactions with all four NTPs (lane 1) were diluted tenfold prior to loading for ease of analysis.

**A**

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 19727 | 772 | -990 | -898 | -1485 | -1462 | 31273 | -1829 | 3253 | -1389 | 866 | 2544 | 37018 | -287 | -125 | -1024 | -1357 | 1866 | 37980 | -180 | -1848 | 313 | 1243 | 2822 |
| B | 60654 | 418 | 1283 | 1567 | 1170 | 320 | 53132 | -52 | 232 | 1154 | 958 | 2812 | 32087 | -534 | 21 | -66 | -55 | -1316 | 41828 | -413 | -43 | -646 | 538 | 2293 |
| C | 9729 | -612 | 353 | 2314 | -2160 | 2293 | 52306 | -1230 | 2942 | -55 | 940 | 4790 | 27742 | -1568 | -653 | -754 | -409 | -1303 | 21799 | -67 | 2978 | 1365 | -386 | 1202 |
| D | 60654 | 2802 | 1153 | 1915 | 178 | 2444 | 60647 | 669 | -282 | 756 | -116 | 2721 | 26725 | -577 | -1473 | -1591 | -505 | -803 | 29685 | -1174 | 138 | -529 | 549 | 139 |
| E | 1621 | 1251 | 469 | -1928 | 532 | -1945 | 1124 | -1289 | -1532 | -1640 | 209 | 198 | 4383 | -277 | -504 | -373 | -127 | -256 | 5688 | 533 | 1063 | 599 | 1130 | 1559 |
| F | 8065 | 1931 | 1999 | 467 | 497 | 1205 | 7001 | 663 | -633 | -2 | -550 | 550 | 4912 | 752 | -797 | -2988 | 471 | -93 | 7783 | 669 | 508 | -357 | 2996 | 576 |
| G | 5772 | 1308 | -500 | -1504 | -626 | -679 | -588 | -1529 | -1605 | -1773 | -1626 | 944 | 4865 | -689 | -187 | -2020 | -367 | -1914 | 3244 | -716 | 270 | 1663 | 425 | 881 |
| H | 7085 | 2051 | 5897 | 1468 | 732 | 2944 | 8405 | 1491 | 480 | 1428 | 300 | 903 | 2449 | -410 | -1681 | -669 | -133 | 191 | 4876 | 181 | -1051 | -1113 | 118 | 1354 |
| I | 5874 | 249 | 1999 | -926 | 1158 | -1761 | 2918 | -1527 | -1540 | -1526 | -115 | 619 | 6813 | -673 | 606 | -876 | 1112 | 1026 | 5479 | 14 | 711 | 1555 | 484 | 1002 |
| J | 10286 | 2013 | 2608 | 1579 | 649 | 1244 | 9548 | 628 | 1167 | -207 | 155 | 1353 | 3584 | 975 | 1016 | 50 | 921 | 1150 | 7489 | 159 | -716 | 17 | -123 | 871 |
| K | 2993 | -60 | 1331 | 36 | 590 | -1220 | 677 | -1830 | -1274 | 460 | 274 | -1352 | 6077 | -399 | 195 | 539 | -89 | 490 | 5195 | -147 | 963 | 1018 | 857 | 2535 |
| L | 9590 | 2809 | 3061 | 1199 | 1042 | 1779 | 7250 | 286 | -607 | 143 | -623 | 316 | 4110 | -73 | -970 | 45 | -1023 | -375 | 7128 | -439 | 288 | -784 | -342 | 1216 |
| M | -68 | 3847 | -700 | -1289 | -710 | -585 | -421 | -2025 | -1301 | -1604 | -1009 | 343 | -893 | -1083 | -481 | -695 | -350 | -219 | -255 | 437 | 734 | -322 | 215 | 951 |
| N | 1209 | -222 | 369 | 2553 | 1028 | -345 | 1562 | 415 | 378 | 776 | -442 | -184 | -923 | -394 | 493 | 47 | 57 | -811 | -433 | -85 | -739 | -1055 | -669 | 361 |
| O | 476 | -1003 | 74 | 960 | 1889 | -548 | -905 | 716 | 773 | 1256 | 1478 | 576 | -465 | -1071 | -1023 | -1236 | -546 | -1345 | -248 | -952 | 215 | 533 | 352 | 712 |
| P | 619 | 1233 | -192 | 1469 | 2045 | 1532 | 583 | 852 | 1382 | 1132 | 122 | 2625 | -452 | -792 | 471 | 6495 | 932 | -174 | 607 | -65 | -358 | -63 | -488 | 950 |

**B**

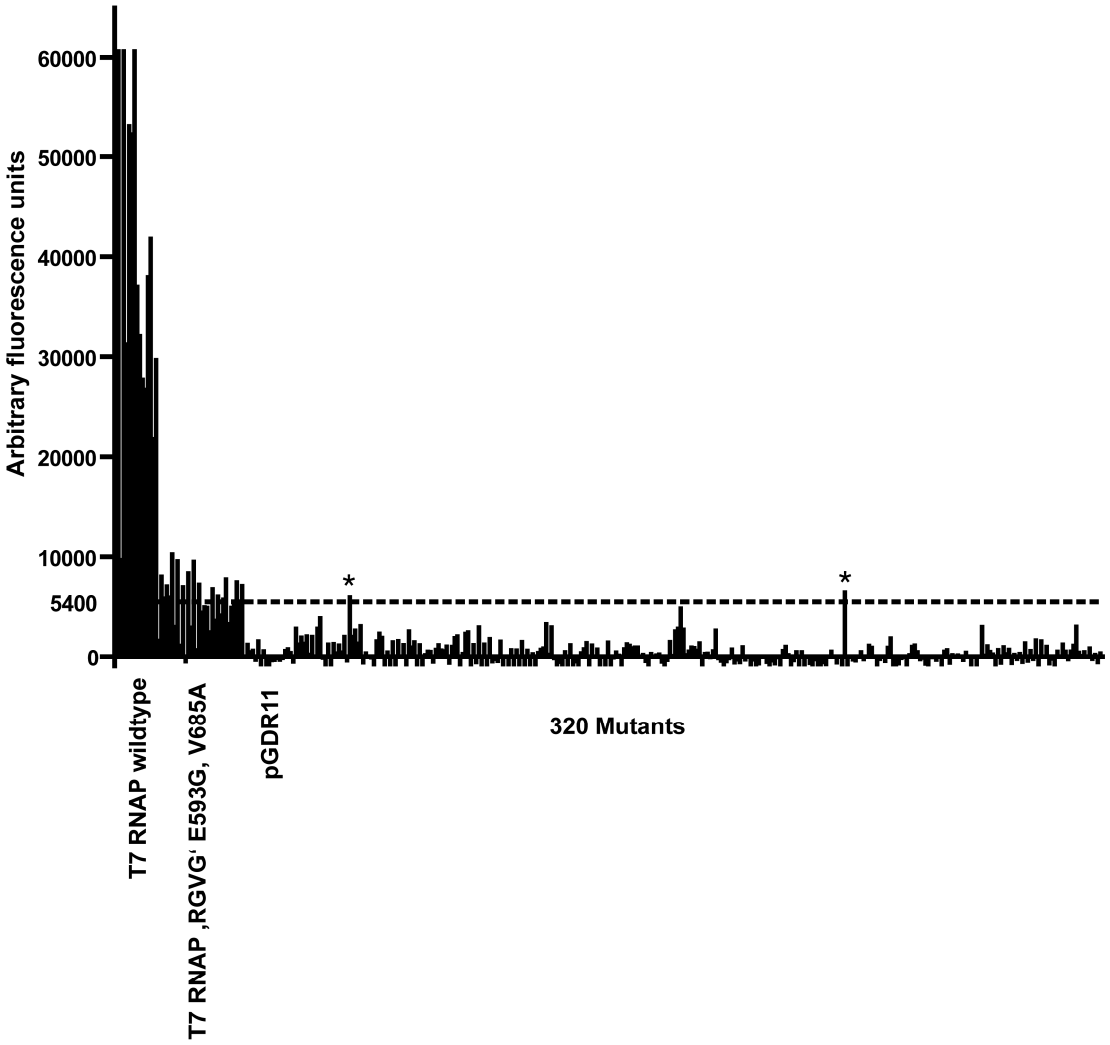



**Figure S3**: Fluorescence readings from a representative 384-well plate resulting from the screening of the second mutant library that was constructed by randomizing the gene of mutant 'RGVG' E593G, V685A by error-prone PCR. Fluorescence intensities were determined by fluorescence readout using a microplate reader with excitation at 485 nm and emission at 520 nm. (A) Represented values show arbitrary fluorescence units of a 384-well plate that were calculated by subtracting the mean fluorescence of the assay blank (pGDR11) from all readings. Controls with T7 RNA polymerase wildtype were present in columns 1, 7, 13 and 19 in rows A-D. Positive controls (mutant 'RGVG' E593G, V685A) were present in columns 1, 7, 13 and 19 in rows E-L. Negative controls (pGDR11) were present in columns 1, 7, 13 and 19 in rows I-P. 320 mutants were present in the remaining wells of the 384-well plate. Positive hits that showed at least the same level of activity as the mean value of the positive control with parental mutant 'RGVG' E593G, V685A were selected and are highlighted in gray. (B) Bar chart depicting the results from the representative 384-well plate. Mutants labeled with an asterisk (*) showed at least the same level of activity as the parental mutant 'RGVG' E593G, V685A and were selected.

## References

1 T. Kovács and L. Ötvös, *Tetrahedron letters,* 1988, 29, **36**, 4525-4528.

2 V. Siegmund, T. Santner, R. Micura and A. Marx, *Chem Sci*, 2011, **2**, 2224-2231.

3 R. Peist, A. Koch, P. Bolek, S. Sewitz, T. Kolbus and W. Boos, *J Bacteriol*, 1997, **179**, 7679-7686.

4 (a) K. B. Sauter and A. Marx, *Angew Chem Int Ed Engl*, 2006, **45**, 7633-7635; (b) S. Gieseking, K. Bergen, F. Di Pasquale, K. Diederichs, W. Welte and A. Marx, *J Biol Chem*, 2011, **286**, 4011-4020.