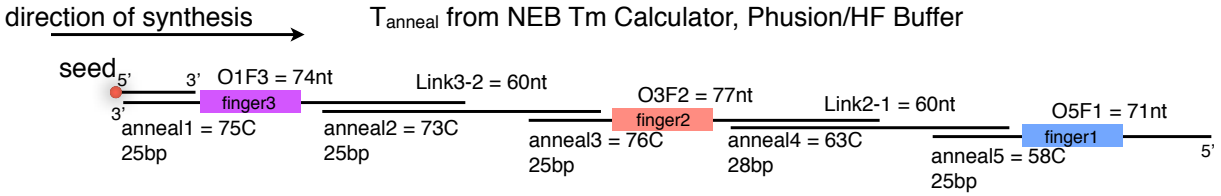
```
gat ctt aag gct aga gta cTA ATA CGA CTC ACT ATA GGG AGA CCA CAA CGG TTT CCC TCT AGA AAT AAT T  <70
            10          20          30          40          50          60

TT GTT TAA Ctt aag aag gag gaa aaa aaa atg gaa cgt ccg tac gct tgc ccg gtt gaa tct tgc gac cg  <140
                                      M   E   R   P   Y   A   C   P   V   E   S   C   D   R
            80          90          100         110         120         130

              finger1
t cgt ttc tct cgt tct gac gaa ctg acc cgt cat att aga att cat act gga caa aAA CCA TTC CAa tgt  <210
  R   F   S   R   S   D   E   L   T   R   H   I   R   I   H   T   G   Q   K   P   F   Q   C
            150         160         170         180         190         200

                                  finger2
aga att tgt atg aga aat ttc tct cgt tct gac cac ctg acc acc cac atc cgt acc cac acc ggt gaa a  <280
R   I   C   M   R   N   F   S   R   S   D   H   L   T   T   H   I   R   T   H   T   G   E   K
            220         230         240         250         260         270

                                                  finger3
aa ccg ttc gct tgc gac atc tgc ggt cgt aaa ttc gct cgt tct gac gaa cgt aaa cgt cac acc aaa at  <350
   P   F   A   C   D   I   C   G   R   K   F   A   R   S   D   E   R   K   R   H   T   K   I
            290         300         310         320         330         340

c cac ctg cgt cag aaa gac cca gcg cca gcg cca tct aaa ggt gaa gaa tta ttc act ggt gtt gtc cca  <420
  H   L   R   Q   K   D   P   A   P   A   P   S   K   G   E   E   L   F   T   G   V   V   P
            360         370         380         390         400         410

att ttg gtt gaa tta gat ggt gat gtt aat ggt cac aaa ttt tct gtc tcc ggt gaa ggt gaa ggt gat g  <490
I   L   V   E   L   D   G   D   V   N   G   H   K   F   S   V   S   G   E   G   E   G   D   A
            430         440         450         460         470         480

ct act tac ggt aaa ttg acc tta aaa ttt att tgt act act ggt aaa ttg cca gtt cca tgg cca acc tt  <560
   T   Y   G   K   L   T   L   K   F   I   C   T   T   G   K   L   P   V   P   W   P   T   L
            500         510         520         530         540         550

a gtc act act tta act tat ggt gtt caa tgt ttt tct aga tac cca gat cat atg aaa caa cat gac ttt  <630
  V   T   T   L   T   Y   G   V   Q   C   F   S   R   Y   P   D   H   M   K   Q   H   D   F
            570         580         590         600         610         620

ttc aag tct gcc atg cca gaa ggt tat gtt caa gaa aga act att ttt ttc aaa gat gac ggt aac tac a  <700
F   K   S   A   M   P   E   G   Y   V   Q   E   R   T   I   F   F   K   D   D   G   N   Y   K
            640         650         660         670         680         690

ag acc aga gct gaa gtc aag ttt gaa ggt gat acc tta gtt aat aga atc gaa tta aaa ggt att gat tt  <770
   T   R   A   E   V   K   F   E   G   D   T   L   V   N   R   I   E   L   K   G   I   D   F
            710         720         730         740         750         760

t aaa gaa gat ggt aac att tta ggt cac aaa ttg gaa tac aac tat aac tct cac aat gtt tac atc atg  <840
  K   E   D   G   N   I   L   G   H   K   L   E   Y   N   Y   N   S   H   N   V   Y   I   M
            780         790         800         810         820         830

gct gac aaa caa aag aat ggt atc aaa gtt aac ttc aaa att aga cac aac att gaa gat ggt tct gtt c  <910
A   D   K   Q   K   N   G   I   K   V   N   F   K   I   R   H   N   I   E   D   G   S   V   Q
            850         860         870         880         890         900

aa tta gct gac cat tat caa caa aat act cca att ggt gat ggt cca gtc ttg tta cca gac aac cat ta  <980
   L   A   D   H   Y   Q   Q   N   T   P   I   G   D   G   P   V   L   L   P   D   N   H   Y
            920         930         940         950         960         970

c tta tcc act caa tct gcc tta tcc aaa gat cca aac gaa aag aga gac cac atg gtc ttg tta gaa ttt  <1050
  L   S   T   Q   S   A   L   S   K   D   P   N   E   K   R   D   H   M   V   L   L   E   F
            990         1000        1010        1020        1030        1040

gtt act gct gct ggt att acC CAT GGT ATG GAT GAA TTG TAC AAA ggg ggt tct cat cat cat cat cat c  <1120
V   T   A   A   G   I   T   H   G   M   D   E   L   Y   K   G   G   S   H   H   H   H   H   H
            1060        1070        1080        1090        1100        1110

at TAA TAA CGA CTC AGG CTG CTA CCT AGC ATA ACC CCT TGG GGC CTC TAA ACG GGT CTT GAG GGG TTT TTT GA
   *   *    1130        1140        1150        1160        1170        1180
```

Features :
T7 promoter: [19 : 42], RBS: [86 : 92]
finger1: [151 : 171], finger2: [235 : 255], finger3: [319 : 339]
proline linker: [365 : 384], yEGFP: [385 : 1095], 6x Histidine tag: [1105 : 1122], T7 terminator: [1145 : 1192]

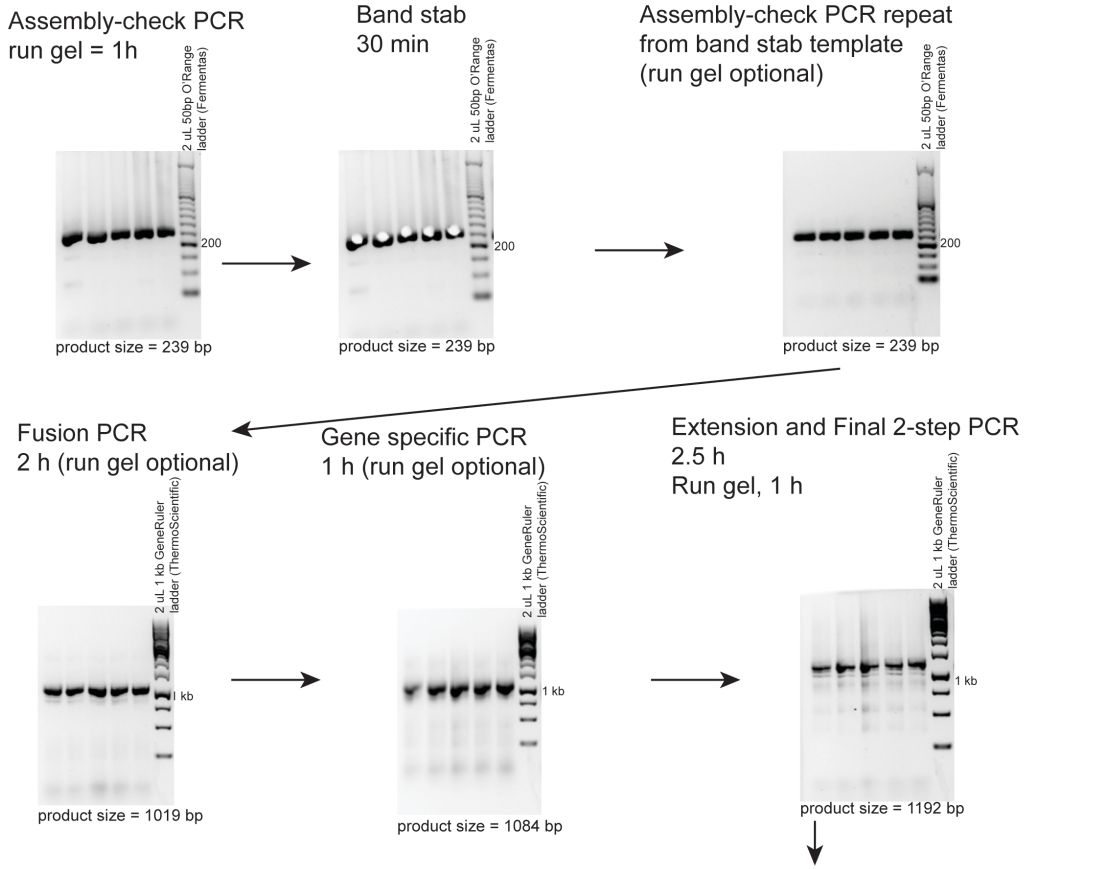
Supplementary Figure 1
Complete annotated DNA sequence of expression-ready linear template for WT zif268. To create ZF variants, substitute in the coding sequence (21 nt, in blue) for residues -1 to 6 of recognition helix into the positions of finger 1, 2 and 3.



Supplementary Figure 2

Schematic of the oligomers used in APE assembly of tridactyl zinc finger transcription factors. Generic sequences are given in Supplementary Table 1. Synthesis occurs from the 3'end of the gene (Finger3) to the 5'end (Finger1). O1F3, O3F2, and O5F1 are unique oligos with 21nt in the colored regions coding for recognition helix variants that target different DNA triplets. Link3-2 and Link2-1 are universal, and used in all APE assemblies.

Time Estimates (for maximum 12 unique constructs, starting with identical 1st and 2nd oligos)	Time
Aliquoting beads, washing, resuspending, mixing with initiator oligo	5 min
Incubation on rotisserie	15 min
Buffer exchange/wash, resuspend in Oligo1 Mix, aliquot, thermal cycle	10 min
Pool, wash, resuspend in NaOH	15 min
Buffer exchange/wash, resuspend in Oligo2 Mix, aliquot, thermal cycle	10 min
Pool, wash, resuspend in NaOH	15 min
Partition, buffer exchange/wash, resuspend in Oligo3 Mix, thermal cycle	15 min
Wash, resuspend in NaOH	15 min
Buffer exchange/wash, resuspend in Oligo4 Mix, thermal cycle	15 min
Wash, resuspend in NaOH	15min
Partition, buffer exchange/wash, resuspend in Oligo5 Mix, thermal cycle	20 min
Wash, resuspend in 1xTris-Cl	10 min
	Total = 2.5-3 h

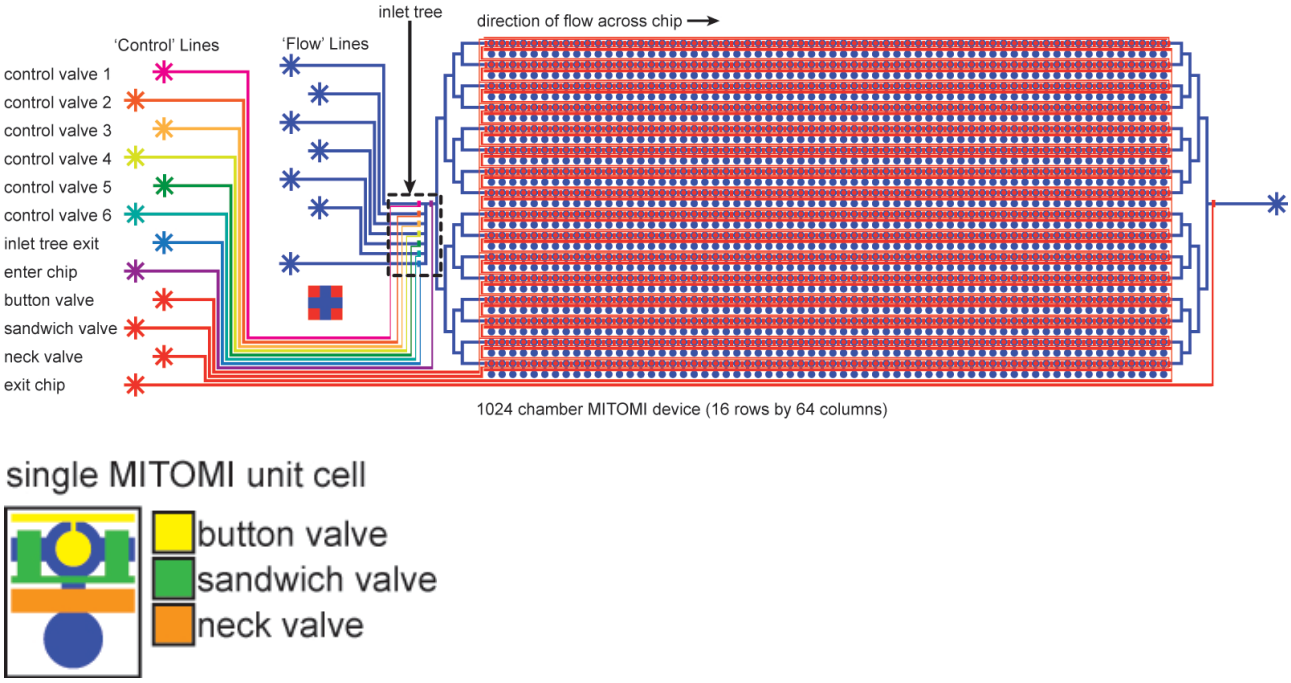


template used directly for microarray printing

Microarray spotting: 4 slides	Time estimate
2% BSA in water prespot	1.5h
Template spotting	2.75h
2% BSA in water overspot	1.5h
Klenow target spotting	5.5h

MITOMI Experiments
6h to prepare and incubate
1h to scan

Supplementary Figure 3
Time estimates for performing up to 12 unique (no reaction pooling and splitting) APE assembly reactions by hand. Following the assembly reaction, several PCRs are performed to confirm full-length assembly, and to add on sequence necessary for on-chip expression and detection (5' and 3'UTRs, EGFP tag). Representative gels for the assembly of 5 templates during each step are shown. Finally, another list of time estimates for the microarraying process and running MITOMI experiments is given.



Supplementary Figure 4
A schematic of the 1024 chamber MITOMI device used for all of the experiments described in this publication. On the far left are several ports where PBS-filled control lines are inserted to actuate microfluidic valves on the device with compressed air. The last 4 lines (in red) control the button valve, sandwich valve, neck valve and chip exit valve, from top-to-bottom, respectively. An enlarged image of a single MITOMI unit cell displays where each of the valves are located. The flow lines are where experimental buffers/reagents are inserted to flow across the chip. The operation of the device is detailed in the methods section.

Command line:

```
rosetta_scripts.linuxgccrelease -s 1A1L_0001.pdb  -parser:protocol prot-dna_script.xml -nstruct 3 -
ignore_unrecognized_res
```

Rosettascripts code:

```
<ROSETTASCRIPTS>
        <TASKOPERATIONS>
                <InitializeFromCommandline name=IFC/>
                <IncludeCurrent name=IC/>
                <RestrictDesignToProteinDNAInterface name=DnaInt base_only=1 z_cutoff=3.0 dna_defs=B.1.GUA/>
                <OperateOnCertainResidues name=AUTOprot>
                        <AddBehaviorRLT behavior=AUTO/>
                        <ResidueHasProperty property=PROTEIN/>
                </OperateOnCertainResidues>
                <OperateOnCertainResidues name=ProtNoDesign>
                        <RestrictToRepackingRLT/>
                        <ResidueHasProperty property=PROTEIN/>
                </OperateOnCertainResidues>
                <OperateOnCertainResidues name=DnaNoPack>
                        <PreventRepackingRLT/>
                        <ResidueHasProperty property=DNA/>
                </OperateOnCertainResidues>
        </TASKOPERATIONS>
        <SCOREFXNS>
                <DNA weights=dna/>
        </SCOREFXNS>
        <FILTERS>
        </FILTERS>
        <MOVERS>
                <DnaInterfacePacker name=score scorefxn=DNA task_operations=IFC,IC,AUTOprot,ProtNoDesign,DnaInt
probe_specificity=1 binding=1/>
        </MOVERS>
        <PROTOCOLS>
                <Add mover_name=score/>
        </PROTOCOLS>
</ROSETTASCRIPTS>
```

Protein stability calculations
Command line:

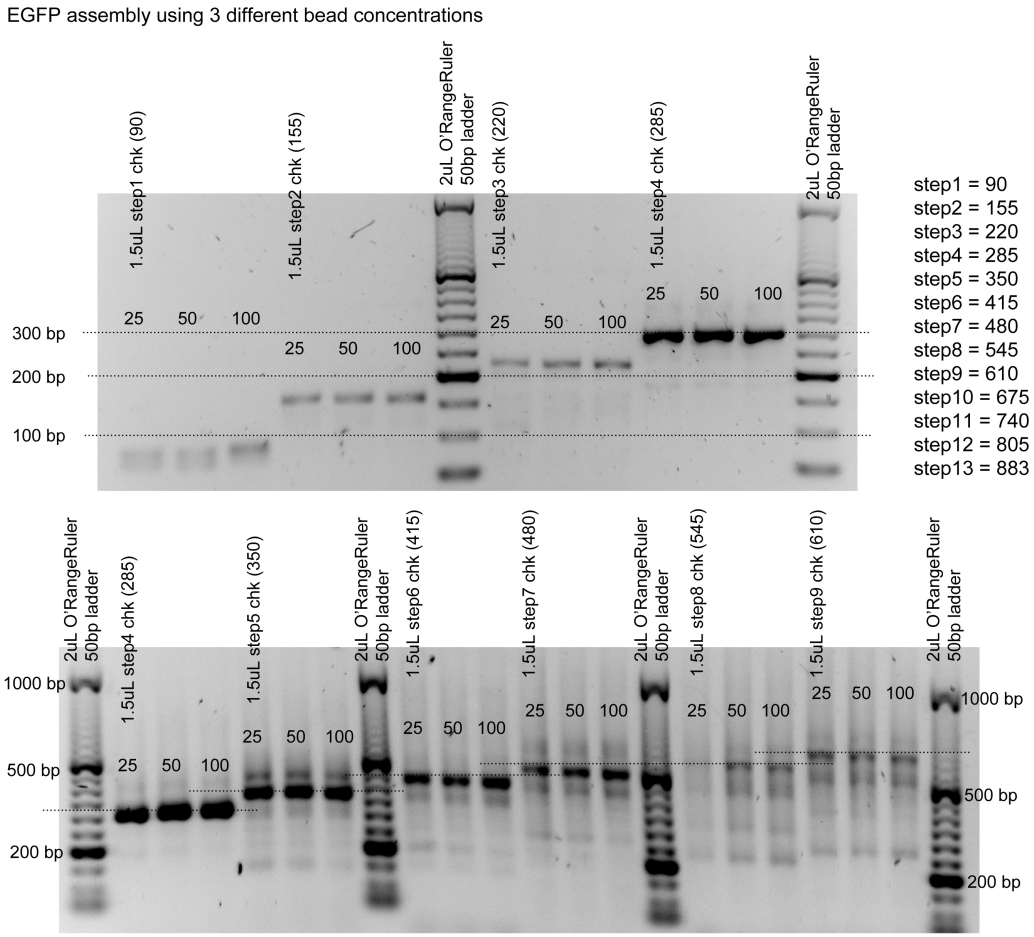
```
rosetta_scripts.linuxgccrelease -s 1A1L_noDNA.pdb -parser:protocol prot_stab.xml -nstruct 50 -
ignore_unrecognized_res -ex1 -ex2 -extrachi_cutoff 5 -in:auto_setup_metals
```

Rosettascripts code:

```
<ROSETTASCRIPTS>
        <TASKOPERATIONS>
                <ReadResfile name=rrf filename=resfile/>
        </TASKOPERATIONS>
        <SCOREFXNS>

                <scorefxn1 weights=talaris2013>
                <Reweight scoretype="atom_pair_constraint" weight=1.0/>
                <Reweight scoretype="angle_constraint" weight=1.0  />
                </scorefxn1>

        </SCOREFXNS>
        <FILTERS>
        </FILTERS>
        <MOVERS>
                <PackRotamersMover name=packrot  task_operations=rrf scorefxn= scorefxn1/>
                <Prepack name=ppk jump_number=0 scorefxn= scorefxn1/>
                <MinMover name=sc_bb_min bb=0 chi=1 scorefxn= scorefxn1/>
        </MOVERS>
        <PROTOCOLS>
                <Add mover_name=packrot />
                <Add mover_name=ppk />
                <Add mover_name=sc_bb_min />
        </PROTOCOLS>
</ROSETTASCRIPTS>
```

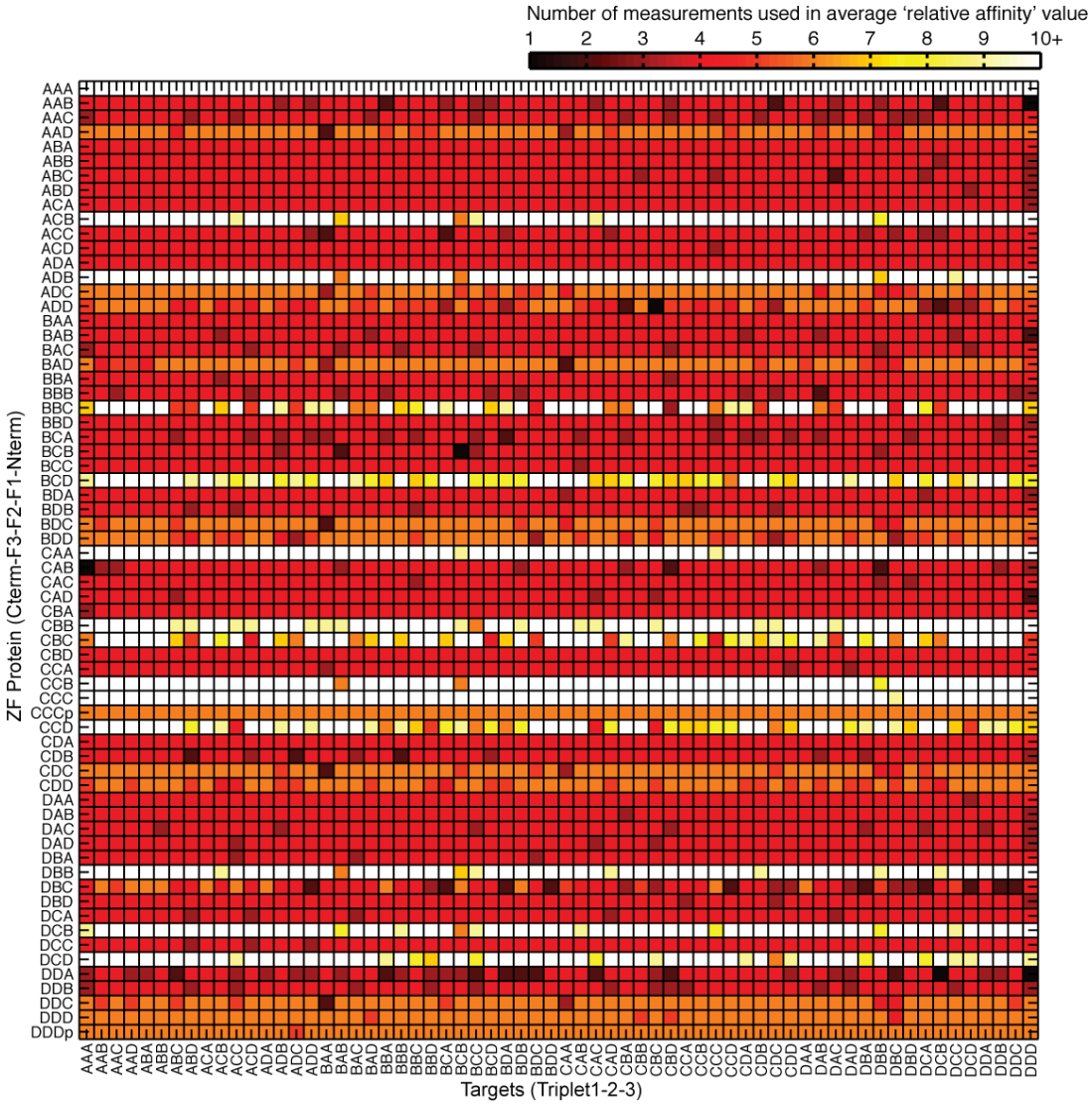
Supplementary Figure 5

Protein-DNA binding energy calculations performed using Rosetta command lines and scripts.
Rosetta version from Github repository-2662b747e67cf11cd76e6dedf2e9ff48cfefcd7c



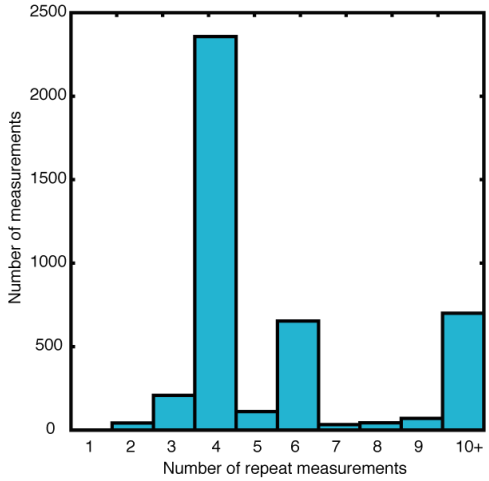
Supplementary Figure 6

Agarose gel images of PCR amplifications of APE assemblies attached to magnetic beads using 90mers with 25bp overlap to construct a linear template of EGFP. Here we demonstrate that up to 9 consecutive APE steps can be performed.



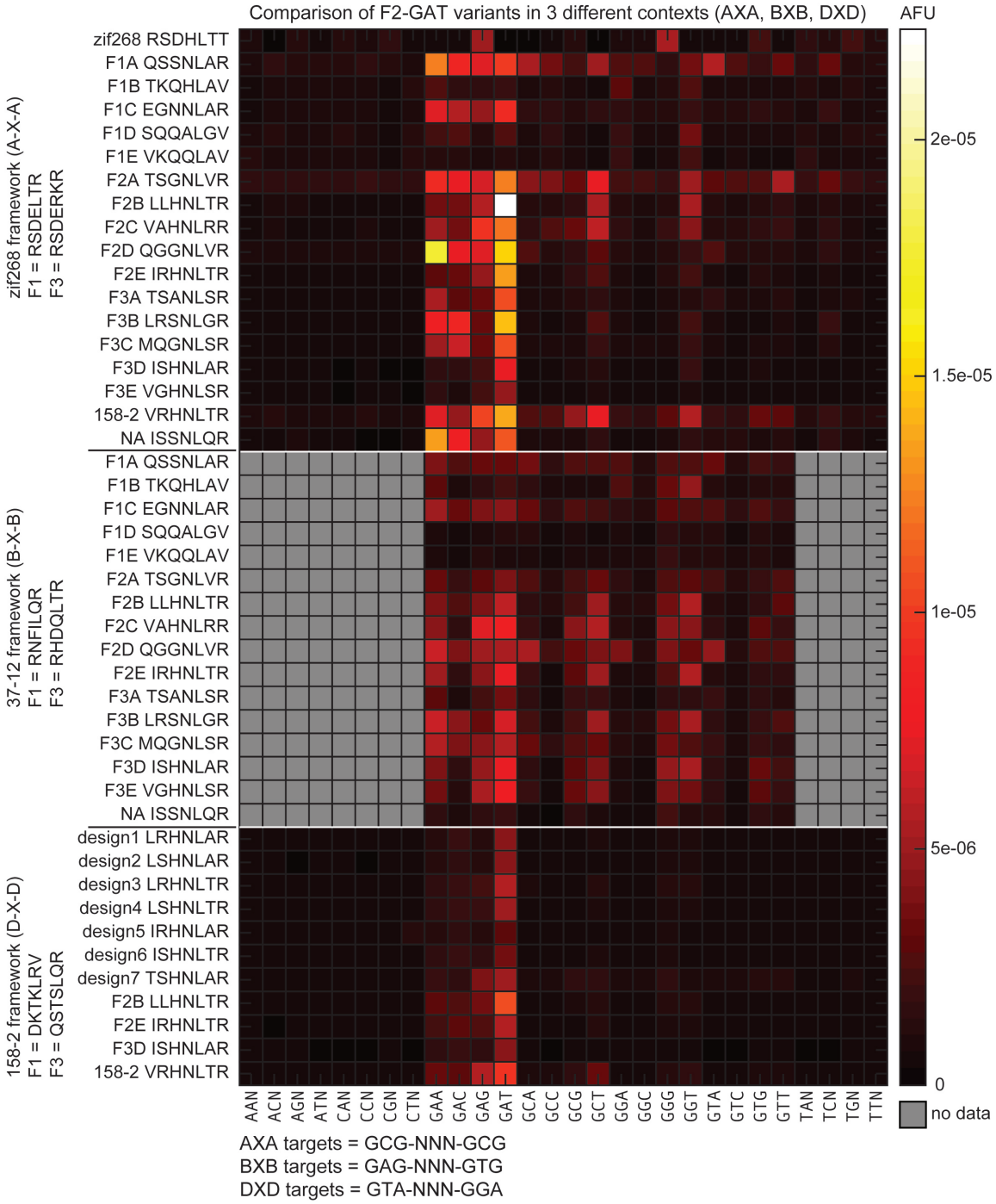
Supplementary Figure 7

Tabulation of number of repeated measurements used in generating an average 'relative affinity' value reported in a given tile of Figure 1f.



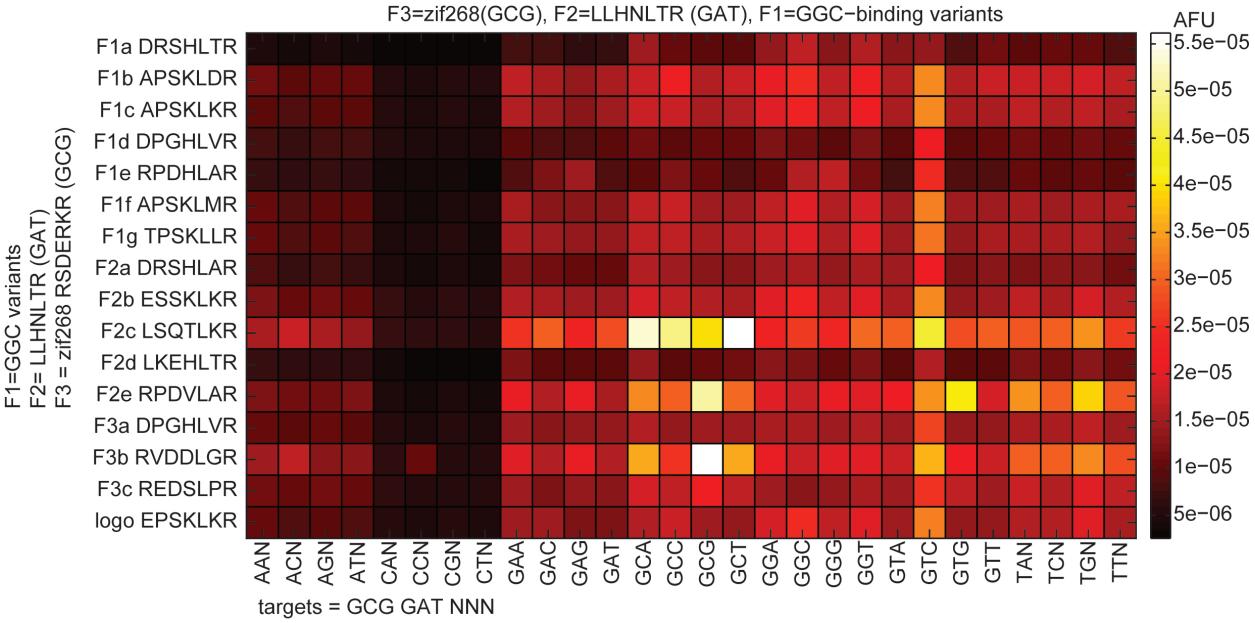
Supplementary Figure 8

Histogram of number of repeated measurements used in average values of Figure 1f.



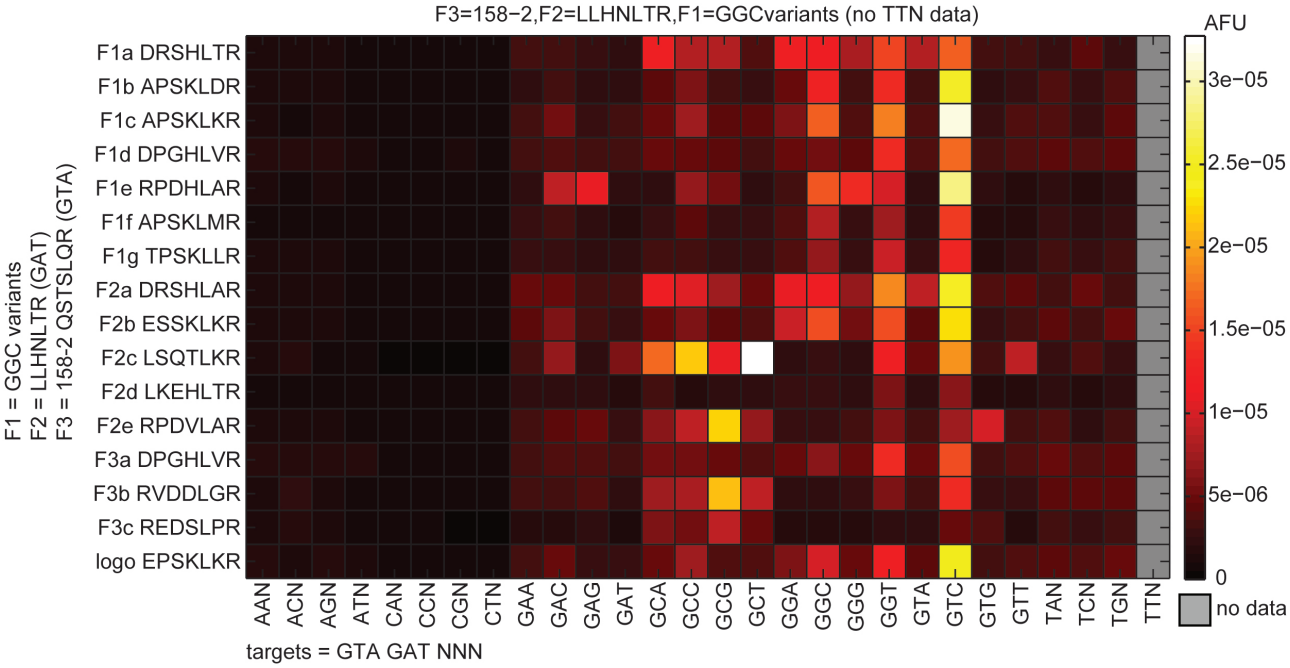
Supplementary Figure 9

Complete heat map data (subset of data in Figure 2a) from all ZF TFs containing F2 variants that were selected to bind the triplet GAT in different contexts (different F1/F3 combinations). The topmost 18 RH's placed in the Zif268 F1/F3 context exhibited the highest affinities for the GAT target, whereas the same set placed within the 37-12 F1/F3 context exhibited weakened affinities with no clear affinity for GAT. As a final screen, the highest affinity variants from the Zif268 screen were placed into the 158-2 F1/F3 context, in addition to seven 'designed' RHs based on residue combinations from the highest affinity/lowest non-specific variants (F2B, F2E and F3D). As observed in the Zif268 context, the RH F2B (LLHNLTR) had the highest affinity, and was used as the F2 variant in subsequent screens for the other finger positions.



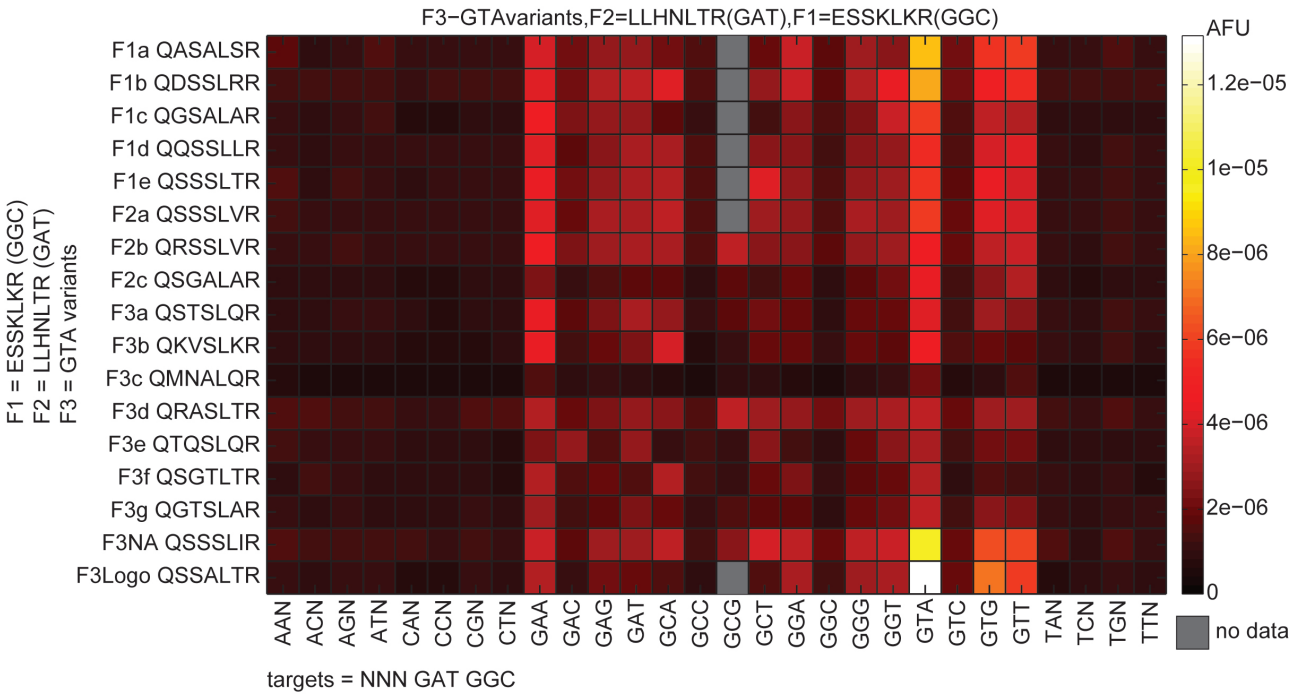
Supplementary Figure 10

Complete heat map data from F1 RH variants selected to bind GGC with F3 from Zif268 and F2-LLHNLTR from the GAT selection screen. Due to nonspecific binding for nearly all GNN targets, there is no clear RH with high specificity for GGC, and so a second screen was performed with a different F3 (Supplementary Figure 11).



Supplementary Figure 11

Complete heat map data from F1 RH variants selected to bind GGC with F3 from 158-2 (subset of data in Figure 2a), which resulted in a weakened affinity across the entire target range, but also reduced the non-specific binding 'noise' seen in Supplementary Figure 8. In this screen, while there is no high-specificity variant for GGC, by comparing the relative specificities for GGC and GTC, F2B ESSKLKR was selected. The 'logo' RH at the bottom of the heat map was generated by taking the highest frequency residue in each RH position from all available GGC variants listed in the Zinc Finger Database.



Supplementary Figure 12

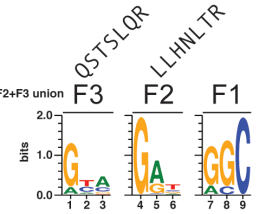
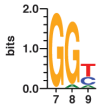
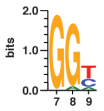
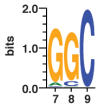
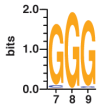
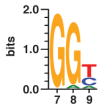
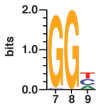
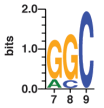
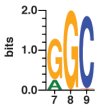
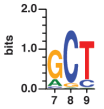
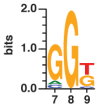
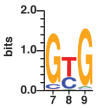
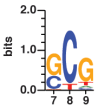
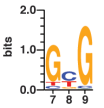
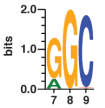
Full heat map data from F3 RH variants selected to bind GTA (subset of data in Figure 2a), with the chosen RH from the F2 and F1 screening rounds. In this screen, the highest affinity variant was the 'logo' design, which was generated by taking all of the available GTA variants listed in the Zinc Finger Database and selecting the highest frequency residue at each position (QSSALTR). This RH was chosen to complete the 3 selection rounds towards developing a ZF TF that recognizes the sequence GTA GAT GGC.

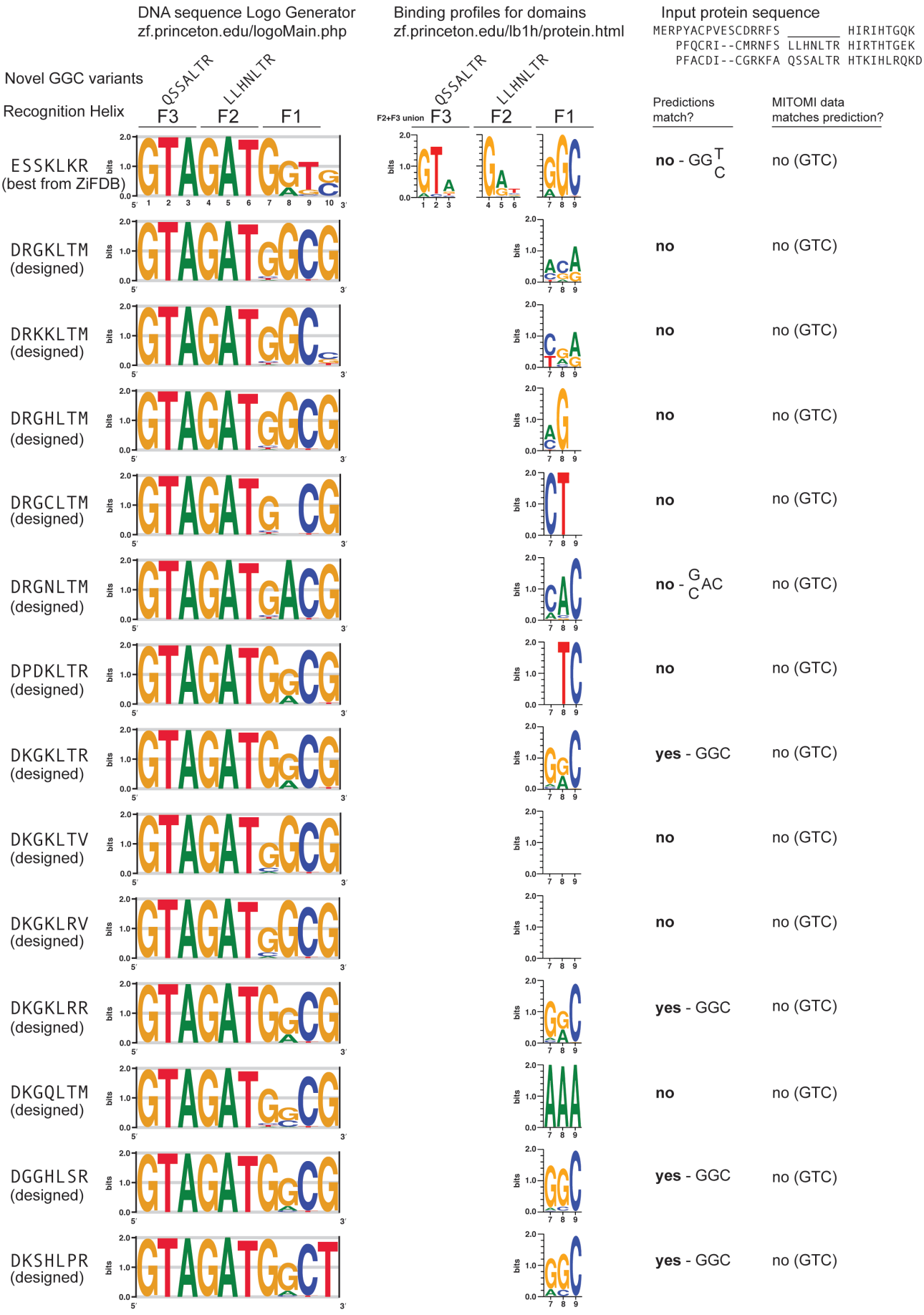
Input protein sequence

```
MERPYACPVESCDRRFS _______ HIRIHTGQK
   PFQCRI--CMRNFS LLHNLTR HIRTHTGEK
   PFACDI--CGRKFA QSTSLQR HTKIHLRQKD
```

DNA sequence Logo Generator: zf.princeton.edu/logoMain.php

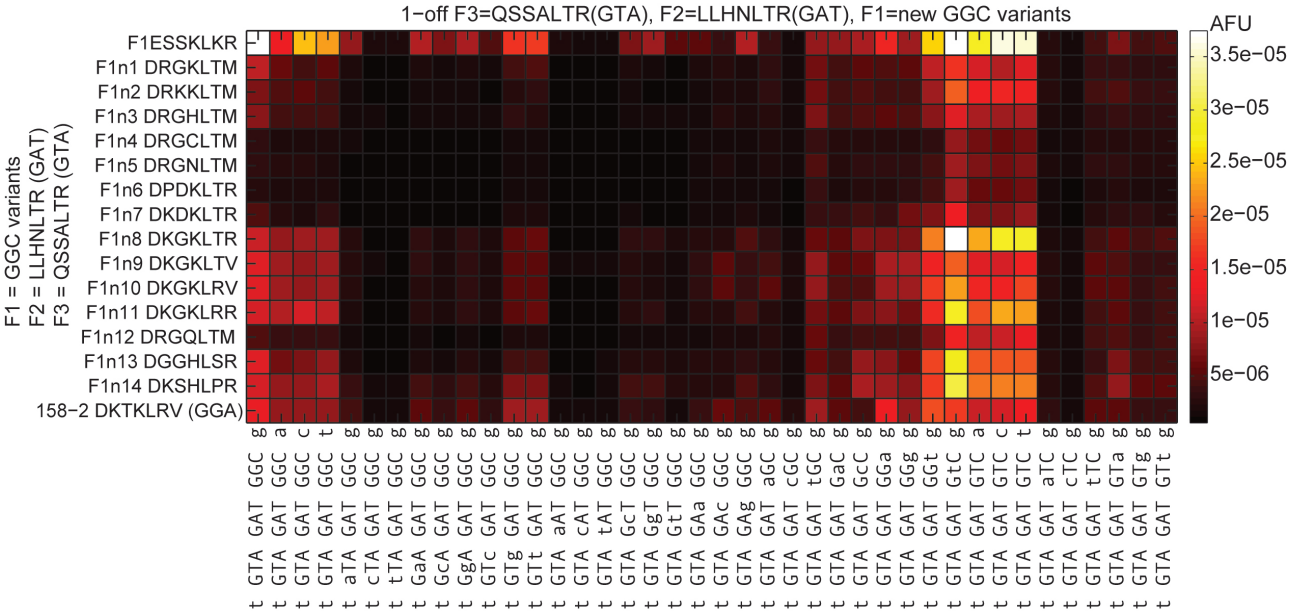
Binding profiles for domains: zf.princeton.edu/lb1h/protein.html

GGC binding variants from ZiFDB Recognition Helix	DNA sequence Logo Generator: zf.princeton.edu/logoMain.php	Binding profiles for domains: zf.princeton.edu/lb1h/protein.html	Predictions match?	MITOMI data matches prediction?
F1a DRSHLTR			**yes** - GGC	no (GTC)
F1b APSKLDR			**yes** - GGT	no (GTC)
F1c APSKLKR			**yes** - GGT	no (GTC)
F1d/F3a DPGHLVR			**yes** - GGC	no (GTC)
F1e RPDHLAR			**yes** - GGG	no (GTC)
F1f APSKLMR			**yes** - GGT	no (GTC)
F1g TPSKLLR			**yes** - GGT	no (GTC)
F2a DRSHLAR			**yes** - GGC	no (GTC)
F2b ESSKLKR			**no** - GG$^{T}_{C}$	no (GTC)
F2c LSQTLKR			**yes** - GCT	yes (GCT)
F2d LKEHLTR			**yes** - GGT	no (GTC)
F2e RPDVLAR			**no** - G$^{C}_{T}$G	yes (G$^{C}_{T}$G)
F3b RVDDLGR			**yes** - GCG	yes (GCG)
F3c REDSLPR			**yes** - GCG	yes (GCG)
logo EPSKLKR			**no** - GG$^{T}_{C}$	no (GTC)


DNA sequence Logo Generator
zf.princeton.edu/logoMain.php
Binding profiles for domains
zf.princeton.edu/lb1h/protein.html
Input protein sequence
MERPYACPVESCDRRFS _____ HIRIHTGQK
PFQCRI--CMRNFS LLHNLTR HIRTHTGEK
PFACDI--CGRKFA QSSALTR HTKIHLRQKD
Novel GGC variants
Recognition Helix
QSSALTR
LLHNLTR
F3
F2
F1
F2+F3 union
Predictions match?
MITOMI data matches prediction?
ESSKLKR (best from ZiFDB)
no - GG T C
no (GTC)
DRGKLTM (designed)
no
no (GTC)
DRKKLTM (designed)
no
no (GTC)
DRGHLTM (designed)
no
no (GTC)
DRGCLTM (designed)
no
no (GTC)
DRGNLTM (designed)
no - G C AC
no (GTC)
DPDKLTR (designed)
no
no (GTC)
DKGKLTR (designed)
yes - GGC
no (GTC)
DKGKLTV (designed)
no
no (GTC)
DKGKLRV (designed)
no
no (GTC)
DKGKLRR (designed)
yes - GGC
no (GTC)
DKGQLTM (designed)
no
no (GTC)
DGGHLSR (designed)
yes - GGC
no (GTC)
DKSHLPR (designed)
yes - GGC
no (GTC)

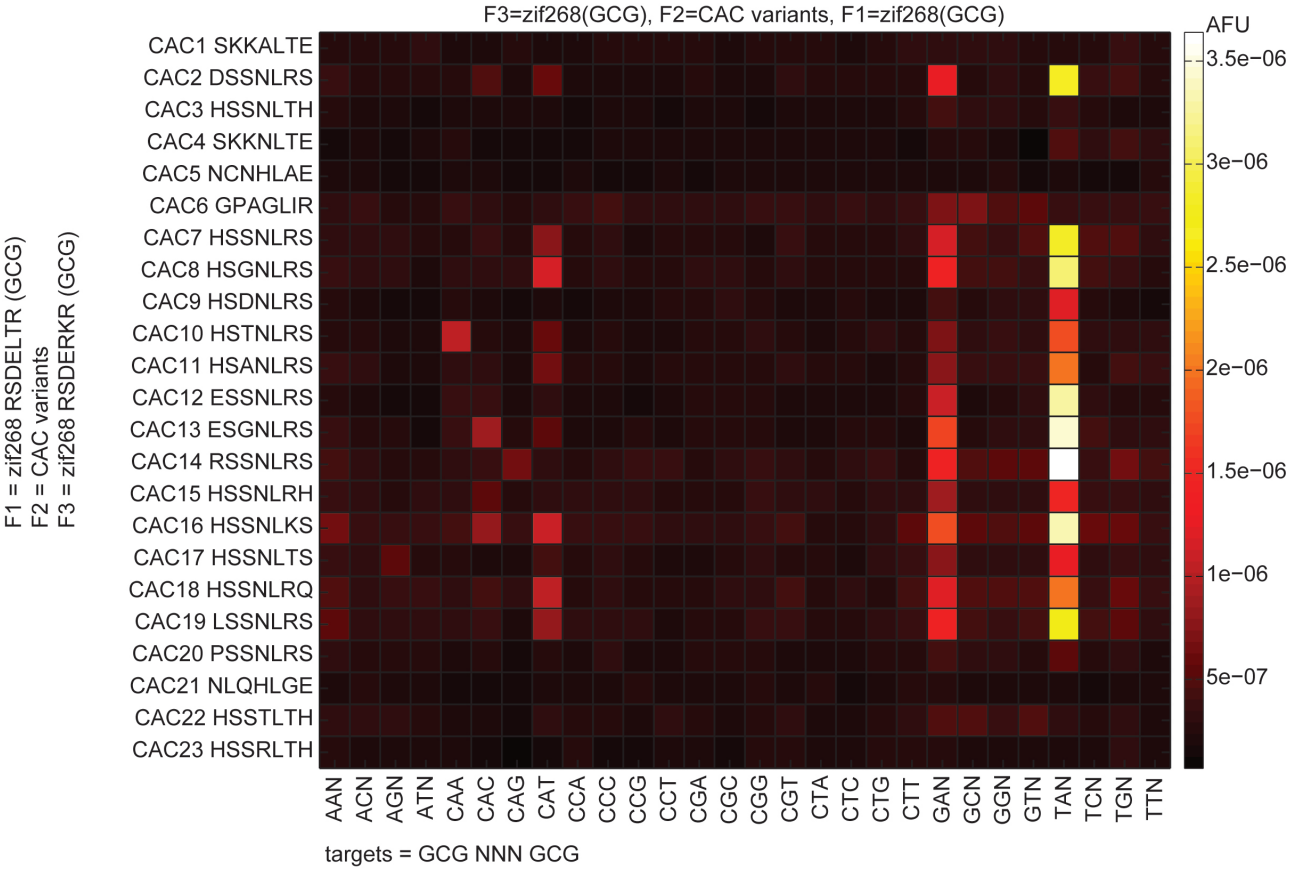
Supplementary Figure 13

Two tables of predicted DNA binding specificities for GGC-binding variants taken from the ZF Consortium Database, and novel variants created by substituting in new residues, using online prediction programs (20-23). The amino acid sequence of the complete ZF TF sequence (containing all three ZF domains) is given as input to the programs, which detect the ZF RH (residues -1 to 6) then predict the DNA binding site of each ZF domain. The first, leftmost column indicates the RH variant of interest. The output of each program is given in either the second or third column, respectively, as sequence logos. The fourth column indicates whether the two predictions agree with each other, and the final column compares the prediction with the observed MITOMI binding preference (target bound with the highest affinity).



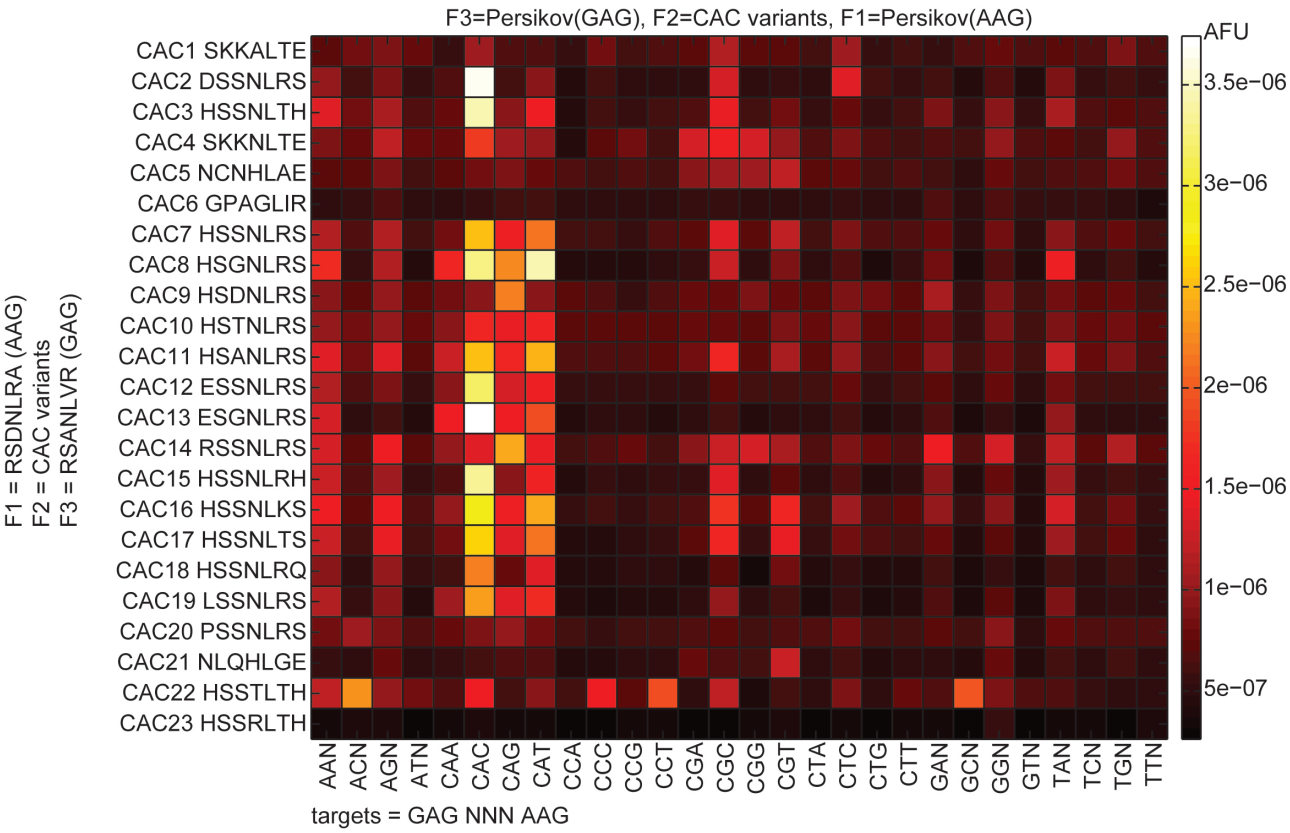
Supplementary Figure 14

Complete heat map data from the final engineered variant selected to bind GTA GAT GGC, in addition to several other F1 variants which were predicted (Supplementary Figure 13) to bind more specifically to GGC than those tested in the earlier F1 selection round (Supplementary Figure 11). Here all of the variants are tested against a 1-off target library to generate a detailed summary of specificity towards the target of interest. Despite the DNA specificity predictions given for these F1 variants, most of them have a preference for GTC rather than the desired GGC.



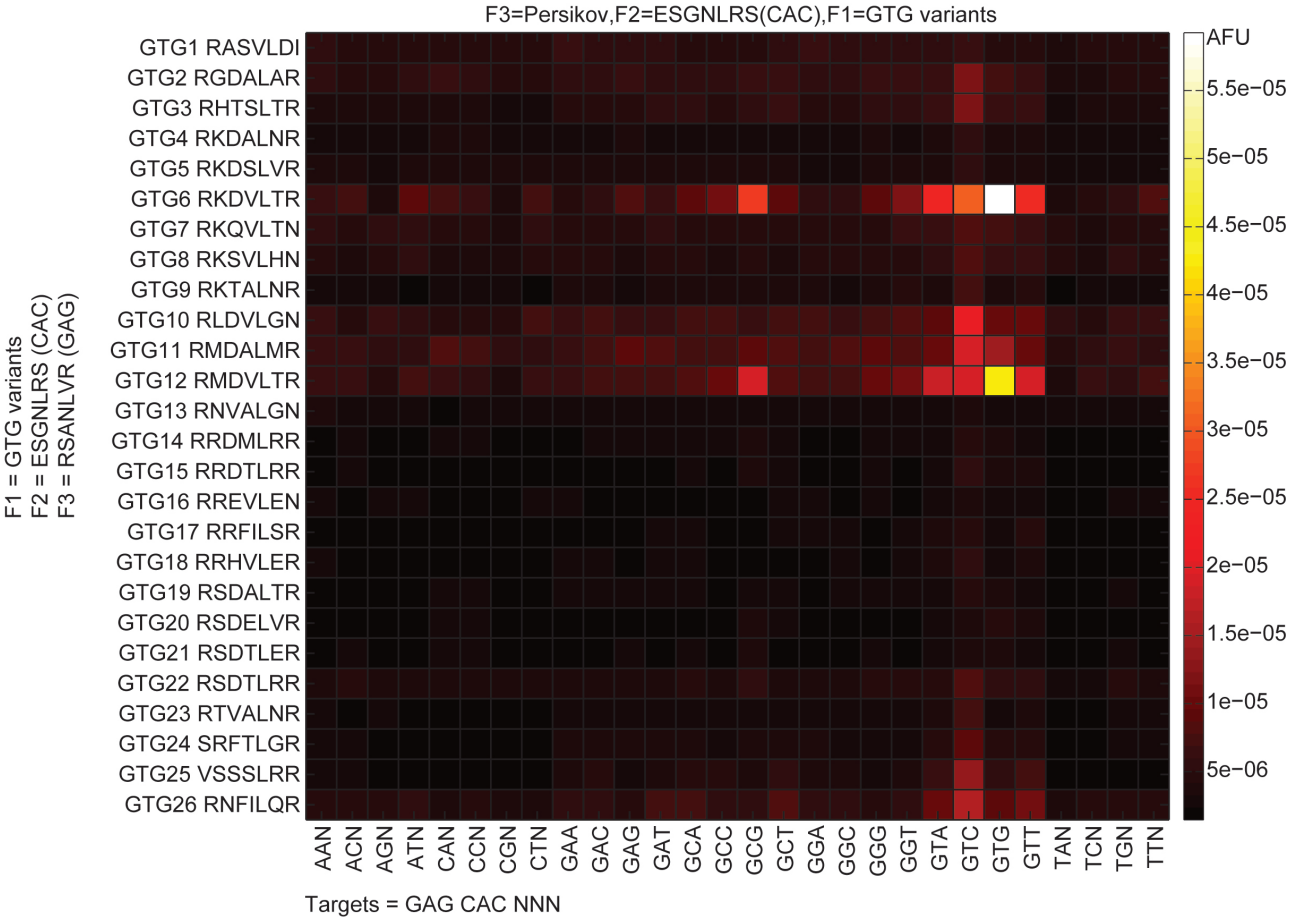
Supplementary Figure 15

Complete heat map data from all ZF TFs containing F2 variants that were selected or designed to bind the triplet CAC with F1/ F3 from Zif268. Due to a low list of options from the Zinc Finger Consortium Database, RHs were taken from recent publications, which reported CAC-binding ZF domains. CAC1, 5 and 6 came from reference 16 (Drier et al, 2005), CAC2 and 3 came from reference 22 (Persikov et al, 2014), and CAC21 was taken from a patent application (2004, EP1421177A2). The remaining RHs were designed around the amino acid residue logos presented for CAC (ref. 22) or from half-site designs reported in reference 11 (Gupta et al, 2012). In this initial screen, using F1/F3 from Zif268, it appeared that none of the RHs were functional for binding CAC, and instead we observed strong affinity for GAN or TAN targets. We believe this result can be explained by the strong cross-site interaction of the Zif268 F3 aspartic acid in position 2 of the RH, which had also been observed by ref 16. The natural F2 target in Zif268 is TGG/GGG, and so the cross-site interaction would prefer an A or C in the first base of the F2 target complement, which was only available in the GNN and TNN targets. Since this screen with F1/F3 from Zif268 did not produce any high affinity CAC binding variants, we performed a second screen (Supplementary Figure 16) using F1/F3 from ref 22.



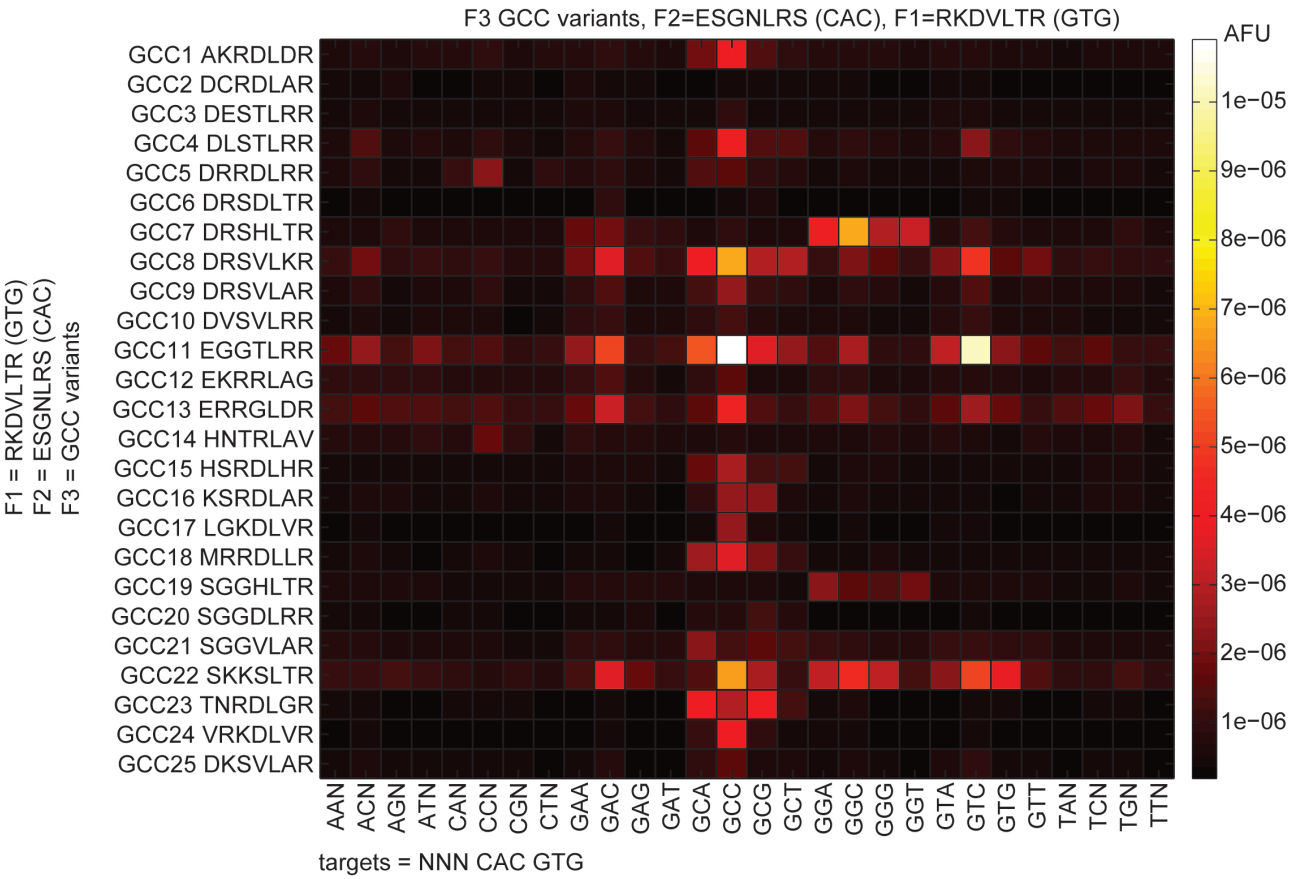
Supplementary Figure 16

Complete heat map data (subset of data in Figure 2b) from all ZF TFs containing F2 variants that were selected or designed to bind the triplet CAC with F1/ F3 from ref 22. This F1/F3 set is the context within which CAC2 and 3 were tested, so we knew at least these variants should be specific for CAC. As reported in ref 22, CAC1 displayed only weak binding towards CAC. While most of the variant designs did not come from published examples, many of them were capable of binding CAC, but with reduced specificity. In the end, variant CAC13 was chosen for subsequent screens since it was a novel design and exhibited relatively high specificity and affinity towards CAC.



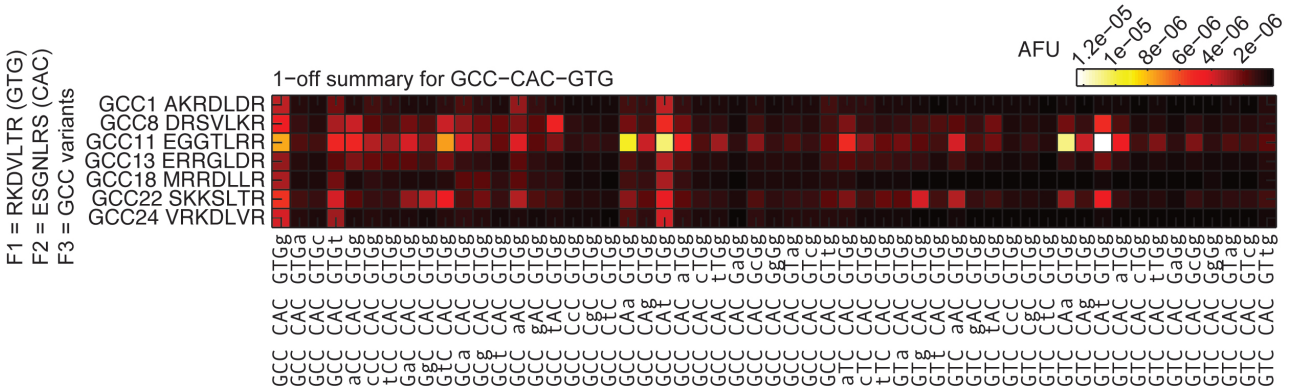
Supplementary Figure 17

Full heat map data from F1 RH variants selected to bind GTG (subset of data in Figure 2b) with F3 from ref 22 and F2 (CAC13, ESGNLRS) from the CAC selection in Supplementary Figure 16. A large majority of the RHs tested were not functional, or bound with very low affinity, and many of the variants had a binding preference for GTC in addition to the desired target GTG. Variant GTG6 and GTG12 exhibited the highest affinity for GTG, and differ from each other in a single residue position (position 1 in the RH). GTG6 (RKDVLTR) was selected for the subsequent screens in spite of its secondary binding preference for GTC.



Supplementary Figure 18

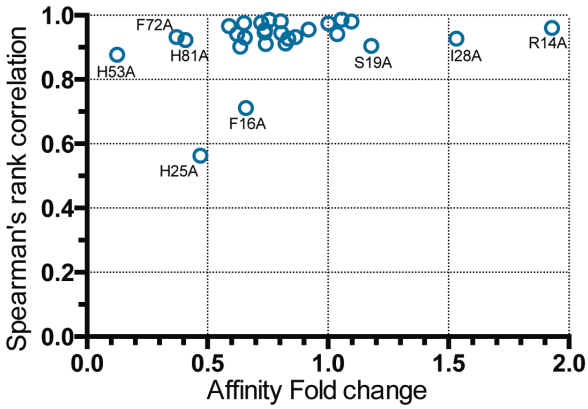
Full heat map data from F3 RH variants selected to bind GCC (subset of data in Figure 2b) with F1 from the GTG selection in Supplementary Figure 17 and F2 from the CAC selection in Supplementary Figure 16. In this selection, the majority of the variants were functional and displayed high specificity for GCC, but the highest affinity variants also displayed non-specific binding to other triplets. A selection of these variants was characterized against a 1-off target library.

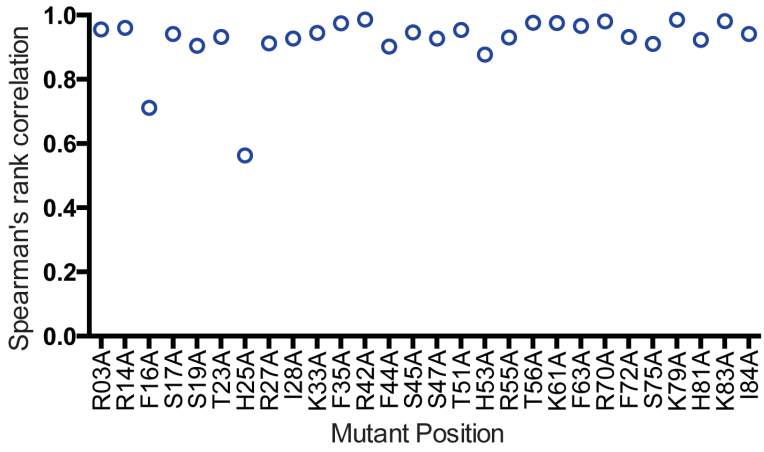


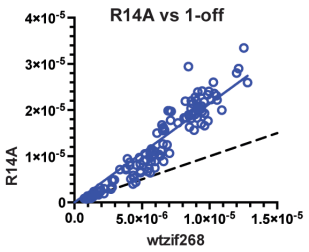
Supplementary Figure 19

Complete heat map data from the final engineered ZF TF selected to bind GCC CAC GTG, in addition to several other F3 (GCC) variants, which displayed lower affinity but higher specificity to GCC. Here all of the variants are measured against a 1-off target library to generate a detailed summary of specificity towards the target of interest. All of the variants have some affinity for CAT rather than CAC, and it becomes clear that a few of the F3 variants have a preference for GTC rather than GCC. In most cases, the target of interest (GCC CAC GTG) is bound, but with equivalent or lower affinity than other targets. The highest affinity variant (F1 = RKDVLTR, F2 = ESGNLRS, F3 = EGGTLRR) is reported in the vertically-oriented heat map of Figure 2b.

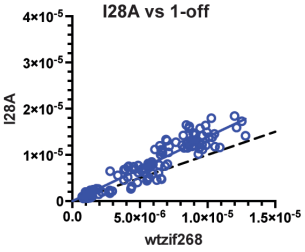
name	Fold change	Spearmans r
R03A	0.91975	0.95553
R14A	1.92895	0.96043
F16A	0.65925	0.71152
S17A	1.03805	0.94123
S19A	1.17970	0.90451
T23A	0.86478	0.93176
H25A	0.47010	0.56274
R27A	0.82351	0.91201
I28A	1.53270	0.92675
K33A	0.80639	0.94420
F35A	1.00195	0.97462
R42A	1.05607	0.98630
F44A	0.63581	0.90239
S45A	0.73692	0.94611
S47A	0.83495	0.92732
T51A	0.73796	0.95412
H53A	0.12509	0.87681
R55A	0.65476	0.93045
T56A	0.72269	0.97588
K61A	0.65126	0.97577
F63A	0.58931	0.96605
R70A	1.09723	0.98021
F72A	0.37127	0.93171
S75A	0.74140	0.91020
K79A	0.75666	0.98539
H81A	0.40791	0.92325
K83A	0.80416	0.98101
I84A	0.62245	0.94153



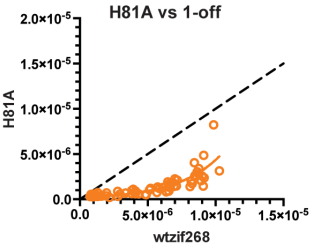




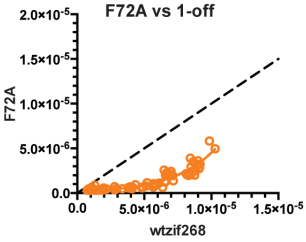
y = A * x
Slope 2.140
R square 0.8913



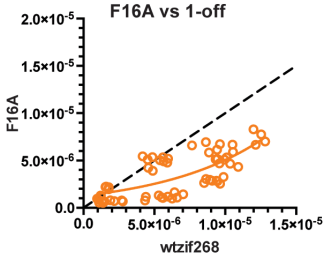
y = A * x
Slope 1.390
R square 0.8680



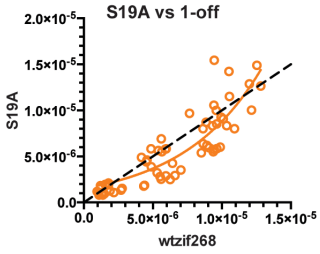
y = A * exp(B * x)
A 1.922e-007
B 311842
R square 0.6771



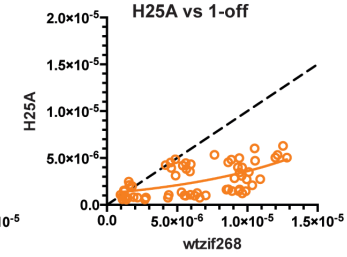
y = A * exp(B * x)
A 1.872e-007
B 318330
R square 0.8612



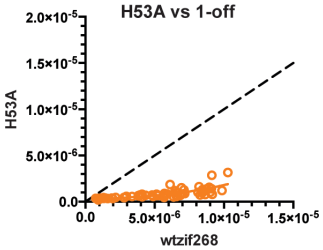
y = A * exp(B * x)
A 1.268e-006
B 138257
R square 0.5222



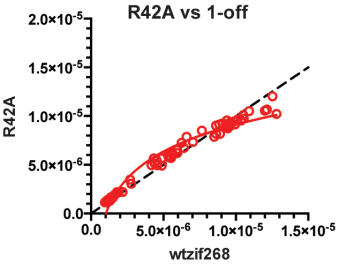
y = A * exp(B * x)
A 1.536e-006
B 174653
R square 0.7653



y = A * exp(B * x)
A 1.231e-006
B 107192
R square 0.3217



y = A * exp(B * x)
A 2.547e-007
B 195522
R square 0.5812



y = A* ln(x)+B
A 3.996e-006
B 5.514e-005
R square 0.9552

Supplementary Figure 20

Analysis of data obtained from examining Zif268 point-mutation affinity variants (Figure 2c) tested against a 1-off target library for the Zif268 consensus target (GCG TGG GCG). To visualize the effect of the mutation in each variant, the affinities of the wt Zif268 protein for each DNA target are plotted against the measured affinities of the mutated protein. To show that target specificity has not changed as a result of tuning the affinity, the Spearman's rank correlation coefficient was calculated for each variant (see table of values above). These coefficients were plotted against the fold-change in affinity calculated using only the Zif268 consensus target (Figure 2d), and also as a function of position along the Zif268 protein sequence. The majority of variants have a Spearman's rank correlation greater than 0.9, except for 2 variants, F16A and H25A, which displayed severe departures from the specificity of the wildtype protein as a result of the mutation. These variants have a higher than expected affinity for certain targets, and lower than expected affinity for others as seen in the individual scatter plots. For each variant, we fit a curve to the data set, using a linear (blue), exponential (yellow) or log (red) equation to capture the behavior of the data. The black dashed line signifies the behavior of wildtype Zif268 plotted against itself, to compare against the behavior of the mutant variants.