Supplementary

Structures of CTCF-DNA complexes including all eleven zinc fingers

Jie Yang[1,$], John R. Horton[1,$], Bin Liu[1], Victor G. Corces[2], Robert M. Blumenthal[3], Xing Zhang[1,*], and Xiaodong Cheng[1,*]

[1]Department of Epigenetics and Molecular Carcinogenesis, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

[2]Department of Human Genetics, Emory University School of Medicine, Atlanta, Georgia, USA

[3]Department of Medical Microbiology and Immunology, and Program in Bioinformatics, The University of Toledo College of Medicine and Life Sciences, Toledo, OH 43614, USA

[$]equal contribution

* Correspondence: XZhang21@mdanderson.org (XZ); XCheng5@mdanderson.org (XC)

Email addresses of other authors:

JY (jieyang301@gmail.com); JRH (JRHorton@mdanderson.org);

BL (BLiu1@mdanderson.org); VGC (vgcorces@gmail.com);

RMB (Robert.Blumenthal@utoledo.edu)

Figures S1- S5

Tables S1-S2

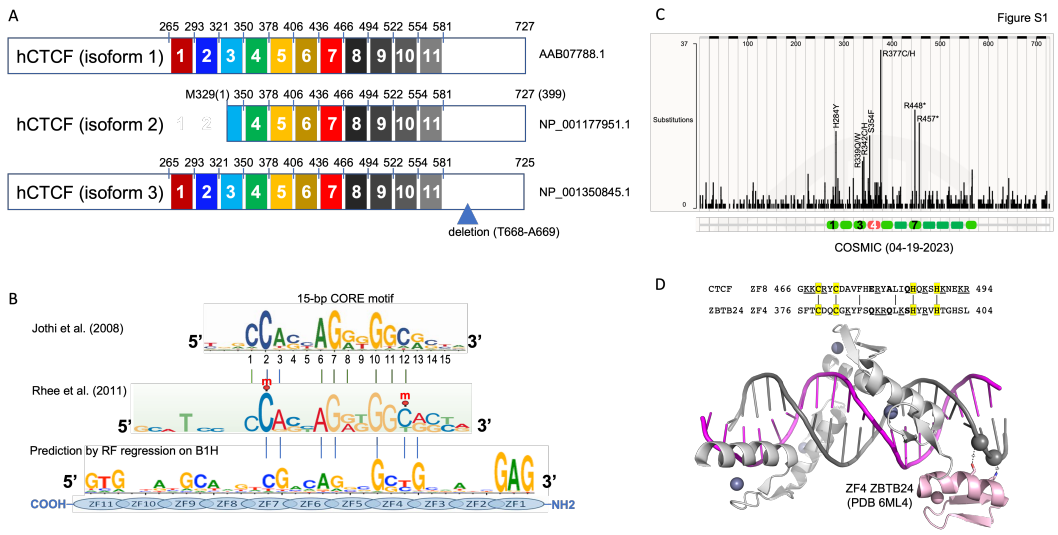


Figure S1. (A) Humans have three isoforms of CTCF, one of which is missing ZFs 1, 2, and half of 3. (**B**) Two examples of CTCF-binding consensus CORE sequence (base pairs 1-15) as determined by ChIP-seq (*1*) and ChIP-exo (*2*). DNA cytosine methylation (indicated by red circles and letter m) occurs at positions 2 and 12 of the consensus sequence in a subset of CTCF-binding sites (*3*). The experimentally determined consensus is aligned with a predicted CTCF DNA-binding specificity of ZF3-7. (**C**) CTCF mutations found in the Catalogue of Somatic Mutations in Cancer (COSMIC). The highest rate of mutations is C>T substitutions at codons of H284, R339, R342, S354, R377, R448 and R457. H284 of ZF1 is a zinc ligand. R339 of ZF3 is a base-interacting residue (see Figure 2F). R342 of ZF3 interacts with a DNA phosphate group. Ser354 is next to zinc-ligand Cys353 of ZF4. R377 of ZF4 interacts with a DNA phosphate group. R448* and R457* result in two deletions which eliminate translation of ZF7 and beyond. (**D**) Comparison between CTCF ZF8 and ZBTB24 ZF4 which holds two large and changed residues at positions -6 and -5. A spacer of ZF4 of ZBTB24 across DNA major groove.

Figure S2. CTCF is highly conserved among vertebrates, especially in the ZF DNA binding Domain. Human CTCFL is shown for comparison. The green dots (•) indicate the two sites of sumoylation (*4*) – both lysine residues are highly conserved among CTCF orthologs, but both are missing in CTCFL. The maroon white dots (•) indicate sites of poly-ADP-ribosylation (*5*). The magenta dots (•) indicate sites of phosphorylation during mitosis that reduce DNA binding (*6*). The insertions in Danio rerio CTCF are indicated by cyan triangles.

Numbering above the sequence indicates residues receiving particular attention in the text, and refer to the human ortholog. Letters in **bold red** are positions making base-specific contacts that are substituted in cases of human CTCF-related disorder (*7*); all are fully-conserved, even in CFCTL (shown for comparison). Grey shading is where ≥5 of the six sequences are identical. The ZFs are shaded yellow, with substitutions in cyan, with Zn-coordinating residues in bold. The highly-conserved Met in white on black (**M**) in ZF3 indicates the initiation codon for CTCF isoform 2 (NP_001177951.1), which in humans has a somewhat broader sequence specificity and competes with canonical CTCF, disrupts CTCF/cohesin binding, alters CTCF-mediated chromatin looping and promotes IFI6 activation and apoptosis (*8*). The white on black Thr-Ala (**T-A**) indicates two residues missing in CTCF isoform 3 (NP_001350845.1).

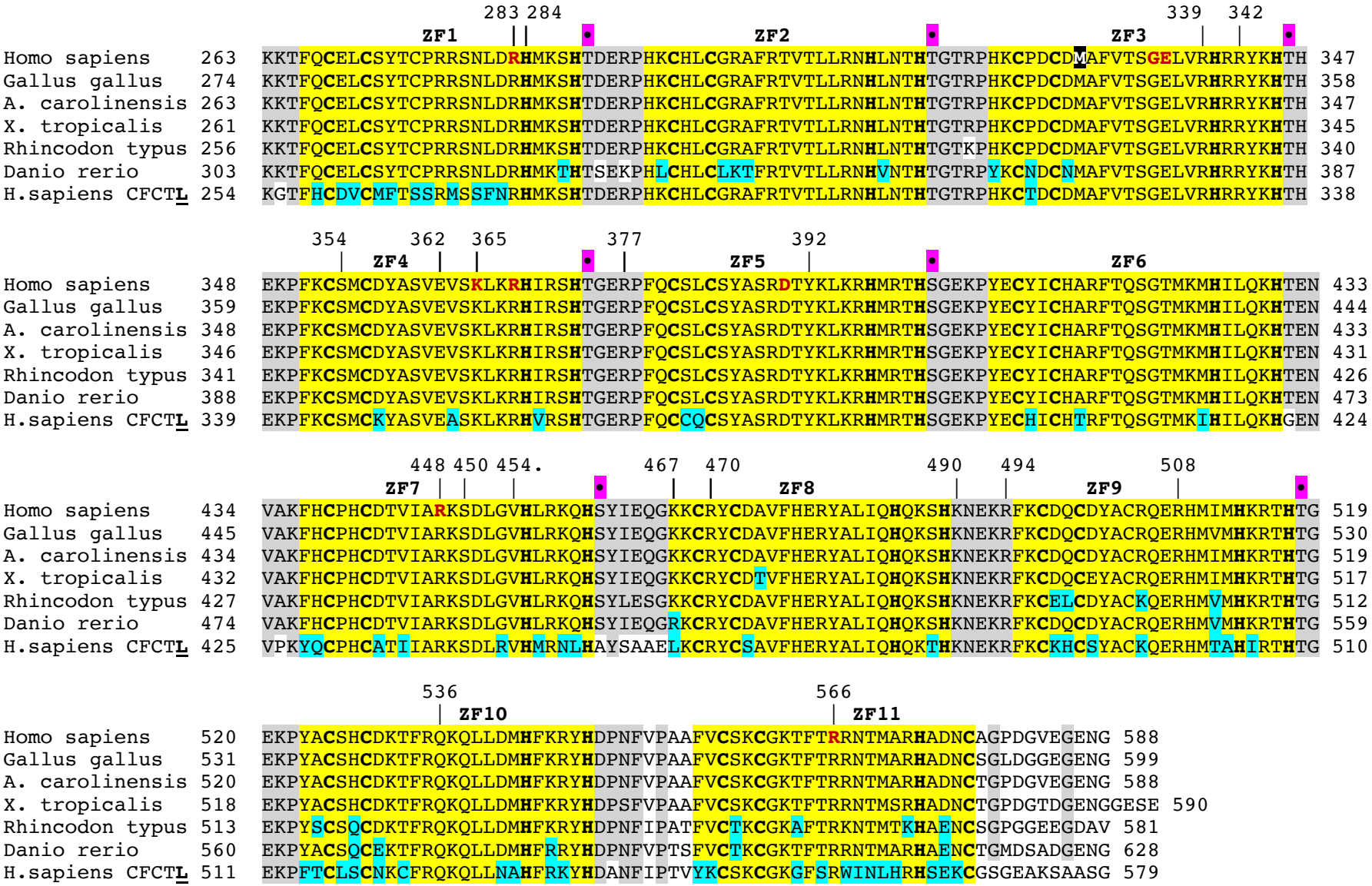
While all 11 CTCF ZFs are highly conserved among vertebrates, the number of substitutions is highest in ZFs 9-11. In particular, the whale shark *Rhincodon typus* has 12 substitutions relative to the human ortholog in ZFs 9-11, twice as many as the runner-up (*Danio rerio*). Three-quarters of the *R. typus* substitutions are fully conserved among other members of the class Chondrichthyes (not shown).

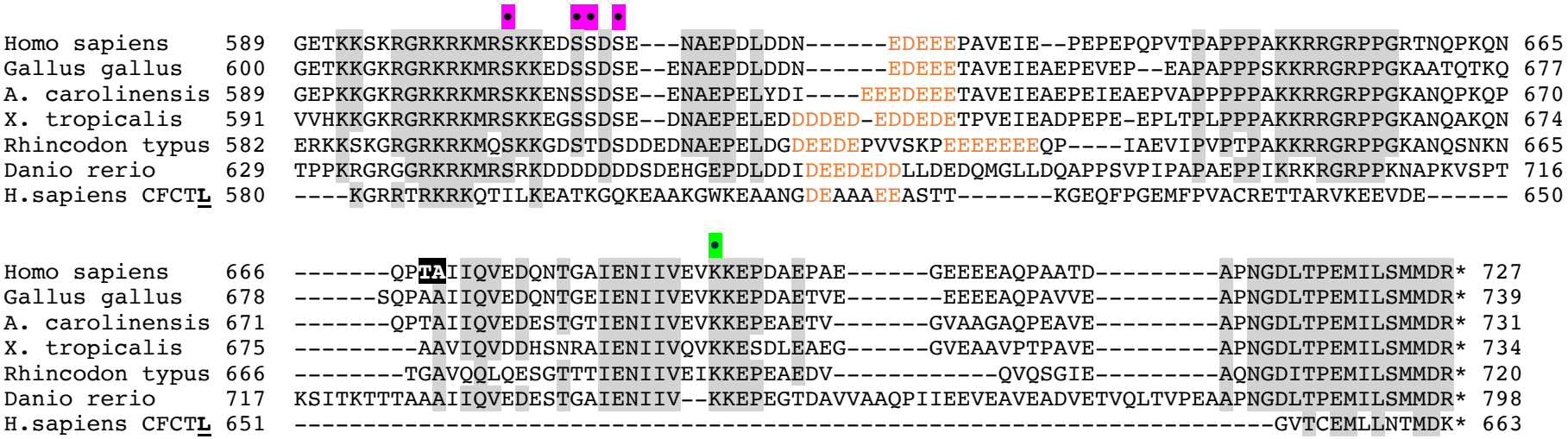
Figure. S2. CTCF is highly conserved among vertebrates, especially in the DBD. Human CTCFL is shown for comparison.

Description	Query cover	E Value	Per. Ident	Acc. Len	Accession	
transcriptional repressor CTCF isoform X1 [Homo sapiens]	100%	0.0	100.00	727	AAB07788.1	Mammalia
transcriptional repressor CTCF isoform X1 [Gallus gallus]	100%	0.0	93.55	739	XP_015134605.1	Aves
PREDICTED: transcriptional repressor CTCF [Anolis carolinensis]	100%	0.0	92.08	731	XP_003225381.1	Reptilia
transcriptional repressor CTCF [Xenopus tropicalis]	100%	0.0	85.93	734	NP_001116268.1	Amphibia
transcriptional repressor CTCF [Rhincodon typus]	100%	0.0	76.55	720	XP_048462054.1	Chondrichthyes
transcriptional repressor CTCF [Danio rerio]	100%	0.0	71.94	798	NP_001001844.1	Osteichthyes
transcriptional repressor CTCF**L** isoform 1 [Homo sapiens]	85%	2e-167	47.51	663	NP_001373922.1	

The green dots (•) indicate the two sites of sumoylation – both lysines are highly conserved among CTCF orthologs, but both are missing in CTCFL.
The maroon white dots (•) indicate sites of poly-ADP-ribosylation.
The magenta dots (•) indicate sites of phosphorylation during mitosis.
The insertions in Danio rerio CTCF are indicated by cyan triangles.

```
Homo sapiens     1    MEGDAVEAIVEESETFIKGKERKTYQRRREGGQEEDACHLPQNQTDGGEVVQDVNSSVQMVMMEQLDPTLLQMKTEVMEGTVAPEAEAAVDDTQ  94
Gallus gallus    12   MEGEAVEAIVEESETFIKGKERKTYQRRREGGQEDEACHIAPNQADGGEVVQDVNSGVQMVMMEQLDPTLLQMKTEVMEGAVPQETEATVDDTQ 105
A. carolinensis  1    MEGEVVEAIGEESETFIKGKERKTYQRRREGGQEEDVCSMPPNQADGTEVVQDVNTGVQMVMMEQLDPTLLQMKTEVMEGAVQQEAEATVDDTQ  94
X. tropicalis    1    MESEMAEAVVEDSETFMKRKETKTYQRRREGGVDEDNCVIVQSQTDISEVPHDVNSNVQMVMMEQLDPTLLQMKTEVMEGVVSQEGDPTVDDTQ  94
Rhincodon typus  1    MATESENVAVEESEMFTKVKEHKTYQRRREGGQPEDADGQQKESDGAEVAAQDMNSNVQVVMMQPLDPTLLQMKQPVEGG--AHESDT-VDDTQ  91
Danio rerio      1    MEGGPTEAVVEDAGDAFKAKECKTYQRRREGAELLQAAVIEQAQAEVEESVVSVNSSVDMMMMETLDPALLQMKTEVMEAAVGAPVAVTVDDTQ 118
                                                  DEEV            PVVEAQQQLV                                   AGAAHEATVT

H.sapiens CFCTL 1        MAATEISVLSEQFTKIKELELMPEKGLKEEEKDGVCREKDHRSPSELEAERTSGAFQDSVLEEEVELVLAPSEESEKYILTLQTVHFTSEA  91

Homo sapiens     95   IITLQVVNMEEQPINIGELQLVQVPVPVTVPVATTSVEELQGAYENEVSKEGLAESEPMICHTLPLPEGFQVVKVGANGEVETLEQG--ELPPQ 186
Gallus gallus    106  IITLQVVNMEEQPINLGELQLVQVPVPVTVPVATTSVEELQGAYENEVSKGGLQEGEPMICHTLPLPEGFQVVKVGANGEVETLEQG--ELQPQ 197
A. carolinensis  95   IITLQVVNMEEQPINLGELQLVQVPVPVTVPVATTSVEDLQGAYENEVSKGGLQEGEPMICHTLPLPEGFQVVKVGANGEVETLEQ-A-ELQPQ 186
X. tropicalis    95   IITLQVVNMEEQPINLGELQLVQVPV--AVPMATTSVGELHAAFENEVSKEGLQEGEPMICHTLPLPEGFQVVKVGANGEVETLEQ-A-ELQQQ 184
Rhincodon typus  92   IITLQVVNMEEQPLNLGELQLVQVAQ--------SSIDELQGNYENEAPKDGLQEGDPVICHTLPLPEGFQVVKVGANGEVETVEDGAIEVREN 177
Danio rerio      119  IITLQVVNMEEQQLGLGELQLVQVPV-SAVPVTAATVEELQGTLVDATAM--PKDGEPVICHTLPLPEGFQVVKVGANGEVETVEQDEMAEPQN 225
                                                                                                LQPQDDQPPHQEEEEE

H.sapiens CFCTL 92   -VELQDMSLLSIQQQEGVQVVVQQPGPGLLWLEEGPRQSLQQCVAISIQQELYSPQEMEVLQFHALEENVMVASEDSKLAVSLAETTGLIKL   182

Homo sapiens     187  ED--PSWQKDPDYQPPAKKTKKTKKSKLRY-TEEGKDVDVSVYDFEEEQQEGLLSEVNAEKVVGNMKPPKPTKIKKKGV 262
Gallus gallus    198  ED--PNWQKDPDYQPPAKKTKKNKKSKLRY-TEEGKDVDVSVYDFEEEQQEGLLSEVNAEKVVGNMKPPKPTKIKKKGV 273
A. carolinensis  187  ED--PGWQKDPDYQPPAKKTKKTKKSKLRY-TEEGKDVDVSVYDFEEEQQEGLLSEVNAEKVVGNMKPPKPTKIKKKGV 262
X. tropicalis    185  EE--PGWQKDPDYVPPIKKTKKTKKSKLRY-TEEGKDVDVSVYDFEEEQQEGLLSDVNAEKVVGNMKPPKPTKIKKKGV 260
Rhincodon typus  178  EDASAAWQKDPDYQPPVKKVKK-KKNKLRYKVDDSKDVDLSVYDFEEEQQEGLLSEVNVEKAVGTMKPPKPMKIKKKGV 255
Danio rerio      226  ED--PAWSKDPDYTPPVKKVKKTKKSKLRYNTEGDKDMDVSVYDFEEEQQEGLLSEVNAEKVVGNMKPPKPTKIKKKGV 302
H.sapiens CFCTL 183  EE---EQEKNQLLAERTKEQLFFVETMSGDERSDEIVLTVSNSNVEEQEDQPTAGQADAEKA-----KSTKNQRKTKGA 253
```


283 284
339 342
ZF1
ZF2
ZF3
Homo sapiens 263 KKTFQCELCSYTCPRRSNLDRHMKSHTDERPHKCHLCGRAFRTVTLLRNHLNTHTGTRPHKCPDCDMAFVTSGELVRHRRYKHTH 347
Gallus gallus 274 KKTFQCELCSYTCPRRSNLDRHMKSHTDERPHKCHLCGRAFRTVTLLRNHLNTHTGTRPHKCPDCDMAFVTSGELVRHRRYKHTH 358
A. carolinensis 263 KKTFQCELCSYTCPRRSNLDRHMKSHTDERPHKCHLCGRAFRTVTLLRNHLNTHTGTRPHKCPDCDMAFVTSGELVRHRRYKHTH 347
X. tropicalis 261 KKTFQCELCSYTCPRRSNLDRHMKSHTDERPHKCHLCGRAFRTVTLLRNHLNTHTGTRPHKCPDCDMAFVTSGELVRHRRYKHTH 345
Rhincodon typus 256 KKTFQCELCSYTCPRRSNLDRHMKSHTDERPHKCHLCGRAFRTVTLLRNHLNTHTGTKPHKCPDCDMAFVTSGELVRHRRYKHTH 340
Danio rerio 303 KKTFQCELCSYTCPRRSNLDRHMKTHTSEKPHLCHLCLKTFRTVTLLRNHVNTHTGTRPYKCNDCNMAFVTSGELVRHRRYKHTH 387
H.sapiens CFCTL 254 KGTFHCDVCMFTSSRMSSFNRHMKSHTDERPHKCHLCGRAFRTVTLLRNHLNTHTGTRPHKCTDCDMAFVTSGELVRHRRYKHTH 338
354 362 365 377 392
ZF4
ZF5
ZF6
Homo sapiens 348 EKPFKCSMCDYASVEVSKLKRHIRSHTGERPFQCSLCSYASRDTYKLKRHMRTHSGEKPYECYICHARFTQSGTMKMHILQKHTEN 433
Gallus gallus 359 EKPFKCSMCDYASVEVSKLKRHIRSHTGERPFQCSLCSYASRDTYKLKRHMRTHSGEKPYECYICHARFTQSGTMKMHILQKHTEN 444
A. carolinensis 348 EKPFKCSMCDYASVEVSKLKRHIRSHTGERPFQCSLCSYASRDTYKLKRHMRTHSGEKPYECYICHARFTQSGTMKMHILQKHTEN 433
X. tropicalis 346 EKPFKCSMCDYASVEVSKLKRHIRSHTGERPFQCSLCSYASRDTYKLKRHMRTHSGEKPYECYICHARFTQSGTMKMHILQKHTEN 431
Rhincodon typus 341 EKPFKCSMCDYASVEVSKLKRHIRSHTGERPFQCSLCSYASRDTYKLKRHMRTHSGEKPYECYICHARFTQSGTMKMHILQKHTEN 426
Danio rerio 388 EKPFKCSMCDYASVEVSKLKRHIRSHTGERPFQCSLCSYASRDTYKLKRHMRTHSGEKPYECYICHARFTQSGTMKMHILQKHTEN 473
H.sapiens CFCTL 339 EKPFKCSMCKYASVEASKLKRHVRSHTGERPFQCCQCSYASRDTYKLKRHMRTHSGEKPYECHICHTRFTQSGTMKIHILQKHGEN 424
448 450 454. 467 470 490 494 508
ZF7
ZF8
ZF9
Homo sapiens 434 VAKFHCPHCDTVIARKSDLGVHLRKQHSYIEQGKKCRYCDAVFHERYALIQHQKSHKNEKRFKCDQCDYACRQERHMIMHKRTHTG 519
Gallus gallus 445 VAKFHCPHCDTVIARKSDLGVHLRKQHSYIEQGKKCRYCDAVFHERYALIQHQKSHKNEKRFKCDQCDYACRQERHMVMHKRTHTG 530
A. carolinensis 434 VAKFHCPHCDTVIARKSDLGVHLRKQHSYIEQGKKCRYCDAVFHERYALIQHQKSHKNEKRFKCDQCDYACRQERHMIMHKRTHTG 519
X. tropicalis 432 VAKFHCPHCDTVIARKSDLGVHLRKQHSYIEQGKKCRYCDTVFHERYALIQHQKSHKNEKRFKCDQCEYACRQERHMIMHKRTHTG 517
Rhincodon typus 427 VAKFHCPHCDTVIARKSDLGVHLRKQHSYLESGKKCRYCDAVFHERYALIQHQKSHKNEKRFKCELCDYACKQERHMVMHKRTHTG 512
Danio rerio 474 VAKFHCPHCDTVIARKSDLGVHLRKQHSYIEQGRKCRYCDAVFHERYALIQHQKSHKNEKRFKCDQCDYACRQERHMVMHKRTHTG 559
H.sapiens CFCTL 425 VPKYQCPHCATIIARKSDLRVHMRNLHAYSAAELKCRYCSAVFHERYALIQHQKTHKNEKRFKCKHCSYACKQERHMTAHIRTHTG 510
536 566
ZF10
ZF11
Homo sapiens 520 EKPYACSHCDKTFRQKQLLDMHFKRYHDPNFVPAAFVCSKCGKTFTRRNTMARHADNCAGPDGVEGENG 588
Gallus gallus 531 EKPYACSHCDKTFRQKQLLDMHFKRYHDPNFVPAAFVCSKCGKTFTRRNTMARHADNCSGLDGGEGENG 599
A. carolinensis 520 EKPYACSHCDKTFRQKQLLDMHFKRYHDPNFVPAAFVCSKCGKTFTRRNTMARHADNCTGPDGVEGENG 588
X. tropicalis 518 EKPYACSHCDKTFRQKQLLDMHFKRYHDPSFVPAAFVCSKCGKTFTRRNTMSRHADNCTGPDGTDGENGGESE 590
Rhincodon typus 513 EKPYSCSQCDKTFRQKQLLDMHFKRYHDPNFIPATFVCTKCGKAFTRKNTMTKHAENCSGPGGEEGDAV 581
Danio rerio 560 EKPYACSQCEKTFRQKQLLDMHFRRYHDPNFVPTSFVCTKCGKTFTRRNTMARHAENCTGMDSADGENG 628
H.sapiens CFCTL 511 EKPFTCLSCNKCFRQKQLLNAHFRKYHDANFIPTVYKCSKCGKGFSRWINLHRHSEKCGSGEAKSAASG 579



Numbering above the sequence indicates residues receiving particular attention in the text, and refer to the human ortholog.
Letters in **bold red** are positions making base-specific contacts that are substituted in cases of human CTCF-related disorder; all are fully-conserved, even in CFCTL (shown for comparison).
Grey shading is where ≥5 of the six sequences are identical.
The ZFs are shaded yellow, with substitutions in cyan, with Zn-coordinating residues in bold.
The highly-conserved Met in white on black (M) in ZF3 indicates the initiation codon for CTCF isoform 2 (NP_001177951.1)
The white on black Thr-Ala (T-A) indicates two residues missing in CTCF isoform 3 (NP_001350845.1).

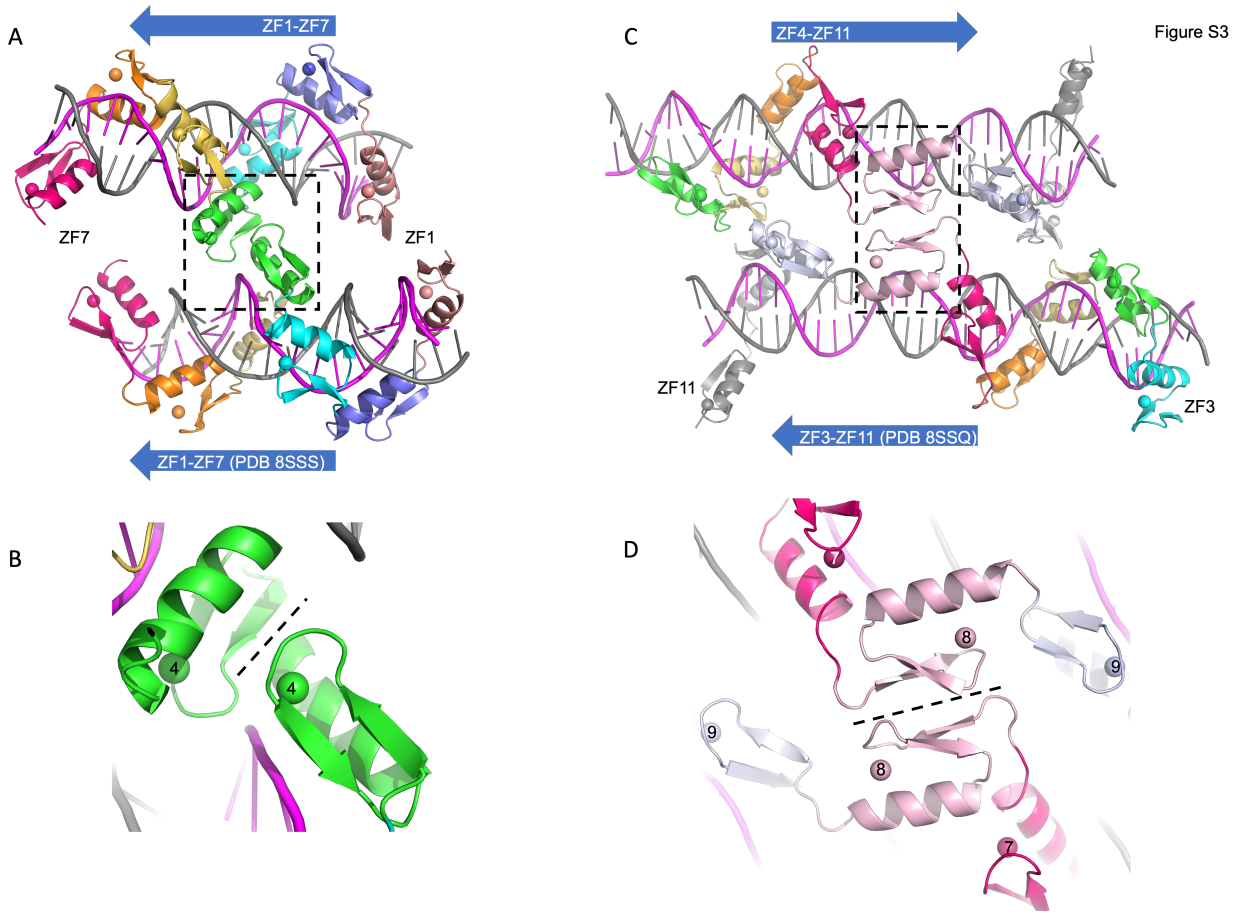


Figure S3. (**A**) Crystal of ZF1-7 in complex with DNA, with two protein-DNA complexes per crystallographic asymmetric unit. The two complexes are arranged in parallel. (**B**) The interface of two molecule is mainly mediated by ZF4 involving the hair loop between the antiparallel strands. (**C**) Crystal of ZF3-11 in complex with DNA, with two protein-DNA complexes per crystallographic asymmetric unit. One of the complexes has a missing ZF3 (top). The two complexes are arranged in antiparallel. (**D**) The interface of the two complexes is mainly mediated by hairpin β-strands of ZF8 (related to Figure 8A).

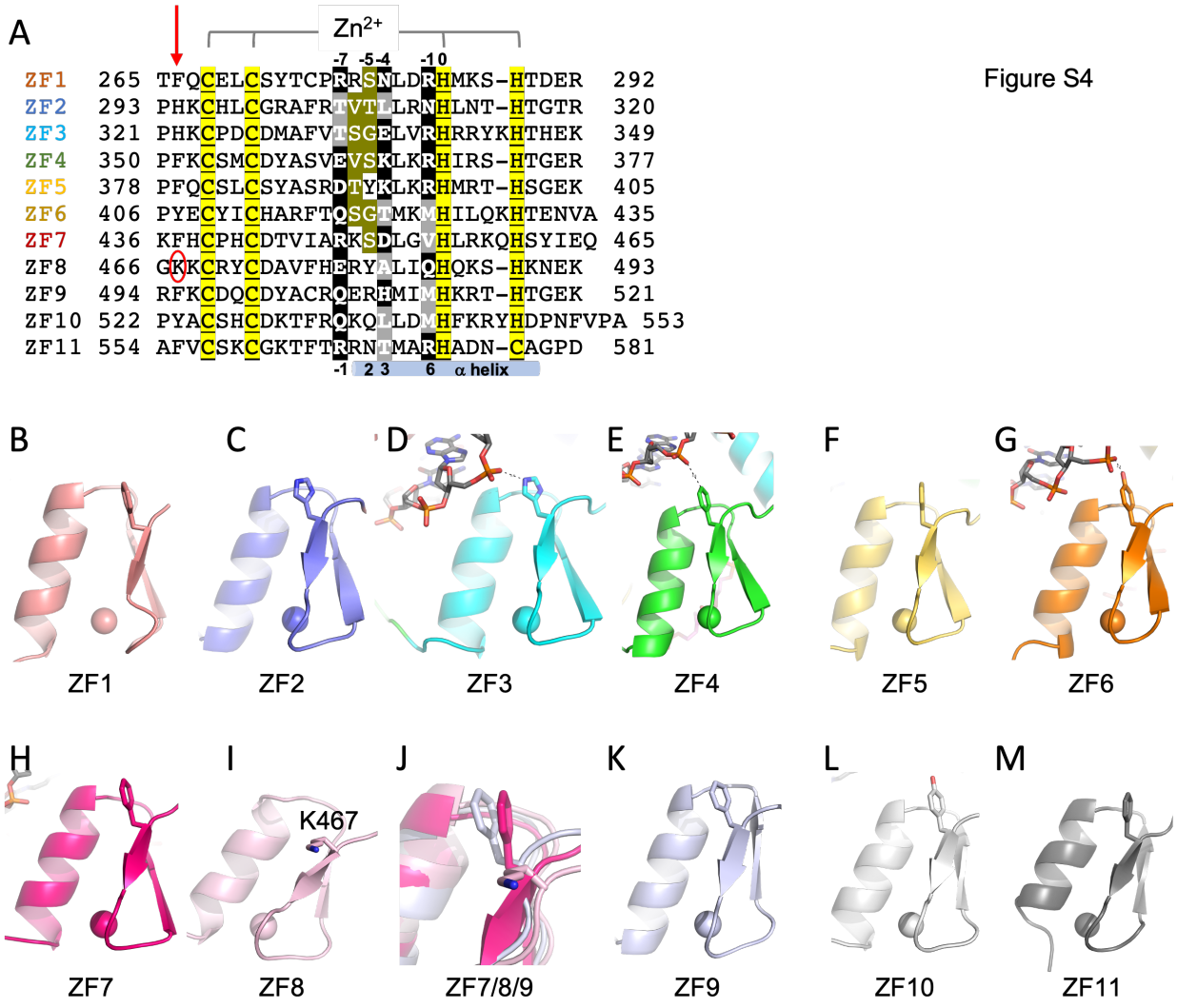


Figure S4

Figure S4. (A) In the corresponding position to Lys467 of ZF8 (indicated by a red arrow), all the other ten fingers in CTCF have a Phe, Tyr or His as the first residue of the first β strand. (**B-H** and **K-M**) This side chain packs between the β strand and the helix and provides stability for the finger. (**D-E** and **G**) In three fingers (ZF3, ZF4 and ZF6), when the finger gets deep into the DNA major groove, the corresponding His, Phe or Tyr provides an H-bond or van der Waals contact to the phosphate group of the non-recognition strand. (**I**) Lys467 of ZF8 points to DNA. (**J**) Superimposition of three figures, ZF7, ZF8 and ZF9.

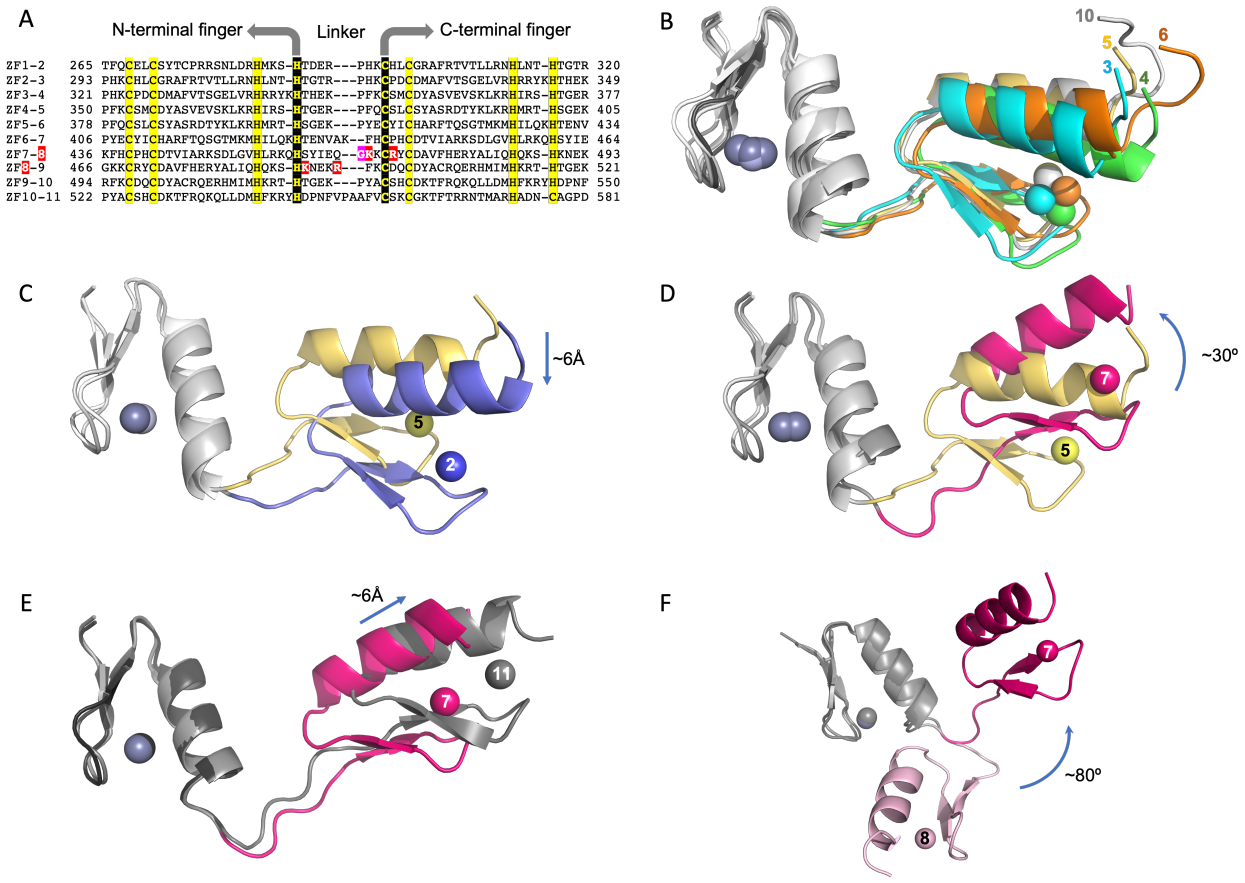


Figure S5. Effect of inter-finger linker length on the relative conformation of associated fingers. (**A**) we examined the effect of inter-finger linker on the relative positions of two neighboring fingers. We generated ten pairs of two-finger element. Seven pairs have the inter-finger length of seven residues between the last Zn-ligand His of the N-terminal finger and the first Zn-ligand Cys of the C-terminal finger (ZF1-2, ZF2-3, ZF3-4, ZF4-5, ZF5-6, ZF8-9 and ZF9-10), two pairs have the length of eight residues (ZF6-7 and ZF7-8), and one pair has ten residues between ZF10 and ZF11. (**B**) We superimposed the corresponding N-terminal fingers and resulting C-terminal fingers adopted at least four varied conformations. First, five pairs of seven-residue linker have similar conformations of the linker as well as the C-terminal finger, including ZF2-3, ZF3-4, ZF4-5, ZF5-6 and ZF9-10. Both fingers of these five pairs are involved in intimate DNA base contacts, as discussed in text (Figures 2 and 7). (**C**) Using ZF4-5 as a

reference, superimposition of ZF1-2 onto ZF4-5 resulted in a movement of ZF2 as well as the linker, away from the DNA base interface. Though this movement (~6Å) was sufficient to distance the base-interacting residues but till placed ZF2 in the DNA major grove. (**D**) Superimposing ZF6-7, which has eight-residue linker, onto ZF4-5, seen the rotation of ZF7 helix of ~30º in reference to ZF5 helix. (**E**) Further increasing the linker distance to ten residues between ZF10 and ZF11 placed ZF11 additional ~6Å away, which might correlate to the observation that the base-interacting residues at -1 and -4 positions of ZF11 are farthest away from their corresponding DNA bases. (**F**) Superimposition of ZF7-8 and ZF6-7, where both pairs include an eight-residue inter-finger linker, positioned ZF7 and ZF8 in opposite directions. The two conformations are approximately ~80º of rotation apart. Apparently, the inter-finger length alone is not the primary reason that that ZF8 spans the minor groove. As noted in the text, the linkers prior to and after ZF8 harbor unique basic residues (white letters against red background in panel A) which form interactions with DNA phosphate groups.

References

1. R. Jothi, S. Cuddapah, A. Barski, K. Cui, K. Zhao, Genome-wide identification of in vivo protein-DNA binding sites from ChIP-Seq data. *Nucleic Acids Res* **36**, 5221-5231 (2008).
2. H. S. Rhee, B. F. Pugh, Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. *Cell* **147**, 1408-1419 (2011).
3. H. Wang *et al.*, Widespread plasticity in CTCF occupancy linked to DNA methylation. *Genome Res* **22**, 1680-1688 (2012).
4. M. J. MacPherson, L. G. Beatty, W. Zhou, M. Du, P. D. Sadowski, The CTCF insulator protein is posttranslationally modified by SUMO. *Mol Cell Biol* **29**, 714-725 (2009).
5. D. Farrar *et al.*, Mutational analysis of the poly(ADP-ribosyl)ation sites of the transcription factor CTCF provides an insight into the mechanism of its regulation by poly(ADP-ribosyl)ation. *Mol Cell Biol* **30**, 1199-1216 (2010).
6. T. Sekiya, K. Murano, K. Kato, A. Kawaguchi, K. Nagata, Mitotic phosphorylation of CCCTC-binding factor (CTCF) reduces its DNA binding activity. *FEBS Open Bio* **7**, 397-404 (2017).
7. H. G. Valverde de Morales *et al.*, Expansion of the genotypic and phenotypic spectrum of CTCF-related disorder guides clinical management: 43 new subjects and a comprehensive literature review. *Am J Med Genet A* **191**, 718-729 (2023).
8. J. Li *et al.*, An alternative CTCF isoform antagonizes canonical CTCF occupancy and changes chromatin architecture to promote apoptosis. *Nat Commun* **10**, 1535 (2019).

Table S1. Summary of X-ray data collection and refinement statistics, beamline 22-ID (APS) and 1.0000Å wavelength

Protein Construct	ZF3-ZF11	ZF3-ZF11	ZF1-ZF7	ZF1-ZF7 (K365T)	ZF3-ZF11
DNA (see Table S2)	35-4	35-20	23-bp	23-bp	19-bp
PDB code	8SSQ	8SSR	8SSS	8SST	8SSU
Date Collected	07-17-20, 08-07-20, 09-26-20	07/10/22, 11/7/2022	09-26-2020	03-24-2021	02/06/2020
Space group	*C*2	*C*2	*C*2	*C*2	$P3_221$
Cell dimensions (Å)	352.1, 67.8, 61.2	362.2, 68.1, 62.0	161.0, 41.2, 135.7	160.6, 41.6, 135.2	80.5, 80.5, 187.6
α, β, γ (°)	90, 92.6, 90	90, 94.9, 90	90, 105.2, 90	90 ,105.6, 90	90, 90, 120
Resolution (Å)	41.86-3.12 (3.23-3.12)	37.59-3.14 (3.26-3.14)	43.18-2.30 (2.37-2.30)	40.17-2.19 (2.27-2.19)	38.34-2.89 (2.99-2.89)
[a] R_{merge}	0.398 (2.173)	0.254 (1.068)	0.090 (0.989)	0.121 (1.315)	0.203 (2.56)
R_{pim}	0.076 (0.703)	0.061 (0.338)	0.038 (0.649)	0.053 (0.753)	0.065 (0.860)
$CC_{1/2}$	0.985 (0.397)	0.940 (0.883)	1.000 (0.513)	0.992 (0.387)	0.974 (0.706)
[b] <I/σI>	11.0 (1.2)	10.9 (1.1)	17.0 (1.1)	13.3 (1.0)	12.8 (1.5)
Completeness (%)	100.0 (99.8)	95.7 (73.6)	94.0 (74.4)	97.2 (87.4)	99.9 (100.0)
Redundancy	27.1 (11.5)	15.8 (7.1)	6.0 (2.5)	5.8 (3.2)	10.3 (9.7)
Observed reflections	701,822	391,247	220,378	250,177	168,662
Unique reflections	25,877 (2,583)	24,779 (1,894)	36,617 (2,757)	43,468 (3,842)	16,333 (1,605)
Refinement					
Resolution (Å)	3.12	3.14	2.30	2.19	2.89
No. reflections	25,644	24,639	35,515	43,313	16,037
[c] R_{work} / [d] R_{free}	0.238 / 0.268	0.253 / 0.285	0.205 / 0.231	0.220 / 0.243	0.205 / 0.240
No. Atoms					
Protein	4051	4051	3150	3140	1,888
DNA	2856	2856	1874	1874	802
Zn	17	17	14	14	8
Solvent	15	2	111	121	14
B Factors (Å^2)					
Protein	177.9	190.6	68.7	85.9	97.0
DNA	187.5	195.1	70.5	88.7	83.1
Zn	185.7	238.5	69.4	86.7	102.2
Solvent	100.7	144.4	49.2	54.8	73.2
R.m.s. deviations					
Bond lengths (Å)	0.003	0.003	0.003	0.003	0.003
Bond angles (°)	0.5	0.5	0.5	0.5	0.5

* Values in parenthesis correspond to highest resolution shell.

[a] $R_{merge} = \Sigma \mid I - \langle I \rangle \mid / \Sigma I$, where I is the observed intensity and <I> is the averaged intensity from multiple observations.

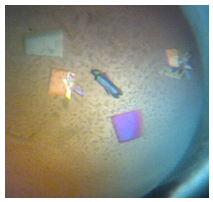
[b] <I/σI> = averaged ratio of the intensity (I) to the error of the intensity (σI).

[c] $R_{work} = \Sigma \mid Fobs - Fcal \mid / \Sigma \mid Fobs \mid$, where Fobs and Fcal are the observed and calculated structure factors, respectively.

[d] R_{free} was calculated using a randomly chosen subset (5%) of the reflections not used in refinement.

Table S2. Oligonucleotides used in the study

Residues	name	pXC#	DNA oligos for crystallization	Crystallization conditions
263-465	ZF1-ZF7	1564	23-bp: 5′-G CCA GCA GGG GGC GCT AGT GAG G-3′ 3′-C GGT CGT CCC CCG CGA TCA CTC C-5′	260 mM DL-Malic acid 7.0 26.5% PEG 3350
263-465	K365T	2332	23-bp: 5′-G CCA GCA GGG GGC GCT AGT GAG G-3′ 3′-C GGT CGT CCC CCG CGA TCA CTC C-5′	200 mM ammonium acetate 100 mM BIS-TRIS pH 5.5 25% (w/v) PEG 3350
321-581	ZF3-ZF11	1566	35-4: 5′-GTG CAG TACC ACATTTAA CCA GCA GGG GGC GCT AA-3′ 3′-CAC GTC ATGG TGTAAATT GGT CGT CCC CCG CGA TT-5′	200 mM sodium malonate pH 5.0 20% (w/v) PEG 3350
			35-20: 5′-GTG CAG TACC ACATTTAA CCA GCA GG**T** GGC GCT AA-3′ 3′-CAC GAC TTGG TGTAAATT GGC CGT CCA CCG CGA TT-5′	20% (w/v) PEG 3350 100 mM BIS-TRIS pH 5.0 50 mM NaCl
			19-bp: 5′-TGC GCC CCC TGC TGG TCC T-3′ 3′-ACG CGG GGG ACG ACC AGG A-5′	40 mM citric acid, 60 mM BIS-TRIS propane pH 6.4 20% (w/v) PEG 3350
			Primers for mutagenesis:	
	K365T		FP: 5′-GAA GTC AGC ACA TTA AAA CGT CAC ATT CG-3′ RP: 5′-CGT TTT AAT GTG CTG ACT TCT ACA CTG GC-3′	
			For FP binding:	
			5'-A GGA CCA GCA GGG G**X**C GCA-3' 3'-T CCT GGT CGT CCC C**Y**G CGT-5'-FAM X:Y=G:C, A:T, T:A, C:G	



ZF3-ZF11 + oligo 35-4

ZF3-ZF11 + 19 bp