

Table S1. List of high-confidence data listed by sources.

Source:	Positives	Negatives	Kd<200	Kd>200	Comparat.	Total	Ref.
DBSFB01	28	292	5	1	105	431	[1]
DFSLBHKB05	88	133	0	0	200	421	[2]
SDBB99	69	103	25	0	294	491	[3]
DSB00	70	388	0	0	351	809	[4]
BMB03	9	27	7	12	0	55	[5]
BFS02	1024	0	0	0	0	1024	[6]
BJC02	0	0	25	99	0	124	[7]
BKSHRP03	0	0	33	0	0	33	[8]
CGU99	5	0	0	0	0	5	[9]
CK94a	67	0	0	0	0	67	[10]
CK94b	0	9	10	0	0	19	[11]
DB92	0	0	0	0	9	9	[12]
DB93	0	1	5	3	0	9	[13]
GP97	21	0	3	0	0	24	[14]
ICK97	0	0	4	0	0	4	[15]
IKC01	0	0	7	0	42	49	[16]
JKW94	0	0	18	0	0	18	[17]
KFM05	40	0	0	0	0	40	[18]
LXC02	32	0	0	0	0	32	[19]
NGC92	0	0	40	0	0	40	[20]
PDB*	16	0	0	0	0	16	[21]
RP94	0	0	12	0	0	12	[22]
RUMIWCKC03	0	0	8	0	0	8	[23]
TB90	11	0	0	0	0	11	[24]
WGRP99	0	0	6	0	0	6	[25]
WYB95	0	0	24	2	0	26	[26]
Sum:	1480	953	232	117	1001	3783	
Filtered:	1086	835	226	117	914	3178	

\* List of co-crystal structures obtained from PDB: 1A1F, 1A1G, 1A1H, 1A1I, 1A1J, 1A1K, 1A1L, 1AAY, 1G2D, 1G2F, 1JK1, 1JK2, 1MEY, 1P47, 1ZAA, 1F2I, 1LLM, 2DRP, 1UBD, and 2GLI.

1. Dreier, B., et al., *Development of zinc finger domains for recognition of the 5'-ANN-3' family of DNA sequences and their use in the construction of artificial transcription factors.* J Biol Chem, 2001. **276**(31): p. 29466-78.
2. Dreier, B., et al., *Development of zinc finger domains for recognition of the 5'-CNN-3' family DNA sequences and their use in the construction of artificial transcription factors.* J Biol Chem, 2005. **280**(42): p. 35588-97.

3. Segal, D.J., et al., *Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5'-GNN-3' DNA target sequences*. Proc Natl Acad Sci U S A, 1999. **96**(6): p. 2758-63.
4. Dreier, B., D.J. Segal, and C.F. Barbas, 3rd, *Insights into the molecular recognition of the 5'-GNN-3' family of DNA sequences by zinc finger domains*. J Mol Biol, 2000. **303**(4): p. 489-502.
5. Blancafort, P., L. Magnenat, and C.F. Barbas, 3rd, *Scanning the human genome with combinatorial transcription factor libraries*. Nat Biotechnol, 2003. **21**(3): p. 269-74.
6. Benos, P.V., A.S. Lapedes, and G.D. Stormo, *Probabilistic code for DNA recognition by proteins of the EGR family*. J Mol Biol, 2002. **323**(4): p. 701-27.
7. Bulyk, M.L., P.L. Johnson, and G.M. Church, *Nucleotides of transcription factor binding sites exert interdependent effects on the binding affinities of transcription factors*. Nucleic Acids Res, 2002. **30**(5): p. 1255-61.
8. Bae, K.H., et al., *Human zinc fingers as building blocks in the construction of artificial transcription factors*. Nat Biotechnol, 2003. **21**(3): p. 275-80.
9. Cook, T., B. Gebelein, and R. Urrutia, *Sp1 and its likes: biochemical and functional predictions for a growing family of zinc finger transcription factors*. Ann N Y Acad Sci, 1999. **880**: p. 94-102.
10. Choo, Y. and A. Klug, *Toward a code for the interactions of zinc fingers with DNA: selection of randomized fingers displayed on phage*. Proc Natl Acad Sci U S A, 1994. **91**(23): p. 11163-7.
11. Choo, Y. and A. Klug, *Selection of DNA binding sites for zinc fingers using rationally randomized DNA reveals coded interactions*. Proc Natl Acad Sci U S A, 1994. **91**(23): p. 11168-72.
12. Desjarlais, J.R. and J.M. Berg, *Toward rules relating zinc finger protein sequences and DNA binding site preferences*. Proc Natl Acad Sci U S A, 1992. **89**(16): p. 7345-9.
13. Desjarlais, J.R. and J.M. Berg, *Use of a zinc-finger consensus sequence framework and specificity rules to design specific DNA binding proteins*. Proc Natl Acad Sci U S A, 1993. **90**(6): p. 2256-60.
14. Greisman, H.A. and C.O. Pabo, *A general strategy for selecting high-affinity zinc finger proteins for diverse DNA target sites*. Science, 1997. **275**(5300): p. 657-61.
15. Isalan, M., Y. Choo, and A. Klug, *Synergy between adjacent zinc fingers in sequence-specific DNA recognition*. Proc Natl Acad Sci U S A, 1997. **94**(11): p. 5617-21.
16. Isalan, M., A. Klug, and Y. Choo, *A rapid, generally applicable method to engineer zinc fingers illustrated by targeting the HIV-1 promoter*. Nat Biotechnol, 2001. **19**(7): p. 656-60.
17. Jamieson, A.C., S.H. Kim, and J.A. Wells, *In vitro selection of zinc fingers with altered DNA-binding specificity*. Biochemistry, 1994. **33**(19): p. 5689-95.
18. Kaplan, T., N. Friedman, and H. Margalit, *Ab initio prediction of transcription factor targets using structural knowledge*. PLoS Comput Biol, 2005. **1**(1): p. e1.
19. Liu, Q., et al., *Validated zinc finger protein designs for all 16 GNN DNA triplet targets*. J Biol Chem, 2002. **277**(6): p. 3850-6.
20. Nardelli, J., T. Gibson, and P. Charnay, *Zinc finger-DNA recognition: analysis of base specificity by site-directed mutagenesis*. Nucleic Acids Res, 1992. **20**(16): p. 4137-44.
21. Berman, H.M., et al., *The Protein Data Bank*. Nucleic Acids Res, 2000. **28**(1): p. 235-42.
22. Rebar, E.J. and C.O. Pabo, *Zinc finger phage: affinity selection of fingers with new DNA-binding specificities*. Science, 1994. **263**(5147): p. 671-3.
23. Reynolds, L., et al., *Repression of the HIV-1 5' LTR promoter and inhibition of HIV-1 replication by using engineered zinc-finger transcription factors*. Proc Natl Acad Sci U S A, 2003. **100**(4): p. 1615-20.

24. Thiesen, H.J. and C. Bach, *Target Detection Assay (TDA): a versatile procedure to determine DNA binding sites as demonstrated on SPI protein*. Nucleic Acids Res, 1990. **18**(11): p. 3203-9.
25. Wolfe, S.A., et al., *Analysis of zinc fingers optimized via phage display: evaluating the utility of a recognition code*. J Mol Biol, 1999. **285**(5): p. 1917-34.
26. Wu, H., W.P. Yang, and C.F. Barbas, 3rd, *Building zinc fingers by selection: toward a therapeutic application*. Proc Natl Acad Sci U S A, 1995. **92**(2): p. 344-8.