Risk assessment of Cytotoxicity of DNA binding proteins

Christian Bach, Hassan Bajwa, Kris Erodula, Prabir Patra
Department of Biomedical Engineering
University of Bridgeport, Bridgeport, CT

Introduction

The Role of Transcription Factors

For clinical applications the biological functions of DNA-binding proteins require that they interact with their target binding site with high affinity and specificity. Advances in randomized production and target-oriented selection of engineered artificial DNA binding domains incited a rapidly expanding field of designer transcription factors (TFs). Zinc-finger binding domains fabricated via modular assembly display an unexpectedly high failure rate having either a lack of activity as ZNFs in human cells or activity at “off-target” binding sites on the human genome causing cell death. To address these shortcomings we created new binding domains using a targeted modification strategy. Instead of tethering single fingers together to assemble zinc finger binding domains, we exchanged amino acid residues in the alpha-helical region of the transcription factor Sp1. To illustrate the technique, we produced two Sp1 mutants and identified their best target binding sites. Our research concludes that the biological functions of DNA-binding proteins require factors and mechanisms beyond the ones of affinity and specificity.

Alterations in 2nd finger of Sp1

Figure 1 shows the three finger binding domain of Sp1. The underlined/blue amino acids represent the alpha-helical region in the second finger. Figure 2 and Figure 3 show the altered amino acids of CB1 and MR14 in italic/red.

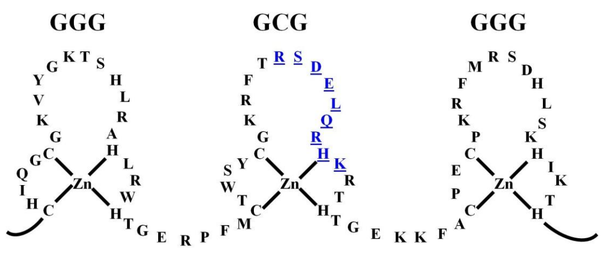

Figure 1: Three finger of Sp1 binding domain and consensus binding site GGGGCGGGG

Sp1-mutants CB1 and MR14

With site directed PCR mutagenesis (Thiesen & Bach, 1991) Sp1 mutants CB1 (Figure 4) and MR14 (Figure 5) were created.

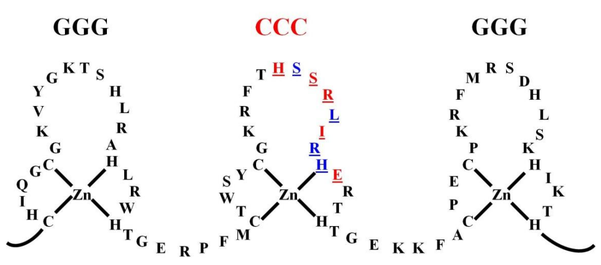

Figure 2: Three finger of CB1 binding domain and best binding site GGGCCCGGG

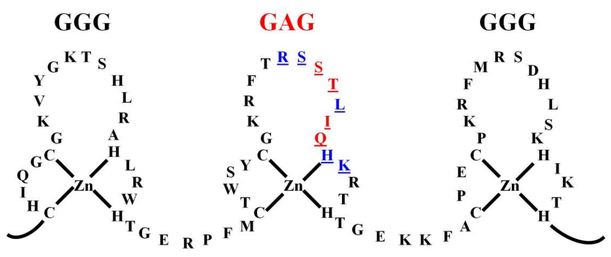

Figure 3: Three finger of MR14 binding domain and best binding site GGGGAGGGG

The new binding domains were produced via a targeted modification strategy that differs from the modular assembly strategy in the way that instead of tethering single fingers together to assemble zinc finger arrays, we exchanged amino acid residues in the alpha-helical region of the transcription factor Sp1.

3finger/9bp Paradox and Paradigms of DNA/protein Interactions

Our findings shown in Figure 4 display the DNA recognition of the 2nd finger changes from GCG for Sp1 to CCC for CB1 and GAG for MR14. It indicates that the framework of the Sp1 binding domain has the structural stability and flexibility that can be employed for a targeted modification strategy.

We further investigate the feasibility to use three finger binding domains for clinical applications because it is known that engineered three finger domains display cytotoxicity when used in zinc finger nuclease (ZNF). Cytotoxicity is linked to the ability of engineered DNA binding proteins to bind at the same or similar binding sites throughout the genome and exert influence on cells functions. Such binding events at “off-target” sites can cause cell death (Cathomen, 2009). We searched the NCBI HuRef genome for exact matches 3 finger / 9 base pair binding sites.

The results are displayed in Table 1. We found a high frequency of nine base pair target binding sites for three finger binding domains throughout the human genome, which is a significant imbalance between the number of genes a natural zinc finger protein regulates and the actual number of locations that exist on the human genome. In the current predominant notion, affinity and specificity (E. J. Rebar & Pabo, 1994) are the factors that determine binding, which consequently means that Sp1 binds to nearly twenty thousand locations on the human genome.

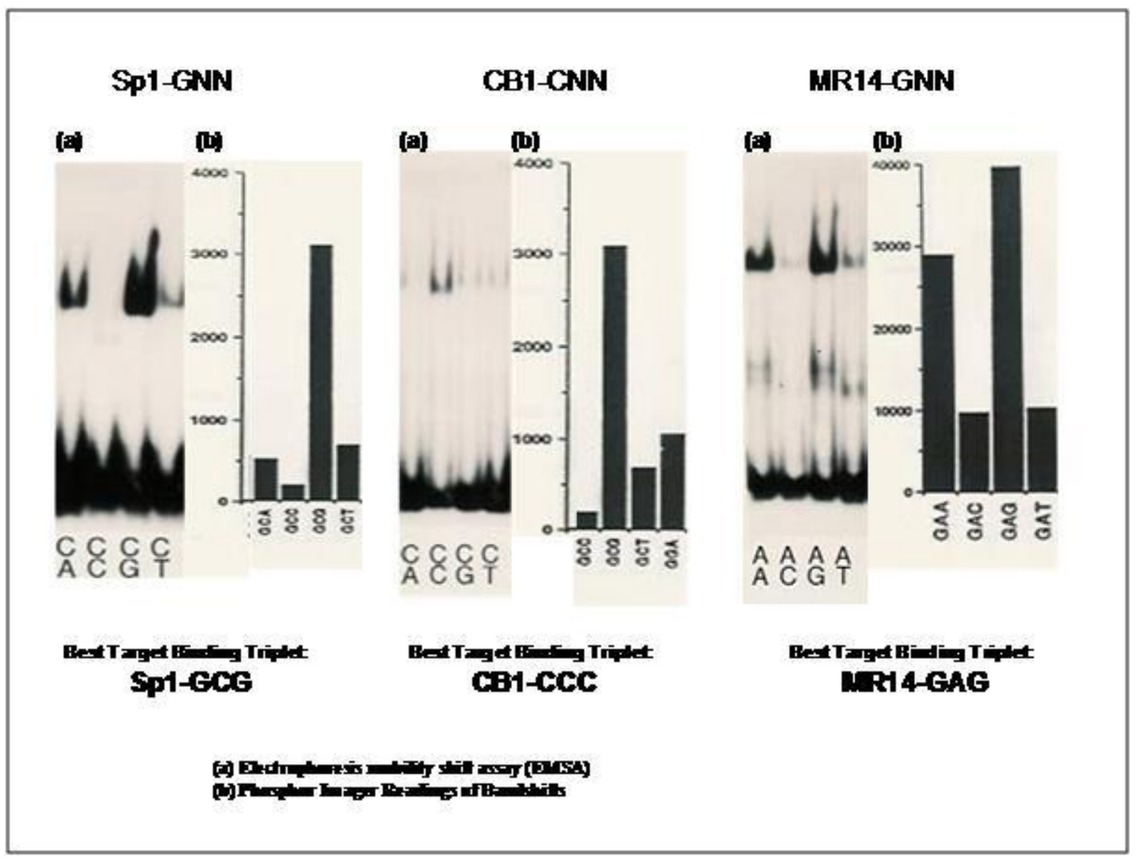

Figure 4: Three Three helical region of the second sp1 finger, RSDELQRH, was replaced by the amino acids KSSALISH creating zinc finer mutant CB1 by performing PCR based site directed mutagenesis. The non conserved amino acids, arginine (R), aspartic acid (D), glutamic acid (E), glutamine (Q) and arginine (E) at zinc finger positions 15,17,18,20 and 21, were replaced by lysine (K), serine (S), alanine (A), isoleucinc (D) and serine (S) (Thiesen et al. 1991). Additional 30 sp1 MQ- and MR- mutants have been created. Recombinant mutated sp1 proteins were purified by FPLC Mono S chromatography. The DNA binding activity of sp1 and mutated sp1 proteins was assessed by incubating the 64 labeled double-stranded oligonucleotides (n = 3) with chimaeric proteins CB1 and MR14 and performing electrophoretic mobility shift assays (EMSA) (Thiesen & Bach, 1991).

The 3finger/9bp Paradox

A more feasible strategy for nature to avoid cytotoxicity would be the use of a six finger domain structure binding an eighteen base pair DNA sequence that with high probability occurs only a single time on the genome. Therefore, it is a 3finger/9bp paradox that the evolutionary process came up with a three finger domain solution that will bind to 5,000 – 50,000 identical DNA sequences on the human genome.

	GGGGCGGGG (Sp1 consensus)	GGGCGGGGG (Sp1 Best)	GGGCCCGGG (CB1 Best)	GGGGAGGGG (MR14 Best)	CGGCCCCAG (MA1 Best)
chr1	1667	1281	487	4461	392
chr2	1279	1077	360	3685	319
chr3	959	795	250	2857	219
chr4	703	561	199	2430	186
chr5	798	683	213	2620	198
chr6	918	783	216	2710	232
chr7	967	771	293	2586	290
chr8	722	575	192	2283	193
chr9	830	669	259	2173	226
chr10	822	634	260	2422	212
chr11	1078	758	311	2757	238
chr12	882	735	242	2421	207
chr13	422	337	100	1186	98
chr14	669	547	162	1609	141
chr15	553	429	175	1483	149
chr16	898	618	295	1779	255
chr17	1077	756	317	2203	286
chr18	352	294	103	1049	77
chr19	1555	887	445	2145	280
chr20	596	504	227	1418	149
chr21	244	182	91	595	79
ch22	567	420	196	1158	145
chX	757	624	161	2819	147
chY	30	21	4	119	7
Total	19345	14941	5558	50968	4725

Table 1: List of exact matches on each chromosome for each of the 9bp binding target sites of Sp1, CB1, MR14.

If sequence specific DNA binding of a zinc finger is solely based on specificity and affinity (Cornu et al., 2007), but do not control engineered zinc fingers, Sp1 is prone to indiscriminately bind all twenty thousand identical consensus sequences (GGGGCGGGG) on the human genome and would probably cause some undesired interactions. Because Sp1 induced negative effects are not observed, leads us to the conclusion that other factors (besides specificity and affinity) and location specific mechanisms (besides sequence specific) that control recognition play a role. Three possible targets could be involved in controlling location specific binding of Sp1:

1. Regulation of the whole zinc finger protein
2. Regulation of the binding domain
3. Regulation of the binding sites

Regulatory Cell Mechanisms

Our findings indicate the existence of yet unknown regulatory mechanisms in the cell to control binding natural zinc fingers. It seems that such regulatory cell mechanisms do not function with artificial engineered zinc finger binding domains. This leads to the strong argument that with targeted modification of natural zinc fingers the function of the regulatory cell apparatus might be preserved and the apparatus would be functioning at new target binding sites. It also might require tagging new target sites with factors that are recognized by regulatory elements. We summarize our findings with seven predominant and seven complementary paradigms to help manufacture improved zinc finger proteins for clinical applications.

Predominant Paradigms	Complementary Paradigms
1. Paradigm of **equal functionality** of each zinc finger	Paradigm of **preservation** of evolutionary trades inherent in natural zinc fingers
2. Paradigm of **independency**	Paradigm of **interdependency**
3. Paradigm of **modular assembly**	Paradigm of **targeted modifications**
4. Paradigm of **randomness**	Paradigm of **evolution**
5. Paradigm of **overlap problem**	Paradigm of **precision**
6. Paradigm of **best target sites**	Paradigm of **complete recognition code**
7. Paradigm of **clinical compatibility**	Paradigm of **complete “off-target” cleavage pattern**

Table 2: List of resulting paradigms.

Conclusion

Our results show that Sp1 has a robust three-dimensional structure and can be used for the 3x2 strategy (Klug, 2010). We propose to use targeted modification as strategy to gradually alter Sp1 to recognize new GC and AT rich target binding sites.

Oka, S., Shiraishi, Y., Yoshida, T., Ohkubo, T., Sugiura, Y., and Kobayashi, Y. "NMR Structure of Transcription Factor Sp1 DNA Binding Domain," *Biochemistry* (43:51) 2004, pp 16027-16035.
Klug, A. (2010). The Discovery of Zinc Fingers and Their Applications in Gene Regulation and Genome Manipulation. *Annual Review of Biochemistry, 79*, 213-231.
Cathomen, T. (2009). Zinc-finger nucleases: finding the balance between activity and toxicity. *Human Gene Therapy, 20*(11), 1356.
Thiesen, H.-J., and Bach, C. "Determination of DNA Binding Specificities of Mutated Zinc Finger Domains," *FEBS LETTERS* (283:1) 1991, pp 23-26.
Rebar, E. J., & Pabo, C. O. (1994). Zinc finger phage: affinity selection of fingers with new DNA-binding specificity. *Science, 263*, 671-673.
Cornu, T. I., Thibodeau-Beganny, S., Guhl, E., Alwin, S., Eichtinger, M., Joung, J., et al. (2007). DNA-binding Specificity Is a Major Determinant of the Activity and Toxicity of Zinc-finger Nucleases. *Molecular Therapy, 16*(2), 352-358.