行政院國家科學委員會專題研究計畫　成果報告

以噬菌體基因庫選殖對酵母菌基因啟動子有專一性的鋅指蛋白建構重組轉錄因子來探討對基因表現和外表型的影響

計畫類別：個別型計畫
計畫編號：NSC92-2311-B-040-007-
執行期間：92 年 08 月 01 日至 93 年 07 月 31 日
執行單位：中山醫學大學生命科學系

計畫主持人：謝家慶

報告類型：精簡報告

處理方式：本計畫涉及專利或其他智慧財產權，2 年後可公開查詢

中　華　民　國 93 年 11 月 2 日

行政院國家科學委員會專題研究計畫報告

計畫編號：計畫編號：NSC 92－2311－B－040 －007－
執行期限：92 年 08 月 01 日至 93 年 07 月 31 日
主持人：謝家慶　中山醫學院生命科學系
計畫參與人員：鄭予哲，呂詩偉，劉乃甄，呂坤松，林盈君，王仲安

一、中文摘要

Cys_2His_2 鋅指結構域（domain）乃真核生物蛋白質中最常見的去氧核醣核酸結合基序（motif），可辨識多種不同組合的去氧核醣核酸序列。鋅指蛋白模組（module）因其結構的穩定性可於蛋白質工程中做為多元變異的台架。轉錄因子 Zif268 的三指模組已被用於構築由噬菌體呈現之組合基因庫的框架來闡釋鋅指 ── DNA 相互辨識的理論。近年來研究者開發對鋅指選擇及設計的方法已可構築幾乎能標的於雙螺旋 DNA 任何位置的鋅指蛋白質。然而以鋅指蛋白質為基礎由体外選擇及設計的人工重組轉錄因子與其在生物體內活性的關係仍無完整有系統的分析。

本計劃的目的是利用出芽酵母菌為模式系統研究量身設計訂做的鋅指蛋白質對引導控制基因表達的能力。我探討分屬兩個不同裂殖原活化蛋白質激酵素路徑而涉及高張壓甘油反應和侵略性生長的基因。我將建構量身訂做的轉錄因子以活化下述基因之表達。其一為磷酸三甘油去氫酵素（*GPD1*）基因使得剔除 *HOG1* 基因的出芽酵母菌能於高張壓條件下仍能生長。另一為 musin-like（*MUC1*）基因使單倍體的出芽酵母菌在未缺乏氮養分來源時能呈侵略性生長。為了達到能識別各種組合的 DNA 序列，我採取一種新穎的同時並行（bipartite）選擇方法──同時以兩個互補基因庫選擇兩個 DNA 結合結構域，再經重組產生能識別一九個鹼基對的三指胜肽。已篩選出對 *MUC1* 或 *GPD1* 啟動子序列有專一性的鋅指結合基序，三指胜肽。其中一個三指胜肽被聯結成能識別一複合十八個鹼基對的六指胜肽。目前完成鋅指結合基序與激活結構域 VP16 或 VP64 的粘接於載體的構築，且已將鋅指結合基序與激活結構域轉錄因子構築於能在酵母菌並可調節表達蛋白質的載體。重組轉錄因子在酵母菌細胞內對 *GPD1* 或 *MUC1* 轉錄活化及表現型的影響將被檢視。

此研究的完成會成就一個模式系統－能深入且仔細分析依據不同 DNA 結合力的鋅指和不同功能結構域構築成的轉錄因子且在不同蛋白表達量下與生物體內基因調節之生物功能性的關係。這是以建構的鋅指蛋白質為基因開關來了解基因行為和分子醫學中臨床研應用的一大步。

關鍵詞：出芽酵母菌，鋅指蛋白質，噬菌體呈現，組合式分子庫，基因表現開關

Abstract

Cys_2His_2 zinc finger domain represents the most common eukaryotic DNA binding motif capable of recognizing an assorted set of DNA sequences. The structural stability of zinc finger motif makes it an exceptionally adaptable scaffold for protein engineering. The three-finger module from the transcription factor Zif268 has been used as a framework to construct combinatorial libraries displayed on bacteriophage to elucidate the principles of zinc finger-DNA

recognition. Recent progress in selection and design of strategies for zinc fingers has allowed construction of zinc finger proteins targeting at nearly any desired site on double-stranded DNA. Nevertheless, the *in vivo* biological activity of artificial recombinant transcription factors-based on zinc fingers in relation to their *in vitro* selection and design remains to be investigated systematically.

I have used budding yeast *Saccharomyces cerevisiae* as a model system to study the ability of tailor-made zinc finger proteins in regulation of gene expression. I exploited two gene mucin-like (*MUC1*) gene to promote invasive growth in haploid cells without depletion of nitrogen source. To accomplish comprehensive DNA recognition, I have adopted a novel approach of bipartite selection where two complementary libraries are used in parallel to select two DNA-binding domains, and whose products are recombined to produce a three-finger peptide that recognizes a 9two of which have been linked to generate a six-finger peptide that recognize a composite 18-bp sequence. These zinc s each of which codes for protein that is target in one of the two mitogen-activated protein kinase pathways, involved in either high-osmolarity glycerol response (HOG) or invasive growth. I aim to engineer tailor-made transcription factors that activate the expression of either glycerol-3-phosphate dehydrogenase (*GPD1*) gene to restore growth of cells deleted in *HOG1* in condition of high osmolarity or finger motifs have been attach to the activation domain VP16 from the herpes simplex virus or its derivative VP64 and have been constructed into a repressible yeast expression vector. This will allow me to examine the effects of the resulting recombinant transcription factors in yeast cells on transcriptional activation of *GDP1* or *MUC1* and accompanying phenotypic alterations-bp site. Several three-finger peptides recognizing promoter sequence of either *MUC1* or *GPD1* gene have been identified,.

Completion of this study will result in a model system that permits in depth and detailed analysis of biological functionality for gene regulation *in vivo* in relation to differently modified functional domains by engineering zinc finger-based transcription factors. It is a major step forward in the use of engineered zinc finger proteins as gene switches in the understanding of gene behavior and in clinical applications in molecular medicine.

Keywords: budding yeast, zinc finger, phage display, combinatorial libraries, gene switches

二、前言

The Cys_2His_2 zinc finger domain represents the most common DNA-binding motif in eukaryotes and the second most frequent encoded protein domain in the human proteome, around 2% of all our proteins (1). The individual fingers of about 30 amino acids generally function by binding with DNA. Each zinc finger module holds both an anti-parallel β-sheet and a

recognition α-helix that interacts directly with 3 base pairs of DNA in the major groove. This structurally simple ββα domain is stabilized by hydrophobic interactions and the chelation of single zinc ion between a pair of cysteines from the β-sheet and a pair of histidines from α-helix. Cys_2His_2 zinc finger domains are exceptionally suitable for the construction of artificial TFs as they generally set as covalent tandem repeats, allowing the recognition of extended asymmetrical DNA sequences. The modularity of the zinc finger proteins both in structure and function serves greater advantage in comparison with other types of DNA-binding domains that normally recognize DNA as dimmers and use non-modular recognition domains (reviewed in 2). The X-ray crystal structure of three-zinc finger protein Zif268 that binds to DNA revealed that the sequence-specific interactions are made by amino acids protruding from the N-terminus of the α-helix that fits into the major groove of the DNA double helix. Contacts are made mainly with one strand of the DNA, with positions –1, 3, and 6 of each finger (see Fig. 1) contacting the 3'-, middle, and 5'-nucleotides of a 3 bp subsite, respectively. The three fingers of Zif268 are oriented in such a way that finger 1 is at the 3' end of the primary strand and finger 3 is at the 5' end (Fig. 1). The structural stability, simple mode of DNA-recognition, and independent modularity of individual finger, Zif268 has subsequently provided a framework for the most design and selection studies of zinc fingers.

The selection of combinatorial libraries based on the display of three-finger Zif268 protein with randomized DNA-contacting residues from the surface of filamentous bacteriophage (3, 4) has been used to elucidate the principles of zinc-finger DNA recognition. Such libraries have been used to select for zinc finger proteins with novel DNA binding specificity. Recent progress of methods for the effective and efficient selection of zinc finger DNA-binding domains allows *de novo* design of tailored site-specific transcription factors (TFs) used as gene switches that up- or down-regulate endogenous genes.

三、研究目的

Research in the selection and design of tailor-made TFs for the use in controlling gene expression has been fruitful in recent years. However, a comprehensive analysis of the effect in functionality of the recombinant TFs and the accompanying phenotypic outcomes in relation to combination of the *in vitro* selected zinc finger DBDs with different affinity, increased finger numbers, the diverse effector domains, and the level of expression of the recombinant TFs has not been fully established. To establish the above mentioned, I took advantage of the well-established budding yeast *Saccharomyces cerevisiae* as a model system to study the ability of tailor-made zinc finger TFs in regulation of the expression of endogenous genes. Glycerol-3-phosphate dehydrogenase (*GPD1*) and mucin-like (*MUC1*) genes that are the targets on

pathways of either high-osmolarity glycerol response (HOG) or invasive growth respectively have chosen for the study as the pathways and the genes being well-characterized (see 49 for a recent review). Overproduction of GPD1 partially suppresses the osmosensitive phenotype of *HOG1* null mutant (8). The ability to invade by diploid cells of yeast can be achieved by overexpressing *MUC1* without the requirement of nitrogen depletion (9). Hence, I engineered tailor-made TFs that activate the expression of either *GPD1* to restore growth of cells deleted in *HOG1* under high osmolarity or *MUC1* to advance invasive growth in cells devoid of nitrogen source depletion. I aim at the result of wide-ranging analysis of the tailor-made TFs based on zinc finger with respect to biological functionality for *in vivo* gene regulation upon completion of this study. This will lay down a step stone for my future goal in establishing sequence-specific TFs as gene switches for the use in functional genomics and transcriptional therapy in clinical application when post-genome era can indeed offer such a great opportunity.

四、研究方法

Five and four target sites each of 9 bp from the 5' un-translated region or close to the start of translation of either *GPD1* or *MUC1* respectively have been identified from *Saccharomyces* Genome Database (SGD) for the use in the selection by phage display. These are all close to or within the cis-acting elements where the natural transcriptional regulators bind (10, 11). The target sites each of which have been undergone extensive search for their uniqueness in the budding yeast genome. The sequences of target sites chosen were essentially well balanced between purine and pyrimidine whenever possible. In order to test the feasibility of generating longer array of zinc finger peptides that target unique sequences when application of zinc fingers in large genome size such as human is a major concern, six-finger TFs have been sought to generate. This was achieved by fusing two selected three-fingers whose target sites are 0-2 bps apart as the optimal linker sequences for linking two three-finger peptides to become one six-finger peptide have been established (12).

Target sites were made with 5'-end being biotyinylated and were subject for selection by the bipartite approach where two complementary phage libraries are used in parallel to select two DNA-binding domains each of which recognizes a given 5 bp site, and whose products are recombined to produce a single three-finger peptide that recognizes a composite 9 bp site (Fig. 2). Only those selected three-finger peptides exhibit binding affinity to their target sequences at least as good as three-finger peptide of wild type Zif268 to its cognate target sequence determined by the ELISA-based assay were kept for further study. The selected phage clones expressing three-finger peptides each of which recognizing a 9-bp site and having a gap of 0-2 bp were linked to generate clones expressing six-finger peptides that recognize

a composite 18-bp sequence. These clones together with three-finger clones were then used to make constructs whose products are finger peptides attached to the minimal acidic activation domain of the herpes simplex virus virion transactivator VP16 or a tetrameric repeat of VP16's minimal activation domain, termed VP64. Such constructs were made to allow expressing recombinant protein in an *in vitro* transcription/translation system. To ensure that the six-finger peptides retaining binding affinity after adding the effector domain and to compare the binding affinity of three-finger peptides with that from whom six-finger peptides were made to perform gel retardation assay. Constructs that allow expressing finger peptides recombined with effector domains repressible by tetracycline were eventually made into a yeast expression vector (13). The effects of the resulting recombinant TFs in yeast cells on transcriptional activation of *GDP1* or *MUC1* under different concentration of tetracycline will be determined by Northern analysis or RT PCR. Genes with similar sequences to those of *MUC1* and *GPD1* (i.e. alteration of purine to purine or pyrimidine to pyrimidine in one base of the entire target site) will be sought to determine changes of their transcripts. The phenotypic consequences such as extent of invasiveness for cells carrying recombinant TFs targeting *MUC1* gene will be examined microscopically and by agar penetration of cells. Growth of *HOG1* mutants in high osmolarity will be observed in cells bearing recombinant TFs targeting *GPD1* on agar plate.

五、結果與討論

Selection of sequence-specific phage clones has resulted in identification of a number of zinc finger clones expressing peptides that bind 5-bp sites from 5' un-translated region of *MUC1* gene (Figure 2). These selected peptides exhibit binding affinity at least as good as three-finger peptide of wild type Zif268 by the ELISA-based assay (14). Among all target sites, a single three-finger peptide that recognizes a composite 9-bp site of 5'un-translated region of *MUC1* by recombination has been selected (Figure 3; Figure 4B, 4C; Figure 5). Another three-finger peptide that recognizes a 9-bp site of 5' un-translated region of *MUC1* by direct selection has been identified (Figure 2A). Both peptides have the binding affinity of nanomolar range, similar to that of the wild type Zif268 to its cognate binding sequence. These two selected three-finger peptides each of which recognizing a 9-bp site and having a gap of 2-bp have been linked to generate a six-finger peptide that recognizes a composite 18-bp sequence (Figure 6). All but one selected and one recombined three-finger motifs as well as one six-finger motif derived from those two have been subject to attachment of vector carrying either VP16 or VP64 sequence. The resulting plasmids can make recombinant TF (rTF) proteins carrying zinc finger motif and VP16 (or VP64) effector domain in an *in vitro* transcription/translation system. The recombinant TF proteins will be subject to

gel retardation analysis. Constructs bearing zinc finger motif and VP16 (or VP64) have been constructed into yeast expression vector pCM185 that is tetracycline repressible. Functional study of the ability of rTFs trans-activating the expression of *MUC1* gene and the accompanying increase of invasive growth will be accessed if further funding can be granted.

The selection of zinc finger clones using *GPD1* target sites has essentially been followed accordingly. Preliminary results indicate that the clones specifically recognizing promoter sequences of *GPD1* gene is obtainable. A three-finger peptide that recognizes a 9-bp site of 5' un-translated region of *GPD1* by direct selection has been identified (14). Several zinc finger clones expressing peptides that bind 5-bp sites from 5' un-translated region of *GPD1* gene have been isolated. These selected peptides exhibit binding affinity comparable with the three-finger peptide of wild type Zif268 as determined by the ELISA-based assay (14). Contrast to *MUC1*, single three-finger peptides that recognize composite 9-bp sites of 5' un-translated region of *GPD1* by recombination have not yet been identified. Nevertheless, I will carry on to make constructs of existing clone to GPD1 target site that allow conducting gel retardation and functional assay before any recombined clones being identified.

六、計劃成果自評

It has been slow in the first quarter of the year conducting the project, as setting up a workable environment is time-consuming. However, I have been able to complete all the constructs necessary for gel retardation analysis and functional study. Further financial support will warrant submitting a manuscript to journal of high impact factor on systemic analysis of the artificial transcription factor-based on C_2H_2 zinc finger in the control of gene expression in association the effect of phenotypic alteration.

七、參考文獻

1. Tupler, R., Perini, G. & Green, M.R. (2001) Expressing the human genome. *Nature* **409**, 832-833.
2. Klug, A. (1999) Zinc finger peptides for the regulation of gene expression. *J. Mol. Biol.* **293**, 215-218.
3. Choo, Y. & Klug, A. (1995) Designing DNA-binding proteins on the surface of filamentous phage. *Curr. Opin. Biotechnol.* **6**, 431-436.
4. Joung, J.K., Ramm, E.I. & Pabo, C.O. (2000) A bacterial two-hybrid selection system for studying protein-DNA and protein-protein interactions. *Proc. Natl. Acad. Sci. USA* **97**, 7382-7387.
5. Pavletich, N.P. & Pabo, C.O. (1991) Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science* **252**, 809-817.
6. Wolfe, S.A., Nekludova, L. & Pabo, C.O. (2000) DNA recognition by Cys_2His_2 zinc finger proteins. *Annu. Rev. Biophys. Biomol. Struc*. **29**, 183-212.
7. Pabo, C.O., Peisach, E. & Grant, R.A.

(2001) Design and Selection of Novel Cys_2His_2 zinc finger proteins. *Annu. Rev. Biochem.* **70**, 313-340.

8. Marquez, J.A., Pascul-Ahuir, A., Proft, M. & Serrano, R. (1998) The Ssn6-Tup1 repressor complex of *Saccharomyces cerevisiae* is involved in the osmotic induction of HOG-dependent and –independent genes. *EMBO J.* **17**, 2543-2553.
9. Lo, W.S. & Dranginis, A.M. (1998) The cell surface flocculin Flo11 is required for Pseudohyphae formation and invasion by *Saccharomyces cerevisiae*. *Mole. Biol. Cell* **9**, 161-171.
10. Steffen, R., Summers, E., Lo, H.-J., Madhani, H. & Fink, G. (1999) MAP kinase and camp filamentation signaling pathways converge on the unusually large promoter of the yeast FLO11 gene. *EMBO J.* **18**, 1257-1269.
11. Eriksson, P., Alipour, H., Adler, L. & Blomberg, A. (2000) Rap1p-binding sites in the Saccharomyces cerevisiae GPD1 promoter are involved in its response to NaCl. *J. Biol. Chem.* **275**, 29368-29376.
12. Moore, M., Choo, Y. & Klug, A. (2001) Design of polyzinc finger pepetides with structured linkers. *Proc. Natl. Acad. Sci. USA* **98**, 1432-1436.
13. Gari, E., Piedrafita, L., Aldea, M. & Hierrero, E. (1997) A set of vectors with a tetracycline-regulatable promoter system for modulated gene expression in *Saccharomyces cerevisiae*. *Yeast* **13**, 837-848.
14. *Shieh, J-C.*, Choo, Y., Isalan, M., Moore, M., Choo, K-B., Hong S-Y. & Klug A. (2002) Engineering zinc finger proteins to target at promoters of endogenous genes for regulating their expression in yeast. The 4th Asian-Pacific Organization for Cell Biology Congress-Cell Biology in the Post-Genomic Era.

附圖

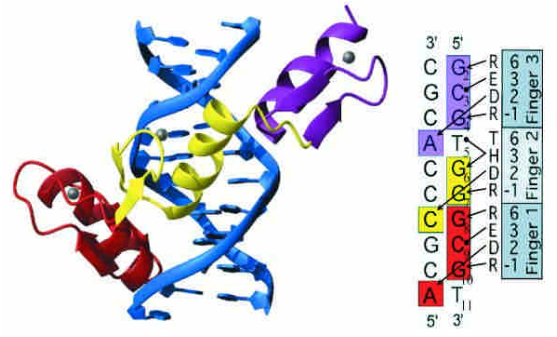

Figure 1. The Zif268-DNA complex (5), showing the three zinc fingers of Zif268 bound in the major groove of the DNA. Fingers are spaced at 3-bp intervals. The DNA is blue; fingers 1, 2, and 3 of Zif268 are red, yellow, and purple, respectively; and the coordinated zinc ions are represented as silver spheres. The DNA sequence of the Zif268 site is color-coded and base contacts made from positions –1,2 3, and 6 of each α-helix are indicated schematically to the right of the structure. *Arrows* indicate contacts mediated by hydrogen bonds; *closed circles* indicate hydrophobic interactions. For reference the base pairs are numbered (2 through 11) as in the original reference (5). (Adopted, combined, and modified from references 6 and 7.)

Figure 2 Sequences of the target sites used in the bipartite selection. Upper panel (A) shows the full length sequence of each target site. Lower panel (B) shows sequence of "half site' in each target used in the selection in that the 'GCGG' (in green) needs to be introduced in the sequence along with the 'GT' as the fixed flanking bases. Actual target sequences (blue and red) have a single base overlapping (in purple). Flanking sequences are shown in gray. The position where each finger peptide binds is indicated on the top of the target sequence.

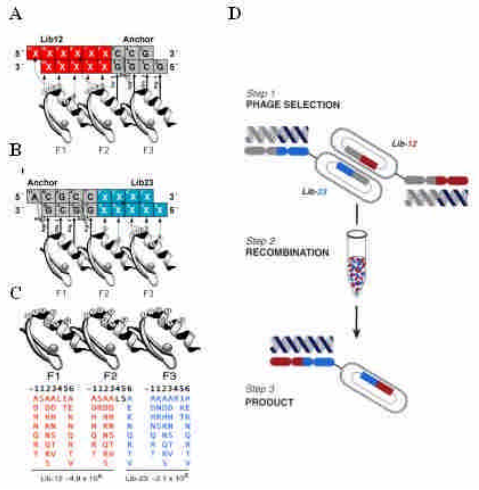

Figure 3 Phage display libraries and selection approach used in this project. Schematic model of zinc finger binding modes in (A) Lib-12 and (B) Lib-23. The binding scheme is based on the Zif268 co-crystal structure. Each library contains anchor region, based on Zif268 scaffold, which binds to the DNA bases shown in gray. The randomized library regions bind to bases ^{6}X-^{10}X for Lib-12 (red) and bases ^{2}X-^{6}X for Lib-23 (blue). Randomized amino acid residues are shown as circles numbered relative to their α-helical positions. *Broken arrows* tentatively indicate the geometry of potential amino acid-base interactions. (C) Figure showing potential amino acid composition of randomized zinc finger residues in lib-12 (red) and Lib-23 (blue). The distribution of these positions in the Zif268-framework is shown shematically by circles. The theoretical sizes of each half-library are shown below the figure. A *DdeI* restriction site has been introduced into the DNA sequence coding for helical residues 4 and 5 (Leu and Ser) in finger 2. An *in vitro* recombination of the two half-libraries may be carried out by cutting and religating at this *DdeI* site. (D) Schematic diagram representing the bipartite selection strategy. Two libraries. Lib-12 and Lib-23 contain randomized amino acid positions in finger 1 and 2 (red) or finger 2 and 3 (blue). Non-randomized 'anchor' regions of the Zif268 scaffold are shown in grey. After selection against DNA 'half-sites' (Step 1), the genes coding for the two selected half-portions are recombined *in vitro* (Step 2). In this way, aproduct that binds the full length DNA target site may be cloned back into the phage vector (Step 3).

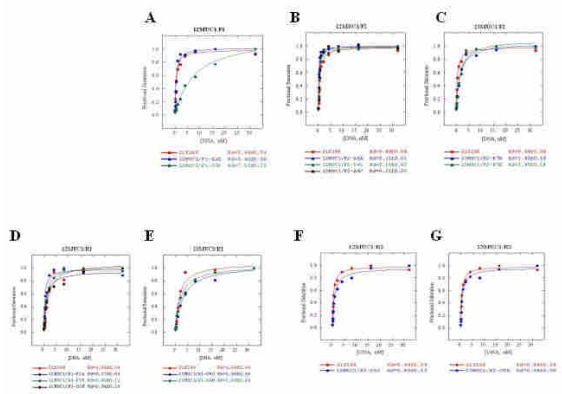

Figure 4 Apparent equilibrium binding affinities of Zif268 and selected phage clones for *MUC1* promoter using either direct selection of MUC1/F1 (A) or bipartite system, in which phage clones target each "half (the 5-bp) site" for MUC1/F2 (B & C), MUC1/R1 (D & E), or (G & F). Number 12 and 23 shown on top of each Binding curves of zinc finger phage for their appropriate DNA target sites as determined by ELISA. Dissociation constants derived from binding curves using program KALEIDAGRAPH 3.0.5.

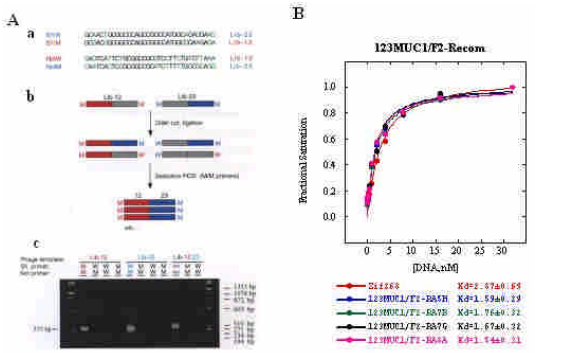

Figure 5 Phage clone expressing a three-finger peptide targeting the MUC1/F2 site. A. Selective PCR of phage- selected and 'anchor' portions of the bipartite libraries. Throughout this figure 5A, primers complementary to Lib12 and Lib23 are shown in red and blue, respectively. A(a) PCR primers carry mutations that are complementary to individual library cassettes. Primers with wild-type Zif268-phage sequences are marked, "W". Mutant primers are marked, "M". The mutations (shown in green) control PCR specificity but do not alter protein sequence when translated. A(b) Recombination scheme for recovering selected mutant (M) portions from each library. Note that the process allows selective recovery of recombined cassettes from a mixed pool of templates, thereby eliminating the need for purification after phage selection. A(c) Agarose gel showing selective PCR operating on a mixture of templates, using different combinations of mutant (M) and wild-type (W) primers. PCR products comprising the ~300 bp zinc finger cassette are only produced when appropriate pairs of Sfi and Not primers are used. B. Apparent equilibrium binding affinities of Zif268 and selected phage clones for MUC1/F2 from *MUC1* promoter using bipartite selection system. These four recombined clones were the results by screening 48 clones from recombination process of three Lib12 derived phage clones (Fig. 4B) and two Lib23 derived phage clones (Fig. 4C) targeting each of their "half (the 5-bp) site" of the "full MUC1/F2 (the 9-bp) site''.

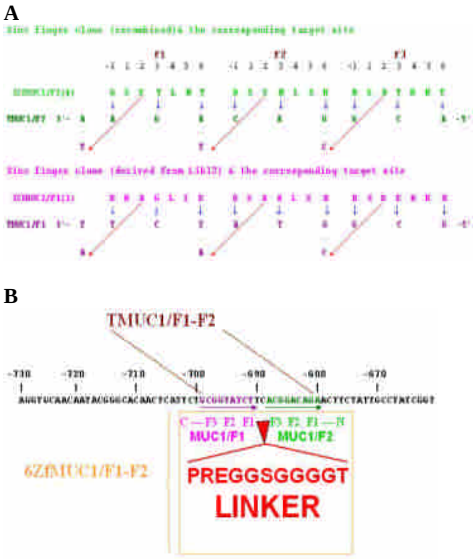

Figure 6 Peptide sequences of phage clones and their corresponding DNA target sites. A. Three-finger 3ZfMUC1/F1 and 3ZfMUC1/F2 targeting TMUC1/F1 and TMUC1/F2, respectively. The numbers in the round bracket of clones named donate times clones being isolated. Note that blue arrows represent base-amino acid residue contacts on one DNA strand, whereas the extended red arrows represent cross-strand contacts. B. Six-finger 6ZfMUC1/F1-F2 has an 11- residues linker that join 3ZfMUC1/F1 and 3ZfMUC1/F2 together. Target site of 6ZfMUC1/F1-F2 comprises of combined sites of TMUC1/F1 and TMUC1/F2 with a 2-bp gap in between.