Available online at www.sciencedirect.com
SCIENCE DIRECT
Biochemical and Biophysical Research Communications 339 (2006) 263–270
BBRC
www.elsevier.com/locate/ybbrc

Effects of different zinc finger transcription factors on genomic targets

Leon W. Neuteboom, Beatrice I. Lindhout, Ingrid L. Saman, Paul J.J. Hooykaas, Bert J. van der Zaal *

Institute of Biology Leiden, Department of Molecular and Developmental Genetics, Clusius Laboratory, Leiden University, Wassenaarseweg 64, 2333 AL, Leiden, The Netherlands

Received 28 October 2005
Available online 10 November 2005

Abstract

We have developed a novel vector system for the efficient assembly of polydactyl zinc fingers. Next to proteins that possess short canonical TGEKP linkers between all constituting zinc fingers we constructed proteins with longer, 12 amino acid linkers between two three-finger (3F) units and between three two-finger (2F) units. Fusions of these zinc finger domains with the VP16 activation domain were tested for their ability to regulate a repressed genomic locus containing contiguous or noncontiguous zinc finger binding sites in yeast. In contrast to other studies, which were mostly confined to in vitro tests, we did not obtain evidence that superior artificial zinc finger transcription factors need to contain longer linkers between individual fingers. For the regulation of genomic loci, canonical linkers within a highly regular backbone in combination with a contiguous 18 base pair DNA target site were found to provide a sound base for polydactyl zinc finger design.

Keywords: Polydactyl zinc finger; Artificial transcription factor; Genomic target site; Linker design; DNA binding; In vivo regulation; Yeast (*Saccharomyces cerevisiae*); *MEL1*; VP16; Mig1p

Artificial transcription factors consisting of a designed polydactyl zinc finger (PZF) domain for DNA binding, linked to an activator or repressor domain, have recently gained enormous attention as novel tools in molecular biology [1,2]. In theory, it should now be possible to regulate at will any gene from any organism for which sufficient genomic sequence information has been obtained. PZF domains are most successfully constructed using the highly modular Cys_2-His_2 zinc finger (ZF) domains. Just a few amino acid residues within the α-helix of a Cys_2-His_2 ZF domain are crucial for the recognition of a particular triplet of a DNA sequence. Via elegant experimental protocols optimized ZF sequences for triplets starting with 5′ G, A and C have now been elucidated [3–6].

The most common strategy for the construction of PZF domains involves PCR-mediated grafting of modules into a regular scaffold with a canonical 'TGEKP' five amino acid linker sequence between the different ZFs [2]. Assembly of six-finger (6F) domains in this way has been shown to lead to a relatively modest increase in affinity for the corresponding 18 bp target sites compared to the constituting 3F domains for their 9 bp target sites [7,8]. Since on theoretical grounds a much higher gain in affinity should be possible, it was hypothesized that the helical periodicity of ZFs connected via canonical linkers is just too short to match the helical periodicity of B-form DNA [9]. Binding of more extended PZF moieties would therefore increase DNA unwinding, resulting in a structural tension that counteracts the stability of the complex. For this reason, the use of longer linkers at one or more positions within 6F proteins was expected to increase the stability of the protein–DNA complex and binding affinity. Indications to support this view have been published by Kim and Pabo

* Corresponding author. Fax: +31 71 527 4999.
E-mail address: Zaal@Rulbim.LeidenUniv.nl (B.J. van der Zaal).

0006-291X/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2005.11.011

[10], who inserted LRQKDGERP and LRQKDGGGSERP linkers between two 3F domains and calculated a 6000-fold tighter interaction with the target site than for its 3F components. Interestingly, these 6F proteins were able to span one or two extra base pairs present between the two 9 bp sequences of the half sites. In addition, Moore et al. [11] found that insertion of TGGEKP linkers between fingers 2 and 3 and between fingers 4 and 5 (3× 2F design) resulted in high affinity binding of a 18 bp contiguous target site, while TGGGSEKP linkers allowed recognition of noncontiguous target sites with one base pair inserted between positions 6 and 7 and/or 12 and 13 of the original 18 bp sequence. Unfortunately, neither Kim and Pabo nor Moore et al. tested the performance of a 6F protein with only canonical linkers and all conclusions were solely based on in vitro binding assays and transient in vivo assays using plasmid-based reporter genes.

Apart from the possibility to vary linker length, the use of different frameworks to construct PZF domains can lead to different affinities for the binding sites [8]. So far there is no consensus in the field and even mixtures of different backbones within a single PZF protein are frequently encountered in literature. In this work, we developed a cloning vehicle for rapid and flexible production of PZF proteins. To avoid any unnecessary bias and target site overlap, PZF proteins were in principle constructed using optimized 5′-GNN recognition helices [3] within a strict Sp1C framework and cognate binding sites comply with $(GNN)_6G$. The system allows a choice of flexible linkers of different length at various positions in the multifinger protein. The sequence of each linker, including the longer linkers, complies with sequences known to lead to stabilizing α-helix C-capping interactions upon DNA binding [9,12,13]. We systematically compared proteins arranged in a 2× 3F and 3× 2F design with highly regular 6F proteins for optimal DNA binding. Although PZF modules are intended for manipulating expression of genomic loci, systematic analyses of different types of PZF modules for their ability to be used on chromatin-embedded DNA are lacking. Therefore, we introduced series of VP16-PZF fusions into yeast (*Saccharomyces cerevisiae*) strains that carry a stably integrated, repressed reporter gene with various ZF target sites. The in vivo gene expression data were combined with in vitro analyses for binding characteristics of the different proteins. We found that strong in vitro binding does not guarantee a good performance of the corresponding artificial transcription factor in vivo. Remarkably, the most optimal PZF modules for in vivo transcriptional regulation possessed canonical linkers between all fingers. These modules were also highly specific for their corresponding uninterrupted target sites. 6F proteins also performed much better than four-finger (4F) and five-finger (5F) proteins for in vivo gene activation, a feature not observed previously. We did not find any indications that 6F proteins with short, canonical linkers impose strain on the DNA helix that counteracts DNA binding affinity.

Materials and methods

Construction of vectors and generation of PZF proteins. The pBluescript II SK+-derived plasmid pSKN-*Sgr*AI (Fig. 1A, cloning details in Supplementary material) was used for the construction of ZF modules. Oligonucleotides with the coding region optimized for recognition of a particular 5′-GNN-3′ DNA target sequence [3] were sequentially ligated (see Supplementary material) into the (reoccurring) *Sgr*AI site to obtain the 6F modules PTF1 and E2C1 (Fig. 2B). PTF2 and E2C2 (Fig. 2B) were constructed by fusing two 3F-encoding domains. DNA encoding the C-terminal half of the 6F protein was isolated as a *Sgr*AI–*Bsp*EI fragment and cloned into a *Bsp*EI-digested plasmid encoding the N-terminal half of

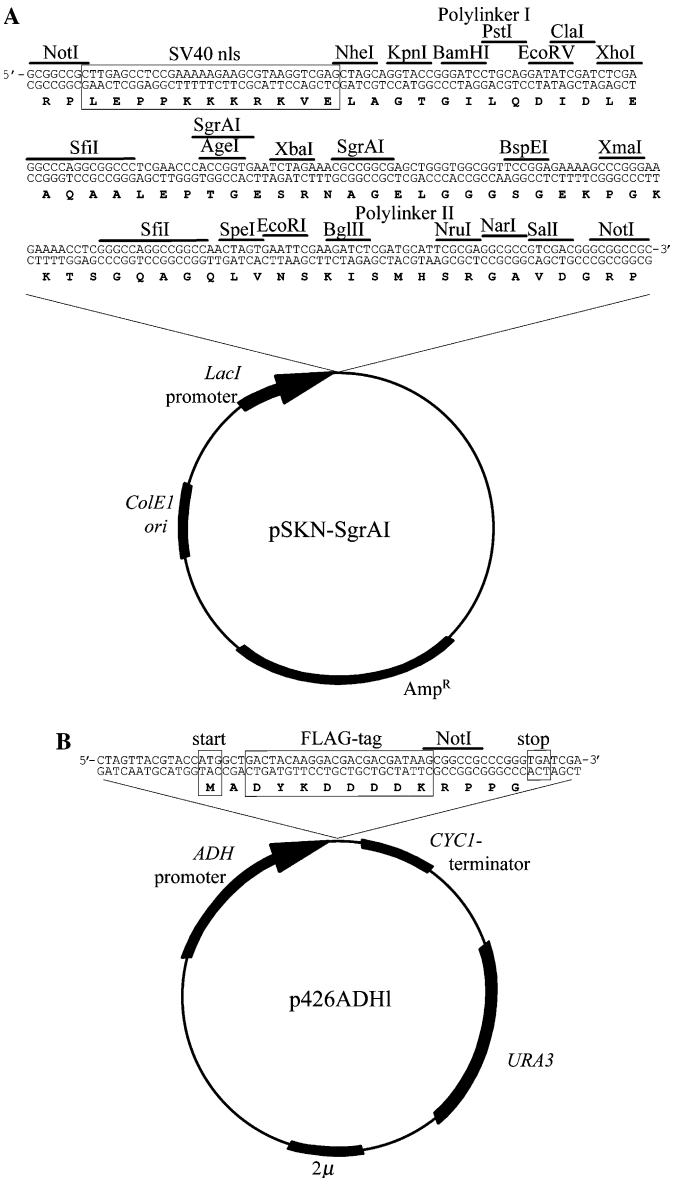

Fig. 1. Plasmids used in this study. (A) pSKN-*Sgr*AI for the construction of PZF-transcription factors. Relevant DNA sequences and the derived amino acid sequence, which includes the SV40 nuclear localization signal are shown. *Sfi*I sites can be used for the excision of constructed ZF modules, while a variety of unique restriction sites in polylinkers I and II allows translational fusion(s) of additional regulatory effector domains at the N- and/or the C-terminus of the ZF moiety. (B) p426ADHl for the expression of PZF-transcription factors in yeast. A linker, which adds an N-terminal FLAG-tag to the expressed proteins, was inserted in the *Spe*I and *Xho*I sites of p426ADH [14].

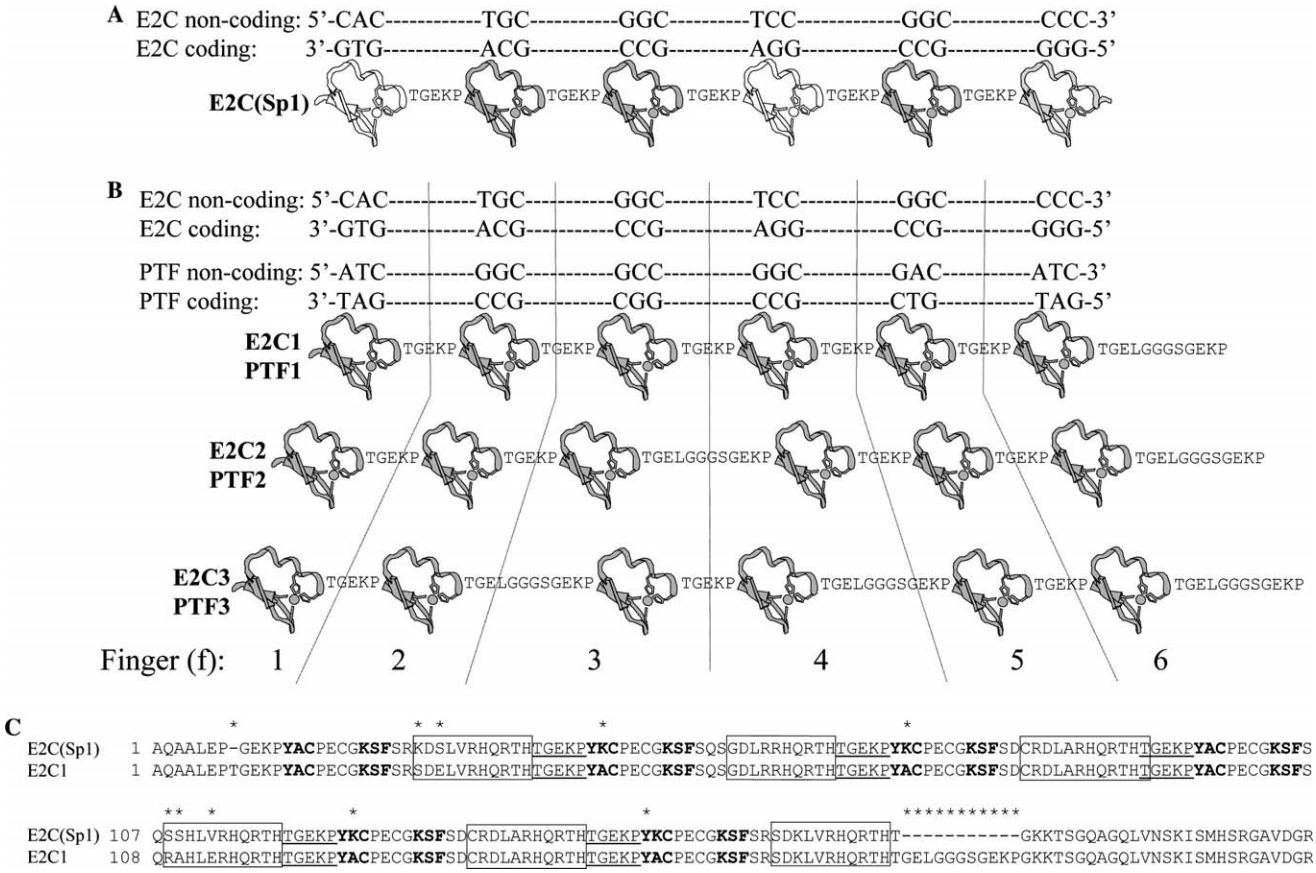

Fig. 2. Overview of the 6F protein designs used in this study. (A) The E2C(Sp1) 6F construct with its cognate binding site. (B) 6F constructs developed with pSKN-*Sgr*AI. The PTF and E2C contiguous binding sites are shown above the constituting fingers to indicate the design of the corresponding individual fingers f1–f6. (C) Alignment of E2C(Sp1) and E2C1. Linkers are underlined, β-sheets are shown in boldface and α-helices are boxed. Differences caused by applying the optimized GNN code [3] and pSKN-*Sgr*AI for construction of E2C1 are highlighted by asterisks.

the 6F protein. PTF3 and E2C3 were similarly constructed by sequential ligation of 2F modules. In addition, a sequence encoding E2C(Sp1) was generated in the pSKN-*Sgr*AI-like vector pSKNN (see Supplementary material) via stepwise PCR protocols using long overlapping primers (Figs. 2A and C). The herpes simplex virus VP16 transcriptional activation domain was cloned upstream from the PZF-encoding sequence (see Supplementary material). *Not*I sites were used to excise the fusion protein-coding sequences for cloning into the p426ADH-derived [14] yeast expression vector p426ADHl (Fig. 1B).

Construction of yeast strains. A *Not*I–*Kpn*I fragment of plasmid pMELα UASc-10bp-UASc (kind gift of Dr. Karsten Melcher, University of Frankfurt, Germany) [15,16], harbouring the *MEL1* gene driven by the *MEL1* core promoter was used for the cloning of either one or three ZF binding sites in the *Aat*II site (Fig. 3), which is located 79 bp upstream from the nearest TATA box (cloning details in Supplementary material). After subcloning of *Not*I–*Kpn*I fragments into pSKNN, constructs could be inserted as *Not*I fragments into the yeast integration vector pINT1 [18]. Recombinant pINT1 plasmids harbouring the *MEL1* gene in the same orientation as the *APT1* gene were linearized with *Sac*I and *Nco*I and the DNA was used to transform yeast strain 21RΔmel1 [17] by a PEG/LiAc transformation method [18]. Yeast colonies containing the *MEL1* gene at the *PDC6* locus were selected. Plasmid p426ADHl with the zinc finger-containing modules (or control constructs) were subsequently transferred into the reporter strains. The levels of FLAG-tagged proteins were verified by Western blotting (see Supplementary material).

Galactosidase assays. Galactosidase assays were performed in triplicate with three independently grown cultures for each yeast target strain/p426ADHl-derived expression plasmid combination (see Supplementary material).

Gel shift assays and determination of dissociation constants (see also Supplementary material). Crude yeast cell extracts were prepared using CelLyticTM Y Cell Lysis Reagent (Sigma), including 2 mM DTT, according to the manufacturer's instructions. Samples were assayed for DNA-binding activity using [α-^{32}P]dCTP end-labelled synthetic oligonucleotide duplexes (Fig. 3) essentially as described by Moore et al. [19] except that 20 mM Hepes (pH 7.2) was used as a buffer component instead of Bis-Tris propane (pH 7.0). Apparent dissociation constants (K_d) for binding reactions were determined by Scatchard analysis essentially as described [20], with a constant amount of labelled binding site (15 pM) and increasing amounts of unlabelled binding sites in each reaction. After initial trials, conditions were—when possible—chosen in such a manner that at the lowest probe concentration used (about 10-fold lower than the estimated K_d) at least 50–60% was bound in complex with the PZF protein. Assays were performed in triplicate.

Results

Synthesis of PZF proteins and DNA binding sites

The vector pSKN-*Sgr*AI (Fig. 1A) was used for the construction of two different series of 6F proteins (Fig. 2). The PTF series can be regarded as a random choice out of all possible 6F proteins that could have been made. Although not relevant for the present study, the unique 18 bp PTF recognition sequence is present in the first exon of the

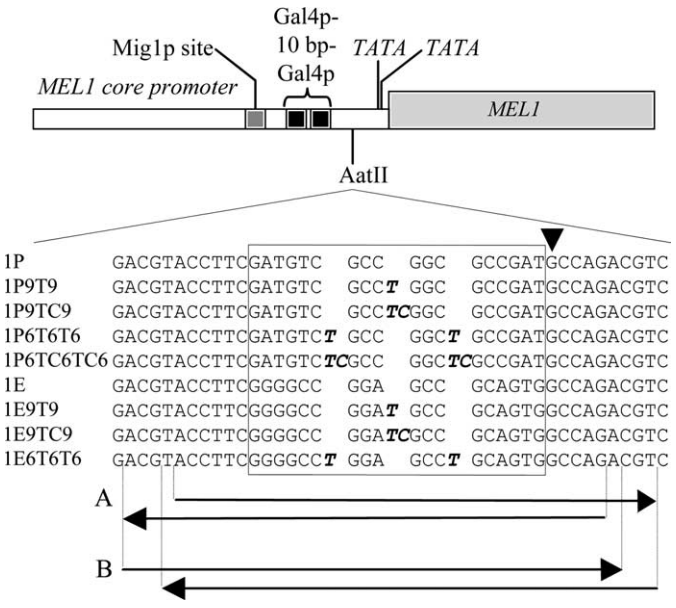

Fig. 3. Organisation of the chromosomal reporter for in vivo assays. The *MEL1* gene is driven by the *MEL1* core promoter in which one Mig1P and two Gal4p sites are present. The AatII site was used for the cloning of ZF binding sites shown in the sequence alignment underneath. They were cloned using primer pairs indicated by arrows in (A) directly below the corresponding sequence. The arrow to the left represents the opposite strand of the sequence that is directly above it. The triplets for which ZFs were designed are boxed and extra insertions within this 18 bp sequence are shown in italics. An extra G, following the last GNN triplet is indicated by a black triangle. Constructs with three ZF target sites were obtained by ligation of three sets of primer pairs indicated in (A). Primer pairs used for end-labelling are shown in (B).

PINOID gene (AF232236, At2g34650) of the model plant species *Arabidopsis thaliana*. The E2C proteins are based on E2C(Sp1) as studied by Beerli et al. [8]. E2C(Sp1), which differs from E2C1 in fingers 1 and 4 (Fig. 2C), was included in our study as a reference PZF module. Data from the E2C series served predominantly to corroborate the 'random' PTF data. For the PTF series the individual ZFs matched with GNN triplets occurring in the 18 bp sequence 5′-GAT GTC GCC GGC GCC GAT-3′, for the E2C series the GNN triplets matched with 5′-GGG GCC GGA GCC GCA GTG-3′. PTF1 and E2C1 represent 6F modules with canonical TGEKP linkers connecting all individual ZF moieties, while PTF2 and E2C2 contain TGELGGGSGEKP linkers between fingers 3 and 4 and PTF3 and E2C3 between fingers 2 and 3 as well as fingers 4 and 5. Apart from these 6F final versions, precursors of these proteins were also used for experimentation. Binding sites were designated 1P and 1E for single contiguous binding sites corresponding to PTF and E2C, respectively (Fig. 3) and 3P and 3E for three direct repeats of the same binding site. We accounted for any possible target site overlap problem that is inherent to some of the GNN triplet-binding ZFs, which require a G or T to be present at the 5′ position in the neighbouring finger's subsite [1]. Apart from a G residue downstream from the target site sequence, extra internal nucleotides were either T or TC (Fig. 3). Insertions were placed after every 6 bases and after 9 bases in the 18 bp PTF binding site, hence giving rise to the single target sites 1P9T9, 1P9TC9, 1P6T6T6, and 1P6TC6TC6 as indicated in Fig. 3 and the triple target sites 3P9T9, 3P9TC9, 3P6T6T6 and 3P6TC6TC6. Correspondingly, interrupted target sites for E2C (Fig. 3) comprised 1E6T6T6, 1E9T9, 1E9TC9, 3E6T6T6, and 3E9T9.

In vivo activity of VP16-PZF proteins

The ability of various VP16-PZF fusions to induce expression of a reporter gene on a chromosomal locus was tested in yeast. All single and triple contiguous and noncontiguous target sites were cloned into the modified promoter of the *MEL1* reporter gene (Fig. 3). The different reporter constructs obtained in this way were all integrated into the *PDC6* locus [17]. Growth of the 'target' strains in the presence of glucose results in active repression of the *MEL1* gene through the presence of a Mig1p site [15,21] in the *MEL1* core promoter (Fig. 3). The actively repressed state of the genomic locus can be overcome by *ADH*-promoter-driven expression of the appropriate VP16-PZF fusion construct from the 2μ plasmid p426ADHl. Our assay provides an excellent model for the most common application of designed PZF-TFs. Moreover, the presence of two Gal4p sites (Fig. 3) allows a direct comparison between the levels of VP16-PZF-mediated reporter gene expression in the presence of glucose and the 'maximum' level of MEL1 activity observed after galactose induction.

VP16-PTF1 and shorter precursors were tested for their ability to activate MEL1 expression in yeast strains possessing either one (Fig. 4A) or three (Fig. 4B) uninterrupted binding sites in the *MEL1* promoter. The in vivo test system clearly demonstrated that at least five ZFs are needed within a PTF module to activate expression of the reporter gene under repressed conditions. None of the VP16 constructs with four (VP16-PTFf3-6 and VP16-PTFf1-4) or three ZFs (VP16-PTFf1-3 and VP16-PTFf4-6) was able to induce detectable levels of α-galactosidase. As expected, neither PTF1 nor VP16 alone were able to trans-activate gene expression (Figs. 4A and B); at the same time neither VP16-PTF1 nor any of the other *ADH* promoter-driven constructs mentioned in this paper were able to activate MEL1 expression in a control strain, lacking a ZF binding site in the *Aat*II site of the modified *MEL1* promoter (data not shown).

The successful trans-activation of *MEL1* expression via the VP16-PTF1 protein in our test system allowed us to compare characteristics of the PTF1 module with those of differently designed 6F modules with longer linkers on both contiguous and interrupted binding sites. As shown in Figs. 4C and D, the introduction of a longer linker between finger 3 and finger 4 (VP16-PTF2) dramatically reduced *MEL1* activation compared to VP16-PTF1, which possesses TGEKP canonical linkers between all fingers.

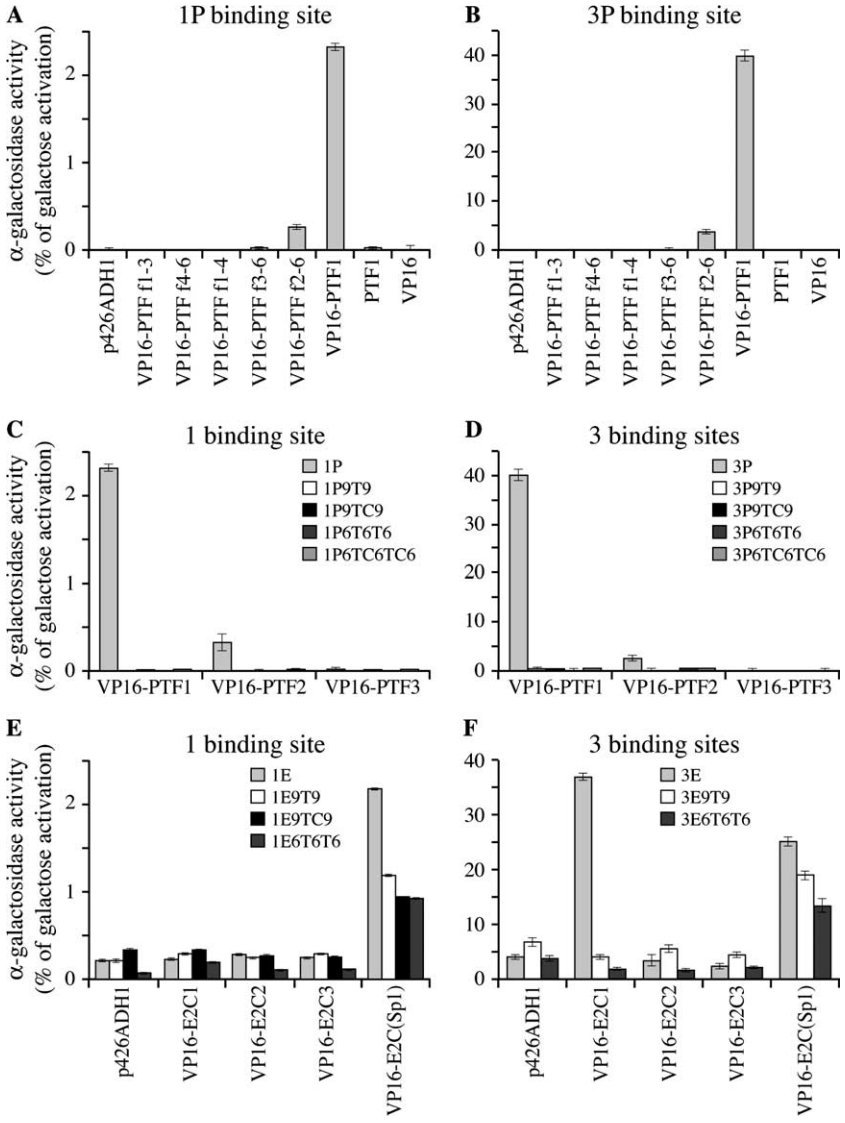

Fig. 4. Effect of different PTF and E2C constructs and controls on MEL1 expression in yeast strains carrying one or three binding sites. MEL1 expression is shown as the percentage of expression measured in corresponding yeast target strains harbouring p426ADHl and grown in the presence of 2% galactose. For 3P9T9-containing strains the galactose-induced expression was found to be relatively low, but this did not affect results. (A) Strains containing the 1P binding site with various PTF-derived constructs. VP16-PTFf4-6 possesses the last three fingers (f4–f6) of the complete 6F PTF1 protein as indicated in Fig. 2B. VP16-PTFf1-3 possesses the first three fingers. Similarly, the individual ZFs present in the four finger proteins (VP16-PTFf3-6 and VP16-PTFf1-4) and the five finger protein VP16-PTFf2-5 can be derived from Fig. 2B. (B) strains containing the 3P binding site with various PTF-derived constructs. (C) Strains with various single binding sites with PTF-derived constructs. (D) Strains with various triple binding sites with PTF-derived constructs. (E) Strains with various single binding sites with E2C-derived constructs. (F) Strains with various triple binding sites with E2C-derived constructs.

VP16-PTF3, with 2 longer linkers dividing the 6F protein in a 3× 2F configuration, was completely inactive. Surprisingly, none of the VP16-PTF constructs was able to induce reporter gene expression in yeast strains with interruptions in the ZF binding site(s) (Figs. 4C and D). Thus the PTF1 interaction with the contiguous 18 bp target site is the most effective interaction and is also very specific by not tolerating extra nucleotides enlarging the binding site.

Results obtained with 6F PTF modules were extended by experiments performed with E2C(Sp1) and pSKN-*Sgr*AI-constructed E2C versions of PZF domains. Strains carrying E2C binding sites or derivatives thereof showed relatively high background levels, indicating interference of other yeast factors. Nevertheless, VP16-6F proteins were able to further activate expression of the reporter gene (Figs. 4E and F). The E2C(Sp1) PZF module was somewhat less effective than the E2C1 module for *MEL1* activation on the 3E target sequence, but on a single target site it was much more effective. However, reporter genes endowed with other target sites than the contiguous 18 bp sequence were also activated by VP16-E2C(Sp1). Although E2C(Sp1) might be prone to some nonspecific interactions compared to E2C1, due to the presence of two less specific ZFs in E2C(Sp1) (Fig. 2C), the reason why more extended target sites seem to be recognized is not easily explained. The observed stimulation is specific

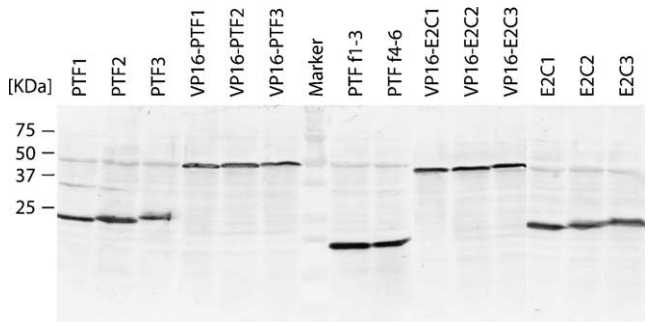

Fig. 5. Western blot analysis of expression of a selection of PZF constructs in strains carrying three corresponding contiguous binding sites.

Table 1
Apparent K_d values for E2C and PTF proteins and their target sites

Protein	Binding site	Apparent K_d (nM)
E2C(Sp1)	1E	0.4 (±0.1)
E2C1	1E	7.1 (±1.9)
	1E9T9	>20
E2C2	1E	>20
	1E9T9	>20
PTF1	1P	1.5 (±0.7)
	1P9T9	1.9 (±0.2)
PTF2	1P	5.4 (±0.9)
	1P9T9	2.4 (±0.8)

Measurements were performed in triplicate. Standard deviations are shown in brackets.

for the cloned binding sites, because VP16-E2C(Sp1) did not stimulate expression in yeast strains harbouring PTF target sites (data not shown). In any case, for the pSKN-*Sgr*AI-generated E2C series as well as for the PTF series the interaction between a PZF domain with canonical linkers and a noninterrupted 18 bp target site proved to be the most efficient and specific for VP16-mediated trans-activation of a repressed genomic locus in yeast.

To exclude that the rather unexpected results of the in vivo gene activation assay were caused by differences in protein synthesis we checked the presence of FLAG-tagged PZF proteins by immunodetection (Fig. 5). All proteins were produced at comparable levels, irrespective of the presence of longer linkers and a VP16 domain.

In vitro DNA-binding properties of PZF proteins

Extracts from the yeast strains tested in the in vivo system were used for gels shift analysis. Although not affecting binding parameters, the presence of the fused VP16 domain consistently led to the appearance of supershifted bands (data not shown). This is probably caused by interactions of VP16 with components of the basal transcription machinery and/or other chromatin components [22]. With a single shifted complex the bound fraction is much easier to quantify, therefore we used extracts from yeast strains producing 'free' PZF domains for a detailed in vitro characterization. Calculated dissociation constants are shown in Table 1.

The yeast system was validated by the observation that the apparent K_d for E2C(Sp1) with 1E was calculated to be 0.4 nM. This value was very close to the value of 0.5 nM reported by Beerli et al. [8], who used purified *Escherichia coli*-produced protein. The difference in K_d value for E2C1 with 1E (7.1 nM) is most likely explained by the use of two different individual ZFs in the E2C1 version (Fig. 2C). A drawback of the yeast system is that the PZF protein concentrations are relatively low, which hampers any accurate determination of K_d values above 20 nM. This was the case for all pSKN-*Sgr*AI-derived 6F proteins with binding sites not listed in Table 1, including the 6F proteins with two longer linkers (PTF3 and E2C3). In addition, none of the 3F proteins mentioned in our study bound any of the target sites with a K_d value in the lower nanomolar range (data not shown). Instead, all estimated K_d values for 3F proteins, which can be regarded as constituents of the more complex PZF domains used in our study, were estimated to be well over 20 nM. Again this agrees with data reported by Beerli et al. [8], who found K_d values ranging from 25 to 75 nM for a variety of 3F proteins. In all those cases, where accurate measurements were not feasible due to the combination of high K_d values and low protein concentrations as present in yeast extracts we did not pursue further analysis of in vitro DNA binding.

The 6F protein PTF1 bound with nanomolar affinity to the 1P and 1P9T9 site, with only a slightly higher K_d for the 1P9T9 site (Table 1). Surprisingly, in our in vivo test system we observed a substantial difference in MEL1 expression; while the contiguous sites mediated efficient transactivation of MEL1 expression, no MEL activity was detected in strains carrying 1P9T9 and 3P9T9 target sites.

Just as PTF1, the 2× 3F version of PTF (PTF2) exhibited nanomolar K_d values for the interaction with 1P (5.4 nM) and 1P9T9 (2.4 nM). The introduction of the longer TGELGGGSGEKP linker between the two 3F peptides thus does not have a major negative effect on the binding of the 6F protein in vitro; the 2× 3F K_d values are still much lower than measured for the two 3F components separately. The in vitro data indicate that introduction of a single TGELGGGSGEKP linker causes a modest preference of the resulting 2× 3F protein for the interrupted 1P9T9 binding site over the 1P site. Remarkably, this was again not apparent in the in vivo assay (Figs. 4C and D) since PTF2 could trans-activate reporter gene expression via the contiguous 1P and 3P sites, but not the 1P9T9 and 3P9T9 sites.

While the in vivo and the in vitro data regarding PTF proteins and their derivatives could be compared, a similar comparison proved to be problematic for the E2C series. As mentioned above, E2C(Sp1) is constructed with different ZFs (Fig. 2C) with the purpose of serving as an internal standard, but not as a logical component of the series of E2C pro-

teins. Of this series, only E2C1 was effective in vivo and had a sufficiently low K_d value to be determined in our system. The relatively high K_d value of 7.1 nM of E2C1 for binding to 1E explains why K_d values for E2C2 just fell outside of the measurable range, assuming that E2C1 and E2C2 follow the same trend as PTF1 and PTF2 (Table 1). For the 2× 3F E2C2 protein, an in vitro preference for the 9T9 binding site might also exist, but could not be established under the experimental circumstances.

Discussion

Successful use of PZF-containing artificial transcription factors to modulate expression of eukaryotic genes requires that the PZF domains can interact efficiently with cognate recognition sites that are present within chromatin. The yeast system developed in this study offers the possibility to test transcription factors based on zinc finger technology for their potential to regulate a genomic locus and—at the same time—allows the production of different types of PZF proteins for in vitro analysis. Our study describes, for the first time, a systematic comparison of different 6F designs, including a 'straight' 6F peptide with only canonical linkers between the individual ZFs. The well established 6F peptide E2C(Sp1) was included a reference.

While several studies have indicated that artificial transcription factors with shorter PZF modules can influence gene expression in vivo, even on genomic loci [23–25], we did not find any transcriptional activation of the repressed genomic yeast reporter locus via VP16-PZF modules with less than five ZF moieties. Apparently, our in vivo assay system puts up a high threshold for establishing contact with the binding site. Regarding the fact that most PZF modules to be constructed will be intended for in vivo applications, most notably on a genomic locus, a robust assay to test a logical series of PZF designs to function on a genomic locus was advantageous. Because of the challenging in vivo requirements in our experimental setup, several practical conclusions regarding PZF design can now be drawn with more certainty. First of all, efficient transcriptional stimulation can only be achieved when a PZF domain possesses an apparent K_d value for its cognate binding site that is below 10 nM under our experimental conditions. Remarkably, as a second point, it is absolutely clear that a low K_d value for a PZF–DNA interaction is no guarantee that this interaction will be useful in vivo. In addition to these conclusions it should be remarked that an apparent K_d value is a rather arbitrary unit as there is no strict consensus on a uniform method for establishing these values for protein–DNA interactions. K_d values can vary considerably when determined under different conditions [19] and a calculated value can therefore not direct PZF design for in vivo applications. The third and most important conclusion to be drawn from our experiments is that a very strict 6F design with only TGEKP canonical linkers between the individual fingers performs optimally in vivo with great specificity for contiguous 18 bp target sites.

Our 2× 3F designs (PTF2 and E2C2), with the longer TGELGGSGEKP linker between finger 3 and finger 4, showed a dramatic reduction in the trans-activating potential compared to the 6F modules with canonical linkers between all constituting ZFs. The presence of extra base pairs within the binding site completely abolished trans-activation. Despite our efforts to comply with insights based on structural analysis of linker properties [9,12,13,26], it might be argued that our longer linker does not have the most optimal design. However, in vitro the TGELGGSGEKP linker containing PTF2 protein showed strong interactions with both the 1P and 1P9T9 target sites (Table 1), similar as observed for PZFs with alternative designs of longer linkers in comparable in vitro test systems [10,11,19,26]. Although we cannot exclude that superior longer linkers can be designed, it seems unlikely that a particular need for longer linkers exists in vivo in order to meet any progressing unwinding of the DNA caused by an increasing number of zinc fingers. This conclusion is likely to be a universal lead for PZF design. Recent structural data do indeed suggest that PZF domains with TGEKP linker sequences should provide ample flexibility to bind DNA without any destabilizing effects [27].

The discrepancies between in vitro and in vivo data are not easily explained, but must originate from a rather basal difference between naked DNA and similar sequences presented as chromatin-embedded DNA. The helical periodicity of B-form DNA in solution is ~10.5 bp per turn, but it has been shown that DNA unwinding can be actively induced upon binding of PZFs [28]. However, DNA that is organised within a nucleosome has a different helical periodicity [29] and is probably less prone to allow conformational changes. Altogether, our data indicate that the helical periodicity of a longer PZF domain with only short canonical TGEKP linkers fits best with the helical periodicity of contiguous target sites present in chromatin-embedded DNA. The demonstration that the most regular PZF design has in fact all valuable traits for gene regulation in vivo strongly indicates that designing biologically active PZF domains in such a manner will be a successful and reproducible strategy.

Acknowledgments

This work was supported by Grant 050-10-123 from the NWO Programma Genomics. We are grateful to Dr. Karsten Melcher (University of Frankfurt, Germany), Dr. Pieter Ouwerkerk and Dr. Paul van Heusden (University of Leiden, the Netherlands) for materials and valuable discussions. We thank Martin Brittijn (University of Leiden) for assistance in figure preparations and Dr. Marianne Rots (Therapeutic Gene Modulation, University of

Groningen, the Netherlands) for critical reading of the manuscript.

Supplementary data

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.bbrc.2005.11.011.

References

[1] P. Blancafort, D.J. Segal, C.F. Barbas III, Designing transcription factor architectures for drug discovery, Mol. Pharmacol. 66 (2004) 1361–1371.
[2] A.C. Jamieson, J.C. Miller, C.O. Pabo, Drug discovery with engineered zinc-finger proteins, Nat. Rev. Drug Discov. 2 (2003) 361–368.
[3] D.J. Segal, B. Dreier, R.R. Beerli, C.F. Barbas III, 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. USA 96 (1999) 2758–2763.
[4] B. Dreier, D.J. Segal, C.F. Barbas III, Insights into the molecular recognition of the 5′-GNN-3′ family of DNA sequences by zinc finger domains, J. Mol. Biol. 303 (2000) 489–502.
[5] B. Dreier, R.R. Beerli, D.J. Segal, J.D. Flippin, C.F. Barbas III, 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. 276 (2001) 29466–29478.
[6] B. Dreier, R.P. Fuller, D.J. Segal, C. Lund, P. Blancafort, A. Huber, B. Koksch, C.F. BarbasI III, 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. 280 (2005) 35588–35597.
[7] Q. Liu, D.J. Segal, J.B. Ghiara, C.F. Barbas III, Design of polydactyl zinc-finger proteins for unique addressing within complex genomes, Proc. Natl. Acad. Sci. USA 94 (1997) 5525–5530.
[8] R.R. Beerli, D.J. Segal, B. Dreier, C.F. Barbas III, Toward controlling gene expression at will: specific regulation of the *erbB-2/HER-2* promoter by using polydactyl zinc finger proteins constructed from modular building blocks, Proc. Natl. Acad. Sci. USA 95 (1998) 14628–14633.
[9] D.S. Wuttke, M.P. Foster, D.A. Case, J.M. Gottesfeld, P.E. Wright, Solution structure of the first three zinc fingers of TFIIIA bound to the cognate DNA Sequence: determinants of affinity and sequence specificity, J. Mol. Biol. 273 (1997) 183–206.
[10] J.S. Kim, C.O. Pabo, Getting a handhold on DNA: design of poly-zinc finger proteins with femtomolar dissociation constants, Proc. Natl. Acad. Sci. USA 95 (1998) 2812–2817.
[11] M. Moore, A. Klug, Y. Choo, Improved DNA binding specificity from polyzinc finger peptides by using strings of two finger units, Proc. Natl. Acad. Sci. USA 98 (2001) 1437–1441.
[12] J.H. Laity, H.J. Dyson, P.E. Wright, DNA-induced α-helix capping in conserved linker sequences is a determinant of binding affinity in Cys_2-His_2 zinc fingers, J. Mol. Biol. 295 (2000) 719–727.
[13] M. Elrod-Erickson, M.A. Rould, L. Nekludova, C.O. Pabo, Zif268 protein–DNA complex refined at 1.6 Å: a model system for understanding zinc finger–DNA interactions, Structure 4 (1996) 1171–1180.
[14] D. Mumberg, R. Müller, M. Funk, Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds, Gene 156 (1995) 119–122.
[15] K. Melcher, B. Sharma, W.V. Ding, M. Nolden, Zero background reporter plasmids, Gene 247 (2000) 53–61.
[16] K. Melcher, H.E. Xu, Gal80–Gal80 interaction on adjacent Gal4p binding sites is required for complete *GAL* gene repression, EMBO J. 20 (2001) 841–851.
[17] A.H. Meijer, P.B.F. Ouwerkerk, J.H.C. Hoge, Vectors for transcription factor cloning and target site identification by means of genetic selection in yeast, Yeast 14 (1999) 1407–1416.
[18] D. Gietz, A. St. Jean, R.A. Woods, R.H. Schiestl, Improved method for high efficiency transformation of intact yeast cells, Nucl. Acids Res. 20 (1992) 1425.
[19] M. Moore, Y. Choo, A. Klug, Design of polyzinc finger peptides with structured linkers, Proc. Natl. Acad. Sci. USA 98 (2001) 1432–1436.
[20] K.R. Clemens, P. Zhang, X. Liao, S.J. McBryant, P.E. Wright, J.M. Gottesfeld, Relative contribution of the zinc fingers of transcription factor IIIA to the energetics of DNA binding, J. Mol. Biol. 244 (1994) 23–35.
[21] J.O. Nehlin, M. Carlberg, H. Ronne, Control of yeast *GAL* genes by MIG1 repressor: a transcriptional cascade in the glucose response, EMBO J. 10 (1991) 3373–3377.
[22] F.J. Herrera, S.J. Triezenberg, VP16-dependent association of chromatin-modifying coactivators and underrepresentation of histones at immediate-early gene promoters during herpes simplex virus infection, J. Virol. 78 (2004) 9689–9696.
[23] P.Q. Liu, E.J. Rebar, L. Zhang, Q. Liu, A.C. Jamieson, Y. Liang, H. Qi, P.X. Li, B. Chen, M.C. Mendel, X. Zhong, Y.L. Lee, S.P. Eisenberg, S.K. Spratt, C.C. Case, A.P. Wolffe, Regulation of an endogenous locus using a panel of designed zinc finger proteins targeted to accessible chromatin regions. Activation of vascular endothelial growth factor A, J. Biol. Chem. 276 (2001) 11323–11334.
[24] L. Zhang, K. Spratt, Q. Liu, B. Johnstone, H. Qi, E.E. Raschke, A.C. Jamieson, E.J. Rebar, A.P. Wolffe, C.C. Case, Synthetic zinc finger transcription factor action at an endogenous chromosomal site, J. Biol. Chem. 275 (2000) 33850–33860.
[25] R.R. Beerli, B. Dreier, C.F. Barbas III, Positive and negative regulation of endogenous genes by designed transcription factors, Proc. Natl. Acad. Sci. USA 97 (2000) 1495–1500.
[26] W. Nomura, Y. Sugiura, Effects of length and position of an extended linker on sequence-selective DNA recognition of zinc finger peptides, Biochemistry 42 (2003) 14805–14813.
[27] E. Peisach, C.O. Pabo, Constraints for zinc finger linker design as inferred from X-ray crystal structure of tandem Zif268–DNA complexes, J. Mol. Biol. 330 (2003) 1–7.
[28] Y. Shi, J.M. Berg, DNA unwinding induced by zinc finger protein binding, Biochemistry 35 (1996) 3845–3848.
[29] J.J. Hayes, D.J. Clark, A.P. Wolffe, Histone contributions to the structure of DNA in the nucleosome, Proc. Natl. Acad. Sci. USA 88 (1991) 6829–6833.