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

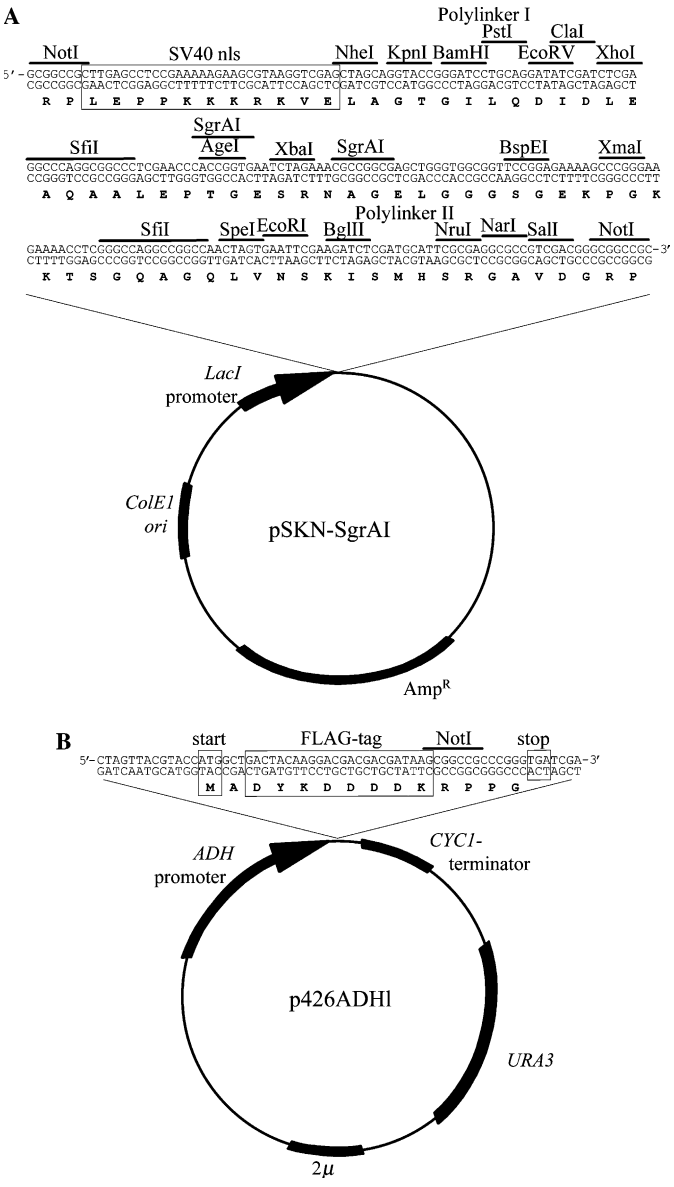

Fig. 1. Plasmids used in this study. (A) pSKN-*Sgr*AI for the construction of PZF-transcription factors. Relevant DNA sequences and the derived amino acid sequence, which includes the SV40 nuclear localization signal are shown. *Sfi*I sites can be used for the excision of constructed ZF modules, while 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].