Biochimica et Biophysica Acta 1799 (2010) 768–774

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagrm

Review

Small-molecule regulators that mimic transcription factors

José A. Rodríguez-Martínez, Kimberly J. Peterson-Kaufman, Aseem Z. Ansari *

Department of Biochemistry and the Genome Center, University of Wisconsin-Madison, Madison, WI 53706, USA

ARTICLE INFO

Article history:
Received 25 April 2010
Received in revised form 17 August 2010
Accepted 22 August 2010
Available online 6 September 2010

Keywords:
Chemical biology
Synthetic biology
Transcription factor mimic
Gene regulation
Cooperative binding
Protein–DNA interaction

ABSTRACT

Transcription factors (TFs) are responsible for decoding and expressing the information stored in the genome, which dictates cellular function. Creating artificial transcription factors (ATFs) that mimic endogenous TFs is a major goal at the interface of biology, chemistry, and molecular medicine. Such molecular tools will be essential for deciphering and manipulating transcriptional networks that lead to particular cellular states. In this minireview, the framework for the design of functional ATFs is presented and current challenges in the successful implementation of ATFs are discussed.

1. Introduction

Most cells in multicellular organisms carry the same genome, yet are able to produce a wide range of phenotypes which gives rise to sets of specialized cells that differ in morphology and function. This diversity is in part attributed to differences in tightly regulated gene expression patterns, with some genes being actively transcribed and others repressed. Transcription factor (TF) proteins are active participants in the regulation of specific gene-expression programs in response to cellular needs. Therefore, it is not surprising that the malfunctioning of TFs has been directly linked to many disease states [1]. This link has turned TFs into attractive therapeutic targets for treating a wide range of diseases, including cancer [2–4].

In response to a specific signal, TFs target particular genes within the genome. Once localized to the targeted genes, TFs recruit macromolecular machines to modify chromatin and initiate transcription [5]. Over several decades, much effort has been invested in the identification of the components of the transcriptional machinery targeted by TFs [6,7]. Transcription factors have been shown to interact with RNA polymerase II, the general transcription factors (GTFs) [5], coactivators such as components of the Mediator protein complex [8,9], and TBP-associated factors [10,11]. TFs also recruit nucleosome remodeling complexes such as the Swi/Snf complex and histone acetyltransferases, such as the SAGA complex [12,13]. Components of the proteasome have also been identified as targets of transcriptional activators (Fig. 1) [7].

Natural transcription factors can be minimally composed of two functional domains: a DNA-binding domain (DBD) and a regulatory domain (RD) [5]. The DBD determines which genes will be activated or repressed by selectively targeting specific DNA sequences within the *cis*-regulatory motifs associated with the target genes; the RD dictates whether to activate or repress transcription by recruiting components associated with the transcriptional machinery or the repression machinery, respectively. The magnitude of the response is encoded within the regulatory domain.

An important feature of natural TFs is that the DBD and the RD function independently from each other, as demonstrated by domain swapping experiments in yeast and other eukaryotes [14]. The modular nature of TFs highlights the possibility of exchanging the DBD and RD for synthetic counterparts to engineer artificial transcription factors (ATFs). Engineering replacements for the DBD and RD has been the most used strategy for creating TF mimics (Fig. 1) [15].

The potential benefits of implementing ATF-based tools are extensive [16]. These molecular tools could be used to dissect genome-wide transcriptional cascades, yielding fundamental insights on developmental processes. Diseases based on malfunctioning transcription factors could be treated or prevented with ATFs. The metabolic pathways of an organism could be engineered to produce valuable compounds. ATFs would also be invaluable tools for the emerging field of synthetic biology, as they could be used to control synthetic cellular circuits [17].

* Corresponding author. 315C Biochemistry Addition, Department of Biochemistry, 433 Babcock Drive, Madison, WI 53706-1544, USA. Tel.: +1 608 265 4690; fax: +1 608 262 3453.
E-mail address: ansari@biochem.wisc.edu (A.Z. Ansari).

1874-9399/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbagrm.2010.08.010

2. DNA binding domains

The information contained within the DBD dictates which DNA sequence is targeted and therefore determines which genes are regulated by a given transcription factor. Similarly the DBD confers specificity on a given ATF. Different types of binding domains have been employed in ATFs to target specific DNA sequences. Examples of DNA binding domains used for ATF construction include protein-based zinc fingers, oligonucleotides and oligonucleotide analogs, as well as synthetic small molecules (Fig. 2).

The zinc finger (ZF) domain is one of the most represented DBD in the human genome [18,19]. A zinc finger module is composed of 30 amino acids assembled in a ββα fold stabilized by a zinc ion. Each ZF recognizes and binds to three base pairs in the target DNA (Fig. 2). ZF modules can be strung together to recognize larger unique sequences in the genome. For example, three consecutive ZFs target a 9 bp sequence, and a polydactyl ZF consisting of six ZFs targets an 18 bp sequence [20]. The complexity of sequences that can be recognized by ZFs has been expanded through a variety of strategies, including structure-guided methods, phage display screens, and the bacterial one-hybrid system [21–23]. The most successful artificial ZF modules target sequences containing GNN triplets [20,24]. A detailed protocol for the modular construction of ZF libraries was recently published by the Barbas group [25]. Zinc fingers have been widely employed as DNA-binding domains in the construction of ATFs [26]. However, it has been shown that in some cases the targeted sequences of individual ZFs are not completely separable and that DNA binding is influenced by the neighboring ZFs as well [27,28].

Two recent reports described the DNA recognition "code" of the transcription activator-like (TAL) effectors of bacteria from the genus *Xanthomonas* [29,30]. TAL effectors are DNA binding proteins from plant pathogenic bacteria [31]. Members of the TAL effectors family possess a characteristic central domain of tandem repeats of 34 amino acids. In each repeat, the amino acids located in positions 12 and 13 are hypervariable and referred to as the repeat-variable diresidue (RVD). The DNA binding specificity of TAL effectors is determined by the tandem repeat region [32]. Specifically, a one-to-one correspondence was found between the identity of the RVD and target DNA [29,30].

The deciphering of the DNA binding code of TAL effectors highlights the possibility of engineering TAL effectors with custom DNA sequence specificity. However, the molecular details on how the repeat domain of TAL effectors recognizes targeted DNA are currently lacking. Although more work is needed to support the generality of the proposed DNA binding code, TAL effectors could potentially be utilized as DBD in designing transcription factor mimics.

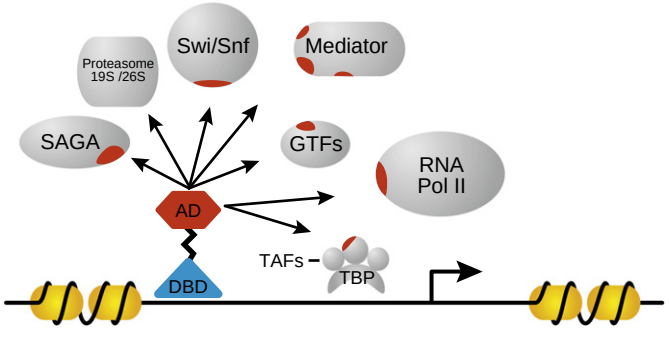

Fig. 1. Transcription activation by transcription factors. TFs are minimally composed of a DNA binding domain (DBD) and an activator domain (AD). The DBD recognize and binds to a DNA sequence to activate the targeted gene(s). The AD recruits the transcriptional machinery components through interactions with RNA polymerase II (RNApol), general transcription factors, (GTFs), (TBP)-associated factors (TAFs), the Mediator complex, chromatin remodeling complexes such as SAGA and Swi/Snf complexes and/or the 19S and 26S components of the proteasome.

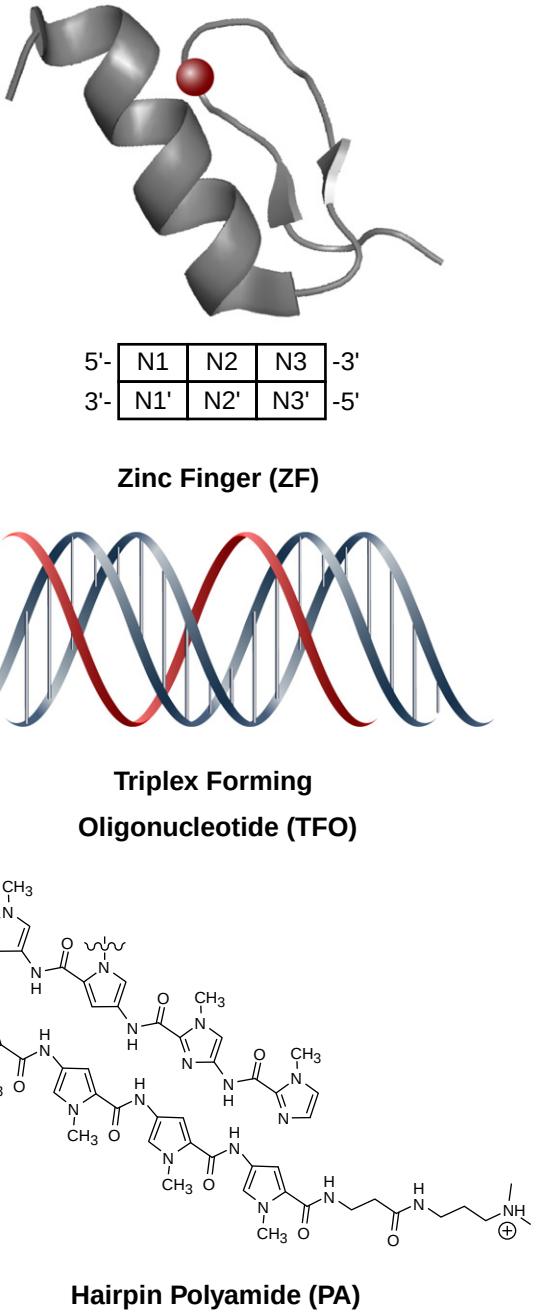

Fig. 2. DNA binding domains commonly used in ATFs. Zinc fingers (ZF) recognize and bind to 3 bp (N1–3) in dsDNA; triplex forming oligonucleotides (TFOs); hairpin polyamides (PAs).

DBDs have also been constructed from oligonucleotides [33,34] as well as oligonucleotide analogs, such as locked-nucleic acids (LNAs) [35] and peptide nucleic acids (PNAs) [36]. These molecules recognize and bind to DNA by forming a triple helix DNA strand (referred to as triplex-forming oligos (TFO)), or by strand invasion of double-stranded DNA [37]. An ATF consisting of a triplex-forming oligonucleotide DBD linked to a minimal VP16 peptide AD was first reported by Kuznetsova et al. [38]. This work was later extended by Young and colleagues to create TFO-based ATF that induced the expression of a reporter gene in tissue culture cells [39].

The most effective small-molecule DBDs to date are based on *N*-methylpyrrole and *N*-methylimidazole polyamides (PA). These molecules bind in the minor groove of dsDNA [40]. When engineered to form hairpins, PAs are capable of binding to targeted DNA