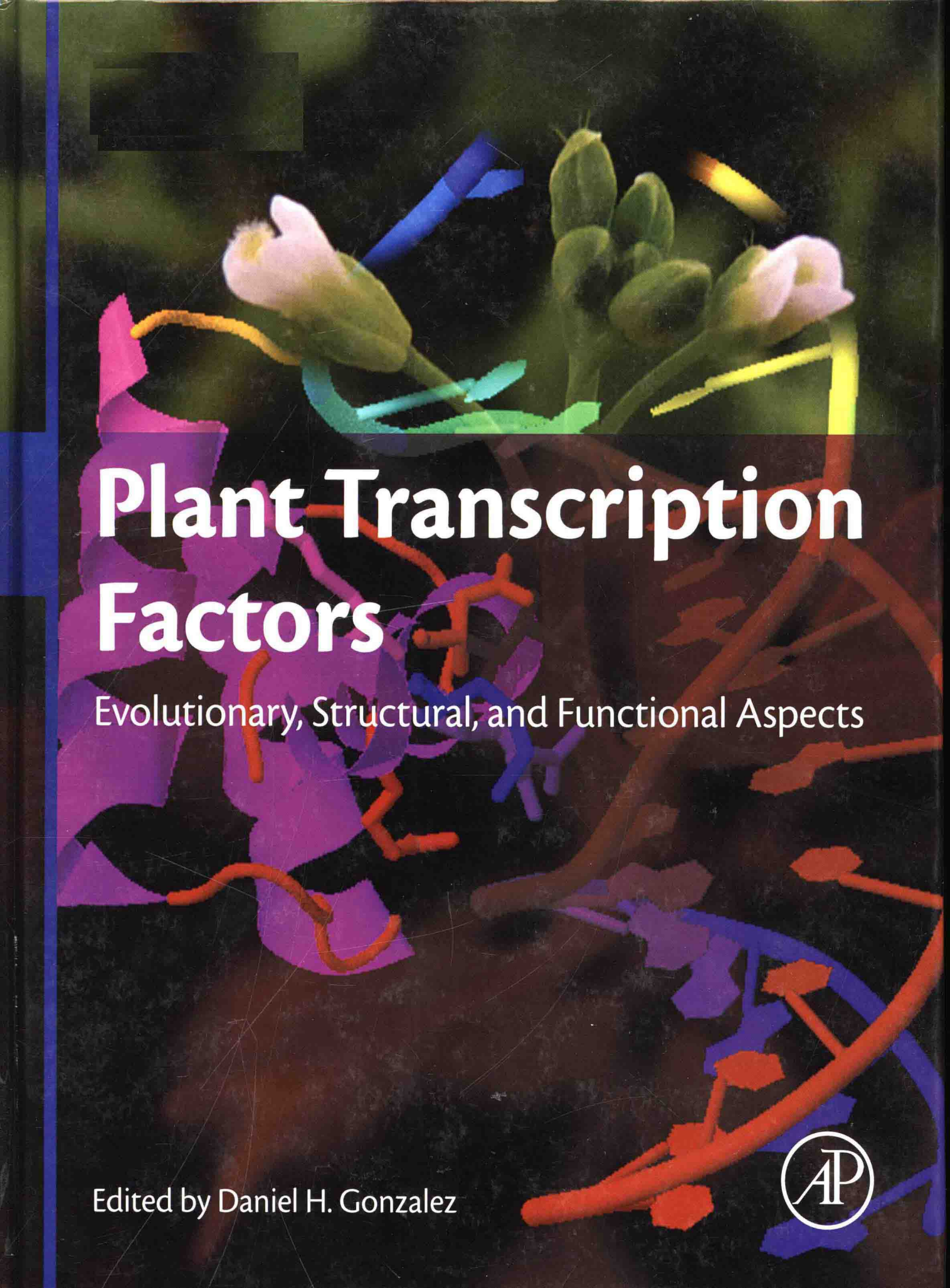

Plant Transcription Factors

Evolutionary, Structural, and Functional Aspects

Edited by Daniel H. Gonzalez

PLANT TRANSCRIPTION FACTORS

EVOLUTIONARY, STRUCTURAL, AND FUNCTIONAL ASPECTS

Edited by

DANIEL H. GONZALEZ
Universidad Nacional del Litoral, Santa Fe, Argentina

ELSEVIER

AMSTERDAM • BOSTON • HEIDELBERG • LONDON
NEW YORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO

Academic Press is an Imprint of Elsevier

Academic Press is an imprint of Elsevier
125, London Wall, EC2Y 5AS, UK
525 B Street, Suite 1800, San Diego, CA 92101-4495, USA
225 Wyman Street, Waltham, MA 02451, USA
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK

Notices
Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress

ISBN: 978-0-12-800854-6

Typeset by Thomson Digital
Printed and bound in the United States of America

PLANT TRANSCRIPTION FACTORS

List of Contributors

Agustín L. Arce Instituto de Agrobiotecnología del Litoral, Universidad Nacional del Litoral, CONICET, Santa Fe, Argentina

Nicolas Arnaud INRA, UMR1318, AgroParisTech, Institut Jean-Pierre Bourgin, Versailles, France

Cordelia Bolle Department Biologie I, Lehrstuhl für Molekularbiologie der Pflanzen (Botanik), Biozentrum der LMU München, Planegg-Martinsried, Germany

Matías Capella Instituto de Agrobiotecnología del Litoral, Universidad Nacional del Litoral, CONICET, Santa Fe, Argentina

Raquel L. Chan Instituto de Agrobiotecnología del Litoral, Universidad Nacional del Litoral, CONICET, Santa Fe, Argentina

Pilar Cubas Department of Plant Molecular Genetics, Centro Nacional de Biotecnología (CNB-CSIC), Madrid, Spain

Juan Manuel Debernardi Instituto de Biología Molecular y Celular de Rosario (IBR), CONICET, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Rosario, Argentina

Farah Deeba Department of Biology, University of Copenhagen, Copenhagen, Denmark; Department of Biochemistry, PMAS Arid Agriculture University Rawalpindi, Rawalpindi, Pakistan

María Florencia Ercoli Instituto de Biología Molecular y Celular de Rosario (IBR), CONICET, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Rosario, Argentina

Marçal Gallemí Centre for Research in Agricultural Genomics (CRAG), Consortium CSIC-IRTA-UAB-UB, Barcelona, Spain

Maria Dolores Gomez Instituto de Biología Molecular y Celular de Plantas (IBMCP), Universidad Politécnica de Valencia-Consejo Superior de Investigaciones Científicas, Ciudad Politécnica de la Innovación (CPI), Valencia, Spain

Beatriz Gonçalves INRA, UMR1318, AgroParisTech, Institut Jean-Pierre Bourgin, Versailles, France

Daniel H. Gonzalez Instituto de Agrobiotecnología del Litoral (CONICET-UNL), Cátedra de Biología Celular y Molecular, Facultad de Bioquímica y Ciencias Biológicas, Universidad Nacional del Litoral, Santa Fe, Argentina

Eduardo González-Grandío Department of Plant Molecular Genetics, Centro Nacional de Biotecnología (CNB-CSIC), Madrid, Spain

Lydia Gramzow Department of Genetics, Friedrich Schiller University Jena, Jena, Germany

Sarah Hake U.S. Department of Agriculture-Agricultural Research Service, Plant and Microbial Biology Department, Plant Gene Expression Center, University of California at Berkeley, Berkeley, CA, USA

David J. Hannapel Plant Biology Major, Iowa State University, Ames, Iowa, USA

Jong Chan Hong Division of Life Science, Applied Life Science (BK21 Plus Program), Plant Molecular Biology and Biotechnology Research Center, Gyeongsang National University, Jinju, Korea; Division of Plant Sciences, University of Missouri, Columbia, MO, USA

Aeni Hosaka-Sasaki Plant Genome Research Unit, Agrogenomics Research Center, National Institute of Agrobiological Sciences (NIAS), Tsukuba, Ibaraki, Japan

Yuji Iwata Graduate School of Life and Environmental Sciences, Osaka Prefecture University, Nakaku, Sakai, Osaka, Japan

Shoshi Kikuchi Plant Genome Research Unit, Agrogenomics Research Center, National Institute of Agrobiological Sciences (NIAS), Tsukuba, Ibaraki, Japan

Nozomu Koizumi Graduate School of Life and Environmental Sciences, Osaka Prefecture University, Nakaku, Sakai, Osaka, Japan

Patrick Laufs INRA, UMR1318, AgroParisTech, Institut Jean-Pierre Bourgin, Versailles, France

Yuan Li Institute of Molecular Plant Sciences, School of Biological Sciences, University of Edinburgh, King's Buildings, Edinburgh, UK

Gary J. Loake Institute of Molecular Plant Sciences, School of Biological Sciences, University of Edinburgh, King's Buildings, Edinburgh, UK

Leila Lo Leggio Department of Chemistry, University of Copenhagen, Copenhagen, Denmark

Jaime F. Martínez-García Centre for Research in Agricultural Genomics (CRAG), Consortium CSIC-IRTA-UAB-UB; Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

Aude Maugarny INRA, UMR1318, AgroParisTech, Institut Jean-Pierre Bourgin, Versailles, France

Toshifumi Nagata Plant Genome Research Unit, Agrogenomics Research Center, National Institute of Agrobiological Sciences (NIAS), Tsukuba, Ibaraki, Japan

Michael Nicolas Department of Plant Molecular Genetics, Centro Nacional de Biotecnología (CNB-CSIC), Madrid, Spain

Javier F. Palatnik Instituto de Biología Molecular y Celular de Rosario (IBR), CONICET, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Rosario, Argentina

Miguel A. Perez-Amador Instituto de Biología Molecular y Celular de Plantas (IBMCP), Universidad Politécnica de Valencia-Consejo Superior de Investigaciones Científicas, Ciudad Politécnica de la Innovación (CPI), Valencia, Spain

Roel C. Rabara Texas A and M AgriLife Research and Extension Center, Dallas, TX, USA

Pamela A. Ribone Instituto de Agrobiotecnología del Litoral, Universidad Nacional del Litoral, CONICET, Santa Fe, Argentina

Charles I. Rinerson Texas A and M AgriLife Research and Extension Center, Dallas, TX, USA

Ramiro E. Rodriguez Instituto de Biología Molecular y Celular de Rosario (IBR), CONICET, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Rosario, Argentina

Paul J. Rushton Texas A and M AgriLife Research and Extension Center, Dallas, TX, USA

Qingxi J. Shen School of Life Sciences, University of Nevada, at Las Vegas, NV, USA

Karen Skriver Department of Biology, University of Copenhagen, Copenhagen, Denmark

Sophia L. Stone Department of Biology, Dalhousie University, Halifax NS, Canada

Günter Theißen Department of Genetics, Friedrich Schiller University Jena, Jena, Germany

Prateek Tripathi Molecular and Computational Biology Section, Dana and David Dornsife College of Letters, Arts and Sciences, University of Southern California, Los Angeles, CA, USA

Katsutoshi Tsuda U.S. Department of Agriculture-Agricultural Research Service, Plant and Microbial Biology Department, Plant Gene Expression Center, University of California at Berkeley, Berkeley, CA, USA

Francisco Vera-Sirera Instituto de Biología Molecular y Celular de Plantas (IBMCP), Universidad Politécnica de Valencia-Consejo Superior de Investigaciones Científicas, Ciudad Politécnica de la Innovación (CPI), Valencia, Spain

Ivana L. Viola Instituto de Agrobiotecnología del Litoral (CONICET-UNL), Cátedra de Biología Celular y Molecular, Facultad de Bioquímica y Ciencias Biológicas, Universidad Nacional del Litoral, Santa Fe, Argentina

Jia-Wei Wang National Key Laboratory of Plant Molecular Genetics, Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Shanghai, China

Ditte H. Welner Joint BioEnergy Institute, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Kazuhiko Yamasaki Biomedical Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan; RIKEN Quantitative Biology Center, Tsurumi-ku, Yokohama, Japan

Shuichi Yanagisawa Laboratory of Plant Functional Biotechnology, Biotechnology Research Center, The University of Tokyo, Yayoi 1-1-1, Bunkyo-ku, Tokyo, Japan

Preface

Transcription factors (TFs) are central regulators of gene expression and, as such, modulate essential aspects of organismal function, including cell differentiation, tissue and organ development, responses to hormones and environmental factors, metabolic networks, and disease resistance, among others. Intrinsic to TF action is their capacity to specifically interact with DNA sequences and with other proteins as part of transcriptional complexes involved in the regulation of gene expression. Accordingly, knowledge of the structure of TFs and of the molecular mechanisms involved in the establishment of these interactions is essential to understand TF action. Plants contain a vast number of TFs (about 10% of plant genes encode TFs) that were acquired at different stages of evolution and were adapted to perform specific functions.

This book is intended as a source of information for those interested in the study of plant TFs and the many processes they regulate. It contains basic information that can be useful to students and researchers entering the field as well as more specific chapters devoted to plant TF families. These specific chapters do not constitute a comprehensive list of what is known about the different TF families but are rather examples of how the study of the different aspects of a specific TF family can be useful to establish the molecular aspects of TF function. Section A deals with general aspects of plant TFs. It contains two introductory chapters that describe the basics of TF action and methods usually employed to study TFs, followed by chapters that discuss structural and evolutionary aspects of plant TF families and plant-specific TF DNA-binding domains. Sections B and C present information about the structure, evolution, and functional aspects of several plant TF families, with examples of families that arose at different stages of organismal evolution and were adapted to modulate specific aspects of plant developmental programs and responses. Finally, Section D contains chapters that discuss aspects of the posttranslational regulation of plant TF action by either intra- or intercellular movement, proteolytic processing, ubiquitination, or redox interconversions. The book is centered on TFs rather than on processes, understanding that there are excellent books that already describe plant regulatory networks and the TFs involved. These books, however, often describe interactions established by TFs in regulatory networks but do not deepen into the structural aspects of the TFs involved. I hope that through this book readers will acquire a general view of different aspects of plant TFs that eventually will help to fill the gap existing between the knowledge of the participation of a TF in a defined process and the establishment of the structural properties related to TF functions. In other words, to establish structural–functional relationships that explain in detail the molecular mechanisms involved in TF action.

I would like to thank all authors who generously contributed their chapters. As experts in the field, their contribution was essential for the assembly of this book. I would also like to thank Mary Preap, from Academic Press/Elsevier, for her valuable assistance, and people from my lab for their contributions along many years. Finally, I also acknowledge support from the Argentine Research Council (Consejo Nacional de Investigaciones Científicas y Técnicas, CONICET) and Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT) for their support to my research activities.

Daniel H. Gonzalez

Contents

D

MODULATION OF PLANT TRANSCRIPTION FACTOR ACTION

SECTION A

GENERAL ASPECTS OF PLANT TRANSCRIPTION FACTORS

CHAPTER

1

Introduction to Transcription Factor Structure and Function

Daniel H. Gonzalez

Instituto de Agrobiotecnología del Litoral (CONICET-UNL), Cátedra de Biología Celular y Molecular, Facultad de Bioquímica y Ciencias Biológicas, Universidad Nacional del Litoral, Santa Fe, Argentina

OUTLINE

1.1 INTRODUCTION: TRANSCRIPTION IN EUKARYOTES

In eukaryotes, various RNA polymerases are responsible for the transcription of nuclear genes (Roeder and Rutter, 1969). Particularly, RNA polymerase II is in charge of transcribing all protein-coding genes, in addition to several genes that encode noncoding RNAs (Kornberg, 2007; Cramer et al., 2008). RNA polymerase II promoters are composed of a big number of discrete DNA sequences (also named boxes or elements), usually located upstream of the transcription start site, but also within and downstream of transcribed regions (Lenhard et al., 2012). These sequences can be classified as basal promoter elements, upstream promoter elements, and enhancers, and are the binding sites of proteins named transcription factors, which influence the transcription of genes linked to them (Figure 1.1). Basal promoter elements are usually located near the transcription start site (Juven-Gershon et al., 2008) and are the binding site of general transcription factors that participate in the expression of most genes by promoting the binding of RNA polymerase II (Li et al., 1994; Orphanides et al., 1996; Roeder, 1996; Conaway and Conaway, 1997; Reese, 2003). The most common basal promoter element is the TATA box, recognized by TATA-box binding protein (TBP) (Peterson et al., 1990; Burley, 1996), a component of the general transcription factor II D (TFIID) (Horikoshi et al., 1990). Upstream promoter elements are very diverse. They are located further upstream of basal elements (up to several hundred base pairs of the transcription start site) and are recognized by specific transcription factors according to the type of elements present in each gene (Mitchell and Tjian, 1989; Ptashne and Gann, 1997; Lee and Young, 2000). In most genes, the presence of these elements is necessary for efficient transcription since the sole interaction of general transcription factors with the basal promoter is not enough to assemble a stable transcriptional complex (Gill, 1996; Stargell and Struhl, 1996; Struhl et al., 1998). In addition, many of these elements are required for the transcriptional regulation of gene expression under different circumstances, thus receiving the name of response elements. Enhancers are regions of the genome that affect the expression of particular genes linked to them (Stadhouders et al., 2012; Smallwood and Ren, 2013;

Plant Transcription Factors. http://dx.doi.org/10.1016/B978-0-12-800854-6.00001-4

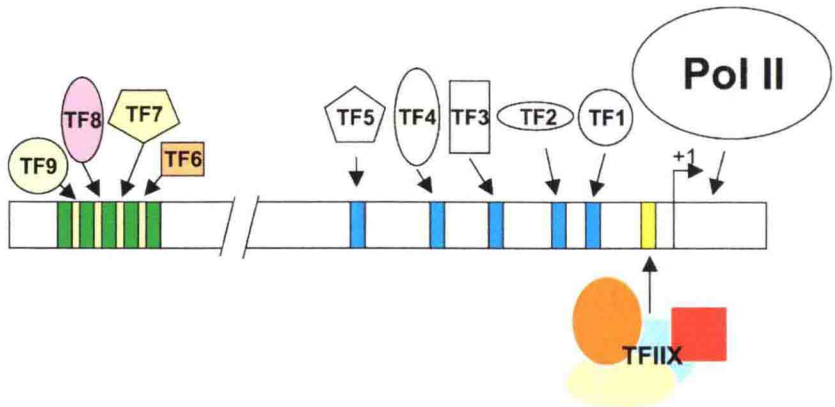

FIGURE 1.1 **The structure of eukaryotic promoters.** Eukaryotic gene promoters are composed of discrete binding sites for multiple transcription factors dispersed over long distances (usually several thousands of base pairs). General transcription factors (TFIIX) for RNA polymerase II (Pol II) interact with sequences located near the transcription start site (yellow). Specific transcription factors (TF1 to TF9) recognize particular sequences located in proximal promoter regions (blue; at hundreds of base pairs of the start site) or in enhancers (green; at thousands of base pairs of the start site). The transcriptional activity of a gene will be defined by the nature of the transcription factors bound in different regions of its promoter. The transcription start site is indicated by +1.

Levine et al., 2014). Usually, they are not classified as part of the promoter, although their action is required for the correct transcriptional activity of the corresponding genes. Enhancers contain groups of response elements and have the peculiarity of acting at long distances (up to several thousands of base pairs), through the formation of loops in DNA. The transcription characteristics of a gene will then be established by the nature of the different elements that compose its promoter region, including enhancers, and the interactions established by different proteins with these elements and among themselves. Apart from the obvious presence of the appropriate partners, the interaction of promoter elements with the corresponding binding proteins will be influenced by the structure of the chromatin in that particular region of the genome, which leads to an additional source of complexity (Li et al., 2007; Cairns, 2009; Venters and Pugh, 2009; Voss and Hager, 2014).

1.2 STRUCTURE OF TRANSCRIPTION FACTORS

Transcription factors are proteins that influence the transcription of genes by binding to defined regions of the genome (Latchman, 1997). Genes encoding transcription factors constitute 3–10% of all genes in eukaryotic genomes (Levine and Tjian, 2003; Harbison et al., 2004; Reece-Hoyes et al., 2005). A basic feature of transcription factors is that they contain DNA-binding domains that recognize specific sequences within the promoter regions of the genes they regulate (Figure 1.2A; Kummerfeld and Teichmann, 2006). By binding to these sequences they either increase or decrease the transcription of their target genes, thus acting as activators or repressors, respectively. Activation or repression is usually achieved through interaction with other components of the transcription apparatus (Figure 1.2B,C; Takagi and Kornberg, 2006), which makes protein–protein interactions another important feature of transcription factor action (Walhout, 2006). Activation or repression of gene expression can also be achieved through interaction with chromatin-modifying enzymes, which are then directed to modulate the accessibility of the transcription machinery to specific regions of the genome (Figure 1.2B,C). The capability of a transcription factor to act as a repressor or an activator is, in most cases, dependent on domains that are located outside, and act independently of, the DNA-binding domain (Figure 1.2A). This brings transcription factors a modular structure and the possibility of acquiring new properties by domain mixing or shuffling, a process used by evolution and researchers to generate new mechanisms of transcriptional regulation (Gossen and Bujard, 1992; Morgenstern and Atchley, 1999; Beerli et al., 2000; Ansari and Mapp, 2002; Traven et al., 2006; Liu et al., 2013).

1.3 DNA RECOGNITION BY TRANSCRIPTION FACTORS

The recognition of specific DNA sequences by transcription factors is achieved by interactions established between the side chains of amino acids of the DNA binding domain with nucleotides of the target site (Figure 1.3). For specific recognition, interactions must be established with the nucleotide bases that are located inside the DNA double-helical structure. For this reason, most transcription factors establish connections with DNA by binding to the major groove, although interactions through the

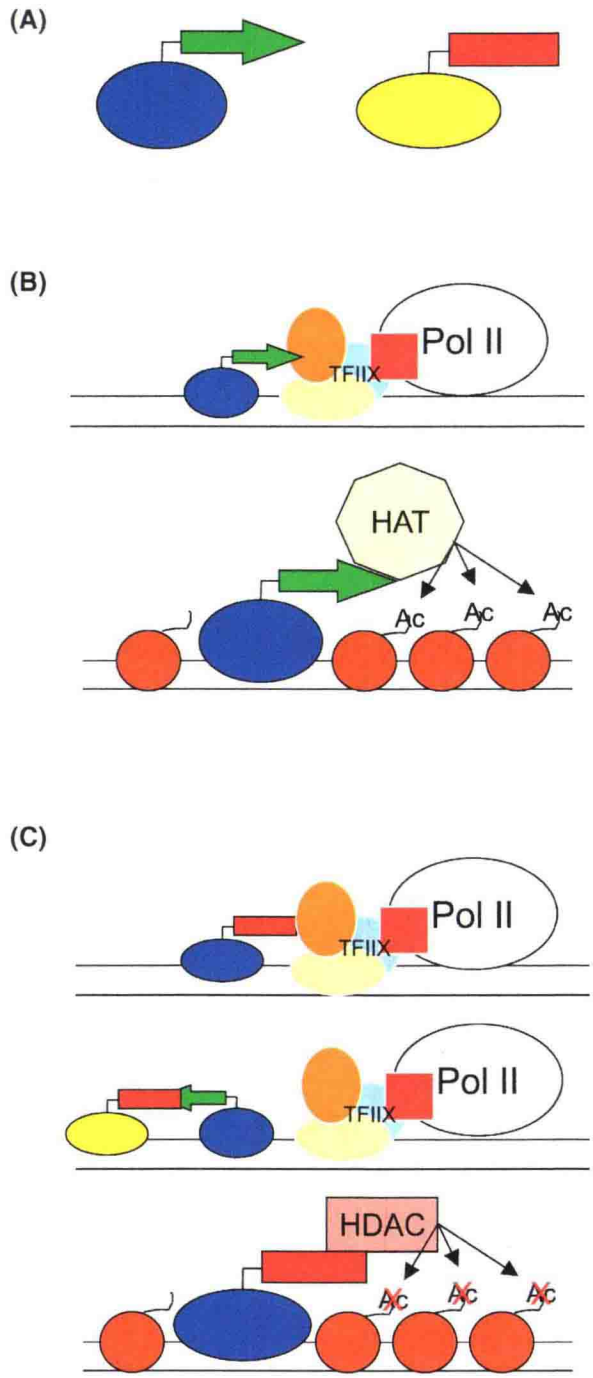

FIGURE 1.2 **The general structure of transcription factors.** (A) Most transcription factors have a modular structure. They contain a DNA binding domain involved in the recognition of specific DNA sequences (blue and yellow ovals). Usually, they also contain activation (green arrow) or repressor (red rectangle) domains that increase or decrease, respectively, the transcription of genes to which they bind. (B) Activation can be achieved through stabilizing protein–protein interactions with other components of the transcription machinery or the recruitment of chromatin modifying enzymes, like HAT, that relax nucleosome structure. (C) Repression can be achieved by interfering with activators or recruiting chromatin-modifying enzymes, like histone deacetylases (HDAC), that increase compaction of nucleosomes.

minor groove have also been reported in several cases (Figure 1.3). Amino acid side chains of transcription factor DNA-binding domains can establish specific interactions with bases through hydrogen bonding and van der Waals contacts (Figure 1.3B,C; Shimoni and Glusker, 1995; Suzuki et al., 1995; Luscombe et al., 2001; Rohs et al., 2010). While these interactions determine the specificity, the strength, or affinity of the interaction is additionally determined by unspecific contacts established with the sugar phosphate backbone, including ionic interactions between DNA phosphates and positively charged residues of the DNA-binding domain. Another factor that influences the strength and specificity of the interaction is the topology of the DNA around the transcription factor-binding site (Pan et al., 2010). Curvatures in DNA are often required by transcription factors to bind their target genes efficiently (Rohs et al., 2009), and some transcription factors induce DNA bending upon binding (van der Vliet and Verrijzer, 1993), thus leading to changes that facilitate other processes, like DNA melting or the binding of additional proteins.

1.4 DNA-BINDING DOMAINS

DNA-binding domains adopt different structures, and the interaction of these domains with DNA can be established through alpha helices, beta sheets, or disordered regions (Figure 1.3; Pabo and Sauer, 1992). Usually, the DNA-binding domain forms a module that can be separated from the rest of the transcription factor structure without losing activity. This facilitates structural studies of the isolated DNA-binding domains or their complexes with DNA using techniques that require low molecular weight proteins, like crystallization or NMR. DNA-binding domains are named according to their structural characteristics, and most organisms contain several transcription factors that share the same type of DNA-binding domain. Accordingly, transcription factors are classified in families that usually receive the name of the respective DNA-binding domain (Table 1.1; Stegmaier et al., 2004; Vaquerizas et al., 2009; Charoensawan et al., 2010). Transcription factors that share the same type of DNA-binding domain (in other words, transcription factors from the same family) tend to have more similar DNA-binding specificities than those that belong to different families. In any case, variations in DNA-binding specificity are often observed within the same family, and these are most often due to changes in specific residues of the DNA-binding domain (Berger et al., 2008; Noyes et al., 2008; Badis et al., 2009). Thus, changes in the amino acid residues of the DNA-binding domain are also used by evolution and researchers to create transcription factors with novel DNA-binding characteristics (Blancafort et al., 2004; Amoutzias et al., 2007; Joung and Sander, 2013).

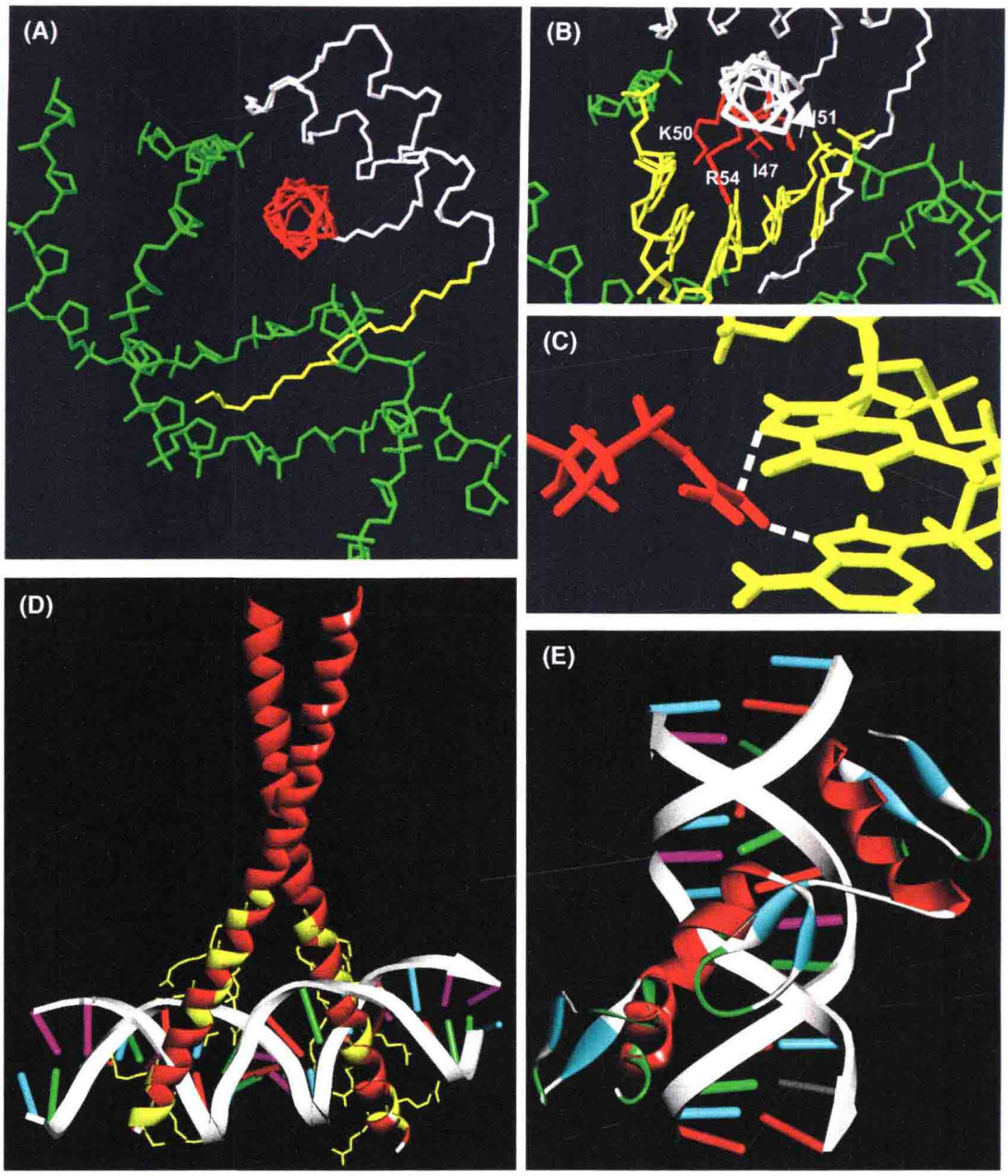

FIGURE 1.3 **DNA recognition by transcription factors.** (A) Specific interactions of the *Drosophila* bicoid homeodomain transcription factor with DNA (Baird-Titus et al., 2006) are established by an alpha helix (helix III, in red) that inserts into the major groove of DNA (green). Additional interactions are established by a disordered N-terminal arm (yellow) along the adjacent minor groove (only the path of backbone atoms is shown). (B) For specific recognition, the side arms of bicoid helix III amino acids Ile47, Lys50, Asn51, and Arg54 (red) establish specific contacts with bases of the DNA (yellow). (C) For example, bicoid Arg54 forms hydrogen bonds with G and A of the recognition sequence GGATTA. (D) A dimer of the yeast bZip transcription factor GCN4 (Keller et al., 1995, in red) interacts with DNA through alpha helices that run across two adjacent major grooves. Specific contacts are established mainly by basic amino acids (yellow). (E) Zinc finger transcription factors contain DNA-binding modules formed by adjacent alpha helices and beta hairpins. Each module binds DNA through the major groove and is connected to adjacent modules by a disordered region. The different modules wrap along the major groove of DNA (Kim and Berg, 1996).

1.5 PROTEIN–PROTEIN INTERACTIONS

Many transcription factors require the formation of dimers to bind DNA (Figure 1.4A,B). These dimers can form between two identical molecules (homodimers) or between different molecules (heterodimers), usually from the same transcription factor family (Amoutzias et al., 2008; Funnell and Crossley, 2012; Pogenberg et al., 2014). Typical transcription factors that require dimerization for binding are those of the bZip and bHLH families (Figure 1.4A,B), which use a basic region to interact with DNA and in which the dimerization domain