FEBS 22810 FEBS Letters 460 (1999) 401–410

Minireview

Transcriptional activation by the Myb proteins requires a specific local promoter structure

Brigitte Ganter, Samuel T. Chao, Joseph S. Lipsick*

Department of Pathology, Stanford University, School of Medicine, Stanford, CA 94305-5324, USA

Received 28 July 1999; received in revised form 30 September 1999

Abstract The biological effects of the cellular c-Myb and the viral v-Myb proteins are strikingly different. While c-Myb is indispensable for normal hematopoiesis, v-Myb induces acute leukemia. The v-Myb DNA-binding domain (DBD) differs from that of c-Myb mainly by deletion of the first of three repeats which correlates with efficient oncogenic transformation and a decrease in DNA-binding activity. To investigate the difference in DNA-binding and transcriptional activation, oligonucleotide selection and electrophoretic mobility shift assays were employed. The v-Myb DBD (R2R3) shows an intrinsic DNA-binding specificity for an AT-rich downstream extension of the Myb recognition element (MRE) PyAAC$^{T}/_{G}$G for efficient binding to this site, whereas R1 within the c-Myb DBD allows for more flexibility for this downstream extension. Therefore, due to the presence of repeat R1, c-Myb can bind to a greater number of target sites. The intrinsic DNA-binding specificity of R2R3 is further supported with the results from in vivo transcriptional activation experiments which demonstrated that both the v-Myb and c-Myb DBDs require an extension of the MRE (motif #1) by a downstream T-stretch (motif #2) for full activity. Surprisingly, the T-stretch improves binding when present on either strand, but is required on a specific strand for transcriptional activation.

© 1999 Federation of European Biochemical Societies.

Key words: Myb DNA-binding domain; Transcriptional activator; T-stretch; Upstream activation sequence; Repeat R1

1. Introduction

The *myb* oncogene, which is unique in its hematopoietic specificity of transformation, was first discovered in two independently derived avian retroviruses that cause acute leukemia (for review, [1]). Both AMV and the E26 virus encode v-Myb proteins that are truncated versions of the cellular c-Myb protein. The normal *c-myb* proto-oncogene encodes a 75 kDa nuclear protein that is expressed at high levels in immature hematopoietic cells. Its expression decreases as these cells differentiate [2,3]. Lack of *c-myb* results in failure of intrafetal hematopoiesis [4]. On the other hand, the AMV *v-myb* oncogene specifically blocks monoblast differentiation which results in transformation of macrophage precursors [5]. The fact that these two related proteins cause such dramatic cellular effects provides a framework for studying how the viral and the cellular Myb proteins interact with DNA target sites.

The c-Myb DNA-binding domain (DBD) consists of three tandem imperfect repeats of approximately 50 amino acids (R1, R2 and R3) [6]. During the genesis of v-Myb, deletion of most of the first repeat (R1) has occurred ([7–9]. The minimal region of c-Myb that is both necessary and sufficient for sequence-specific DNA-binding has been narrowed down to repeats R2 and R3 [7,10–12] and unlike many other sequence-specific DNA-binding proteins, the Myb proteins appear to bind to DNA as monomers [10,11]. Though repeat R1 was not shown to be involved in sequence-specific recognition, it has been suggested to increase the affinity of the protein for DNA and the stability of the Myb-DNA complex [13–17].

Selection of a pool of chicken genomic DNA fragments led to the identification of the v-Myb recognition sequence PyAAC$^{T}/_{G}$G (Myb recognition element = MRE) [7]. Three such MREs were found 5′ of the transcriptional start site of the E26-inducible *mim-1* gene [18]. Recently, three independent groups have reported the results of NMR-based studies of the minimal c-Myb R2R3 DBD [19,20] and the minimal B-Myb R2R3 DBD [21,22]. Each repeat contains three conserved tryptophans spaced 18 or 19 amino acids apart [23] that form a hydrophobic core [24] and that play a critical role in sequence-specific DNA-binding [23,25,26]. Each repeat can fold into three helices that form a helix-turn-helix-related structural motif as was shown by NMR spectroscopy [20–22,24]. Though there is a high degree of amino acid conservation between individual repeats, the structural data imply that repeats R2 and R3 of both the c-Myb and the B-Myb DBD do not fold into the same exact tertiary structure [19,21,22]. R3 folds into a helix-turn-helix-related motif of which the C-terminally located third helix is the recognition helix which makes specific contacts within the major groove of the DNA double helix. Repeat R2 does not contain such a stable preformed recognition helix at its C-terminus. Therefore, it has been suggested that the conformational instability of the DNA-binding site on the protein may be necessary to allow for binding to a number of different specific DNA target sites.

Besides the N-terminal truncation which deletes most of repeat R1 of the Myb DBD, four amino acid substitutions are present within the v-Myb DBD. As expected from the similarity within the DBD region, the only observed difference in terms of binding of these two proteins to the MRE is a reduced affinity of the v-Myb protein by a factor of 5–6 [13,14,17] when compared to the c-Myb protein ($K_d \approx 4 \times 10^{-10}$ M, [17]). Two extended consensus sequences were derived from an amplification-based selection of a pool of random sequences: the 9 bp YAACBGYCR [27] and the 8 bp YAACKGHH [28]. These two extended sites were found

*Corresponding author. Fax: (1) (415) 725-6902.
E-mail: lipsick@stanford.edu

0014-5793/99/$20.00 © 1999 Federation of European Biochemical Societies. All rights reserved.
PII: S0014-5793(99)01373-3

both for the v-Myb and the chicken [28] and mouse [27] c-Myb protein.

Consistent with this finding, purified c-Myb R2R3 and c-Myb R1R2R3 protect the same target sites with the same positions of the borders in exonuclease-III footprinting assays [14]. In contrast, a large difference was observed in how the border in the two complexes withstood exonuclease attack [14] which was attributed to the presence of repeat R1 in the c-Myb DBD. A stabilizing effect was observed for the downstream border once repeat R1 was present in the DNA-protein complex [14].

To better understand how individual repeats of the c-Myb and the v-Myb protein function in binding to specific DNA targets in order to activate transcription, different repeat combinations were used in this study to investigate the c-Myb and v-Myb-DNA interactions in relation to different MRE flanking sequences.

2. Materials and methods

2.1. Random site selection

Selection and amplification of high-affinity binding sites for the Myb proteins were performed essentially as described elsewhere [29]. For selection, the 60 base oligonucleotide PCRP-RSS#6 (5′-CGGGATATCTCTAGAGATGG(N_{20})GAGGCCTGCAGGATATCGTA-3′) was used, which contains 20 random nucleotides flanked on either side by 20 bases of defined sequence to permit PCR amplification. PCR amplification involved the primers PCRP-RSS#4 (5′-CGGGATATCTCTAGAGATGG-3′) and PCRP-RSS#5 (5′-TACGATATCCTGCAGGCCTC-3′) in a 50 μl total reaction volume for 20 cycles of 95°C for 45 s/59°C for 1 min/73°C for 30 s. After precipitation of the reaction mix, the DNA was labelled using T4 DNA polymerase (New England Biolabs) and [α-^{32}P]dCTP. Binding reactions of end-labelled DNA and purified proteins were carried out in a final volume of 20 μl containing 20 mM Tris-Cl (pH 7.5), 3 mM $MgCl_2$, 50 mM NaCl, 15% (v/v) glycerol and 2 μg poly [dI-dC]. The binding mixture was incubated on ice for 10 min, separated using a 6% polyacrylamide gel (29:1 acrylamide:bisacrylamide) and radiolabelled DNA-binding sites were excised from the gel. After elution and purification, half of the DNA was used for the next round of DNA amplification and selection. After the fifth round of PCR amplification, the gel-purified DNA fragments were cloned into the TA-cloning vector pCRII (Invitrogen) and the inserts were sequenced by the dideoxy sequencing method [63].

2.2. Electrophoretic mobility shift assay (EMSA)

The methods used for EMSAs are essentially as described elsewhere for other DNA-binding proteins [64] and include separating free DNA from protein-complexed DNA on electrophoretic mobility shift gels. Double-stranded DNA probes with four nucleotide overhangs were 3′ end-labelled by end-filling with [α-^{32}P]dATP and Klenow enzyme (New England Biolabs). Binding reactions were carried out in a final volume of 10 μl containing 20 mM Tris-Cl (pH 7.5), 3 mM $MgCl_2$, 50 mM NaCl, 15% (v/v) glycerol and 2 μg poly [dI-dC], end-labelled DNA fragment and protein sample. Bacterially expressed Myb DBDs were used for the EMSA. The binding mixture was incubated for 10 min on ice before separated on a 6% polyacrylamide gel (29:1 acrylamide:bisacrylamide).

2.3. Cell culture, transient DNA transfection, nuclear extracts and β-galactose and luciferase assays

Quail QT6 cells were grown in Dulbecco's modified essential Eagle's medium (DMEM) supplemented with 5% fetal calf serum, glucose (4.5 g/l), 1×non-essential amino acids, 2 mM L-glutamine, 1 mM sodium pyruvate, streptomycin (100 μg/ml) and penicillin (100 U/ml) in a humidified 10% CO_2-90% air incubator at 37°C. Transient transfections were performed by a modified calcium phosphate precipitation method [65,66]. Myb-expressing activator plasmids or control vector-only plasmid N-Cla (3 μg), luciferase reporter plasmid (1 μg), carrier tRNA (to give a total of 10 μg of nucleic acid for transfection) and 0.5 μg of internal control plasmid expressing β-galactosidase from the cytomegalovirus promoter (CMV-β-gal) were co-transfected into approximately 10^6 QT6 cells per 10 cm diameter plate. Two days later, the cells from each transfection plate were scraped in phosphate-buffered saline. Half of these cells were used for a luciferase assay [67,68] and a β-galactosidase assay [69], the other half was solubilized by boiling for 3 min in sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) loading dyes and subjected to immunoblotting to demonstrate expression of Myb in transfected cells.

The method used to prepare nuclear cell extracts which include Myb proteins is essentially as described by others [70].

2.4. Western blot analysis

Normalized volumes of QT6 lysates (based on internal β-Gal activity for transient transfections) were subjected to SDS-PAGE (10–15% polyacrylamide) and then, the transiently expressed proteins were transferred to a nitrocellulose membrane (BA-S 83; Schleicher and Schuell) and detected with anti-Myb antibody (mouse monoclonal Myb 5E antibody, [71]). Immunodetection was performed using an enhanced chemiluminescence system (Pierce).

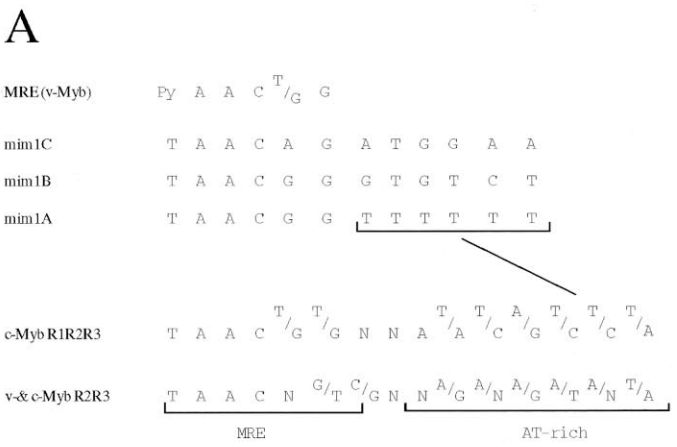
3. Results and discussions

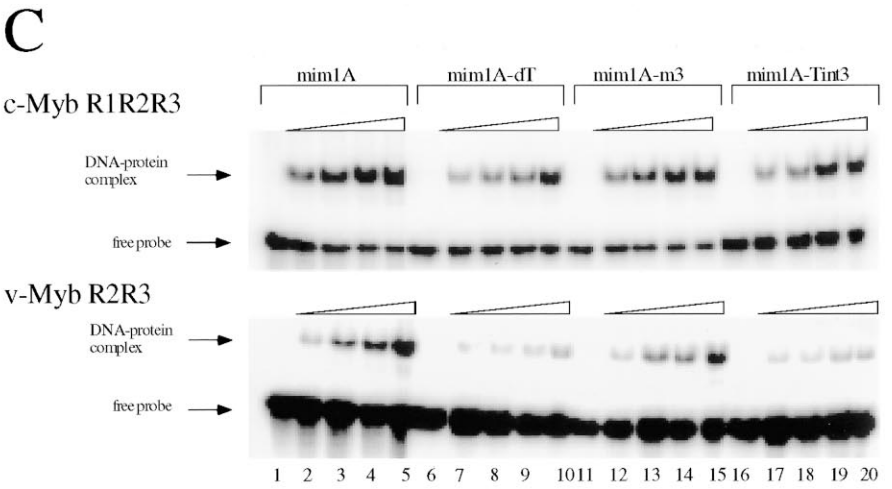
3.1. v-Myb requires an AT-rich downstream extension of the MRE for efficient binding

In an effort to clarify whether the sole function of repeat R1 is to contribute to the stability of the c-Myb-DNA complex as previously suggested [13,14] and not to be involved in sequence-specific DNA-binding, we employed the random oligonucleotide site selection strategy [29] using different combinations of purified Myb DBD repeats (c-Myb R1R2R3, c-Myb R2R3, c-Myb R1R2 and v-Myb R2R3). From the obtained selected sequences, several conclusions can be drawn. As shown in Fig. 1A, the majority of the sequences obtained for c-Myb R1R2R3 and v-Myb R2R3 contain the proposed v-Myb consensus site TAACN$^{G}/_{T}$ with the MRE core element AAC which has been shown before by others

→

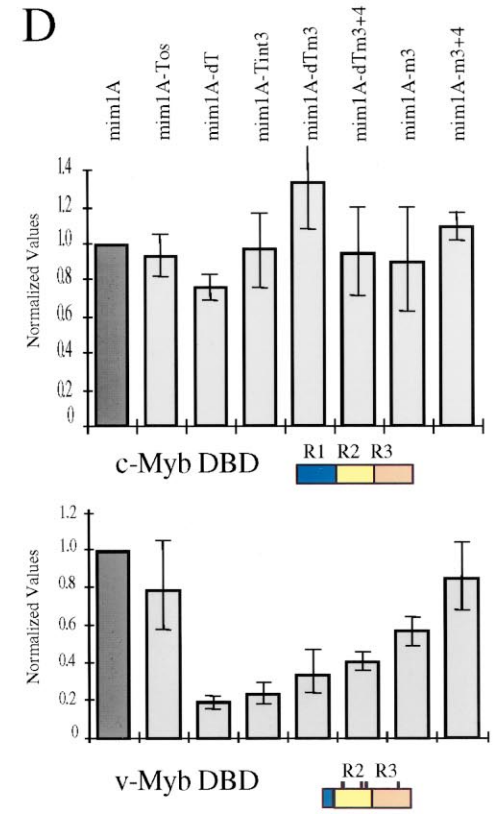
Fig. 1. A: Binding site selection data with purified Myb DBDs. Consensus Myb-binding sites derived for c-Myb R1R2R3, c-Myb R2R3 and v-Myb R2R3 from analysis of selected sequences. The three MREs of the 5′ non-coding region of the *mim-1* gene are shown as a comparison. B–D: DNA-binding assays using EMSA. B: Sequence of double-stranded DNAs used in the DNA-binding assays. All the oligonucleotides used were 26 bp long. The MRE region along with the downstream flanking T-stretch are shown in bold. Every base that is different in comparison to the original mim1A site is underlined. * Highlights the bases that represent a mismatch at position 3 or 4 within the T or the dT sequence (mim1A-dTm3, mim1A-dTm3+4, mim1A-m3 or mim1A-m3+4). C: The v-Myb DBD shows a dependency for the downstream AT-rich sequence. Analyses of c-Myb R1R2R3 and v-Myb R2R3 -mim1A, -mim1A-dT, -mim1A-m3 and -mim1A-int3 complexes were performed (only some examples are shown). Free DNA is shown in lanes 1, 6, 11 and 16. Binding reactions with an increasing amount of the c-Myb R1R2R3 protein (upper gel) or v-Myb R2R3 (lower gel) are shown in lanes 2–5, 7–10, 12–15 and lanes 17–20. Myb-specific DNA-protein complexes and free DNA probes are marked on the left of the gel. D: Quantitative analysis of EMSAs done in B. Shown are averages of three experiments. For quantitative analysis of the DNA-binding assays, we compared each fraction of DNA bound by equal amounts of protein and measured the relative amount of radioactivity in free versus bound DNA using the phosphorimager (Molecular Dynamics). The relative values were normalized against the corresponding mim1A value for each protein concentration. For each DNA fragment, the normalized values from the various protein concentrations were averaged to assign a value for protein affinity.

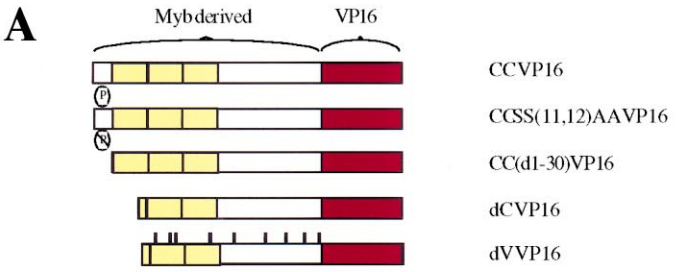


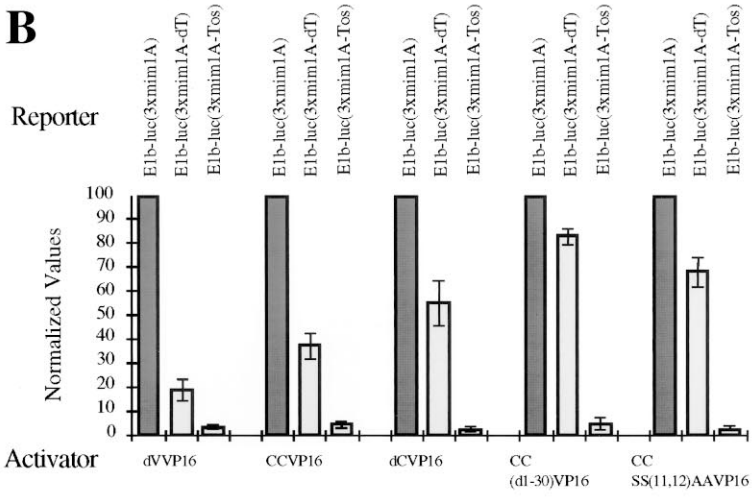


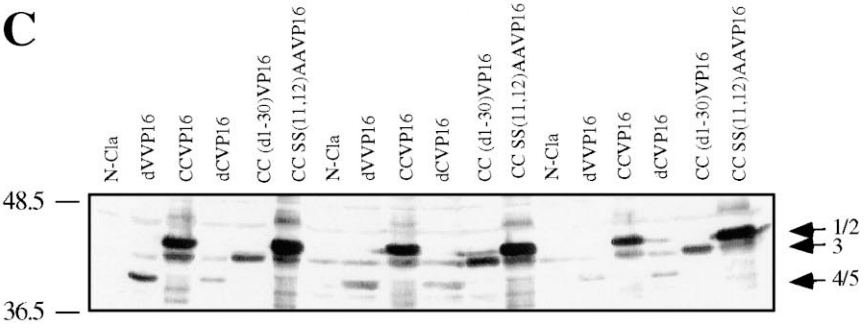
B

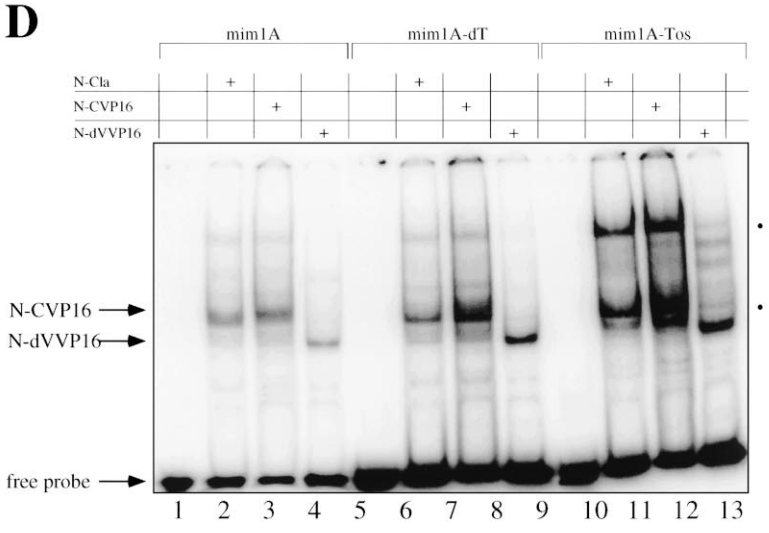
```
mim1A          5'-TCGANNNTAACGGTTTTTTNNN-3'
                  3'- NNNATTGCCAAAAAANNNAGCT-5'

mim1A-Tos            /NNNTAACGGAAAAAANNN
                      NNNATTGCCTTTTTTNNN/

mim1A-dT             /NNNTAACGGTTGACCNNN
                      NNNATTGCCAACTGGNNN/

mim1A-Tint3          /NNNTAACGGTTGTTTNNN
                      NNNATTGCCAACAAANNN/
                               *
mim1A-dTm3           /NNNTAACGGTTGACCNNN
                      NNNATTGCCAAATGGNNN/
                                **
mim1A-dTm3+4         /NNNTAACGGTTGACCNNN
                      NNNATTGCCAAACGGNNN/
                                *
mim1A-m3             /NNNTAACGGTTGTTTNNN
                      NNNATTGCCAAAAAANNN/
                                **
mim1A-m3+4           /NNNTAACGGTTGGTTNNN
                      NNNATTGCCAAAAAANNN/
```




A
Myb derived
VP16
CCVP16
CCSS(11,12)AAVP16
CC(d1-30)VP16
dCVP16
dVVP16


B
Reporter
E1b-luc(3xmim1A)
E1b-luc(3xmim1A-dT)
E1b-luc(3xmim1A-Tos)
Normalized Values
100
90
80
70
60
50
40
30
20
10
0
Activator
dVVP16
CCVP16
dCVP16
CC (d1-30)VP16
CC SS(11,12)AAVP16


C
N-Cla
dVVP16
CCVP16
dCVP16
CC (d1-30)VP16
CC SS(11,12)AAVP16
48.5
36.5
1/2
3
4/5


D
mim1A
mim1A-dT
mim1A-Tos
N-Cla
N-CVP16
N-dVVP16
free probe
1 2 3 4 5 6 7 8 9 10 11 12 13

[7,27,28]. Second, roughly 75% of the MRE containing sequences selected with c-Myb R1R2R3 and 85% of the sequences selected with v-Myb R2R3 showed a region rich in As and/or Ts downstream of the MRE. The c-Myb DBD showed a preference for Ts, whereas the v-Myb DBD showed a preference for As. Similar results were obtained for v-Myb R2R3 and c-Myb R2R3 without the amino acid substitutions of AMV. Third, site selection experiments with the c-Myb R1R2 protein, which binds only weakly to the mim1A site [13], did not yield any specific target DNA (results not shown), which tells us that these two repeats are not sufficient to bind to DNA specifically. This is particularly interesting because it suggests that even though there is a high degree of sequence conservation between individual c-Myb DBD repeats (R1, R2 and R3), they do not seem to fold into the same tertiary structure. As mentioned above, this has already been shown with the NMR structure for repeats R2 and R3 of the minimal c-Myb R2R3 protein [19] and the minimal B-Myb R2R3 protein [21]. In addition, several different transcriptional regulators have been identified that contain more distantly related Myb repeats [30]. These Myb repeat containing proteins can be grouped into different families, including the telobox family [31], the SANT domain family of transcriptional regulators [32] and the transcription terminator family [33]. Except for the SANT domain containing proteins, all of these Myb domain containing proteins do have intrinsic DNA-binding activity. Therefore, the Myb repeats within the SANT domain containing proteins must have another function. This function could either be involved in protein-protein interaction such as has been shown for cyclin D [34] or it could be involved in unspecific protein-DNA interaction as is suggested for repeat R1 [13,14]. From these site selection data, we conclude that repeat R1 does not contribute to specific DNA-binding, but increases the affinity of c-Myb for the MRE, especially in the presence of an AT-rich downstream sequence.

As shown in Fig. 1A the mim1A element of the 5′ region of the *mim-1* gene contains a stretch of six Ts downstream of the MRE TAACGG [18]. This is particularly interesting since the mim1A element is a naturally occurring MRE for which the c-Myb protein shows the highest affinity [18]. Given the close resemblance of the selected sequences for both the c-Myb and v-Myb DBD with the strong mim1A site, we decided to investigate the relationship of this T-stretch to the MRE in terms of DNA-binding and subsequent *trans*-activation.

Certain nucleotide sequences spontaneously impart a preferred direction of curvature on a DNA molecule, for example by originating an intrinsic bend [35,36]. Although A or T tracts are believed to be major determinants of DNA curvature, the sequence context also plays a critical role [37]. PolyA/polyT tracts are particularly rigid and bending may occur preferentially at the A/T boundaries [38–40]. In particular, polyA/polyT tracts adopt a Z-like DNA conformation and this gets separated from normal B-like DNA by a local bend at the junction of the two structures [40]. Therefore, the stretch of six Ts downstream of the MRE in the mim1A site might have a dramatic effect on the local promoter structure. In addition, sequences that are rich in ATs compared to GC-rich sequences have a lower melting temperature and are therefore less stable. In other words, if binding to a protein to DNA requires melting of one or more base pairs, then, the binding of mismatched base pairs should be favored over AT pairs that in turn should bind better than GC pairs. To gain insight into the structural effects of DNA sequences in the formation of specific Myb-MRE complexes, we used EMSAs to test the effect of different downstream flanking sequences upon Myb-binding to the same core MRE. We designed different double-stranded DNAs (see Fig. 1B) to be tested with quantitative EMSA to be able to discriminate between the two different possibilities, such as bending or melting.

Interestingly, when the v-Myb DBD was tested in comparison to the c-Myb DBD, we observed that the v-Myb R2R3 protein shows a much greater requirement for the downstream AT-rich sequence than the c-Myb R1R2R3 protein (see Fig. 1C,D). This preference was also reflected in the pool of selected oligonucleotides: 85% of oligonucleotides selected by v-Myb R2R3 contain an AT-rich downstream sequence compared to only 75% selected by c-Myb R1R2R3. Deletion of the T-stretch and introduction of a random sequence (mim1A-

←

Fig. 2. *Trans*-activation by the Myb proteins requires an extension of the MRE by a T-stretch. A: Structure of the different Myb-VP16 fusion proteins used in this study. The c-Myb VP16 protein (CCVP16), v-Myb VP16 protein (dVVP16) and various mutant forms are diagrammed. The gray boxes indicate the highly conserved Myb DBD, which is composed of three imperfect direct repeats of approximately 50 amino acids each. The black box represents the 78 amino acid transcriptional activation domain of the herpes simplex virus VP16 which is fused to various Myb N-termini as diagrammed. The different activator constructs used have been described elsewhere: CCVP16 (c-Myb VP16), dCVP16 and dVVP16 (v-Myb VP16) [5,13]. The truncations of Myb are named according to the sequence present in c-Myb (C) or absent in v-Myb (D) [13]. The vertical bars within dVVP16 indicate amino acid substitutions in v-Myb relative to c-Myb. The c-Myb mutant designated CC SS(11,12)AAVP16 represents site-directed mutagenesis of the CKII phosphorylation sites, whereas the c-Myb mutant designated d1–30 (CC(d1–30)VP16) has a N-terminal deletion of 30 amino acids including the CKII sites and an adjacent acidic region just upstream of the Myb DBD [13]. B: *Trans*-activation by different Myb VP16 fusion proteins. The animal cell reporter constructs polyA-E1b-luc(3×mim1A) (three tandem TAACCG**TTTTTT** sites), polyA-E1b-luc(3×mim1A-dT) (three tandem TAACCG**TTGACG** sites) and polyA-E1b-luc(3×mim1A-Tos) (three tandem TAACCG**AAAAAA**T sites) all harbor a simian virus 40 polyA site and three mim1A Myb-binding sites (underlined) with different downstream flanking sequences (bold) upstream of an E1b TATA box and the luciferase gene followed by another SV40 polyA site. Activities are shown for each protein relative to the E1b-luc(3×mim1A) reporter construct, which was assigned a value of 100 (black bar) after normalizing for the transfection efficiency using a CMV-β-gal reporter. Data are averages of three experiments. Error bars indicate S.D.s. C: Immunoblot analysis of Myb proteins of the same transfected cell extracts. Relative mobilities of transfected Myb VP16 fusion proteins are labelled on the right of the gel: 1=CCVP16 (c-Myb VP16), 2=CC SS(11,12)AAVP16, 3=CC(d1–30)VP16, 4=dVVP16 (v-Myb VP16) and 5=dCVP16. N-Cla is the retroviral vector [13] which does not encode a functional activator protein. D: Binding of Myb VP16 fusion proteins to different reporter genes. Free DNA is shown in lanes 1, 5 and 10. Incubation of binding reactions with control cell extracts (N-Cla=retroviral vector [13] which does not encode a functional activator protein) is shown in lanes 2, 6 and 11. Binding reactions with the c-Myb VP16 (N-CVP16) cell extract are shown in lanes 3, 8 and 12, whereas the binding reactions with the v-Myb VP16 (N-dVVP16) cell extract are shown in lanes 4, 9 and 13. Myb-specific DNA-protein complexes and free DNA probes are marked on the left of the gel. Background bands observed after the incubation of probes with nuclear extracts are marked on the left of the gel (●).

dT) greatly affected binding of the v-Myb DBD domain (repeats R2 and R3) to the MRE, whereas the c-Myb DBD (repeats R1, R2 and R3) did not seem to be affected. An interruption of the T-stretch by placing a G at position T(3) (mim1A-Tint3) or placing a mismatch at position T(3) or positions T(3)+T(4) both in the wild-type mim1A sequence (mim1A-Tm3/mim1A-Tm3+4) or in the mim1A sequence that has the downstream T-stretch deleted (mim1A-dTm3/mim1A-dTm3+4) again affected binding of v-Myb to the MRE, but not of c-Myb (see Fig. 1D). To some extent, the mismatches can compensate for the absence of AT-rich DNA for the v-Myb DBD (mim1A-dTm3/mim1A-dTm3+4). Interestingly, the c-Myb DBD seemed not be affected to the same degree and the only difference between these two proteins is the presence of four amino acid substitutions and the absence of most of repeat R1 in the v-Myb DBD. Since we obtained very similar results for the c-Myb R2R3 protein as for the v-Myb R2R3 protein (results not shown), we can rule out that the four amino acid substitutions have an effect in these gel shift assays. This clearly points to a new and important function for repeat R1. In addition, neither c-Myb R1R2R3 nor v-Myb R2R3 seem to care on which strand the T-stretch is located. A quantitative analysis of the EMSAs shown in Fig. 1C is summarized in Fig. 1D. Relative affinities were determined by analyzing binding of various protein concentrations using the same amount of wild-type or mutant DNA. All the measurements were done within the linear range of the assay to avoid protein excess which would result in saturation. These results suggest that for the Myb proteins to form a stable Myb-MRE complex, a certain tertiary complex has to be formed which can be facilitated by the presence of a downstream AT-rich sequence. Since the c-Myb protein can bind to a bigger pool of target sites and shows less of a dependency on the downstream flanking sequence, we believe that repeat R1 can function as a 'wedge' in the formation of the Myb-MRE complex. This wedge would open up neighboring downstream sequences in the presence and absence of AT-rich sequences so that the same stable complex can be formed. This conclusion strongly supports the idea that R1 is not involved in specific DNA-binding, but rather in increasing the stability of the Myb-DNA complex [13–16].

3.2. *In vivo trans-activation by the Myb proteins requires an extension of the MRE by a downstream T-stretch*

Further insight about the contribution of downstream flanking sequences and the successful formation of the Myb-MRE complex can be gained by comparing the binding capabilities of different Myb proteins to the MRE to their ability to activate transcription in vivo. The results of the *trans*-activation assays in QT6 cells are summarized in Fig. 2B. All the Myb proteins tested were VP16 fusion proteins (see Fig. 2A) which allowed us to study the N-terminal contribution of the different proteins in DNA-binding and *trans*-activation without the complications and specific intramolecular actions within full-length Myb proteins. We tested three different Myb-responsive reporters in this type of assay: wild-type mim1A (three mim1A sites), mim1A with a deletion of the downstream T-stretch (three mim1A-dT sites) and mim1A with the T-stretch on the opposite strand (three mim1A-Tos sites) (see Fig. 2).

Three main conclusions can be drawn from these sets of assays. First, the v-Myb DBD (dVVP16) is more sensitive than the c-Myb DBD (CCVP16) to whether the T-stretch is present or not (compare mim1A with mim1A-dT), which supports our in vitro DNA-binding studies (see above and Fig. 1A,D). Once the T-stretch was deleted and replaced by a random sequence (mim1A-dT), transcriptional activity of dVVP16 was decreased the most (by 80%), whereas CCVP16 was less affected (by 60%). This is in agreement with what we have seen in the EMSAs. The v-Myb DBD shows a higher dependency for the downstream T-stretch compared to the c-Myb DBD for DNA-binding. In addition, dCVP16, which is the c-Myb DBD without repeat R1, is less affected than dVVP16 and the only difference is the presence of four amino acid substitutions in the v-Myb DBD. Our in vitro EMSAs did not reveal a difference for the c-Myb R2R3 compared to the v-Myb R2R3 DBD. Therefore, *trans*-activation of these two proteins is again not just dependent on DNA-binding, but rather on the formation of the proper protein-DNA complex. These QT6 assays with the different reporter constructs can therefore be used as an assay to study the contribution of the different amino acid substitutions and their contribution in *trans*-activation. We also tested the mutant CC SS(11,12)AA, in which both serines, the CKII phosphorylation sites at position 11 and 12 [41], were substituted with alanines [13]. Further, we tested the deletion mutant CC d1–30, which lacks the first 30 amino acids including serine 11 and 12 and an adjacent stretch of acidic amino acids [13]. Phosphorylation of serine 11 and 12 was reported to reduce DNA-binding activity [42], whereas experiments with a shorter form of c-Myb that lacks the entire C-terminus implied just the opposite, namely that phosphorylation of Ser 11 and 12 by CKII increases DNA-binding activity [43]. In addition, our lab showed that deletion of the first 30 amino acids (d1–30) upregulates both DNA-binding and transcriptional activation [13]. In general, it seems that CKII (Ser-11 and Ser-12, absent in v-Myb) potentially regulates c-Myb function through the N-terminal domain [13,41,43]. When tested with different reporter constructs, *trans*-activation by CC(d1–30)VP16 and CC SS(11,12)AAVP16 were the least affected (Fig. 2B).

Strikingly, all of the Myb proteins tested showed a strict requirement for the T-stretch to be located on the top strand, just downstream of the MRE. If placed on the bottom strand (mim1A-Tos), no transcriptional activation was measured which is a new and surprising observation.

Given the absence of other well-developed experimental systems for addressing DNA-bending and transcriptional activation in vivo, our approach seems to be sufficiently sensitive to directly analyze the influence of nearby DNA structures on a given DNA-protein complex. We are quite comfortable with the idea that the differences in *trans*-activation observed for the different reporter constructs is due to the influence of the neighboring flanking DNA sequence and not due to the actual MRE. We also tested the DNA-binding activity of the Myb proteins from transfected QT6 nuclear extracts to the different reporter sequences using EMSA. Fig. 2D shows binding of the two proteins v-Myb VP16 (N-dVVP16) and c-Myb VP16 (N-CCVP16) to the mim1A, the mim1A-dT and the mim1A-Tos sites. It is clear that both proteins still show specific binding to the MRE with different downstream flanking sequences even in the presence of high amounts of competitor (2 μg poly [dI-dC] in a 20 μl reaction volume). Therefore, these gel shift experiments with nuclear extracts support our in vitro experiments with purified proteins that DNA-binding

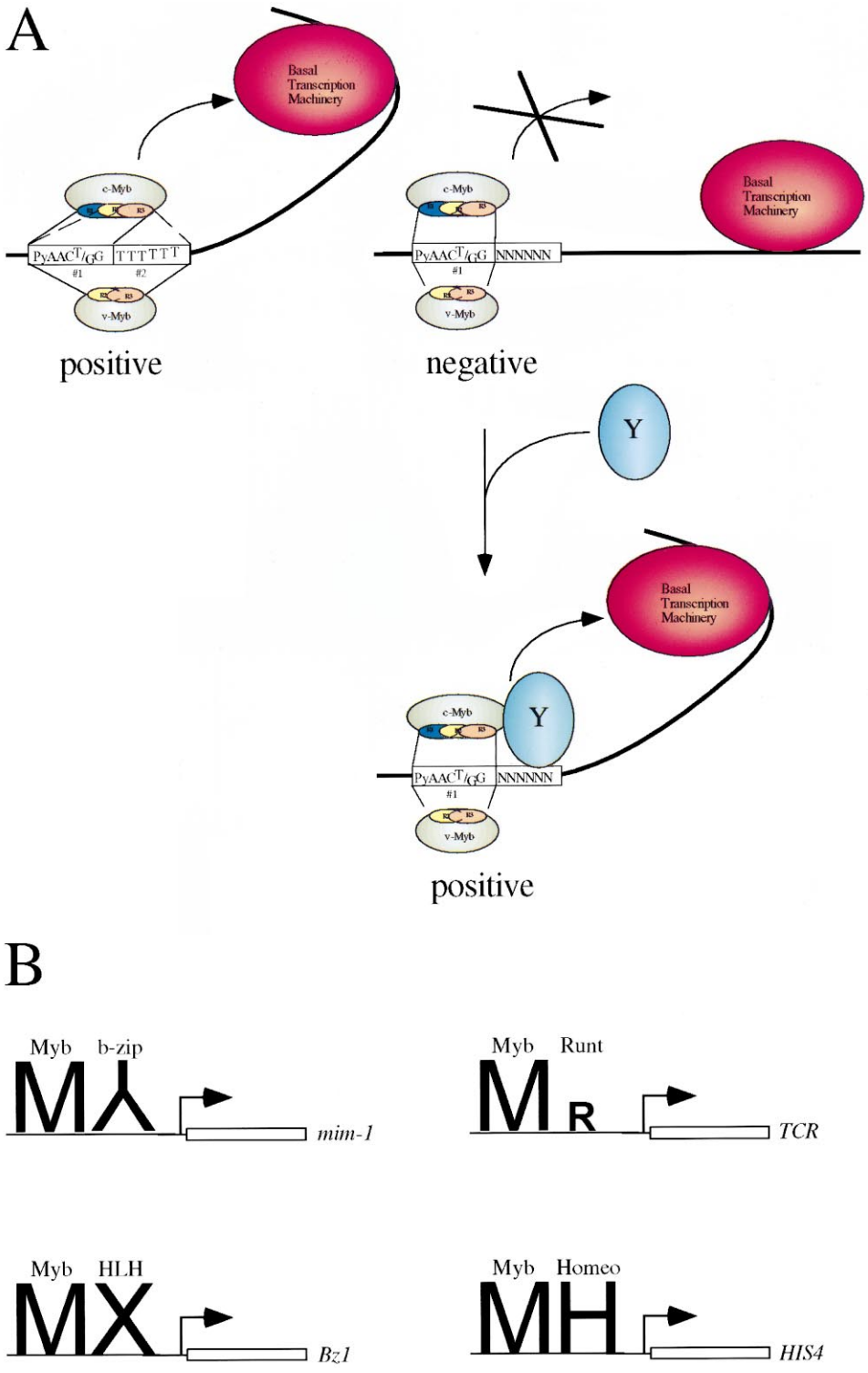


Fig. 3. A Model of Myb-specific transcriptional activation. Induction of transcription by the Myb proteins either through direct or indirect contacts with the basal transcription machinery requires a specific promoter structure. A specific Myb-MRE complex needs to be formed. This promoter structure can be obtained by two different mechanisms: (1) specific sequence-induced bend or (2) protein-induced bend. Both types of mechanisms would result in a positive induction of Myb-specific transcription. Therefore, a Myb-specific promoter should include, besides the MRE (motif #1), either a motif #2, which is a stretch of six Ts, or another proximal located UAS element for another factor that induces the same kind of structural changes into the DNA which is required for transcription to occur. The exact location of the 'Y' UAS element relative to the MRE PyAAC$^{T}/_{G}$G is unclear and is drawn immediately following the MRE as an example in this figure. B: Cooperation of Myb with other transcription factors. In each case, a Myb family protein and a DNA-binding protein of another family have been shown to co-regulate a specific target gene by occupation of adjacent binding sites within the promoter. c-Myb and C/EBP (b-ZIP) regulate the chicken *mim-1* gene [58], c-Myb and AML1 (Runt) regulate the murine T-cell receptor [59], C1 (Myb) and R (helix-loop-helix) regulate the maize *Bronze1* gene [60], Bas1 (Myb) and Bas2/Pho2 (Homeo) regulate the yeast *HIS4* gene [61].

occurs regardless of which strand the T-stretch is on. In addition, the observed difference in *trans*-activation (Fig. 2B) is not due to a difference in protein level within the QT6 cells as is shown in the immunoblot in Fig. 2C.

From the in vitro DNA-binding studies, we have learned that there is a dependency of the Myb DBD, in particular of the R2R3 v-Myb DBD, for the downstream AT-rich flanking sequence, though it seemed not to matter what the exact sequence was. Therefore, we concluded that DNA melting or in other words DNA 'breathing' is likely to be an important mechanism involved in the formation of a stable Myb-MRE complex. The in vivo *trans*-activation experiments extended this model by including another level of DNA structure, possibly DNA-bending. As mentioned above, a stretch of six Ts downstream of the MRE is required for the Myb proteins to form a stable DNA-protein complex that is capable of initiating transcription. This means that the DNA must adopt a certain conformation such as a local bend with a given orientation which influences binding of the Myb proteins to the neighboring MRE.

3.3. *The Myb proteins require a local DNA distortion for efficient trans-activation*

The present study establishes the requirement of an extension of the MRE (motif #1) PyAAC$^{T}/_{G}$G [7] by a downstream stretch of six Ts (motif #2), such as is found in the naturally occurring mim1A site within the *mim-1* promoter [18], for efficient binding and transcriptional activation by the Myb proteins. In terms of DNA-binding, the presence or absence of repeat R1 to some extend influences the intrinsic specificity of the Myb DBD for the MRE. N-terminal deletion of the c-Myb DBD to the v-Myb DBD equivalent showed a greater dependency for the T-stretch for efficient Myb-DNA complex formation, whereas the c-Myb DBD showed a relaxed dependency. The exact sequence of the downstream MRE element seems not to be important, as long as it is rich in A and T nucleotides, which suggests that melting of the downstream base pairs is a requirement for the transcriptional active protein-DNA complex formation. Experiments with the mismatched double-stranded DNA mim1A-m3 and mim1A-m3+4 support this notion. Random oligonucleotide selection and EMSAs revealed that repeat R1 within the c-Myb DBD does not contribute to specific DNA-binding, but seems to increase the affinity of the c-Myb protein for an MRE in the presence of an AT-rich sequence. In addition, R1 seems to be necessary and sufficient for the c-Myb DBD to be able to bind to a MRE without a downstream T-stretch. This implies that repeat R1 stabilizes the c-Myb-DNA complex by functioning as a 'wedge' which opens up neighboring DNA sequences other than AT-rich stretches. In summary, this explains why the c-Myb protein can bind to a greater number of different target sites.

Astonishingly, once specific Myb-DNA complexes were formed, an additional level of specificity was observed for transcriptional activation by the Myb proteins to occur. Transcriptional activation by the Myb proteins showed a strict requirement for the T-stretch (motif #2) downstream of the MRE (motif #1) TAACGG, which suggests that a specific Myb-DNA complex had to be formed that is highly dependent on the local DNA structure. As mentioned above, at the boundary of a stretch of six Ts, a bend can occur [38–40]. Bendable DNA can allow for a mixture of many different conformational states, the equilibrium of which can be displaced towards one specific form by external forces such as proteins interacting with them [44,45]. The specificity of DNA-protein contacts is not only the result of direct recognition of certain bases within the major groove of the DNA with specific amino acids of the protein, but also of an array of interactions between the protein and DNA surfaces not necessarily linked to a particular nucleotide sequence. This type of interaction has first been described for the *trp* repressor/operator complex [46] and has been termed indirect readout as opposed to the direct interaction of individual base pairs. Indirect readout involves protein contacts with the sugar-phosphate backbone of DNA through ionic bridges, hydrogen bonds and hydrophobic interactions. Both indirect and direct interactions contribute to the specificity and stability of a particular DNA-protein complex and require the protein and the DNA to accommodate each other in order to set up adequate contacts such as conformational distortions, which may be required in both partners to achieve an optimal fit [38,47]. The extent to which sequence-dependent bendability of DNA plays a key role in the formation of a functional Myb-DNA complex is clearly reflected in our *trans*-activation data. A Myb-responsive reporter in which the downstream T-stretch (motif #2) was placed on the opposite strand (E1b-luc(3×mim1A-Tos)) leaving an A-stretch downstream of the MRE (motif #1) essentially abolished the promoter activity. Since it is known that polyA/polyT tracts are particularly rigid and that bending may occur at the boundary of such a tract [38–40], it is likely that changing the sequence from a T-stretch to an A-stretch would change the orientation of a bend by 180° around the axis of the MRE double helix. In other words, it would completely distort the local promoter structure. This is what we have observed with the E1b-luc-(3×mim1A-Tos) reporter construct, because we did not obtain any transcriptional activity for the Myb proteins. Further, DNA transiently transfected into animal cells is partially organized into chromatin [48]. Strand specificity of the T-stretch could influence the position of a nearby nucleosome which in turn affects binding of adjacent proteins. The in vitro DNA-binding assays were all done with 'naked' DNA, whereas transcriptional activation assays in transiently transfected QT6 cells reflects a more in vivo-like situation. Recent data from our lab already demonstrated that the quality of the MRE determines the readout of Myb transcriptional activity. In particular Myb proteins, even when tethered to the strong VP16 activation domain, did not activate but rather repressed the *N-ras* promoter [49]. These results support a model in which the DNA recognition site itself can act as an allosteric effector of transcriptional regulation [50].

The promoter region of the E26-inducible *mim-1* gene contains three closely spaced Myb-binding sites [18] and is strongly activated by v-Myb and c-Myb in transient transfection assays. The mim1A site is the strongest of the three Myb-binding sites and is required for this activation in transient assays. Recently, additional Myb-regulated genes have been identified, such as the chicken lysozyme gene [51], the *tom-1* gene [52], the homeobox gene *GBX2* [53], the *c-kit* gene [54] and the Bcl-2 gene [55,56]. The Myb-binding sites defined within these promoter regions all contain an AAC core element which is required for the Myb proteins to bind efficiently to the MRE [17]. Nevertheless, none of the Myb-binding sites other than the mim1A site in the *mim-1* promoter contains a

downstream extension of the MRE by a stretch of six Ts. Our binding site selection data, gel shift data and QT6 *trans*-activation data clearly show that the mim1A site is the most ideal MRE known. Since none of the MREs described so far contains an additional motif #2 (T-stretch) downstream of the MRE, we believe that nature has other ways to assemble strong promoter elements. For example, other regulatory factors may bind to adjacent specific DNA elements or bind to Myb itself and function as co-activators that regulate the expression of target genes in conjunction with the Myb transcription factors. Such a regulatory factor would distort the local promoter structure to a similar extent like the T-stretch in the mim1A site so that a functional Myb-MRE complex can be formed. This would also provide an additional level of regulation on Myb-responsive genes. Interestingly, several of the Myb target genes described so far require another myeloid-specific transcription factor, NF-M, a homologue of the mammalian C/EBP protein [57], for full induction. This was shown for the *mim-1* and the *tom-1* gene [52,58]. Our favored model is outlined in Fig. 3A, which suggests that formation of a functional Myb-MRE complex requires a certain tertiary local promoter structure for direct or indirect interaction with the basic transcription machinery. This complex can either be obtained through the rigid structure of a downstream T-stretch or through a regulatory factor Y that induces a similar kind of bend. In the case of the *mim-1* promoter, two potential NF-M-binding sites are found besides the three MREs. One is located at position −60, just upstream of the mim1A site at the proper site relative to the MRE, and the second downstream, closer to the start site [58]. Both NF-M sites are required for full induction of the *mim-1* gene [58], but whether NF-M has factor 'Y' function within the context of the *mim-1* promoter is not clear yet and has to be investigated.

Besides c-Myb and NF-M (b-ZIP) that cooperate to regulate the chicken *mim-1* gene [58], there are additional examples of other Myb-related proteins that must bind directly adjacent to transcription factors with other classes of DBDs for specific target gene expression. In each case, a Myb family protein and a DNA-binding protein of another family have been shown to co-regulate a specific target gene by occupation of adjacent binding sites within the same promoter region. For example, c-Myb and AML1 (Runt) regulate expression of the murine T-cell receptor delta enhancer [59], C1 (Myb) and R (helix-loop-helix) regulate expression of the maize *Bronze1* gene [60], Bas1 (Myb) and Bas2/Pho2 (Homeo) regulate expression of the yeast *HIS4* gene [61]. These four given examples are summarized in Fig. 3B. Whether any of these other classes of DNA-binding proteins have DNA-bending ability is not clear, although recent evidence suggests that the Runt domain proteins can bend DNA [62]. A major challenge for future research is to elucidate the required conformations of Myb-MRE-Y complexes in greater detail. Such insights may provide a crucial key to the elusive goal of Myb-specific target gene expression.

Acknowledgements: We would like to thank Phil Verzola for help with figures and all the members of our laboratory for helpful discussions. This work was supported in part by the Swiss National Science Foundation (B.G.), a summer fellowship from the URO office of Stanford University (S.T.C.) and the USPHS Grant R01 CA43592.

References

[1] Introna, M., Golay, J., Frampton, J., Nakano, T., Ness, S.A. and Graf, T. (1994) Semin. Cancer Biol. 5, 113–124.
[2] Westin, E.H., Gallo, R.C., Arya, S.K., Eva, A., Souza, L.M., Baluda, M.A., Aaronson, S.A. and Wong-Staal, F. (1982) Proc. Natl. Acad. Sci. USA 79, 2194–2198.
[3] Gonda, T.J. and Metcalf, D. (1984) Nature 310, 249–251.
[4] Mucenski, M.L. et al. (1991) Cell 65, 677–690.
[5] Graesser, F.A., Graf, T. and Lipsick, J.S. (1991) Mol. Cell. Biol. 11, 3987–3996.
[6] Sakura, H., Kanei-Ishii, C., Nagase, T., Nakagoshi, H., Gonda, T.J. and Ishii, S. (1989) Proc. Natl. Acad. Sci. USA 86, 5758–5762.
[7] Biedenkapp, H., Borgmeyer, U., Sippel, A.E. and Klempnauer, K.-H. (1988) Nature 335, 835–837.
[8] Rosson, D. and Reddy, E.P. (1986) Nature 319, 604–606.
[9] Gerondakis, S. and Bishop, J.M. (1986) Mol. Cell. Biol. 6, 3677–3684.
[10] Howe, K.M., Reakes, C.F.L. and Watson, R.J. (1990) EMBO J. 9, 161–170.
[11] Garcia, A., Lamontagne, K., Reavis, D., Stober-Graesser, U. and Lipsick, J.S. (1991) Oncogene 6, 265–274.
[12] Gabrielsen, O.S., Sentenac, A. and Fromageot, P. (1991) Science 253, 1140–1143.
[13] Dini, P.W. and Lipsick, J.S. (1993) Mol. Cell. Biol. 13, 7334–7348.
[14] Ording, E., Kvavik, W., Bostad, A. and Gabrielsen, O.S. (1994) Eur. J. Biochem. 222, 113–120.
[15] Ebneth, A., Schweers, O., Thole, H., Fagin, U., Urbanke, C., Maass, G. and Wolfes, H. (1994) Biochemistry 33, 14586–14593.
[16] Ogata, K. et al. (1995) Nat. Struct. Biol. 2, 309–320.
[17] Tanikawa, J., Yasukawa, T., Enari, M., Ogata, K., Nishimura, Y., Ishii, S. and Sarai, A. (1993) Proc. Natl. Acad. Sci. USA 90, 9320–9324.
[18] Ness, S.A., Marknell, A. and Graf, T. (1989) Cell 59, 1115–1125.
[19] Jamin, N., Gabrielsen, O.S., Gilles, N., Lirsac, P.-N. and Toma, F. (1993) Eur. J. Biochem. 216, 147–154.
[20] Ogata, K. et al. (1994) Cell 79, 639–648.
[21] Carr, M.D., Wollborn, U., Mcintosh, P.B., Frenkiel, T.A., Mccormick, J.E., Bauer, C.J., Klempnauer, K.-H. and Feeney, J. (1996) Eur. J. Biochem. 235, 721–735.
[22] McIntosh, P., Frenkiel, T., Wollborn, U., McCormick, J., Klempnauer, K., Feeney, J. and Carr, M. (1998) Biochemistry 27, 9619–9629.
[23] Anton, I.A. and Frampton, J. (1988) Nature 336, 719.
[24] Ogata, K., Hojo, H., Aimoto, S., Nakai, T., Nakamura, H., Sarai, A., Ishii, S. and Nishimura, Y. (1992) Proc. Natl. Acad. Sci. USA 89, 6428–6432.
[25] Saikumar, P., Murali, R. and Reddy, E.P. (1990) Proc. Natl. Acad. Sci. USA 87, 8452–8456.
[26] Kanei-Ishii, C., Sarai, A., Sawazaki, T., Nakagoshi, H., He, D.-N., Ogata, K., Nishimura, Y. and Ishii, S. (1990) J. Biol. Chem. 265, 19990–19995.
[27] Howe, K.M. and Watson, R.J. (1991) Nucleic Acids Res. 19, 3913–3920.
[28] Weston, K. (1992) Nucleic Acids Res. 20, 3042–3049.
[29] Blackwell, T.K. and Weintraub, H. (1990) Science 250, 1104–1110.
[30] Ganter, B. and Lipsick, J.S. (1999) Adv. Cancer Res. 76, 21–60.
[31] Bilaud, T., Koering, C.E., Binet-Brasselet, E., Ancelin, K., Pollice, A., Gasser, S.M. and Gilson, E. (1996) Nucleic Acids Res. 24, 1294–1303.
[32] Aasland, R., Stewart, A.F. and Gibson, T. (1996) TIBS 21, 87–88.
[33] Reeder, R.H. and Lang, W.H. (1997) TIBS 22, 473–477.
[34] Ganter, B., Fu, S.-F. and Lipsick, J.S. (1998) EMBO J. 17, 255–268.
[35] Hagermann, P.J. (1990) Annu. Rev. Biochem. 59, 755–781.
[36] Trifonov, E.N. (1985) CRC Crit. Rev. Biochem. 19, 89–106.
[37] Dlakic, M. and Harrington, R.E. (1996) Proc. Natl. Acad. Sci. USA 93, 3847–3852.
[38] Suzuki, M., Loakes, D. and Yagi, N. (1996) Adv. Biophys. 32.
[39] Suzuki, M., Yagi, N. and Finch, J.T. (1996) FEBS Lett. 379, 148–152.

[40] Wu, H.-M. and Crothers, D.M. (1984) Nature 308, 509–513.
[41] Luscher, B., Christenson, E., Litchfield, D.W., Krebs, E.G. and Eisenman, R.N. (1990) Nature (Lond.) 344, 517–521.
[42] Oelgeschlaeger, M., Krieg, J., Luescher-Firzlaff, J.M. and Luescher, B. (1995) Mol. Cell. Biol. 15, 5966–5974.
[43] Ramsay, R.G., Morrice, N., Van Eeden, P., Kanagasundaram, V., Nomura, T., De Blaquiere, J., Ishii, S. and Wettenhall, R. (1995) Oncogene 11, 2113–2120.
[44] Travers, A.A. (1989) Annu. Rev. Biochem. 58.
[45] Flashner, Y. and Gralla, J.D. (1988) Cell 54, 713–721.
[46] Otwinowski, Z., Schevitz, R.W., Zhang, R.G., Lawson, C.L., Joachimiak, A., Marmorstein, R.Q., Luisi, B.F. and Sigler, P.B. (1988) Nature 355.
[47] Suzuki, M. and Yagi, N. (1996) J. Mol. Biol. 255, 677–687.
[48] Reeves, R., Gorman, C. and Howard, B. (1985) Nucleic Acids Res. 10, 3599–3615.
[49] Ganter, B. and Lipsick, J.S. (1997) Oncogene 15, 193–202.
[50] Lefstin, J. and Yamamoto, K. (1998) Nature 392, 885–888.
[51] Burk, O. and Klempnauer, K.-H. (1991) EMBO J. 10, 3713–3720.
[52] Burk, O., Worpenberg, S., Haenig, B. and Klempnauer, K.-H. (1997) EMBO J. 16, 1371–1380.
[53] Kowenz-Leutz, E., Herr, P., Niss, K. and Leutz, A. (1997) Cell 91, 185–195.
[54] Yamamoto, K., Tojo, A., Aoki, N. and Shibuya, M. (1993) Jpn. J. Cancer Res. 84, 1136–1144.
[55] Taylor, D., Badiani, P. and Weston, K. (1996) Genes Dev. 10, 2732–2744.
[56] Frampton, J., Ramqvist, T. and Graf, T. (1996) Genes Dev. 10, 2720–2731.
[57] Sterneck, E., Mueller, C., Katz, S. and Leutz, A. (1992) EMBO J. 11, 115–126.
[58] Ness, S.A., Kowenz-Leutz, E., Casini, T., Graf, T. and Leutz, A. (1993) Genes Dev. 7, 749–759.
[59] Hernandez-Munain, C. and Krangel, M.S. (1994) Mol. Cell. Biol. 14, 473–483.
[60] Roth, B., Goff, S., Klein, T. and Fromm, M. (1991) Plant Cell 3, 317–325.
[61] Tice-Baldwin, K., Fink, G. and Arndt, K. (1989) Science 246, 931–935.
[62] Golling, G., Li, L., Pepling, M., Stebbins, M. and Gergen, J. (1996) Mol. Cell. Biol. 16, 932–942.
[63] Sanger, F., Nickeln, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[64] Treisman, R. (1986) Cell 46, 567–574.
[65] Chen, C.H. and Okayama, H. (1987) Mol. Cell. Biol. 7, 2745–2752.
[66] Ibanez, C.E. and Lipsick, J.S. (1990) Mol. Cell. Biol. 10, 2285–2293.
[67] Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1989) Current Protocols in Molecular Biology, pp. 9.7.12–9.7.14, John Wiley and Sons, New York.
[68] de Wet, J., Wood, K.V., DeLucas, M., Helsinki, D.R. and Subramani, S. (1987) Mol. Cell. Biol. 7, 725–737.
[69] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[70] Andrews, N.C. and Faller, D.V. (1991) Nucleic Acids Res. 19, 2499–2500.
[71] Sleeman, J.P. (1993) Oncogene 8, 1931–1941.