Characterization of Randomly-Obtained Matrix Attachment Regions (MARs) from Higher Plants[†,‡]

Susan M. Michalowski,[§,||] George C. Allen,[⊥] Gerald E. Hall, Jr.,[§,#] William F. Thompson,[§,⊥] and Steven Spiker*[,§]

Department of Genetics and Department of Botany, North Carolina State University, Raleigh, North Carolina 27695

Received May 18, 1999; Revised Manuscript Received July 20, 1999

ABSTRACT: Matrix attachment regions (MARs) can be operationally defined as DNA fragments that bind to the nuclear matrix. We have created a library of randomly obtained MARs from tobacco (*Nicotiana tobacum*) by cloning DNA fragments that co-isolate with nuclear matrixes prepared by a method involving lithium diiodosalicylate. The interactions of several of the cloned MARs with nuclear matrixes were tested by an in vitro binding assay in which genomic DNA was used as competitor. Based on this assay, the MARs were classified as strong, medium, and weak binders. Examples of each of the binding classes were further studied by in vitro binding using self- and cross-competition. Estimates of dissociation constants for several MARs ranged from 6 to 11 nM and correlated inversely with binding strength. The number of binding sites per matrix for several MARs ranged from 4×10^5 to 9×10^5 and correlated directly with binding strength. We conclude that binding strength, as we have measured it, is a function of both numbers of binding sites and affinity for the sites. The tobacco MARs were sequenced and analyzed for overall AT content, for distribution of AT-rich regions, and for the abundance of several MAR-related motifs. Previously identified MAR motifs correlate to various degrees with binding strength. Notably, the *Drosophila* topoisomerase II motif does not correlate with binding strength of the tobacco MARs. A newly identified motif, the "90%AT Box," correlates better with binding strength than any of the previously identified motifs we investigated.

The idea that chromatin is organized into loop domains is as old as the original observation of loops in lampbrush chromosomes by Flemming in 1882 (*1*, *2*). A reincarnation of that idea came from the electron micrographs of Paulson and Laemmli (*3*), which showed loops of DNA attached at their bases to the proteinaceous scaffold of metaphase chromosomes. A natural extension of the observations of Paulson and Laemmli is that the DNA−chromosome scaffold interaction is not random. Indeed, abundant evidence has been presented for interactions between specific DNA fragments and the metaphase chromosome scaffold and the interphase nuclear matrix [reviewed in Bode et al. (*4*, *5*)]. The specific DNA sequences have been termed MARs (matrix associated regions) (*6*) or SARs (scaffold attachment regions) (*7*, *8*). Both terms apparently describe the same biological entity (*9*). We use the term "MAR" because of its prevalence in the literature (*10*).

MARs have a typical size of around 1 kb and have been postulated to form the anchorage points of torsionally constrained loop domains ranging in size from a few kb to more than 100 kb [reviewed by Gasser et al. (*11*) and Bode et al. (*4*, *5*)]. These putative structural elements of chromatin have attracted considerable attention because of evidence for their involvement in the regulation of transcription, including stabilization and increased levels of transgene expression (*12*−*27*) [reviewed by Bode et al. (*4*)]. The comapping of MARs and origins of replication and the functioning of MARs as ARS elements suggest a role for MARs in DNA replication (*28*−*34*).

MARs can be operationally defined by two criteria: first, as endogenous DNA fragments that copurify with the nuclear matrix (i.e., remain bound to the nuclear matrix after chromatin proteins and DNA in the presumptive loops have been removed); or, second, as exogenously added DNA fragments that bind to the purified nuclear matrix in the presence of competitor DNA (*6*, *35*, *36*).

The most common method of identifying MARs is to assay a large region of genomic DNA for fragments that bind to the nuclear matrix DNA. In many cases, the large region of genomic DNA assayed for MARs was chosen because it harbors a gene or genes of particular interest. Examples of this approach include a 10.5 kb region containing *Drosophila* heat shock genes and a 5 kb region containing *Drosophila* histone gene repeats (*37*), a 19 kb region containing the mouse kappa immunoglobulin gene (*6*), a 36 kb region

[†] This work was supported by grants to S.S. from the National Science Foundation and the U.S. Department of Agriculture Research Grants Office and by a grant to W.F.T. and G.C.A. from the U.S. Department of Agriculture Research Grants Office. S.M.M. was supported by a fellowship from the U.S. Department of Education Patricia Roberts Harris Fellowship Program.

[‡] The nucleotide sequences reported in this work are deposited in GenBank under Accession Numbers AF065877, AF065878, AF065879, AF065880, AF065881, AF065882, AF065883, AF065884, AF065885, AF065886, AF065887, AF065888, and AF065889.

* To whom correspondence should be addressed. Telephone: 919-515-5760. Fax: 919-515-3350.

[§] Department of Genetics.

[||] Present address: Department of Microbiology & Immunology, University of North Carolina, Chapel Hill, NC 27599.

[⊥] Department of Botany.

[#] Present address: Cyanamid Agricultural Research Center, Quakerbridge Rd., PO Box 400, Princeton, NJ 08543.

10.1021/bi991142c CCC: $18.00
Published on Web 09/03/1999

containing the human interferon-β gene (*9*), a 35 kb region containing the Chinese hamster dihydrofolate reductase gene (*38*), a 90 kb region containing a human globin gene complex (*39*), a 19 kb region containing the chicken lysozyme gene (*40*), a 23 kb region containing a root-specific gene from tobacco (*41*), a 3.6 kb region containing the β-phaseolin gene of bean (*42*), a 7 kb region containing a tomato heat shock cognate gene (*43*), a 17 kb region containing a seed-specific lectin gene from soybean (*26*), and a 16 kb region containing the *Arabidopsis* plastocyanin gene (*44*).

In the examples mentioned above, the regions chosen for study, although relatively large, are smaller than 100 kb. Examples also exist in which the region studied is larger than 100 kb. In many of these cases, the region was chosen for study not because it harbored a specific gene, but rather because it contained a number of unrelated genes or represented a large region of a chromosome. Such examples include an 800 kb region of the *Drosophila* X chromosome (*45*, *46*), a 320 kb region of *Drosophila* chromosome 3 (*8*), a 100 kb region of Chinese hamster chromosome 1q (*47*), a 280 kb region of maize chromosome 1 containing the alcohol dehydrogenase 1 gene (*48*, *49*), and a 270 kb region representing an amplicon in Chinese hamster ovary cells (*50*).

The methods mentioned above are valuable for studying MARs that are associated with particular genes and for studying the loop domain organization of defined regions of the genome. However, the selection of MARs by these methods may be biased by selection based on particular genes and thus be inappropriate for study of the general characteristics of MARs. In a few studies, more general approaches to obtaining MARs have been used. Nikolaev and co-workers (*51*) selected matrix binding fragments from a human chromosome 19 library and used these to construct a chromosome-specific library of MARs. Boulikas and Kong (*52*, *53*) used a different procedure and went a step further toward creating a random MAR library. They used the high-salt method (*6*) and micrococcal nuclease in the preparation of nuclear matrixes from human erythroleukemia cells. The DNA fragments associated with the purified nuclear matrix (MARs by operational definition) were used to generate a human MAR library. This seems to be a logical approach for the random acquisition of MARs. As of yet, the sequence analysis of only a limited number of the MARs obtained by this method has been presented (*52*, *53*).

In the work presented here, we have generated a library of MARs from tobacco (*Nicotiana tabacum)*. The approach we have used is similar in concept to that of Boulikas and Kong (*52*, *53*) except that we have used the lithium diiodosalicylate (LIS) method (*37*) and a battery of restriction enzymes for making nuclear matrixes. The use of low concentrations of LIS to remove chromosomal proteins is viewed as a milder extraction procedure that circumvents sliding of DNA−matrix attachment points and precipitation of transcription complexes that may occur in the presence of 2 M NaCl (*37*). We have characterized several of the cloned DNA fragments for in vitro binding to the nuclear matrix. From these cloned fragments, we have selected representative strong, medium, and weak binders for sequencing and further study.

From the earliest studies of MARs, investigators have attempted to learn what properties are responsible for MARs binding to the nuclear matrix. Several DNA sequence motifs have been identified as being characteristic of MARs. The most widely considered motifs include A-boxes and T-boxes (*54*), the consensus sequence for *Drosophila* topoisomerase II (*6*, *7*), the ARS consensus sequence (*55*), and the "MAR" sequence (*6*). In addition to these, several other MAR-related motifs have been proposed as reviewed by Boulikas (*56*−*59*). Despite the documented occurrence of these motifs in several MARs, it is clear that the simple presence of one or more of these motifs is not sufficient to establish a DNA fragment as a MAR. Thus, more subtle structural elements are inferred to define the essentials of MARs (*4*, *58*). Boulikas and Kong have proposed that a high density of inverted repeats is involved in matrix binding (*53*). Another of the suspected defining structural elements of MARs is the narrow minor grooves that occur in AT-rich DNA (*4*, *58*). Nearly all characterized MARs are AT-rich (greater than 65% AT) (*58*), but simply being AT-rich does not make a DNA fragment a MAR (*44*, *47*, *60*, *61*). DNA sequences that bestow a propensity for bending have been associated with MARs (*62*). A role for unwinding potential as a defining element for MARs has been suggested by Bode and co-workers (*63*). These workers have used a 9 bp sequence, the base-unpairing region (BUR), as a preliminary indicator of unwinding potential, but in collaboration with Benham (*64*) they have devised a much more sophisticated method of predicting stress-induced unwinding.

One of the goals of the work by Benham et al. (*64*) is to identify MARs on the basis of sequence alone. This goal has been shared by many others, and a variety of predictive procedures have been devised using previously suggested MAR motifs in mathematical models and computer-assisted searches (*53*, *56*−*58*, *65*−*68*). All these procedures are designed to predict if a DNA sequence will or will not bind to the nuclear matrix. Binding of DNA to the matrix is, however, a quantitative phenomenon. A continuum of binding strengths exists from strong to moderate to weak to barely detectable (*4*, *13*, *60*, *69*, *70*). With this in mind, we have assessed correlations between the abundance of a variety of MAR motifs and binding strength in a number of the randomly obtained MAR sequences we have obtained. We find that binding strength correlates best with a motif we call the "90%AT Box," a sequence of 20 nucleotides of which 18 or more are A or T. The *Drosophila* topoisomerase II motif does not correlate with binding strength of the tobacco MARs.

EXPERIMENTAL PROCEDURES

Isolation of Nuclei and Preparation of Nuclear Matrixes. Nuclei were isolated from a *Nicotiana tabacum* cell line (NT-1) as previously described (*35*, *41*). Nuclear matrixes were prepared from the isolated nuclei by a modification (*35*, *69*) of the method of Mirkovitch et al. (*37*), which employs lithium diiodosalicylate (LIS). Separate nuclear matrix preparations were made with each of the restriction enzymes *Hin*dIII, *Eco*RI, *Taq*I, and *Rsa*I.

Isolation and Cloning of MAR DNA. Restriction enzyme treatment solubilizes DNA in the presumptive loop domains of chromatin. The DNA fragments remaining associated with the purified nuclear matrixes are MARs by operational definition. These DNA fragments were purified by overnight treatment with proteinase K followed by phenol/chloroform

extraction and ethanol precipitation (*69*). The DNA fragments obtained from matrixes made with *Hin*dIII, *Eco*RI, *Taq*I, or *Rsa*I were cloned into the *Hin*dIII, *Eco*RI, *Cla*I, or *Eco*RV site (respectively) of pBluescript II SK+ (Stratagene).

Binding Assays. The binding of the cloned fragments to isolated nuclear matrixes was performed by the *exogenous binding assay* as previously described (*13*, *35*, *41*, *69*). To generate radioactively labeled probes for the binding assays, plasmids were cut with either *Eco*RI or *Hin*dIII to release the insert. Both vector and insert were ^{32}P end-labeled with Klenow. The labeled vector was included in the binding assays and provides a convenient internal control of a non-MAR DNA fragment. A typical binding assay contained approximately 300 000 nuclear matrixes (determined by counting in a hemocytometer) and 5 ng (2 fmol, 50 000 cpm) of probe. Two types of competitor DNA were used. In the standard assay, tobacco genomic DNA was used as competitor. In this assay, DNA released by restriction enzymes during the isolation of nuclear matrixes was not separated from the matrixes. Thus, a constant ratio of genomic competitor DNA to nuclear matrix was assured from assay to assay. The excess of genomic competitor DNA to probe is approximately 1000-fold. The assays using genomic DNA as competitor were used to evaluate the relative binding strength of the MARs we have studied. In other assays, we used gel-isolated MAR inserts for self-competition or cross-competition. For all assays, nuclear matrixes, probe, and competitor were mixed together and incubated at 37 °C for 3 h with resuspension every 20 min. Following incubation, a pellet (bound DNA fragments) and a supernatant (unbound DNA fragments) fraction were formed by centrifugation. DNA from the pellet and supernatant fractions was subjected to electrophoresis and autoradiography. In assays involving quantification, the amounts of DNA in bands in the supernatant and pellet fractions were determined by using the Radioanalytic Imaging System (AMBIS Co.). In all experiments presented here, we have used the "equal fractions" method of presenting data (*36*), i.e., an equal fraction (usually 20%) of the total volume of each of the supernatant and resuspended pellet fractions is subjected to electrophoresis. This approach has the advantage that the amount of radioactivity in a band from the pellet or supernatant is directly proportional to the amount of probe DNA in each fraction. Thus, the ratio of matrix-bound probe to unbound probe can be directly compared from lane to lane.

Subcloning, Sequencing, and Sequence Analysis. Some of the clones we obtained were large and contained DNA not involved in binding to the nuclear matrix. To effectively analyze the sequence of MAR DNA, we wanted to exclude the nonbinding DNA in our clones. Thus, we treated each of the large clones with an array of restriction enzymes and tested the resulting fragments for matrix binding by the standard exogenous assay using genomic DNA as competitor. Core binding fragments were subcloned into pBC KS+ (Stratagene) for further analysis. The resulting subclones are identified with a hyphen and a number following the designation of the original clone from which the subclone was derived. Thus, pS211-1 is an *Eco*RI−*Hin*dIII fragment of pS211; pS220-1 an *Xho*I−*Hin*dIII fragment of pS220; pS206-1 a *Bam*HI−*Hin*dIII fragment of pS206; pS205-2 an *Eco*RI−*Hin*dIII fragment of pS205; pS217-1 a *Bam*HI fragment of pS217; pS202-1 and pS202-2 *Eco*RI−*Hin*dIII fragments of pS202. The core binding fragment we refer to here as pS116-1 is a subclone of pS116 listed in GenBank as pS116-1.1B (a *Mun*I−*Pst*I fragment of pS116 ligated into the *Eco*RI−*Pst*I sites of pBluescript II SK+).

Clones were sequenced in stages by primer walking at the Iowa State University Nucleic Acids Facility. Sequences were analyzed for the occurrence and distribution of motifs, base composition, and base composition within a sliding window using software from the Genetics Computer Group, Wisconsin Package Version 9.1, Madison, WI, September 1997. Sequences for the DNA molecules studied here can be found in GenBank according to the following accession numbers: pS1, AF056877; pS115, AF056878; pS116-1, AF056879; pS202-1, AF05680; pS202-2, AF05681; pS205-2, AF05682; pS206-1, AF05683; pS211-1, AF05684; pS217-1, AF05685; pS218, AF05686; pS220-1, AF05687; pS4, AF05688; pS8, AF05689.

RESULTS

Clones. From the library of DNA sequences that copurified with the nuclear matrix, 34 clones were tested for binding to the tobacco nuclear matrix by our standard exogenous assay. In this assay, DNA of the presumptive loop domains, which is released from the nuclear matrix by nucleases, is not separated from nuclear matrixes until after binding of the radiolabeled probes has been carried out. Thus, the genomic DNA serves as competitor. We chose this method of standardization because it ensures a constant ratio of competitor DNA to nuclear matrix. We have assessed the binding strength of the 34 clones based on the percentage of the radiolabeled probe that partitions with the nuclear matrix in our standard binding assay. Clones in which 70−100% of the probe DNA partitions with the nuclear matrix are classified as "strong" binders. Clones in the 40−70% range are "medium" binders, and clones in the just-detectable to 40% range are "weak" binders. Of the 34 clones, we chose 3 strong binders, 6 medium binders, and 2 weak binders for further study. We also included in our study two clones for which no binding was detected in the standard assay. Binding of these clones to the nuclear matrix can be detected on overloaded gels or overexposed autoradiographs, but less than 1% partitions with the nuclear matrix in our standard binding assay. In addition to the 13 randomly isolated tobacco MARs, we included in our studies a previously characterized strong tobacco MAR that is part of a genomic clone containing the root-specific gene RB7 (*13*, *41*, *71*).

The tested clones ranged in size from 380 bp to 7 kb. It has long been recognized that fragments larger than about 4.5 kb do not bind well to the nuclear matrix in exogenous binding assays, presumably because of steric hindrance (*8*). Because of this, we used additional restriction enzymes to further cleave large inserts. All resulting fragments were assayed for binding to the nuclear matrix. This allowed us to identify fragments that do not contribute to the binding and to concentrate our studies on binding fragments that ranged in size from 306 to 1499 bp. All subcloned fragments are identified by a hyphen and a number following the original designation of the clone.

Examples of standard endogenous binding assays (genomic DNA as competitor) are shown in Figure 1. The MAR designated pS116-1 is a strong nuclear matrix binder. More

Table 1: Properties of 14 Tobacco MARs[a]

	binding	length	% AT	90% AT	A box	T box	ARS	Topo II	BUR	MRS	ORI	TG-rich	curved	kinked
pS116-1	90	998	76.6	20	11	14	1	2	17	3	34	0	8	0
RB7-6	80	1103	73.2	18	13	6	3	2	3	1	46	0	5	2
pS211-1	70	685	71.7	10	7	11	2	4	0	2	43	0	3	2
pS220-1	70	1499	73.1	23	21	10	5	2	3	3	73	1	4	2
pS206-1	60	306	77.1	6	5	2	1	1	5	1	18	1	0	0
pS205-2	50	704	71.4	12	9	5	4	2	2	0	42	2	4	2
pS217-1	40	899	65.2	3	10	5	2	1	1	1	19	0	3	0
pS202-1	40	635	68.3	4	7	4	0	1	0	0	31	0	1	0
pS202-2	40	1087	64.2	4	7	4	2	2	0	1	20	0	1	6
pS115	40	866	65.7	6	4	5	2	0	5	0	27	2	3	0
pS4	20	587	72.2	4	7	4	2	2	0	0	20	0	2	1
pS1	20	437	71.8	9	9	3	0	1	5	0	38	0	1	3
pS8	0	383	61.6	2	3	1	0	1	0	0	12	0	1	2
pS218	0	999	62.0	0	5	5	2	4	0	0	22	2	0	3

[a] Binding = % of MAR partitioning with the nuclear matrix in a standard exogenous assay. Length = length of MAR in bp. Last 11 columns of the table are the numbers of occurrences of each of the motifs in each MAR (both strands). If motifs of the same type overlap, only one instance is counted. Sequences of each of the motifs are given below followed by the number of mismatches allowed in the searches. 90%AT = W_{20} (2 mismatches). A box = AATAAAYAA (2 mismatches). T box = TTWTWTTWTT (1 mismatch). ARS = WTTTATRTTTW (1 mismatch). Topo II = GTNWAYATTNATNNR (2 mismatches but no mismatches in WAYATT). BUR = AATATATTT (1 mismatch). MRS = TAWAWWWN-NAWWRTAANNWWG (full MRS with 2 mismatches but must end in G) or TAWAWWW (first half MRS with no mismatches) + AWWRTAANNWWG (second half MRS with 1 mismatch but must end in G) (first half and second half counted with up to 4 intervening nucleotides or up to 3 overlapping nucleotides). ORI = ATTA or ATTTA or ATTTTA (no mismatches). TG-rich = TGTTTTG or TGTTTTTTG or TTTTGGGG (no mismatches). Curved = $AAAAN_7AAAAN_7AAAA$ or TTTAAA (no mismatches). Kinked = TANNNTGNNNCA or TANNNCANNNTG or TGNNNTANNNCA or CANNNTANNNTG (no mismatches).

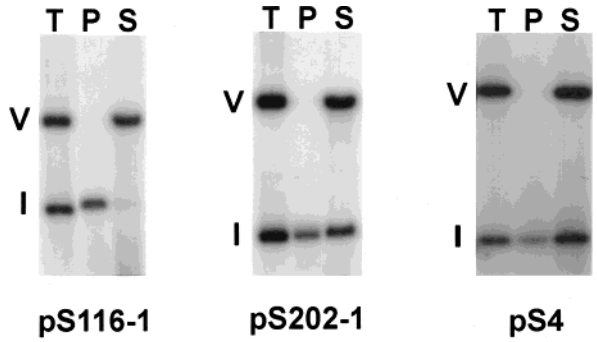

FIGURE 1: Standard exogenous matrix binding assay for strong, medium, and weak MARs. Lanes marked "T" represent aliquots of the total radiolabeled probe used in the assay. Lanes marked "P" represent aliquots of the probe found in the pellet, i.e., bound to the nuclear matrix. Lanes marked "S" represent aliquots of the probe found in the supernatant, i.e., not bound to the nuclear matrix. Bands marked "V" are the DNA fragments from the vector. Bands marked "I" are the inserts (MARs). pS116-1 is a strong nuclear matrix binder. More than 90% of the MAR partitions with the pellet. pS202-1 is a medium binder. Approximately 40% partitions with the pellet. pS4 is a weak binder. Approximately 20% partitions with the pellet. Vector fragments do not bind to the nuclear matrix and are found in the supernatants.

than 90% of the MAR partitions with the pellet (nuclear matrix-bound) fraction. pS202-1 is a weak-to-medium binder. Approximately 40% partitions with the pellet. pS4 is a weak binder. Approximately 20% partitions with the pellet. Vector fragments serve as negative controls as they do not bind to the nuclear matrix under these conditions and are found exclusively in the supernatants. The binding strengths of the randomly obtained MARs that we have investigated are presented in Table 1 and Figure 6.

Basis of Binding Strength. The method we have used to assess binding strength is a convenient way of comparing the binding of several MARs to the nuclear matrix. Because of the way the assay is carried out, the observed differences in binding strength among the MARs we have isolated could be due to the following: (a) differences in the numbers of binding sites in the nuclear matrix for each MAR; (b) differences in the affinity for individual binding sites; or (c) a combination of numbers of sites and affinities. To investigate these possibilities, we have carried out binding studies using isolated MAR DNA fragments as competitors in the absence of genomic DNA. An autoradiograph of such a binding study is shown for the RB7-6 MAR in Figure 2. The top panel shows (in lane "T") the total labeled probe used in the study. "V" is the vector and "I" the MAR insert. Also in the top panel are the fragments that partition with the pellet (i.e., bind to the nuclear matrix) in the assay. In the bottom panel are fragments that partition with the supernatant in the binding assay. Lane "G" shows the results of a standard assay using genomic DNA as competitor (as in Figure 1). Under these conditions, the vector does not bind to the nuclear matrix, and essentially all of the vector is found in the supernatant. In the lanes marked "self-competitor," increasing amounts of nonradiolabeled RB7-6 fragments (0.01−10 pmol/assay) are used as competitor. Increasing levels of competitor shift the MAR probe from the pellet to the supernatant.

Similar self-competition studies were carried out with pS116-1, pS211-1, and pS202-1. In our study we also included the yeast ARS-1 (*55*), which binds weakly to tobacco nuclear matrixes (*12*, *13*, *41*). The binding data for these five MARs are presented as Scatchard plots (*72*) in Figure 3. Plotting the data in this manner allows us to calculate a dissociation constant (K_d) and the number of apparent binding sites per nuclear matrix. As can be seen from the values presented in Table 2, there is some degree of correspondence between the calculated number of binding sites per nuclear matrix and the binding strength of the MARs. There is also some degree of correspondence between binding strength and the inverse of the dissociation constants of the MARs. A tighter correspondence (correlation coefficient of 0.90) is found between the binding strength and

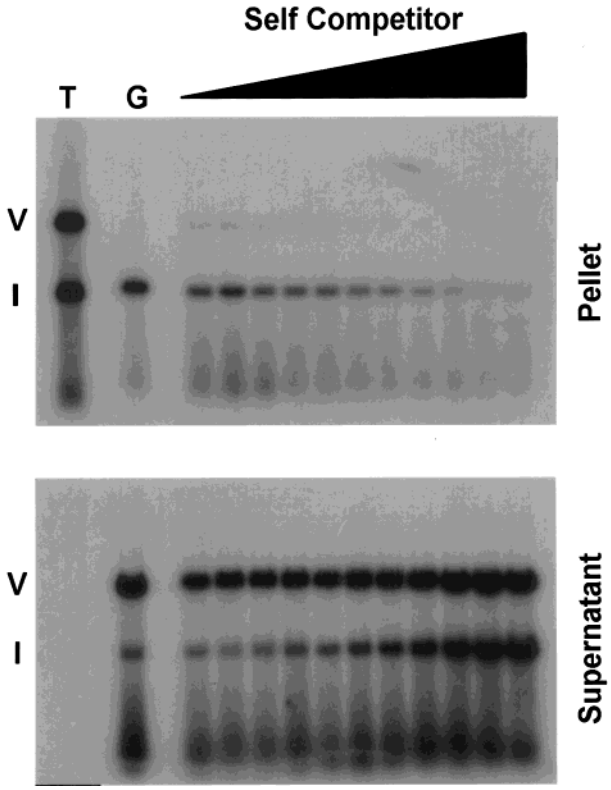

FIGURE 2: Self-competition of the tobacco RB7-6 MAR in an exogenous matrix binding assay. V = vector, I = insert (MAR), T = aliquot of total labeled probe used in the assay. The aliquot of the total labeled probe used in the assay is shown only in the top panel. G = genomic DNA as competitor (standard assay as in Figure 1). Self Competitor = increasing levels of unlabeled RB7-6 from 0.01 to 10 pmol per assay with 300 000 nuclear matrixes. DNA partitioning with the pellet (i.e., bound to the nuclear matrix) is in the top panel. DNA partitioning with the supernatant (i.e., not bound to the nuclear matrix) is in the bottom panel.

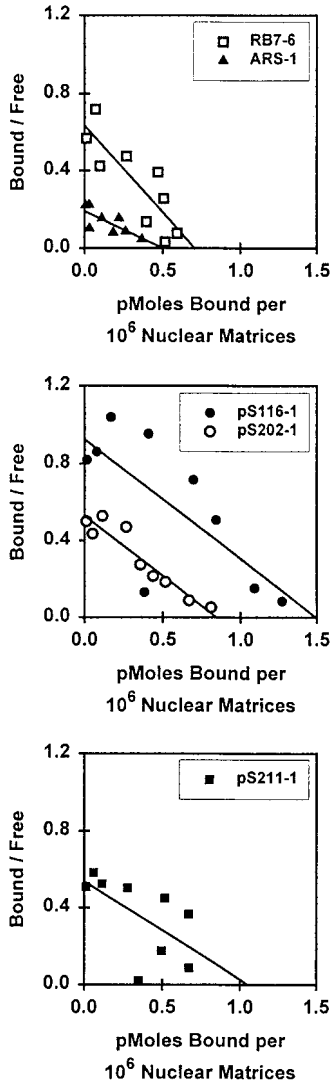

FIGURE 3: Scatchard plots of strong, medium, and weak MARs. In each case, exogenous matrix binding assays were carried out as in Figure 2 with isolated MAR inserts used as self-competitors for radioactively labeled MARs. The ratio of bound/free MARs is plotted versus the picomoles of bound MAR per 10^6 nuclear matrixes. In addition to the plant MARs, data for the weak binding yeast MAR ARS-1 are included. The MARs are classified as strong, medium, and weak based on the percent of probe found in the pellet in the *standard exogenous binding assay* using genomic tobacco DNA as competitor as in Figure 1. The strong binders are tobacco pS116-1 (90% in pellet in *standard exogenous binding assay)* and RB7-6 (80%). A medium binder is tobacco pS211-1 (70%). A weak to medium binder is tobacco pS202-1 (40%), and yeast ARS-1 is a weak binder (<10%).

the product of the association constant and the number of binding sites (last column of Table 2). Thus, it appears that the differences in binding strength of the MARs we have investigated are due to a combination of differences in numbers of sites available and affinity for the sites. The results of our cross-competition studies indicate that binding sites for the various MARs overlap extensively. We have compared saturation curves for the binding of pS202-1 with self-competitor and pS202-1 with RB7-6 as competitor. The curves are essentially superimposable (data not shown). Thus, the two MARs are competing for many of the same binding sites.

Sequence Analysis. One of our goals was to use sequence information to enable us not only to *identify* a MAR but also to *predict the strength* of binding of a MAR to the nuclear matrix. Thus, we have sequenced the 13 randomly obtained MARs and analyzed these sequences (as well as that of the RB7-6 MAR) in a number of ways. We have (a) calculated the percentages of AT nucleotides in each of the MARs, (b) plotted the distribution of AT content in sliding windows across each of the MARs, (c) determined the number and distribution of several motifs that have previously been associated with MARs, and (d) correlated the numbers of these motifs in each of the MARs with the binding strength of the MARs.

Of the 12 MARs that show binding in the standard exogenous assay, 8 are 70% or greater in overall AT content (Table 1, Figure 6). The remaining four MARs that show binding in the standard exogenous assay have lower than 70% overall AT content. These MARs, however, have extensive regions of greater than 70% AT content as shown in the sliding window plots of AT content in Figure 4. Thus, the often cited characteristic of high AT content (*4*, *5*, *58*) applies to the plant MARs. Plots in Figure 4 also show that the weakest binders (i.e., those that do not have detectable binding by the standard exogenous assay) have only very limited regions in which the local AT content is greater than 70%.

We have used the Genetics Computer Group Wisconsin Package software to identify in the plant MARs a number of motifs that have been associated with previously identified

Table 2: Dependence of MAR Binding Strength on Dissociation Constants and Binding Sites per Nuclear Matrix[a]

MAR	binding strength	K_a (M^{-1})	sites/matrix	(K_a × sites) × 10^{-12}
pS116-1	90	1.1×10^8	900 000	99
RB7-6	80	1.7×10^8	400 000	67
pS211-1	70	0.9×10^8	600 000	55
pS202-1	40	1.1×10^8	500 000	56
ARS-1	10	0.7×10^8	300 000	21

[a] Binding strength = % of MAR partitioning with the nuclear matrix in a standard exogenous assay. Dissociation constants (K_d) and binding sites per matrix were calculated from plots in Figure 3. K_a (association constant) = $1/K_d$. (K_a × sites) × 10^{-12} = association constant times binding sites per nuclear matrix for each MAR multiplied by 10^{-12}.

MAR sequences. We have identified A-box and T-box motifs (*54*), the yeast ARS consensus sequence (*55*), the consensus sequence for *Drosophila* topoisomerase II (*7*, *73*), the base unpairing region (BUR) (*4*, *5*, *63*, *64*), the MAR recognition sequence (MRS) (*44*), origin of replication and homeotic protein recognition motifs (ORI), TG-rich sequences, curved DNA motifs, and kinked DNA motifs (*56*, *57*, *65*, *66*, *68*). The frequency of occurrence of these motifs in each MAR is listed in Table 1.

In Figure 5 the length of the MAR sequences, their AT content, and their content of seven MAR motifs are plotted versus binding strength. The highest correlation between binding strength and numbers of sequence motifs is found for the motif that we call the "90%AT box" ($r = 0.81$). This motif is defined simply as 20 consecutive base pairs of which 18 or more are AT base pairs.

Several studies have led to the conclusion that sequences shorter than 300 bp do not bind strongly to the nuclear matrix [reviewed in Mielke et al. (*60*) and Gasser et al. (*11*)]. Our work supports the conclusions drawn from these studies (data not shown). Binding strength, however, does not appear to be strongly dependent upon length per se. The correlation between binding strength and length of the MARs we have studied is weak ($r = 0.41$). Longer MARs, of course, have the potential for extended regions that bind to the nuclear matrix and thus increase the strength of binding through cooperativity. Bode and co-workers (*63*) have shown that oligomerizing DNA fragments containing base unpairing regions enables binding to the nuclear matrix.

The correlation between the MAR recognition sequence (MRS) of van Drunen and co-workers (*44*) and binding strength is quite high ($r = 0.77$). Despite the strong correlation, this motif may not be useful as a predictor of matrix binding strength. Several of our MARs that bind in the 40−50% range do not have any of the MRS sequences. This is true even though gaps, overlaps, and up to two mismatches were allowed in the identification of the motif. None of our tobacco MARs had an MRS motif with no mismatches. It should be noted, however, that van Drunen and co-workers proposed this sequence only to be an *identifier* of MARs in *Arabidopsis*. They did not claim that this sequence would correlate with binding strength.

The overall %AT and the number of T-boxes correlate moderately well with binding strength ($r = 0.74$ and 0.72, respectively). A-boxes, BUR, and ARS motifs have considerably lower correlations with binding strength ($r = 0.58$, 0.55, and 0.39, respectively). There is essentially no correlation between *Drosophila* topoisomerase II consensus sequences and binding strength of the tobacco MARs ($r = 0.08$). Thus, this consensus sequence does not appear to be useful for the assessment of MAR affinity for the nuclear matrix in tobacco.

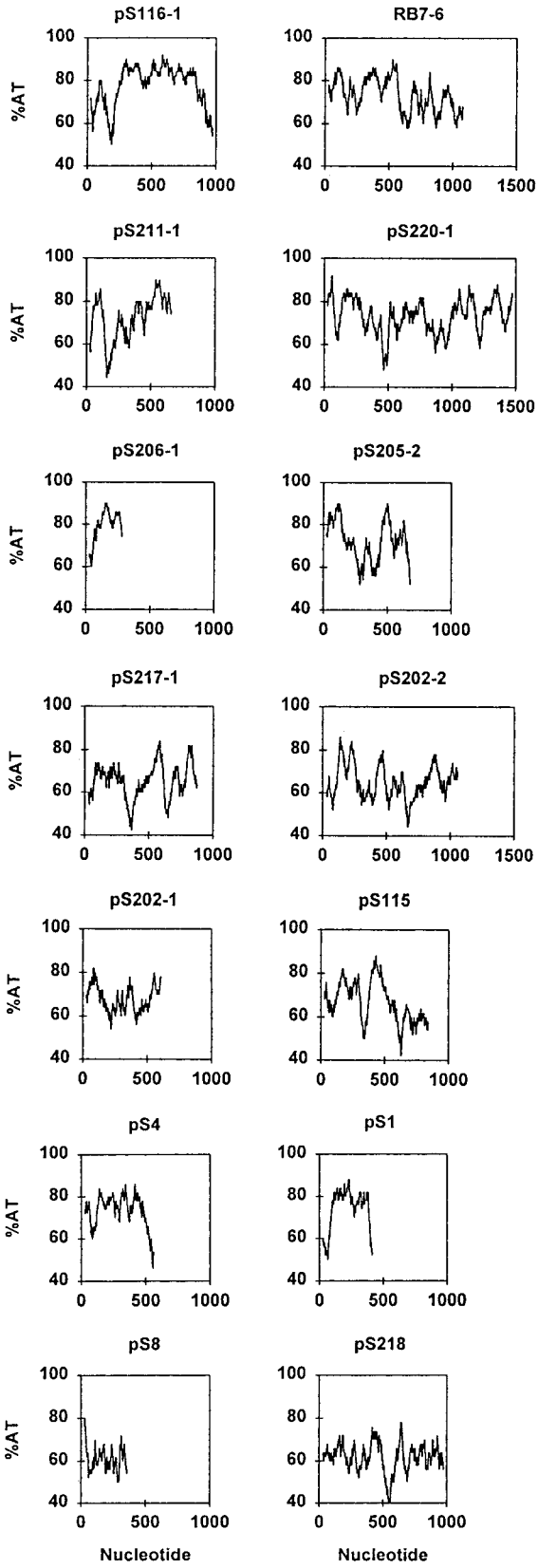

FIGURE 4: Sliding window of %AT across the plant MARs. The %AT in windows of 50 bp sliding at 1 bp intervals is plotted versus nucleotide position in each of the plant MARs.

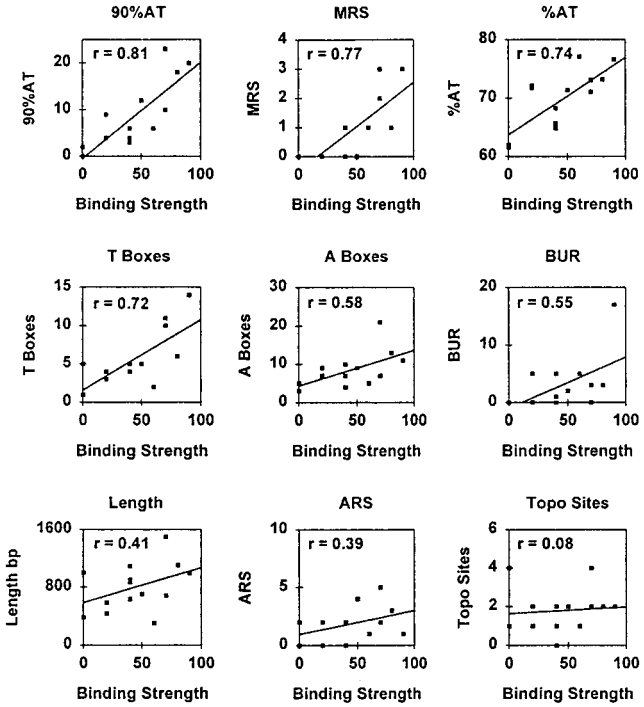

FIGURE 5: Correlation of MAR-related motifs and binding strength. The abundance of MAR-related motifs is plotted versus binding strength for the 14 plant MARs. Motifs are defined in the legend to Table 1. The lengths of the MARs in bp and their overall AT content (%AT) are also plotted versus binding strength. r = regression coefficient.

Additional motifs, ORI, TG-Rich, Curved, and Kinked, have been suggested as diagnostic of MARs by Boulikas (*57*) and used by Krawetz and co-workers (*65*, *66*, *68*) in their MAR-Finder program. The plots for the occurrence of these motifs vs binding strength are not included in Figure 5, but the Curved and ORI motifs correlated positively with binding strength ($r = 0.74$ and 0.58, respectively). The TG-Rich and Kinked motifs correlate negatively ($r = -0.14$ and -0.25, respectively).

Figure 6 is a graphic presentation of the length of the plant MARs and the distribution of the following motifs: 90%AT boxes, A-boxes, T-boxes, ARS motifs, Topo II motifs, and base unpairing regions. Visual inspection of the data presented in this way makes it obvious that the overall density of motifs correlates with binding strength. Certainly the strongest binding MAR, pS116-1, has the highest density of total motifs. The density of the 90%AT boxes is particularly noticeable. At the other end of the scale, pS218 (a nonbinder) has no 90%AT boxes. This DNA fragment does have examples of all the other motifs considered in the figure (except for base unpairing regions), but at low density.

DISCUSSION

We have isolated several plant MARs by cloning DNA fragments that are MARs by operational definition, i.e., DNA fragments that co-isolate with tobacco nuclear matrixes. A caveat must be introduced here, namely, that we have no direct proof that any one operationally defined MAR is actually associated with the nuclear matrix in vivo. The LIS method, which we have used, was devised to overcome potential artifacts associated with the high-salt method (*37*). But once the system is perturbed with either method, DNA

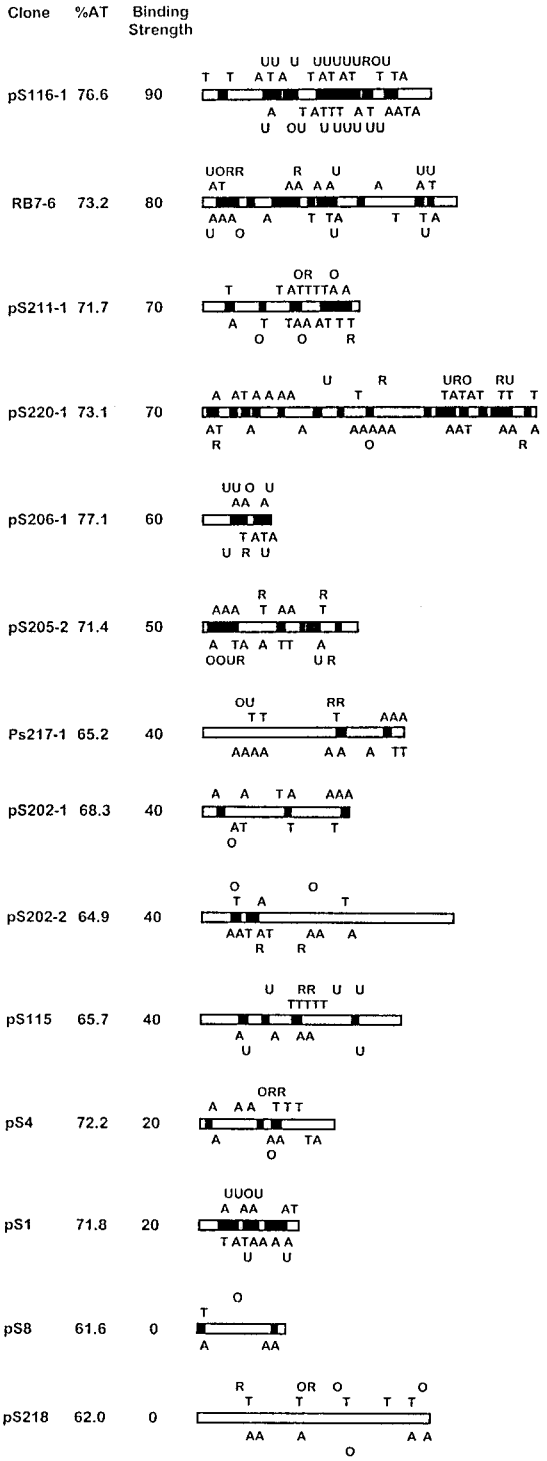

FIGURE 6: Distribution of motifs in the plant MARs. For each MAR, the overall %AT and binding strength are listed. Bars are proportional to the length of the MARs in base pairs. Filled boxes represent 90%AT boxes (many are contiguous). Symbols representing the other motifs are placed above or below the bar at the positions where they occur in the MARs. A = A-box motif, T = T-box motif, R = ARS motif, O = Topo II motif, and U = base unpaired region (BUR) motif. The sequences of the motifs are specified in the legend of Table 1.

sequences with the *potential* to bind to the nuclear matrix may do so, regardless of whether they are actually associated with the nuclear matrix in vivo or not.

The MARs we have studied were obtained randomly in the sense that they originated from the entire population of

DNA sequences rather than from a subset of sequences preselected for proximity to well-characterized genes. Thirty-four of the clones were characterized for strength of binding to the nuclear matrix by the *standard exogenous binding assay*, in which genomic DNA is used as competitor. Thirteen clones, representing strong, medium, and weak binders, were selected for sequencing and further studies.

We have asked if the binding strength of the MARs as assayed using genomic DNA as competitor is a function of (a) numbers of binding sites in the nuclear matrix for a specific MAR, (b) affinity of a specific MAR for its cognate sites in the nuclear matrix, or (c) a combination of both properties. To approach this question, we have carried out binding studies with five MARs using the isolated MARs as self-competitors. For these studies, we have used a strong, a medium, and a weak binder from the library of randomly isolated tobacco MARs as well as two MARs we have worked with previously, the strong tobacco MAR RB7-6 and the weak yeast MAR ARS-1 (*12*, *13*, *41*). Numbers of binding sites per matrix and dissociation constants were calculated using Scatchard plots. There is only a weak correspondence between rank order of binding strength and association constant ($1/K_d$) or between rank order of binding strength and numbers of binding sites per matrix. As shown in Table 2, however, there is good correspondence between rank order of binding strength and the product of association constant and numbers of binding sites per matrix (note that this product is equal to the *y*-intercept of the Scatchard plot). This suggests that the differences in binding strength of the MARs tested by the *standard exogenous binding assay* are due to both numbers of sites and affinities for these sites. This conclusion is bolstered by the observation that the strong MAR, RB7-6, and the weaker MAR, pS202-1, compete for many of the same binding sites. Saturation binding curves for self-competition and cross-competition of these MARs overlap (data not shown).

Our estimates of the numbers of binding sites for the tobacco MARs range from 400 000 to 900 000 per matrix (Table 2). This is higher than the estimates that have been made for MAR binding sites in human nuclei. Gross and Garrard (*75*) have estimated between 10 000 and 100 000, while Mielke et al. (*60*) estimate between 3000 and 14 000. Although our estimates are higher, they seem quite reasonable based on genome size and loop domain size considerations. The tobacco genome is somewhat larger than the human genome, and the average size of the loop domains may be smaller. We have used flow microfluorometry to determine the number of base pairs of DNA in the genome of tobacco NT-1 cells in culture to be $\sim 15 \times 10^9$ bp per 2C nucleus (data not shown). Dividing this value by our highest estimate of binding sites (9×10^5 per nucleus) results in an average size of loop domain plus MAR of about 17 kb. This result compares favorably with the domain structure reported for regions in the rice and sorghum genomes (*61*) in which seven putative loop domains have an average size of about 9 kb. Thus, the average size of loop domains in plants may be smaller than in humans. Gross and Garrard (*75*) estimate an average loop domain size of 70 kb for humans, while Jackson et al. (*76*) have estimated 86 kb, and Vogelstein et al. (*77*) have estimated $\sim 50-100$ kb. Paul and Ferl (*78*) have estimated the mean size of putative loops delimited by "loop basements" to be 45 kb for maize and 25 kb for *Arabidopsis*. The relationship between MARs and "loop basements" has yet to be determined.

Several attempts have been made to identify MARs based on sequence alone [reviewed by Benham et al. (*64*)]. The most common approach is to identify sequence motifs that occur in known MARs and suggest that these motifs may be useful for predicting matrix binding for an untested DNA fragment of known sequence. Krawetz and co-workers (*65*, *66*, *68*) have combined a number of these motifs into a program, which they call "MAR-Finder." This approach appears to have merit, and in many cases there is a correlation between sequences identified by the program and sequences that have been experimentally verified as binding to the nuclear matrix. Avramova et al. (*61*), however, report that none of the sequence motifs can be used as a reliable probe for predicting matrix binding in rice or sorghum.

Even though sequence motifs have been used as potential predictors of matrix binding, it is widely accepted that general features of the structure of the DNA, e.g., narrow minor groove, tendency to unwind, tendency to form bent DNA, tendency to form cruciforms), are the determining factors (*64*). Boulikas and Kong (*53*) have proposed that a high density of inverted repeats is diagnostic of MARs. None of the plant MARs we have isolated, however, has a high density of inverted repeats. The most extensive stem structure found was 20 base pairs in length. We have assessed the relationship between binding strength and numbers of potential cruciforms with stems of 9 or more base pairs. The correlation coefficient for these two parameters is only 0.26 (data not shown).

The procedures devised to identify MARs by sequence information alone have the goal of determining if a specific DNA fragment is or is not a MAR. This is a good first step, but it is widely recognized that there is no sharp distinction between MAR DNA and non-MAR DNA as determined by in vitro binding [reviewed by Bode et al. (*5*)]. There is a continuum of affinities from nonbinders to very strong binders. In the work presented here, we have attempted to correlate binding strength with the abundance of a number of motifs and structural characteristics that previously have been considered diagnostic for MARs. We have assessed their potential for identifying MARs by determining the correlation between the numbers of the various motifs and the strength of binding of the MAR.

Four of the most widely considered motifs are A-boxes and T-boxes (*54*), the ARS motif (*55*), and the *Drosophila* topoisomerase II motif (*7*, *73*). A-boxes and T-boxes appear to have moderate to good value for predicting matrix binding strength ($r = 0.58$ and 0.72, respectively). The ARS motif appears to have little predictive value ($r = 0.39$). There is essentially no correlation between binding strength of the tobacco MARs and the presence of the *Drosophila* topoisomerase II motif ($r = 0.08$). This is in accordance with the data of Avramova et al. (*61*), who found none of these motifs in the nine rice and sorghum MARs they studied, and with the data of van Drunen et al. (*44*), who found that these motifs did not correlate with the presence of the MARs they studied in *Arabidopsis*. These observations may or may not be surprising, depending on the basis for the occurrence of topoisomerase II sites in animal MARs. If the presence of topoisomerase II motifs in MARs were directly connected to the necessity of MARs to interact with topoisomerase II

in vivo, the lack of *Drosophila* topoisomerase II motifs in plant MARs would not be considered surprising. We might reason that even though *Drosophila* topoisomerase II motifs are not found in plant MARs, if *plant* topoisomerase II motifs were known, they would be found in plant MARs. It is quite possible, however, that the sequence contained in the *Drosophila* topoisomerase II motif contributes to some structural feature of DNA that is necessary for binding to the nuclear matrix but only fortuitously corresponds to topoisomerase II interaction sites. This possibility is bolstered by the fact that *Drosophila* topoisomerase II motifs have been considered diagnostic for vertebrate MARs [reviewed in Boulikas (*58*)] and by the contention that topoisomerase II cleavage−ligation sites in vitro do not correspond to those used in vivo (*74*). Another possibility is that even though *Drosophila* topoisomerase II motifs have been found in MARs from various vertebrate animal sources, they may not differentiate MARs from other DNA sequences. More widespread analysis of large blocks of sequence containing MAR and non-MAR sequences, such as the analysis carried out by Avramova et al. (*61*), would seem to be in order.

None of the four commonly considered MAR motifs discussed above correlate with binding strength as well as %AT ($r = 0.74$). Thus, we might be tempted to think of %AT as a good predictor of binding strength. However, high AT content alone is not sufficient for MAR function (*44*, *47*, *60*, *61*). Although %AT per se may not be useful, we propose that the following refinement may be useful. Inspection of the tobacco MAR sequences revealed the occurrence of a number of blocks of 20 nucleotides of which 18 or more (90%) were A or T. The number of such "90%AT blocks" correlates better with binding strength ($r = 0.81$) than any of the other motifs or structural features we have investigated. Thus, the number of 90%AT blocks may be a useful predictor of DNA affinity for the nuclear matrix.

Bode and co-workers (*63*) have proposed that unwinding tendency is the prime determinant of matrix binding. The base unpairing sequence (BUR) proposed by these workers to be diagnostic of MAR sequences was not highly correlated with binding strength in our tobacco MARs ($r = 0.55$). However, this group has recently developed a more sophisticated computational analysis which emphasizes the presence and arrangement of stress-induced base unpairing regions (*64*). In the sequences they have analyzed to date, there is good correspondence between the parameters measured and sequences with MAR activity measured by in vitro binding assays. Using the stress-inducing duplex destabilization calculations, they have predicted the relative binding strength of the tobacco MARs we have studied in this work. These predictions correlate well ($r = 0.89$) with our experimentally determined binding strengths (*79*). Thus, it appears as though unwinding tendency (as determined by stress-induced duplex destabilization profiles) will be a good predictor of MAR binding strength as well as the presence or absence of MARs.

REFERENCES

1. Flemming, W. (1882) *Zellsubstanz, Kern und Zelltheilung*, F. C. W. Vogel, Leipzig, Germany.
2. Macgregor, H. C. (1986) in *Chromosome Structure and Function* (Risley, M. S., Ed.) pp 152−186, Van Nostrand & Reinhold, New York.
3. Paulson, J. R., and Laemmli, U. K. (1977) *Cell 12*, 817−828.
4. Bode, J., Schlake, T., Ríos-Ramírez, M., Mielke, C., Stengert, M., Kay, V., and Klehr-Wirth, D. (1995) *Int. Rev. Cytol. 162A*, 389−454.
5. Bode, J., Stengert-Iber, M., Kay, V., Schlake, T., and Dietz-Pfeilstetter, A. (1996) *Crit. Rev. Eukaryotic Gene Expression 6*, 115−138.
6. Cockerill, P. N., and Garrard, W. T. (1986) *Cell 44*, 273−282.
7. Gasser, S. M., and Laemmli, U. K. (1986) *EMBO J. 5*, 511−518.
8. Mirkovitch, J., Spierer, P., and Laemmli, U. K. (1986) *J. Mol. Biol. 190*, 255−258.
9. Bode, J., and Maass, K. (1988) *Biochemistry 27*, 4706−4711.
10. Spiker, S., and Thompson, W. F. (1996) *Plant Physiol. 110*, 15−21.
11. Gasser, S. M., Amati, B. B., Cardenas, M. E., and Hofman, J. F.-X. (1989) *Int. Rev. Cytol. 119*, 57−96.
12. Allen, G. C., Hall, G. E., Jr., Childs, L. C., Weissinger, A. K., Spiker, S., and Thompson, W. F. (1993) *Plant Cell 5*, 603−613.
13. Allen, G. C., Hall, G., Jr., Michalowski, S., Newman, W., Weissinger, A. K., Spiker, S., and Thompson, W. F. (1996) *Plant Cell 8*, 899−913.
14. Phi-Van, L., von Kries, J. P., Ostertag, W., and Strätling, W. H. (1990) *Mol. Cell. Biol. 10*, 2302−2307.
15. Phi-Van, L., and Strätling, W. H. (1996) *Biochemistry 35*, 10735−10742.
16. Poljak, L., Seum, C., Mattioni, T., and Laemmli, U. K. (1994) *Nucleic Acids Res. 22*, 4386−4394.
17. Schöffl, F., Schröder, G., Kliem, M., and Rieping, M. (1993) *Transgenic Res. 2*, 93−100.
18. Stief, A., Winder, D. M., Strätling, W. H., and Sippel, A. E. (1989) *Nature (London) 341*, 343−345.
19. Thompson, E. M., Christians, E., Stinnakre, M.-G., and Renard, J.-P. (1994) *Mol. Cell. Biol. 14*, 4694−4703.
20. Xu, M., Hammer, R. E., Blasquez, V. C., Jones, S. L., and Garrard, W. T. (1989) *J. Biol. Chem. 264*, 21190−21195.
21. Kalos, M., and Fournier, R. E. K. (1995) *Mol. Cell. Biol. 15*, 198−207.
22. Klehr, D., Maass, K., and Bode, J. (1991) *Biochemistry 30*, 1264−1270.
23. McKnight, R. A., Shamay, A., Sankaran, L., Wall, R. J., and Henninghausen, L. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 6943−6947.
24. Mlynárová, L., Loonen, A., Heldens, J., Jansen, R. C., Kelzer, P., Stiekema, W. J., and Nap, J.-P. (1994) *Plant Cell 6*, 417−426.
25. Mlynárová, L., Jansen, R. C., Conner, A. J., Stiekema, W. J., and Nap, J.-P. (1995) *Plant Cell 7*, 599−609.
26. Breyne, P., Van Montagu, M., Depicker, N., and Gheysen, G. (1992) *Plant Cell 4*, 463−471.
27. Liu, J.-W., and Tabe, L. M. (1998) *Plant Cell Physiol. 39*, 115−123.
28. Amati, B., Pick, L., Laroche, T., and Gasser, S. M. (1990) *EMBO J. 9*, 4007−4016.
29. Amati, B., and Gasser, S. M. (1990) *Mol. Cell. Biol. 10*, 5442−5454.
30. Dijkwel, P. A., and Hamlin, J. L. (1995) *Int. Rev. Cytol. 162A*, 455−484.
31. Carri, M. T., Micheli, G., Graziano, E., Pace, T., and Buongiorno-Nardelli, M. (1986) *Exp. Cell Res. 164*, 426−436.
32. Hozak, P., and Cook, P. R. (1994) *Trends Cell Biol. 4*, 48−52.
33. Brun, C., Dang, Q., and Miassod, R. (1990) *Mol. Cell. Biol. 10*, 5455−5463.
34. Mah, D. C., Dijkwel, P. A., Todd, A., Klein, V., Price, G. B., and Zannis-Hadjopoulos, M. (1993) *J. Cell Sci. 105*, 807−818.
35. Hall, G. E., Jr., and Spiker, S. (1994) in *Plant Molecular Biology Manual* (Gelvin, S., Schilperoort, R., and Verma, D. P., Eds.) pp D2: 1−12, Kluwer Academic Publishers, Dordrecht, The Netherlands.

36. Kay, V., and Bode, J. (1995) in *Methods in Molecular and Cellular Biology: Methods for studying DNA−protein interactions: an overview* (Papavassiliou, A. G., and King, S. L., Eds.) pp 186−194, Wiley-Liss, Inc., New York,.
37. Mirkovitch, J., Mirault, M.-E., and Laemmli, U. K. (1984) *Cell 39*, 223−232.
38. Käs, E., and Chasin, L. A. (1987) *J. Mol. Biol. 198*, 677−692.
39. Jarman, A. P., and Higgs, D. R. (1988) *EMBO J. 7*, 3337−3344.
40. Phi-Van, L., and Strätling, W. H. (1988) *EMBO J. 7*, 655−664.
41. Hall, G., Jr., Allen, G. C., Loer, D. S., Thompson, W. F., and Spiker, S. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 9320−9324.
42. van der Geest, A. H. M., Hall, G. E., Jr., Spiker, S., and Hall, T. C. (1994) *Plant J. 6*, 413−423.
43. Chinn, A. M., and Comai, L. (1996) *Plant Mol. Biol. 32*, 959−968.
44. van Drunen, C. M., Oosterling, R. W., Keultjes, G. M., Weisbeek, P. J., van Driel, R., and Smeekens, S. C. M. (1997) *Nucleic Acids Res. 25*, 3904−3911.
45. Surdej, P., Got, C., Rosset, R., and Miassod, R. (1990) *Nucleic Acids Res. 18*, 3713−3722.
46. Iarovaia, O., Hancock, R., Lagarkova, M., Miassod, R., and Razin, S. (1996) *Mol. Cell. Biol. 16*, 302−308.
47. Fernandez, M. A., Baron, B., Prigent, M., Toledo, F., Buttin, G., and Debatisse, M. (1997) *J. Cell. Biochem. 67*, 541−551.
48. Avramova, Z., SanMigel, P., Georgieva, E., and Bennetzen, J. L. (1995) *Plant Cell 7*, 1667−1680.
49. Avramova, Z., and Bennetzen, J. L. (1993) *Plant Mol. Biol. 22*, 1135−1143.
50. Dijkwel, P. A., and Hamlin, J. L. (1988) *Mol. Cell. Biol. 8*, 5398−5409.
51. Nikolaev, L. G., Tsevegiyn, T., Akopov, S. B., Ashworth, L. K., and Sverdlov, E. D. (1996) *Nucleic Acids Res. 24*, 1330−1336.
52. Boulikas, T., and Kong, C. F. (1993) *Int. J. Oncol. 2*, 325−330.
53. Boulikas, T., and Kong, C. F. (1993) *J. Cell. Biochem. 53*, 1−12.
54. Gasser, S. M., and Laemmli, U. K. (1986) *Cell 46*, 521−530.
55. Amati, B., and Gasser, S. M. (1988) *Cell 54*, 967−978.
56. Boulikas, T. (1993) *J. Cell. Biochem. 52*, 23−36.
57. Boulikas, T. (1993) *J. Cell. Biochem. 52*, 14−22.
58. Boulikas, T. (1995) *Int. Rev. Cytol. 162A*, 279−388.
59. Boulikas, T. (1992) *J. Cell. Biochem. 50*, 111−123.
60. Mielke, C., Kohwi, Y., Kohwi-Shigematsu, T., and Bode, J. (1990) *Biochemistry 29*, 7475−7485.
61. Avramova, Z., Tikhonov, A., Chen, M., and Bennetzen, J. L. (1998) *Nucleic Acids Res. 26*, 761−767
62. Homberger, H. P. (1989) *Chromosoma 98*, 99−104.
63. Bode, J., Kohwi, Y., Dickinson, L., Joh, R. T., Klehr, D., Mielke, C., and Kohwi-Shigematsu, T. (1992) *Science 255*, 195−197.
64. Benham, C., Kohwi-Shigematsu, T., and Bode, J. (1997) *J. Mol. Biol. 274*, 181−196.
65. Kramer, J. A., Singh, G. B., and Krawetz, S. A. (1996) *Genomics 33*, 305−308.
66. Singh, G. B., Kramer, J. A., and Krawetz, S. A. (1997) *Nucleic Acids Res. 25*, 1419−1425.
67. Strausbaugh, L. D., and Williams, S. M. (1996) *J. Theor. Biol. 183*, 159−167.
68. Walter, W. R., Singh, G. B., and Krawetz, S. A. (1998) *Biochem. Biophys. Res. Commun. 242*, 419−422.
69. Michalowski, S. M. (1997) Ph.D. Dissertation, Department of Genetics, North Carolina State University, Raleigh, NC.
70. Strissel, P. L., Dann, H. A., Pomykala, H. M., Diaz, M. O., Rowley, J. D., and Olopade, O. I. (1998) *Genetics 47*, 217−229.
71. Conkling, M. A., Cheng, C.-L., Yamamoto, Y. T., and Goodman, H. M. (1990) *Plant Physiol. 93*, 1203−1211.
72. Conners, K. A. (1987) *Binding Constants*, John Wiley & Sons, New York.
73. Sander, M., and Hsieh, T. S. (1985) *Nucleic Acids Res. 13*, 1057−1072.
74. Käs, E. (1992) *EMBO J. 11*, 705−716.
75. Gross, D. S., and Garrard, W. T. (1987) *Trends Biochem. Sci. 12*, 293−297.
76. Jackson, D. A., Dickinson, P., and Cook, P. R. (1990) *EMBO J. 9*, 567−571.
77. Vogelstein, B., Pardoll, D. M., and Coffey, D. S. (1980) *Cell 22*, 79−85.
78. Paul, A. L., and Ferl, R. J. (1998) *Plant Cell 10*, 1349−1359.
79. Bode, J., Bartsch, J., Boulikas, T., Iber, M., Mielke, C., Schübeler, D., Seibler, J., and Benham, C. (1998) *Gene Ther. Mol. Biol. 1*, 551−580.

BI991142C