BIOLOGIJA. 2005. Nr. 4. P. 1–8

DNA sequences involved in yeast tightly bound DNA–protein complexes

L. Bagdoniene*,

K. Bonikataite,

V. Borutinskaite,

D. Labeikyte,

B. Juodka

Department of Biochemistry and Biophysics, Vilnius University, Èiurlionio 21, LT-03101 Vilnius, Lithuania

A computer nucleotide sequence analysis of fourteen of DNA fragments involved in yeast *Saccharomyces cerevisiae* tightly bound DNA–protein complexes shows that these sequences are heterogeneous, but have some motifs common to all of them: topoisomerase II binding sites, short A + T-rich stretches, kinked DNA streches. These sequences are rich in transcription factor binding sites and origin of replication sites. DNA fragment localization sites in the genome are different. There are neither telomeric nor centromeric sequences. Our results show that the study sequences have some characteristic of S / MAR sequences and may be important for DNA structurization and for the regulation of gene expression and DNA replication.

Key words: chromosomal DNA loops, nuclear matrix DNA, S/MAR sequences

Abbreviations: TBD – eukaryotic non-histone proteins tightly bound to DNA; S / MARs – conserved DNA sequences associated with matrix and / or scaffold.

INTRODUCTION

The eukaryotic genome is organized in a hierarchical fashion within the nucleus. Eukaryotic chromosomes are thought to be organized into a higher-order structure consisting of discrete and topologically independent loop domains attached to the nuclear matrix by S / MAR sequences. The loop organization of chromosomes may be important not only for compaction of the chromatin fiber, but also for regulation of gene expression and replication [1, 2]. It is now generally assumed that mammalian origins of replication are not exclusively determined by their DNA sequence, but their function relies on epigenetic principles, such as the presence of bound transcription factors, chromatin structure, or nuclear localization [3, 4].

Like in higher eukaryotes, the organization of DNA into chromatin and chromosomal structures plays a central role in many aspects of yeast cell biology. The processes ranging from chromosome stability and segregation to gene expression are intimately linked to chromatin configuration [5]. The yeast genome compared with that of multicellular organisms is more compact. Non coding regions in it occupy very little space, and genes lie in apparent promiscuity [6]. Recent data suggest that yeast gene regulation may be more susceptible to long-range chromatin interaction than previously thought. Therefore, yeast may also have evolved efficient mechanisms for insulating the genes from each other [7]. What proteins are involved in the 3-dimensional genome organization is not completely comprehensible. In this context, those polypeptides that are able to form either permanent or transient tightly bound or covalent complexes with DNA are of special interest [8–10, 13, 17].

Tightly bound proteins – the protein group that remains attached to DNA after its deproteinisation – have been found in numerous eukaryotically distant species [8–10]. They are released from DNA only after DNA digestion with DNase I. These proteins are part of insoluble nuclear matrix proteins [17]. Only a few of these proteins have been characterized. Recently it has been demonstrated that some of such peptides tightly bound to DNA, separated from Erlich ascite and yeast cells, show phosphatase and kinase activity [14]. C1D first expressed 16 kDa tightly bound to DNA protein. C1D functions as a component of a complex involved in transcriptional repression [9].

It has been suggested that proteins co-isolated with DNA under conditions that release other peptide materials from DNA are involved into DNA

* Corresponding author. E-mail: lida.bagdoniene@gf.vu.lt

structurisation and participate in the regulation of gene expression and replication [8, 9]. TBD protein distribution in the genome is site-specific. They are enriched in several reiterated sequences in the attachment sites to the nuclear skeleton [10–12], but their distribution in unique genes reflects the type of cell differentiation [13].

Recently tightly bound DNA–protein complexes have been isolated from yeast *Sacharomyces cerevisiae* cells and the composition and some properties of proteins in these complexes have been determined [14].

The aim of the current research was to investigate the properties of the DNA sequences isolated from the tightly bound DNA–protein complexes of yeast *Saccharomyces cerevisiae* cells.

MATERIALS AND METHODS

Materials

[α-^{32}P]dATP (3000 Ci/mmol: NEN) and [γ-^{32} P]dATP were from Amersham; ligation and sequencing kits were from FERMENTAS; other chemicals were of possibly highest quality.

Cell lysis and DNA salting-out procedure were described in detail elsewhere [14].

Separation of DNA protected by co-isolated polypeptides

Aliquots containing 6 mg of DNA with tightly bound proteins were precipitated with 2 vol. ethanol. DNA pellets were collected, dried and dissolved in 50 µl of DNase I buffer (50 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$). DNase I (Pharmacia, RNase free) (10 u/mg DNA) were added and, after overnight incubation at room temperature under dialysis conditions against DNase I buffer, DNA-polypeptide complexes released by DNase I digestion were subjected to 1% agarose gel electrophoresis. Material from 1% agarose gel start was recovered and treated with proteinase K (1mg/ml) in TE 0.1 (10 mM Tris-HCl, pH 7.5; 0.1 mM EDTA), containing 0.5% SDS (3 h, 37 °C). The digest was twice phenolized, precipitated with ethanol, and the resulting DNA fragments were blunt-ended and ligated into pUC19 / *Sma*I vector. Transformation of *Escherichia coli* XLI-Blue competent cells resulted in colonies containing recombinant plasmids with inserts which were analyzed by conventional techniques [15]. Fourteen clones were sequenced by the dideoxy chain-termination method. The nucleotide sequences of the clones have been deposited in the Genbank under the accession numbers: Seq-1-DQ206873, Seq-2 - DQ206874, Seq-3 - DQ206875, Seq-4 - DQ206876, Seq-5 - DQ206877, Seq-6 - DQ206878, Seq-7 - DQ206879, Seq-8 - DQ206880, Seq-9 - DQ206881, Seq-10 - DQ206882, Seq-11- DQ206883, Seq-12 - DQ206884, Seq-13 - DQ206885, Seq-14-DQ206886.

Bioinformatics resources

Fourteen DNA sequences were compared with known sequences from the Genbank by the BLAST program (http:// www.ncbi.nlm.nih.gov/BLAST/), Saccharomyces genome database (http:// www.yeastgenome.org/), Fungal Genomes database (http://seq.yeastgenome.org/cgi-bin/blast-fungal.pl), MAR discovery program http://www.futuresoft.org/ MAR-Wiz/, program OMIGA 2.0 (Oxford Molecular), transcription factor binding site localization in DNA sequences was analyzed with (https:// www.genomatix.de/cgi- bin/matinspector_prof/).

RESULTS AND DISCUSSION

Our goal was to further characterize tightly DNA-bound complexes from budding yeast. Polypeptides copurifying with chromosomal DNA have been described in various eukaryotic systems such as Erlich ascite cells, rat liver cells, mouse mammary carcinoma cells, plants, but only little is known about protein composition and functions and about DNA sequences involved in these tightly bound DNA–protein complexes [8–14, 18].

Tightly bound proteins are released from DNA only after DNA digestion with DNase I. However, part of the DNA–polypeptide complexes were not completely dissociated during nuclease digestion. After a prolonged digestion of yeast *Saccharomyces cerevisiae* DNA with DNase I, 0.05–0.1% of undigested (protected) DNA was recovered. Following additional deproteinization of the nuclease-released undigested DNA, 200–300 bp DNA fragments were obtained. Blunt-ended DNA fragments were cloned into SmaI site of pUC19 vector (see Materials and Methods). Fourteen clones of the transformants containing inserts were picked at random and sequenced. Figure represents the sequences of these DNA inserts. Southern blot hybridization of total yeast DNA with sequences 1–14 revealed that these sequences are present in middle-copy numbers (results not shown).

No significant similarities were found among the different DNA clones (Figure). DNA fragment localization sites in the genome were also different. They were localized in different chromosomes and in different parts of individual chromosomes. There were no telomeric and no centromeric sequences. Computer analysis has also revealed that DNA fragments are localized both in the regions between genes and inside structural genes.

It has been shown earlier that DNA presented in tightly bound DNA–protein complexes in higher eukaryotic organisms exhibits some properties of S / MAR. These sequences have a high AT content and belong to the Alu family of sequences [10, 16].

Analysis of the nucleotide composition of budding yeast tightly bound DNA has shown that most

Seq-1 (1 – 557)
TTTCCACACTTTCCACGTTGGTTATCTCTACGTCTTCGTCTTCGTTCTCTCAGTTTCCATTATTTCCTTTCTTAACGCTTTG
ATTTTCTCTCCATTTCTTTATTTGCTTGCTCTTAAGGGTTAGGAATCTAATCCTCTTCAACTTCATCTCCAGTTTAC
TATTTTAATTCAAATTGGTCATTAAGTACATTCATCATCTCATCCTCATGTTCTTATTA**TATCCAAATTTTTCCATT**
CTTAAATTAACGTTTCCCTCATCATCATTCACTTCCCAATCTTCATCGTCTTCTCCATCTTCTTTATCTTATTTCACCCTCTC
CACATCTTCCTCTTCATCCATATACTCGTCTTCGTCATATCCTTCATTTAACCTTATACCTTAAGTAGATATTCTAGAT
TGCAGAGGAATAAACGACATGATCTGATTAATGGTTTtCTtAGTTTCGGTGAGAATATTATTATTCYTTATTCTC
TATTCTCTcATCTCTTCTAAATCATCTTCATCACTGCTATCGTTTAAAGTTGAATGCTCcTGATTATTGTAACCATTTTATTTT
Seq-2 (1 – 378)
TTTTATTCTTGACAAGCAGGAGTAGGTAGTTTATAACAATTAAATAATCCAG**AAACAAAGCTGTTACTACTTAGTCTTTTA**
TGAGAAGCTGAACTAGGTTGGAGAATATTCAAGGAAGATTGTGCAGGACTATTAGGCCTCTCCCCACTCCCCCCAAGGT
AATACTGGAGAAAAA**TGAAGCCAAGTGAAATGCTATTCTTGACAAGCAGGAGTAGGTAGTTTATAACAATTAAATAATC**
CAGAAACAAAGCTGTTACTACTTAGTCTTTTATGAGAAGCTGAACTAGGTTGGAGAATATTCAAGGAAGATTGTGCAGGACTATTA
GGCCTCTCCCCACTCCCCCCAAGGTAATACTGGAGAAAAATGAAGCCA
Seq-3 (1 – 492)
TCGCGAAAGCTCTAGTTCTGTAAAGTTTGTAGTCTAACATGTTATGTAAATATAATAGTTACTAGTTACTGCTTTTGTAC
ATAGAAACATGAGAATGTATTAAGTACtACCGGATTAGATGTGTATTCCTATATGTTTTCCTTGTCTCGAGTTTCCCATGAAGCGACTCTGAA
ACAATACTATTCACAAAATATATCATGCTTCTGCTCTGACACATGCATTAAAATCACTGAGTAAATTTATAAAA
AAAACTAACAA**TGAGCTTTTGCATAAAACCTACACATAGCAAGCCTTCATTTCATAGAAATATTAGACTCTCAAATAATTAT**
TATAAATAATAATAAGTATTTTCTGTTTATTTATCTCTATCTTTCGTCTTCGGCTTCTCTCAGTTTCCAGCGTTATCTTTCTTAACG
CTTTCTTTTTCTCTCCATTTCTTTATTTGCTTGCTCTTAAGGGTTAGGAATCTAATCCTCTTTTGCAAA
Seq-4(1–129)
GCACAACCTGTTCAACCAACTCAACCAACTCAACCTGTTCAGCCAATCAACCTGTTCAGCCAACTCAACCTGTTCA
ACCAACTCCGTATCGCCGATCACCGGGCGAATCACCGGCTGGCTGTCGTAGGC
Seq-5 (1 –131)
AGTGTTAGCCGAATTGAAAGTGGCCGTCGGTGACAAAGTCAGCGAAGGTTGCAGGTAGTACGGAGCAGAAGC
AGAAGCGGCTGCAGGAGTAGCTGCAGGAGCTCAGCGGCAGCGGAAGCCTGTCCCTGCTT
Seq-6 (1 –179)
TTCCCCGCTGATTCTGCCAAGCCCGTTCCCTTGGCTGTGGTTTCGCTAGATAGTAGATAGGGACAGTGGGAATCTCGT
TAATCCATTCATGCGCGTCACTAATTAGATGACGAGGCATTTGGCTACCTTAAGAGAGTCATAGTTACTCCCGCCG
TTTACCCGCGCTTGGTTGAATTTCT
Seq-7 (1 –267)
CAGCGTAGTACATAGACACACTACGGCACGTCGCGTAGCTCATGCCTGTAGCCTGTAGTACATGAATGTCAAAGCTC
AACAGGGTCTTCTTTCCCCGCTGATTCTGCCGCCCGTTCCCTTGGCTGTGG**TTTCGCTAGATAGTAGATAGGGACAGTGG**
GAATCTCGTTAATCCATTCATGCGCGTCACTAATTAGATGACGAGGCATTTGGCTACCTTAAGAGAGTCATAGTTACTC
CCGCCGTTTACCCGCGCTTGGTTGAATTTCT
Seq-8 (1 –178)
GGCTACGTCCACGAGCAGCTTGTAGTGCTCGCTATCGCGGA**TGAGCAACCCAAAACAGCTGCACAGCGTAGCAGCACCA**
CTGTACGAGTCACACTACGTCACACTACGTCACACGACATGAGATACGTACGTACGTACCCAGACGATTCGTCTCATACGT
GCCGTCTGAGACAGCGAC
Seq-9 (1 –153)
CTAGAGTGAG**CATAGGCAGCAWGCGTATGCATCGTAACGTGACTAGNGGYCADATAGCGYAGGTCGACTAGCGTGAT**
CAACTTTGCTATGATCCATACTGATCGTCGTTCTAAGGATCTGCTATTCAGTCAGTCGACTGCTAGCAGCTGACGT
Seq-10 (1 –235)
ACACTATGAAAAAGTATTTT**GTATTATTGCTTTGACTGCTTGCATGCCTGGCATGGTTACGACTCTAGATGGATGGCAAGAAAGTTTCGTTA**
TCGACCGATATTATTGAGGAAGAACCTGATGACTGCGTAGCTTTGTACGAGTCAGTGCAGTGCGCATACACtGAGACAGACCTG
GCCTGTACGAGCCATAGGCTCAACCAAGTACTGATCAGCAGTGCTCAGACCGATTACAGTCCCCG
Seq-11 (1 – 177)
GACACTATAGAATACGAATT**AATTCGAGCTCTTCGGTACCGTTGTCTCTTTTGCAGAAGCTCAGAATCAAACGTCACT**
TTGCAGATCTACASTGCTCGCTGATCTGTGTGTGTGTCATCATTGCAGTGTACCGTAGGTACGTGGACGACGCGGTCG
TCGATCATCACGACATGGTAGC
Seq-12 (1 –167)
AAATATTACT**GCATAATAGTCCAATAGCAAGGATATAGTCGCAAACATGTATAGCATTTCGAGATGCTTTCATCCCTCTG**
ACGATCGAGCTCCCTACGTCCCAGACAGTGGTSTSTSATCCGGACGATGCTCTCGATGCCTGGTACGTGTGTGACAAGATGCTACAT
Seq-13 (1 –157)
CTTTATTGAGGGAAATCGGTGGANGTGGACAGCGAGCGCCAGCGTTCAAGATCGGCGCGAACCTGGTCCAGCTTGG
CTTCNACGACACTGAGCGCATCGCTGCCAGCGACGAAACGCAAGGGCGGTGCTGCCANANCCGCCAATTGCACAACCGCCAGG
Seq-14 (1 –199)
GACCGAATCAACTCGGCGCCACGCTCGAACACGCGCGCACGGTCGTACCGTCAACCGGCCGTCTCGTTAGC
TTC**AACACACCGCTGGAAGGTCCTCCGGTAGGCTAGGCGGCGGAGGTACCCGCCATAAGCTCCGATCCTGTGAATATGTT**CGAGACAAT
GCTGGCTAACCCTATTCAAGTGTCTTCTATGTTCTTGCA

Figure. Composition of yeast *Sacharomyce cerevisiae* DNA fragments released from tightly bound DNA–protein complexes. According to MAR discovery program http://www.futuresoft.org/MAR-Wiz/ the underlined bold fragments of yeast tightly bound DNA sequences (1–14) show the MAR potential higher than 60%

Table 1. **Properties of yeast DNA fragments involved in tightly bound DNA protein complexes. (+) – direct strand, (–) – reverse strand**

SeqNr	ORI	Kinked DNA	Topo II site	AT sites
1(+)	58, 160, 179, 214, 241, 364	164		212, 223, 224, 225, 525
1(–)	424, 389, 314, 127, 62, 2 326, 314, 3 13, 11, 3, 2	381	304, 289, 211, 10	
2(+)	73, 166, 264, 357	63, 136, 254, 327		64, 255
2(–)	329, 326, 211, 138, 135, 20		389	326, 135
3(+)	101, 116, 222, 240, 311, 327, 330, 360		234, 333	222–224, 307–309, 322–28, 331–335, 340, 341
3(–)	438, 433, 262, 250, 163, 153, 150, 147, 17			249–245, 164–160, 151–145, 138, 129, 128
4(+)			1, 80, 105, 360	
4(–)			480, 443, 385, 6	
5(+)				
5(–)				
6(+)	101		100, 107	
6(–)	97, 76			
7(+)	189			
7(–)	97, 76		194	
8(+)			18, 48	
8(–)			0	
9(+)		29	24	54, 82
9(–)		112	28, 118	58
10(+)	23, 26	152	26, 36, 42, 154	
10(–)		71	176, 145, 95, 72, 12	
11(+)	17	114	44, 88, 150	
11(–)	154		132, 144	
12(+)	4		134, 164	
12(–)	150		50, 10	
13(+)			1, 16	
13(–)			130	
14(+)			49, 184	
14(–)		58	67	

Table 2. **Transcription factor binding sites in yeast TBD cloned sequences 1–14. (+) – direct strand, (–) – reverse strand**

Position	Strand	Sequence	Transcription factor family and characterization
Seq-1 (1–557)			
79–99	(–)	AGAAatggagagaaaatcaaa	HSF, heat shock factor
223–243	(–)	atttaAGAAtggaaaaatttg	HSF, heat shock factor
227–241	(+)	tttttcCATTcttaa	ABAA, contains an ATTS DNA-binding motif
322–336	(+)	catCTTCctcttcat	GCR1, activator of glycolytic enzymes genes
384–390	(–)	TATCtac	NIT2, activator of nitrogen-regulated genes
459–479	(–)	AGAAtagagaataargaataa	HSF, heat shock factor
524–544	(+)	TGAAtgctcctgattattgta	HSF, heat shock factor
Seq-2 (1–378)			
138–158	(+)	AGAAtattcaaggaagattgt	HSF, heat shock factor
175–197	(+)	tctccccactCCCCccaaggtaa	ZBP89, zinc finger transcription factor
211–231	(+)	TGAAgccaagtgaaatgctat	heat shock factor
218–234	(+)	aagtGAAAtgctattct	PRDM1, PRDI binding factor 1

Table 2 continued

Seq-3 (1–492)			
144–156	(–)	aacTCGAgacaag	XBP1, stress induced expression
210–222	(+)	ctcTGACacatgc	GCN4, dimerizing leucine zipper protein
251–269	(–)	aaagctCATTgttagtttt	STE11, needs cooperative Mat1-Mc binding
275–293	(+)	aaaaCCTAcacatagcaag	Yeast MADS-Box RLM1 transcription factor
275–289	(+)	aaaACCTacacatag	RAP1 (TUF1), activator or repressor
276–294	(–)	gcttGCTAtgtgtaggttt	Yeast MADS-Box RLM1 transcription factor
364–370	(+)	TATCtct	NIT2, activator of nitrogen-regulated genes
405–423	(+)	ttatctttcttaACGCttt	ABF1, binding to ARS sequences
421–441	(–)	AGAAatggagagaaaaagaaa	HSF, heat shock factor
Seq-4 (1–129)			
38–50	(+)	ttcagCCAAtcaa	HAP2, yeast factor complex HAP2/3/5
8–103	(–)	cccggtgatcgGCGAta	UAY involved in purine utilization
Seq-5 (1–31)			
107–121	(–)	aggCTTCcgctgccg	GCR1, activates glycolytic enzyme genes
Seq-6 (1–179)			
90–98	(+)	tttcCCCGc	REB1, *S. cerevisiae*
91–107	(–)	gcagaatcagcGGGGaa	MIG1, mediates glucose repression
111–127	(–)	cacaGCCAagggaacgg	PACC, pH-responsive regulator
134–140	(–)	TATCtag	NIT2, activator of nitrogen-regulated genes
218–230	(–)	ctaTGACtctctt	GCN4, dimerizing leucine zipper protein
243–251	(+)	tttaCCCGc	REB1, multifunctional transcription factor
Seq-7 (1–267)			
90–98	(+)	tttcCCCGc	REB1, multifunctional transcription factor
91–107	(–)	gcagaatcagcGGGGaa	MIG1, mediates glucose repression
111–127	(–)	cacaGCCAagggaacgg	PACC, pH-responsive regulator
133–145	(+)	gctaGATAgtaga	GATA1, GATA-binding factor 1
134–140	(–)	TATCtag	NIT2.activator of nitrogen-regulated enes
167–177	(+)	taatCCATtca	CCAAT, cellular and viral CCAAT box
179–195	(+)	gcgcgtcactAATTaga	ATBF, AT-binding transcription factor 1
218–230	(–)	ctaTGACtctctt	GCN4, leucine zipper protein
243–252	(+)	tttaCCCGc	REB1, multifunctional transcription factor
Seq-8 (1–178)			
121–133	(–)	gtgTGACtcgtac	GCN4, leucine zipper protein
120–139	(+)	acgacatgagatACGTacg	ABF1, ARS binding factor 1
158–164	(–)	TATCtca	NIT2, activator of nitrogen-regulated genes
153–172	(+)	gattcgtctcatACGTgcc	ABF1, binding to ARS sequences
Seq-9 (1–153)			
112–132	(–)	ctgacTGAAtagcagatcctt	HSF, heat shock factor
Seq-10 (1–235)			
131–143	– (–)	cacTGACtcgtac	GCN4, dimerizing leucine zipper protein
189–201	(+)	ctcaaCCAAgtac	HAP2, yeast factor complex HAP2/3/5
Seq-11 (1–177)			
1o8–112	(–)	tgaTGACacacac	GCN4, dimerizing leucine zipper protein
154–172	(+)	tcgtcgatcatcACGAcat	ABF1, binding to ARS sequences
Seq-12 (1–167)			
130–142	(+)	ctcTCGAtgcctg	XBP1, stress-induced expression
Seq-13 (1–157)			
121–141	(–)	tggCGGNtntggcagcaccgc	LAC9, homologous to GAL4
Seq-14 (1–199)			
22– 34	(+)	cgcTCGAacacgc	XBP1, stress-induced expression
109–125	(–)	ggcgggtacctCCGCcg	UAY, involved in purine utilization
110–126	(+)	ggcggaggtacCCGCca	UAY, involved in purine utilization
116–124	(+)	ggtaCCCGc	REB1, multifunctional transcription factor
145–157	(–)	gtcTCGAacatat	XBP1, stress-induced expression
168–188	(–)	AGAAgacacttgaatagggtt	HSF, heat shock factor

of the sequences (1–14) are rich in GC, and only two sequences (seq-10, 11) contain more AT than GC. Such composition is not typical of S / MAR sequences which are generally AT-rich, and this composition differs from mammalian DNA fragments tightly bound to proteins [10, 16].

The *Saccharomyces* genome is more compact compared with that of higher order eukaryotic organisms. In the *Saccharomyces* genome, no short interspersed nuclear elements (SINEs) or long interspersed nuclear elements (LINEs) were found [6]. We can expect that in the yeast genome other sequences with a different nucleotide composition can perform the same function as in the mammalian genome.

However, some other S / MAR properties are shared by yeast TBD sequences. It is known that DNA replication is associated with nuclear matrix attachment sites [15, 16]. The origins of replication motifs ATTA, ATTTA and ATTTTA are abundant in S / MARs [17] and were present in eight of the sequenced DNA clones (Table1).

A bipartite consensus sequence characteristic of S / MARs [17] is absent in all studied clones. But computer analysis using the MAR finder program (http://www.futuresoft.org/MAR-Wiz/) has revealed that ten from fourteen of yeast tightly bound DNA sequences (seq-1, 2, 3, 7, 8, 9, 10, 11, 12, 14) have some similarity with S / MAR sequences (underlined bold fragments in Figure). The MAR potential for these sequences exceeds 60%.

Also, these sequences have some other elements that are typical of S / MAR sequences. All S / MARs have been supposed to have numerous topoisomerase II cleavage and binding sites [21–23]. Yeast tightly bound DNA sequences are rich in topoisomerase II binding sites (Table 2). Topoisomerase II can lead to a relaxation of both positive and negative supercoils. These observations suggested the possibility that these sequences might be involved in quite diverse processes, all of them changing the topology of chromatin domains. It is of interest to note that a computer search revealed that five of the study sequences contain kinked DNA stretches (seq-9, 10, 11, 14) and Z DNA stretches (seq-11, 14). These data suggest also that DNAs presented in tightly bound DNA–protein complexes have a characteristic conformation instead of consencus sequences.

Computer analysis [24] has also revealed that cloned DNA sequences contain multiple motifs which are recognized by known transcription factors (Table 2). Most of these transcription factors are general regulator factors (GRFs) controlling the development processes and cycle of *Saccharomyces cerevisiae* cells.

In five of the sequences analyzed (seq-1, 2, 3, 9, 14) there were found binding sites of heat shock transcription factor (HSF) which is necessary for heat-induced transcription of not only HSP, but also of genes encoding proteins involved in diverse cellular processes such as protein degradation, detoxification, energy generation, carbohydrate metabolism, and maintenance of cell wall integrity. HSF is conserved from yeast to humans. The different stresses also turn on different subsets of heat shock proteins, suggesting that there is a specificity in regulating HSF activity [25].

Sequences 3, 12, 14 contain an Xbp1 binding site which shows homology to the DNA-binding domain of the *Saccharomyces cerevisiae* cell cycle regulating transcription factors Swi4 and Mbp1. Xbp1 acts as a negative regulator of transcription induced by stress [26]. Sequence 6 contains binding sites of the zinc finger transcription factor PacC homologous to the *Saccharomyces cerevisiae* Rim1p transcription factor, which is mediating the control of meiosis and invasiveness [27], and the zinc finger protein Mig1, which represses gene transcription during growth in glucose. The activity of Mig1 is regulated by phosphorylation and subcellular localization. In high glucose, Mig1 is dephosphorylated and located in the nucleus where it can repress transcription; upon glucose removal, Mig1 is rapidly phosphorylated and translocated into the cytoplasm [28].

The sequences 1–14 have also binding sites for the major nitrogen regulatory protein NIT2 (seq-1, 6, 7), a member of the GATA family of transcription factors [29], the HAP2/3/4/5 complex of *Saccharomyces cerevisiae* (seq-4, 10), which represents a family of CCAAT binding proteins and is a key player in a wide array of development processes [30], yeast transcription regulator MAPKKK Ste11 (seq-3), which is mainly involved in the protein–protein interactions of cell signaling and transcriptional repression [31].

In yeast tightly bound DNA sequences 1–14 there were also detected binding sites of multifunctional regulators Rap1p, Abf1p, Reb1p, Adr1p,Gcn4p, Gcr1 (Table 3, seq-1, 3, 5, 6, 7, 8, 10, 11, 14). These proteins cannot only activate or repress individual genes or gene groups, but they are also endowed with a potent insulating capacity [7]. Heterochromatic domains of the genome are usually transcriptionally repressed, while euchromatic regions are transcriptionally competent. It has been shown recently that yeast may also have evolved efficient mechanisms for insulating genes from each other [32]. Insulators are stretches of DNA located between the enhancer(s) and promoter or the silencer(s) and promoter of adjacent genes or clusters of adjacent genes [33]. Their function is to prevent a gene from being influenced by the activation (or repression) of its neighbours.

The DNA barrier elements that are able to block the propagation of transcriptional silencing in yeast are functionally similar to chromatin boundary / insulator elements in the metazoans that delimit func-

tional chromosomal domains. These barriers share no sequence homology but all consist of multiple binding sites for various regulatory proteins [33]. Upstream, the activating sequences of many highly expressed yeast ribosome protein genes and glycolytic genes exhibit barrier activity [34]. It has been shown that the transcription factors RAP1 and REB1 can be directly involved in barrier activity and their barrier function correlates with local changes in chromatin structure. The Reb1 protein is found associated with several proteins implicated in chromatin structure or remodeling, including Rsc2 and Rsc3, components of a histone deacetylase complex, and Eaf1, part of a complex with the histone acetyltransferase Esa1[35]. Autonomously replicating sequence (ARS) binding factor 1 (ABF1) and the repressor / activator protein 1 (RAP1) play an important role in transcriptional activation and repression, gene silencing, recombination and in telomere structure, and both are abundant and essential for cell growth. Both RAP1 and ABF1 can synergize with T-rich elements and both proteins can function as insulator elements to establish boundaries between regions of silent (heterochromatic) and permissive (euchromatic) chromosomal regions. Furthermore, direct observations show that both ABF1 and RAP1 can remodel chromatin near their binding sites [36].

Finding of transcription factors Rap1p, Abf1p, Reb1p, Adr1p,Gcn4p, GCR1 binding sites in most of yeast TBD sequences (Table 3) and recent data on a possible involvement of these transcription factors in genome barrier activity [7, 33–35] allow to propose that yeast DNA presented in tightly DNA-bound complexes may function as boundary / insulate elements. Multiple topoisomerase II binding sites revealed in these regions also agree with the assumption that yeast TBD sequences may participate in chromatin remodeling and gene activity regulation.

The structure and function of tightly bound DNA complexes are far from being quite clear. Nevertheless, the proposed model may constitute a starting point in designing further experiments aimed at a better understanding of the functions of tightly bound DNA complexes.

ACKNOWLEDGEMENT

This work was partly supported by the European Commission, Program of Centres of Excellence, project CEBIOLA, ICA1-CT-2000-70027.

Received 7 October 2005
Accepted 14 December 2005

References

1. Paulson JR, Laemmli UK. Cell 1977; 12: 817–28.
2. Jenke AC, Stehle IM, Herrmann F, Eisenberger T, Baiker A, Bode J, Fackelmayer F, Lipps H J. PNAS 2004; 101: 11322–7.
3. Zaidi SK, Javed A, Choi J, van Wijnen J, Stein JL, Lian JB, Stein GS. J. of Cell Science 2001; 114: 3093–102.
4. Jackson DA. Chromosome Res 2003; 11: 387–401.
5. Ehrenhofer-Murray AE. Eur J Biochem 2004; 271: 2335–49.
6. Sandmeyer S. Genome Res 1998; 8: 416–8.
7. Qiu Q, Foland T, Griesen D, Galloway C, Chiu Y, Sandmeier J, Broach J, Bi X. Nucleic Acids Res 2003; 31: 224–1233.
8. Tsanev R, Avramova Z. Chromosoma 1994; 103: 293–301.
9. Rothbarth K, Spies E, Juodka B, Yavuzer U, Nehlis S, Stammer H, Werner D. J Cell science 1999; 112: 2223–32.
10. Sjakste N. Proc Latv Acad Sci B 1997; 51: 61–3.
11. Neuer-Nitche B, Lu X, Werner D. Nucleic Acids Res 1988; 16: 8351–60.
12. Pfutz M, Gileadi O, Werner D. Chromosoma 1992; 101: 607–17.
13. Razin S, Chernokhvostov V, Vazetzky E. Nucleic Acids Res 1988; 16: 3617–33.
14. Borutinskaite V, Bagdoniene L, Labeikyte D, Juodka B. Biologija 2004; 2: 11–4.
15. Current Protocols in Molecular Biology. John Wiley & Sons, Inc. 1994.
16. Neuer-Nitsche B, Lu X, Werner D. Nucleic Acids Res 1988; 16: 8351–60.
17. Avramova Z, Georgiev O, Tsanev R. Cell Biol 1994; 13: 539–48.
18. Razin S. Critical Reviews in Eukaryotic Gene Expression 1996; 6: 247–69.
19. Dickinson L, Dickinson CD, Kohwi-Shigematsu T. JBC 1997; 272: 11463–70.
20. Lamond AI, Earnshaw WC. Science 1998; 280: 547–53.
21. Frisch M, Frech K, Klingenhoff A, Cartharius K, Liebich I, Werner T. Genome Res 2002; 12: 349–54.
22. Liebich I, Bode J, Reuter I, Wingender E. Nucleic Acids Res 2002; 30: 3433–42.
23. Jack RS, Eggert H. Eur J Biochem 1992; 209: 503–9.
24. Cartharius K, Frech K, Grote K, Klocke B, Haltmeier M, Klingenhoff A, Frisch M, Bayerlein M, Werner T. Bioinformatics 2005; 21: 2933–42.
25. Yamamoto A, Mizukami Y, Sakurai H. J Biol Chem 2005; 280: 11911–9.
26. Mai B, Breeden L. Mol Cell Biol 1997; 17: 6491–501.
27. Denison S, Negrete-Urtasun S, Mingot J, Tilburn J, Mayer W, Goel A, Espeso E, Peñalva M, Arst H. Mol Microbiol 1998; 30: 259–63.
28. Ahuatzi D, Herrerd P, de la Cera T, Moreno F. J Biol Chem 2004; 279: 14440–6.
29. Feng, B, Marzluf G. Mol Cell Biol 1998; 18: 3983–90.
30. Chen H, Crabb J, Kinsey A. J Genetics 1998; 148: 123–30.

31. Bhattacharjya S, Xu P, Chakrapani M, Johnston L, Ni F. Protein Science 2005; 14: 828–35.
32. Kimura A, Horikoshi M. Genes to Cells 2004; 9: 499–508.
33. Bi X, Broach JR. Current Opinion in Genetics and Development 2001; 11: 199–204.
34. Fernandez LA, Winkler M, Grosschedi R. Mol Cell Biol 2001; 21: 196–208.
35. Oki M, Valenzuela L, Chiba T, Ito T, Kamakaka R T. Mol Cell Biol 2004; 24: 1956–67.
36. Yarragudi A, Miyake T, Li R, Morse RH. Mol Cell Biol 2004; 24: 9152–64.

L. Bagdonienë, K. Bonikataitë , V. Borutinskaitë, D. Labeikytë, B. Juodka

SACHAROMYCES CEREVISIAE DNR BALTYMINIØ KOMPLEKSØ DNR SEKØ SAVYBIØ TYRIMAS

S a n t r a u k a

Tiriant *Sacharomyces cerevisiae* mieliø làsteliø tvirtø DNR baltyminiø kompleksø DNR sekas (1–14) nustatyta pastarøjø heterogeninë sudëtis. Tiesioginës homologijos tarp jø nerasta. DNR fragmentø lokalizacijos vietos genome yra ávairialypës, tarp jø nëra centromeriniø ir telomeriniø srièiø. Pagal *S. cerevisiae* mieliø làsteliø DNR fragmentø sudëtyje aptinkamus motyvus (trumpi AT pasikartojimai, ori, topoizomerazës II, transkripcijos faktoriø prisijungimo vietos) ðias sekas galima priskirti savitai MAR / SAR sekø grupei. Manome, kad tvirti mieliø DNR baltyminai kompleksai gali dalyvauti pertvarkant chromatino struktûrà ir atskiriant aktyviai ekspresuojamø genø sritis nuo neaktyviø genø.