The EMBO Journal vol.11 no.10 pp.3787-3796, 1992

Zuotin, a putative Z-DNA binding protein in *Saccharomyces cerevisiae*

Shuguang Zhang, Curtis Lockshin, Alan Herbert, Edward Winter[1] and Alexander Rich

Department of Biology 16-739, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[1]Present address: Department of Biochemistry and Molecular Biology, Thomas Jefferson University, 233 South 10th Street, Philadelphia, PA 19107-6799, USA

Communicated by A.Rich

A putative Z-DNA binding protein, named zuotin, was purified from a yeast nuclear extract by means of a Z-DNA binding assay using [^{32}P]poly(dG-m^{5}dC) and [^{32}P]oligo(dG-Br^{5}dC)$_{22}$ in the presence of B-DNA competitor. Poly(dG-Br^{5}dC) in the Z-form competed well for the binding of a zuotin containing fraction, but salmon sperm DNA, poly(dG-dC) and poly(dA-dT) were not effective. Negatively supercoiled plasmid pUC19 did not compete, whereas an otherwise identical plasmid pUC19(CG), which contained a (dG-dC)$_7$ segment in the Z-form was an excellent competitor. A Southwestern blot using [^{32}P]poly(dG-m^{5}dC) as a probe in the presence of $MgCl_2$ identified a protein having a molecular weight of 51 kDa. The 51 kDa zuotin was partially sequenced at the N-terminal and the gene, *ZUO1*, was cloned, sequenced and expressed in *Escherichia coli*; the expressed zuotin showed similar Z-DNA binding activity, but with lower affinity than zuotin that had been partially purified from yeast. Zuotin was deduced to have a number of potential phosphorylation sites including two CDC28 (homologous to the human and *Schizosaccharomyces pombe* cdc2) phosphorylation sites. The hexapeptide motif KYHPDK was found in zuotin as well as in several yeast proteins, DnaJ of *E.coli*, csp29 and csp32 proteins of *Drosophila* and the small t and large T antigens of the polyoma virus. A 60 amino acid segment of zuotin has similarity to several histone H1 sequences. Disruption of *ZUO1* in yeast resulted in a slow growth phenotype.

Key words: DNA binding protein/DnaJ similarity/histone H1 similarity/KYHPDK/Z-DNA.

Introduction

DNA is capable of undergoing a number of conformational changes; the most dramatic of these is from right-handed B-DNA to left-handed Z-DNA. There are several conditions that are known to stabilize Z-DNA. For example, poly(dG-m^{5}dC) converts readily to left-handed Z-DNA *in vitro* in the presence of millimolar concentrations of divalent metals and polyamines, as well as small peptides (Behe and Felsenfeld, 1981; Rich *et al.*, 1984; Takeuchi *et al.*, 1991). Certain DNA sequences, especially alternating purines and pyrimidines can adopt the Z-conformation in response to negative supercoiling (Peck *et al.*, 1982). Inside the cell, negative supercoiling can be generated during transcription (Liu and Wang, 1987; Tsao *et al.*, 1989). Furthermore, the equilibrium between B- and Z-DNA can be influenced by proteins that preferentially bind one of the two conformations (Lafer *et al.*, 1985).

A number of studies suggest that Z-DNA may exist *in vivo* (Jaworski *et al.*, 1987; Rahmouni and Wells, 1989; Wittig *et al.*, 1989), however, the extent of its occurrence is yet to be determined. Z-DNA has been implicated in some important biological processes, such as general DNA recombination (Bullock and Botchan, 1986; Treco and Arnheim, 1986; Blaho and Wells, 1987; Wahls *et al.*, 1990), and both positive and negative transcriptional regulation (Nordheim and Rich, 1983; Naylor and Clark, 1990). There are many approaches that can be employed to further our understanding of the biological significance of Z-DNA *in vivo*. One approach is to purify a protein that has an enhanced affinity for the left-handed DNA and to characterize this protein *in vitro*.

We have been interested in determining whether left-handed Z-DNA binding proteins occur in the yeast *Saccharomyces cerevisiae*, in part, because of the genetic methods available. We developed a gel shift assay for detecting Z-DNA binding proteins and used it to identify a putative Z-DNA binding protein in yeast extracts. In this work, we report the partial purification of the putative Z-DNA binding protein from yeast *S.cerevisiae*, the cloning of its gene and partial functional analysis.

Results

Detection of a poly(dG-m^{5}dC) binding protein in S.cerevisiae

Two probes that can be stabilized in the Z-form were used to detect potential Z-DNA binding proteins. One is an ~600 bp fragment of ^{32}P-labelled poly(dG-m^{5}dC) that is stabilized in the Z-DNA form by millimolar concentrations of $MgCl_2$ (Behe and Felsenfeld, 1981); the other is an oligonucleotide [^{32}P](dG-Br^{5}dC)$_{22}$ (A.Herbert and A.Rich, in preparation) that can be stabilized by millimolar concentration of $MgCl_2$ or μM concentration of $Co(NH_3)_6^{3+}$. When phosphocellulose column fractions of yeast nuclear extracts were assayed using either of these ^{32}P-labelled DNA fragments, a distinctive band shift was detected (Figure 1). Both whole yeast cell extracts and a yeast nuclear extract produced a similar band shift. Fractions from a phosphocellulose column that did not bind to B-form DNA (Winter and Varshavsky, 1989) showed significant gel retardation when using [^{32}P]poly(dG-m^{5}dC) in the Z-DNA form even in the presence of a 400-fold molar excess of sheared salmon sperm DNA (Figure 1, fractions 12 and 13). The band shift resulting from a polyclonal anti Z-DNA antibody is also

© Oxford University Press

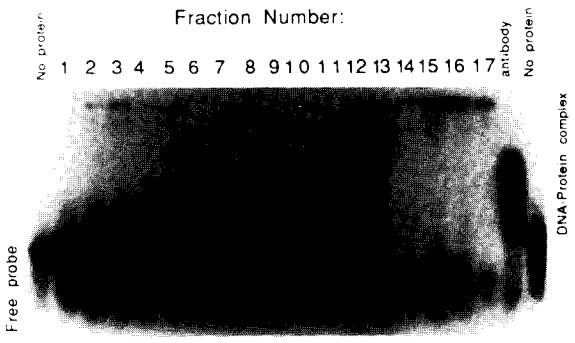

Fig. 1. Gel retardation assay. Yeast nuclear extract was fractionated using a phosphocellulose column and fractions were assayed using a 600 bp fragment of ^{32}P-poly(dG-m5dC) as a probe in the presence of 10 mM $MgCl_2$ and a 400-fold excess of sheared salmon sperm B-DNA. Under these conditions, the polymer assumes the Z-DNA conformation. A Z-DNA specific antibody (Ab) was used as positive control. 1 μl of each fraction (numbers indicated) obtained by salt elution from 0.2–0.5 M potassium phosphate (pH 7.4) were added to the assays. Fractions 12 and 13 showed a positive retarded DNA–protein complex.

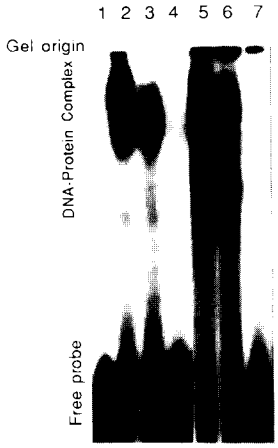

Fig. 2. Gel retardation competition assay of (dG-Br^{5}dC)$_{22}$ binding. [^{32}P](dG-Br^{5}dC)$_{22}$ was incubated in the presence of a 2000-fold excess of sheared salmon sperm DNA with additions as follows: lane 1, no addition; lane 2, a monoclonal anti-Z-DNA antibody (mAb); lane 3, mAb plus additional 50 ng of negatively supercoiled plasmid pUC19 (without a Z-DNA insert); lane 4, mAb plus 25 ng negatively supercoiled pUC19(CG) (containing a Z-DNA segment); lane 5, fraction FI yeast protein; lane 6, FI plus additional 50 ng pUC19; lane 7, FI plus pUC19(GC). The plasmid pUC19(GC) was assayed for its resistance to restriction endonuclease *Bss*HII digestion to confirm the presence of the Z-DNA prior to the assay.

shown in Figure 1. The pooled fractions (FI) also showed binding activity to the oligonucleotide probe, [^{32}-P](dG-Br^{5}dC)$_{22}$.

In order to determine if these band shifts were the result of authentic Z-DNA binding, negatively supercoiled plasmids pUC19 and pUC19(GC) were used as competitor DNAs. pUC19(GC) (B.Johnston, unpublished) contains a 14 bp [(dG-dC)$_7$] insert that can adopt the Z-conformation upon negative supercoiling (Vardimon and Rich, 1983; Azorin *et al.*, 1984; data not shown). Monoclonal anti Z-DNA antibody (Möller *et al.*, 1982) used as a positive control showed specific complex formation in the presence of competing plasmid pUC19, but not when the supercoiled plasmid pUC19(GC) containing Z-DNA was used (Figure 2). Similar binding specificity was observed when a partially purified yeast fraction (FI) were used instead of the anti-Z-DNA antibody. However, the complex observed with the protein fraction was more heterogeneous than that seen with the antibody (Figure 2). The Z-binding activity of fraction FI was further purified using affinity chromatography to a poly(dG-m^{5}dC)–agarose column (FII) followed by Superose 12 (FIII) and Mono-S chromatography. The resultant active fraction (FIV) included a prominent 51 kDa protein but was still quite complex.

Identification of zuotin by Southwestern blotting

In order to identify the specific protein that interacts with the Z-DNA probe, a Southwestern blot was employed. Proteins were transferred from a SDS–polyacrylamide gel to an Immobulon P membrane and exposed to conditions that favor renaturation. Subsequently, the filter was incubated in the presence of [^{32}P]poly(dG-m^{5}dC) in the Z-form with a 300-fold excess of B-DNA. The probe bound a single polypeptide of ~51 kDa (Figure 3B) that spanned fractions 12 and 13, both of which were active in the band shift assay. There was also a weak signal in fraction 14. Although the fractions were quite complex, only the 51 kDa protein was detected by autoradiography (Figure 3A). We have named this putative Z-DNA binding protein as zuotin (from the Chinese, zuo, meaning left).

Purification of zuotin and cloning of ZUO1

Approximately 5 μg of zuotin was gel purified for amino acid composition analysis and N-terminal sequencing. The amino acid composition of zuotin is shown in Table I. N-terminal sequencing yielded the following: MFSLPTL-TSDI(E/D)V[EV](N)(H/S)(D) where [ ] and ( ) indicate moderate and low confidence assignments respectively.

A degenerate 32mer oligonucleotide was designed using 'reverse translation' (Figure 4A) and a yeast genomic Southern hybridization analysis revealed a single hybridizing 2.4 kb *Hin*dIII fragment (data not shown). A yeast genomic EMBL3A library was subsequently screened and 14 clones isolated. Restriction mapping and Southern hybridization using the oligonucleotides revealed that one of the isolates contained a 2.4 kb *Hin*dIII hybridizing fragment (data not shown). This *Hin*dIII fragment was subcloned and the nucleotide sequence determined. The DNA sequence revealed an open reading frame (ORF), whose translated N-terminal sequence corresponded exactly to that of the N-terminal sequence determined from purified zuotin. To obtain the entire coding sequence of the ORF, a 3.1 kb *Bam*HI–*Eco*RI fragment was subcloned into pBluescript vector and the nucleotide sequence was determined.

ZUO1 encodes a 433 residue protein

The 3.1 kb *Bam*HI–*Eco*RI fragment contains the entire zuotin coding region (1291–2590), as well as 5′ and 3′ non-transcribed regions (Figure 5). There is a long ORF encoding a 433 amino acid protein that corresponds to a calculated molecular weight of 49 020 Da. A second ORF in the same orientation within *ZUO1* is able to encode a 168 amino acid polypeptide. It remains to be seen if there is a translated product from this reading frame. The 5′ region of the 3.1 Kb fragment also has another ORF containing 210 codons. The translated product shares extensive homology with the C-terminal part of the *E.coli BioB* gene, one of the

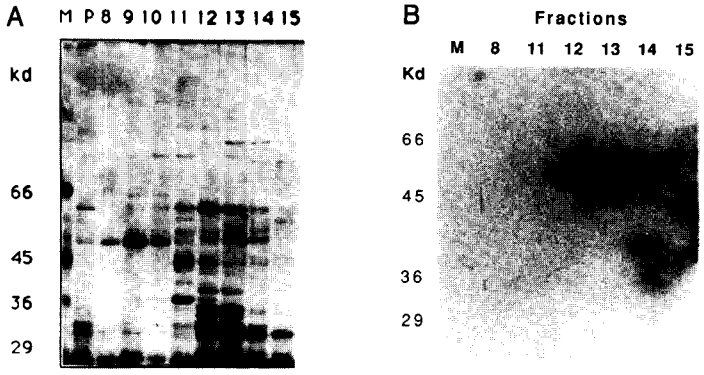

Fig. 3. Southwestern blot analysis of mono-S column fractions using [^{32}P]poly(dG-m^{5}dC) in 15 mM $MgCl_2$ as a probe. Protein fractions were resolved on a polyacrylamide–SDS gel and blotted onto a filter (see Materials and methods). The silver stained mono-S column protein fractions are shown in **(A)**. The filter was incubated with ^{32}P- poly(dG-m^{5}dC) DNA in the presence of 300-fold excess of sheared salmon sperm DNA. The autoradiogram is shown in **(B)**.

biotin synthetases. It has a homology with >50% identity and 75% similarity. The 5′ non-transcribed region contains three A/T-rich segments and two alternating AT segments that may act as regulatory domains for transcription of the gene. There is also a purine-rich tract in the coding region that could adopt a DNA conformation different from conventional B-DNA (McCarthy and Heywood, 1987). The coding region comprises 1299 base pairs and the transcript of *ZUO1* is ~1.7 kb (data not shown). *ZUO1* is localized on yeast chromosome VII near *ADE3* (data not shown).

Zuotin has several interesting features: it consists of 13% alanine, 20.6% positively charged (lysine, arginine and histidine) and 18.5% negatively charged (aspartic and glutamic acids) amino acids and has a pI 8.8. (Table I). The charged amino acid residues are clustered at the C-terminal end and there is one segment with 12 charged residues in a row (Figure 5). There are two continuous perfect and one imperfect octad tandem repeats of alternating alanine and charged amino acids (lysine and glutamic acid) in the alanine/lysine and arginine rich region. It also contains several potential phosphorylation sites including sequences recognized by protein kinase C, casein kinase II, cAMP-dependent protein kinase and tyrosine kinase as predicted by Prosite (Figure 5) (Bairoch, 1991). There are also two potential CDC28 (or cdc2) phosphorylation sites (KTPFVRR from 21–27 and KTPIP from 201–205) (Moreno and Nurse, 1990). It has a bipartite nuclear targeting sequence *KKKAK*EAAKAA*KKKNKR* from 340–356 (Robbins *et al.*, 1991). There are several regions that are predicted to form an α-helix and this includes the repeated octad segment (Chou and Fasman, 1978). However, when a 16 residue peptide of the repeated segment was synthesized and examined by circular dichroism, a distinctive β-sheet structure was observed (Zhang,S., Holmes,T., Lockshin,C. and Rich,A., in preparation).

Zuotin shares sequence similarities with DnaJ protein, which is involved in DNA replication of bacteriophages λ and P1 (Liberek *et al.*, 1988). There are also similarities with several other yeast proteins: YDJ1 (Caplan and Douglas, 1991), SIS1 (Luke *et al.*, 1991), SCJ1 (Blumberg and Silver, 1991) and NPL1/SEC63 (Sadler *et al.*, 1989), and they all include the hexapeptide motif KYHPDK (Figure 6A). This hexapeptide is also present in both the small t and large T antigens of avian budgerigar fledgling disease virus (Rott *et al.*, 1988) and csp29 and csp32 proteins expressed in retina and brain in *Drosophila* (Zinsmaier *et al.*, 1990). It will be of interest to determine if this hexapeptide mediates a common function in this group of proteins. There are similarities between another region of zuotin and several histone H1 variants from different sources. The region of similarity lies in the alanine/lysine/arginine rich domain. For example, the sequences from 304–363 in zuotin and 146–205 in sea urchin histone H1 are 64% similar and 46% identical (Figure 6B). Similarities of zuotin with other histone H1 variants also occur in the same region (Figure 6B).

Table I. Amino acid composition analysis of zuotin

	Hydrolysis		Deduced from DNA (%)
	E.coli (%)	Yeast	
Arg	5.7	7.6	6.0
Lys	9.2	11.9	12.9
His	3.1	1.4	1.6
Glu	(11.5)	(12.1)	9.7
Gln			2.3
Asp	(11.0)	(11.5)	8.8
Asn			3.2
Ser	7.0	7.4	6.9
Thr	4.4	5.0	5.1
Tyr	2.4	1.8	2.1
Ala	11.2	13.0	12.9
Val	6.3	4.8	5.5
Leu	6.8	6.6	5.8
Ile	3.6	3.0	3.2
Phe	3.8	4.5	5.1
Trp	0.0	0.0	1.4
Pro	4.0	3.2	3.5
Gly	5.0	6.1	3.0
Met	0.7	0.1	0.7
Cys	0.0	0.0	0.2

The composition of hydrolyzed zuotin is obtained from the purified yeast protein and from the *E.coli* expressed zuotin. Both were gel purified and subjected to HCl hydrolysis, then analyzed by HPLC with internal controls. The deduced composition of zuotin is derived from the DNA sequence of the ORF of *ZUO1*. The numbers in parenthesis are combined percentages of Glu/Gln, and Asp/Asn as they are individually indistinguishable in the hydrolytic analysis. The amino acid composition of zuotin was 20.5% (Arg, Lys and His), 18.5% (Glu and Asp), 14.1% (Thr, Ser and Tyr) and 68.16 (Arg, Lys, Asp, Glu, Ser, Thr, Ala and Leu).

Construction and analysis of zuo1 mutants

In order to generate an interrupted *zuo1* allele, the 1.2 kb *Hin*dIII fragment containing the *S.cerevisiae URA3* gene was inserted at a unique *Hin*dIII site of the *ZUO1* coding region (Figure 7A). The plasmid pZUO1::URA3 was then linearized and used to transform DM27, a diploid *ura3* yeast strain. Diploid Ura$^+$ transformants, expected to be heterozygous at the *ZUO1* locus, were selected and confirmed to harbor disruption at the *ZUO1* locus. The heterozygous diploid strains were subsequently sporulated and subjected to tetrad analysis. Tetrads yielded four viable colonies, two large and two small in which Ura$^+$ phenotypes co-segregated with the small colonies (i.e. a slow growth phenotype) (Figure 7C). Southern blot analysis using DNA from four tetrads revealed that all clones with the slow growth phenotype harbor the 1.2 kb insertion (Figure 7B). The results showed that the insertion of *URA3* at the *ZUO1* locus produces a similar slow growth phenotype.

Expression of ZUO1 in E.coli

In order to verify that the cloned *S.cerevisiae ZUO1* gene encodes the putative Z-DNA binding protein, *ZUO1* was expressed in *E.coli* using a T7 expression system (Studier *et al.*, 1990). *ZUO1* was cloned in pET8c at the unique *Nco*I and *Bam*HI sites. When the *lac*UV5 promoter was induced with IPTG, a protein band with an apparent molecular weight of ~51 kDa was detected. This protein was not seen in the cell extract with the pET8c plasmid induced by IPTG nor in the cell extract with pETZUO but without IPTG induction (data not shown). Analysis of the protein composition and the N-terminal sequence of the purified and the expressed zuotin are essentially the same (Table I and Figure 4B), thus indicating that zuotin was expressed correctly.

In attempting to purify zuotin expressed in *E.coli*, we found that the exogenously expressed zuotin was sequestered in inclusion bodies. However, enough material was in solution so that a crude cell extract and a partially purified preparation was used in a band shift assay using [^{32}P]poly(dG-m^{5}dC) in the Z-form as the probe. Stepwise elution of *E.coli* expressed zuotin from phosphocellulose yielded a fraction with the band shift shown in Figure 8. A strong shifted band bound to labelled poly(dG-m^{5}dC) even in the presence of a 40-fold excess of salmon sperm DNA. On adding a 40-fold excess of poly(dG-dC) that can form Z-DNA under certain conditions, a somewhat weaker band was visible. Similar results were found with the yeast zuotin. It is interesting that the band shift with zuotin from yeast migrated slightly further towards the positive side of the gel (Figure 8, lane 6) than the *E.coli* expressed zuotin (Figure 8, lane 5). It is possible that this difference could be due to differences in phophorylation or other post-translational modifications that were not carried out in *E.coli*. Furthermore, the yeast zuotin binds Z-DNA more tightly than the *E.coli* expressed zuotin (Figure 8). It is known, for example, that phosphorylation modifies the DNA binding activity of the yeast centromere binding protein CBF3 (Lechner and Carbon, 1991). Analysis of the zuotin sequence using the Prosite computer program (Bairoch, 1991) suggests that zuotin may be phosphorylated. Attempts are now being made to express zuotin in other systems so that it can be more fully characterized.

A

```
N-met-phe-ser-leu-pro-thr-leu-thr-ser-asp-ile...

5'-ATG-TTT-TCT-TTG-CCA-ACT-TTG-ACT-TCT-GAT-AT-3'
            C       T   C       C       C
```

B

N-met-*val*-ser-leu-pro-thr-leu-thr-ser-asp-ile...

C

```
    C CAA AGT GAA TGA TAT GGG GCT AGA AAC GTG TGT TAC TTT AGG TAT GGT TGA TCA AGA TCA AGC AAA GCA ATT GAA AGA TGC AGG   85

TTT GAC TGC ATA CAA CCA TAA CAT CGA CAC TTC CAG AGA ACA CTA TAG TAA GGT CAT CAC CAC GAG AAC CTA CGA CGA CAG GTT ACA GAC  175

CAT CAA GAA TGT CCA AGA ATC TGG AAT AAA AGC CTG TAC CGG TGG TAT TTT GGG TCT CGG TGA AAG CGA AGA CGA CCA TAT AGG ATT CAT  265

CTA CAC ATT ATC CAA TAT GTC TCC TCA TCC TGA GTC CCT ACC AAT TAA TAG ACT AGT TGC TAT CAA AGG GAC TCC AAT GGC TGA GGA ACT  355

TGC CGA TCC AAA GAG TAA AAA GTT GCA ATT CGA CGA AAT TTT GAG AAC CAT TGC CAC AGC GAG AAT AGT TAT GCC AAA GGC CAT TAT AAG  445

ACT TGC CGC TGG TCG TTA TAC AAT GAA AGA AAC AGA GCA ATT TGT CTG TTT CAT GGC AGG TTG TAA CAG TAT CTT CAC CGG TAA GAA AAT  535

GCT GAC GAC AAT ATA TAA CGG TTG GGA CGA AGA CAA GGC AAT GTT GGC TAA ATG GGG ATT GCA ACC TAT GGA GGC ATT TAA GTA CGA CAG  625

ATC TTG AAG ATA GGG ATA TGT GGA TAA TTC TAC GAT TCT AAC TGT ACA TTT CTC CCT TAT TTA TTA AGA AAA CCT ATA TAT ATA TAT ATT  715
                                                                                                    (AT)7
TAC CTA TTT ATT CTG CCA TCG TTA GCT GGC GTT TTA TCT TTT ATG CAT CCA ATA TCT AAT ATT ACT TCC GAT CAC GCA TTT AGT TCT GAT  805

TAC AGC AGA AAT CGT AGC GCG ATG AGA CAT TTC ATC AAA TGG CCT TTT TTT TTT GGG CAA TTT TTT TAT ATC TTG AAA TGA TAG TTG CCT  895
                                                          T10
TGT ACT TTC AAC CGT TCA TTT CAT TAA GAA CTT GAC TAA ATA TGA ACA TTT CTT AAA AAA AAA GGT TGA CAT ATA AAA ATA ATC GAA TAT  985
                                                                          A9
AAA CGA TGG AAT TTT TAT AAA ATT AAA CAC ATA TAT ATA TAT ATA TTA ACT ATA AAT ATG TCA AAG AAA CCA TAC AAT CAT AGA TTT ATA 1075
                                        (AT)8
ACT ATC TTT TGG ATG ACA TTA ATG AAC ATA ACG CTC CTA ATA CAA ATG TCA AAA AAT ATT ACC CGC AAA TAC GAA TCT TTT TTT TTT CTC 1165
                                                                                                    T10
GAT GAA ATT TTG CAA AGA GTT CGA AAT TTT TAT TTC AAG AGC TGG TAG AGA AAA TTT CAT AAG GTT TTC CTA CCG ATG CTT TTA TAA AAT 1255

                                                Met Phe Ser Leu Pro Thr Leu Thr Ser Asp Ile Thr Val Glu Val Asn Ser Ser    18
CTT CGT TTT GTC TCA CAT ATA CCA ACA AGA GTA ACG ATG TTT TCT TTA CCT ACC CTA ACC TCA GAC ATC ACT GTT GAA GTC AAC AGT TCC 1345

Ala Thr Lys Thr Pro Phe Val Arg Arg Pro Val Glu Pro Val Gly Lys Phe Phe Leu Gln His Ala Gln Arg Thr Leu Arg Asn His Thr    48
GCT ACC AAA ACC CCA TTC GTC CGT CGT CCG GTC GAA CCG GTT GGT AAG TTC TTT TTG CAA CAT GCT CAA AGA ACT TTG AGA AAC CAC ACC 1435

Trp Ser Glu Phe Glu Arg Ile Glu Ala Glu Lys Asn Val Lys Thr Val Asp Glu Ser Asn Val Asp Pro Asp Glu Leu Leu Phe Asp Thr    78
TGG TCT GAA TTT GAA AGA ATT GAA GCT GAA AAG AAC GTC AAA ACC GTT GAT GAA TCC AAT GTC GAC CCA GAT GAG TTG TTA TTC GAC ACT 1525
```

```
Glu Leu Ala Asp Glu Asp Leu Leu Thr His Asp Ala Arg Asp Trp Lys Thr Ala Asp Leu Tyr Ala Ala Met Gly Leu Ser Lys Leu Arg   108
GAA TTG GCC GAT GAA GAT TTA CTG ACT CAT GAT GCT AGA GAC TGG AAA ACT GCC GAT TTG TAT GCT GCT ATG GGT TTG TCT AAG TTG CGT 1615

Phe Arg Ala Thr Glu Ser Gln Ile Ile Lys Ala His Arg Lys Gln Val Val Lys Tyr His Pro Asp Lys Gln Ser Ala Ala Gly Gly Ser   138
TTC AGA GCT ACT GAA AGT CAA ATC ATC AAG GCT CAC AGA AAA CAA GTT GTC AAG TAC CAT CCA GAC AAG CAA TCT GCT GCT GGT GGT AGT 1705

Leu Asp Gln Asp Gly Phe Phe Lys Ile Ile Gln Lys Ala Phe Glu Thr Leu Thr Asp Ser Asn Lys Arg Ala Gln Tyr Asp Ser Cys Asp   168
TTG GAC CAA GAT GGC TTT TTC AAG ATT ATT CAA AAG GCC TTT GAA ACT TTG ACT GAT TCC AAC AAG AGA GCT CAG TAC GAC TCA TGT GAT 1795

Phe Val Ala Asp Val Pro Pro Pro Lys Lys Gly Thr Asp Tyr Asp Phe Tyr Glu Ala Trp Gly Pro Val Phe Glu Ala Glu Ala Arg Phe   198
TTT GTT GCC GAT GTT CCT CCT CCA AAG AAG GGT ACC GAT TAT GAC TTT TAT GAA GCT TGG GGC CCC GTT TTC GAA GCT GAA GCT CGT TTT 1885

Ser Lys Lys Thr Pro Ile Pro Ser Leu Gly Asn Lys Asp Ser Ser Lys Lys Glu Val Glu Gln Phe Tyr Ala Phe Trp His Arg Phe Asp   228
TCT AAG AAG ACT CCT ATT CCT TCT CTA GGT AAC AAA GAT TCT TCC AAG AAG GAA GTT GAA CAA TTC TAT GCT TTC TGG CAC AGA TTT GAC 1975

Ser Trp Arg Thr Phe Glu Phe Leu Asp Glu Asp Val Pro Asp Asp Ser Ser Asn Arg Asp His Lys Arg Tyr Ile Glu Arg Lys Asn Lys   258
TCC TGG AGA ACC TTT GAG TTC TTG GAC GAA GAT GTC CCA GAT GAC TCT TCT AAC AGA GAC CAC AAG CGT TAC ATT GAA AGA AAG AAC AAG 2065

Ala Ala Arg Asp Lys Lys Lys Thr Ala Asp Asn Ala Arg Leu Val Lys Leu Val Glu Arg Ala Val Ser Glu Asp Pro Arg Ile Lys Met   288
GCC GCA AGA GAC AAG AAG AAG ACT GCT GAT AAC GCT AGA TTG GTC AAA CTT GTT GAA AGA GCT GTC AGT GAA GAT CCC CGT ATC AAA ATG 2155

Phe Lys Glu Glu Glu Lys Lys Glu Lys Glu Arg Arg Lys Trp Glu Arg Glu Ala Gly Ala Arg Ala Glu Ala Glu Ala Lys Ala Lys Ala   318
TTC AAA GAA GAA GAG AAG AAG GAA AAG GAA AGA AGA AAA TGG GAA AGA GAA GCC GGT GCC AGA GCT GAA GCT GAA GCT AAG GCC AAG GCC 2245
                    Homopurine tract (49)         *
Glu Ala Glu Ala Lys Ala Lys Ala Glu Ser Glu Ala Lys Ala Asn Ala Ser Ala Lys Ala Asp Lys Lys Lys Ala Lys Glu Ala Ala Lys   348
GAA GCT GAA GCG AAG GCT AAA GCT GAA TCT GAA GCC AAG GCT AAC GCC TCC GCA AAA GCT GAC AAA AAG AAG GCT AAG GAA GCT GCT AAG 2335

Ala Ala Lys Lys Lys Asn Lys Arg Ala Ile Arg Asn Ser Ala Lys Glu Ala Asp Tyr Phe Gly Asp Ala Asp Lys Ala Thr Thr Ile Asp   378
GCC GCC AAG AAA AAG AAC AAG AGA GCC ATC CGT AAC TCT GCT AAG GAA GCT GAC TAC TTT GGT GAT GCT GAC AAG GCC ACC ACG ATT GAC 2425

Glu Gln Val Gly Leu Ile Val Asp Ser Leu Asn Asp Glu Glu Leu Val Ser Thr Ala Asp Lys Ile Lys Ala Asn Ala Ala Gly Ala Lys   408
GAA CAA GTT GGT TTG ATC GTT GAC AGT TTG AAT GAC GAA GAG TTA GTG TCC ACC GCC GAT AAG ATC AAG GCC AAT GCT GCT GGT GCC AAG 2515

Glu Val Leu Lys Glu Ser Ala Lys Thr Ile Val Asp Ser Gly Lys Leu Pro Ser Ser Leu Leu Ser Tyr Phe Val TER                 433
GAA GTT TTG AAG GAA TCT GCA AAG ACT ATT GTC GAT TCT GGC AAA CTA CCA TCC AGC TTG TTG TCC TAC TTC GTG TGA ATA CCG TAA GAA 2605

ATG GAA TAG AAT ATA TAC GAA TGT ATA CGA ATA TTA TAG AGA ACG TTC TCT TTT ATT TCT ATA ATG AAT AGG TTC GGG TAA CGG TTC CCT 2695

TTT TAG GTA TTT CTA GAA GAT GAG AGA AGA GGG AAT AAT GAG AAA GGC GAA AAA TAA AGA CAC CTT TAA CGA AAG ATC AAA GGT GTC CTT 2785
                                                        Poly (A)
ATT TAC TTA CAA TAG CTG CAA TTA GTA CGA CTC AAA AAA AGT GAA AAC AAA ACT GAA AGG ATA GAT CAA TGT CTT ACA GAG GAC CTA TTG 2875

GAA ATT TTG GCG GAT AGC CAA TGT CAT CAT CGC TTG GAC CAT ACT CTG GCG GTG CAC AAT TCC GAT CAA ACC AGA ACC AAT CCA CTT CTG 2965

GCA TCT TAA AGC AAT GGA AGC ATT CTT TTG AAA AGT TTG CCT CCA GAA TTG AGG GGC TCA CTG ACA ATG CAG TTG TTT ATA AAT TGA AGC 3055

CTT ACA TTC CAA GTT TGT CAA GAT TTT T   3083
```

Fig. 4. **(A)** Sequence of synthetic oligonucleotides used to isolate *ZUO1*. The first 11 amino acids of N-terminal sequence of zuotin are shown along with the reverse translation of the sequence. Alternative nucleotides were introduced at five positions. Nucleotides which were later determined to be mismatched are underlined. **(B)** N-terminal sequence of *E.coli* expressed zuotin. Val differs from Phe in the second codon of ZUO1 due to the introduction of a *NcoI* site for cloning and expression. **(C)** A 3.1 kb EcoRI−BamHI fragment from both strands were sequenced with the dideoxy chain termination method. The 5′ nontranscribed region containing alternating $(AT)_n$ rich and (A/T) regions that could serve as regulatory sites are underlined, ... and ____, respectively. There is also a homopurine/pyrimidine tract with only one exception (====) in the coding region which can adopt an alternative DNA conformation. The potential polyadenylation site at the 3′ end of the gene is indicated. The sequence data reported here are available from the EMBL sequence data bank under the accession number X63612.

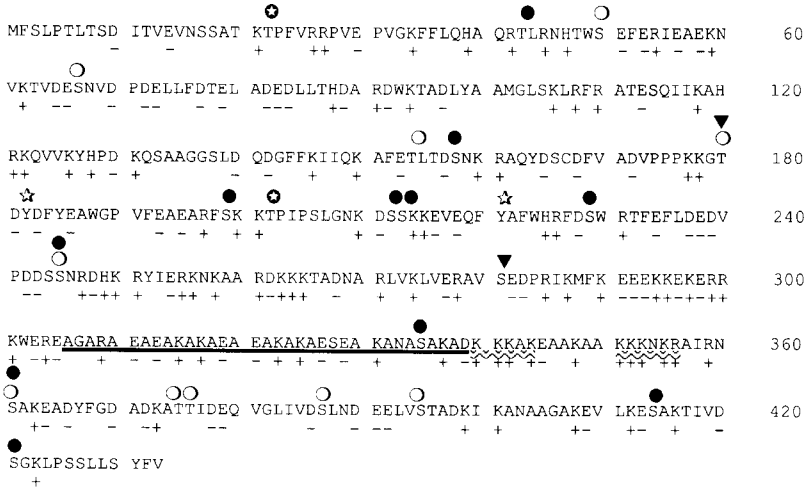

Fig. 5. **(A)** The single letter amino acid sequence of zuotin is shown with a number of features. Numerous potential phosphorylation sites are present. ✪ CDC28 (cdc2) phosphorylation sites; ○, casein kinase II sites; ▼, cAMP-dependent protein kinase sites; ☆, tyrosine kinase sites; and ●, protein kinase C sites. Unusual sequence features include: ___, alternating alanine and charged amino acid repeats; ≈≈≈≈≈ the potential bipartite nuclear targeting sequences. The signs + and − under some amino acids indicate positively and negatively charged residues, respectively. There are segments with di-, tri- and tetra-positively charged residues as well as di- and tri-negatively charged residues. There is a segment having 15 charged residues in a row with only one exception (289−305) in one highly charged cluster area. The regions from 65−84 and from 364−392 are highly negatively charged.

A

```
zuotin      111 ATESQIIKAHKRKQVVKYHPDKQSAAGGSLDQDGFFKIIQKAFETLTDSNKRAQYD 165
DnaJ         16 AEEREIRKAY.KRLAMKYHPDRNQGDKEA.EAKFKEIKEAYEV..LTDSQKRAAYD  67
SCJ1         61 ATEKEIKSAY.RQLSKKYHPDKNAGSEEAHQ.KFIEVGEAYDVLSDPE..KKKIYD 112
csp32        28 ATGDDIKKTY.RKLALKYHPDKNPDNVDAAD.KFKEVNRAHSILSDQT..KRNIYD  79
SIS1         17 ANEQELKKGY.RKAALKYHPDK.PTG.DT.E.KFKEISEAFEILNDP..QKREIYD  65
YDJ1         17 ATDVEIKKAY.RKCALKYHPDKNPSEEAA.E.KFKEASAAYEILSDP..EKRDIYD  67
Avian t&T    17 ATAADIKTAY.RRTALKYHPDKGGDEEKMKELNTLMEEFRETEGLRAD.ETLEDSD  71
SEC63       144 ASDRDIKSAY.RKLSVKFHPDKLAKGLTPDE.K.VQITKAYESLTD.E.LVRQNYL 196
HDJ-1        14 ALGRGDQAGL.PPPGLRYHPDLNLEPG.AEE.LFLEIAEAYDVLSDP..RLREIFD 66

Consensus       AT-.-IKKAY.RKΦALKYHPDKN....-A.E.KF.EΦ..AY-ΦLSDP-..KR.ΦYD
```

B

```
Zuotin      304 REAGARAEAEAKAKAEAEAKAKAESEA.KANASAKADKKKAKEAAKAAKKKNKRAIRNSAK 363
                : |. |  | || || | || || . | ||    |  ||| .|| : ||.|:..: |.|
Sea urchin146 KAAAKRKAALAKKKAAA.AKRKAAAKAKKAKKPKKKAAKKAKKPAKKSPKKAKKPAKKSPK 205


Zuotin      295 KEKERRKWEREAGARAEAEAKAKAEAEAKAKAESEAKANASAK.ADKKKAKEAAKAAKKKNKRA 357
                |||  ||  | ::|:.   |  |  | || |.   | :  |.| | |  |  . ||||   | |
ChickenH1   113 KEKAPRK..RATAAKPKKPAAKKPAAAAK.KPKKAAAVKKSPKKAKKPAAAATKKAAKSPKKAA 152
```

C

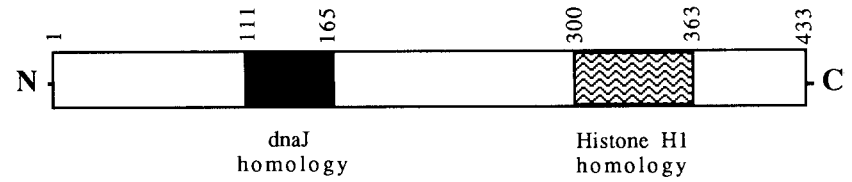

Fig. 6. Sequence comparisons of zuotin and other proteins. **(A)** A segment of zuotin (111−165) has 46% identity and 72% similarity with DnaJ. Other homologous yeast proteins include YDJ1 and SCJ1, SIS1, SEC63. csp32 is a *Drosophila* protein expressed in neuron cells, avian t and T are small and large T antigens of an avian polyoma virus. HDJ-1 is a human protein that shares DnaJ homology. In addition, all of these proteins contain a KYHPDK motif with few conserved amino acids substitutions. The consensus comes from computer analysis. The consensus means that at least five identical amino acids are aligned in a row, _, the negatively charged amino acids, Φ, the nonpolar amino acids. The KYHPDK motif is highlighted in bold. The dots indicate amino acids that are not conserved. **(B)** Different segment of zuotin compared with sea urchin and chicken histone H1 sequences. The identity is 46% and the similarity is 64% with sea urchin H1. The identity is shown by a line, similarity are shown by : or a dot. **(C)** The relative positions of these homologous sequences are shown in the zuotin sequence.

Discussion

DNA binding properties of zuotin

Yeast protein fractions containing zuotin are able to bind both poly(dG-m^{5}dC) and oligo(dG-Br^{5}dC)$_{22}$ in the Z-form as well as negatively supercoiled pUC19(GC) containing a Z-form segment in the presence of competitor B-DNA (Figures 1 and 2). Since zuotin has not been purified to homogeneity from the yeast extract or from the expressed *E.coli* cell extract, it is difficult to obtain a precise Z-DNA binding constant. We estimate that the yeast zuotin has a several hundred-fold enhanced affinity for Z-DNA relative to B-DNA under the experimental conditions used while the *E.coli* expressed zuotin binds less tightly. It would not be surprising if Z-DNA interacting proteins use binding motifs different from those found in several known binding motifs interacting with B-DNA. B-DNA has distinct major and minor grooves and B-DNA binding proteins tend to either anchor their binding motifs in the major grooves or lie along the minor grooves (reviewed in Seeman *et al.*, 1976; Pabo and Sauer, 1984; Churchill and Travers, 1991). Such binding to B-DNA is relatively tight and in some cases very specific. On the other hand, Z-DNA does not have a distinct major groove nor a highly accessible minor groove. Thus, it is possible that a Z-DNA binding protein would not be able to anchor its binding motif to the region corresponding to the major groove. It is possible that Z-DNA binding constants from proteins are lower than their B-DNA binding counterparts. It has been shown that many DNA sequences can adopt the left-handed Z-conformation (Rich *et al.*, 1984). Thus, proteins that recognize Z-DNA may be conformationally specific and/or sequence specific.

Zuotin binding to poly(dG-m^{5}dC) from both yeast cells and expressed in *E.coli* cannot be competed by 40-fold excess of poly(dG-dC) (Figure 8, lanes 11−12 and 18), 200-fold excess of poly(dA-dT) (data not shown) and several thousand-fold of salmon sperm DNA (Figure 2). However, a mere 4-fold excess of poly(dG-Br^{5}dC) in the Z-form completely competed the probe (data not shown). The results suggest that zuotin may recognize the conformation of DNA rather than sequences *per se*. However, additional experiments using a eukaryotic expressed protein will be required to characterize zuotin further. Another example of a protein that recognizes DNA conformation specifically is HMG1 (high mobility group protein). HMG1 and proteins containing HMG1 domains bind DNA not by its sequence but rather by the DNA conformation at the crossing of two duplexes (Bianchi *et al.*, 1992; Lilley, 1992).

The possible biological function of zuotin

Previous studies have suggested that potential Z-forming sequences, i.e. (GC/GC)n, (GT/AC)n and other alternating purine/pyrimidine segments exist in the intergenic regions

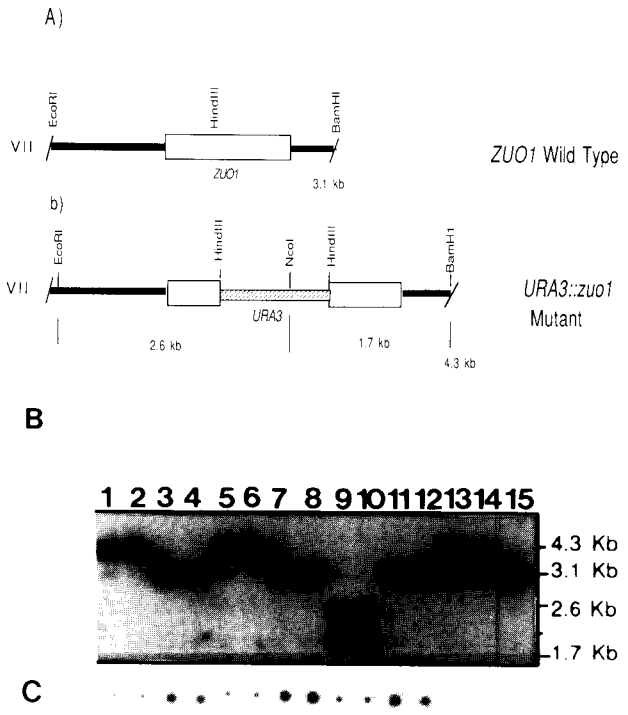

Fig. 7. Disruption of *ZUO1*. **(A)** Restriction map of the wild type *ZUO1* locus and the *URA3::zuo1* disrupted locus. Wild type DNA has a 3.1 kb *Eco*RI−*Bam*HI fragment, whereas, the disrupted mutant has a 4.3 kb fragment. **(B)** Southern blot of wild type and disrupted *ZUO1* locus probed with the entire coding region. Lanes 1−4 and lanes 5−8 are two sets of tetrad DNA digested with *Eco*RI and *Bam*HI. Lanes 9−12, DNA from the first set of tetrads digested with *Eco*RI, *Bam*HI and *Nco*I. Lanes 13, (clockwise insertion) and 14 (anti-clockwise insertion) of *URA3* are diploids disrupted in both orientations digested with *Eco*RI and *Bam*HI. Lane 15, DM27, the wild type diploid *ZUO1* locus digested with *Eco*RI and *Bam*HI. **(C)** The phenotype of the tetrad cell colonies correlates with the disruption and Southern blot analysis, i.e. the 3.1 kb DNA fragments are associated with large wild type colony size.

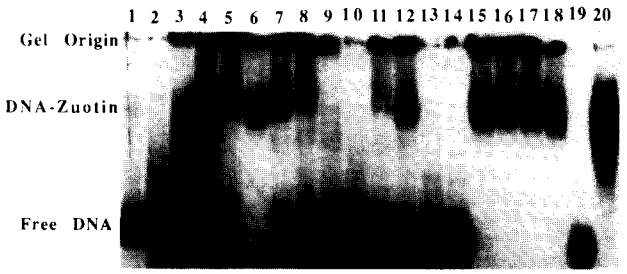

Fig. 8. Gel retardation assay using *E.coli* expressed zuotin and partially purified yeast zuotin. Both the crude *E.coli* extract and a partially purified zuotin from a phosphocellulose fraction were used for the assay. ^{32}P-labelled poly(dG-m^{5}dC) in the Z-form was incubated with the *E.coli* cell extract or with partially purified zoutin in the presence of sheared salmon sperm DNA, poly(dG-dC), or poly(dG-Br^{5}dC). All lanes have [^{32}P]poly(dG-m^{5}dC) in the Z-form: lane 1, labelled probe alone; lane 2, IPTG induced *E.coli* cell extract transformed with the parental plasmid pET8c without the *ZUO1* gene; lane 3, cell extract transformed with expression plasmid pETZUO1, but not induced by IPTG; lane 4, cell extract containing pETZUO1 induced with IPTG; lane 5, partially purified *E.coli* expressed zuotin; lane 6, partially purified yeast zuotin; lanes 7−10, partially purified expresed zuotin with 20-, 40-, 100-, 200-fold excess of salmon sperm DNA; lanes 11−12, 40- and 20-fold excess of poly(dG-dC); lanes 13−14, 20- and 40-fold excess of poly(dG-Br^{5}dC); lanes 15−19, all have partially purified yeast zuotin; lanes 15−17 with 20-, 40-, and 200-fold excess of salmon sperm DNA; lane 18, 40-fold excess of poly(dG-dC); lane 19, 40-fold excess of poly(dG-Br^{5}dC); lane 20, an anti-Z-antibody. The gel origin, DNA−zuotin complex and free DNA are indicated.

of many organisms, including those of yeast (Hamada *et al.*, 1982). These GT/AC segments have been implicated in inducing homologous DNA recombination *in vivo* (Bullock *et al.*, 1986; Treco and Arnheim, 1986; Wahls *et al.*, 1990). Also, a DNA strand transferase from human cells has been partially purified using a Z-DNA affinity column (Fishel *et al.*, 1988). Recently, it has been shown that specific alternating purine/pyrimidine segments in the upstream region of c-*myc* form Z-DNA during active transcription (Witting,B., Wölfl,S., Dorbic,T., Vahrson,W. and Rich,A., in preparation).

The biological function of zuotin is not known at the present time. Since zuotin appears to be of nuclear origin, binds to DNA, is relatively abundant and may potentially be phosphorylated by protein kinases and dephosphorylated by phosphatases during the cell cycle, it could be involved in chromosome organization. The threonine within the KTPFVRR and KTPIP sequences may be phosphorylated by the *S.cerevisiae* CDC28−CLN complex during the cell cycle. Experiments are underway to test this possibility.

Computer sequence comparison analysis revealed that two different domains of zuotin have similarities with three classes of proteins. The first class comprises the DnaJ protein. One domain of zuotin (residues 111−165) has 46% identity and an overall 70% similarity with the N-terminus of *E.coli* DnaJ protein (residues 16−67) (Figure 6A). *E.coli* DnaJ protein is a heat shock protein involved in protein folding and it is also active in phage λ and P1 replication *in vivo* and *in vitro* through its interaction with DnaB helicase (Liberek *et al.*, 1988; Zylicz *et al.*, 1989). The same domain also shares homology with several other yeast proteins: YDJ1 (a yeast DnaJ homolog), which may be involved in protein assembly into the endoplasmic reticulum and nucleus (Caplan and Douglas, 1991); SCJ1, which is involved in protein sorting (Blumberg and Silver, 1991); and SIS1, which is an essential protein and may be involved in yeast DNA replication by mediating a specific protein−protein interaction (Luke *et al.*, 1991). Other homologies involve the protein SEC63/NLS1, which is important for protein assembly into the endoplasmic reticulum and the nucleus (Sadler *et al.*, 1989); both YDJ1 and SIS1 have several cysteines that could potentially form a zinc finger DNA binding motif, but zuotin has only one cysteine and no zinc finger motif could be found. All the above proteins have a conserved hexapeptide KYHPDK except SEC63, in which F has replaced Y (Figure 6A). This peptide motif may play an essential role in these diverse proteins. Moreover, both small t and large T antigens of avian polyomavirus, budgerigar fledgling disease virus, as well as csp29 and csp32 proteins expressed in *Drosophila* retina and brain have the identical hexapeptide motif (Figure 6A) (Zinsmaier *et al.*, 1990). Recently, a human nuclear protein HDJ-1 has also been shown to share homology at both N- and C-termini with DnaJ protein (Raabe and Manley, 1991).

A second homology region is related to histone H1 and some of its variants, such as human H1a, H1b and H1c, chicken H1.11L and H1.11R, and sea urchin H1β and H1δ. The region of similarity is in the extended C-terminal tail of histone H1, a region rich in lysine and alanine residues. As shown in Figure 6B, over a 60 amino acid segment the

identity is 46% and the similarity is 64% (Figure 6B). It is significant that histone H2A, H2B, H3 and H4 do not have regions similar to zuotin. Calf thymus histone H1 has been shown to have a higher affinity to Z-DNA than to B-DNA and it is able to convert Z-DNA to B-DNA, a transition that can be measured using circular dichroism (Russell *et al.*, 1983; Mura and Stollar, 1984). Also, *Drosophila* histone H1 has been previously purified using a Z-DNA affinity column and Z-DNA binding assays in this laboratory (K.Lowenhaupt and A.Rich, unpublished observations). From these observations, we wonder whether zuotin has some elements of histone H1 activity in yeast. We are currently testing this hypothesis.

In an attempt to see whether zuotin is found in other organisms, we carried out a Southern 'zooblot' that probed various DNAs with zuotin DNA. Of 12 plant and animal species that were probed under low stringency, all were negative except yeast (results not shown). This suggests that zuotin may be a yeast-specific protein.

From the experiments in which *ZUO1* was disrupted, the mutants cells have a slow growth phenotype. Thus, its function appears not to be essential, but it may be involved in some activity that is needed to maintain rapid cell growth.

What DNA sequences zuotin binds to *in vivo*, where it binds on chromosomes, when it is most active in the cell cycle, how it affects cell growth, and how it is regulated (at both the RNA and protein levels) remain to be determined by using genetic and biochemical approaches.

Materials and methods

Yeast strains and media

The genotype and sources of yeast *S.cerevisiae* and media used in this work are as follows: DB2670 (*MAT*α, his3-Δ200, ura3-52 can1 pep4::HIS3, prb1-Δ1.6R) was obtained from D.Botstein; 20B-12 (*MAT*α, pep4-3, trp1) has been previously described (Jones, 1977); DM27 (*MAT*α/α, his3/HIS3, leu2/LEU2, ade2/ADE2 ura3/ura3, trp1/trp1, cyh/CYH) was obtained from D.Dawson. Cells were grown in YPD medium (1% yeast extract, 2% bactopeptone and 2% glucose). SD medium contained 0.67% Difco yeast nitrogen base without amino acids and 2% glucose. Nutrients essential for auxotropic strains were supplied at concentrations recommended by Sherman *et al.* (1986). The plasmid, pUC19(GC), was obtained from B.Johnston. Characterization of the anti-Z-DNA Z-22 monoclonal antibody and polyclonal goat anti-Z-DNA antibody have been described (Möller *et al.*, 1982).

Preparation of the poly(dG-m^{5}dC) affinity matrix

Poly(dG-m^{5}dC) DNA (Pharmacia) (1.6 mg in 3 ml) was digested to an average size of ~600 bp using DNase I in 50 mM Tris–HCl (pH 7.5), 30 μg/ml BSA in the presence of 2 mM $MnCl_2$ to produce blunt ends (Maniatis *et al.*, 1982). The digested DNA was deproteinized and resuspended in T4 DNA polymerase buffer in the absence of dNTP and incubated with T4 DNA polymerase at 10°C for 10 min. Subsequently dGTP and biotinylated dCTP (ENZO) were added to 1 mM final concentration and incubation was continued at 37°C for 2 h. DNA was then separated from unincorporated nucleotides by phenol and chroloform extraction followed by two ethanol precipitations. Then DNA was dissolved in 0.1 M NaCl, 1 mM EDTA, 10 mM TrisH–Cl, (pH 7.5) and incubated with 1 ml of streptavidin–agarose (BRL) overnight by gentle inversion. Under these conditions, >60% of the input DNA was bound to streptavidin–agarose as determined by A_{260} measurement after pelleting the agarose. The DNA matrix was then washed extensively with 40 column volumes of buffer [10 mM Tris–HCl (pH7.5), 50 mM KCl and 15 mM $MgCl_2$]. The column wash was assayed for unbound DNA (BluGene non-radioactive nucleic acid detection system, BRL) to assess the column stability.

Purification and sequencing of zuotin

The preparation of crude nuclear fractions has been described previously (Winter and Varshavsky, 1989). Crude total cell extract (from strain DB2670) was prepared from mid-log phase yeast cells (18 l at ~2.4×10^7 cells/ml). Yeast cells were collected at 4500 *g* for 10 min. at 4°C and washed twice with water. The cell pellet was resuspended in 600 ml of 0.2 M potassium phosphate (pH 7.5), 5 mM EDTA, 1 mM phenylmethylsulfonyl flouride (PMSF), 0.8 μg/ml pepstatin A and 10% glycerol. The final volume was ~800 ml. The cell suspension was processed through a high pressure compressor at 900 psi for ~50 passages to break the cells. The suspension was then centrifuged at 12 000 *g* for 10 min at 4°C. The supernatant (~720 ml) was frozen at −80°C. Fractionation of whole cell extract on phosphocellulose P-11 column and Mono-S column used the methods described by Winter and Varshavsky (1989). Pooled fraction (FI) containing binding activity to poly(dG-m^{5}dC) in the presence of sheared salmon sperm DNA was diluted 6-fold to achieve a potassium phosphate concentration of ~50 mM (420 ml). $MgCl_2$ was then added to a final concentration of 15 mM for loading on the poly(dG-m^{5}dC) affinity column. The column was washed with 10 vol of the column buffer and eluted with a linear gradient of 25 ml 0.1–1.0 M KCl without $MgCl_2$ (FII).

The eluted proteins were then loaded on a Superose-12 gel filtration column. One ml fractions were collected and assayed for the Z-DNA binding activity (FIII). These fractions were pooled, diluted 3-fold and loaded on a Mono-S column. The Mono-S column was washed extensively and eluted with a linear 47 ml gradient of buffer, 0–1.0 M NaCl, in 10 mM sodium phosphate (pH 7.2). The 1 ml fractions containing Z-DNA binding activity (FIV) were analysed by SDS–PAGE.

For protein composition and sequence analysis, the pooled Mono-S column fractions 12 and 13 were resolved on a 9% polyacrylamide–SDS gel. After electrophoresis, the protein was electroblotted on Immobilon (Millipore), briefly stained with Coomassie blue and washed. The band with an apparent molecular weight of 50 kDa was excised and 1/4 of the sample was used for amino acid composition analysis. The remaining sample was N-terminally sequenced by automated Edman degradation in an Applied Biosystems 470A Protein Sequencer equipped with on-line 120A PTH analyser.

Band shift and competition experiments

A gel retardation assay was employed to detect proteins with affinity to left-handed Z-DNA. In these assays, two kinds of left-handed DNA probes were used. One was poly(dG-m^{5}dC) (Pharmacia) stabilized in the left-handed Z-form by 15 mM $MgCl_2$. The DNA probe was made as follows: DNA polymer was digested with DNase I in the presence of 2 mM $MnCl_2$ and fragments of ~600–1000 bp were gel purified and labelled with T4 DNA polymerase in the presence of dCTP and [α-^{32}P]dGTP (3000 Ci/mM) (Maniatis *et al.*, 1982). The other probe was derived from a synthetic 44mer oligo $(dG\text{-}Br^5dC)_{22}$ labelled using Klenow polymerase with [α-^{32}P]dGTP and stabilized in the Z-DNA conformation by 10 mM $MgCl_2$ or 0.1 mM $Co(NH_3)_6^{3+}$ (A.Herbert and A.Rich, in preparation). In the assay reaction, 2 μl of diluted fraction samples were incubated for 20 min at room temperature in 20 μl of 50 mM Tris–HCl (pH 8.0), 15 mM $MgCl_2$, 5% sucrose, 0.1% Triton X-100, 10 mM β-mercaptoethanol, a 2000-fold excess of sheared salmon sperm DNA and the labelled Z-DNA probe. In the competition assays, supercoiled pUC19 and pUC19(CG) were added to the reaction samples separately. The samples were then electrophoresed in 1.5% agarose for poly(dG-m^{5}dC) (10 mM $MgCl_2$, 1×TBE pH 8.4) or 4% polyacrylamide for $(dG\text{-}Br^5dC)_{22}$ (1×TBE pH 8.4). After electrophoresis the gel was dried and exposed to X-ray film.

Southwestern blotting

Mono-S column fractions containing proteins were electrophoresed on 9% polyacrylamide–SDS gel as described by Laemmli (1970). After electrophoresis the gel was soaked in the running buffer without SDS but with 20% methanol at room temperature for 45 min with agitation. The proteins were then blotted onto two sheets of Immobilon P membrane (Millipore) at 1.5 mA/cm^2 at room temperature for 60 min. After transfer, the membrane was washed once with 5% milk powder, 30 mM HEPES (pH 7.4) (Celenza and Carlson, 1986), then three times (3 min each) with 10 mM Tris–HCl, 50 mM NaCl, 1 mM EDTA, 10 mM $MgCl_2$ (STEM). After washing, the membrane was incubated with 15 ml of the above buffer containing 10 μg/ml sheared salmon sperm DNA. A DNA probe (average size 1300 bp) [at 2.55×10^6 c.p.m./ml of poly(dG-m^{5}dC)] was added to a final concentration of 30 ng/ml and the incubation continued for 1 h at room temperature with gentle agitation (50 r.p.m.). In this buffer with 10 mM $MgCl_2$, the polymer is in Z-DNA conformation (Behe and Felsenfeld, 1981). The membrane was then washed four times (8 min each) with STEM at room temperature. The membrane was air-dried and exposed to X-ray film.

Cloning and sequencing of ZUO1

The first 11 amino acids of the N-terminal sequence of zuotin (MFSLPTLTSDI) were used to design oligonucleotides, with the yeast codon

usage as a guideline (Sharp *et al.*, 1986). The pools contained an equal molar mixture of 64 different 32mer sequences as follows:

5'-ATGTTTTCTTTGCCAACTTTGACTTCTGATAT-3'
C T C C C

These were gel purified and labelled using [γ-^{32}P]ATP and T4 polynucleotide kinase to a specific activity 200–500 μCi/μg. The labelled oligonucleotides were first used in genomic Southern blot hybridization to determine optimal conditions for screening a yeast phage λ library. For this, yeast DNA was digested with *Hin*dIII, separated on a 1% agarose gel and blotted onto a Gene Screen filter (New England Nuclear). The filter was then hybridized at 40°C in 4.5% SDS, 0.34 M NaCl, 1 mM Na–EDTA, 10 mg/ml BSA and 0.16 M sodium phosphate buffer (pH7.0) overnight. The filters were then washed at temperatures between 40 and 80°C with steps of 5°C in 3 M tetramethylammonium chloride, 2 mM EDTA, 0.1% SDS and 50 mM Tris–HCl (pH 8.0) as described by Wood *et al.* (1985). At the washing temperature of 65°C, there was a single hybridization band ~2.4 kb in size. This temperature was chosen for screening a *S.cerevisiae* genomic phage λ EMBL3A DNA library. Screening of the phage EMBL3A library was essentially the same as described by Winter and Varshavsky (1989). Thirteen positive phage plaques were isolated. Phage DNAs from 11 clones were purified from the confluent plate lysates. The DNA was then digested with several restriction enzymes and a Southern blot was performed as described previously. The restriction pattern and hybridization analysis revealed that the clones fell into three classes. One phage clone with a 3.1 kb *Eco*RI and *Bam*HI fragment was chosen for subcloning into pBluescript. Sequence analysis was carried out as described by Sanger *et al.* (1977). The 3.1 kb DNA fragment was sequenced on both strands with synthetic oligonucleotides as primers using the USB Sequenase kit (Version 2.0).

Expression of zuotin in E.coli

Zuotin was expressed in *E.coli* using the T7 pET expression system (Studier *et al.*, 1990). In order to insert *ZUO1* into the *Nco*I site of the pET8c vector, two bases flanking the initiation ATG were modified: −1 (GC) and at +4 (TG). This produced an amino acid change immediately after Met (PheVal). An oligonucleotide of 25 bases, CAAGAGTAA*CCATGG*TTTCTTTACC, was synthesized and used as primer for PCR amplification. A fragment of 1.8 kb containing the entire coding region of zuotin was amplified using PCR from pSKIIZUO1 and ligated into pET8c which was previously digested with *Nco*I and *Bam*HI and dephosphorylated. The pETZUO1 clones were isolated by colony hybridization using the coding region of *ZUO1* as a probe. Extensive restriction mapping and sequencing using the T7 primer and several internal primers confirmed the correct in-frame cloning of *ZUO1*. Furthermore, the zuotin expressed in *E.coli* had its N-terminus sequenced and its composition analysed by hydrolysis in order to confirm the correct expression (Table I and Figure 4B).

E.coli strain BL21E3LysS, which carries T7 RNA polymerase in the chromosome of the host under the control of the lacUV5 promoter, was transformed with pETZUO1. The transformants were induced with 0.5 mM IPTG after the cells had reached a density of 0.5 O.D. (A_{600}). Cells were harvested 3 h after induction, lysed, treated and analysed as described by Sambrook *et al.* (1989).

After cells were lysed by sonication, the cell suspension was centrifuged at 8000 *g* for 20 min at 4°C, and both supernatant and pellet were saved. The pellet was resuspended in buffer containing 10 mM Tris–HCl, 50 mM NaCl and 1 mM PMSF, and then urea was added to a final concentration of 4 M in order to denature the inclusion bodies. The suspension was stirred at 4°C for 4 h. The suspension was then dialysed overnight at 4°C with three changes of buffer. The dialysed suspension was centrifuged at 10 000 *g* for 30 min. The supernatant was then loaded on the phosphocellulose column. The column was washed in 100 mM KH_2PO_4 buffer and eluted with 1.0 M KH_2PO_4. The eluent was dialysed and used for characterization.

Construction of zuo1 disruption mutants

The gene disruption method of Rothstein (1983) was used for generating the *zuo1* mutants in the yeast genome. A 1.17 kb *Hin*dIII fragment containing *URA3* was inserted at the unique *Hin*dIII site within the coding region of *ZUO1* (corresponding to amino acid position 186). The DNA was then cut with *Eco*RI and *Bam*HI and the released fragment was used to transform a diploid *S.cerevisiae* DM27. Standard techniques were used for yeast transformation, sporulation and tetrad dissection (Sherman *et al.*, 1986). Both orientations of the *URA3* insert were used and yielded the same results.

Computer analysis of zuotin

The predicted structure of zuotin was analysed by computer algorithms using Pepplot, FastA, BLAST and others, in the Genetics Computer Group (GCG) package, as installed at the Whitaker College Computer Facility at M.I.T. GeneWorks Version 2.0 (1991, Intelligenetics, Inc., Mountain View) was used for zuotin alignment with DnaJ and other proteins.

Acknowledgements

We thank David Botstein for yeast strain DB2070, Richard Young for the yeast EMBL3A DNA library, and William Lane (Harvard University) and Richard Cook (M.I.T.) for protein sequencing and amino acid composition analysis, Xing Su (Harvard University) and Stefan Wölfl for PCR reactions. We also thank Dr Dean Dawson's laboratory at Tufts University Medical School for assisting in our yeast work and Dr Martin Egli for sequencing through the homopurine region of DNA. Diana Marmorstein, Daniel Peisach and Gabriella Bulboaca of the M.I.T. Undergraduate Research Opportunity Program are also acknowledged for excellent technical assistance. Linda Riles (Washington University) is thanked for helping to localizing *ZUO1* on yeast chromosome VII. Michael Hengartner for zuotin alignment. We also thank members of Dr Alexander Varshavsky's laboratory for helpful discussions and supplying materials for experiments. We thank Drs Ky Lowenhaupt, Kiran Madura, Alexander Varshavsky and Stefan Wölfl for critically reading this manuscript. This work is supported by grants from National Institute of Health, National Science Foundation, National Aeronautics and Space Administration, Office of Naval Research and American Cancer Society. SZ was an American Cancer Society postdoctoral fellow.

References

Azorin,F., Hahn,R. and Rich,A. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 5714–5718.
Bairoch,A. (1991) *Nucleic Acids Res.*, **19 supp**, 2241–2245.
Behe,M. and Felsenfeld,G. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 1619–1623.
Bianchi,M.E., Falciola,L., Ferrari,S. and Lilley,D.M.J. (1992) *EMBO J.*, **11**, 1055–1063.
Blaho,J.A. and Wells,R.D. (1987) *J. Biol. Chem.*, **262**, 6082–6088.
Blumberg,H. and Silver,P.A. (1991) *Nature*, **349**, 627–629.
Bullock,P., Miller,J. and Botchan,M. (1986) *Mol. Cell. Biol.*, **6**, 3948–3953.
Caplan,A. and Douglas,M. (1991) *J. Cell Biol.*, **114**, 609–621.
Celenza,J.L. and Carlson,M. (1986) *Science*, **233**, 1175–1180.
Chou,P.Y. and Fasman,G.D. (1978) *Annu. Rev. Biochem.*, **47**, 251–276.
Churchill,M.E.A. and Travers,A.A. (1991) *Trends Biochem. Sci.*, **183**, 92–97.
Fishel,R., Detmer,K. and Rich,A. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 36–40.
Hamada,H., Petrino,M.G. and Kakunaga,T. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 6465–6469.
Jaworski,A., Hsieh,W.T., Blaho,J.A., Larson,J.E. and Wells,R.D. (1987) *Science*, **238**, 773–777.
Jones,E. (1977) *Genetics*, **85**, 23–33.
Kmiec,E.B. and Holloman,W.K. (1986) *Cell*, **44**, 545–554.
Laemmli,U.K. (1970) *Nature*, **227**, 680–682.
Lafer,E.M., Sousa,R. and Rich,A. (1985) *EMBO J.*, **4**, 3655–3660.
Lechner,J. and Carbon,J. (1991) *Cell*, **64**, 717–725.
Liberek,K., Georgopoulos,C. and Zylicz,M. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 6632–6636.
Lilley,D.M.J. (1992) *Nature*, **357**, 282–283
Liu,L.F. and Wang,J.C. (1987) *Proc. Natl. Acad. Sci. USA.*, **84**, 7024–7028.
Luke,M.M., Sutton,A. and Arndt,K.T. (1991) *J. Cell Biol.*, **114**, 623–638.
Maniatis,T., Fritch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
McCarthy,J.G. and Heywood,S.M. (1987) *Nucleic Acids Res.*, **15**, 8069–8085.
Möller,A., Gabriels,J.E., Lafer,E.M., Nordheim,A., Rich,A. and Stollar,B.D. (1982) *J. Biol. Chem.*, **257**, 12081–12085.
Moreno,S. and Nurse,P. (1990) *Cell*, **61**, 549–551.
Mura,C.V. and Stollar,B.D. (1984) *Biochemistry*, **23**, 6147–6152.
Naylor,L.H. and Clark,E.M. (1990) *Nucleic Acids Res.*, **18**, 1595–1601.
Nordheim,A. and Rich,A. (1983) *Nature*, **303**, 674–679.
Pabo,C.O. and Sauer,R.T. (1984) *Annu. Rev. Biochem.*, **53**, 293–322.
Peck,L.J., Nordheim,A., Rich,A. and Wang,J.C. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 4560–4564.
Raabe,T. and Manley,J.L. (1991) *Nucleic Acids. Res.*, **19**, 6645.
Rahmouni,A.R. and Wells,R.D. (1989) *Science*, **246**, 358–363.

Rich,A., Nordheim,A. and Wang,A.H.-J. (1984) *Annu. Rev. Biochem.*, **53**, 791–864.
Robbins,J., Dilworth,S.M., Laskey,R.A. and Dingwall,C. (1991) *Cell*, **64**, 615–623.
Rothstein,R. (1983) *Methods Enzymol.*, **101**, 202–211.
Rott,O., Kroeger,M., Mueller,H. and Hobom,G. (1988) *Virology*, **165**, 74–86.
Russell,W.C., Precious,B., Martin,S.R. and Bayley,P.M. (1983) *EMBO J.*, **2**, 1647–1653.
Sadler,I., Chiang,A., Kurihara,T., Rothblatt,J., Way,J. and Silver,P. (1989) *J. Cell Biol.*, **109**, 2665–2675.
Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning. A Laboratory Manual, 2nd Ed.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sanger,F., Nicklen,S. and Coulson,A.R., (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Seeman,N.C., Rosenberg,J.M. and Rich,A. (1976) *Proc. Natl. Acad. Sci. USA*, **73**, 804–808.
Sharp,P.M., Tuohy,T.M. and Mosurski,K.R. (1986) *Nucleic Acids Res.*, **14**, 5125–5143.
Sherman,F., Fink,G.R. and Hicks,J.B. (1986) *Methods in Yeast Genetics.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Studier,F.W., Rosenberg,A.H., Dunn,J.J. and Dubendorff,J.W. (1990) *Methods Enzymol.*, **185**, 60–89.
Takeuchi,H., Hanamura,N., Hayasaka,H. and Harada,I. (1991) *FEBS Lett.*, **279**, 253–255.
Tsao,Y-P., Wu,H-Y. and Liu,L.F. (1989) *Cell*, **56**, 111–118.
Treco,D. and Arnheim,N. (1986) *Mol. Cell. Biol.* **6**, 3943–3947.
Vardimon,L. and Rich,A. (1983) *Proc. Natl. Acad. Sci. USA*, **81**, 3268–3272.
Wahls,W.P., Wallace,L.J. and Moore,P.D. (1990) *Mol. Cell. Biol.*, **10**, 785–793.
Winter,E. and Varshavsky,A. (1989) *EMBO J.*, **8**, 1867–1877.
Wittig,B., Dorbic,T. and Rich,A. (1989) *J. Cell Biol.*, **108**, 755–761.
Wood,W.I., Gitschier,J., Lasky,L. A. and Lawn,R.M. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 1585–1588.
Zinsmaier,K.E., Hofbauer,A., Heimbeck,G., Pflugfelder,G.O., Buchner,S. and Buchner,E. (1990) *J. Neurogenet.*, **7**, 15–29.
Zylicz,M., Ang,D., Liberek,K. and Georgopoulos,C. (1989) *EMBO.J.*, **8**, 1601–1608.

Received on February 25, 1992; revised on June 16, 1992