Vol. 201, No. 1, 1994 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
May 30, 1994 Pages 310-317

A POSSIBLE YEAST HOMOLOG OF HUMAN ACTIVE-GENE-REPAIRING HELICASE *ERCC6*+

Meng-Er HUANG, Jean-Claude CHUAT, and Francis GALIBERT*

Laboratoire de Biochimie et Biologie Moléculaire, Unité de Recombinaisons Génétiques (UPR 41 - CNRS), Faculté de Médecine, 2 avenue du Professeur Léon Bernard, 35043 Rennes, France

Received April 12, 1994

Summary: We report here the sequencing and identification on the chromosome X of *S. cerevisiae* of an open reading frame, designated *GTA1085*, encoding a protein 1085 amino acids in size that displays significant homology to a of helicase subfamily. The highest similarity score is with *ERCC6*, a human putative helicase involved in the repair of active genes, with 53.3 % identity over a stretch of 589 amino acids. This putative protein contains all seven consecutive domains conserved among DNA and RNA helicases. Thus, it apparently constitutes a novel member of this subfamily and might be involved, like *ERCC6*, in the preferential repair of active genes in yeast.

Helicases have been implicated in numerous cellular processes that require DNA or RNA unwinding, including DNA recombination and repair, initiation and progression of DNA replication, cell proliferation, RNA splicing, ribosome assembly, initiation of translation and transcription (1, 2). Sequence analysis has provided grounds for recognizing among these enzymes a subfamily whose members share a degree of homology (about 30 % identity and up to 57 % similarity) by far exceeds the resemblance normally present among helicases (3). This suggests a specific type of helicase function(s) common to all the members of this gene subfamily. This subgroup encompasses proteins implicated in transcription regulation, such as *SNF2* (4), *MOT1* (5) and *brm* (6); in preservation of chromosome stability, such as *lodestar* (7), and in DNA repair, such as *ERCC6* (3), *RAD16* (8), *RAD54* (9) and *RAD5* (10). This subfamily also includes proteins of unknown function, such as *STH1* (11) and *FUN30* (12).

+ Sequence data from this article have been deposited with the GSDB/EMBL/DDBJ Data Libraries under Accession No. L26910.

* To whom correspondence should be addressed. Fax: 99336200.

0006-291X/94 $5.00

Within the framework of the European Community project "Sequencing the yeast genome", we were assigned a fragment of chromosome X situated between genes *MET3* and *CDC8*. Sequencing this region led us to identify an open reading frame (ORF) specifying a protein significantly homologous to this helicase subfamily.

MATERIALS AND METHODS

Sequencing strategy and methods: DNA sequencing was performed by using the shotgun procedure on a whole recombinant cosmid consisting of a fragment approximately 40 kb in size inserted at the *Eco*RI site of cosmid vector pWE15. This cosmid insert originated from a cosmid library prepared by *Sau*3AI partial digestion of genomic DNA from a diploid strain, *FY1679*. The latter is issued from the cross between strains *FY23 (MATa, ura3-52, trp1Δ63, leu2Δ1, GAL2)* and *FY73 (MATα, ura3-52, his3Δ200, GAL2)* (13)

The DNA of the whole cosmid was sonicated, repaired with T4 DNA polymerase and migrated in a 1% low-melting-point agarose gel. Fragments with a size ranging from 0.5 to 1.0 kb were subcloned into the *Eco*RV site of a modified M13 vector derived from M13mp89 (14). Single-strand DNA was sequenced by means of the ABI 373A automatic sequencer using dye primers. Oligonucleotide-directed sequencing with dye terminators was also employed to complete the sequence on both strands and to resolve all the ambiguities.

DNA sequence analysis: DNA sequence assembly, DNA and protein sequence analysis were performed by using the different programs of the DNASTAR software (Lasergene Version). The GenBank (release 80) and EMBL (release 36) nucleic acid sequence databanks were searched with cosmid nucleic acid sequences and PIR (release 38) and Swiss-Prot (release 27) with translated amino acid sequences, using the FASTA algorithm (15) as implemented in the GCG software.

RESULTS AND DISCUSSION

Sequence analysis revealed an ORF of 3600 base pairs (bp) (databank accession no. L26910). The first ATG codon localized within the ORF is situated 346 bp downstream of the in-frame stop codon of *PET191* gene (16). Figure 1 shows the predicted amino acid sequence of this ORF, designated *GTA1085*, specifing a protein of 1085 residues with a predicted molecular size of 124 KDa. It comprises 35 % charged residues, although its net charge is essentially neutral. The sequence contains no motifs or domains such as an signal sequence, putative membrane-spanning helices or zinc-fingers. However, there is an acidic part between residue 173-228, harboring 48.2 % glutamic or aspartic acid and including a stretch of 7 acidic residues. Acidic regions have been found in a number of nuclear proteins that associate with chromatin or histones (3, 17). A systematic computer search with the FASTA program revealed 11 proteins with

```
   1  MEDKEQQDNAKLENNESLKDLGVNVLSQSSLEEKIANDVTNFSNLQSLQQEETRLERSKTALQRYVNKKNHLTRKLNNTT

  81  RISVKQNLRDQIKNLQSDDIERVLKDIDDIQSRIKELKEQVDQGAENKGSKEGLQRPGETEKEFLIRTGKITAFGHKAGF

 161  SLDTANREYAKNDEQKDEDFEMATEQMVENLTDEDDNLSDQDYQMSGKESEDDEEEENDDKILKELEDLRFRGQPGEAKD
                                                 acidic
 241  DGDELYYQERLKKWVKQRSCGSQRSSDLPEWRRPHPNIPDAKLNSQFKIPGEIYSLLFNYQKTCVQWLYELYQQNCGGII

 321  GDEMGLGKTIQVIAFIAALHHSGLLTGPVLIVCPATVMKQWCNEFQHWWPPLRTVILHSMGSGMASDQKFKMDENDLENL
           I                                Ia
 401  IMNSKPSDFSYEDWKNSTRTKKALESSYHLDKLIDKVVTDGHILITTYVGLRIHSDKLLKVKWQYAVLDEGHKIRNPDSE
                                                                              II
 481  ISLTCKKLKTHNRIILSGTPIQNNLTELWSLFDFIFPGKLGTLPVFQQQFVIPINIGGYANATNIQVQTGYKCAVALRDL
                              III
 561  ISPYLLRRVKADVAKDLPQKKEMVLFCKLTKYQRSKYLEFLHSSDLNQIQNGKRNVLFGIDILRKICNHPDLLDRDTKRH

 641  NPDYGDPKRSGKMQVVKQLLLLWHKQGYKALLFTQSRQMLDILEEFISTKDPDLSHLNYLRMDGTTNIKGRQSLVDRFNN
                                                 IV
 721  ESEDVELLTTRVGGLGVNLTGANRIIIFDPDWNPSTDMQARERAWRIGQKREVSIYRLMVGGSIEEKIYHRQIFKQFLTN
                    V                            VI
 801  RILTDPKQKRFFKIHELHDLFSLGGENGYSTEELNEEVQKHTENLKNSKSEESDDFEQLVNLSGVSKLESFYNGKEKKEN

 881  SKTEDDRLIEGLLGGESNLETVMSHDSVVNSHAGSSSSNIITKEASRVAIEAVNALRKSRKKITKQYEIGTPTWTGRFGK

 961  AGKIRKRDPLKNKLTGSAAILGNITKSQKEASKEARQENYDDGITFARSKEINSNTKTLENIRAYLQKQNNFFSSSVSIL

1041  NSIGVSLSDKEDVIKVRALLKTIAQFDKERKGWVLDEEFRNNNAS.
```

Figure 1. Deduced amino acid sequence of *GTA1085*. The acidic region is indicated. Numbers I, Ia, and II–VI refer to the corresponding helicase domains. The nucleotide sequence has been deposited in the Genome Sequence DataBase (accession number L26910).

significant homology to protein *GTA1085* (optimized FASTA score > 200) (Table 1). All these proteins are in a recently identified subfamily of helicases. The best score is with *ERCC6*, a human putative helicase involved in the preferential repair of active genes. It corrects the nucleotide excision repair (NER) defect of Cockayne's syndrome complementation group B (3)

Protein alignment of *GTA1085* and *ERCC6* revealed several regions of interest (Figure 2). Although the N-terminus and the C-terminus between the two proteins are respectively divergent, the central parts share strikingly identical amino acid, i.e. 53.3 % identity over a stretch of 589 amino acid (residue 239 through 828 of protein *GTA1085*). Similarity reaches 68.5 % if conservative amino acid substitutions are taken into consideration. This region contains the seven consecutive domains (I, Ia, II, III, IV, V, VI) conserved in RNA and DNA helicases (2). In domain I, both proteins contain the Walker Type A NTP-binding motif GXGKT (18). Domain II contains the Walker Type B box, which is characterized by several hydrophobic residues followed by the highly conserved DE pair (18). This domain is thought to be an Mg^{2+}-binding domain and to interact with Mg-nucleoside triphosphates through a conserved aspartic acid residue (19). The role of the hydrophobic residues of this sequence is to exclude water from the ATP reaction center (19). A special version of the Walker type B

Table 1 : Protiens showing significant homology to GTA1085 protien

Accession no. (database)	Gene name	amino acid	FASTA Score
A44224 (PIR)	ERCC6 (*Homo sapiens*)	1493	1330
S15047 (PIR)	SNF2 (*S. cerevisiae*)	1703	701
S28406 (PIR)	NPS1(*S. cerevisiae*)	1359	698
S22777 (PIR)	STH1 protein (*S. cerevisiae*)	1352	696
A42091 (PIR)	SNF2/SWI2 homolog-brm (*Drosophila*)	1638	630
S22775 (PIR)	MOT1 protein (*S. cerevisiae*)	1867	422
S21568 (PIR)	KYBP protein (*Mus musculus*)	940	361
JH0440 (PIR)	RAD54 protein (*S. cerevisiae*)	898	359
S22266 (PIR)	SNF2 homolog (*S. cerevisiae*)	1131	320
S31301 (PIR)	RAD5 protein (*S. cerevisiae*)	1169	259
P28370 (SW)	hSNF2L (*Homo sapiens*)	976	217

sequence, the DEAD box, is found in *eIF-4A* and in other RNA helicases (20). In domain VI, protein *GTA1085* has the sequence WRIGQ, whereas the sequence HRIGR is invariant in *eIF-4A* and other RNA helicases (20). The absence of the DEAD box and of the HRIGR sequence suggests that protein *GTA1085*, like *ERCC6*, might possess DNA but not RNA helicase activity. As shown in Figure 2, the conserved regions extend to either side of most of the above-mentioned domains, suggesting that these amino acids, although located outside the previously identified domains, have functional significance.

In addition, the central region of protein *GTA1085* exhibits high homology with other proteins (Table 1) containing the same seven domains: *SNF2* (4), *NPS1* (21), *STH1* (11), *brm* (6), *MOT1* (5), *RAD54* (9), *FUN30* (12), *RAD5* (10) and a presumptive human global transcription activator, *hSNF2L* (22). *NPS1* (21) and *STH1* (11) are apprently the same protein. As for mouse KYBP protein (unpublished, accession no. S21568), a search for characteristic helicase domains showed the presence of domains IV, V, VI, but information concerning the other domains, viz I, Ia, II, III, is not available because the sequence upstream of domain IV is missing. In spite of this, it is quite possible that the whole KYBP

```
239  KDDGDELYYQERLKKWVKQRSCGSQRSSDLPEWRRPHPNIPDAKLNSQFKIPGEIYSLLFNYQKT  GTA1085
453  RDDGDEDYYKQRLRRWNKLRLQDKEKRLKLED----DSEESDAEFDEGFKVPGFLFKKLFKYQQT  ERCC6
                    I                                          Ia
304  CVQWLYELY-QQNCGGIIGDEMGLGKTIQVIAFIAALHHSGLLT----------GPVLIVCPATV  GTA1085
514  GVRWLWELHCQQ-AGGILGDEMGLGKTIQIIAFLAGLSYSKIRTRGSNYRFEGLGPTVIVCPTTV  ERCC6
358  MKQWCNEFQHWWPPLRTV-ILHSMGSGMASDQKFKMDENDLENLIMNSKPSDFSYEDWKNSTRT-  GTA1085
578  MHQWVKEFHTWWPPFR-VAILHETGS--------------------------Y--------TH  ERCC6
                                                  II
421  KKALESSYHLDKLI-DKVVTDGHILITTYVGLRIHSDKLLKVKWQYAVLDEGHKIRNPDSEISLT  GTA1085
605  KK--------EKLIRD-VAHCHGILITSYSYIRLMQDDISRYDWHYVILDEGHKIRNPNAAVTLA  ERCC6
                  III
485  CKKLKTHNRIILSGTPIQNNLTELWSLFDFIFPGKLGTLPVFQQQFVIPINIGGYANATNIQVQT  GTA1085
661  CKQFRTPHRIILSGSPMQNNLRELWSLFDFIFPGKLGTLPVFMEQFSVPITMGGYSNASPVQVKT  ERCC6
550  GYKCAVALRDLISPYLLRRVKADVAKDL--PQKKEMVLFCKLTKYQRSKYLEFLHSSDLNQIQNG  GTA1085
726  AYKCACVLRDTINPYLLRRMKSDVKMSLSLPDKNEQVLFCRLTDEQHKVYQNFVDSKEVYRILNG  ERCC6
613  KRNVLFGIDILRKICNHPDLLDRDTKR-HN-PD-------YGDPKRSGKMQVVKQLLLLWHKQGY  GTA1085
791  EMQIFSGLIALRKICNHPDLFSGGPKNLKGLPDDELEEDQFGYWKRSGKMIVVESLLKIWHKQGQ  ERCC6
       IV                                                          
669  KALLFTQSRQMLDILEEFISTKDPDLSHLNYLRMDGTTNIKGRQSLVDRFNNE-S-FDVFLLTTR  GTA1085
856  RVLLFSQSRQMLDILEVFLRAQ-----KYTYLKMDGTTTIASRQPLITRYNEDTSIF-VFLLTTR  ERCC6
       V                    VI                                      
732  VGGLGVNLTGANRIIIFDPDWNPSTDMQARERAWRIGQKREVSIYRLMVGGSIEEKIYHRQIFKQ  GTA1085
915  VGGLGVNLTGANRVVIYDPDWNPSTDTQARERAWRIGQKKQVTVYRLLTAGTIEEKIYHRQIFKQ  ERCC6
797  FLTNRILTDPKQKRFFKIHELHDLFSLGGENG                                   GTA1085
980  FLTNRVLKDPKQRRFFKSNDLYELFTLTSPDA                                   ERCC6
```

Figure 2. Alignment of translated protein *GTA1085* with *ERCC6*. This alignment was generated by using the FASTA algorithm and further optimized by visual inspection. Two dots indicate identity, and one dot indicates conserved amino acid changes (residues belonging to one of the following groups are considered to be similar: L,V,I,M; F,Y,W; S,T,A,P,G; K,R,H; E,D,Q,N, ref. 2). Bold overlines indicate the seven helicase domains (I, Ia, II–VI). Amino acids at the N- and C-termini are not shown because there are few homologies.

protein will prove a member of this helicase subfamily. *RAD16* (8), another member of this subfamily involved in the NER pathway in *S. cerevisiae*, has a homology score just below the threshold (FASTA score: 193). Figure 3 shows a comparison domain by domain of *GTA1085* with the other related helicases. It can be seen that domains I, II, III, V and VI are the most conserved, while domains Ia and IV are less conserved.

Previous sequence analyses have revealed striking amino acid homology between human NER protein *ERCC1* and yeast NER protein *RAD10* (23), as well as between *ERCC2* and *RAD3* (24). These observations suggest that human and yeast NER proteins are functionally related and that *GTA1085* probably encodes

		I		Ia		II
		** ***** *		*** * * *		** **** ***
GTA1085	317	GGIIG**DEMGLGKT**IQV	347	GP-VLIV**C**P-ATVMK**QW**	462	KWQ-YAVL-**DEGH**KIRN
ERCC6	527	GGILG**DEMGLGKT**IQI	567	GPT-VIV**C**P-TTVMH**QW**	639	DW-HYVIL-**DEGH**KIRN
SNF2	787	NGILA**DEMGLGKT**IQT	817	GPYLVIV-**PLSTLS-NW**	887	KWVH-MII-**DEGH**RMKN
NPS1	490	NGILA**DEMGLGKT**IQS	520	GPFLVIV-**PLSTIT-NW**	590	DWAH-MII-**DEGH**RMKN
STH1	490	NGILA**DEMGLGKT**IQS	520	GPFLVIV-**PLSTIP-NW**	590	DWAH-MII-**DEGH**RMKN
brm	793	NGILA**DEMGLGKT**IQT	823	GPYLIIV-**PLSTLP-NW**	893	QW-KYMII-**DEGH**IMKN
MOT1	1292	HGILC**DDMGLGKT**LQT	1333	LPSLII-**C**P-PSLTG**HW**	1401	EY-NYCVL-**DEGH**RIKN
RAD54	330	GCIMA**DEMGLGKT**LQC	364	DKC-IIV**C**P-SSLVN**NW**	448	NVG--LML**ADEGH**MLKN
FUN30	592	SCILA**DDMGLGKT**CQV	621	GPHLVVV-**PSSTLE-NW**	696	NF-N-VVVY**DEGH**NLKN
RAD5	527	GGILS**DEMGLGKT**VAA	586	KTTLIVV-**PMSLLT-QW**	674	NF-YRIII-**DEGH**NIRN
hSNF2L	125	NGILA**DEMGLGKT**LQT	155	GPHMVLV-**PKSTL-HNW**	226	HW-RYLVI-**DEAH**RIKN
RAD16	205	GGVLA**DEMGMGKT**IQT	231	SPS-LVV**AP**-TVALM**QW**	315	DF-YRVIL-**DEAH**NIKD

		III		IV
		* ******* * * * ** *		** * ***
GTA1085	493	RIILS**GTPIQNNLTELWSLFDF**	669	KALLFTQ-SRQMLDIL**EE**
ERCC6	670	RIILS**GSPMQNNLRELWSLFDF**	857	RVLLFSQ-SRQMLDIL**EV**
SNF2	919	RLIL**TGTPLQNNLPELWALLNF**	1105	RVLIFFQMT-QIMDIM**ED**
NPS1	622	RLIL**TGTPLQNNLPELWALLNF**	809	RVLMFFQMT-QVMDIM**ED**
STH1	622	RLIL**TGTPLQNNLPELWALLNF**	809	RVLMFFQMT-QVMDIM**ED**
brm	925	RLLL**TGTPLQNKLPELWALLNF**	1112	RVLLFCQMT-QCMTII**ED**
MOT1	1432	RLIL**TGTPIQNNVLELWSLFDF**	1648	RALIFCQL-KDMLDMV**EN**
RAD54	479	RVIL**SGTPIQNDLSEYFALLSF**	674	KIVLISNYT-QTLDLI**EK**
FUN30	727	RLLL**TGTPLQNNLKELMSLLEF**	968	KVLIFSLFT-QVLDIL**EM**
RAD5	705	KWVL**TGTPIINRLDDLYSLVKF**	1014	QVVIFSQFST-YLDIL**EK**
hSNF2L	257	RLLL**TGTPLQNNLHELWALLNF**	426	RVLIFSQMTR-LLDIL**ED**
RAD16	346	RWC**LSGTPLQNRIGEMYSLIRF**	639	KSIVFSQFTS-MLDLV**EW**

		V		VI
		*** * ** **** *		* * * ** * * **
GTA1085	725	VFLLTTRV**GGLGVNLTGA**NRIIIF	751	DWN**PSTDMQARERAWRIGQK**
ERCC6	909	VFLLTTRV**GGLGVNLTGA**NRVVIY	935	DWN**PSTDTQARERAWRIGQK**
SNF2	1158	CFILSTRA**GGLGLNLQTA**DTVIIF	1184	DWN**PHQDLQAQDRAHRIGQK**
NPS1	862	CFLLSTRA**GGLGLNLQTA**DTVIIF	888	DWN**PHQDLQAQDRAHRIGQK**
STH1	862	CFLLSTRA**GGLGLNLQTA**DTVIIF	888	DWN**PHQDLQAQDRAHRIGQK**
brm	1165	VFLLSTRA**GGLGLNLQTA**DTVVIF	1191	DWN**PHQDLQAQDRAHRIGQR**
MOT1	1703	CLLLTTKV**GGLGLNLTGA**DTV-IF	1729	DWN**PMNDLQAMDRAHRIGQK**
RAD54	727	IFLLSSKA**GGCGINLIGA**NRLILM	753	DWN**PAADQQALARVWRDGQK**
FUN30	1020	IFILSTKA**GGFGINLVCA**NNVIIF	1046	SFN**PHDDRQAADRAHRVGQT**
RAD5	1072	ILLLSLKA**GGVGLNLTCA**SHA--Y	1098	WWS**PSMEDQAIDRLHRIGQT**
hSNF2L	491	IFMLSTRA**GGLGINLASA**DVVILY	517	DWN**PQVDLQAMDRAHRIGQK**
RAD16	691	VFLVSLKA**GGVALNLCEA**SQ--VF	717	WWN**PSVEWQSGDRVHRIGQY**

Figure 3. Comparison of the helicase domains of *GTA1085*, *ERCC6*, *SNF2*, *NPS1*, *STH1*, *brm*, *MOT1*, *RAD54*, *FUN30*, *RAD5*, *hSNF2L* and *RAD16*. Gene designation is on the far left. Numbers indicate the positions of the leftmost amino acid. Positions of identical or similar amino acid in the 12 proteins are indicated by asterisks. Identical residues among the 12 proteins are printed in boldface.

the yeast homolog of *ERCC6*. The *ERCC6* protein specifically corrects the NER deficiency of Cockayne's syndrome complementation group B. Accordingly, *GTA1085* protein might have a specific role in preferential repair of active genes, contrary to *RAD16*, which is involved in NER of the inactive HML locus (25). The NER pathway has been extensively studied in yeast *S. cerevisiae*. Sixteen genes, known as the *RAD3* epistasis group, have been identified as having a role in NER (26), and protein *GTA1085* is most probably a new member of this epistasis group. However, its definitive status will require further biochemical studies.

ACKNOWLEDGMENTS

We wish to thank Alan Kay for help with computer analysis. We also thank Géraldine Le Provost and Nicole Le Moignier for secretarial assistances. This work was supported by the GIP GREG, the BIOTECH program of the European Community and the Centre National de la Recherche Scientifique (C. N. R. S).

REFERENCES

1. Hodgman, T.C. (1988) Nature 333, 22–23.
2. Gorbalenya, A.E., Koonin, E.V., Donchenko, A.P. and Blinov, V.M. (1989) Nucleic Acids Res. 17, 4713–4730.
3. Troelstra, C., van Gool, A., de Wit, J., Vermeulen, W., Bootsma, D. and Hoeijmakers, J.H. (1992) Cell 71, 939–953.
4. Laurent, B.C., Treitel, M.A. and Carlson, M. (1991) Proc. Natl. Acad. Sci. USA 88, 2687–2691.
5. Davis, J.L., Kunisawa, R. and Thorner, J. (1992) Mol. Cell. Biol. 12, 1879–1892.
6. Tamkun, J.W., Deuring, R., Scott, M.P., Kissinger, M., Pattatucci, A.M., Kaufman, T.C. and Kennison, J.A. (1992) Cell 68, 561–572.
7. Girdham, C.H. and Glover, D.M. (1991) Gene Dev. 5, 1786–1799.
8. Mannhaupt, G., Stucka, R., Ehnle, S., Vetter, I. and Feldmann, H. (1992) Yeast 8, 397–408.
9. Emery, H.S., Schild, D.S.,Kellogg, D.E. and Mortimer, R.K. (1991) Gene 104, 103–106.
10. Johnson, R.E., Henderson, S.T., Petes, T.D., Prakash, S., Bankmann, M. and Prakash, L. (1992) Mol. Cell. Biol. 12, 3807–3818.
11. Laurent, B.C., Yang, X. and Carlson, M. (1992) Mol. Cell. Biol. 12, 1893–1902.
12. Clark, M.W., Zhong, W.W., Keng, T., Storms, R.K., Barton, A., Kaback, D.B. and Bussey, H. (1992) Yeast 8, 133–145.
13. Huang, M.E., Chuat, J.C., Thierry, A., Dujon, B. and Galibert, F. (1994) DNA Sequence in press.
14. Huang, M.E., Gobet, M. and Galibert, F. (1989) Meth. Mol. Cell. Biol. 1, 161–164.
15. Pearson, W.R. and Lipman, D.J. (1988) Proc. Natl. Acad. Sci. USA 85, 2444–2448.
16. Mc Ewen, J.E., Hong, K.H., Park, S. and Preciado, G.T. (1993) Curr. Genet. 23, 9–14.
17. Sung, P., Prakash, S. and Prakash, L. (1988) Genes Dev. 2, 1476–1485.
18. Walker, J.E., Saraste, M., Runswick, M.J. and Gay, N.J. (1982) EMBO J. 1, 945–951.
19. de Vos, A.N., Tong, L.M., Milburn, M.V., Natias, P.M., Jancarik, J., Noguchi, S., Nishimura, S., Miura, K., Ohtsuka, E. and Kim, S.H. (1988) Science 239, 888–893.
20. Linder, P., Lasco, P.F., Ashburner, M., Leroy, P., Nielson, K., Schier, J. and Slonimski, P.P. (1989) Nature 337, 121–122.
21. Tsuchiya, E., Uno, M., Kiguchi, A., Masuoka, K., Kanemori, Y., Okabe, S. and Mikayawa, T. (1992) EMBO J. 11, 4017–4026.

22. Okabe, I., Bailey, L.C., Attree, O., Srinivassan, S., Perkel, J.M., Laurent, B.C., Carlson, M., Nelson, D.L. and Nussbaum, R.L. (1992) Nucleic Acids Res. 20, 4649–4655.
23. van Duin, M., de Wit, J., Odijk, H., Westerveld, A., Yasui, A., Koken, M.H.M., Hoeijmaker, J.H.J. and Bootsma, D. (1986) Cell 44, 913–923.
24. Weber, C.A., Salazar, E.P., Stewart, S.A. and Thompson, L.H. (1990) EMBO J. 9, 1437–1447.
25. Terleth, C., Schenk, P., Poot, R., Bronwer, J. and van der Putte, P. (1990) Mol. Cell. Biol. 10, 4678–4684.
26. Friedberg, E.C., Siede, W. and Cooper, A.J. (1991) In The Molecular and Cellular Biology of the Yeast Saccharomyces. Genomes dynamics, Protein synthesis and energetics (Broach, J.R., Pringle, J.R. and Jones, E.W. Eds), pp. 147–192. Cold Spring Harbor Laboratory Press.