

Duplicated region of sequence similarity to the human *XRCC1* DNA repair gene in the *Schizosaccharomyces pombe rad4/cut5* gene

A.R.Lehmann
MRC Cell Mutation Unit, University of Sussex, Falmer, Brighton BN1 9RR, UK

Received August 31, 1993; Accepted September 22, 1993 EMBL accession no. X62676

Mutation in the *rad4* gene of *Schizosaccharomyces pombe* confers sensitivity to both UV and ionizing radiation damage, as well as a temperature-sensitive phenotype (1). We previously reported the cloning of the *rad4* gene (EMBL Accession no. X62676) and identified an ORF of 579 aa's (2), which contained two very short regions of sequence identity to the human *XRCC1* gene (3). The product of this human gene is known to be involved in the repair of strand-breaks following treatment with ionizing radiation and alkylating agents (4). We also showed that the *rad4* gene was essential for viability of the yeast cells. Recently Saka and Yanagida isolated the *S.pombe cut5* gene which is required for the onset of S-phase and for maintaining the dependency of mitosis on correct progression through the cell cycle (5). They found that the *cut5* and *rad4* genes were identical and they independently detected sequence similarities to *XRCC1*. They also showed that the *rad4/cut5* gene contained two additional short exons at the N-terminus. The correct length of the predicted protein is therefore 648 aa's (5). Close examination of the sequence alignments between the products of the complete *rad4/cut5* gene and the *XRCC1* gene reveals an unusual duplication in the *rad4/cut5* gene. Alignment of the first 58 amino acids, encoded by the two short exons, with aa's 314–371 of *XRCC1* (Figure 1) shows 41% sequence identity (64% similarity) including a stretch of 36 aa's corresponding to the second exon with 53% identity (69% similarity). The sequence similarity ends abruptly at the end of exon 2. However aa's 96–153 which are contained in the large (3rd) exon of *rad4* show 33% identity (53% similarity) to exactly the same region of *XRCC1* (aa's 314–371) (Figure 1). The two regions of *rad4/cut5* are 34% identical (53% similar) to each other. Further regions of similarity between *rad4/cut5* and *XRCC1* have been described by ourselves (2) and Saka and Yanagida (5). The duplicated region is likely to be important for the function of the *rad4/cut5* and *XRCC1* genes.

In order to determine if the duplicated region contained domains found in other proteins, exon 2 of the *rad4/cut5* protein was compared with other proteins in the PIR, SwissProt and GBTRAN databases using the PROSCAN and PATSCAN programmes of the DNASTAR package. No dramatic similarities were found. The highest scoring alignments were for diverse proteins with little in common and the conserved residues between these proteins and *rad4/cut5* were not the same as those conserved in the three sequences shown in Figure 1A. The most highly conserved region TKDVTHLIAG (residues 41–50) was also used in a PROSCAN search. The top scoring proteins with 6 or 7 identical (8 or 9 similar) residues were glycoprotein gp63 precursor, hemolysin activator protein, a yeast pre-mRNA splicing factor and ice nucleation protein. These again form a very diverse set of proteins, and the sequence similarity does not extend beyond these ten amino acids. It seems likely therefore that these sequence alignments with other proteins in the database are probably not of any biological significance and that the domains conserved in *rad4/cut5* and in *XRCC1* are not found in other proteins currently in the PIR, SwissProt and GBTRAN databases.

ACKNOWLEDGEMENT

I am indebted to Dr M.Yanagida for communication of his results on the *rad4/cut5* gene prior to publication.

REFERENCES

1. Duck,P., Nasim,A. and James,A.P. (1976) *J. Bacteriol.* **128**, 536–539.
2. Fenech,M., Carr,A.M., Murray,J.M., Watts,F.Z. and Lehmann,A.R. (1991) *Nucleic Acids Res.* **19**, 6737–6741.
3. Thompson,L.H., Brookman,K.W., Jones,N.J., Allen,S.A. and Carrano,A.V. (1990) *Mol. Cell. Biol.* **10**, 6160–6171.
4. Thompson,L.H., Brookman,K.W., Dillehay,L.E., Carrano,A.V., Mazrimas,J.A., Mooney,C.L. and Minkler,J.L. (1982) *Mutation Res.* **95**, 427–440.
5. Saka,Y. and Yanagida,M. (1993) *Cell* **74**, 383–393.

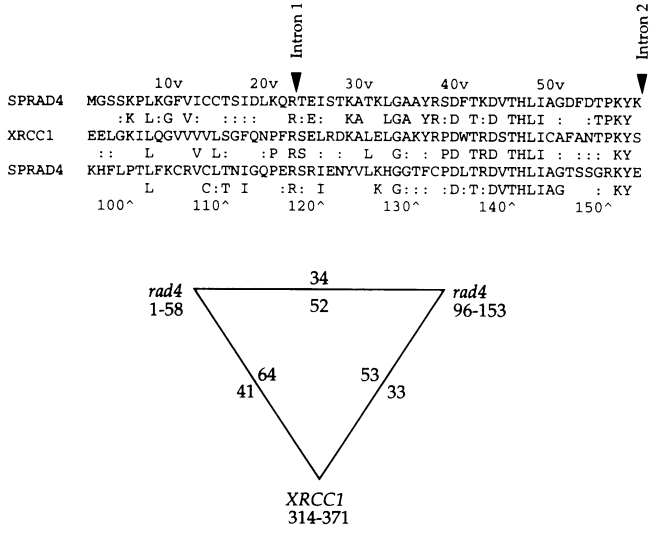

Figure 1. A. Alignments of products predicted from sequences of *S.pombe rad4* and human *XRCC1*. Aa's 1–58 of *rad4*, 314–370 of *XRCC1* and 96–153 of *rad4* are aligned. Amino acid identities and similarities are shown in the intervening lines, identity being indicated by the appropriate amino acid and conservative changes by a colon. On the bottom line identities and similarities between the two regions of *rad4* are indicated. **B.** 3-way comparison showing identities (%) outside triangle, similarities inside.