Identification of a chicken *RAD52* homologue suggests conservation of the *RAD52* recombination pathway throughout the evolution of higher eukaryotes

Olga Y. Bezzubova, Henning Schmidt¹, Kai Ostermann¹, Wolf-Dietrich Heyer² and Jean-Marie Buerstedde

Basel Institut of Immunology, Grenzacherstraße 487, CH-4005 Basel, Switzerland, ¹Institut of Genetics, TU Braunschweig, Spielmannstraße 7, D-38106 Braunschweig, Germany and ²Institute of General Microbiology, University of Bern, Baltzer-Strasse 4, CH-3012 Bern, Switzerland

Received August 27, 1993; Revised and Accepted November 16, 1993

EMBL accession no. U01047

ABSTRACT

Degenerate oligonucleotides encoding conserved regions of the Rad52 protein of *S.cerevisiae* and its homologue, the Rad22 protein of *S.pombe*, were used to clone a chicken *RAD52* counterpart by the polymerase chain reaction. Sequence comparison of the chicken and yeast proteins reveals a strongly conserved region between positions 40 and 178 of the chicken Rad52 sequence indicating that this part of the protein is under strong evolutionary pressure. The first 39 amino acids and the 3' end of the chicken Rad52 homologue does not share significant similarity with the yeast proteins. High abundance of the mRNA in testis makes it likely that the chicken Rad52 protein plays a role in meiotic recombination.

The steady rain of mutants, clones, and genetic and physical assays will keep us moving steadily ahead, as long as we can stay off the rock called *rad52* and avoid the whirlpool of DSBs.

Jeffrey N. Strathern (1)

INTRODUCTION

Genetic analyses of mutations conferring an increase in sensitivity to a variety of DNA-damaging agents in the yeast *Saccharomyces cerevisiae* have defined three epistasis groups. These groups were named after the *RAD3*, *RAD6*, and *RAD52* genes and are thought to represent three different DNA repair pathways (for reviews see ref. 2, 3). The *RAD52* group, made up of the genes *RAD50* to *RAD57* and most likely including *XRS2* (4) and *MRE11* (5), is required for resistance to ionizing radiation (6). Mutations in genes belonging to this group affect mitotic and meiotic homologous recombination leading to the postulate that the *RAD52* pathway repairs DNA damage by recombination (for review see ref. 2, 7). Mutations of the *RAD52* gene decrease the frequency of mitotic spontaneous recombination, particularly gene conversion (8–11). *rad52* mutants do not produce viable spores (12,8), which probably reflects their inability to complete induced meiotic recombination events (13, 14). The *RAD52* gene is also required for mating-type switching (15, 16) initiated by

a double-strand break (DSB) in the MAT locus (17). Moreover, plasmids containing DSBs recombine less efficiently with homologous chromosomal regions in *rad52* mutants as compared to *RAD52* wild-type strains (18). These properties of the *rad52* mutants and the observation that *RAD52* deficient cells accumulate DSBs following irradiation (19, 20) suggest an essential role of the Rad52 protein in the repair of DSBs.

Although the mechanism of action of the Rad52 protein is not understood, it has been reported that it binds directly to the *RAD51* gene product (21) which has been shown to be a RecA-like protein (21, 22). Recently chicken and mammalian homologues of the yeast *RAD51* gene have been cloned (23, 24). Exploiting the amino acid homology between the Rad52 protein of *S.cerevisiae* (25) and its homologue, the Rad22 protein of *S.pombe* (see accompanying paper) we have now identified a chicken *RAD52* gene counterpart. The sequence conservation of the Rad51 and Rad52 proteins between yeast and vertebrates indicates that the *RAD52* recombination pathway is conserved throughout eukaryotic evolution.

MATERIALS AND METHODS

Gene amplification

The sequences of the upstream and downstream degenerate primers are AA(A/G)CTIGGICCGA(A/G)TA and CT(C/T)C-T(G/A)TAICCIAT(G/A)C, where I is the symbol used for inosine, and alternative nucleotides are given in brackets. In addition, sequences corresponding to EcoRI restriction sites were attached at the 5' end of the primers to facilitate subcloning. Reverse transcription of the DT40 mRNA by random priming and the set-up of the PCR was done according to the GeneAmp[®] RNA PCR Kit (Perkin-Elmer Cetus) protocol. After an initial incubation of the reaction mix for 2 min. at 94°C, the following conditions were used for 30 cycles of PCR amplification: 5 sec. at 94°C, 15 sec. at 42°C, and 1 min. at 72°C (Protocol, AMS biotechnology Ltd). The PCR products were analysed on a 4% NuSieve[®] agarose gel (FMC BioProducts), purified and subcloned into M13 for sequence analysis.

Figure 1. a. Alignment of the predicted amino acid sequence (single-letter amino acid code) of the chicken Rad52 homologue with that of the *S.cerevisiae* Rad52 and *S.pombe* Rad22 proteins. Boxed are conserved stretches used for design of degenerate oligonucleotides (see text). Vertical lines indicate invariant amino acids, whereas colons or dots, structurally similar ones. Dashes are introduced to achieve the best alignment. **b.** Diagrammatic alignment of the three proteins. Amino acid sequences are illustrated by open boxes. The regions conserved among the Rad52 homologues are shaded.

The 5' RACE System Kit (GIBCO BRL) was used for the amplification of the 5' end of the chicken RAD52 message. The following three specific antisense primers were derived from the already known RAD52 cDNA sequence: G1(CCAGACACGT-AAGCACTTTCACGTCG), G2(ACCCGAGTAAGTCAGTG-AGTTGT) and G3(CAGTGTTCATTAGTCAGACCG). Poly(A)⁺ RNA from chicken testis was either primed with random primers or the G1 oligonucleotide. Purified cDNA was tailed with dCTP using terminal deoxynucleotidyl transferase and PCR amplified using the upstream anchor primer provided in the Kit and either the G1 primer (for random primed cDNA) or the nested G2 primer (for G1 primed cDNA). Aliquots of these PCR reactions were further amplified using the anchor primer together with the G3 primer. The PCR products were analysed on a 4% NuSieve^R agarose gel (FMC BioProducts). A single prominent band was seen in each reaction, purified and subcloned into M13 for sequencing. Single-stranded sequencing of M13 inserts was performed using Sequenase (United States Biochemical) according to the manufacturers instructions.

Characterisation of the chicken RAD52 cDNA

The cDNA library from the testis of 7 month old cock was provided by AMS biotechnology Ltd. Ten clones were isolated by standard filter hybridisation using the 252 bp PCR derived fragment as a probe. The cDNA inserts were subcloned into M13 phage for sequencing.

Sequence comparison

Sequence similarities were detected and aligned according to the GCG package version 7.1 (36). Amino acid sequence of the yeast *S.cerevisiae* Rad52 protein was taken from the ref. 25.

Southern blot analysis

10 mg of genomic DNA from an outbred white leghorn chicken was digested to completion by the indicated restriction enzymes. The digested DNA was separated on a 0.8% agarose and transferred to a Duralon UVTM membrane (Stratagene). The zoo-blot was purchased from Clontech. Hybridisation and washing was done according to the manufacturers instructions using low stringency washing conditions.

Northern blot analysis

The organs were taken from 3 week old (Fig.3, lane 1–6) and 7 month old (Fig.3, lane 7–10) chickens. Total RNA was isolated by extraction with guanidinium thiocyanate following by centrifugation through a caesium chloride gradient. Poly(A)⁺ RNA was then purified using PolyATtract^R mRNA Isolation System (Promega), 3 µg of Poly(A)⁺ RNA were applied to each lane and separated on a 1.5% agarose gel under denaturing conditions. Transfer to a Duralon UVTM membrane (Stratagene), hybridisation and washing was done according to the manufacturers instructions. The probe represented the antisense strand of the 252 bp PCR amplified fragment. The β-

Figure 2. Southern blot analysis using a chicken *RAD52* cDNA fragment as a probe. **a.** Chicken genomic DNA. Abbreviations for the used restriction enzymes are indicated above the lanes: B, BamHI; E, EcoRI; H, HindIII; P, PstI; S, SacI; X, XbaI. **b.** Zoo-blot. Lanes 1–9 contain, in order, EcoRI digested genomic DNA samples from human, monkey, rat, mouse, dog, cow, rabbit, chicken, and the yeast *Saccharomyces cerevisiae*. In the lanes 'M' the sizes of the different bands of a DNA marker are indicated in kilobases.

actin probe was derived from the 3' untranslated region of the chicken β -actin gene (37).

RESULTS

Cloning of a chicken *RAD52* homologue

Mutant phenotype analysis and sequence comparison suggested that the Rad22 protein of *S.pombe* and the Rad52 of *S.cerevisiae* are homologues (see accompanying paper). Strong sequence similarities between these proteins were limited to the region between position 54 and 192 of the Rad52 and the region between positions 29 and 169 of the Rad22. Two conserved stretches of six amino acids (see boxed aa in Fig.1a) were chosen for the design of degenerate oligonucleotides to clone the chicken *RAD52* homologue by the polymerase chain reaction (PCR). As a source of messenger RNA, we used the avian leukosis virus (ALV)-induced chicken B-cell line DT40 (26). Agarose gel electrophoresis of the PCR product revealed two prominent bands of about 700 bp and 250 bp, but only the smaller fragment was purified, as it had the size expected from the position of our primers in the yeast genes. Sequence analysis of the subcloned fragment revealed an open reading frame (ORF) of 252 bp with high homology to the yeast Rad52 and Rad22 proteins (Fig.1a).

Figure 3. Expression pattern of the chicken *RAD52* gene. Lane 1, Thymus; Lane 2, Bursa; Lane 3, Spleen of 3 week old chicks; Lane 4, Proventriculus; Lane 5, Brain; Lane 6, Heart; Lane 7, Liver; Lane 8, Testis; Lane 9, Ovary; Lane 10, Spleen of adult animals. **a.** hybridisation with the 252 bp *RAD52* probe. Exposure time was 3 hours. **b.** hybridisation with a β -actin probe. In the line marked 'M' the sizes of the different bands or a RNA size standard (RNA Ladder, GIBCO BRL) are indicated in kilobases.

To determine the whole coding sequence, we used the 252 bp fragment as a probe to screen a cDNA library derived from chicken testis. The insert of a hybridising clone contained an ORF of 1074 bp including part of the sequence of the previously amplified fragment. This ORF was followed by a 767 bp sequence with frequent stop codons in all reading frames ending with a Poly(A) tail. To obtain the sequence of the missing 5' end of the chicken *RAD52* mRNA we applied the 5' RACE method using three different oligonucleotides from the 5' end of our cDNA clone as specific antisense primers. The final PCR amplification with the most upstream antisense primer and the anchor primer yielded a single prominent band of about 250 bp size. This fragment was subcloned and a total of seven subclones from two independent cDNA amplifications were sequenced. These sequences extended the previously isolated cDNA clone for 192–212 bp offering an in-frame methionine start codon at position 32–35 of the longest 5' RACE sequence which was added to the sequence of the testis cDNA clone. Combining the sequence obtained by the 5' RACE method with the one of the cDNA clone gives a cDNA of 2052 bp total length, which is in agreement with the size of the most abundant mRNA species detected by Northern blotting analysis (see below).

Sequence comparison with the yeast Rad52 and Rad22 proteins

The ORF of 1,266 bp following the first methionine codon in the isolated cDNA predicts a protein of 422 amino acids. Comparison of the encoded polypeptide sequence with the amino acid sequence of the *S.cerevisiae* Rad52 protein revealed a region of strong homology between position 40 and 178 of the chicken sequence (Fig.1a and b). Within this part the chicken protein shares 51% of identical amino acids with the *S.cerevisiae* homologue. This region corresponds to the one conserved between the *S.cerevisiae* Rad52 and *S.pombe* Rad22 proteins (see accompanying paper), and the alignment of the chicken counterpart with the Rad22 yields an equally high value of 55% identical amino acids. Outside this conserved region the homology among the three proteins is low and probably insignificant, since gaps are introduced at different positions to allow for optimal

alignment (data not shown). These findings define a region within the Rad52 homologues, which is under strong evolutionary pressure and is most likely essential for their putative enzymatic function (Fig.1b). The proteins encoded by the three genes, *S.cerevisiae* RAD52, *S.pombe* rad22⁺, and chicken RAD52, are all characterised by a basic pI of ≥ 8.2 . However, no significant structural motifs could be identified from the protein sequences.

The chicken Rad52 is encoded by a single copy gene

Genomic DNA from chicken was digested with different enzymes and analysed by Southern blotting using a probe derived from the 252 bp PCR amplified fragment (Fig. 2a). The probe hybridised under low stringency conditions to a single fragment of each digest indicating that the RAD52 cDNA was encoded by a single copy gene. To determine whether homologous sequences exist in the genomes of other vertebrates, a zoo-blot was incubated with the same probe (Fig. 2b). The probe crosshybridised to limited number of fragments in the genomic DNA of each analysed vertebrate species suggesting that it corresponds to a sequence which is conserved during vertebrate evolution. On the other hand, no signal was detected in the genome of the yeast *S.cerevisiae*.

Expression of the chicken RAD52 homologue

To determine the expression pattern of the chicken RAD52 homologue, a probe derived from the 252 bp PCR amplified fragment was used for Northern blotting analysis. The probe hybridises to an RNA species of about 2,000 bp expressed at very high level in testis, at moderate levels in ovary, thymus and bursa, but only at low levels in brain, heart, liver, proventriculus and spleen (Fig.3a). Shorter exposure of the blot reveals several bands of larger size present in the lane with testis mRNA that may result from alternative splicing or polyadenylation of the chicken RAD52 transcript (data not shown).

DISCUSSION

Comparison of the chicken Rad52 counterpart with the yeast homologues indicates that the homology is restricted to the region of about 140 amino acids located in the N-terminus half of the protein. In this region 62% of all residues are either identical or strongly conserved (indicated as colons in Fig.1) among all three sequences. It remains speculative how this conserved sequence is related to the function of the Rad52-like proteins, since we were unable to detect significant sequence similarities with any other peptides in the databases. It is, however, possible that this region is important for the interaction with Rad51 (21). Further support of the view that the conserved region is critical for the function of the Rad52 protein comes from mutational analyses in *S.cerevisiae*. The *rad52-1* mutation, a missense point mutation in the codon 90 (25), causes a phenotype indistinguishable from the null phenotype (27). The *rad52-2* mutation, a missense point mutation in the codon 64(27), produces a phenotype very similar to the null phenotype, although it increases rather than decreases recombination (28). On the other hand, a RAD52 allele encoding only the N- terminal 65 % of the wild-type protein retains activities in repair and recombination (27).

The function of the chicken RAD52 homologue needs to be clarified, but its expression pattern is compatible with its role as a putative recombination enzyme. High abundance of the chicken RAD52 mRNA in testis and its moderate level in ovary

may reflect participation in meiotic recombination. Presence of the chicken RAD52 gene product in the chicken Bursa of Fabricius and thymus may be due to its involvement in lymphoid recombination such as gene conversion (29) and V(D)J recombination (30, 31).

Identification of RAD51 (23, 24) and RAD52 homologues in higher eukaryotes indicates that the RAD52 repair pathway is preserved throughout evolution. Although we report here only the characterisation of a chicken RAD52 homologue, the presence of crosshybridising fragments on the zoo-blot suggests that other vertebrate genomes contain genes homologous to the conserved chicken RAD52 sequence. Using the described degenerate primers, we indeed could amplify a sequence from mouse testis cDNA encoding an ORF which shares over 90% of identical amino acids with the chicken Rad52. Moreover, we have recently identified *bona fide* chicken and mouse homologues of the RAD54 gene, another member of the RAD52 group (unpublished data).

The mechanism of DNA repair seems to undergo little changes from yeast to mammals. It has been already shown that members of the RAD3 epistasis group involved with excision repair are strikingly conserved during evolution (32). Mutations in human homologues cause UV sensitivity associated with certain forms of Xeroderma pigmentosum (32). Chromosomal mapping of the members of the RAD52 pathway will clarify whether they are related to genetic diseases being accompanied by increased sensitivity to X-rays like the murine SCID mutation (33) or Ataxia telangiectasia in humans (34, 35).

ACKNOWLEDGEMENTS

We would like to thank Dr Vladimir I.Bashkirov for useful discussions and Drs Rodney Rothstein, Shunichi Takeda, and Michael V.Wiles for critical reading of the manuscript. W.-D.H. was supported by a career development grant (START 31-29254.90) and by grant 31-30202.90 from the Swiss National Science Foundation. The Basel Institute of Immunology was founded and is supported by F.Hoffmann-La Roche & Co. Ltd.

REFERENCES

1. Strathern, J.N. (1992) *Curr.Biol.*, **2**, 691–697.
2. Game, J.C. (1983) In Spencer, J.F.T., Spencer, D. and Smith, A.R.W. (eds.), *Yeast genetics: fundamental and applied aspects*. Springer Verlag, New York, pp. 109–137.
3. Haynes, R.M. and Kunz, B.A. (1981) In Strathern, J.N., Jones, E.W. and Broach, J.R. (eds.), *The molecular biology of the yeast Saccharomyces cerevisiae: Life cycle and inheritance*. Cold Spring Harbor Lab., Cold Spring Harbor, New York, pp. 371–414.
4. Ivanov, E.L., Korolev, V.G. and Fabre, F. (1992) *Genetics*, **132**, 651–664.
5. Ajimura, M., Leem, S.-H. and Ogawa, H. (1993) *Genetics*, **133**, 57–66.
6. Game, J.C. and Mortimer, R.K. (1974) *Mutat. Res.*, **24**, 281–292.
7. Resnick, M.A. (1987) In Moens, P.B. (ed.), *Meiosis*. Academic Press, New York, pp. 109–137.
8. Prakash, S., Prakash, L., Burke, W. and Monteleone, B.A. (1980) *Genetics*, **94**, 31–50.
9. Jackson, J.A. and Fink, G.R. (1981) *Nature*, **292**, 306–311.
10. Haber, J.E. and Hearn, M. (1985) *Genetics*, **111**, 7–22.
11. Bailis, A.M. and Rothstein, R. (1990) *Genetics*, **126**, 535–547.
12. Game, J.C., Zamb, T.J., Braun, R.J., Resnik, M. and Roth, R.M. (1980) *Genetics*, **94**, 51–68.
13. Borts, R.H., Lichten, M. and Haber, J.E. (1986) *Genetics*, **113**, 551–567.
14. Resnick, M.A., Nitiss, J., Edwards, C., Game, J.C. and Malone, R. (1986) *Genetics*, **113**, 531–550.
15. Malone, R.E. and Esposito, R.E. (1980) *Proc.Natl.Acad.Sci.USA*, **77**, 503–507.

16. Weiffenbach, B. and Haber, J.E. (1981) *Mol. Cell. Biol.*, **1**, 522–534.
17. Strathern, J.N., Klar, A.J.S., Hicks, J.K., Abraham, J.A., Ivy, J.M., Nasmyth, K.A. and McGill, C. (1982) *Cell*, **31**, 183–192.
18. Orr-Weaver, T.L., Szostak, J.W. and Rothstein, R.J. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 6358–6362.
19. Ho, K.S.Y. and Mortimer, R.K. (1975) *Mutat. Res.*, **30**, 327–334.
20. Resnick, M.A. and Martin, P. (1976) *Mol. Gen. Genet.*, **143**, 119–129.
21. Shinohara, A., Ogawa, H. and Ogawa, T. (1992) *Cell*, **69**, 457–470.
22. Ogawa, T., Yu, X., Shinohara, A. and Egelman, E.H. (1993) *Science*, **259**, 1896–1898.
23. Bezzubova, O., Shinohara, A., Mueller, R.G., Ogawa, H. and Buerstedde, J.-M. (1993) *Nucleic Acids Res.*, **21**, 1577–1580.
24. Shinohara, A., Ogawa, H., Matsuda, Y., Ushio, N., Ikeo, K. and Ogawa, T. (1993) *Nature Gen.*, **4**, 239–243.
25. Adzuma, K., Ogawa, H. and Ogawa, T. (1984) *Mol. Cell. Biol.*, **4**, 2735–2744.
26. Buerstedde, J.-M., Reynaud, C.-A., Humphries, E.H., Olson, W., Ewert, D.L. and Weill, J.-C. (1990) *EMBO J.*, **9**, 921–927.
27. Boundy-Mills, K.L. and Livingston, D.M. (1993) *Genetics*, **133**, 39–49.
28. Malone, R.E., Montelone, B.A., Edwards, C., Carney, K. and Hoekstra, M.F. (1988) *Curr. Genet.*, **14**, 211–223.
29. Thompson, C.B. (1992) *Trends Genet.*, **8**, 416–422.
30. Roth, D.B., Nakajima, P.B., Menetski, J.P., Bostma, M.J. and Gellert, M. (1992) *Cell*, **69**, 41–53.
31. Lieber, M.R. (1992) *Cell*, **70**, 873–876.
32. Hoeijmakers, J.H.J. and Bostma, D. (1990) *Cancer Cells*, **2**, 311–320.
33. Hendrickson, E.A., Qin, X.-Q., Schatz, D.G., Oettinger, M. and Weaver, D.T. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 4061–4065.
34. Thacker, J. (1989) *Bioessays*, **11**, 58–62.
35. Meyn, M.S. (1993) *Science*, **260**, 1327–1330.
36. Devereux, J., Haerberli, P. and Smithies, O. (1984) *Nucleic Acids Res.*, **12**, 397–395.
37. Buerstedde, J.-M. and Takeda, S. (1991) *Cell*, **67**, 179–181.