2094–2101 Nucleic Acids Research, 1994, Vol. 22, No. 11

Nonhomologous DNA end joining in *Schizosaccharomyces pombe* efficiently eliminates DNA double-strand-breaks from haploid sequences

Wolfgang Goedecke*, Petra Pfeiffer and Walter Vielmetter
Insitut for Genetics, Weyertal, 121, 50931 Cologne, Germany

Received February 11, 1994; Accepted May 5, 1994

ABSTRACT

Cells of higher eucaryotes are known to possess mechanisms of illegitimate recombination which promote the joining between nonhomologous ends of broken DNA and thus may serve as basic tools of double-strand-break (DSB) repair. Here we show that cells of the fission yeast *Schizosaccharomyces pombe* also contain activities of nonhomologous DNA end joining resembling the ones found in higher eucaryotes. Nonhomologous end joining activities were detected by transformation of linearized self-replicating plasmids in yeast cells employing a selection procedure which only propagates transformants carrying recircularized plasmid molecules. Linear plasmid substrates were generated by duplicate restriction cuts carrying either blunt ends or 3′ or 5′ protruding single strands (PSS) of 4 nt which were efficiently joined in any tested combination.Sequence analysis of joined products revealed that junctional sequences were shortened by 1 to 14 nt. Two mechanisms may account for junction formation (i) loss of terminal nucleotides from PSS tails to produce blunt ends which can be joined to abutting ends and (ii) interactions of DNA termini at patches of sequence homologies (1 – 4 bp) by formation of overlap intermediates which are subsequently processed. A general feature of the yeast joining system is that end joining can only be detected in the absence of sequence homology between the linear substrate and host genome. In the presence of homology, nonhomologous DNA end joining is efficiently competed by activities of homologous recombination.

INTRODUCTION

Events of illegitimate recombination may proceed by random DNA breakage and subsequent rejoining of broken DNA at nonhomologous sequence positions. While DNA double-strand-breaks (DSBs) may arise fortuitously by a variety of cell internal and external influences (1, 2), the rejoining of DNA demands specific cellular mechanisms able to link unrelated DNA ends into covalent duplex structures by reactions more complex than mere ligation (2, 3, 4).

Several different illegitimate recombination systems promoting e.g. viral integration into host genomes (5, 6), immunoglobulin rearrangements (3, 7, 8, 9) as well as chromosomal rearrangements (10, 11) have been proposed to proceed by similar end joining mechanisms. This notion was supported by the repeated detection of characteristic mutations such as base pair substitutions, small deletions or insertions found at junctional breakpoints of illegitimate recombination events in these different systems (2). The underlying nonhomologous DNA end joining mechanisms may also be involved in the elimination of lethal DNA DSBs and may provide an alternative to known pathways of homologous recombination outlined in the double-strand-break-repair model (12, 13, 14). While the latter mechanism strictly depends on extensive sequence homologies and reliably restores the original sequences at the breaks, illegitimate DSB removal via nonhomologous DNA end joining is independent of sequence homology but a mutagenic process (15, 16). Mechanisms of nonhomologous DNA end joining have been studied in detail *in vivo* in cultured mammalian cells (17, 16) and *in vitro* and *in vivo* in extracts from *Xenopus laevis* eggs (4, 15, 18, 19, 20, 21). These systems were shown to efficiently join nonhomologous terminus configurations containing either blunt or short protruding single strand (PSS) ends. Whereas only few junctional sequences exhibited larger deletions, the majority of junctions suggested a direct interaction of DNA ends tending to preserve the sequences of parental PSS ends.

While systems promoting nonhomologous DNA end joining predominate in higher eucaryotes comparable systems have been reported to occur only at low frequencies in *E.coli* and *S.cerevisiae* (2). Experimental evidence that the two distantly related yeast species *S.cerevisiae* and *S.pombe* can promote nonhomologous DNA end joining was derived from recircularisation events of transformed linearized plasmids and from illegitimate integration of exogenous DNA into chromosomes (22–29). The comparison of preinsertion sequences with the termini of linearized transformed plasmid molecules indicate the presence of short patches of sequence homologies at integration sites (22, 27, 28) resembling integration

*To whom correspondence should be addressed at: Institut Curie, Pavillon Troullet–Rossignol, 26 rue d'Ulm, 75231 Paris, France

events observed in mammalian cells (30, 31, 32). However, in contrast to the situation in higher eucaryotes, in yeasts nonhomologous integration events occur at exceedingly lower frequencies compared to integrations at homologous sites.

Here we present a systematic study of DNA end joining activities using various nonhomologous terminus configurations. We show that cells of the fission yeast *Schizosaccharomyces pombe* in fact contain an efficient system of nonhomologous DNA end joining which shares several common features with the above mentioned systems in vertebrate cells. However, the efficiency of this activity depends on the absence of homologous sequences flanking the original DSB.

MATERIALS AND METHODS

Strains

Schizosaccharomyces pombe h^- strains carrying the ura4-D18 or ura4-294 mutation (23) were gifts from J.Kohli (Bern) and used as hosts in transformation assays to detect activities of nonhomologous DNA end joining. The pEG1 and pEG114-i2 shuttle vectors were grown in *E.coli* recA$^-$ strain DH5a (BRL).

Plasmid construction

Standard cloning techniques were used to generate the recombinant plasmids pEG1 and pEG1D. Plasmid pEG1 (Fig. 1A) is based on the pSP65 plasmid (Promega) which contains a 1.2kb BamH1−SalI fragment of lambda-DNA in the equivalent sites of its polylinker (33). In this construct the HindIII fragment of the *S.pombe* ura4 gene (23) and the 1.2kb EcoRI fragment carrying the *S.pombe* ars1 element (34) were inserted at equivalent sites of the polylinker. pEG1D was obtained from pEG1 by excision of the lambda-fragment with SmaI and HincII and subsequent blunt end ligation. This plasmid is of similar size as the in *S.pombe* recircularized junctional plasmids and serves as control for the efficiency of transformation and joining.

pEG114 is a plasmid that was recovered from a joining experiment using the SacI−PstI terminus configuration. Due to a deletion this plasmid lacks all restriction sites of the original polylinker except the EcoRI site (Fig. 2A upper right corner). A 77mer double stranded synthetic oligonucleotide carrying a new polylinker flanked by *S.pombe* intron consensus sequences (35) was inserted by blunt end ligation in the single StuI site of the plasmid to yield pEG114-i2. The sequence of top strand of the artificial intron was: 5'-GTATGAATTCTCGAGCCCG-GGATCCGAGCTCCCTGCAGCCCGTCGACAAGCTTACT-AACTGGCCGTCGTTTTACTAG-3'. As a result of the cloning procedure the StuI site of the ura4 gene is disrupted (Fig. 4). The intron-containing allele, designated ura4-i77, is functional in *S.pombe* as demonstrated by efficient transformation of strains which carry the ura4-D18 mutation.

Substrate preparation

Linear DNA substrates carrying nonhomologous terminus configurations for the investigation of DNA end joining activities were prepared from pEG1 and pEG114-i2 by duplicate cuts with two different restriction enzymes in the plasmids' polylinkers. 5 different pEG1-derived substrates with 3'PSS/3'PSS (SacI−PstI), 5'PSS/5'PSS (BamHI−SalI), blunt/3'PSS (SmaI−PstI), 5'PSS/blunt (BamHI−HincII) and 5'PSS/3'PSS (BamHI−PstI) terminus configurations were generated by the excision of the diagnostic lambda-fragment. One pEG114-i2-derived substrate was obtained by double cutting within the intron sequence with SacI and PstI to produce a 3'PSS/3'PSS terminus configuration. The transformed fragments were purified by gel electrophoresis.

Transformation of *S.pombe* cells

In *S.pombe* transformation assays 100ng of linearized plasmids were employed following the protocol of Bröker (36) using standard fission yeast media as described in Gutz *et al.* (37).

Preparation of total nucleic acids from transformed yeast cells

Total DNA and RNA was isolated from *S.pombe* cultures grown under selective conditions by use of glass bead methods. Harvested cells from 50ml cultures were washed twice with sterile water and resuspended in 100µl lysis buffer (1% SDS; 5mM EDTA). After addition of 0.5ml glass beads (0.45 µm in diameter), samples were vortexed twice for 1min, interrupted by a 2min cooling step at 4°C. Each sample was extracted twice with phenol/chloroform (1:1), once with chloroform and precipitated with ethanol. Pellets containing total nucleic acids were resuspended in 50µl water and dialysed against water on

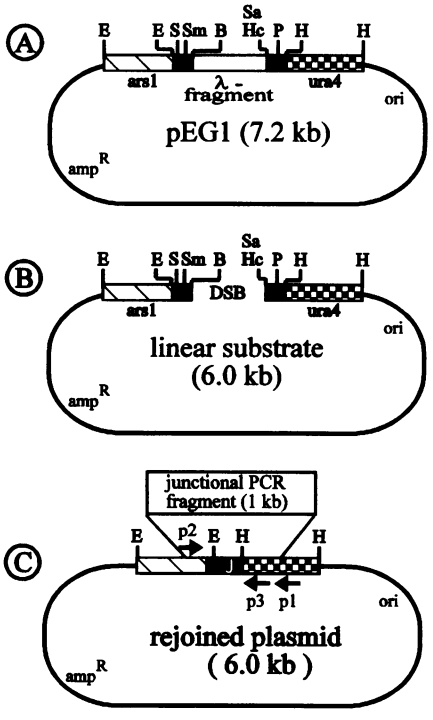


Figure 1. Strategy for substrate preparation and analysis of junctional sequences. **(A)** Structure of the self replicating pEG1 plasmid used for substrate preparation. The pSP65 (line) derived construct contains an ars1 fragment (hatched box) supporting extrachromosomal replication in *S.pombe*, the *S.pombe* ura4 gene (checkered box) as selection marker and a 1.2kb fragment of l-DNA (empty box) that is excised by successful double cleavage during substrate preparation (see Materials and Methods). The ars1 element was cloned into the single EcoRI site (E) and the ura4 gene into the HindIII site (H) of pSP65. An ampicillin resistance gene (amp^R) and a replication origin (ori) provide selection and propagation in *E.coli*. The pSP65 polylinker carries 6 unique restriction sites (SacI (S); SmaI (Sm); BamHI (B); SalI (Sa); HincII (Hc); PstI (P)) to generate various nonhomologous terminus configurations by restriction cleavage. **(B)** Excision of the diagnostic l-fragment during the preparation of linear substrates monitors successful double cutting with two different restriction enzymes. **(C)** Nonhomologous DNA end joining in *S.pombe* produces circular molecules containing a junction (J in black box) which may be amplified by PCR using primers p1 and p2 to yield a 1kb junctional fragment. Sequencing of the 3'−5'-strand was performed using primer p3.

0.2μm filters (Millipore) for at least h. One μl of this DNA solution employed in a PCR reaction is sufficient to yield approximately 1μg of junctional PCR fragment.

PCR and sequence analysis of junctional PCR fragments

For the amplification of junctions obtained from pEG1-derived substrates an ars1-specific p1: (5'-GCTCATATGTTATG-AGTATACCTA) and a ura4-specific p2: (5'-GGGTATTATA-CAAGGCCTCAAGAA) primer (Fig. 1C) were used in PCR assays at a final concentration of 25μM per primer. Junctions obtained from pEG114-i2 derived substrates were amplified using a ura4-specific primer pair p4:(5'-TCAGCTAGAGCTGAGGG-ATG); p5:(5'-ATATCTCTT GGCTTCGACAACAGG) to yield PCR fragments specific for the ura4-294 and ura4-i77 alleles (Fig. 5). Allele specific amplification was performed to obtain distinguishable fragments from each allele permitting sequence analysis of the two different ura4 alleles present in the same cell. The ura-294 allele was amplified with primer pa3: (5'-GTATTATACAAGGGTATGAATTC) and the ura4-i77 allele with pa2: (5'-GTATTATACAAGGCCTCAAAGAA-GTT). p5 served as reverse primer in both cases (Fig. 5).

Reaction conditions were 50mM KCl; 10mM Tris−HCl pH 8.4; 0.25mM $MgCl_2$; 0.01% (w/v) gelatine; 0.001% (v/v) Tween 20; 0.001% (v/v) NP40 and 2.5U Taq polymerase (Beckman) per 100μl sample. Usually 40 cycles (1min 92°C, 2min 56°C and 2min 72°C) were used. After purification of the linear PCR fragment from unincorporated primers by agarose gel electrophoresis followed by elution with DEAE (NA-45) membranes (Schleicher and Schüll) or the Quiaex glass milk method (Diagen), DNA was subjected to sequence analysis according to Sanger employing a T7 DNA polymerase sequencing kit (Pharmacia). Sequencing was performed according to the supplier's protocol using 2pmoles of sequencing primers per reaction. p3 (5'-TGAAAGATGTATGTAGATGA) was used for sequencing of pEG1-derived junctions, p6 (5'-GACGGTATTT-CCAATGT) for pEG114-i2 derived junctions. Both primers anneal approximately 100bp downstream of the junction.

RESULTS

Experimental system

To detect activities of nonhomologous DNA end joining in cells of S.*pombe* suitable linearized pEG1 plasmids bearing two nonhomologous restriction termini with either blunt ends or 4nt long 5'- or 3'PSS were introduced by transformation (Fig. 1A and B). The genetic constitution of plasmids and ura4-recipient cells was chosen so as to propagate only cells maintaining successfully recircularized and therefore replicating plasmids (Fig. 1B and C). To exclude that ura4$^+$ transformants resulting from homologous recombination between the plasmid and the host genome an *S.pombe* strain which carries the ura4-D18 deletion was used as recipient (23). This strain lacks the entire chromosomal copy of the ura4 selection marker, nereby necessitating that ura4$^+$ transformants contain the extrachromosomal copy of the ura4 gene on the plasmid which has been recircularized by nonhomologous DNA end joining. Single ura4$^+$clones were used to isolate rejoined plasmids whose junctions were amplified by the polymerase chain reaction (PCR) method and subjected to sequence analysis. This experimental approach is superior to amplification of junctional products by cloning in *E.coli* because it avoids possible artifacts resulting from bacterial DNA repair and recombination activities.

The extrachromosomal maintenance of transformed plasmids which is supported by the ars1 element, was verified by standard plasmid stability tests (38). After growth for 48h under nonselective conditions and plating on rich medium 80−90% of the cells of 31 independently picked clones had lost their plasmids as detected by loss of the ura4 marker after replica plating on selective medium. This demonstrates that the ura4$^+$ phenotype did not result from nonhomologous integration of the plasmid into the host genome.

Transformation of *S.pombe* cells with linear substrates was found to be almost as efficient as control transformations with the equally sized covalently closed circular (ccc) pEG1D plasmid (Tab. 1). Since the formation of ura4$^+$ clones after transformation with linear plasmid molecules depends on their successful recircularization we conclude, that nonhomologous DNA end joining is an efficient process in *S.pombe*. All 5 different linear substrates tested yielded similar transformation frequencies indicating that nonhomologous DNA end joining manages to process different terminus configurations with comparable efficiencies.

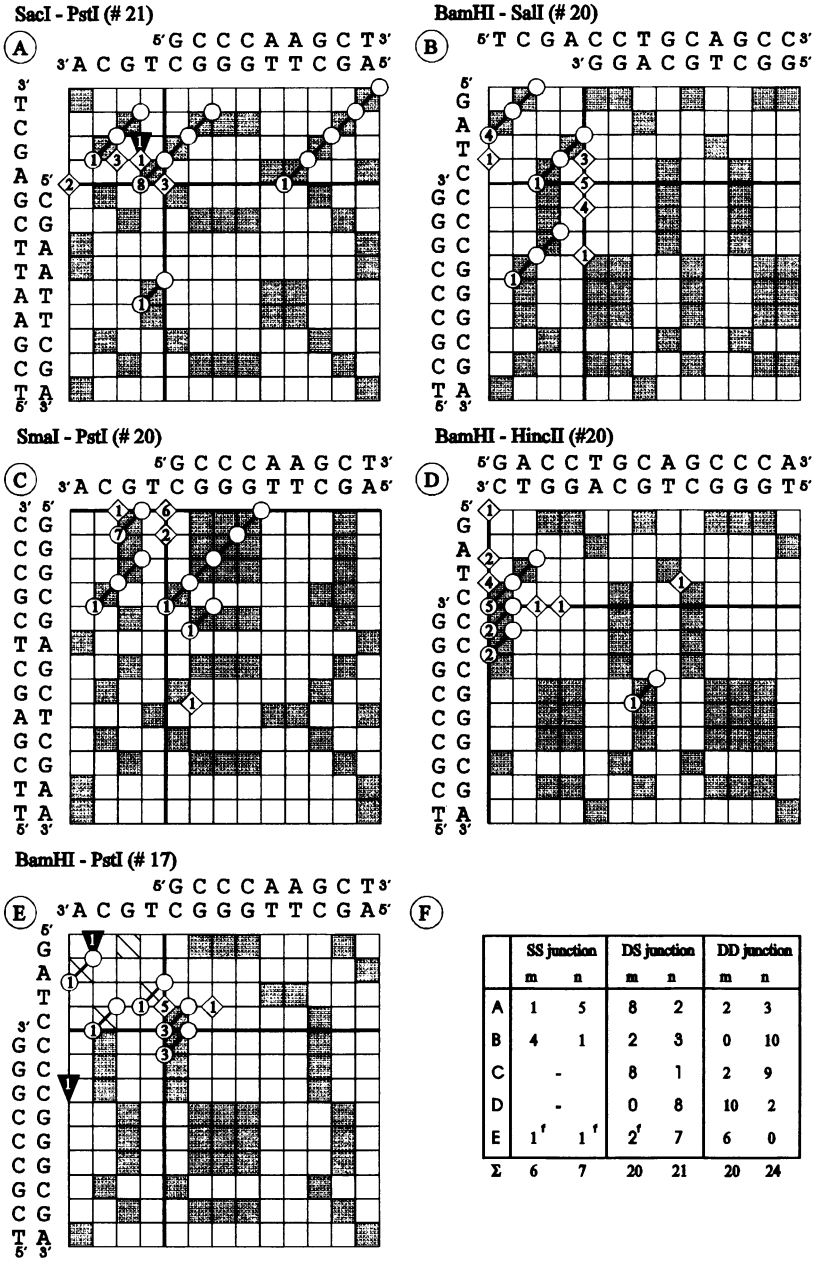
Junctional sequence analysis

The junctions of 106 independently picked ura4$^+$ clones were amplified by PCR using primers p1 and p2 (Fig. 1C). As judged from gel assays the resulting junctional PCR fragments were 1kb in length in agreement with the distance of primer positions (data not shown). The constant size of these junctional PCR fragments indicates that in general no extended deletions or insertions are generated during junction formation. Only 3 clones with reduced fragment sizes were observed due to extended deletions affecting one or the other terminus of the corresponding substrate. Junctions of 5 further ura4$^+$ clones failed to be amplified by the combined use of primers p1 and p2. In these cases, larger deletions probably account for the loss of at least one primer binding site. The relative low frequency of only 7.5% clones (8/106) with large deletions might be selectively biased by the fact, that junctional positions are located in close of vicinity to essential sequences of the ura4 gene and the ars1 element, functionally restricting deletions within these regions to a minimum. The 8 deletion clones were excluded from further analysis. The sequences of the remaining 98 junctional PCR fragments of 1kb unit length are summarized in the matrix schemes of Fig. 2. This kind of display has been used before (16) and was chosen here, because it allows one to directly correlate parental terminal sequences with breakpoint positions

Table 1. Transformation frequencies as obtained from 3 independent experiments using 100 ng substrate DNA and 10^8 *S.pombe* cells per sample

Substrate	Polarity of DNA termini	Transformation frequency [#/μg] 10^3
ccc	−	(4.0 ± 1.6)
SacI−PstI	3'PSS/3'PSS	(4.7 ± 1.0)
BamHI−SalI	5'PSS/5'PSS	(3.5 ± 0.7)
SmaI−PstI	blunt/3'PSS	(2.6 ± 1.3)
BamHI−HincII	5'PSS/blunt	(3.5 ± 0.5)
BamHI−PstI	5'PSS/3'PSS	(2.9 ± 1.4)

The data give the transfomation frequencies per μg DNA. The pEG1D plasmid used for control transformations is employed as (ccc) molecule. Substrate designations refer to the combination of restriction enzymes used for generation of the DSBs.



	SS junction		DS junction		DD junction	
	m	n	m	n	m	n
A	1	5	8	2	2	3
B	4	1	2	3	0	10
C	-		8	1	2	9
D	-		0	8	10	2
E	1[f]	1[f]	2[f]	7	6	0
Σ	6	7	20	21	20	24

Figure 2. Junctional sequence patterns derived by joining of five different terminus configurations given in the grids: **(A)** SacI−PstI (3′PSS/3′PSS); **(B)** BamHI−SalI (5′PSS/5′PSS); **(C)** SmaI−PstI (blunt/3′PSS); **(D)** BamHI−HincII (5′PSS/blunt); **(E)** BamHI−PstI (5′PSS/3′PSS). Nucleotide sequences of the original terminus configurations are drawn along the axes of the matrices with their strand polarities are indicated. Grid lines represent phosphodiester bonds between bases and grid squares represent base positions. Shaded squares indicate possible base pair positions between both ends. Hatched ones indicate, homology patches produced by fill-in synthesis. Open squares indicate mismatch positions between both termini. Bold grid lines assign blunt end positions of parental termini and thus subdivide matrices in 4 (panels A, B, E) or 2 quadrants (C, D) respectively. Junctional breakpoints are marked by diamonds and at homology patches as circles connected by a line representing the patch length. It is emphasized, that breakpoint positions within homology patches cannot be assigned more precisely than between the patch flanks, leading to ambiguities of the location of these break points. Junctions containing a single nucleotide insertion are marked by wedge symbols. Numaral indices within breakpoint symbols represent total numbers of junctions displaying this particular sequence pattern. Continuous junctional sequences can be determined from matrices by reading the vertical strand from bottom to top until a breakpoint position is encountered in the matrix grid to proceed from this position to the horizontal sequence reading from the left to the right. In case of patch homologies, any breakpoint assignment within the patch is for disposition without changing the sequence pattern. For instance, the sequence of the four clones in the upper left corner of panel B is read as GGGGATCGACCTGC. The table in panel **(F)** displays the distribution of breakpoints of each terminus configuration (A−E) in the 3 junction types. Numbers of breakpoints at base positions are listed below 'm', those at positions exhibiting no homology are listed below 'n'. 'f' indicates that the match position is only generated by fill-in of the 5′PSS.

in relation to base matches and other features of junctional sequences.

Inspection of junctional breakpoints

Analysis of the observed patterns of breakpoints (Fig. 2) reveals that most of them are positioned proximal of the ultimate nucleotides of the parental DSB ends. Thus, in most cases (97/98) terminal sequences are shortened by a few nucleotides during junction formation. In 3 exceptional rare cases single extra nucleotides are added at breakpoint positions of otherwise shortened terminal sequences (Fig. 2A and E). As shown in the histogram of Fig. 3 the frequency of terminal nucleotide loss depends on the structure of parental DNA termini. While more than half of all parental blunt ends remained intact, almost all parental 5′ and 3′PSS ends (96/98) were shortened, suggesting the presence of 5′ and 3′ exonucleolytic activities. Approximately 90% of these deletions affect terminal single strand portions (1−4 nt) wheras only 10% reach further into the adjacent duplex region (1−9bp). The frequency of nucleotide loss from PSS tails steadily increases from distal to proximal to reach its maximum at the blunt end position (approximately 30%) (Fig. 3). This suggests that the system prefers to eliminate PSS termini to produce blunt ends which appear to be much more resistant towards exonucleolytic attack than PSS tails. Parental blunt ends are rarely resected by more than 1bp (25%) so that larger deletions affecting adjacent duplex regions occur infrequently. The different stabilities of blunt ends versus PSS may explain the observed clustering of breakponts at blunt end positions indicated by bold lines (Fig. 2 A−E). The obtained junctions (Fig. 2 A−E) include those which result from joinings of PSS tails (SS junction), joinings of PSS tails with double-stranded regions (DS junction) and joinings of two double stranded regions (DD junction).

Apart from the frequent terminal nucleotide loss, another outstanding feature of junctional patterns is reflected by the fact that roughly one half of all junctions (46/98) exhibit short patch homologies of one to four basepairs in length while the other part highly coincides with blunt end positions of at least one partner terminus (18/21 DS junctions and 23/24 DD junctions).

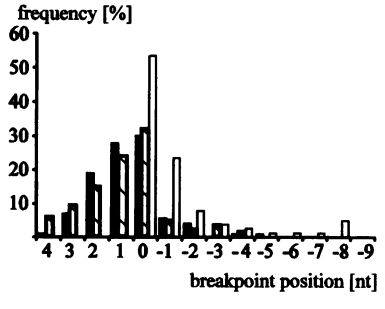


Figure 3. Statistic distribution of terminal nucleotide loss derived from the positions of junctional breakpoints shown in Fig. 2. The abscissa indicates nucleotide positions in the parental strands: positive values indicate nucleotide positions of parental PSS tails, 0 indicates original blunt end positions and PSS-duplex transitions, negative numbers indicate positions in the duplex region. Percentages displayed on the ordinate represent fractions of breakpoints located at a given nucleotide position which were calculated individually for parental 3′PSS (black bars), 5′PSS (hatched bars) and blunt ends (empty bars). Breakpoints of junctions exhibiting homologies are calculated as the ratio of the number of clones with a given junctional sequence to the number of possible breakpoints to take into account that the assignment to the parental termini is ambiguous.

Thus, the analysis of the distribution of junctional breakpoints revealed two types of junctions which may account for two different mechanisms is involved in junction formation. One mechanism that preferentially joins abutting ends (two blunt ends or a blunt and a PSS tail) whereas the other mechanism involves the recognition of patch homologies by overlap formation between fortuitously matching basepairs of interacting DNA termini. Our experiments indicate, that both junction types appear with about the same frequency in *S.pombe* cells.

Nonhomologous DNA end joining is rarely detected in the presence of extensive chromosomal sequence homologies

Due to the absence of extensive sequence homologies between the broken plasmid substrate and the host genome any interference between nonhomologous DNA end joining and homologous recombination were excluded in the experimental conditions used so far. To address the question whether nonhomologous DNA end joining can succeed in the presence of homologous chromosomal sequences we employed a linear substrate with 3′PSS/3′PSS terminus configuration (SacI−PstI) derived from pEG114-i2 to transform a *S.pombe* strain with the ura4-294 point mutation (Fig. 4). In contrast to the strain used before, this strain contains an inactivated chromosomal copy of the ura4 gene and therefore provides extended sequence homology between vector and host. The pEG114-i2 plasmid used for substrate preparation carries a ura4 allele modified by the insertion of a synthetic intron which contains unique restriction sites (Fig. 4). The intron containing ura4 allele, designated as ura4-i77, is functional and

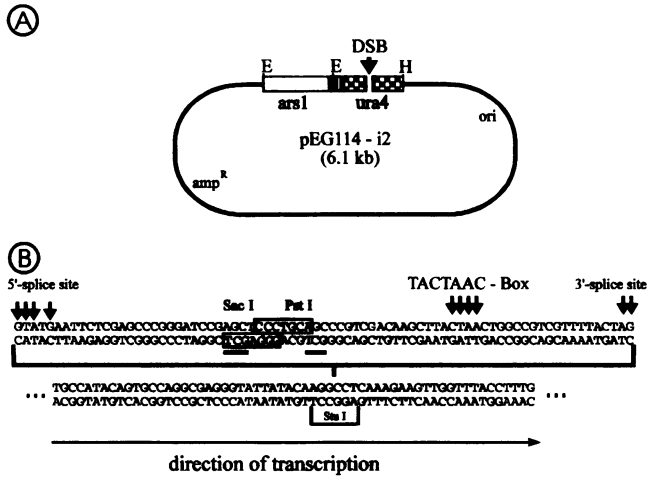


Figure 4. Structure of the pEG114-i2 plasmid used for analysis of end joining activities in the presence of extensive sequence homology located on the host genome. **(A)** pEG114-i2 is based on a junctional plasmid derived from joining of the SacI−PstI terminus configuration (for junctional sequence see upper right corner of matrix in Fig. 2A and materials and methods section). Symbols are as in Fig. 1A. The ura4 gene was modified by insertion of a 77mer oligonucleotide containing a new polylinker flanked by intron consensus sequences. Linear DNA substrates derived from this plasmid were generated by duplicate restriction cuts within the new polylinker. **(B)** Sequence of the synthetic 77mer intron which was cloned into the single StuI site of the ura4 gene. Arrows indicate the basepairs of the consensus sequence of the splice acceptor and donor site and the TACTAA-C-box (9). The boxed sequence is excised from the pEG114-i2 plasmid by double cutting with SacI and PstI. The underlined nucleotides indicate a 3 nucleotide patch homology coinciding with the breakpoints identified after sequencing the 401 bp fragment (Fig. 6, lane 10).

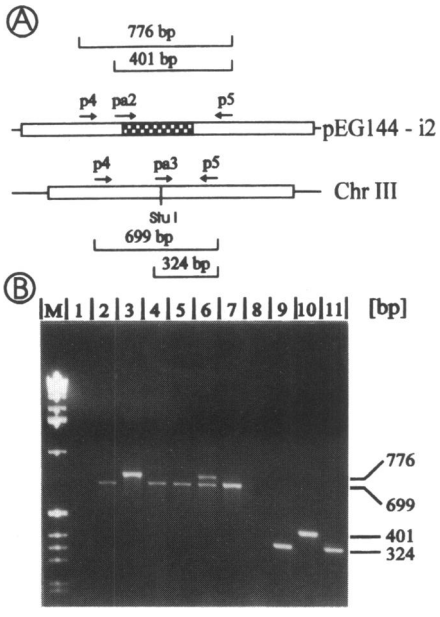


Figure 5. Genotype of partially diploid *S.pombe* cells and allele specific PCR analysis. **(A)** The endogenous inactivated copy of the ura4 gene in ura4-294 host cells on chromosome III provides extensive sequence homology for the repair of the extrachromosomal DSB by homologous recombination. The DSB is located in the intron (checkered) of the functional ura4 gene of the pEG114-i2 plasmid. To analyse by which mechanism DSBs were repaired we used a PCR assay, which can distinguish between the two ura4 alleles. Since both alleles differ in size by the intron (77bp), two different PCR fragments (776bp and 699bp) are obtained with the primer combination p4 and p5. Allele specific amplification of the intron containing ura4 allele is achieved with the primer combination pa2 and p5, which yields a 401bp fragment and of the chromosomal ura4 allele with the combination of pa3 and p5, which yields a 324bp fragment. **(B)** Agarose gel separation of PCR fragments (EthBr stained) amplified from ura4$^+$ clones obtained from transformation assays of the ura4-294 host with linearized pEG114-i2 plasmid (see panel A). (M) marker lane; Lanes 1–3 control experiments: (1) PCR amplification of untransformed ura4-D18 strain yields no product. (2) shows the 699bp product obtained from the chromosomal copy of the untransformed ura4-294 host. (3) PCR product of a rejoined pEG114-i2 plasmid obtained after transformation of the ura4-D18 strain. Lanes 4–8: Analysis of transformants using primers p4 and p5; (4) ura4$^+$ clone obtained after transformation of the ura4-294 strain with pEG114-i2 derived substrate. The size of the PCR product (699bp) indicates that DSB repair occurred by homologous recombination. This happened in 25 out of 26 clones. (5) the same clone as in (4) has lost its ura4$^+$ phenotype under nonselective conditions but still exhibits the 699bp fragment derived from the chromosomal ura4-249 copy, indicating the presence of the ura4$^+$marker maintained on the plasmid (6) PCR products (699bp and 776bp) obtained from a single out of 26 ura4$^+$ transformants indicating coexistence of the extrachromosomal ura4 copy which was repaired by DNA end joining in the presence of the chromosomal ura4-294 allele. (7) after nonselective growth of the clone analyzed in lane 6 the extrachromosomal copy is lost, whereas the chromosomal copy still yields the 699bp signal. Lanes 8–12: Allele specific PCR amplification of the clone tested in lane 4 yields no fragment with primer pair pa2 and p5 (8) and a 324bp fragment with pa3 and p5 (9). (11, 12) show the same experiment for the clone of lane 6: here both primer combinations yield distinguishable fragments of 324bp and 401bp. Sequence analysis of the 401bp fragment revealed a genuine SacI–PstI junction containing a 3bp patch homology at its breakpoint (Fig. 5).

stably propagated in *S.pombe* as demonstrated by efficient transformation of *S.pombe* strains carrying the ura4-D18 deletion and the ura4294 point mutation, respectively.

The ura4-i77 allele differs in size from the endogenous ura4-294 allele of the host which allows distinction of allelic DNAs by PCR analysis using the primer configurations outlined in Fig. 5A. PCR amplification of the ura4-gene of untransformed strains with the ura4-294 mutation results in a single PCR fragment of 699bp reflecting the distance between primers p4 and p5. Control transformations of covalently closed circular pEG114-i2 in the ura4-D18 deletion strain yield a larger fragment of 776bp reflecting the presence of the additional 77bp intron sequence. As expected, control amplifications of untransformed ura4-D18 strains with these primers yield no fragment.

This assay was used to test 26 independently formed ura4$^+$ clones which emerged after transformation of the host carrying the ura4-294 mutation with SacI–PstI linearized pEG114-i2 plasmid. Two types of transformants were detected by PCR analysis (Fig. 5B): a major type (25 clones), which exclusively yielded a unique band corresponding to the 699bp fragment and a minor type represented by a single clone (1/26), which yielded a significant double band corresponding to fragment lengths of 699bp and 776bp. To exclude that the ura4$^+$ phenotype of these transformants was not simply due to chromosomal integration of parts or the entire intact ura4 gene transformants were submitted to segregation analysis in nonselective medium. This resulted in the loss of the intact ura4$^+$ gene in the descendants of all 26 clones showing that the gene resided on the extrachromosomally maintained plasmid (data not shown). This therefore excludes both chromosomal integration and mutational reversion as cause for the generation of the ura4$^+$ phenotype. PCR analysis of the clones which had lost their plasmids after segregation in all tested cases exhibited only the 699bp fragment while the 776bp fragment was lacking.

To verify that the 776bp fragment really resulted from SacI–PstI joined molecules and not from contaminating uncut (ccc) plasmid allele specific primers pa2 and pa3 were employed for selective amplification of alleles ura4-i77 and ura4-294 (Fig. 5A). The allele specific primers were directed against the exon–intron border at the 5′-splice site of the ura4-i77 allele and the StuI site of the chromosomal ura4-294 allele, respectively. Thus, they in fact allow the selective amplification of alleles containing or lacking the intron. The longer PCR product (approx. 401bp) is diagnostic for the intron sequence whereas the shorter 324bp fragments are indicative of the StuI containing chromosomal ura4 copy (Fig. 5). Sequencing of the larger fragment revealed that the presence of a 3nt long patch homology was recognized in this junction (Fig. 4B) while the smaller ones contained intact StuI sites. No PCR product was obtained from amplification of ura4$^+$ transformants which otherwise yielded exclusively the 699bp p4 and p5 primed PCR product. This demonstrates that the smaller fragment size in these clones is not due to deletion events affecting most of the intron sequence.

DISCUSSION

In this study we examined activities of nonhomologous DNA end joining in the fission yeast *Schizosaccharomyces pombe*. Using a transformation assay we introduced linearized self replicating plasmid molecules into host cells of suitable genotype and showed that the *S.pombe* joining system manages to join any tested pair of nonhomologous restriction termini with comparable efficiencies. The two most obvious features deduced from junctional sequences are (i) the loss of terminal nucleotides and (ii) the presence of patch homologies at junctional breakpoints.

The frequent loss of terminal nucleotides is reflected in the distribution of junctional breakpoints. Only one single junction was detected in which both internal sequences had been

completely preserved (Fig. 2D). All other junctions contained small deletions ranging in size from 1−13nt. The clustering of junctional breakpoints at the borders of single to double strands in DS- and DD junctions (bold lines in matrix schemes) indicates that terminal nucleotide loss preferentially generates blunt ends. This is also reflected in the higher stability of parental blunt ends relative to PSS ends. The increased sensitivity towards nucleolytic degradation of PSS versus blunt ends is also indicated by the relative low frequency of deletions reaching into terminal double strands which rapidly decreases with increasing distance from the blunt end position. The fact that terminal nucleotides were removed from both 5′ and 3′PSS with comparable probabilities suggests the existence of 5′ and 3′ single strand specific exonucleolytic activities. The rare addition of single nucleotides to terminal sequences is consistent with previous findings in nonhomologous end joining systems, but the mechanism by which this addition occurs is not clear (39, 17).

The high coincidence of approximately 40% of all junctional breakpoints with blunt end positions (either original ones or generated by loss of entire parental PSS tails) suggests that these junctions arise from interactions of two blunt ends with each other or with a shortened PSS tail. The joining of a PSS to a blunt end requires fill-in DNA synthesis to yield continuous duplex DNA. While fill-in of 5′PSS is easily achieved due to the presence of a recessed 3′end, fill-in of 3′PSS must be necessarily primed at the 3′end of the juxtaposed partner terminus. This joining mechanism resembles the 'fill-in' mode found in *Xenopus* egg extracts which joins abutting terminus configurations (blunt/3′PSS; 5′PSS/3′PSS) by reliably preserving complete PSS sequences by fill-in DNA synthesis (4). In this joining system, fill-in of the 3′PSS was shown to be primed at the partner terminus before the ends had been linked by ligation. This topological problem was postulated to be solved by terminal DNA binding proteins, so-called 'alignment factors', which achieve precise structural alignment of juxtaposed noncovalently linked partner termini to facilitate subsequent fill-in and ligation reactions (4). A similar mechanism could also account for the formation of junctions involving blunt ends in the *S.pombe in vivo* joining system suggesting the existence of alignment factors.

One outstanding feature of the *S.pombe* system is the frequent nucleotide loss from PSS tails to preferentially produce blunt ends whereas only few larger deletions occur. This contrasts to the situation in some other *in vivo* systems where terminal degradation of DNA ends mostly leads to larger deletions affecting proximal duplex regions. In the *Xenopus in vitro* system, PSS tails are generally preserved and larger deletions are relatively rare events. This might reflect one of the differences between the *in vivo* and *in vitro* state of joining systems. While nucleases in the *in vivo* systems may develop their full activities at physiological temperatures they may be more restricted at the lower incubation temperatures and buffer conditions used in the *Xenopus in vitro* system. However, due to the frequent loss of 1−4nt from PSS tails, complete PSS sequences are only rarely preserved by the yeast joining system, which contrasts to the situation in the *Xenopus in vitro* system where PSS tails mostly remain intact and deletions affecting PSS tails and adjacent double strand regions occur only infrequently (15, 21). Like the *S.pombe* system other joining systems also exhibit higher rates of terminal nucleotide loss and larger deletions (11, 40−45). This might reflect one of the differences between the *in vivo* and *in vitro* state of joining systems. While nucleases in the *in vivo* systems may develop their full activities at physiological temperatures, they may be more restricted due to lower incubation temperatures and buffer conditions used in the *Xenopus in vitro* system.

The second feature of the *S.pombe* joining system is the frequent coincidence of junctional breakpoints with microhomology patches that range in size from 1−4bp. Due to the small size of the homology patches it could be argued that this coincidence is by chance. However, the frequency calculations confirm that the observed values of breakpoint coincidence with patch homologies lies above the expected level for chance coincidence which strongly suggests an active interaction of termini at positions of patch homologies by overlap formation. The presence of patch homologies is found not only in SS but also in DS- and DD junctions suggesting recognition of homology patches within duplex DNA. Two different mechanisms could account for the generation of these junctions: (i) homology patches could be exposed by premature single strand degradation of duplex DNA and subsequent annealing of the single strands at regions of microhomologies (resection−annealing model) (19, 45). (ii) the other model suggests single strand invasion in duplex DNA, annealing at regions of microhomologies followed by nucleolytic trimming of the displaced single strand(s) (strand invasion model). Although we cannot conclusively discern between these two possibilities some of our results rather support the strand invasion model.

Overlap formation induced by single fortuitous base matches should preferentially occur between two antiparallel PSS tails. However, only 17% of SS junction breakpoints of the 3′PSS/3′PSS terminus configuration coincide with a patch homology (Fig. 2A) whereas 80% of all SS junction breakpoints of the 5′PSS/5′PSS terminus configuration contain a patch homology (Fig. 2B). If the degradation model accounted for the process of overlap formation both antiparallel PSS configurations should yield similar frequencies of homology containing SS junctions. The observed difference could be due to the polarities of the involved PSS tails. In contrast to 3′PSS, 5′PSS can be readily filled in to yield blunt ends which then could participate as double strands in overlap formation with another PSS tail. Interestingly, several of the SS- and DS junctional breakpoints of the 5′PSS/3′PSS terminus configuration also exhibit homology patches which only could be formed after primary fill-in of the 5′PSS tail (Fig. 2E). That 3′PSS readily form base match induced overlaps with duplex DNA is seen in the DS junctions of the 3′PSS/3′PSS and blunt/3′PSS terminus configurations (Fig. 2A and C). All cases discussed here support the strand invasion model, in which duplex DNA is a prerequisite to induce the formation of overlaps. Formation of overlaps could be facilitated by a helicase mediated process. Deletions whose junctional breakpoints were located at positions of microhomologies were also detected in other joining systems (16, 40, 41, 43, 45). Unfortunately, in none of the systems the joining of 5′PSS/5′PSS terminus configurations was directly compared with that of 3′PSS/3′PSS configurations so that there are not enough data to support either of the proposed models.

Although *S.pombe* possesses an efficient system capable of promoting nonhomologous DNA end joining, this activity does not operate under partially diploid conditions. This was shown in transformation assays using a host strain providing chromosomal sequences that are homologous to the broken molecule. Under these conditions only 1 out of 26 ura4$^+$ transformants resulted from nonhomologous DNA end joining, whereas all other clones were generated by homologus DSB repair, converting the plasmid bound intron allele into the

sequence encoded by the chromosomal ura4 sequence. These data underscore the predominance of homologous recombination pathways over nonhomologous DNA end joining in *S.pombe*. A similar situation is observed in bacteria and *S.cerevisiae* (22, 40, 41, 45). In contrast to higher eucaryotes, these organisms have small genome sizes (2) and contain little repetitive and unessential DNA. This could explain the stronger requirement for high fidelity mechanisms which reliably restore the original sequences at the break and thus avoid mutations and genomic rearrangements jeopardizing cell survival. On the other hand, mechanisms of nonhomologous DNA end joining could represent alternative pathways to eliminate highly lethal DNA DSB if no extensive sequence homologies are available to facilitate homologous recombination. An attractive hypothesis is the coexistence of two double-strand-break repair pathways one of which eliminates DSBs from haploid sequences whereas the other deals with multicopy sequences.

ACKNOWLEDGEMENTS

This work was funded by a grant (Vi 23/1-2) from the Deutsche Forschungsgemeinschaft given to W.V.. P.P. was supported by a fellowship of the Lise Meitner program of the Ministerium für Wissenschaft und Forschung des Landes Nordrhein−Westfalen. W.Goedecke was granted by a Graduierten-stipendium of the Land Nordrhein−Westfalen. We thank Dr K.P.Reiners for helpful discussions and critically reading the manuscript.

REFERENCES

1. Bryant,P.E. (1986) In Simic,M.G., Grossman,L., Upton,A.C. (ed.) Mechanisms of DNA damage and repair, Plenum Press, New York pp. 171−180.
2. Roth,D. and Wilson,J.H. (1988) In Kucherlapati,R. and Smith,G. (ed.) Genetic Recombination, Am. Society for Microbiology, Washington D.C. 20006, pp. 621−653.
3. Lieber,M. (1992) Cell, 70, 873−876.
4. Thode,S., Schäfer,A., Pfeiffer,P. and Vielmetter,W. (1990) Cell, 60, 921−928.
5. Gahlmann,R. and Doerfler,W. (1983) Nucleic Acids Res., 11, 7347−7361.
6. Munz,P.L. and Young,C.S.H. (1991) Virology, 183, 160−169.
7. Gu,H., Förster,I. and Rajewsky,K. (1990) EMBO Journal, 9, 2133−2140.
8. Roth,D.B., Menetski,J.P., Nakajima,P.B., Bosma, M.J. and M.Gellert. (1992) Cell, 70, 983−991.
9. Roth,D.B., Chang,X.-B. and Wilson,J.H. (1989) Mol.Cell.Biol., 9, 3049−3057.
10. Meuth,M. (1989) In Berg,D.E. and Howe,M.M. (ed.), Mobile DNA, American Society for Microbiology, Washington D.C., pp. 833−860.
11. Winegar,R.A., Lutze,L.H., Rufer,J.T. and Morgan W.F. (1992) Mutagenesis, 7, 439−445.
12. Orr-Weaver,T.L., Szostak,J.W. and Rothstein,R.J. (1983) Methods Enzymol., 101,228−245.
13. Resnick,M.A. and Martin,P. (1976) Mol.Gen.Genet.,143, 119−129.
14. Szostak,J.W., Orr-Weaver,T.L., Rothstein,R.J. and Stahl,F.W. (1983) Cell, 33, 25−35.
15. Pfeiffer,P. and Vielmetter,W. (1988) Nucleic Acids Res., 16, 907−924.
16. Roth,D.B. and Wilson,J.H. (1986) Mol. Cell. Biol., 6, 4295−4304.
17. Roth,D.B., Porter,T.M. and Wilson,J.H. (1985) Mol.Cell.Biol., 5, 2599−2607.
18. Goedecke,W., Vielmetter,W. and Pfeiffer,P. (1992) Mol. Cell. Biol., 12, 811−816.
19. Lehmann,C.W. and Carroll,D. (1991) Proc.Natl.Acad.Sci.,88, 10840−10844.
20. Lehmann, C.W., Clemens,M., Worthylake, D.K., Trautmann, J.K. and Carroll, D. (1993) Mol. Cell. Biol., 13, 6897−906.
21. Pfeiffer,P., Thode,S., Hancke,J., and Vielmetter,W. (submitted) Mol.Cell. Biol..
22. Grimm,C. and Kohli,J. (1988) Mol. Gen. Genet., 215, 87−93.
23. Grimm,C., Kohli,J., Murray,J. and Maundrell,K. (1988) Mol. Gen. Genet., 215, 81−86.
24. Kunes,S., Botstein,D. and Fox,M. (1990) Genetics, 124, 67−80.
25. Kunes,S., Botstein,D. and Fox,M. (1985) J. Mol. Biol., 184, 375−387.
26. Orr-Weaver,T.L. and Szostak,J.W. (1983) Proc. Natl. Acad. Sci., 80, 4417−4421.
27. Schiestl,R.H., Dominska,M. and Petes,T. (1993) Mol. Cell. Biol., 13, 2697−2705.
28. Schiestl,R.H. and Petes,T.D. (1991) Proc. Natl. Acad. Sci., 88, 7585−7589.
29. Suzuki,K., Imai,Y., Yamashita,I. and Fukui,S. (1983) J. Bacteriol., 155, 747−754.
30. Brinster,R.L., Chen,H.Y., Trummbauer,M.E., Yagle,M.K. and Palmiter,R.D. (1985) Proc. Natl. Acad. Sci., 82, 4438−4442.
31. Folger,K.R., Wong,E.A., Wahl,G. and Capecchi,M.R. (1982) Mol. Cell. Biol., 2, 1372−1387.
32. Robins,D.M., Ripley,S., Henderson,A. and Axel,R. (1981) Cell, 23, 29−39.
33. Melton,D.A., Krieg,P.A., Rabagliati,M.R., Maniatis,T., Zinn,K. and Greene, M.R. (1984) Nucl. Acids Res., 12, 3278−3287.
34. Maundrell,K., Hutchinson,A. and Shall,S. (1988) EMBO Journal, 7, 2203−2209.
35. Russel,P. (1989) In Nasim,A., Young,P., Johnson,B.F. (ed.) Molecular biology of the fission yeast, Academic press, new York pp. 243−271.
36. Bröker,M. (1987) Biotechniques, 5, 516−517.
37. Gutz,H., Heslot,H., Leupold,U. and Loprieno,N. (1974) In King (ed.) Handbook of Genetics, Volume 1: Bacteria, Bacteriophages and Fungi Chapter 25 pp. 395−445.
38. Heyer,W.D., Sipiczki,M. and Kohli,J. (1986). Mol. Cell. Biol., 6, 80−89.
39. Clark,J.M. (1988) Nucl. Acids Res., 16, 9677−9688.
40. Conley,E.D., Saunders,V.A., Jackson,V. and Saunders,J.R. (1986) Nucl. Acids Res., 14, 8905−8917.
41. Conley,E.D., Saunders,V.A. and Saunders,J.R. (1986) Nucl. Acids Res., 14, 8905−8932.
42. Fairman,M.P., Johnson,A.P. and Thacker,J. (1993) Nucl. Acids Res., 16, 4145−4152.
43. King,J.S., Valcarcel,E.R., Rufer,J.T., Phillips,J.W. and Morgan,W. (1993) Nucl. Acids Res., 21, 1055−1059.
44. North,P., Ganesh,A. and Thacker,J. (1990) Nucl. Acids Res., 18, 6205−6210.
45. Thacker,J., Chalk,J., Ganesh,A. and North,P. (1993). Nucl. Acids Res., 20,6183−6188.