Nucleic Acids Research, Vol. 20, No. 17 4575–4579

Construction of human *XRCC1* minigenes that fully correct the CHO DNA repair mutant EM9

Keith W.Caldecott, James D.Tucker and Larry H.Thompson*
Biology and Biotechnology Research Program, Lawrence Livermore National Laboratory, PO Box 808, Livermore, CA 94551-9900, USA

Received May 7, 1992; Revised and Accepted August 5, 1992

ABSTRACT

The human gene that corrects the DNA repair defect of the CHO cell mutant EM9 is designated *XRCC1* and is the first human gene to be cloned that has an established role in DNA strand-break repair. In this study, either an *XRCC1* cosmid genomic fragment or synthetic oligonucleotides were ligated to an incomplete *XRCC1* cDNA to generate two full-length *XRCC1* cDNA constructs. The ability of these minigene constructs to restore normal levels of sister chromatid exchange (SCE) to EM9 upon transfection was demonstrated, and the transfectants grew at normal rates in selective medium that is fully toxic to EM9 cells. Constructs in which the *XRCC1* open reading frame (ORF) was transcribed from the SV40 early promoter or the genomic *XRCC1 native promoter* were compared in their efficiency of correction. EM9 transfectants derived from the SV40 promoter displayed fewer SCEs and lower sensitivity to CldUrd than either AA8 wild-type cells or transfectants containing the ORF transcribed from the native promoter.

INTRODUCTION

Living cells have adopted a variety of biochemical pathways to repair the damage that arises in DNA both spontaneously from metabolism and from exposure to chemical and physical agents present in the cellular environment. While our understanding of the biochemistry of DNA repair in *E.coli* is considerable, much less is known about repair pathways in eukaryotic cells, particularly mammalian cells. To date, only one human gene affecting ionizing radiation sensitivity has been cloned. This gene, designated *XRCC1* (*x-r*ay repair *c*ross *c*omplementing), was cloned on the basis of its ability to fully correct the repair defect present in the CHO mutant cell line, EM9 (1). EM9 is hypersensitive to ethyl methanesulfonate (10-fold) and ionizing radiation (1.8-fold), and is unable to grow in medium containing chlorodeoxyuridine (CldUrd) under conditions that produce 20% replacement of dThy by ClUra in the DNA (2,3). EM9 repairs γ-ray and ethyl methanesulfonate induced single-strand breaks at a reduced rate and exhibits a ten-fold increase in the occurrence of SCE (1,4). High baseline SCEs are also seen in cells isolated from individuals with Bloom's Syndrome (BS), an autosomal recessive disorder conferring a predisposition to cancer and represented by a single complementation group (5,6).

XRCC1 is 33 kb in length and encodes a 2.2 kb transcript and corresponding putative protein of 633 amino acids (~ 69.5 kDa) (1). A pcD2 human fibroblast cDNA library screened with *XRCC1* cosmid sequences yielded six positive clones, the largest of which contained an insert of 2.3 kb and was termed pXR1-30. This clone only partially corrected the repair defect in EM9 cells, as indicated by the 3-fold elevated SCEs in pXR1-30 transfectants of EM9, compared with wild-type AA8 cells, and by the abnormally slow rate at which these transformants grew under CldUrd selection. A comparison of nucleotide sequences derived from cosmid genomic clones and pXR1-30 showed that the cDNA was missing 26 bp from the 5′ end of the open reading frame (ORF), including the putative ATG initiation codon. It was suggested that partial correction in transfected cells might result from abnormal XRCC1 protein having a truncated N-terminus, resulting from initiation at an internal ATG site within the cDNA, or protein having an abnormal N-terminus resulting from initiation at the ATG that is present in the pcD2 vector upstream of the cDNA insert (1).

Our ultimate goal is to isolate XRCC1 protein and address its biochemical role in DNA strand-break repair. A complete, fully functional *XRCC1* cDNA minigene should facilitate the rapid purification of large quantities of protein from prokaryotic and/or eukaryotic expression systems. Therefore, in the present study we reconstructed the full length *XRCC1* ORF by both oligonucleotide addition and by the addition of a cosmid fragment that spans the region missing from the cDNA. The functionality of these minigene constructs was assessed in transfected EM9 cells.

MATERIALS AND METHODS

Cell lines and culture conditions

Isolation of the CHO EM9 mutant from parental AA8 cells, as well as the conditions used for the propagation of AA8, EM9, and transfectants of EM9, were described previously (1,4).

* To whom correspondence should be addressed

Growth curves were obtained by culturing cells in suspension in screw-cap tubes on a revolving drum at 40 rpm and determining cell concentrations on a daily basis.

Plasmids and oligonucleotides

The cDNA clone pXR1-30, which contains a truncated *XRCC1* insert, was isolated from the pcD2 human expression library as previously described (1). The expression vector pcD2E is a derivative of pcD2 that contains an *Eco*RI restriction site for insertion of cDNA minigenes (Dr. Christine Weber, personal comm.). pcD2E contains only one of the two introns that are present in pcD2 between the SV40 early promoter and the oligodG/oligodC tract on the 5′ side of the cDNA insert. Complementary 54-mer and 46-mer oligonucleotides were generated on an automated synthesizer (see Fig. 1, panel B for nucleotide sequences). The oligonucleotides were purified with NENSORB cartridges as directed by the manufacturer (DuPont/NEN Research Products) and then annealed together, resulting in a double stranded 54-mer with *Eco*RI and *Pst*I 5′ and 3′ overhangs, respectively.

Transfection of plasmids by electroporation

Plasmid DNA was introduced into EM9 cells at a density of 2×10^7 cells in 1 ml of electroporation buffer (specified by the supplier) using a Biorad Cell Porator set at 300 V and 1600 μF. A single type of plasmid was used for each transfection, and, since the plasmids differed in size (4.4−7.4 kb), it was ensured that plasmid DNA concentrations were equimolar (4−6 μg/ml) for each of the transfections. After electroporation, cells were incubated 24−48 h for phenotypic expression, then replated in triplicate into 10 cm petri dishes. For determining transfection frequencies in Geneticin (G418 sulfate, Gibco BRL) at 1.5 mg/ml, cells were plated at 5×10^4 cells per dish, and for determining the frequency of CldUrd resistance, cells were plated at $2-5\times10^5$ cells per dish in medium that additionally contained 32 μM thymidine, 200 μM deoxycytidine, 10 μM FldUrd, and 8 μM CldUrd. After ≈10 days incubation, individual macroscopic colonies were isolated using Pipetman tips. In the transfections with plasmids pcD2EX and pBSX, pooled cultures of ≈250 and ≈300 colonies, respectively, were also produced.

Growth rate in chlorodeoxyuridine

Suspension cultures in exponential phase were diluted to 1×10^4 cells/ml in both normal and CldUrd selective medium (described above). The average doubling time was monitored at 24−48 h intervals for a total period of 200 h. The relative growth rate in CldUrd was calculated for each cell line by dividing the doubling time in the absence of CldUrd by that in its presence. Cultures were diluted to 1×10^4 cells/ml in prewarmed medium to continue the experiment when cell densities approached 2×10^5 cells/ml (late exponential phase).

Plating efficiency in chlorodeoxyuridine

Cells were plated in normal medium at a density of 200 or 300 cells per 10-cm dish and in CldUrd selective medium (8−32 mM thymidine, 8−32 μM CldUrd, 200 μM deoxycytidine, and 10μM FldUrd) at 10^2, 10^3, and 10^4 cells per 10-cm dish. The CldUrd and thymidine concentrations together equalled 40 μM. When colonies growing in CldUrd reached macroscopic size (10−14 days), they were fixed, stained, and counted using a criterion of 50 cells per colony. Cells plated in normal medium were fixed after 9−10 days to determine plating efficiency.

SCE determinations

Suspension cultures were inoculated at a density of 1×10^5 cells/ml and grown in the presence of 10 μM BrdUrd for 23 h, after which Colcemid was added to 0.1 μg/ml. After an additional 5 h, the cells were harvested according to standard techniques and prepared for SCE analysis as described (7,8). Fifty well-spread second division cells with clearly differentiated sister chromatids were scored from each culture.

RESULTS AND DISCUSSION

Construction of minigenes from full-length *XRCC1* cDNA

The human cDNA in pXR1-30 is missing 26 bp of coding sequence from the 5′ terminus of the *XRCC1* ORF, including the putative ATG initiation codon (1). The sequence of the missing 26 bp, as well as the immediate upstream region, was determined from a cosmid genomic clone and was used to design two complementary oligonucleotides that spanned the missing region (Fig. 1, panel B). The longer oligonucleotide (coding strand) extends 37 bp downstream from the presumed ATG initiation codon into a *Pst*I site, and 17 bp upstream from the ATG to provide a segment of leader sequence possibly important for translation (9). A mutant form of the duplex oligonucleotide that was missing a single base pair was also generated for use in control experiments. This duplex was used to construct an *XRCC1* minigene that contained a single base pair deletion at the position shown in Fig. 1, panel B. This Δ 1 bp minigene was ligated into pcD2E as described above to generate plasmid pcD2EXδ.

In addition to the two constructs described above (Fig. 1, panel E), a third construct was generated in which the *XRCC1* ORF was placed immediately downstream of the native promoter. A 2.2 kb *Pst*I fragment was isolated from an *XRCC1* cosmid clone and was ligated to the *Pst*I/*Eco*RI cDNA fragment described above. The cosmid fragment contained the first 37 bp of the ORF and also included the putative promoter elements of the *XRCC1* gene (1). The resulting 4.4 kb fragment containing the full length *XRCC1* ORF was ligated into pBluescript II KS (+) (Stratagene) to generate pBSX (Fig. 1, panel F), a construct containing no introns.

Frequency of corrected EM9 transfectants using the minigene constructs

The constructs described above were transfected into EM9 cells, and repair proficient transformants were selected in CldUrd medium under conditions which produce 20% replacement of thymine by ClUra during DNA replication (see Materials and Methods; 2). While this selection does not retard the growth of AA8 cells, it is highly toxic to EM9 (survival $<10^{-7}$). The frequency at which Geneticin-resistant colonies arose was similar for the three pcD2E-based plasmids (Table 1). No pcD2E transfectants grew in Geneticin/CldUrd selective medium although slow-growing colonies developed at a low frequency when pcD2EXd was transfected into EM9 cells. Slow-growing transformants similarly arose at a low frequency when pXR1-30 was transfected into EM9 cells (1). The *XRCC1* cDNA insert in pXR1-30 is missing the first 26 bp of coding sequence, including the initiation codon, while the 19th coding base pair in pcD2EXd is deleted, resulting in a frame shift. Thus, it is surprising that any transformants arose in transfection experiments involving these constructs. It is possible that translation of the

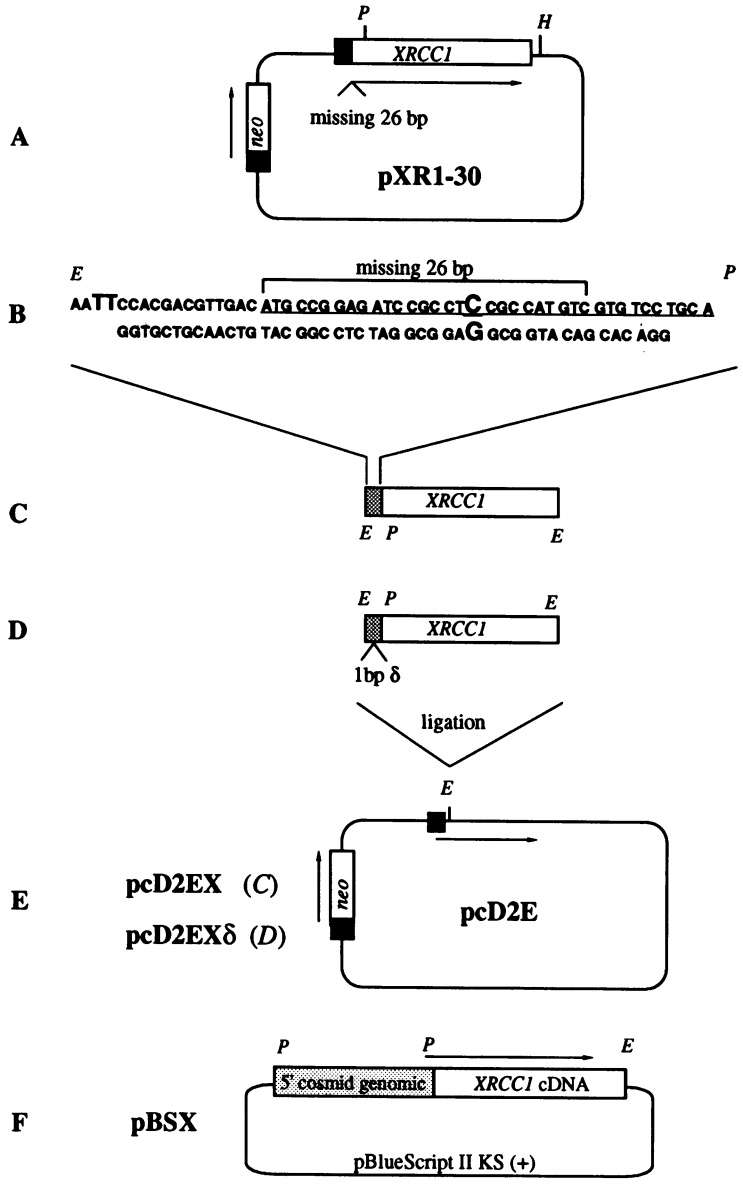

Figure 1. Schematic showing the construction of the three *XRCC1* minigene expression constructs used in this study. (A) pXR1-30, the incomplete *XRCC1* cDNA clone previously isolated from the pcD2 library. (B) The two synthetic oligonucleotides used to reconstruct the full-length *XRCC1* ORF. The upper oligonucleotide encodes 17 bp of leader sequence and the first 37 bp of ORF (underlined), including the 26 bp missing from pXR1-30. The 17 bp leader sequence differs from the native *XRCC1* leader sequence by 2 bp at the 5' end of the oligonucleotide, where GpC was replaced by TpT in the oligonucleotide to generate an *Eco*RI terminus (highlighted). A second oligonucleotide duplex was produced that was lacking the deoxycytidine nucleoside highlighted in **B**. Both oligonucleotide duplexes were ligated to the 2.15 kb *Pst*I/*Hinc*II fragment shown in *A*, which spans the remainder of the *XRCC1* ORF, after the *Hinc*II terminus of this fragment had been converted to an *Eco*RI terminus by subcloning through pBluescribe (Stratagene). The two resulting *XRCC1* cDNA cassettes, shown in **C** and **D**, were inserted into the *Eco*RI site of the expression vector pcD2E to generate the two expression constructs pcD2EX and pcD2EXδ, respectively (E). (F) A 2.2 kb *Pst*I DNA fragment (shaded box), which spanned the 26 bp missing from pXR1-30, was isolated from an *XRCC1* cosmid clone and was ligated to the *Pst*I/*Eco*RI cDNA fragment described above to generate a 4.4 kb minigene in which the *XRCC1* cDNA is situated downstream of its native promoter and leader sequences. This minigene cassette was inserted into the plasmid pBlueScript II KS (+) to generate the third expression construct, pBSX. Symbols: Arrows indicate the direction in which the *neo* gene (conferring Geneticin resistance) and *XRCC1* gene are transcribed from the SV40 early promoters (black boxes). *P*, *H* and *E* represent sites for the restriction endonucleases *Pst*I, *Hinc*II and *Eco*RI, respectively.

XRCC1 mRNAs encoded by these constructs initiates at an internal AUG (the first candidate site is at position 280 in the ORF) producing a partially active protein, or that they can occasionally integrate into an active gene downstream of the deletion, such that the truncated *XRCC1* ORF is in-frame with the host ORF and is consequently translated as a fusion protein. In any case, these results suggest that *XRCC1* proteins possessing an abnormal N-terminus can be partially active.

Transfections with either pcD2EX or pBSX resulted in a high frequency of CldUrd resistant colonies. For pcD2EX, a comparison of the transfection frequency in single and double selection suggests that 64% of all Geneticin-resistant colonies were also resistant to CldUrd. Assuming that there is generally only one randomly integrated copy of the initially circular pcD2EX molecules in each transfectant, this percentage is close to the expected result because the site of integration in pcD2EX

Table 1. Transformation efficiency for phenotypic correction of EM9 cells with *XRCC1* minigene constructs[a]

Plasmid	Geneticin^r	Geneticin^r + CldUrd^r
pcD2E[b]	1.2×10^{-3}	$< 5\times10^{-7}$
pcD2EXd	1.5×10^{-3}	3×10^{-6}
pcD2EX	1.1×10^{-3}	7×10^{-4}
pBSX[c]		9×10^{-4}

EM9 cells were transfected as described in Materials and Methods.
[a]Selection conditions are described in Materials and Methods.
[b]Average of two experiments
[c]Selection in CldUrd only

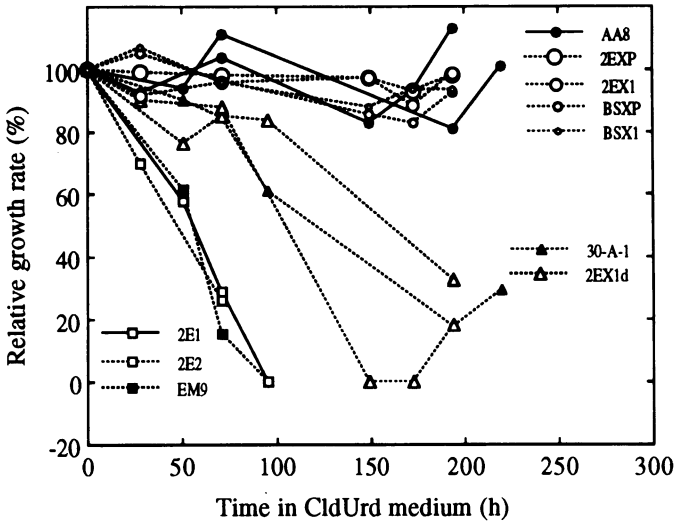

Figure 2. Growth rates of EM9 cDNA transfectants in CldUrd medium. Cell lines were grown in suspension culture in the presence and absence of CldUrd, and their average doubling time during 24–48 h periods was calculated over a total of 200 h. Each data point represents the relative growth rate of the indicated cell line during the preceding 24–48 h period. A single culture is shown for each cell line, with the exception of AA8 and 2EX1δ, for which two determinations are shown. Symbols are as defined in the figure.

should fall outside of the *XRCC1* sequence approximately 60% of the time in such transfectants. The transfection frequency with pBSX, which contains no *neo* gene and has the native *XRCC1* promoter, was very similar to that for pcD2EX. Overall, the results in Table 1 suggest that both pcD2EX and pBSX contain *XRCC1* minigenes that function efficiently.

Growth rate of EM9 transfectants in CldUrd selective medium

It was previously reported that the CldUrd-resistant EM9 transfectants that arise at a low frequency in transfections with pXR1-30 are only partially corrected ($\approx$ 80%) with respect to their SCE level as well as their growth rate in CldUrd (1). To determine whether pcD2EX and pBSX could fully correct EM9 cells, independent transfectants and their pooled populations were isolated from transfections with pXR1-30, pcD2E, pcD2EXd, pcD2EX, and pBSX. The cultures were expanded and compared for their growth rate in CldUrd selective medium as well as for their level of SCE.

The results shown in Fig. 2 indicate that the doubling time of EM9 cells grown in suspension rapidly declined in the presence of CldUrd, with cessation of growth after 100 h. A similar decline in doubling time in the presence of CldUrd was observed for the two independent pcD2E transfectants 2E1 and 2E2. In contrast, the doubling time of AA8 in CldUrd remained the same

Table 2. SCE frequencies in *XRCC1* minigene transfectants of EM9

Cell Line	DNA Construct	SCEs/cell (S.E.)[a]	% correction[b]
AA8	–	8.9	100
EM9	–	82.1	0
2E1	pcD2E	69.6 (2.2)	17
2E2	pcD2E	89.3 (2.7)	0
30-A-1	pXR1-30	28.1 (1.0)	74
2EX1δ	pcD2EXd	21.7 (0.8)	83
2EXP	pcD2EX	6.7 (1.1)	102
2EX1	pcD2EX	6.9 (0.5)[c]	102
2EX2	pcD2EX	6.3 (0.4)[c]	102
2EX3	pcD2EX	7.1 (0.4)[c]	101
2EXPc1	pcD2EX	6.4 (0.4)[c]	104[d]
2EXPc2	pcD2EX	6.8 (0.5)[c]	104[d]
BSXP	pBSX	10.5 (0.9)	97
BSX1	pBSX	9.5 (0.6)	98
BSX2	pBSX	18.2 (0.9)	88
BSX3	pBSX	11.6 (0.5)	96

[a] Each mean and its standard error (in parentheses) are based on measurements from 50 second-division cells of each culture in one experiment, except for AA8 (average values of 3 experiments: 8.9 +/− 0.39, 8.2 +/− 0.39, 9.6 +/− 0.50) and for EM9 (average of 2 experiments: 74.8 +/− 1.9, 89.4 +/− 3.2).
[b] Overcorrection was calculated using the values for AA8 and EM9 cells measured in the same experiment as a transfectant. The degree of correction conferred by different minigene constructs was calculated using the following formula:
% correction in cell line Y = $[SCE_{EM9} - SCE_Y]\times100/[SCE_{EM9} - SCE_{AA8}]$
[c] Values lower than that of AA8 measured within the experiment ($P < 0.05$) based on a 1-tailed t-test.
[d] 2EXPc1 and 2EXPc2 were done in a separate experiment from the other 2EX transfectants, using only AA8 as a control. Since EM9 was not analyzed in this experiment, overcorrection was calculated based on the average EM9 value shown and the AA8 value from the experiment (9.6).

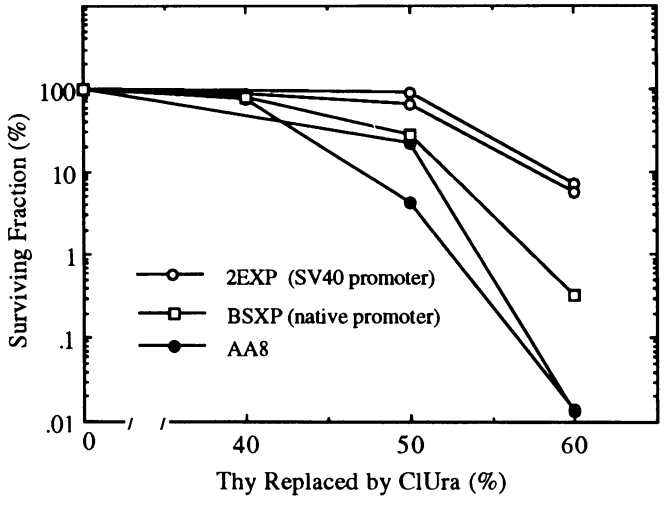

Figure 3. Survival of colony forming ability of *XRCC1* cDNA transfectants in increasing concentrations of CldUrd. The cloning efficiencies of AA8, 2EXP, and BSXP cells were determined as described in Materials and Methods. Symbols are defined in the figure. The two curves shown for AA8 and 2EXP are the result of two independent experiments, and the curve shown for BSXP is from one experiment.

as that in normal medium for greater than 200 h. Growth of the pXR1-30 transfectant 30-A-1 was severely impaired in CldUrd, a result that is consistent with the slow rate at which pXR1-30 transfectant colonies develop in CldUrd. Similar growth curves were observed for the pcD2EXδ transfectant 2EX1δ. In contrast, the pcD2EX transfectants, 2EXP and 2EX1, and the pBSX transfectants, BSXP and BSX1, maintained a growth rate in CldUrd comparable to that of AA8, suggesting that these constructs, which contain the full length *XRCC1* cDNA, were able to fully correct the defect in EM9 cells.

Level of SCE in EM9 transfectants

The level of phenotypic correction in the transfectants was further quantified by measuring the number of SCEs accrued over two cell generations. Average levels of 8.9 and 82.1 SCEs per cell were observed for AA8 and EM9 cells, respectively, in the experiments summarized in Table 2. These results are consistent with those reported in previous studies (1,4). Introduction of the pcD2E vector into EM9 cells (cell lines 2E1 and 2E2) did not appreciably alter its high level of SCE. The value of 28.1 SCEs per cell observed in 30-A-1 was similar to that reported previously for this cell line (1) and was similar to the level of SCE seen in 2EX1δ (21.7 SCEs per cell). The intermediate levels of SCE in these two cell lines are consistent with their slow growth rates in CldUrd and imply that the repair defect of EM9 is only partially corrected by the constructs pXR1-30 and pcD2EXd. In contrast, the pBSX and pcD2EX transfectants exhibited levels of SCE similar to that of AA8. Values ranged from 9.5−18.2 SCEs per cell in pBSX transfectants. With the exception of line BSX2, which had 18.2 SCEs per cell, these values show that pBSX conferred greater than 95% correction with respect to the SCE end point. Moreover, the value of 10.5 SCEs per cell seen in the culture of pooled transfectants (BSXP) suggests that the average level of correction conferred by pBSX is ≈ 97%. It is noteworthy that this near-complete correction is occurring with a construct that contains no introns. The cell line BSX2 could be less corrected than the others of this series because the chromosomal integration event may have disrupted the *XRCC1* reading frame.

All of the five independent transfectant clones derived with plasmid pcD2EX, as well as the pool 2EXP, displayed average SCE levels below that of AA8, raising the possibility of overcorrection in these cells. In addition, a second approach also pointed to the possibility of overcorrection. 2EXP displayed higher survival than did AA8 in the presence of 60% CldUrd in experiments in which these cells were tested for survival to various levels of CldUrd substitution (Fig 3.). These combined results may indicate that the construct pcD2EX can overcorrect the EM9 repair defect. It is possible that expression from the strong SV40 promoter present in pcD2EX (10) results in higher levels of XRCC1 protein than is present in AA8 cells. This interpretation would be consistent with the fact that we did not observe SCE levels lower than that of AA8 in those transfectants of EM9 that contain pBSX, in which the minigene is transcribed from the native promoter. The slightly elevated survival observed for BSXP transfectants with 60% CldUrd is not necessarily inconsistent with this hypothesis (Fig 3.). This elevation may reflect a small subpopulation of hyper-resistant cells resulting from transfectants that integrated multiple copies of the construct, and such a subpopulation would not necessarily affect the average SCE level for the total population. However, physical confirmation of overexpression is required to confirm these conclusions. To this end, we have generated recombinant XRCC1 protein in *E.coli* and are currently attempting to produce rabbit antibodies specific to XRCC1 (unpublished observations). We are also extending these experiments to determine whether the repair proficiency of cultured human cells can be increased and/or decreased by the introduction of sense and antisense *XRCC1* constructs, respectively. These studies should address the possibility that the sensitivity of human cells to DNA-damaging agents could be manipulated by altering the level of individual repair proteins, a possibility that could prove useful in designing novel approaches to cancer therapy.

ACKNOWLEDGMENTS

We thank Dr. Christine Weber for providing the pcD2E vector. This work was performed under the auspices of the U.S. Department of Energy by the Lawrence Livermore National Laboratory under contract No. W-7405-ENG-48.

REFERENCES

1. Thompson, L. H., Brookman, K. W., Jones, N. J., Allen, S. A. and Carrano, A. V. (1990) *Mol. Cell. Biol.*, **10**, 6160−6171.
2. Dillehay, L. E., Thompson, L. H. and Carrano, A. V. (1984) *Mutat. Res.*, **131**, 129−136.
3. Thompson, L. H., Brookman, K. W., Minkler, J. L., Fuscoe, J. C., Henning, K. A. and Carrano, A. V. (1985) *Mol. Cell. Biol.*, **5**, 881−884.
4. Thompson, L. H., Brookman, K. W., Dillehay, L. E., Carrano, A. V., Mazrimas, J. A., Mooney, C. L. and Minkler, J. L. (1982) *Mutat. Res.*, **95**, 247−440.
5. Chaganti, R. S. K., Schonberg, S. and German, J. (1974) *Proc. Natl. Acad. Sci. U.S.A.*, **71**, 4508−4512.
6. Weksberg, R., Smith, C., Anson-Cartwright, L. and Maloney, K. (1988) *Am. J. Hum. Genet.*, **42**, 816−824.
7. Perry, P. and Wolff, S. (1974) *Nature*, **251**, 156−158.
8. Minkler, J., Stetka, Jr D. and Carrano, A. V. (1978) *Stain Technol.*, **53**, 359−360.
9. Kozak, M. (1989) *J. Cell. Biol.*, **108**, 229−241.
10. Okayama, H. and Berg, P. (1983) *Mol. Cell. Biol.*, **3**, 280−289.