Vol. 195, No. 2, 1993 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
September 15, 1993 Pages 643-653

AUGMENTATION OF PROTEIN PRODUCTION BY A COMBINATION OF THE T7 RNA POLYMERASE SYSTEM AND UBIQUITIN FUSION:

Overproduction of the human DNA repair protein, ERCC1, as a ubiquitin fusion protein in *Escherichia coli*

Marcel H.M.Koken*, Hanny H.M.Odijk, Marcel van Duin[1], Maarten Fornerod and Jan H.J.Hoeijmakers

Department of Cell Biology and Genetics, Medical Genetics Centre, Erasmus University, P.O.Box 1738, 3000DR Rotterdam, The Netherlands

Received July 16, 1993

This article presents the development of a set of new expression vectors for overproduction of proteins in *Escherichia coli*. The vectors, pETUBI-ES1, 2 and 3, allow in-frame cloning of any sequence with the ubiquitin gene driven by the strong T7f10 promoter. Combination of the T7 expression system with ubiquitin fusion appears to have a synergistic effect on protein overproduction. Large amounts of stable RNA are produced by T7 RNA polymerase, and fusion of ubiquitin to the N-terminus of target proteins seems to confer more efficient translation, better folding or protection against proteolytic degradation. The ubiquitin part can be utilized for purification of the fusion protein, after which it can be easily removed from the fusion product by ubiquitin-specific proteases. The advantage of combining both systems is demonstrated by the synthesis of large quantities (up to 40-50% of the total protein) of the human ERCC1 protein that hitherto was refractory to overproduction in various other *E.coli* and yeast expression systems.

The human DNA excision-repair gene, *ERCC1*, codes for a protein of 297 amino acids, with a calculated molecular weight of 32.5kD (1). The gene product appears to be conserved in evolution and to be built as a mosaic of a yeast DNA excision-repair protein, RAD10, and parts of two *E.coli* DNA excision-repair enzymes UvrA and UvrC (2). To understand the importance of segments conserved between these polypeptides, of postulated functional domains (1), as well as for crystallization purposes, functional full-length protein has to be prepared in sufficiently large amounts. Many attempts were made to overproduce this polypeptide in *E.coli* and yeast. Yeast systems that were successful in overproduction of yeast RAD10 (3) and many other proteins, failed to produce significant quantities of ERCC1 (results obtained in collaboration with S. and L. Prakash, Rochester). In *E.coli*, the cloning of parts or the complete open reading frame of *ERCC1*

*To whom correspondence should be addressed at present address: Hôpital St. Louis, Laboratoire des Rétrovirus et Rétrotransposons des Vertébrés, Batiment INSERM, 16, Rue de la Grange aux Belles, 75010 Paris, France. Fax: (33)1.42.06.48.57.

[1]Present address: Dept. of Immunology, Organon Int. B.V., P.O.Box 20, 5340BH Oss, The Netherlands.

0006-291X/93 $4.00

behind several strong bacterial promoters (TAC, P_L, Trp, LacZ) in different protease-deficient bacterial strains did not yield noticeable quantities of ERCC1 protein, whereas the same systems tested with other genes in parallel induced large amounts of the corresponding polypeptides. Fusions of ERCC1 with ß-galactosidase (LacZ) or *Staphylococcus aureus* protein A enabled us to overproduce the C-terminal and middle parts of the protein, but only in relatively small amounts (unpublished results). The largest part of ERCC1 that could be synthesized in this way was a protein A fusion-product containing the C-terminal 85% of the protein (i.e. amino acid 33 to C-terminus), against which polyclonal antibodies were made. These were used to check the protein production in this study as well as in all above mentioned experiments. The N-terminal region of the protein could never be expressed successfully in *E.coli*.

Here we report the generation of a set of new bacterial expression vectors consisting of a combination of the T7 RNA polymerase system to ensure high expression at the RNA level and gene fusion with ubiquitin, which is thought to protect against proteolysis. The T7 RNA polymerase system (4) allows inducible expression, driven by the strong T7 f10 promoter, and synthesis of large amounts of mRNA. In addition, the vectors harbour optimal translation initiation signals and efficient transcription terminators for the T7 RNA polymerase. They are used in combination with derivatives of the bacterial strain BL21 that are deficient for the OmpT and Lon protease, and which have integrated the gene for T7 RNA polymerase into their genome under control of the LacZ promoter. Some BL21 derivatives produce T7 lysozyme, an inhibitor of T7 RNA polymerase. Under non-induced conditions the leakiness of the LacZ promoter can cause promiscuous production of T7 RNA polymerase, which is inhibited by very low quantities of T7 lysozyme. Even very toxic proteins can be overproduced in this way (4).

N-terminal ubiquitin fusion is now a frequently used overexpression method in biotechnology (5-16). Ubiquitin, one of the most conserved proteins in eukaryotes, is thought to play an important role in many cellular processes, including specific protein breakdown, chromatin remodelling, protein (re)folding, and unexpectedly "protein stabilization" (for reviews see, (17-19)). The latter two properties of ubiquitin are probably exploited by a number of naturally occurring eukaryotic genes that actually are encoded as ubiquitin fusion proteins (7, 20-31). The protein-stabilizing effect and improved folding make ubiquitin an attractive partner for fusion with proteins to be properly synthesized in large quantities, although the phenomenon of this N-terminal-end protection, improved translation or facilitated folding accomplished by ubiquitin is not understood. One of the additional advantages of the ubiquitin fusion protein system is the occurrence of very accurate ubiquitin-specific proteases that can be used to clip the junction between ubiquitin and the (poly)peptide of interest in a very precise manner. Several of these enzymes have been isolated and some of the corresponding genes have been cloned from yeast and human (32-38). The combination of the T7 and ubiquitin systems, in the case of ERCC1 protein, enabled us to augment the protein yield to a level much higher than achieved by either system alone. It is plausible that the augmentation of production achieved by the combination of these two systems will show to be universally applicable for overproduction of polypeptides. Towards that end vectors allowing the insertion of three reading frames have been constructed.

METHODS AND MATERIALS

DNA manipulations

All DNA manipulations were done essentially as described (39, 40).
Oligonucleotides, used in PCR reactions to create restriction enzyme sites, ubiquitin-ERCC1 joining or polylinkers were synthesized on an Applied Biosystems 392 DNA/RNA synthesizer.
Single-stranded oligonucleotides:

Met Asp Pro Gly Lys
I. 5'*Nde* I site: 5'AGGCTC**CATATG**GACCCTGGGAAG3'

II. 3'*Bam* HI site: 5'CCG**GGATCC**AAGCTTGGCAGCTGGGGTCAT3'

III.3'End ubiquitin (*Bfr* I site)-5'end ERCC1:
Val Leu Arg Leu Arg Gly Gly Met Asp Pro Gly Lys
5'TTGT**CTTAAG**ACTAAGAGGTGGTATGGACCCTGGGAAGG3'.

Restriction sites are indicated in bold, *ERCC1* -encoded regions are underlined, while ubiquitin encoded segments are doubly underlined.
As a template for the PCR reactions involving *ERCC1*, the plasmid pcDE (1) was used.
Double-stranded oligonucleotides used for the construction of pETUBI-ES1, 2, AND 3:

Bam HI *Eco* RI *Pst* I *Sma* I *Bam* HI
ES1: 5'GATCCGGTACCCATC**GAATTC**CTG**CAG**CCC**GGG**G 3'
3' GCCATGGGTAG**CTTAAGGACGTCGGGCCC**CCTAG 5'

Bam HI *Eco* RI *Pst* I *Sma* I *Bam* HI
ES2: 5'GATCCGGGTACCCATCGA**ATT**CCT**GCA**GCC**CGG**G 3'
3' GCCCATGGGTAG**CTTAAGGACGTCGGGCCC**CTAG 5'

Bam HI *Eco* RI *Pst* I *Sma* I *Bam* HI
ES3: 5'GATCCGGGGTACCCATCG**AATTC**CT**GCA**GCC**CGG**G 3'
3' GCCCCATGGGTAG**CTTAAGGACGTCGGGCCC**CTAG 5'.

In bold the three restriction sites that can be used for in frame cloning are indicated. The single underlining indicates the codons which are in frame with ubiquitin. Double underlining shows the difference between the three pETUBI-vectors. The fusion of the C-terminal end of ubiquitin (underlined) into the different vectors looks as follows:

ES1: Arg-Gly-Gly-Ala-Asp-Pro-Val-Pro-Ile-**Glu-Phe***
ES2: Arg-Gly-Gly-Ala-Asp-Pro-Gly-Thr-Asp-A**rg-Ile***
ES3: Arg-Gly-Gly-Ala-Asp-Pro-Gly-Tyr-Arg-Se**r-Asn***.

* The bold amino acids are encoded by the *Eco*RI site.
To check whether the heatshock-induced degradation systems could be responsible for the difference in protein production between bacteria containing the T7 and P_L promoter driven constructs, the *Bam* HI site located directly behind ubiquitin was removed from the original pNMHUB-poly cloningvector (using Mungbean nuclease), and pETUBI vector (filling in by a Klenow reaction). This removal generates in both vectors a stop codon 6 amino acids downstream of the normal C-terminal end of ubiquitin.

Bacterial strains

For all experiments bacterial strains described by Studier *et al.* (4) (BL21 derivatives) and Butt *et al.* (5) (AR58) were used. All BL21 derivatives were purchased from Novagen, Inc. In the BL21 parental strain Studier and coworkers have integrated T7 RNA polymerase under control of the LacZ promoter giving rise to strain BL21(DE3). This strain has subsequently been transformed with the specified plasmids causing the accumulation of different amounts of T7 lysozyme; strains BL21(DE3)LysS and BL21(DE3)LysE. Overnight bacterial cultures were diluted a hundred times in 50 ml Terrific Broth (TB) (39) and incubated at 32°C or 37°C till OD600=0.7. IPTG (final concentration: 0.4mM) was added to the cultures harbouring pET3c-derived constructs, and for the pNMHUB-poly derivatives the temperature was raised to 42°C while shaking in a waterbath and kept at 42°C for the rest of the experiment. Every 30 minutes samples were taken for determination of OD, and, after induction, for protein detection on Western blot.

Protein manipulations

Standard techniques for SDS-PAGE, Coomassie Brilliant Blue staining, Western blotting and antibody incubations were used (41). Bacteria grown as described were pelleted, resuspended in sample buffer, and boiled during 3 minutes. 11% PAGE-Gels (1.5 mm MiniProtean (BioRad)), loaded with about 10-15µg total protein per lane were run under standard conditions. Gels were blotted onto immobilon PVDF (Millipore) membranes, during 1 hour, at 250mA, at 4°C, according to the manufacturers description. The polyacrylamide gels were, after blotting, stained with Coomassie Brilliant Blue. Blots were incubated with anti-ERCC1 antiserum overnight at 4°C, washed, incubated during 1 hour with alkaline phosphatase-conjugated Goat-anti-Rabbit IgG (TAGO, Inc.) at room temperature, washed again, and incubated with substrate (40µl 1M $MgCl_2$, 1 ml 0.1% nitrobluetetrazolium in veronal buffer, 100µl 5-bromo-4-chloro-3-indolylphosphate (5mg/ml dimethylformamide)) in 9 ml 0.15M veronalbuffer (diethylbarbituric acid (Na^+ salt), pH 9.6). Polyclonal antibodies against ERCC1 are described elsewhere (42). The antibodies were prepared as delineated before (41, 43).(Scanning of the CBB gels and blots was performed on a Biorad 2D videodensitometer, model 620, equiped with a 600nm interference filter. The SDS-PAGE gels used for scanning were not blotted, but directly stained and scanned. However, no obvious differences in relative amounts between blotted and non-blotted CBB stained gels could be detected. To determine the relative fraction of ERCC1(-fusion) product the average was calculated from the scanning results of three independent experiments.

The solubility of proteins was checked by making a sonicated extract of the bacteria that overproduced them (5 times 30 seconds, at 4°C). This extract was spun for 15 minutes in an Eppendorf table-top centrifuge at 4°C. The pellet was regarded as the insoluble, and the supernatant as the soluble fraction.

Clipping of the ubiquitin fusion proteins was done as described (36). In brief, 10µl (≈5µg) of a 33% $(NH_4)_2SO_4$-precipitated (44) extract of *E.coli* producing the ubiquitin-ERCC1 fusion protein was mixed with 1 µl of a total sonicated extract of yeast in 10µl clipping buffer (50mM Tris.HCl, pH 7.5, 1mM EDTA, 10 mM DTT) and incubated for 30 minutes at 37°C. The reaction was stopped by adding one volume of 2x sample buffer (0.12M Tris.HCl pH6.8, 2% SDS, 20% glycerol, 0.6M ß-mercaptoethanol, 0.005% Bromophenolblue).

RESULTS AND DISCUSSION

Construction of the expression vectors

The cloning strategy for all the constructs is shown in Figure 1. Using PCR, an *Nde* I site (oligonucleotide I; for the sequence of the primers used, see "METHODS AND MATERIALS") was created at the position of the *ERCC1* ATG initiation codon, and a *Bam* HI site (oligonucleotide II) at the 3' end of the open reading frame. The *Nde* I-*Bam* HI PCR fragment was inserted into the T7 cloning vector, pET3c (4), resulting in pET3c..ERCC1 (Fig.1). An oligonucleotide harbouring the C-terminal part of ubiquitin and the N-terminus of ERCC1 (oligonucleotide III) and a 3' *ERCC1* oligonucleotide (oligonucleotide II) carrying a *Bam* HI site, were used in PCR to link the ubiquitin and *ERCC1* coding regions. After digestion with *Bam* HI and *Bfr* I, that cleaves in the C-terminus of ubiquitin the PCR fragment was introduced between the *Bfr* I and *Bam* HI sites of pNMHUB-poly (5), yielding construct pNMHUB-poly..ERCC1 (Fig.1). To combine the T7 and the ubiquitin systems the complete open reading frame of ubiquitin was released from pNMHUB-poly by an *Nde* I-*Bam* HI digestion, and cloned into the *Nde* I-*Bam* HI sites of pET3c, yielding vector pETUBI. In order to use the *Sal* I site in ubiquitin as a cloning site for fusion, the internal *Sal* I site in the pET3c vector-part was removed by ligation of Mungbean nuclease and Klenow polymerase treated partial *Sal* I digests of pETUBI. The vector that retained the *Sal* I site in the ubiquitin coding region was called pETUBI-S (Fig.1). Subsequently, the *Sal* I-*Bam* HI fragment from pNMHUB-poly..ERCC1 was cloned into pETUBI-S digested with *Sal* I/*Bam* HI, generating pETUBI-S..ERCC1 (Fig.1). The sequence of the whole open reading frame of the ubiquitin-

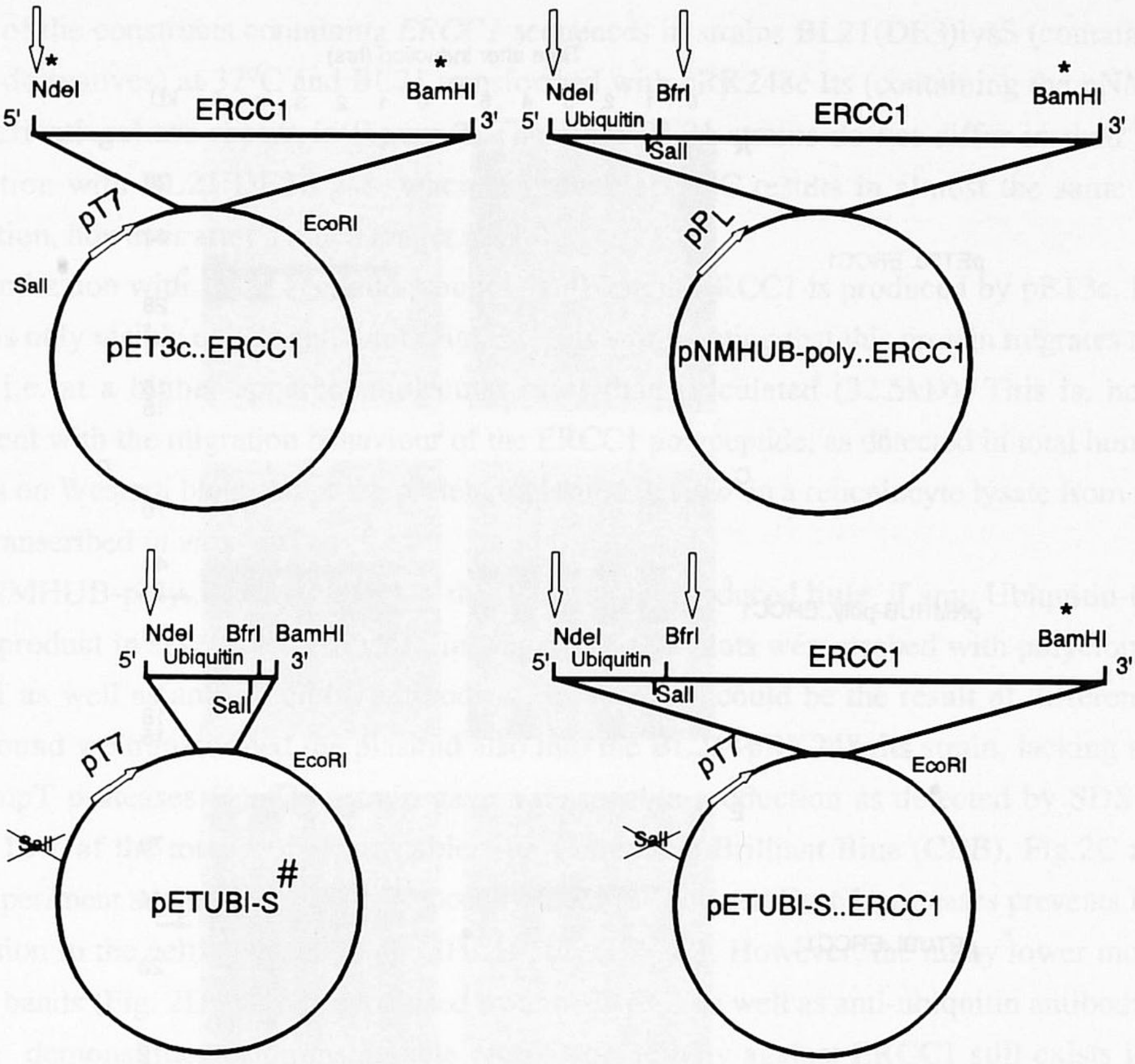

Figure 1.
Construction of the *E.coli* overexpression vector, pETUBI-S.
The exact construction of the different plasmids is described in RESULTS AND DISCUSSION-Construction of the expression vectors. In short: Via PCR cloning sites were created around the open reading frame of *ERCC1* to insert it into pET3c (not shown), resulting in pET3c..ERCC1. With the same technique the C-terminal part of ubiquitin was fused to the coding region of *ERCC1*. This *ERCC1* containing PCR fragment was inserted into the pNMHUB-poly vector (not shown), which resulted in pNMHUB-poly..ERCC1. At the same time pETUBI was made by inserting the fragment of pNMHUB-poly harbouring the ubiquitin gene into pET3c. After some modifications of the pETUBI vector the *Sal* I-*Bam* HI fragment derived from pNMHUB-poly..ERCC1 was inserted giving rise to pETUBI-S..ERCC1.
pT7 and pP_L indicate the type of promoter driving the plasmid-inserts. Arrows above the inserts indicate the position of the ATG of ubiquitin and/or ERCC1. The crossed *Sal* I site is removed as explained in the text. *: sites made by PCR amplification. #: variants of this vector (pETUBI-ES1 to 3), enabling the in frame fusion with ubiquitin were prepared by inserting the in "Experimental protocol" specified double-stranded oligonucleotides into the *Bam*HI site. The indicated *Eco* RI site present in the pET3c backbone was removed for the construction of pETUBI-ES1 to 3.

ERCC1 fusion in this construct was verified to exclude possible PCR-derived mutations. The pETUBI-S plasmid was converted into a more general expression vector by inserting a multiple cloning site in three reading frames behind the ubiquitin sequence. Towards this end, the *Eco* RI site located between the T7 terminator and the ampicillin resistance gene in the pET3c part of the vector was destroyed by a Klenow treatment, yielding pETUBI-ES. Finally, one of the three double-stranded oligonucleotides, ES1, ES2 or ES3 (see "METHODS AND MATERIALS") was inserted into the *Bam* HI site of pETUBI-ES, giving rise to pETUBI-ES 1, 2, and 3 (see legend Fig.1), respectively, permitting in frame fusion with other genes.

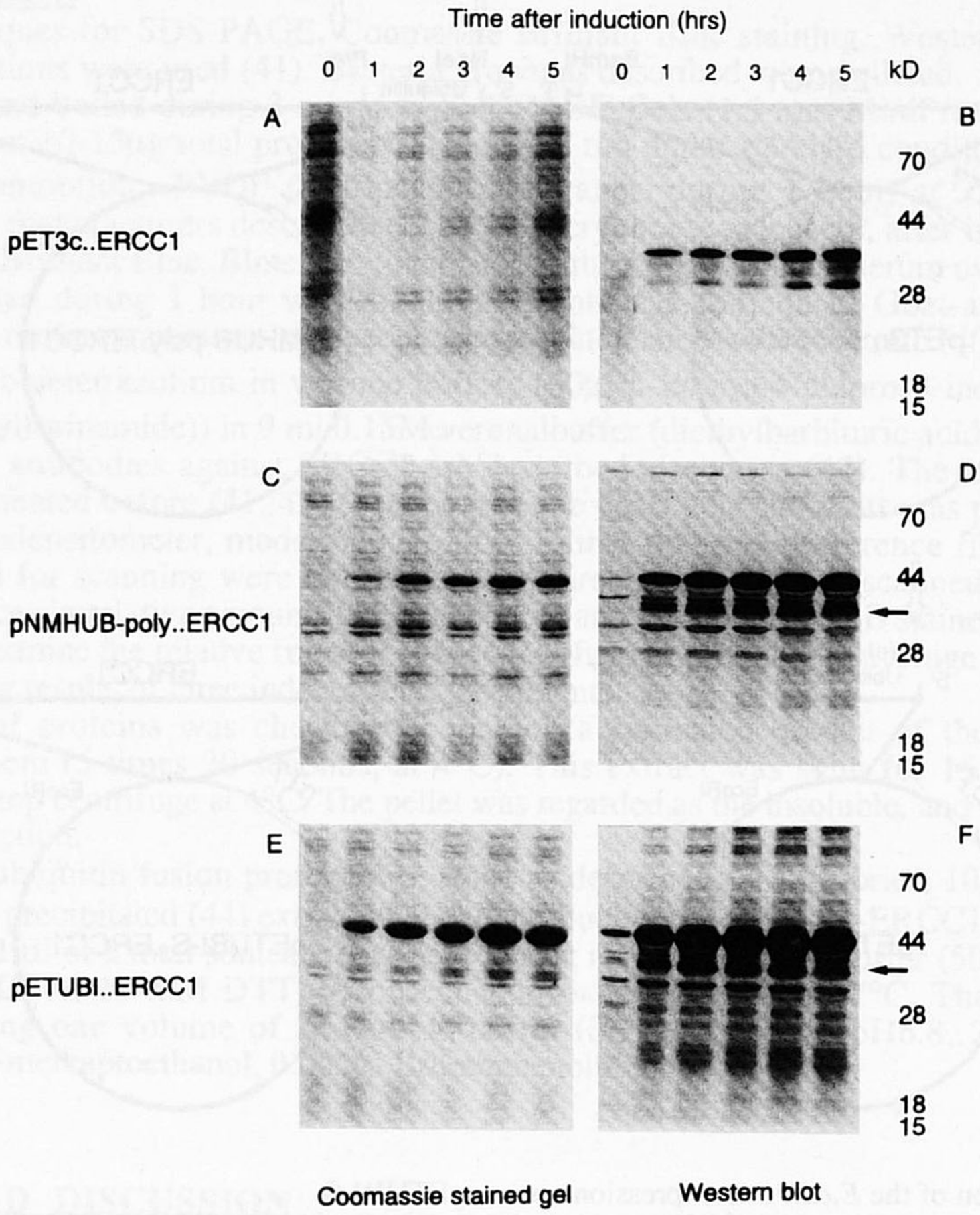

Figure 1.
Augmentation of protein overexpression by combination of the T7 RNA polymerase system with N-terminal ubiquitin fusion.
Bacteria containing the specified constructs were grown and blots were prepared as described in "METHODS AND MATERIALS". Blots D and F were incubated during an equal time span, to allow comparison, whereas the **B** blot was incubated **4** times as long. Left panels: Coomassie Brilliant Blue-stained gels (after blotting). Right panels: Western blots incubated with anti-ERCC1 antiserum. The arrow indicates the *E.coli* processed band. Panels A, B, E and F represent experiments performed at 37°C; Panels C and D at 42°C. The poor visability of the free ERCC1 protein in Fig. 2C/E is probably due to intrinsic properties of the ERCC1 protein which make that it does not easily retain the dye, and quickly loses it during destaining. As shown in Fig. 3 short destaining permits a more clear visualisation of free ERCC1. To circumvent this problem all experiments were double-checked by Western blotting.

Production of ERCC1 protein in the different expression systems

E.coli strains BL21(DE3), BL21(DE3)LysS, and BL21(DE3)LysE (4), were transformed with plasmids pET3c, pETUBI-S, pET3c..ERCC1 and pETUBI-S..ERCC1. The vectors pNMHUB-poly and pNMHUB-poly..ERCC1 were introduced into bacterial strain AR58 (5) carrying the *c*I857 temperature-sensitive mutation of the lambda repressor, and into BL21, which harboured already the plasmid pRK248*c*Its (45, 46), encoding a ts*cIAt* 2 repressor. After induction (see "METHODS AND MATERIALS") of the strains harbouring the pET3c-derived constructs by IPTG, and of the bacteria containing the pNMHUB-poly derivatives by a temperature shift to 42°C, protein production was investigated by Western blot analysis in a time-course experiment. The

results of the constructs containing *ERCC1* sequences in strains BL21(DE3)lysS (containing the pET3c derivatives) at 37°C and BL21 transformed with pRK248*c* Its (containing the pNMHUB-poly derivatives) are shown in Figure 2. The other BL21 strains do not differ in their protein production with BL21(DE3)LysS, whereas growth at 32°C results in almost the same protein production, however after a much longer period.

Upon induction with IPTG a small amount of full-length ERCC1 is produced by pET3c..ERCC1 which is only visible on Western blot (Fig.2B). It is worth noting that this protein migrates at about 39kD, i.e. at a higher apparent molecular mass than calculated (32.5kD). This is, however, consistent with the migration behaviour of the ERCC1 polypeptide, as detected in total human cell extracts on Western blots and of the protein translated *in vitro* in a reticulocyte lysate from *ERCC1* RNA transcribed *in vitro* (47).

The pNMHUB-poly..ERCC1 vector in the AR58 strain produced little, if any, Ubiquitin-ERCC1 fusion product in several experiments, in which Western blots were probed with polyclonal anti-ERCC1 as well as anti-ubiquitin antibodies. Because this could be the result of different strain background we transformed the plasmid also into the BL21+pRK248*c*Its strain, lacking the Lon and OmpT proteases (4). This strain gave a reasonable production as detected by SDS-PAGE (about 10% of the total protein stainable with Coomassie Brilliant Blue (CBB), Fig.2C and D). This experiment suggests that the presence of functional Lon and OmpT proteases prevents ERCC1 expression in the cells containing pNMHUB-poly..ERCC1. However, the many lower molecular weight bands (Fig. 2D and F), recognised by anti-ERCC1 as well as anti-ubiquitin antibodies (not shown) demonstrates that considerable proteolytic activity against ERCC1 still exists in these strains.

By far the highest ERCC1 yield, however, was achieved with construct pETUBI-S..ERCC1 in BL21(DE3)LysS cells as shown by the Coomassie staining (Fig.2E). Scanning of the CBB-stained gel indicates that in different experiments in the order of 40-50% of the total stainable protein is Ubiquitin-ERCC1 (44kD), which is approximately five-fold higher than pNMHUB-poly..ERCC1 in BL21(DE3)LysS.

Because our goal was to produce a native, functional protein, the polypeptide has to satisfy several criteria. One of these is the solubility of the overexpressed protein. A problem often associated with an overproduction of proteins in *E.coli* is the formation of socalled inclusion bodies (48-50), yielding insoluble protein. In several instances growth of the bacteria at lower temperatures has solved this problem (48-50). In the case of the Ubiquitin-ERCC1 fusion, at 37°C indeed about 90% of the protein could be found in the insoluble pellet of the extract, whereas at 32°C almost 90% of the fusion protein appeared to be soluble (results not shown). This demonstrates that in the case of Ubiquitin-ERCC1, the formation of inclusion bodies can be successfully prevented by culturing at lower temperatures.

Another criterium is that the ubiquitin part can be cleaved off the fusion-product thus generating the free ERCC1 protein.To test whether the naturally occurring ubiquitin-specific proteases present in yeast or human extracts can accomplish this, a digestion of the ubiquitin-fusion was performed as described (36). As shown in Figure 3 (right panel lane 1), a partially purified Ubiquitin-ERCC1 fusion product can be cut by one of the ubiquitin-specific proteases, present in total yeast (or

human) protein extracts. In unpublished experiments we have found that the overproduced ubiquitin-specific protease encoded by the cloned yeast gene, UBP1 (32), can clip the fusion product (unpublished results, not shown), whereas the bovine ubiquitin carboxyl terminal hydrolase, UCH-L1 (PGP9.5) (35), or the hydrolase encoded by the yeast YUH1 (36) failed to do so.
Fig.2B shows that the free ERCC1 protein (39kD) produced by *E.coli* strain BL21(DE3)lysS is degraded to a considerable extent. Comparing the intact ERCC1 (-fusion) band (Fig.2B and 2F) with the lower-molecular-weight "degradation products" demonstrates that in the case of Ubiquitin-ERCC1 (Fig.2F) about 50% of the total ERCC1 fusion protein appears not to be degraded, whereas with the non-fused ERCC1 this is about 58% (Fig.2B). In the case of Ubiquitin-ERCC1 (Fig.2F), however, the results are complicated by the presence of a protease in *E.coli* which -albeit slowly- cleaves exactly between ubiquitin and ERCC1, as has been documented for other ubiquitin fusions (5, 51). When this 'non-degradative' activity is taken into account, the lower 39kD band (arrow) in Figure 2F, which represents about 30% of the total stainable protein, is probably free intact ERCC1. In agreement with this interpretation is the notion that the protein has the expected size and reacts with the ERCC1 antiserum. This implies that about 80% of pETUBI-S..ERCC1 is detected as a Ubiquitin-ERCC1 product or free non-degraded ERCC1, in contrast to pET3c..ERCC1 where about 55% is found not to be degraded (i.e. 20% versus 50% degradation). When this 39 kD band indeed represents free ERCC1, its relatve intactness suggests that the prior fusion with ubiquitin induces protection of the free protein against proteolysis. A possible explanation could be that the ubiquitin imposes a correct folding of the protein thus making it less vulnerable to the action of proteases.

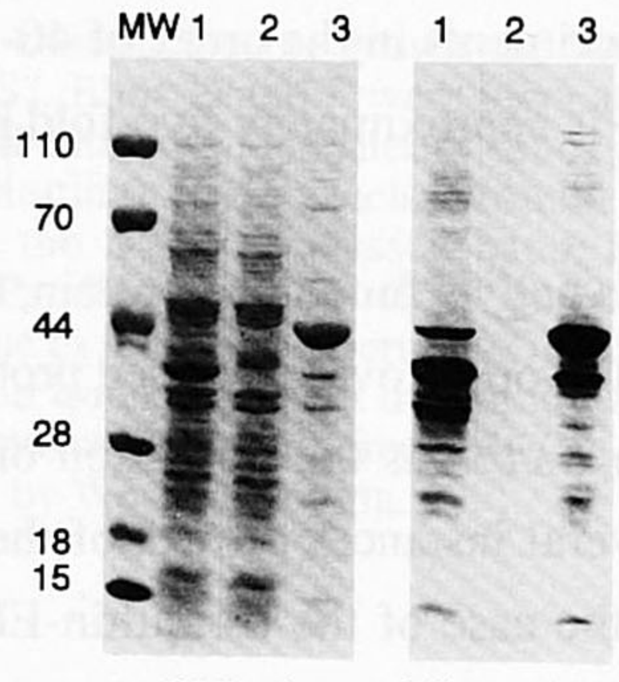

Figure 3.
Cleavage of ubiquitin-ERCC1 by a total *S.cerevisiae* protein extract.
$(NH_4)_2SO_4$ precipitated extracts of BL21(DE3)LysS transformed with pETUBI..ERCC1, treated, or not treated, with a total yeast extract according to the method described in "METHODS AND MATERIALS" were run on an 11% SDS-PAGE gel and blotted. The remaining gel was stained, and the blot was incubated with anti-ERCC1 antiserum overnight. The rest of the procedure is as specified in M&M.
Left panel: Coomassie Brilliant Blue stained gel. Right panel: Western blot incubated with anti-ERCC1 antiserum. Lane 3 (Control): 10 µl of a 33% $(NH_4)_2SO_4$-precipitated(44) extract of *E.coli* producing the ubiquitin-ERCC1 fusion protein. Lane 2 (Control): 1µl of a total yeast extract. Lane 1 (reaction): mixture of 2 and 3, incubated as specified in "MATERIALS AND METHODS". MW: apparent molecular weight. It is remarkable that the free ERCC1 protein is hardly visible on the CBB-stained gels.

When the absolute production of bacteria carrying plasmids pNMHUB-poly..ERCC1 and pETUBI..ERCC1 is compared with that with pET3c..ERCC1 it is apparent that ubiquitin fusions yield much larger amounts. This could be the result of a more efficient translation of the fusion protein, N-end protection and/or better folding. However, also free ERCC1 (*E.coli* processed) seems to be quite stable after initial production as a ubiquitin fusion. This makes the mechanism of N-end protection more unlikely.

The amount produced by BL21(DE3)LysS bacteria harbouring pNMHUB-poly..ERCC1 (Fig.2C) is about one fifth of that with the pETUBI..ERCC1 construct (Fig.2E). This can be explained by differences in promoter strength, mRNA stability or translation efficiency. Alternatively it can be due to the initiation of the heatshock response necessary for induction of the P_L promoter which is known to trigger the protein-degradation systems (52). To address the latter possibility a stopcodon was introduced into both vectors (see "METHODS AND MATERIALS") 6 codons downstream of the normal C-terminal end of ubiquitin. Both constructs were transformed into the appropriate BL21 strains for heatshock (pNMHUB-poly) or IPTG (pETUBI) induction. Like with the ERCC1-containing vectors, an approximately five fold difference in protein induction was detected (data not shown). Ubiquitin itself is a very stable heatshock protein, which is highly resistant to degradation (21). Therefore, it is likely that the above mentioned differences in protein yield are not caused by breakdown, but by differences in promoter strength and/or mRNA stability.

We show here that the combination of the T7 system and ubiquitin fusion permitted the synthesis of large quantities of ERCC1 protein which appeared to be refractory to overproduction using a variety of other procedures. Probably the main reason for failure using other systems is the rapid degradation of the protein. We think that the combination of the strong RNA expression brought about by the T7 RNA polymerase vector and the N-terminal ubiquitin which protects against degradation and promotes proper folding are the major factors contibuting to the successful overproduction of ERCC1. Although, almost all the proteins produced as ubiquitin fusions which have been published in literature were found to be functional (6, 8, 9, 11-14, 16, 51), the functionality of ERCC1 cannot be tested *in vivo* or *in vitro* because efficient assays are not available at present. Moreover, we have recently found using an *in vitro* repair assay based on Manley-type cell-free extracts that the ERCC1 polypeptide resides *in vivo* in a stable protein complex together with the gene products of ERCC4, ERCC11 and the XPFC genes (47). The absence of one of the proteins causes the breakdown/instability of all other constituents of the complex. Therefore, only the cloning of the genes for the other components and the isolation of the corresponding proteins will make it possible to assess the function of the complex in the repair reaction and the role of the individual components.

The vector system developed here may proove to be useful for other proteins which are difficult to overproduce in *E.coli* using the commonly known systems, particularly when the sensitivity to proteolysis is one of the main reasons for the low yield. Since N-terminally fused ubiquitin is a natural substrate for ubiquitin-specific proteases the chance is high that the ubiquitin part can easily and in a highly specific way be clipped off by the purified UBP1 enzyme or by cell extracts, as was found for the ERCC1 protein. The latter property may be very useful when the protein has to be tested in cell-free extracts. Finally, the ubiquitin part can be a convenient tag for the purification of the particular protein using ubiquitin antibody columns.

Acknowledgments: We thank Drs. F.W.Studier (Brookhaven Natl. Lab., NY) and T.R.Butt (Smith Kline Beecham Pharm., Pennsylvania) for providing us with their vectors and cells; Drs. A.Varshavsky and J.W.Tobias (Mass.Inst.Tech., Massachusetts) for their cloned yeast hydrolase, UBP1; Drs. K.Wilkinson (Emory Univ., Atlanta) and W.J.Henzel (Genentech, Inc., California) for their isolated ubiquitin hydrolases; and Dr. R.Crowl for giving us his RR1 strain containing the pRK248*c* Its vector. A.van Oudenaren is acknowledged for help with densitometry; Dr. A.Eker, J.v.d.Tol, I.Jaspers-Dekker, P.Reuvers, and Dr. N.G.J.Jaspers for technical advice and/or assistance; Prof. D.Bootsma for critical reading the manuscript; and Mirko Kuit and Tom de Vries Lentsch for photography and Sjozef van Baal for computer assistance. This work was supported by the Dutch Cancer Society (project IKR 88-2 and 90-20) and the European Community Contract B16-141-NL.

REFERENCES

1.Van Duin, M., De Wit, J., Odijk, H., Westerveld, A., Yasui, A., Koken, M.H.M., Hoeijmakers, J.H.J. and Bootsma, D. (1986) Cell 44, 913-923.
2.Van Duin, M., Van den Tol, J., Warmerdam, P., Odijk, H., Meijer, D., Westerveld, A., Bootsma, D. and Hoeijmakers, J.H.J. (1988) Nucl. Acids Res. 16, 5305-5322.
3.Sung, P., Prakash, L. and Prakash, S. (1992) Nature 355, 743-745.
4.Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990) Meth. Enzymol. 185, 60-89.
5.Butt, T.R., Jonnalagadda, S., Monia, B.P., Sternberg, E.J., Marsh, J.A., Stadel, J.M., Ecker, D.J. and Crooke, S.T. (1989) Proc. Natl. Acad. Sci. USA 86, 2540-2544.
6.Ecker, D.J., Stadel, J.M., Butt, T.R., Marsh, J.A., Monia, B.P., Powers, D.A., Gorman, J.A., Clark, P.E., Warren, F., Shatzman, A. and Crooke, S.T. (1989) J. Biol. Chem. 264, 7715-7719.
7.Monia, B.P., Ecker, D.J., Jonnalagadda, S., Marsh, J., Gotlib, L., Butt, T.R. and Crooke, S.T. (1989) J. Biol. Chem. 264, 4093-4103.
8.Lu, C., Yang, Y.F., Ohashi, H. and Walfish, P.G. (1990) Biochem. Biophys. Res. Comm. 171, 138-142.
9.Power, R.F., Conneely, O.M., McDonnell, D.P., Clark, J.H., Butt, T.R., Schrader, W.T. and O'Malley, B.W. (1990) J. Biol. Chem. 265, 1419-1424.
10.Yoo, Y., Rote, K. and Rechsteiner, M. (1989) J. Biol. Chem. 264, 17078-17083.
11.Sabin, E.A., Lee-Ng, C.T., Shuster, J.R. and Barr, P.J. (1989) Bio/Technology 7, 705-709.
12.Mak, P., McDonnell, D.P., Weigel, N.L., Schrader, W.T. and O'Malley, B.W. (1989) J. Biol. Chem. 264, 21613-21618.
13.Wittliff, J.L., Wenz, L.L., Dong, J., Nawaz, Z. and Butt, T.R. (1990) J. Biol. Chem. 265, 22016-22022.
14.Cherney, B.W., Chaudhry, B., Bhatia, K., Butt, T.R. and Smulson, M. (1991) Biochem. 30, 10420-10427.
15.Barr, P.J., Inselburg, J., Green, K.M., Kansopon, J., Hahm, B.K., Gibson, H.L., Lee-Ng, C.T., Bzik, D.J., Li, W.B. and Bathurst, I.C. (1991) Mol. Biochem. Parasitol. 45, 159-170.
16.Klein-Hitpass, L., Tsai, S.Y., Weigel, N.L., Allan, G.F., Riley, D., Rodriguez, R., Schrader, W.T., Tsai, M.J. and O'Malley, B.W. (1990) Cell 60, 247-257.
17.Hershko, A. (1988) J. Biol. Chem. 263, 15237-15240.
18.Rechsteiner, M. (1988) Ubiquitin. Plenum, New York.
19.Jentsch, S., Seufert, W., Sommer, T. and Reins, H.A. (1990) Trends Biochem. Sci. 15, 195-198.
20.Ozkaynak, E., Finley, D., Solomon, M.J. and Varshavsky, A. (1987) EMBO J. 6, 1429-1439.
21.Baker, R.T. and Board, P.G. (1991) Nucl. Acids Res. 19, 1035-1040.
22.Cabrera y Poch, H.L., Arribas, C. and Izquierdo, M. (1990) Nucl. Acids Res. 18, 3994.
23.Callis, J., Raasch, J.A. and Vierstra, R.D. (1990) J. Biol. Chem. 265, 12486-12493.
24.Finley, D., Bartel, B. and Varshavsky, A. (1989) Nature 338, 394-401.
25.Lee, H.S., Simon, J.A. and Lis, J.T. (1988) Mol. Cell. Biol. 8, 4727-4735.
26.Mezquita, J., Pau, M. and Mezquita, C. (1988) Nucl. Acids Res. 16, 11838.
27.Redman, K.L. and Rechsteiner, M. (1988) J. Biol. Chem. 263, 4926-4931.
28.Redman, K.L. and Rechsteiner, M. (1989) Nature 338, 438-40.
29.Rocamora, N. and Agell, N. (1990) Biochem. J. 267, 821-829.
30.Swindle, J., Ajioka, J., Eisen, H., Sanwal, B., Jacquemot, C., Browder, Z. and Buck, G. (1988) EMBO J. 7, 1121-1127.

31.Taccioli, G.E., Grotewold, E., Aisemberg, G.O. and Judewicz, N.D. (1991) Gene 102, 133-137.
32.Tobias, J.W. and Varshavsky, A. (1991) J. Biol. Chem. 266, 12021-12028.
33.Miller, H.I., Henzel, W.J., Ridgway, J.B., Kuang, W.-J., Chisholm, V. and Liu, C.-C. (1989) Bio/Technology 7, 698-704.
34.Agell, N., Ryan, C. and Schlesinger, M.J. (1991) Biochem. J. 273, 615-620.
35.Wilkinson, K.D., Lee, K.M., Deshpande, S., Duerksen-Hughes, P., Boss, J.M. and Pohl, J. (1989) Science 246, 670-673.
36.Liu, C.C., Miller, H.I., Kohr, W.J. and Silber, J.I. (1989) J. Biol. Chem. 264, 20331-20338.
37.Jonnalagadda, S., Butt, T.R., Monia, B.P., Mirabelli, C.K., Gotlib, L., Ecker, D.J. and Crooke, S.T. (1989) J. Biol. Chem. 264, 10637-10642.
38.Mayer, A.N. and Wilkinson, K.D. (1989) Biochem. 28, 166-172.
39.Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: a laboratory manual. Cold Spring Harbor, NY.
40.Koken, M.H.M., Reynolds, P., Jaspers-Dekker, I., Prakash, L., Prakash, S., Bootsma, D. and Hoeijmakers, J.H.J. (1991) Proc. Natl. Acad. Sci. USA 88, 8865-8869.
41.Harlow, E. and Lane, D. (1988) Antibodies: a laboratory manual. Cold Spring Harbor, NY.
42.Belt, P.B.G.M., Van Oosterwijk, M.F., Odijk, H., Hoeijmakers, J.H.J. and Backendorf, C. (1991) Nucl. Acids Res. 19, 5633-5637.
43.Amero, S.A., James, T.C. and Elgin, S.C.R. (1988) In New Protein Techniques (Walker, J.M., Ed.), 3, 355-362. Humana, Clifton, New Jersey.
44.Pognonec, P., Kato, H., Sumimoto, H., Kretzschmar, M. and Roeder, R.G. (1991) Nucl. Acids Res. 19, 6650.
45.Crowl, R., Seamans, C., Lomedico, P. and McAndrew, S. (1985) Gene 38, 31-38.
46.Bernard, H.-U. and Helinski, D.R. (1979) Meth. Enzymol. 68, 482-492.
47.Van Vuuren, A.J., Appeldoorn, E., Odijk, H., Yasui, A., Jaspers, N.G.J., Bootsma, D. and Hoeijmakers, J.H.J. (1993) EMBO J. (in press).
48.Kane, J.F. and Hartley, D.L. (1988) Tibtech 6, 95-101.
49.Schein, C.H. and Noteborn, M.H.M. (1988) Bio/Technology 6, 291-294.
50.Schein, C.H. (1989) Bio/Technology 7, 1141-1148.
51.Jonnalagadda, S., Butt, T.R., Marsh, J., Sternberg, E.J., Mirabelli, C.K., Ecker, D.J. and Crooke, S.T. (1987) J. Biol. Chem. 262, 17750-17756.
52.Caulcott, C.A. and Rhodes, M. (1986) Tibtech June, 142-146.