Cell-free cloning of highly expanded CTG repeats by amplification of dimerized expanded repeats

Robert J. Osborne[1,2,*] and Charles A. Thornton[1]

[1]Department of Neurology, University of Rochester School of Medicine and Dentistry, 601 Elmwood Avenue, Rochester, NY 14642, USA and [2]Department of Physiology, Anatomy and Genetics, University of Oxford, South Parks Road, Oxford, OX1 3QX, UK

ABSTRACT

We describe conditions for producing uninterrupted expanded CTG repeats consisting of up to 2000 repeats using φ29 DNA polymerase. Previously, generation of such repeats was hindered by CTG repeat instability in plasmid vectors maintained in *Escherichia coli* and poor *in vitro* ligation of CTG repeat concatemers due to strand slippage. Instead, we used a combination of *in vitro* ligation and φ29 DNA polymerase to amplify DNA. Correctly ligated products generating a dimerized repeat tract formed substrates for rolling circle amplification (RCA). In the presence of two non-complementary primers, hybridizing to either strand of DNA, ligations can be amplified to generate microgram quantities of repeat containing DNA. Additionally, expanded repeats generated by rolling circle amplification can be produced in vectors for expression of expanded CUG (CUG^exp) RNA capable of sequestering MBNL1 protein in cell culture. <u>A</u>mplification of <u>d</u>imerized <u>e</u>xpanded <u>r</u>epeats (ADER) opens new possibilities for studies of repeat instability and pathogenesis in myotonic dystrophy, a neurological disorder caused by an expanded CTG repeat.

INTRODUCTION

Microsatellite repeat expansions cause about 20 different neurological disorders (1,2). For a subset of these, the causative mutation involves expansions of several hundreds or thousands of repeats located in non-protein-coding regions of the genome. This group of disorders includes myotonic dystrophy, the most common form of muscular dystrophy in adults. The more prevalent form of myotonic dystrophy, type 1 (DM1), is caused by a CTG repeat expansion in the 3′ untranslated (UTR) region of the *dystrophia myotonica protein kinase* (*DMPK*) gene.

While transgenic and cell culture models have been generated, the repeat expansions are towards the lower end of the disease range or it has been necessary to create interruptions in the repeat sequence (3–7). It has been difficult to create models with larger uninterrupted repeats owing to the technical difficulties associated with cloning CTG repeat expansions. However, these models are important for studies of repeat instability and disease pathogenesis, particularly because the severity of DM1 correlates with repeat length (8). Also, perhaps more importantly, larger repeat expansions of $(CTG)_{500-2000}$ are associated with a congenital-onset form of DM1 that has a qualitatively different phenotype (9).

Traditional methods for cloning expanded repeat tracts have relied on plasmid vectors in *E. coli*. In the case of CTG repeats, despite optimization of host and vector combinations, the upper limit that can be maintained is generally <300, owing to the propensity of *E. coli* to delete repeat tracts [(10), however, see reference (11) for an exception]. *In vitro* ligation of pure CTG repeats is also inefficient owing to strand slippage that changes the overhangs at the ends of the DNA molecules (12). Despite recent technical advances that yield greater ligation efficiencies, it is still difficult to generate ligation products with >300 CTG repeats (12).

Previously RCA has been used to generate artificial repetitive sequences from small circular DNA molecules using *E. coli* DNA polymerase I (13). Herein we describe an alternative strategy to generate highly expanded CTG repeats by a combination of *in vitro* ligation and RCA using φ29 DNA polymerase to amplify DNA. This enzyme has a low error rate, strong strand displacement activity and high processivity. It provides an alternative to cloning using *E. coli* vectors or PCR amplification, especially when the target sequence is repetitive (14,15). RCA amplification products are produced from circular templates and contain many tandem copies of the starting molecule. These tandem copies are then templates for additional DNA synthesis producing a hyperbranched

*To whom correspondence should be addressed. Tel: +44 1865 282655; Fax: +44 1865 272420; Email: robert.osborne@dpag.ox.ac.uk
Correspondence may also be addressed to Charles A. Thornton. Tel: +1 585 275 6377; Fax: +1 585 273 1255; Email: charles_thornton@urmc.rochester.edu

amplification product containing double-stranded DNA (16). We found that CTG repeats are remarkably stable during φ29 DNA polymerase amplification and have been able to generate microgram quantities of DNA containing $(CTG)_{1920-2080}$ repeats, in the range of repeat expansions that lead to congenital DM1.

METHODS

In vitro ligations and cloning

To create $pHC_{120-130}$, the CTG repeat tract from pACT [$(CTG)_{120-130}$ cloned at the BbsI-BsaI sites in pACYC177] was excised by BamHI and EcoRI, blunt ended and ligated into the blunt cloning site in pSMART HCKan (Lucigen, accession no AF532107). Plasmids were selected with repeat tracts orientated so that the lagging strand template encoded CAG repeats. $pHC_{240-260}$ and $pH_{480-520}$ were cloned by sequential dimerization reactions. For dimerization reactions, vectors or amplification products were digested with BbsI-AlwNI. A second aliquot was digested with BsaI-AlwNI. Fragments containing the CTG repeat tract were gel purified from each digest using the Zymoclean™ Gel DNA recovery kit (Zymo Research) and ligated using 400 units of T4 DNA ligase in $1 \times$ ligase buffer [New England Biolabs (NEB), MA] at room temperature for 1 h. When BbsI-AlwNI repeat and non-repeat fragments were of similar length, we supplemented this digestion with PvuI. Vectors containing CTG repeats were sequenced using primers SL1 5′-CAGTCCA GTTACGCTGGAGTC-3′ or SR2 5′-GGTCAGGTA TGATTTAAATGGTCAGT-3′ (Lucigen) to confirm that they contained uninterrupted CTG repeats. For experiments using pHC without repeats, a NotI linker was cloned at the blunt cloning site of pSMART HCKan.

Culture of repeat-containing vectors in *E. coli*

We used previously described conditions to culture transformed HB101 *E. coli* to minimize instability of expanded CTG repeats (12). Cells were grown in Luria–Bertani broth supplemented with $10 \mu g/ml$ kanamycin at $30°C$, and cultures were harvested before stationary phase (OD_{650} of 0.7).

φ29 amplification reactions

A total of 100 ng or 150 ng of template DNA in dH_2O and 100 pmoles of each exonuclease-resistant primer containing phosphothioate linkages (*) 5′-GCGATTCCGACTC GTCCA*A*C-3′ and 5′-AACGGGAAACGTCTTGCT* C*G-3′ in 10 mM Tris, 0.1 mM EDTA, pH 8.0 were heated to $95°C$ for 3 min, and cooled to $4°C$ in a thermocycler. For dimerization of plasmid-derived template we typically used 100 ng of template and for dimerization of RCA products, where more template was available, 150 ng. Template and primers were then added to a $30 \mu l$ reaction containing 1 mM each of the four dNTPs, $6 \mu g$ BSA, $1 \times$ φ29 buffer and 10 units of φ29 DNA polymerase (NEB). For control reactions template DNA was substituted with dH_2O or φ29 DNA polymerase was omitted. Reactions were incubated in a thermocycler at

$37°C$ for 18 h and then at $65°C$ for 10 min to heat inactivate φ29 DNA polymerase before storage at $4°C$.

Analysis of amplification reactions

For restriction endonuclease analysis, $1 \mu l$ of amplification product or 500 ng miniprep DNA was digested with enzyme. One-half of each digest was analysed by electrophoresis on 1% SeaKem LE agarose gels (Cambrex Bio Science, Rockland, ME) and visualized by staining with ethidium bromide. For sequencing of RCA products, we gel purified linearized amplification products after restriction digestion with AlwNI. To sequence across repeat tracts we used primers SL1 and SR2. To sequence across the AlwNI junction RCA products were first digested with XhoI, and then sequenced with 5′-TTCTCCCTTCGG GAAGCGTG-3′.

DNA quantitation

DNA was mixed 1:1 with Power SYBR Green PCR Master Mix (Applied Biosystems), which contains SYBR Green I and internal ROX Reference Dye. Fluorescence intensity was measured on an Applied Biosystems 7900HT Sequence Detection System and normalized to internal ROX Reference Dye. Fluorescent measurements were compared to standard curves constructed using UltraPure™ Salmon Sperm DNA Solution (Invitrogen, Carlsbad, CA).

CUG^{exp} expression construct

The vector for expression of CUG^{exp} RNA was constructed by inserting a linker (5′-GGCCACATGTC CGAATGCTAGCCTGCAGATTCGAGGTAACCGG TCCAAGATCT-3′) at the NotI site of pHC and then inserting the CMV IE promoter from pEGFP-C1 (Clontech) at PciI-NheI and a modified *DMPK* 3′UTR at PstI-BglII. The modified *DMPK* 3′UTR contained non-palindromic sequences for SfiI and BsteII restriction endonucleases inserted at the BsmI site. The SfiI and BsteII sites were separated by a 1.2-kb stuffer fragment that allowed resolution of SfiI-BsteII digested vector from that digested with only one enzyme. We also replaced the cleavage and polyadenylation signals from *DMPK* with those from murine *Acta1*, because the native *DMPK* signals are inefficient (17). We then modified $pHC_{120-130}$ by insertion of SfiI and BsteII sites surrounding the BsaI and BbsI sites to create $pDWD_{120-130}$, and then dimerized repeat tracts and maintained $pDWD_{240-260}$ in *E. coli*. This vector was then dimerized and sequential *in vitro* ligation and RCA amplifications were used to generate $(CTG)_{960-1040}$ repeats. The repeats were ligated into the expression vector using SfiI and BsteII, amplified by RCA and sequence confirmed.

Fluorescence *in situ* hybridization and immunofluorescence (FISH-IF)

HEK293 cells were plated at a density of 1.3×10^5 cells per well into 6-well plates containing pre-sterilized coverslips. Cells were cultured in DMEM supplemented with 1% penicillin–streptomycin (Gibco) and 10% fetal bovine

serum (Invitrogen). Cells were then transfected 24 h after plating with 100 ng of gel-purified AlwNI linearized RCA products using Fugene6 (Roche, Indianapolis, IN) according to the manufacturer's instructions. Twenty-four hours after transfection cells were washed twice with PBS then fixed in 3% paraformaldehyde PBS for 30 min, permeabilized in 0.5% Triton PBS for 5 min and then prehybridized in 30% formamide and 2 SSC at room temperature for 10 min. Next, cells were hybridized with probe (0.2 ng/µl) for 2 h at 37°C in buffer (30% formamide, $2 \times$ SSC, 0.02% BSA, 66 µg/ml yeast rRNA, 2 mM vanadyl complex) and then washed for 30 min in 30% formamide/$2 \times$ SSC at 42°C followed by $1 \times$ SSC for 30 min at room temperature. The CAG-repeat probes for FISH were HPLC purified 2-*O*-methyl oligoribonucleotide 20-mers conjugated at the 3' end with Texas Red (IDT, Coralville, IA). Cells were then incubated with polyclonal anti-MBNL1 antibody in PBS (A2764 at 1:5000) (18) overnight at 4°C, washed five times in PBS for 2 min, then incubated in secondary antibody (Alexa 488-labeled goat anti-rabbit polyclonal) and 33 nM diamidino-2-phenylindole (DAPI) for 30 min at room temperature. Cells were washed five times in PBS prior to mounting. Images were obtained on a Nikon Eclipse E600 microscope fluorescence microscope equipped with a Photometrics Coolsnap HQ camera using Metaview version 6.1r5 (Universal Imaging) software.

RESULTS

Critical factors affecting CTG•CAG repeat stability in plasmid vectors include orienting the repeats relative to the replication origin so that the lagging strand template for DNA synthesis encodes CAG repeats (10) and minimizing transcription through repeat sequences (19). We therefore cloned $(CTG)_{120-130}$ repeats in the stable orientation using a vector that contains transcriptional terminators (pSMART HCKan, Lucigen), to create vector $pHC_{120-130}$. We found that CTG repeats in vector $pHC_{120-130}$, or when dimerized to produce vector $pHC_{240-260}$, could be maintained in HB101 host cells, an *E. coli* strain previously shown to reduce the frequency of CTG repeat deletions (10). Consistent with previous observations (10), larger repeat tracts, from dimerization of $pHC_{240-260}$, were not stable (data not shown). Given the difficulty of maintaining larger CTG repeats in *E. coli* we used a combination of *in vitro* ligation followed by RCA to amplify the circularized ligation products (Figure 1).

RCA reactions often rely on random hexamer primers that contain phosphothioate linkages to prevent degradation by the 3'–5' exonuclease activity of φ29 DNA polymerase. We avoided random hexamer primers because of concerns that initiating DNA synthesis within a CTG•CAG repeat tract would produce repeat tracts that contained phosphothioate linkages, with unknown downstream effects on transcription and repeat stability. Instead, we used two non-complementary primers in a hyperbranched-RCA reaction (16). In this reaction one primer binds to circular DNA template and initiates an
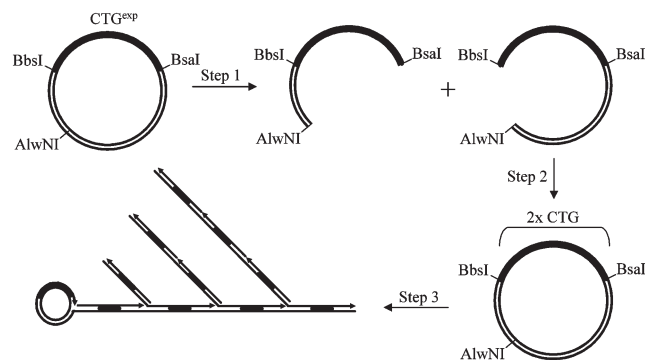

Figure 1. Diagram of the amplification procedure. In step 1, plasmid DNA is digested by BbsI-AlwNI and BsaI-AlwNI. In step 2, the repeat containing fragments are gel purified and ligated to generate a circular double-stranded (ds) DNA template with an uninterrupted dimerized repeat tract. In step 3, dsDNA template is denatured and primers are annealed. Template DNA is then amplified by RCA generating hyperbranched DNA that can be used to initiate the next cycle of repeat dimerization.

RCA reaction. The other primer binds to each copy generated by the rolling circle, initiating sequential primer extension reactions (Figure 1). Because the circular DNA template is generated by denaturation of double-stranded vector DNA, either primer can initiate an RCA reaction.

Initial experiments using the pHC vector without CTG repeats were used to establish conditions for the RCA. Using 100 ng of plasmid DNA as template we were able to detect reaction product in an RCA reaction at 37°C but not at 30°C (data not shown). Restriction endonuclease digestions of amplification products with XhoI or HindIII yielded fragments at the expected size, but no products were produced when template or φ29 DNA polymerase were omitted from the reaction (Figure 2A). We then used RCA to amplify 100 ng of plasmid DNA prepared from a vector containing CTG repeats, $pHC_{120-130}$. Digestion of amplification products with XhoI, HindIII or Acc65I-BsrGI verified that the repeats were not deleted during the RCA reaction (Figure 2B). Digestions also yielded some minor bands which correspond in size to digested amplification products that terminate at one of the primers used for RCA. We confirmed that the $(CTG)_{120-130}$ repeats were faithfully copied by sequencing plasmid DNA and amplification products. These sequencing reads included the entire CTG repeat tract, confirming that the plasmid DNA template and duplicate RCA products contained the same number ($n = 124$) of uninterrupted repeats.

To dimerize CTG repeats we used an approach similar to previously described methods for seamless cloning of CTG repeats (3,20). Two type II restriction endonucleases (BbsI and BsaI) flanked the CTG repeat tract in $pHC_{120-130}$. These enzymes cut outside their recognition site and within the repeat tract to generate complementary overhangs (5'-AGCA-3' or 5'-TGCT-3'). Religation of these overhangs creates an uninterrupted CTG repeat. $pHC_{120-130}$ was digested with BbsI and AlwNI or with BsaI and AlwNI. The repeat-containing fragments
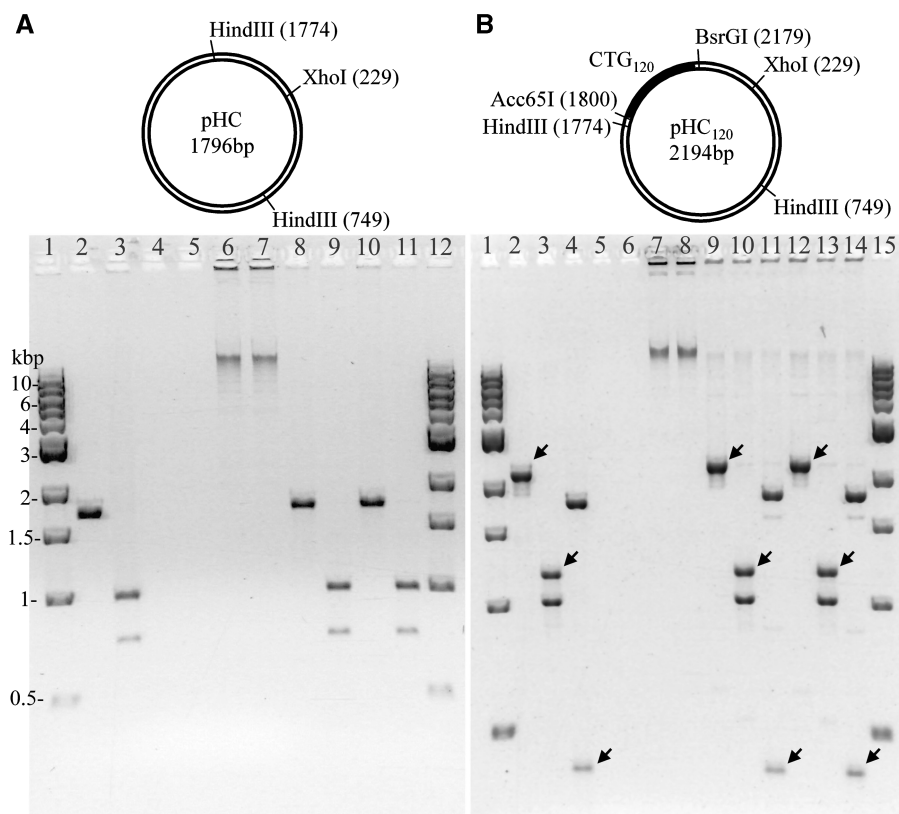
Figure 2. RCA of pHC or pHC$_{120-130}$ plasmid DNA. (**A**) Restriction endonuclease digest of plasmid pHC generated a linear 1796 bp fragment with XhoI (lane 2), or 771 bp and 1025 bp fragments with HindIII (lane 3). RCA of plasmid pHC is shown in lanes 4–11. No RCA product was observed in the absence of template (lane 4) or φ29 DNA polymerase (lane 5). RCA products from duplicate reactions containing template and φ29 DNA polymerase migrated as high molecular weight DNA (lanes 6 and 7), slower than the largest marker DNA (10 kb). Restriction endonuclease digest of duplicate amplification products using XhoI (lanes 8 and 10) or HindIII (lanes 9 and 11) generated the same pattern as plasmid pHC template. Lanes 1 and 12 contain 1 kb DNA ladder (NEB). (**B**) Restriction endonuclease digest of plasmid pHC$_{120-130}$ generated a linear 2194–2224 bp fragment with XhoI (lane 2), 1169–1199 bp and 1025 bp fragments with HindIII (lane 3) and 379–409 bp and 1815 bp with Acc65I and BsrGI (lane 4). Acc65I and BsrGI cleave on either side of the repeat tract. RCA products were observed from duplicate reactions containing both template and φ29 polymerase (lanes 7 and 8) but not in the absence of template (lane 5) or φ29 polymerase (lane 6). Restriction endonuclease digest of duplicate RCA products using XhoI (lanes 9 and 12), HindIII (lanes 10 and 13) or Acc65I-BsrGI (lanes 11 and 14) generated the same pattern as plasmid pHC$_{120-130}$. Fragments containing repeats are marked with arrows. Lanes 1 and 15 contain 1 kb DNA ladder (NEB).

from both digests were isolated, as shown in Figure 1. All overhangs in these fragments are non-palindromic, therefore their ligation can generate only one circular product or concatemers of this product. The circular product contains an uninterrupted dimerized repeat tract and a reconstituted vector backbone (Figure 1). We determined that the ligation reaction generated uninterrupted repeats by sequencing eight pHC$_{240-260}$ clones maintained in *E. coli*. Sequencing reads obtained from either end of the repeat tract extended across the ligation junction [(CTG)$_{>145}$], confirming that BbsI and BsaI have generated the expected overhangs and that ligation of these fragments maintains an uninterrupted CTG repeat.

We next sought to determine whether we could amplify DNA from a ligation reaction containing a larger repeat tract. We first dimerized the repeat tract in vector pHC$_{240-260}$ and then used RCA to amplify the ligation reaction. Analysis of amplification products with restriction endonucleases verified that the size of the repeat tract was consistent with a dimerized repeat containing (CTG)$_{480-520}$ (Figure 3A). For duplicate reactions,

we sequence verified >290 repeats from either end of the repeats, again confirming that the ligation junction within the repeats formed an uninterrupted repeat tract. We then used sequential ligation and RCA reactions to generate increasingly expanded CTG repeat tracts with (CTG)$_{960-1040}$ and (CTG)$_{1920-2080}$. After RCA we quantified double-stranded DNA using SYBR Green I fluorescence, and found that in each case (*n* = 4) ligation reactions were amplified by more than 1000-fold to generate >180 µg of product DNA containing (CTG)$_{960-1040}$ or (CTG)$_{1920-2080}$ in a single reaction. Both (CTG)$_{960-1040}$ and (CTG)$_{1920-2080}$ products were analysed by restriction endonuclease digestion (Figure 3B), sequencing from either end of the repeats, and sequencing across the AlwNI junction. Both of these RCA reactions generated additional minor products with undimerized repeat tracts. These products were obtained in multiple reactions and probably occurred because incompletely digested template was included in the ligation reactions. The most likely source of this template was from failure to resolve BsaI-AlwNI repeat containing fragments from those digested with only one enzyme.
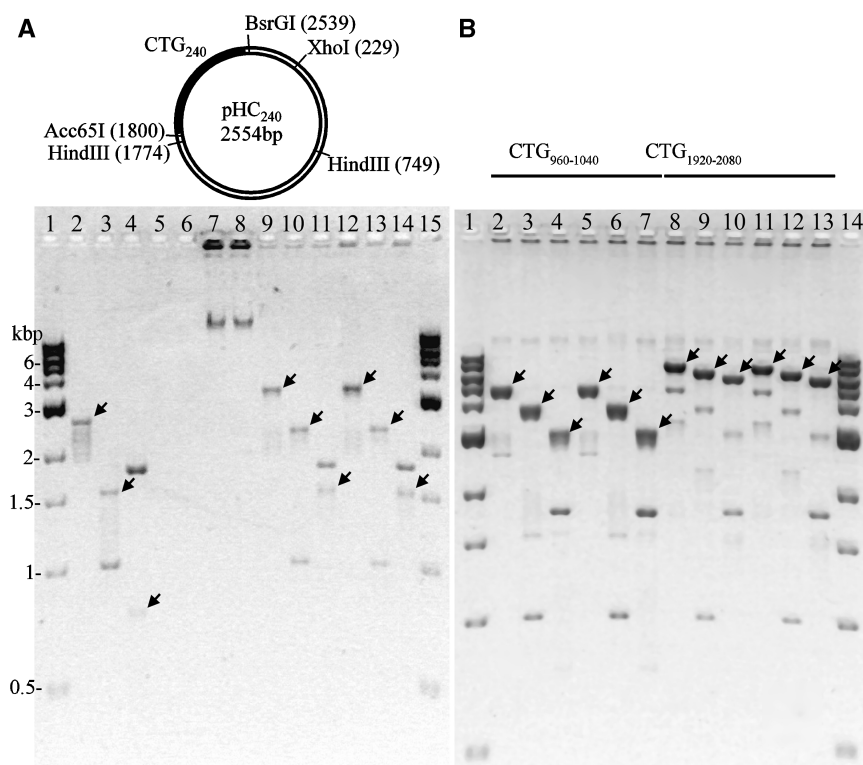
Figure 3. RCA of plasmid or ligation reactions containing highly expanded CTG repeats. (**A**) Restriction endonuclease digest of plasmid $pHC_{240–260}$ generated a linear 2554–2614 bp fragment with XhoI (lane 2), 1529–1589 bp and 1025 bp fragments with HindIII (lane 3), and 739–799 bp and 1815 bp fragments with Acc65I-BsrGI (lane 4). Ligation reactions containing $pHC_{240–260}$ with a dimerized CTG repeat were used as template for RCA. No amplification was observed in the absence of template (lane 5) or $\phi29$ DNA polymerase (lane 6), but products were obtained from duplicate reactions containing both template and $\phi29$ DNA polymerase (lanes 7 and 8). Restriction endonuclease digest of duplicate amplification products using XhoI (lanes 9 and 12), HindIII (lanes 10 and 13) or Acc65I-BsrGI (lanes 11 and 14) generated a vector backbone at the same size as for $pHC_{240–260}$ and a dimerized repeat tract containing $(CTG)_{480–520}$. Fragments containing repeats are marked with arrows. Lanes 1 and 15 contain 1 kb DNA ladder (NEB). (**B**) RCA product containing $(CTG)_{960–1040}$ was dimerized and then amplified by RCA. Restriction endonuclease digests using XhoI (lanes 2 and 5), HindIII (lanes 3 and 6) and BsrGI-Acc65I (lanes 4 and 7) on duplicate RCA reactions containing $(CTG)_{960–1040}$ showed expected fragment sizes. RCA product containing $(CTG)_{960–1040}$ was dimerized and then amplified by RCA. Digests of duplicate RCA product using XhoI (lanes 8 and 11), HindIII (lanes 9 and 12) and BsrGI-Acc65I (lanes 10 and 13) showed expected fragment sizes for $(CTG)_{1920–2080}$. For the $(CTG)_{960–1040}$ and $(CTG)_{1920–2080}$ reactions additional minor products corresponding in size to undimerized repeats were generated. Fragments containing repeats are marked with arrows. Lanes 1 and 14 contain 1 kb DNA ladder (NEB).

Expression of RNA containing expanded CUG repeats in DM1 results in the sequestration of splicing factors in the muscleblind (MBNL) family (21,22). To determine whether DNA generated by RCA would support the synthesis of expanded CUG repeats in cells, we ligated $(CTG)_{960–1040}$ repeats into an expression vector, amplified the ligation products by RCA, linearized the RCA product, and transfected the resultant DNA into HEK293 cells. *In situ* hybridization using a fluorescent-labeled CAG repeat probe showed nuclear foci of CUG^{exp} RNA in cells transfected with RCA product, but not in mock-transfected controls. To determine whether MBNL1 was recruited to nuclear RNA foci we combined fluorescence *in situ* hybridization (FISH) using CAG repeat probes with MBNL1 immunofluorescence. In mock-transfected cells no foci were present and MBNL1 was diffusely distributed in nuclei, whereas in cells transfected with RCA products MBNL1 was recruited into RNA foci (Figure 4). Foci were not observed on hybridization with CUG-(sense) or GUC-repeat probes (data not shown), making it unlikely that foci were the result of probe binding to nuclear proteins or DNA.

DISCUSSION

Generation of expanded DNA tracts of > 300 CTG repeats is an important goal for the development of new models of repeat instability and pathogenesis for DM1. So far these studies have been hampered by the difficulty of maintaining > 300 CTG repeats in *E. coli* and the inefficient ligation of CTG repeat concatemers *in vitro* (10,12). Here we described ADER, a cell-free method for generating hyperexpanded CTG repeats that involves rolling circle amplification of *in vitro* ligated CTG repeats. ADER has several advantages over traditional methods; (i) unlike cloning in *E. coli*, RCA does not delete CTG repeats; (ii) doubling in size of the CTG repeat can be easily monitored at each step; (iii) sufficient DNA is synthesized to allow sequence verification of amplification products and other downstream applications; and (iv) expanded repeats can be produced in vectors for expression of CUG^{exp} RNA.

CTG repeats can form non-classical slipped-strand DNA structures, such as hairpins, which can be bypassed by the DNA replication machinery and thus deleted (23).
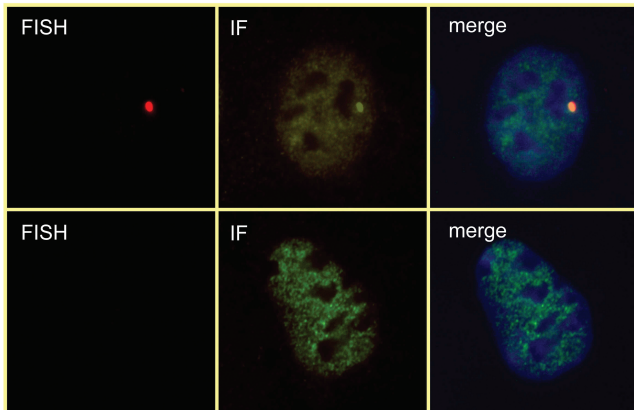
Figure 4. RCA products support synthesis of CUG^{exp} RNA. Fluorescence-labeled CAG repeat probe shows RNA foci (red) in cells transfected with RCA products (top-left) but not mock-transfected cells (bottom-left). Immunofluorescence for MBNL1 (green) shows nuclear foci in cells transfected with RCA products (top-middle) and a speckled distribution of MBNL1 in mock-transfected cells (bottom-middle). Merged images show nuclear foci of RNA and MBNL1 are colocalized in nuclei (blue) of cells transfected with RCA products (top-right) but not in mock-transfected cells (bottom-right). Ribonuclear foci were observed in 5% of cells, consistent with low efficiency from transfection of linear DNA (27).

The accurate copying of expanded CTG repeats is not essential for φ29 replication, since the largest CTG repeat in the φ29 genome is (CTG)$_2$. The φ29 DNA polymerase has intrinsic strand displacement activity which allows it to replicate DNA in the absence of processivity factors and DNA helicases (14). The crystal structure of φ29 DNA polymerase suggests that the downstream DNA template passes through a tunnel before entering the active site (24). The narrow dimensions of the tunnel (<10 Å long and ~10 Å in diameter) make it too small to accommodate double-stranded DNA and require the separation of template and non-template strands. It is possible that the tunnel separates non-classical slipped-strand DNA structures before they enter the active site of the DNA polymerase. This mechanism would prevent bypassing of intramolecular hairpin structures by φ29 DNA polymerase and would thus allow accurate copying of CTG repeats. Intriguingly, this mechanism may also allow φ29 DNA polymerase to accurately copy other repeats, such as trinucleotide (CGG•CCG, GAA•TTC), tetrameric (CCTG•CAGG), pentameric (AATCT•AGATT), hexameric (TTAGGG•CCCTAA) and dodecameric (C$_4$GC$_4$GCG•CGCG$_4$CG$_4$) repeats involved in neurological disease (1,2) or telomere length maintenance (25,26).

ACKNOWLEDGEMENTS

Conflict of interest statement. None declared.

REFERENCES

1. Gatchel,J.R. and Zoghbi,H.Y. (2005) Diseases of unstable repeat expansion: mechanisms and common principles. *Nat. Rev. Genet.*, **6**, 743–755.
2. Mirkin,S.M. (2007) Expandable DNA repeats and human disease. *Nature*, **447**, 932–940.
3. Mankodi,A., Logigian,E., Callahan,L., McClain,C., White,R., Henderson,D., Krym,M. and Thornton,C.A. (2000) Myotonic dystrophy in transgenic mice expressing an expanded CUG repeat. *Science*, **289**, 1769–1773.
4. Amack,J.D. and Mahadevan,M.S. (2001) The myotonic dystrophy expanded CUG repeat tract is necessary but not sufficient to disrupt C2C12 myoblast differentiation. *Hum. Mol. Genet.*, **10**, 1879–1887.
5. Seznec,H., Agbulut,O., Sergeant,N., Savouret,C., Ghestem,A., Tabti,N., Willer,J.C., Ourth,L., Duros,C., Brisson,E. *et al.* (2001) Mice transgenic for the human myotonic dystrophy region with expanded CTG repeats display muscular and brain abnormalities. *Hum. Mol. Genet.*, **10**, 2717–2726.
6. de Haro,M., Al-Ramahi,I., De Gouyon,B., Ukani,L., Rosa,A., Faustino,N.A., Ashizawa,T., Cooper,T.A. and Botas,J. (2006) MBNL1 and CUGBP1 modify expanded CUG-induced toxicity in a Drosophila model of myotonic dystrophy type 1. *Hum. Mol. Genet.*, **15**, 2138–2145.
7. Mahadevan,M.S., Yadava,R.S., Yu,Q., Balijepalli,S., Frenzel-McCardell,C.D., Bourne,T.D. and Phillips,L.H. (2006) Reversible model of RNA toxicity and cardiac conduction defects in myotonic dystrophy. *Nat. Genet.*, **38**, 1066–1070.
8. Logigian,E.L., Moxley,R.T.t., Blood,C.L., Barbieri,C.A., Martens,W.B., Wiegner,A.W., Thornton,C.A. and Moxley,R.T.,3rd. (2004) Leukocyte CTG repeat length correlates with severity of myotonia in myotonic dystrophy type 1. *Neurology*, **62**, 1081–1089.
9. Tsilfidis,C., MacKenzie,A.E., Mettler,G., Barcelo,J. and Korneluk,R.G. (1992) Correlation between CTG trinucleotide repeat length and frequency of severe congenital myotonic dystrophy. *Nat. Genet.*, **1**, 192–195.
10. Kang,S., Jaworski,A., Ohshima,K. and Wells,R.D. (1995) Expansion and deletion of CTG repeats from human disease genes are determined by the direction of replication in E. coli. *Nat. Genet.*, **10**, 213–218.
11. Sarkar,P.S., Chang,H.C., Boudi,F.B. and Reddy,S. (1998) CTG repeats show bimodal amplification in E. coli. *Cell*, **95**, 531–540.
12. Kim,S.H., Cai,L., Pytlos,M.J., Edwards,S.F. and Sinden,R.R. (2005) Generation of long tracts of disease-associated DNA repeats. *BioTechniques*, **38**, 247–253.
13. Fire,A. and Xu,S.Q. (1995) Rolling replication of short DNA circles. *Proc. Natl. Acad. Sci. USA*, **92**, 4641–4645.
14. Blanco,L., Bernad,A., Lazaro,J.M., Martin,G., Garmendia,C. and Salas,M. (1989) Highly efficient DNA synthesis by the phage phi 29 DNA polymerase. Symmetrical mode of DNA replication. *J. Biol. Chem.*, **264**, 8935–8940.
15. Hutchison,C.A.,3rd, Smith,H.O., Pfannkoch,C. and Venter,J.C. (2005) Cell-free cloning using phi29 DNA polymerase. *Proc. Natl. Acad. Sci. USA*, **102**, 17332–17336.
16. Lizardi,P.M., Huang,X., Zhu,Z., Bray-Ward,P., Thomas,D.C. and Ward,D.C. (1998) Mutation detection and single-molecule counting using isothermal rolling-circle amplification. *Nat. Genet.*, **19**, 225–232.
17. Philips,A.V., Timchenko,L.T. and Cooper,T.A. (1998) Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy. *Science*, **280**, 737–741.
18. Lin,X., Miller,J.W., Mankodi,A., Kanadia,R.N., Yuan,Y., Moxley,R.T., Swanson,M.S. and Thornton,C.A. (2006) Failure of

MBNL1-dependent post-natal splicing transitions in myotonic dystrophy. *Hum. Mol. Genet.*, **15**, 2087–2097.

19. Bowater,R.P., Jaworski,A., Larson,J.E., Parniewski,P. and Wells,R.D. (1997) Transcription increases the deletion frequency of long CTG.CAG triplet repeats from plasmids in Escherichia coli. *Nucleic Acids Res.*, **25**, 2861–2868.

20. Tian,B., White,R.J., Xia,T., Welle,S., Turner,D.H., Mathews,M.B. and Thornton,C.A. (2000) Expanded CUG repeat RNAs form hairpins that activate the double-stranded RNA-dependent protein kinase PKR. *RNA*, **6**, 79–87.

21. Miller,J.W., Urbinati,C.R., Teng-Umnuay,P., Stenberg,M.G., Byrne,B.J., Thornton,C.A. and Swanson,M.S. (2000) Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy. *EMBO J.*, **19**, 4439–4448.

22. Mankodi,A., Urbinati,C.R., Yuan,Q.P., Moxley,R.T., Sansone,V., Krym,M., Henderson,D., Schalling,M., Swanson,M.S. and Thornton,C.A. (2001) Muscleblind localizes to nuclear foci of aberrant RNA in myotonic dystrophy types 1 and 2. *Hum. Mol. Genet.*, **10**, 2165–2170.

23. Pearson,C.E., Nichol Edamura,K. and Cleary,J.D. (2005) Repeat instability: mechanisms of dynamic mutations. *Nat. Rev. Genet.*, **6**, 729–742.

24. Kamtekar,S., Berman,A.J., Wang,J., Lazaro,J.M., de Vega,M., Blanco,L., Salas,M. and Steitz,T.A. (2004) Insights into strand displacement and processivity from the crystal structure of the protein-primed DNA polymerase of bacteriophage phi29. *Mol. Cell*, **16**, 609–618.

25. Moyzis,R.K., Buckingham,J.M., Cram,L.S., Dani,M., Deaven,L.L., Jones,M.D., Meyne,J., Ratliff,R.L. and Wu,J.R. (1988) A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. *Proc. Natl. Acad. Sci. USA*, **85**, 6622–6626.

26. Morin,G.B. (1989) The human telomere terminal transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats. *Cell*, **59**, 521–529.

27. Weintraub,H., Cheng,P.F. and Conrad,K. (1986) Expression of transfected DNA depends on DNA topology. *Cell*, **46**, 115–122.