Structures of CUG Repeats in RNA

POTENTIAL IMPLICATIONS FOR HUMAN GENETIC DISEASES*

Received for publication, March 7, 2002, and in revised form, June 17, 2002
Published, JBC Papers in Press, June 19, 2002, DOI 10.1074/jbc.M202235200

Philip Pinheiro‡, Garry Scarlett‡, Alison Rodger§, P. Mark Rodger§, Anna Murray§,
Tom Brown**, Sarah F. Newbury‡‡, and James A. McClellan‡

From the §Biophysics Laboratories, School of Biological Sciences, University of Portsmouth, St. Michael’s Building,
White Swan Road, Portsmouth, PO1 2DT, United Kingdom, the ¶Wessex Regional Genetics Laboratory, Salisbury Health
Care National Health Service Trust, Salisbury District Hospital, Salisbury, Wiltshire SP2 8BJ, United Kingdom,
the **Department of Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ, United Kingdom,
the ‡Department of Biochemistry, University of Warwick, Coventry CV4 7AL, United Kingdom, and
the ‡‡Department of Biochemistry, South Parks Road, Oxford OX1 3QU, United Kingdom

Triple repeat expansion diseases (TREDs)1 are genetic disorders that are caused by expanded tracts of repeated
sequences (usually CNG) (reviewed in Ref. 1). Naturally vari-
able lengths of these tracts occur in several genes involved in
neurological disorders. Up to about 30 trinucleotide repeats do
not cause neurological defects, but once a certain critical length
is exceeded, disease ensues. The disease-causing alleles are
usually dominant. The TREDs are complex syndromes that
may exhibit anticipation (genetic instability leading to longer
tract lengths with each generation). In several cases there is
evidence that the longer the tract, the more severe or the
earlier the onset of disease. Proof that long tracts cause disease
rather than merely being correlated with it was obtained by
artificially introducing long CNG tracts into mice, thereby
inducing a version of spinocerebellar ataxia type I (2). In those
experiments some, but not all, features of the syndrome were
reproduced in the mouse model, and in fact it is known that
other symptoms of Fragile X are absent and that protein
associated with learning difficulties in boys, despite the fact
that other symptoms of CGG TREDs, because if it did
there would be a class of CA(Purine) diseases (CAA is another
codon for glutamine). Furthermore, in the case of Fragile X it
appears that absence of the protein and/or RNA in the cyto-
plasm causes the main features of the disease (9). There are
several mechanisms that can lead to this result. One possibil-
ity is that expanded-tract alleles of FMR-1 are hypermethylated
perhaps because CGG repeats are particularly good substrates
for methyltransferases; (10)), and this is associated with ge-
netic shutdown. That absence of gene function can explain the
condition is shown by the existence of rare deletions in FMR-1,
leading to symptoms of Fragile X (11–13). Interestingly, inter-
mediate-sized alleles (41–60 repeats) of the tract in FMR1 are
associated with learning difficulties in boys, despite the fact
that other symptoms of Fragile X are absent and that protein
levels are normal (14). However, the totality of the syndrome
also includes a DNA effect: the expanded-tract chromosomes
themselves are physically fragile and genetically hypervariable.
This may be a general property of CNG tracts, since these
are known to mediate genetic instability not only in eukaryotes
but also in Escherichia coli (16–18).2 Similar behavior of other
repeated sequences was previously shown to reflect cellular
reaction to the in vivo formation of non-B DNA secondary
structures (19, 20). Less attention has been paid to the role of
the RNA. Nevertheless, since in all cases the CNG tracts are

* This work was supported by the National Health Service South
West Region Research and Development Directorate and the Engineer-
ing & Physical Sciences Research Council Life Sciences Interface: GR/
M91105. The costs of publication of this article were defrayed in part by
the payment of page charges. This article must therefore be hereby
marked “advertisement” in accordance with 18 U.S.C. Section 1734
solely to indicate this fact.

† To whom correspondence should be addressed: Dept. of Chemistry,
University of Warwick, Coventry CV4 7AL, UK. Tel.: 44-24-76523234;
Fax: 44-24-76524112; E-mail: A.Rodger@warwick.ac.uk.

‡ The abbreviations used are: TRED, trinucleotide repeat expansion
disease; NOE, nuclear Overhauser effect.

G. Scarlett and P. Pinheiro, manuscript in preparation.
transcribed but in the loci associated with the two commonest diseases (Fragile X and myotonic dystrophy), they are not translated, it is likely that effects at this level are of central importance, at least for the normal function of the tracts (21, 22). As for how the expanded tracts actually cause disease, the complex nature of the syndromes suggests that this is likely to involve all these aspects. For example, in cases where the tracts are translated, the production of mutant protein may have effects on the cell; if the expanded tracts form a novel structure in DNA, this may act to stimulate recombination or to compromise replication, thus accounting for anticipation and chromosomal fragility; and formation of a new structure in RNA may lead to abnormal stability, splicing, localization, or translation.

The clearest candidate for a CNG disease mediated largely by effects at the RNA level is myotonic dystrophy (21). The shortest repeats that are known from the general population are about 15 nucleotides in length (five repeats). The shorter abnormal alleles are associated with adult-onset myotonic dystrophy and the longer ones with the congenital form. In this case a CTG repeat (CUG in the RNA) is untranslated and located in a locus involved in neuromuscular development. The disease-causing allele is dominant. Long RNA containing the expanded tracts has been shown to be transcribed effectively and to accumulate in foci within the nucleus (23).

Our working hypothesis therefore is that there are special features of CNG RNA structure or biochemistry. These special features mediate some useful function in the short, wild-type alleles and possibly also a deleterious function in the long, disease-causing alleles. In this paper we have used in vitro transcription and chemical synthesis to produce CUG RNA of various sizes and compared it with control RNAs, such as GUC (different polarity) and a randomized but isobasic sequence, as well as RNA markers and triplet repeat RNA transcripts were run on 10% denaturing and denaturing (7 M urea (Sigma)) acrylamide (29:1 acrylamide to bisacrylamide (National Diagnostics)) gels. In all cases the electrophoresis buffer was 1× TBE (90 mM Tris, pH 8.0 (Sigma), 90 mM boric acid (Sigma), 2 mM EDTA (Sigma)). Gels were fixed in 10% acetic acid and dried under vacuum onto Whatman 3MM paper, before autoradiography using Kodak film.

**End Labeling of DNA Markers and of Synthetic RNA—DNA oligonucleotide markers 8–32 were purchased from Amersham Biosciences and labeled using T4 polynucleotide kinase (New England Biolabs) and [γ-32P]ATP (3000 Ci/mmol; PerkinElmer Life Sciences). Synthetic RNA was labeled in a similar manner.

**Circular Dichroism and UV Melting—DNA and RNA oligonucleotides for both CD and UV melting were dissolved to a concentration of 4 μM (for DNA) or 8 μM (for RNA) in 5 mM sodium phosphate buffer, pH 7.5, unless otherwise stated. All solutions for both the CD and UV melting experiments were filter-purified using a 0.2-μm nylon filter (Sigma). CD spectra were gathered on a Jasco J-715 CD spectropolarimeter using a 5-mm path length cell. Data were stored using the supplied software and then exported to Kaleidagraph for manipulation. Data were collected over the range 350–200 nm with a 0.5 nm resolution and 1 nm bandwidth at a speed of 100 nm/min; spectra were averaged over 16 scans. For comparison purposes the RNA CD data have been normalized to 1.2 OD260, while the DNA CD spectra have been normalized to 0.6 OD260.

U2 melting was conducted on a Cary 1 spectrophotometer ramped at 1 degree/min and the data collected with the supplied software, before exporting to Kaleidagraph for manipulation. Prior to the melting curve determinations, the samples were heated to 95 °C and cooled slowly (24). Melting curve measurements were repeated at least three times, and no significant differences were found between each set of data. The data were analyzed following Marky and Breslauer (25). The fraction of unfolded RNA in solution was determined by calculating, where x is the difference between the final absorbance of the unfolded RNA and the absorbance at a given temperature, and y is the difference in absorbance between the temperature-dependent absorbance and the base line corresponding to the temperature dependence of the unmelted RNA absorbance. The midpoint of the plot of a versus 1/T, where T is the absolute temperature (in Kelvin), gives the melting temperature (by definition the temperature at which half of the RNA is melted). The van’t Hoff transition enthalpy of the unimolecular transition is, 

\[ \Delta H_{NH} = B' \left( \frac{1}{T_{max}} - \frac{1}{T_2} \right) \]  

where \( B' = 3.50 \text{ cal K}^{-1} \text{ mol}^{-1} \) (25), and \( T_{max} \) is the absolute temperature of the maximum of the a versus 1/T plot, and \( T_2 \) is the absolute temperature of the high temperature half-height of the curve. The entropy of the transition is given as follows.

\[ \Delta S = \frac{\Delta H_{NH}}{T_2} \]

Molecular Modeling—A variety of techniques were used to perform the calculation. AF国道 search in CHARMM version 25.2. These methods included a grid search based on the backbone torsion angles, a Boltzmann jump algorithm, and molecular dynamics simulations at temperatures of both 300 and 600 K. NOE constraints were used in some of the calculations to improve the chance of locating suitable Watson-Crick C-G base pairings within each triplet and also to locate possible stacked triplets; other calculations were performed without constraints. In all cases, the selected conformations were energy mini-
mized at the end of the conformational searches; where NOE constraints were imposed the minimization was initially done with constraints imposed, but then subsequently with the constraints removed.

RESULTS

Electrophoretic Mobility of CUG RNAs—The mobility of end-labeled chemically synthesized (CUG)$_5$ electrophoresed on denaturing gels together with RNA markers (Fig. 1A) shows that the RNA migrates as expected for an 18-mer. Electrophoresis of CUG repeat RNAs and control RNAs on denaturing gels shows that in vitro transcription generates bands of the expected sizes for all the samples (Fig. 1C). When these CUG repeat RNAs are electrophoresed on non-denaturing gels (Fig. 1, B and D), they migrate significantly faster than the marker RNAs. Randomized and GUC repeat RNAs migrate similarly to controls. The overall conclusion to be drawn from Fig. 1, A–D, is that the CUG RNAs are all unusually compact, whether they are made by chemical synthesis or in vitro transcription and that over this size-range their degree of compaction increases with tract length. This is in accord with the behavior previously observed for single-stranded DNA (26).

For all the CUG and GUC repeats, it is interesting to note that extra bands at low molecular weight are observed. These extra small bands are not observed for the randomized control and probably reflect abortive transcription. DNA polymerase apparently also appears to pause at CNG tracts (27). In addition to bands that move faster than the appropriate markers, some that move more slowly, especially in the six-repeat (19 nucleotides) and seven-repeat (22 nucleotides) tracks, are observed. These are not present in the denaturing gel nor in the chemically synthesized sample, so they are probably higher order complexes of RNA; possibly DNA/RNA hybrids that resist DNase I treatment. Multistrand complexes of CNG DNA have been observed previously by other workers (28).

Mobility of Long RNAs from in Vitro Transcription of Plasmids—Structural work on triplet repeats has been on short tracts that cannot be directly compared with those from TRED patients. Recently we have managed to create RNAs of lengths approaching clinical relevance by in vitro transcription of a range of plasmids containing 24–51 CTG repeats by cloning CNG tracts into bacterial plasmids under conditions where phenomena similar to 5-transduction may be observed. The gel mobility of these RNAs is shown in Fig. 1, E and F. These in vitro transcriptions are not marked by the high levels of premature termination seen for the oligonucleotide templates. On the non-denaturing gel (Fig. 1F) the CNG RNAs migrate as two bands which migrate anomalously slowly (in contrast to the anomalously fast migration of the smaller triplet repeats). Comparison with the mobilities of the markers shows this to be consistent with duplex or higher formation. Some sort of structural transition appears to occur between the 5–7 range and the 36–50 repeat range. Since this structural transition occurs at a length close to the threshold for myotonic dystrophy, this is potentially a result of some significance.

CD Spectra of Triplet Repeat RNA and DNA—To investigate the structures of the unusual compact structures found in the gels, CD spectra were collected for the chemically synthesized (CTG)$_5$ and (CUG)$_5$ samples and randomized controls. The CD spectrum of (CTG)$_5$ in phosphate buffer consists of a positive peak at 285 nm and two negative peaks at 258 and 208 nm (Fig. 2A). The 285 nm maximum is at too long a wavelength for the tract to be considered an A-DNA conformation, and the spectra do not contain a positive peak at around 205 nm, which would be expected if the tract were adopting duplex B-DNA structure (29–33). The negative peak at 208 nm suggests that the tract has triplex character, as a negative peak between 200 and 220 nm is considered a hallmark of triplex DNA (29, 30, 32). Spectra for triplexes are usually able to be approximated as the sum of the CD for the component duplex and single-stranded DNA plus some extra intensity due to the more rigid structure (33). In accord with this the observed spectrum closely resembles that expected for the sum of poly(d(G-C))$_2$ and poly(d(T)) (34). Addition of 20 mM NaCl or 20 mM MgCl$_2$ increases the 265 nm and 258 nm bands to the same intensity and increases the 208 nm band. The 208 nm band is particularly enhanced by MgCl$_2$, which probably correlates with a stabilization of base stacking and more effective reduction of phosphate-phosphate repulsion (35–37).

To investigate whether the observed triplex-like structure was inter or intramolecular, labeled (CTG)$_5$ was titrated with unlabeled (CTG)$_5$. Gel electrophoresis showed that there is no evidence for intermolecular structures (Fig. 1G), which would be expected to shift the band up the gel due to an increased mass of the complex. There was also no evidence of concentration dependence of spectra other than that required by the Beer-Lambert Law. We therefore conclude that (CTG)$_5$ forms a monomolecular base stacking structure that has spectroscopic triplex characteristics.

The CD spectra of (CUG)$_5$ RNA and an isobase randomized control at 20 °C are shown in Fig. 2B. The (CUG)$_5$ spectrum has a positive band at 269 nm, a small negative band at 237 nm, and a large negative band at 208 nm. The 208 nm band, which is indicative of RNA base pairing and stacking (38), is not present in the randomized control, suggesting that (CUG)$_5$ RNA is significantly more base-paired and stacked than the control. The long wavelength peak (at 269 nm rather than the 285 nm of the corresponding DNA and 282 nm for the RNA control) is very sensitive to base composition and RNA conformation (38, 39). The overall spectrum is very similar to those obtained for RNAs containing pseudoknots with a small but significant negative band at ~237 nm (40, 41) (naturally occurring A-form RNAs generally lack any marked band at this wavelength (42–44)). The spectrum of the intermolecular triplex formed by poly(A)-poly(G)-poly(C) is also similar to that of the (CUG)$_5$ repeat, except that the small negative band for this intermolecular triplex is at 245 nm (33). Since an intramolecular pseudoknot might be described as a pseudo-triplex, these two descriptions are equally valid.

The Effect of Temperature on (CUG)$_5$ RNA—To analyze the effect of temperature on (CUG)$_5$ RNA, we measured the CD spectra at a range of temperatures between 5 and 80 °C (Fig. 2C). The negative band at 208 nm decreases markedly with increased temperature as the base pairs melt apart (44). At 65 °C, the 208 nm band is absent and the 267 nm band decreases and shifts to 275 nm resembling the randomized control, indicating that the (CUG)$_5$ structure is completely melted. The series of spectra obtained are almost identical to CD spectra of a 59-nucleotide flavivirus pseudoknot at various temperatures (41).

The melting curve for (CUG)$_5$ (Fig. 2D) was analyzed following (25) to give a single transition with melting temperature of 54.8 °C. The van’t Hoff transition enthalpy of the transition was determined from Fig. 2F ($T_m$ = 328.1 and $T_c$ = 336.6 K) to be $\Delta H_{\text{m}} = (190 \pm 30) \text{kJ mol}^{-1}$ and $\Delta S = (580 \pm 90) \text{J K}^{-1} \text{mol}^{-1}$. A single transition is not inconsistent with the possibility that the structure may resemble a pseudoknot; detailed studies of the thermal melting of pseudoknots show that they can melt with one apparent transient (45, 46). However, $T_m$ and $\Delta H$ values are high for a 15-mer of any kind, especially as this usually means <8 base pairs. For example, the $T_m$ for a duplex with a similar ratio of G/C and A/U in 1 M NaCl (GUCUGAC) (versus 5 mM
Fig. 1. A, synthetic r(CUG)$_n$ electrophoresed on a denaturing gel together with RNA markers. B, synthetic r(CUG)$_n$ electrophoresed on a non-denaturing gel together with RNA markers. C, migration of short CUG tract RNA products transcribed from DNA oligonucleotides electrophoresed on a 10% denaturing gel. Tracts of length 16, 19, and 22 bases with RNA markers are shown. Full-length RNAs are indicated together with prematurely terminated products. Transcriptions from random oligonucleotides (R, randomized C$_5$G$_6$U$_5$, right hand lane) produce very few prematurely terminated products. D, RNA products as in C electrophoresed on a 10% non-denaturing gel. E, gel electrophoresis of in vitro transcripts of a range of plasmids containing 24–51 CUG repeats on a denaturing gel together with RNA markers. A double-stranded DNA marker (M) is shown, and the numbers correspond to numbers of base-pairs. F, transcribed RNA products as in E electrophoresed on a 10% non-denaturing gel. G, DNA (CTG)$_n$ oligonucleotides at concentrations of 1–100 $\mu$M on a 15% non-denaturing polyacrylamide gel.
sodium phosphate in our experiments) is 56.2 °C (47) and the Tm and ΔH values for a 23-mer RNA hairpin (with 7 Watson-Crick base pairs) in 10 mM sodium phosphate are 50.9 °C and 193 kJ mol⁻¹, respectively (48). Thus whatever structure is adopted is very stable and is unlikely to be a hairpin.

The RNA containing (CUG)₅₀ by way of contrast has two
Molecular Modeling — A conformational study of a single strand of RNA was undertaken to assess the viability of the proposed structure (see below). This was not intended to be a definitive modeling study, indeed the scope of this study would not allow for such a major computational undertaking, but rather to ensure that the proposed structure was reasonable; in particular it should be thermally accessible and represent at least a local free energy minimum structure.

Calculations were performed on both 5'-CUG and 5'-CUGCUG sequences. In general, low energy conformations involving the formation of three C-G H-bonds within each triplet was found to occur readily and with distances in the range 2–2.2 Å between the “heavy” (non-hydrogen) atoms. A stable stacking of these Watson-Crick base pairs involving the hairpin turn of the backbone envisaged in Fig. 3A was less easy to locate, but was found to occur with energies comparable with those found in the more conventional helical stacking patterns. An example of such a conformation is shown in Fig. 3B. The distance between the two Watson-Crick base pairs is about 4 Å, and the H-bonds within each C-G pair are all in the range 2–2.3 Å. There is some tilting of the base pairs relative to each other, but it is likely that this would be reduced by π-stacking interactions in an extended oligomer. It is also interesting to note that the uracils tend to orient parallel to each other and outside the C-G stack. The distance between the uracils is about 5 Å, which would be close enough to allow for favorable π-stacking interactions and spectroscopic interactions.

DISCUSSION

In this paper, we have used circular dichroism, UV absorption as a function of temperature, and gel electrophoresis to make some progress in understanding the structure(s) adopted by CNG tracts under our conditions. The spectroscopic data show that both (CUG)$_n$ and (CTG)$_n$ adopt highly ordered structures, which are different from that of the control randomized nucleic acids in accord with the gel mobility data. The CD spectrum for (CUG)$_n$, e.g., is predominantly A-form, but resembles spectra obtained for RNA containing a pseudoknot or triplex structure. The CD spectra for (CTG)$_n$ DNA show that it does not adopt a duplex B-form conformation. The structure that is formed is intramolecular and more stable than a random sequence.

We have also shown that CNG RNA forms compact structures and that the structure becomes relatively more compact as tract length increases for short to medium length repeats.

thermal transitions occurring at −40 °C and 76 °C (Fig. 2H) and a CD spectrum resembling that of A-form RNA (Fig. 2G). It is interesting to note that there are two bands when this sequence is run on a non-denaturing gel (Fig. 1F). It may therefore be that the two transitions reflect two distinct species melting rather than one species undergoing a two stage transition.

Fig. 3. A, schematic of the toblerone structure. B, alternate views of a toblerone energy-minimized structure for 5'-CUGCUG resulting from a conformational search where NOE constraints were initially imposed then subsequently removed. Color coding of the atoms is: medium gray, carbon; black, oxygen; dark gray, nitrogen; light gray, hydrogen.

Structures of CUG Repeats
We suggest that the compact structure seen for medium length tracts is biologically important, and is probably required for normal gene function, and that the inability of very short tracts to adopt it may be a factor in maintaining them above a certain minimum length. In support of this hypothesis, in the locus associated with myotonic dystrophy, for example, tracts shorter than 5 repeats are not observed in the normal population (1). Proteins that bind CNG repeats have been isolated (49, 50), and it is possible that the compact structure formed by these repeats binds to proteins that are involved in nuclear export (23). The gel mobilities of long CNG RNA are also consistent with high order structures, if in addition to being more compact they are also significantly more rigid than duplex RNA (since rigidity reduces gel mobility).

But what are the compact structures associated with CNG tracts in nucleic acids? Although all laboratories working in the area agree that short-to-medium length CNG tracts adopt compact structures, there is significant disagreement about the nature of these structures. The structure adopted probably depends on tract length, sequence context, DNA/RNA concentration, and environmental conditions such as temperature and ionic strength (26, 28, 51). The simplest structural model for a CNG repeat is a mismatched hairpin (52–54), and it is possible that under some conditions such a structure is indeed adopted (16, 55). However, a mismatched hairpin does not readily explain the following facts: (i) CNG tracts readily adopt a highly compact structure but GNC tracts do not, (ii) the structure becomes more compact as its length increases (26), (iii) much better hairpin-forming sequences exist in eukaryotic genomes, (iv) higher order structures may mask the binding of proteins that are involved in nuclear export or in other essential functions such as stability and environmental conditions such as temperature and ionic strength (26, 28, 51). The simplest structural model for a CNG hairpin is illustrated in Fig. 3A.

Acknowledgments—We thank Professor Pat Jacobs for inspiration and constructive criticism and Colin Derrick for expert photographic assistance.

REFERENCES

1. Warren, S. T., and Nelson, D. L. (1993) Curr. Opin. Neurobiol. 3, 752–759
2. Burright, E., Clark, H., Servadio, A., Matilla, T., Feddersen, R., Yunis, W., Duvick, L., Zoghbi, H., and Orr, H. (1995) Cell 82, 975–984
3. Goldberg, V., McMurray, C., Zeiser, J., Almyvist, E., Silence, D., Richards, F., Gacy, A., Buchanan, J., Telenius, H., and Hayden, M. R. (1995) Hum. Mol. Genet. 4, 1911–1918
4. Dijan, P., Hancock, P., and Chana, H. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 417–421
5. Davies, S. W., Turmaine, M., Cozens, B. A., DiFiligo, M., Sharp, A. H., Ross, C. C., Scherzinger, E., Wanker, E. E., Mangiarini, L., and Bates, G. F. (1997) Cell 90, 537–548
6. Li, X.-J., Li, S.-H., Sharp, A. H., Nucifora, F. C. J., Schilling, G., Lanahan, A., Worley, P., Snyder, S. H., and Ross, C. A. (1995) Nature 378, 398–402
7. Scherzinger, E., Luz, R., Mann, M., Mangiarini, L., Helenbch, B., Hasenbank, R., Bates, G. P., Davies, S. W., Lehrach, H., and Wanker, E. E. (1997) Cell 90, 549–558
8. Trottier, Y., Lutz, Y., Stevanin, G., Imbert, G., Devys, D., Cancel, G., Saudou, F., Weber, C., David, G., Tora, L., Agid, Y., Bruce, A., and Mandel, J.-L. (1995) Nature 378, 403–406
9. Koo, R. F., Willemsen, R., and Oostra, B. A. (2000) Mol. Med. Today 6, 194–198
10. Smith, S. S., Laayoun, A., Lingeman, R. G., Baker, D. J., and Riley, J. (1994) J. Mol. Biol. 245, 143–151
11. G.T.V. Luguenheil, K. A., Vockley, J. G., Grody, W. W., and Nelson, D. L. (1994) Hum. Mol. Genet. 3, 705–706
12. Luguenheil, K. A., Pieier, A. M., Carlsen, N. L., Chudley, A. E., and Nelson, D. L. (1995) Nature 378, 483–485
13. Trottier, Y., Imbert, G., Fouvka, A., Fryns, J.-P., and Mandel, J.-L. (1994) Am. J. Med. Genet. 51, 454–457
14. Murray, A., Youings, S., Dennis, N., Latsky, L., Linehan, P., McKechnie, N., MacPherson, J., Pound, M., and Jacobs, P. (1996) Hum. Mol. Genet. 5, 727–735
15. Fry, M., and Loeb, L. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 4950–4954
16. Darlow, J., and Leach, D. (1995) Genetics 141, 825–832
17. Johnson, K. H., Gray, D. M., Morris, P. M., and Sutherland, J. C. (1990) Methods Enzymol. 180, 392–404
18. Laing, L., and Draper, D. E. (1994) J. Biol. Chem. 269, 18553–18556
19. Greaves, D., Patient, R., and Lilley, D. (1985) J. Mol. Biol. 185, 461–478
20. McClellan, J. A., Boublikova, P., Palecek, F., and Lilley, D. M. J. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 8253–8257
21. Hamshere, M. G., and Brook, J. D. (1996) Trends Genet. 12, 332–334
22. McClellan, J. A., and Newbury, S. F. (1996) Nature 379, 396
23. Taneja, K. C., McCurrach, M., Schalling, M., Hausman, D., and Singer, R. H. (1995) J. Cell Biol. 128, 995–1002
24. Fugld, J. D., and Tiino, J. O. (1989) Methods in Enzymol. 180, 304–325
25. Marky, L. A., and Breslauer, K. J. (1987) Biopolymers 26, 1601–1620
26. Mitchell, J. E., Newbury, S. F., and McClellan, J. A. (1996) Nucleic Acids Res. 24, 1876–1881
27. Kang, S., Ohshima, K., Shimizu, M., Amiraeh, S., and Wells, R. D. (1995) J. Mol. Biol. 250, 735–837
28. Smith, G. K., Jie, J., Fox, G. E., and Gao, X. L. (1995) Nucleic Acids Res. 23, 4303–4311
29. Park, Y. W., and Breslauer, K. J. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 5653–6605
30. Gray, D. M., Ratcliffe, R. L., and Vaughan, M. R. (1992) Methods Enzymol. 211, 389–405
31. Shina, P. V., and Shafir, R. H. (1991) J. Biol. Chem. 266, 5417–5423
32. Radakrishnan, I., and Patel, D. J. (1994) Biochemistry 33, 11405–11416
33. Chastain, M., and Tiino, J. O. (1992) Nucleic Acids Res. 20, 315–318
34. Johnson, K. H., Gray, D. M., Morris, P. M., and Sutherland, J. C. (1990) Biopolymers 29, 325–339
35. Schrewe, D. C., Chelladururi, B. S., Nicholson, A. W., and Moore, P. B. (1994) Nucleic Acids Res. 22, 604–612
36. Laiing, L., and Draper, D. E. (1994) J. Mol. Biol. 237, 560–576
37. Cole, P. E., Yang, S. K., and Crothers, D. M. (1972) Biochemistry 11, 4358–4368
38. Gray, D. L., Liu, J., Ratcliffe, R. L., and Allen, F. S. (1981) Biopolymers 20, 1357–1362
39. Cauvel, G. C., and Johnson, W. C. J. (1982) Biopolymers 21, 1763–1780
40. Johnson, K. H., and Gray, D. M. (1992) J. Biomol. Struct. Dyn. 9, 733–746
41. Shi, P.-Y., Brinton, M. A., Veal, J. M., Zhong, Y. Y., and Wilson, W. D. (1996) J. Mol. Biol. 254, 535–539.
Structures of CUG Repeats

Biochemistry 35, 4222–4230
35. Lin, L., Roberts, R., Casket, T., and Swanson, M. S. (1996) Nucleic Acids Res. 24, 4407–4414
36. Kohwi, Y., Wang, H., and Kohwi-Shigematsu, T. (1993) Nucleic Acids Res. 21, 5651–5655
37. Gacy, A., Goellner, G., Juranic, N., Macura, S., and McMurray, C. (1995) Cell 81, 533–540
38. Mariappan, S. V., Catasti, P., Chen, X., Ratliff, R., Moyzis, R. K., Bradbury, E. M., and Gupta, G. (1996) Nucleic Acids Res. 24, 784–792
39. Mitas, M., Yu, A., Dill, J., and Haworth, I. S. (1995) Biochemistry 34, 12803–12811
40. Mitas, M., Yu, A., Dill, J., Kamp, T. J., Chambers, E. J., and Haworth, I. S. (1995a) Nucleic Acids Res. 23, 1050–1059
41. McClellan, J. A., Palecek, E., and Lilley, D. M. J. (1986) Nucleic Acids Res. 14, 9291–9309
42. Kuryavyi, V. V., and Jovin, T. M. (1995) Nat. Genet. 9, 339–341
43. Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., and Turner, D. H. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 9373–9377
44. Sarkar, M., Sigurdsson, S., Tomac, S., Sen, S., Rozners, E., Sjoberg, B.-M., Strumborg, R., and Gradlund, A. (1996) Biochemistry 35, 4678–4688
45. Yano-Yanagisawa, H., Li, Y., Wang, H., and Kohwi, Y. (1995) Nucleic Acids Res. 23, 2654–2660
46. Timchenko, L. T., Miller, J. W., Timchenko, N. A., De Vore, D. R., Datar, K. V., and E. M., and Gupta, G. (1996) Nucleic Acids Res. 24, 784–792
47. Mariappan, S. V., Catasti, P., Chen, X., Ratliff, R., Moyzis, R. K., Bradbury, E. M., and Gupta, G. (1996) Nucleic Acids Res. 24, 784–792
48. Mitas, M., Yu, A., Dill, J., and Haworth, I. S. (1995) Biochemistry 34, 12803–12811
49. Mitas, M., Yu, A., Dill, J., Kamp, T. J., Chambers, E. J., and Haworth, I. S. (1995a) Nucleic Acids Res. 23, 1050–1059
50. McClellan, J. A., Palecek, E., and Lilley, D. M. J. (1986) Nucleic Acids Res. 14, 9291–9309
51. Kuryavyi, V. V., and Jovin, T. M. (1995) Nat. Genet. 9, 339–341
52. Gacy, A., Goellner, G., Juranic, N., Macura, S., and McMurray, C. (1995) Cell 81, 533–540
53. Mariappan, S. V., Catasti, P., Chen, X., Ratliff, R., Moyzis, R. K., Bradbury, E. M., and Gupta, G. (1996) Nucleic Acids Res. 24, 784–792
54. Mitas, M., Yu, A., Dill, J., and Haworth, I. S. (1995) Biochemistry 34, 12803–12811
55. Mitas, M., Yu, A., Dill, J., Kamp, T. J., Chambers, E. J., and Haworth, I. S. (1995a) Nucleic Acids Res. 23, 1050–1059
56. McClellan, J. A., Palecek, E., and Lilley, D. M. J. (1986) Nucleic Acids Res. 14, 9291–9309
57. Kuryavyi, V. V., and Jovin, T. M. (1995) Nat. Genet. 9, 339–341
58. Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., and Turner, D. H. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 9373–9377
59. Sarkar, M., Sigurdsson, S., Tomac, S., Sen, S., Rozners, E., Sjoberg, B.-M., Strumborg, R., and Gradlund, A. (1996) Biochemistry 35, 4678–4688
60. Yano-Yanagisawa, H., Li, Y., Wang, H., and Kohwi, Y. (1995) Nucleic Acids Res. 23, 2654–2660
61. Timchenko, L. T., Miller, J. W., Timchenko, N. A., De Vore, D. R., Datar, K. V., and E. M., and Gupta, G. (1996) Nucleic Acids Res. 24, 784–792
62. Mitas, M., Yu, A., Dill, J., and Haworth, I. S. (1995) Biochemistry 34, 12803–12811
63. Mitas, M., Yu, A., Dill, J., Kamp, T. J., Chambers, E. J., and Haworth, I. S. (1995a) Nucleic Acids Res. 23, 1050–1059
64. McClellan, J. A., Palecek, E., and Lilley, D. M. J. (1986) Nucleic Acids Res. 14, 9291–9309
65. Kuryavyi, V. V., and Jovin, T. M. (1995) Nat. Genet. 9, 339–341
66. Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., and Turner, D. H. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 9373–9377
67. Sarkar, M., Sigurdsson, S., Tomac, S., Sen, S., Rozners, E., Sjoberg, B.-M., Strumborg, R., and Gradlund, A. (1996) Biochemistry 35, 4678–4688
68. Yano-Yanagisawa, H., Li, Y., Wang, H., and Kohwi, Y. (1995) Nucleic Acids Res. 23, 2654–2660
69. Timchenko, L. T., Miller, J. W., Timchenko, N. A., De Vore, D. R., Datar, K. V., and E. M., and Gupta, G. (1996) Nucleic Acids Res. 24, 784–792
70. Mitas, M., Yu, A., Dill, J., and Haworth, I. S. (1995) Biochemistry 34, 12803–12811
71. Mitas, M., Yu, A., Dill, J., Kamp, T. J., Chambers, E. J., and Haworth, I. S. (1995a) Nucleic Acids Res. 23, 1050–1059
72. McClellan, J. A., Palecek, E., and Lilley, D. M. J. (1986) Nucleic Acids Res. 14, 9291–9309
73. Kuryavyi, V. V., and Jovin, T. M. (1995) Nat. Genet. 9, 339–341