Base Stacking and Even/Odd Behavior of Hairpin Loops in DNA Triplet Repeat Slippage and Expansion with DNA Polymerase*

Received for publication, December 23, 1999, and in revised form, March 29, 2000
Published, JBC Papers in Press, April 3, 2000, DOI 10.1074/jbc.M910272200

Michael J. Hartenstein, Myron F. Goodman‡, and John Petruska
From the Department of Biological Sciences, Hedco Molecular Biology Laboratories, University of Southern California, Los Angeles, California 90089-1340

Repetitions of CAG or CTG triplets in DNA can form intrantrad hairpin loops with combinations of normal and mismatched base pairs that easily rearrange. Such loops may promote primer-template slippage in DNA replication or repair to give triplet-repeat expansions like those associated with neurodegenerative diseases. Using self-priming sequences (e.g. (CAG)_{16}(CTG)_{4}), we resolve all hairpin loops formed and measure their slippage and expansion rates with DNA polymerase at 37 °C. Comparing CAG/CTG loop structures with GAC/GTC structures, having similar hydrogen bonding but different base stacking, we find that CAG, CTG, and GTC triplets predominantly form even-membered loops that slip in steps of two triplets, whereas GAC triplets favor odd-numbered loops. Slippage rates decline as hairpin stability increases, supporting the idea that slippage initiates more easily in less stable regions. Loop stabilities (in low salt) increase in the order GTC < CAG < GAC < CTG, while slippage rates decrease in the order GTC > CAG > GAC > CTG. Loops of GTC compared with CTG melt 9 °C lower and slip 6-fold faster. We interpret results in terms of base stacking, by relating melting temperature to standard enthalpy changes for doublet of base pairs and mispairs, considering enthalpy-entropy compensation.

Repetitive DNA sequences such as tandemly repeated triplets of bases are abundant and highly polymorphic in the human genome, probably because of strand slippage promoted by repetition in DNA replication, repair, or recombination (1–6). Occasionally, a repetitive region within a human gene expands sufficiently to cause inherited human disease (7). At least 12 neurological disorders are associated with triplet repeat expansions within human genes (5).

Eight such disorders arise by the expansion of CAG repetitions in gene regions encoding glutamine repeats in protein. In each case, involving a different gene (Huntington’s disease, spinobulbar muscular atrophy, dentatorubral-pallidolysian atrophy, and spino-cerebellar ataxias 1, 2, 3, 6, and 7) the number of CAG repeats in tandem is expanded from a normal range of 5–30 to a disease-causing range of 40–100 (5). The resultant polyglutamine expansion in protein may cause disease by forming insoluble nuclear protein aggregates (8–10), possibly with cross-linking by transglutaminase (11, 12), an enzyme abundant in the brain (13, 14), enabling glutamine reaction with lysine to form a peptide-like cross-link between polypeptides.

Another four disorders involve other triplet repetitions expanded in noncoding regions of genes. Myotonic dystrophy, for example, is associated with repeating CTG triplets expanded in an untranslated 3’-terminal gene region, from a normal range of 5–40 repeats to a disease-causing range of 50–300 (5). Expansions of this magnitude are also observed in CGG repeats and CCG repeats, found in 5’-untranslated regions of genes associated with fragile X syndromes A and E, respectively (5). Similarly large increases are observed for GAA repetitions in an intron of a gene associated with Friedreich’s ataxia (5). In each case, the noncoding region involved is transcribed into RNA but not translated into protein. The increased number of repetitions in RNA may cause disease by interfering with RNA transcription or processing within cell nuclei (15–18).

In DNA undergoing replication or repair, repeating triplets may enable one DNA strand to slip relative to the other so that the number of triplet repeats can be expanded by DNA polymerase (2, 4, 19). Also, because CNG triplets associated with disease can form intrantrad hairpin folds with secondary structure (20–22), it is of interest to determine how such folding affects slippage and expansion with polymerase. Recently, we developed a convenient in vitro assay, using self-priming repeat sequences, to measure rates of slippage and expansion in relation to hairpin structure (4).

In an earlier study of the major (CAG/CTG) class of disease-associated triplets (4), we examined DNA polymerase extension products of the self-priming sequence, (CTG)$_{16}$(CAG)$_{4}$, which forms a well defined series of hairpin loops. Using high resolution gel electrophoresis to analyze product lengths, we found that the CTG repeats form loops that are predominantly even-membered (i.e. have even rather than odd numbers of bases in the hairpin bend). The products of polymerase extension were observed to slip and expand in steps of two triplets at rates of ~1 step/min at 37 °C. In the present study, a similar analysis is made for the complementary case, (CAG)$_{16}$(CTG)$_{4}$, and “sister” cases obtained by replacing CAG/CTG with GAC/GTC. Since these cases exhibit the same hydrogen bonding with different base stacking, they reveal the influence of base stacking on the even-odd character of hairpin folds and slippage rates leading to repeat expansions with polymerase.

EXPERIMENTAL PROCEDURES

Materials

DNA Synthesis—Like the self-priming DNA 60-mer, 5’-(CTG)$_{16}$(CAG)$_{1}$-3’ previously studied (4), the complement (CAG)$_{16}$(CTG)$_{4}$ and sisters, (GTC)$_{16}$(GAC)$_{4}$ and (GAC)$_{16}$(GTC)$_{4}$, were synthesized using the standard phosphite triester method.
each synthesized using β-cyanoethyl phosphoramidites in an Applied Biosystems DNA/RNA synthesizer and purified by electrophoresis on denaturing 12% polyacrylamide gel. Each DNA 60-mer extracted from gel was dialyzed extensively against the same 0.02 M Na+ phosphate buffer, and then allowed to incubate under various (4) 6 M Na2HPO4, 5 M NaH2PO4, 1 M Na2EDTA, pH 7.6, for 24 h at 37°C. DNA Polymerase and Substrates—The KFexo- polymerase used to achieve rapid primer 3’ extension on template, an Escherichia coli DNA polymerase I Klenow fragment mutant (D355A,E357A) devoid of 5’-3’ exonuclease activity, was purified from overproducing strains (23). The dNTP substrates used in extension experiments were purchased from Amersham Pharmacia Biotech, along with dNTPs employed for sequence analysis of extended products.

Methods

Melting Curves—Thermal denaturation profiles were obtained for each DNA 60-mer at the same strand concentration (2 μM) in 0.02 M Na+ phosphate buffer (4), by measuring UV absorbance (A260nm) versus temperature, from 20 to 85 °C at 2 °C/min.

Radiolabeling—For use in extension reactions, DNA 60-mers were 5’-labeled with 32P-ATP and T4 polynucleotide kinase (U.S. Biochemical Corp./Amersham Pharmacia Biotech) in kinase buffer (50 mM Tris-Cl, pH 7.5, 10 mM MgCl2, and 10 mM 2-mercaptoethanol). As before (4), labeled samples at 100 mM sodium concentration, in polymerase reaction buffer (50 mM Tris-HCl, pH 7.5, 40 cm 9 mM borate, pH 8.3, 2 mM Na2EDTA). Gels were dried on paper and scanned by a Molecular Dynamics Storm 860 PhosphorImager. FragmeNT Analysis software (Molecular Dynamics) was used to integrate the intensity of each radioactive band in a gel lane, expressed as a percentage of total integrated band intensities in the lane.

Resolution of Hairpin Loops—In each of the four cases studied here, the initial 60-mer sequence (A) forms a series of self-primed hairpin loops, A, (n = 0, 1, 2, etc.), where n is the number of overhanging template triplets at the 5’-end, as illustrated (Fig. 1, a-d), with an asterisk indicating the 5’-end labeled with 32P. The loops are rapidly extended to their corresponding blunt-end products by adding DNA polymerase and sufficient amounts (0.4 μM each) of the three dNTPs needed to reach near maximum velocity of primer extension on template, with minimal background “pause” banding caused by misinsertion or terminal transferase activity (4). In 15 s of reaction time (t) with polymerase, the primer 3’-end of A, is extended by n primer triplets to form the corresponding blunt-end product, A, plus n triplets, measured as a discrete gel band intensity (I,0) by denaturing gel electrophoresis and PhosphorImager analysis. Each I,0 (n = 0, 1, 2, etc.) changes as reaction time t is increased, because each blunt-end product undergoes slippage and further extension with polymerase. By analyzing I,0 versus t and extrapolating back to t = 0, we evaluate I,0, indicating the approximate amount of loop A, present initially, before polymerase was added. The loops are classified as even-numbered or odd-numbered, depending on whether an even or odd number of bases is enclosed within the loop. The even-numbered loops (n = 0, 2, . . . 10) have the same kind of hairpin bend made with an even number of unpaired bases, minimally four as shown (Fig. 1, a-d). The odd-numbered loops (n = 1, 3, . . . 11) have another kind of bend made with an odd number of unpaired bases (minimally 3). The other loops shown (n = 12–15) have one or more primer triplets in the bend region where there is less opportunity for correct base pairing; these are much less stable loops, indicating how bend regions change as primer triplets enter the bend region in extreme cases of slippage.

Evaluation of Band Intensity I,0—The initial blunt-ended loop, A, which cannot be extended unless slippage occurs, is seen as band intensity I,0 remaining after 15 s of reaction. As A, undergoes slippage to an extendable form (A, A, etc.), I,0 declines in a small manner, suggesting a first order diffusion reaction, dI,0/dt = -k,0I,0, or in terms of measured differences (©), the

\[ \Delta I,0/\Delta t = -k,0I,0 \]  

(Eq. 1)

where k,0 is the slippage rate constant and I,0 is the average intensity value measured in a short time interval, ©t, of 15 or 30 s. For 4t between reaction times t and t, the intensity change is I,0 = I,(t) - I,(t), and the corresponding I,(t) on the right side of Equation 1 is the average value, I,(t) = 0.5(I,(t) + I,(t)).

In our experiments, I,0 decays to a low residual “background” intensity (I,0) that remains nearly constant, indicating that a small fraction (<1%) of initial 60-mers are nonreactive at the 3’-end. To take I,0 into account, we modify Equation 1 as follows.

\[ \Delta I,0/\Delta t = -k,0(I,0 - I,0) \]  

(Eq. 2)

The resultant integrated equation is then the following.

\[ \ln[I,0 - I,0] = \ln[I,0] - k,0t \]  

(Eq. 3)

where I,0 is the initial intensity corrected for background.

To evaluate I,0 and rate constant k,0, we plot \( \Delta I,0/\Delta t \) versus I,0 and fit a straight line by least squares, obtaining slope -k,0 and intercept k,0 according to Equation 2. Then, using Equation 3, we plot \( \ln[I,0 - I,0] \) versus t and linearly extrapolate t to 0 to obtain I,0, indicating the amount of A, present initially, as a percentage of total loops A, (n = 0–11).

Slippage Rate Constant Evaluation—The rate constant k,0 indicates the rate of A, slippage, resulting in extension with DNA polymerase.

Two loops are examined, k,0 contributing to band intensity I,0 (by A, A, slippage and extension by 1 triplet) and k,0 contributing to I,0 (by A, A, slippage and extension by 2 triplets). To determine the relative magnitude of these components, we start with the differential equations, dI,0/dt = k,0I,0 - k,1I,0 and dI,1/dt = k,0I,0 + k,1I,0 - k,2I,0, and replace differentials by measured differences, as in Equation 1 as follows.

\[ \Delta I,0/\Delta t = k,0(I,0) - k,1I,0 \]  

(Eq. 4)

\[ \Delta I,1/\Delta t = k,0(I,0) + k,1I,0 - k,2I,1 \]  

(Eq. 5)

Equation 4 is used to evaluate k,0 and k,1 by a linear least squares fit to a plot of \( \Delta I,0/\Delta t \) versus I,0. To evaluate k,0 and k,1 in a similar manner, we examine linear versions of Equation 5 obtained for three different approximations: (a) k,0 >> k,1, (b) k,0 ~ k,1, and (c) k,0 << k,1. Experimental plots of \( \Delta I,0/\Delta t \) against (a) (I,0) I,0 + I,0 and (c) I,0 are then compared to determine which plot gives the best linear least squares fit for k,0 evaluation. In all cases where even-numbered loops predominate, approximation a gives the best fit, indicating that extension following slippage is more frequent in steps of two triplets than in steps of one triplet. Only in the case of (GAC),4(dTTC), where odd-numbered loops are favored over even-numbered, do we find approximations b or c giving a better fit.

Having evaluated k,0, k,1, and k,2 as described above, we use a similar approach to obtain other k,0 values for n > 2. In each case, starting with the general form of Equation 5,

\[ \Delta I,0/\Delta t = k,0(I,0I,0) + k,1(I,0I,0) - k,2 \]  

(Eq. 6)

we apply approximations (a) k,0 ~ k,1 ~ k,2 ~ k,3 ~ k,4 ~ k,5 ~ k,6, and (c) k,2 ~ k,3 ~ k,4 ~ k,5 ~ k,6 ~ k,7 ~ k,8 ~ k,9 ~ k,10 ~ k,11, to determine which gives the best k,0 estimate by linear least squares fit. In all cases where approximation a applies to n = 2 (i.e. k,0 ~ k,1), in Equation 5, we find that the corresponding general approximation (k,0 ~ k,1 ~ k,2 ~ k,3 ~ k,4 ~ k,5 ~ k,6 ~ k,7 ~ k,8 ~ k,9 ~ k,10 ~ k,11) also applies to Equation 6 for evaluating other even-numbered k,0 values (n = 4, 6, etc.). By using the k,0 values measured with corresponding simplified (linear) versions of Equations 5 and 6, we extrapolate I,0 back from t = 15 s to t = 0, to estimate I,0, indicating the initial amount of A, present before extension with polymerase.

RESULTS

Each of the self-priming sequences examined here, like (CTG)12(dCAG)4 previously studied (4), forms a series of hairpin
loops, $A_n( n = 0, 1, 2, \ldots)$, with stable p/t$^1$ duplex enclosing a less stable t/t loop domain (Fig. 1, $a$–$d$). The subscript $n$ in $A_n$ indicates that there are $n$ overhanging template triplets on which DNA polymerase can rapidly add a complementary primer triplets to the primer 3' end. $a$, loops $A_n( n = 0, 1, 2, \ldots)$ with hydrogen-bonded base pairs for sequence 5'-CTG$_n$(CTG)$_4$-3'; $b$, corresponding loops for “sister” sequence of same base composition, (CTG)$_n$(GAC)$_4$; $c$ and $d$, corresponding loops for respective “complementary” cases, (CAG)$_n$(CTG)$_4$ and (GAC)$_n$(GTC)$_4$. The hairpin bends in even-numbered loops ($n = 0, 2, \ldots 14$) are made with an even number of unpaired bases, minimally four as shown; those in odd-numbered loops ($n = 1, 3, \ldots 15$) are made with an odd number of unpaired bases (at least three). The hydrogen-bonded, Watson-Crick base pairs are indicated by dots in a horizontal series; three dots between triplets in parentheses indicate stably paired primer-template triplets in the p/t duplex domain; two dots between triplets in parentheses indicate less stably paired template triplets in the t/t loop domain. The asterisk indicates the 5'-end of template labeled with $^{32}$P. Note that loops $A_n$ have $n$ template triplets available for extending primer 3' end on template with DNA polymerase (e.g. polymerase KFexo$^\ast$). In the presence of three dNTPs ($n = C, G$, and A or T), polymerase KFexo$^\ast$ rapidly extends $A_n$ by $n$ primer triplets to form product blunt-end hairpins. The products in each case, $A_{n}$ plus (CAG)$_n$, (a), $A_{n}$ plus (GAC)$_n$, (b), $A_{n}$ plus (CTG)$_n$, (c), or $A_{n}$ plus (GTC)$_n$, (d), are observed as a series of band intensities $I_n$ ($n = 0, 1, 2, \ldots$) by denaturing gel electrophoresis (Figs. 2b and 3b). The changes in $I_n$ with increasing reaction time are measured to determine rates at which successive blunt-end product hairpins undergo slippage resulting in further polymerase-catalyzed expansion. Extrapolation of $I_n$ to zero reaction time yields $I_n^0$, the initial amount of $A_n$ formed by each sequence.

In each case (Fig. 1, $a$–$d$), the even- and odd-numbered loops $A_n$ up to $n = 11$ have all four primer triplets correctly hydrogen-bonded and stacked in Watson-Crick bp (indicated by dots) with four antiparallel, template triplets located n triplets from the 5'-$^{32}$P-end (*). The p/t duplex in $A_n( n = 0, 1, 11)$ encloses a t/t loop domain containing $12 - n$ template triplets. The t/t domain has the hairpin bend made with an even or odd number of unpaired bases and also has mispaired bases held between correct bp formed between opposing (antiparallel) template triplets. Opposing triplets of type CTG or GTC form mispairs of type G/C and C/G bp (dots), as shown (Fig. 1, $a$ and $b$) for sequences (CTG)$_n$(GAC)$_4$ and (GTC)$_n$(GAC)$_4$, respectively. On the other hand, opposing triplets of type CAG or GAC form mispairs of type $A$, as shown (Fig. 1, $c$ and $d$) for (CAG)$_n$(CTG)$_4$ and (GAC)$_n$(GTC)$_4$, respectively.

Loops $A_{12}$ to $A_{15}$ (Fig. 1, $a$–$d$), having fewer than four of the primer triplets correctly bound to template triplets, are much less stable and are hardly apparent initially. These loops are included to show how bends change in extreme cases of slippage after more than 11 triplets are added by DNA polymerase in longer reaction times (Figs. 2b and 3b).

**Fig. 1.** Possible hairpin loops formed by self-priming sequences of CAG/CTG and GAC/GTC triplet repeats. $A$, self-priming DNA sequence having 16 repeats of a template triplet followed by four repeats of complementary primer triplet; $A_n$, hairpin fold of sequence $A$, having $n$ overhanging template triplets on which DNA polymerase can rapidly add a complementary primer triplets to the primer 3'-end. $a$, loops $A_n( n = 0, 1, 2, \ldots)$ with hydrogen-bonded base pairs for sequence 5'-CTG$_n$(CTG)$_4$-3'; $b$, corresponding loops for “sister” sequence of same base composition, (CTG)$_n$(GAC)$_4$; $c$ and $d$, corresponding loops for respective “complementary” cases, (CAG)$_n$(CTG)$_4$ and (GAC)$_n$(GTC)$_4$. The hairpin bends in even-numbered loops ($n = 0, 2, \ldots 14$) are made with an even number of unpaired bases, minimally four as shown; those in odd-numbered loops ($n = 1, 3, \ldots 15$) are made with an odd number of unpaired bases (at least three). The hydrogen-bonded, Watson-Crick base pairs are indicated by dots in a horizontal series; three dots between triplets in parentheses indicate stably paired primer-template triplets in the p/t duplex domain; two dots between triplets in parentheses indicate less stably paired template triplets in the t/t loop domain. The asterisk indicates the 5'-end of template labeled with $^{32}$P. Note that loops $A_n$ have $n$ template triplets available for extending primer 3'-end on template with DNA polymerase (e.g. polymerase KFexo$^\ast$). In the presence of three dNTPs ($n = C, G$, and A or T), polymerase KFexo$^\ast$ rapidly extends $A_n$ by $n$ primer triplets to form product blunt-end hairpins. The products in each case, $A_{n}$ plus (CAG)$_n$, (a), $A_{n}$ plus (GAC)$_n$, (b), $A_{n}$ plus (CTG)$_n$, (c), or $A_{n}$ plus (GTC)$_n$, (d), are observed as a series of band intensities $I_n$ ($n = 0, 1, 2, \ldots$) by denaturing gel electrophoresis (Figs. 2b and 3b). The changes in $I_n$ with increasing reaction time are measured to determine rates at which successive blunt-end product hairpins undergo slippage resulting in further polymerase-catalyzed expansion. Extrapolation of $I_n$ to zero reaction time yields $I_n^0$, the initial amount of $A_n$ formed by each sequence.

$^1$ The abbreviations used are: p/t, primer-template DNA duplex; t/t, loop domain of template triplet interactions; bp, base pair(s).
buffer (0.02 M Na\textsuperscript{+}), by plotting UV absorbance (A\textsubscript{260}) versus temperature (Figs. 2a and 3a). In each case, as found previously for (CTG\textsubscript{16})(CAG\textsubscript{4}) (4), two sigmoidal transitions are evident, the first with melting temperature (T\textsubscript{m}) below 60 °C and the second with T\textsubscript{m} near 80 °C.

The first transition indicates the melting of the t/t loop domain containing the base pairs and mispairs of template triplet interactions. As anticipated from previous work (22) and seen by comparing Figs. 2a and 3a, this domain is most stable for CTG triplets (T\textsubscript{m} = 56 °C) followed by GAC (54 °C), CAG (52 °C), and last GTC (47 °C). The second transition indicates the dissociation of stable p/t duplex, which melts at nearly the same temperature (T\textsubscript{m}’ = 79 °C) in all four cases.

If slippage in the stable p/t duplex is promoted by slippage in the less stable t/t domain, as we suggested previously (4), then (GTC\textsubscript{16})(GAC\textsubscript{4}), having the lowest T\textsubscript{m} for the first transition (47 °C, Fig. 3a), should also have the highest rate of slippage and expansion with polymerase. This is evidently the case, as seen by comparing gel band patterns (Figs. 2b and 3b) obtained by primer extension with DNA polymerase KFexo at 37 °C, for reaction times ranging from 0.5 to 16 min.

Polymerase Extension Products Resolved by Gel Electrophoresis and PhosphorImager Analysis—The addition of DNA polymerase KFexo\textsuperscript{-} and appropriate (0.4 μM) dNTP substrates results in rapid extension of loops A\textsubscript{n} to their blunt-end products, A\textsubscript{n} + n triplets, resolved as band intensities I\textsubscript{n} by electrophoresis and \textsuperscript{32}P PhosphorImager analysis (Figs. 2b and 3b). After stopping the reaction at various times and resolving product bands in gel lanes, we evaluate band intensities I\textsubscript{0}, I\textsubscript{1}, I\textsubscript{2}, etc. by integration in each lane, to measure the relative amounts of products A\textsubscript{0}, A\textsubscript{1} + 1 triplet, A\textsubscript{2} + 2 triplets, etc. as a function of time. The I\textsubscript{n} values obtained at times of 0.25 and 0.5 min, before products rearrange significantly by slippage, are used for extrapolation back to 0 time, to obtain I\textsubscript{0}, indicating the initial amounts of A\textsubscript{0} present when polymerase was added.

In the case of (CTG\textsubscript{16})(CAG\textsubscript{4}), whose band patterns are shown in Fig. 2b (left), we see that the loops are predominantly of the even-numbered type, as previously reported (4). The complementary case, (CAG\textsubscript{16})(CTG\textsubscript{4}), shown in Fig. 2b (right), also forms mainly even-numbered loops, as does (GTC\textsubscript{16})(GAC\textsubscript{4}), shown in Fig. 3b (left). After 0.5 min of reaction time, the most intense bands correspond to even numbers of triplets added (0, 2, . . . 10). As reaction time is increased, the band intensities gradually change as blunt-end products rearrange and are expanded further, mainly in steps of two triplets.

Extension by Slippage Compared with Melting Temperature—By examining band intensity changes with reaction time (Fig. 2b), we see that expansion by slippage is several times faster for (CAG\textsubscript{16})(CTG\textsubscript{4}) than for (CTG\textsubscript{16})(CAG\textsubscript{4}). A faster slippage rate is in keeping with the observation (Fig. 2a) that the first transition has a 4 °C lower T\textsubscript{m} value, 52 °C, compared with 56 °C for (CTG\textsubscript{16})(CAG\textsubscript{4}). The second transition in both cases is the same (T\textsubscript{m}’ = 78 °C).

The “sister” sequence (GTC\textsubscript{16})(GAC\textsubscript{4}), which has the same base composition as (CTG\textsubscript{16})(CAG\textsubscript{4}), has a much lower first transition (T\textsubscript{m} = 47 °C), with no significant change in the second transition (T\textsubscript{m}’ = 78 °C), as seen in Fig. 3a. The first T\textsubscript{m} is 9 °C below that for (CTG\textsubscript{16})(CAG\textsubscript{4}), and the rate of expansion by slippage is also much faster as found by comparing the gel band patterns in Figs. 3b (left) and 2b (left). Nevertheless, (GTC\textsubscript{16})(GAC\textsubscript{4}) still forms even-numbered loops preferentially and slips in steps of two triplets, as do (CTG\textsubscript{16})(CAG\textsubscript{4} and
that the first transition for (GAC)_{16}(GTC)_{4} has
a rate is consistent with the melting curve in Fig. 3
a, showing transitions and corresponding band patterns showing trip-
plets added with increasing time of DNA polymerase reaction. a,
melting curves obtained by plotting UV absorbance A_{m}, versus tempera-
ture, at the same (2 mM) strand concentration in low salt buffer, show-
ing that (GAC)_{16}(CTG)_{4} has lower first T_m than (CTG)_{16}(GAC)_{4} (52
versus 56 °C) but the same second T_m (79 °C). b, patterns of bands on
denaturing gel, obtained by reaction with DNA polymerase KFexo 
(79 °C). c, patterns of bands on denaturing gel, obtained by reaction with DNA polymerase KFexo 
(79 °C). d, patterns of bands on denaturing gel, obtained by reaction with DNA polymerase KFexo 
(79 °C). e, patterns of bands on denaturing gel, obtained by reaction with DNA polymerase KFexo 
(79 °C). f, patterns of bands on denaturing gel, obtained by reaction with DNA polymerase KFexo 
(79 °C).

In contrast, (GAC)_{16}(GTC)_{4}, which has the same base com-
position as (CAG)_{16}(CTG)_{4}, tends to form odd-numbered loops. Unlike the other three cases, the most intense bands in this
case (Fig. 3b, right) correspond to odd numbers of triplets added (1, 3, 5, etc.). The changes in band intensity with time indicate that expansion by slippage is much slower than ob-
served for (GTC)_{16}(CAG)_{4} in Fig. 3b (left), being comparable to that of (CAG)_{16}(CTG)_{4} in Fig. 2b (right). A slower slippage
rate is consistent with the melting curve in Fig. 3a, showing that the first transition for (GAC)_{16}(GTC)_{4} has T_m = 54 °C,
about 7 °C higher than observed for (GTC)_{16}(GAC)_{4} (Fig. 3a)
and 2 °C above that found for (CAG)_{16}(CTG)_{4} (Fig. 2a).

**Initial Blunt-end Hairpin Amount and Slippage Rate—** In all four cases, after 15 s of reaction, band intensity I_0, representing blunted hairpin A_0, shows simple exponential decay to a small residual “background” intensity I_0, which we evaluate by Equation 2 under “Experimental Procedures.” The rate con-
stant k_0 for this decay is obtained from a linear least squares fit to a plot of ln(I_n – I_0) versus reaction time, according to
Equation 3. Extrapolation to 0 time yields an I_0^0 value indicating
the initial amount of A_0 present at the time polymerase was added.

By evaluating two components of k_0, namely k_{01} for slippage
by one triplet (A_0 → A_1) and k_{02} for slippage by two triplets (A_0
→ A_2), as described by Equations 4 and 5, and we find that k_{02} is an
order of magnitude smaller than k_{02} in every case except (GAC)_{16}(GTC)_{4}. For the latter case, k_{01} appears equal to or greater than k_{02}, verifying that this case differs from the rest by forming odd-numbered loops equally well or better than even-numbered loops. A plot of k_0 against T_m for the first transition (Fig. 4) shows that (GAC)_{16}(GTC)_{4} has a k_0 value (open circle) somewhat greater than expected from the trend indicated by the other three cases (solid circles). The slippage rate for (GAC)_{16}(GTC)_{4} (open circle) is equivalent to that of (CAG)_{16}(CTG)_{4} (closed circle), although the first transition T_m is 2 °C higher, suggesting that the ability to slip by one triplet as well as two enhances the rate.

**Decline in Slippage Rate with Increasing Number of Triplets**

Added—While k_0 measures the slippage rate within the initial
blunt hairpin A_0, the k_n values obtained for n = 1, 2, etc.
measure the slippage rates within the extended blunt-end hair-
pins, A_n plus n triplets. The k_n values found in each case, by
applying linear versions of Equation 6 under “Experimental
Procedures,” are shown plotted against increasing n (Fig. 5). As
expected, extension increases hairpin stability, so k_n decreases
as n increases. The k_n values decline monotonically to similar
low values at n = 10 for the two complementary cases with
CTG and GAG triplets and to similarly low values at n = 11 for
the sister cases with GTC and GAC triplets. For n > 11, the k_n
value shows a more pronounced decline on a log scale (not shown), in keeping with our expectation that primer triplets

**Fig. 2. Comparison of (CAG)_{16}(CTG)_{4} and (CTG)_{16}(CAG)_{4} melting transitions and corresponding band patterns showing triple-
lets added with increasing time of DNA polymerase reaction. a, melting curves obtained by plotting UV absorbance A_m, versus tempera-
ture, at the same (2 mM) strand concentration in low salt buffer,
showing that (CAG)_{16}(CTG)_{4} has lower first T_m than (CTG)_{16}(GAC)_{4} (52
versus 56 °C) but the same second T_m (79 °C). b, patterns of bands on
denaturing gel, obtained by reaction with DNA polymerase KFexo 
(79 °C).**

**Fig. 3. Comparison of (GAC)_{16}(GTC)_{4} and (GTC)_{16}(GAC)_{4} melting transitions and corresponding band patterns showing triple-
lets added with increasing time of DNA polymerase reaction. a, melting curves obtained by plotting UV absorbance A_m, versus tempera-
ture, at the same (2 mM) strand concentration in low salt buffer,
showing that (GAC)_{16}(GTC)_{4} has higher first T_m than (GTC)_{16}(GAC)_{4} (54
versus 47 °C) but the same second T_m (78 °C). b, patterns of bands on
denaturing gel, obtained by reaction with DNA polymerase KFexo 
(79 °C).**
Base Stacking in Triplet Repeat Slippage and Expansion

estimated amounts of loops present initially—using equation 6 with observed \( k_0 \) values, we extrapolate each band intensity \( I_n \) to time 0 of reaction, in order to estimate \( I_n^0 \), the amount of loop \( A_n \) initially present. The \( I_n^0 \) values shown plotted versus \( n \) in a histogram (Fig. 6) were obtained using gel band patterns like those in Figs. 2b and 3b, but for shorter (15- rather than 30-s) intervals of reaction time, starting at 15 s (data not shown).

loop stability in relation to base stacking—If the stabilities of hairpin loops depended simply on the number of possible hydrogen-bonded bp shown (Fig. 1, a–d), then the predicted order of stability would be \( A_0 \), \( A_2 \), \( A_3 \), etc. In this case, initial band intensities should decrease in the order \( I_0^0 \), \( I_2^0 \), \( I_3^0 \), etc., and slippage by one triplet (\( A_0 \), \( A_1 \) to \( A_2 \)) should require less energy and occur more readily than slippage by two triplets (\( A_0 \), \( A_1 \) to \( A_2 \)). As seen in Fig. 6, none of the four cases examined conform to this simple model. Only in one case, (GAC)\(_6\)GTC, is \( I_0^0 \) greater than \( I_1^0 \), and in this case, \( I_0^0 \) is also greater than \( I_2^0 \). In the other three cases, \( I_0^0 \) is much less than \( I_1^0 \) and also less than \( I_2^0 \), indicating that odd-numbered loops are less stable than even-numbered ones. Since the four cases have the same hydrogen bonding in different nearest neighbor sequence contexts, we see that nearest neighbor base stacking has a strong influence on the even-odd character of hairpin loops and their stabilities.

We note that the even-numbered loops \( A_0 \), \( A_2 \), \( A_4 \), etc. each have a four-base bend held by two stacked bp (two adjacent dots) next to it (Fig. 1, a–d). The odd-numbered loops, \( A_1 \), \( A_3 \), \( A_5 \), etc., however, each have a three-base bend held by only a single bp (lone dot). If this lone bp is unstable because of unfavorable base stacking, there will be seven rather than three unpaired bases in the odd-numbered bend. In this case, with seven versus four unpaired bases in odd versus even bends, \( A_1 \) may become less stable than \( A_2 \), \( A_4 \), etc. so that slippage in steps of two triplets becomes favorable (\( A_0 \), \( A_1 \) to \( A_2 \), etc.). This kind of behavior, which we first observed (4) for (CTG)\(_{16}\)(CAG)\(_4\) (Fig. 1a), we now also observe for (GTC)\(_{16}\)(GAC)\(_4\) and (CAG)\(_{16}\)(CTG)\(_4\) (Fig. 1, b and c) but not for (GAC)\(_{16}\)(GTC)\(_4\) (Fig. 1d). To explain this, we examine the stacking of nearest neighbor doublets of base pairs and mispairs.

evaluation of nearest neighbor doublet stabilities—Both CAG repeats (CAGCGAG...) and CTG repeats (CTGCTG...) contain the nearest neighbor doublet, GC. This self-complementary doublet (5'-GC-3'), when hydrogen-bonded to its own kind in antiparallel, forms the strongly stacked bp doublet, 5'-GC/GC-5', which has the highest \( T_m \) value found for nearest neighbor doublets of DNA base pairs in low to physiological salt concentrations (24). The \( T_m \) value obtained for the GC/GC bp doublet is almost as high in 0.02 M salt (136 °C) as in 1 M salt, 139 °C (Table I).

The strong base stacking within the bp doublet GC/GC is largely responsible for the stability of both the p/t duplex domi nants of (CTG)\(_{16}\)(CAG)\(_4\) and (CAG)\(_{16}\)(CTG)\(_4\) and their t/t loop domains (Fig. 1, a and c). The p/t duplex in each case has the same kind of repeating unit, in which the strongly stacked GC/GC doublet is accompanied by two correct but less strongly stacked bp doublets, CT/GA and CA/GT, with \( T_m \) values of only 58 and 55 °C, respectively, in 0.02 M salt (Table I). The observed \( T_m \) for the p/t duplex in each case, 79 °C (Fig. 2a), is close to the

**Fig. 4.** Rate constant \( k_0 \) for \( A_n \) slippage and expansion in relation to first \( T_m \) in melting profile. Results shown as solid circles are for the three cases that predominantly form even-numbered loops, mainly \( A_n \) with \( k_0 \) measuring slippage by two triplets (\( A_n \), \( A_{n+1} \) to \( A_{n+2} \)). These results indicate that \( k_0 \) declines with increasing \( T_m \) of the first transition seen in Figs. 2a and 3a. The lone open circle is for the case of (GAC)\(_{16}\)(GTC)\(_4\), which favors odd-numbered loops, mainly \( A_1 \) (Fig. 1d). In this case, \( k_0 \) measures slippage by both one triplet (\( A_0 \), \( A_1 \), etc.) and two triplets (\( A_0 \), \( A_1 \)).

**Fig. 5.** Decline in slippage rate with increasing number of triplet repeats added by DNA polymerase. The rate constant \( k_0 \) for slippage and extension of blunt-end hairpin, \( A_n \) plus \( n \) triplets, was evaluated from changes in band intensity \( I_n \) with DNA polymerase reaction time, as described by Equation 6, using gel data shown (Figs. 2b and 3b) and additional data at 0.25-min intervals (not shown).

**Fig. 6.** Initial amounts of even- and odd-numbered loops \( A_n \) obtained for CAG/CTG and GAC/GTC triplet repeat sequences. The results shown are expressed as percentage of total intensity of bands found by extrapolation to zero reaction time with DNA polymerase, using rate constants (Fig. 5) and band intensities observed in short reactions of 0.25 min (not shown) and 0.5 min (Figs. 2b and 3b).
average value for these three doublets, (136 °C + 58 °C +
55 °C)/3 = 83 °C. The latter is the predicted \( T_m \) for an infinitely long duplex of CAG/CTG repeats in 0.02 M salt (24).

In the t/t loop domains of (CAG)\(_{16}\)(CTG)\(_4\), the GC/CG doublet is flanked by A/A mispairs (Fig. 1c). The poor stacking of A/A between C/G and G/C bp yields the weak mispaired doublets, CA/GA and GA/CA, which are equivalent. Considering the observed \( T_m \) of 52 °C (Fig. 2a) as approximately the average value for all of the doublets involved in t/t domain melting, we can evaluate \( T_m \) for CA/GA as follows.

\[
(\text{Tm}(\text{GC/CG}) + 2\text{Tm}(\text{CA/GA})) / 3 = 52 ^\circ \text{C}
\]

This yields \( T_m(\text{CA/GA}) = 10 ^\circ \text{C} \), given \( T_m(\text{GC/CG}) = 136 ^\circ \text{C} \) in 0.02 M salt (Table I).

We can also evaluate \( T_m \) for CA/GA by including the observed \( T_m \) for p/t duplex (79 °C, Fig. 2a), described as follows.

\[
(\text{Tm}(\text{GC/CG}) + \text{Tm}(\text{CA/CT}) + \text{Tm}(\text{CT/CA})) / 3 = 79 ^\circ \text{C}
\]

Subtracting Equation 7 from Equation 8 to eliminate \( T_m(\text{GC/CG}) \), we obtain the following.

\[
\text{Tm}(\text{CA/GA}) = \frac{1}{3}[3\text{Tm}(\text{CT/CA}) - \text{Tm}(\text{CA/CT})] = 79 ^\circ \text{C} - 52 ^\circ \text{C} = 81 ^\circ \text{C}
\]

This yields \( T_m(\text{CA/GA}) = 16 ^\circ \text{C} \), given \( T_m(\text{CT/CA}) = 58 ^\circ \text{C} \) and \( T_m(\text{CA/CT}) = 55 ^\circ \text{C} \) in 0.02 M Na\(^+\) (Table I).

The two methods of evaluation agree that CA/GA has a very low \( T_m \) value, 13 ± 3 °C in 0.02 M salt. From this value, we predict a corresponding low enthalpy change for CA/GA doublet melting, \( \Delta H^\circ = 1.0 \pm 0.3 \) kcal/mol (Table I), on the basis of the enthalpy-entropy compensation formula,

\[
\Delta H^\circ = a \Delta T_m ^\circ \text{(°C)}
\]

where \( T_m ^\circ \) (°C) is the doublet melting temperature in degrees Centigrade and \( a \) is an entropy constant (80 ± 10 cal/mol deg) shown to hold for normal bp doublets and also doublets with single mispairs, for Na\(^+\) salt concentrations up to 1.0 M (24).

The same approach using the observed melting temperatures for the two transitions in the other three cases (Figs. 2a and 3a) yields \( T_m(\text{CT/CT}) = 19 \pm 3 ^\circ \text{C} \), \( T_m(\text{GT/GT}) = 40 \pm 6 ^\circ \text{C} \), and \( T_m(\text{CA/GA}) = 50 \pm 6 ^\circ \text{C} \), with proportional \( \Delta H^\circ \) values calculated by Equation 10, as shown (Table I). Thus, in low salt (0.02 M), we find that GA/CA is the strongest, while CA/GA is the weakest, of the mispaired doublets formed in hairpin folds of CAG/CTG and GAC/GTC triplet repeats.

**DISCUSSION**

The major (CAG/CTG) class of DNA triplet repeats found expanded in neurodegenerative diseases has the ability to form stable hairpin loops on each strand, as shown in previous studies (21, 22, 25). The formation of such loops when repeating sequences are being replicated or repaired in vitro may promote primer-template slippage like that demonstrated here in vitro, enabling DNA polymerase to catalyze repeat expansions (4, 26, 27). In the present study, comparing self-primering loops of (CTG)\(_{16}\)(CTG)\(_4\), we find that even-numbered loops starting with A\(_0\) (Fig. 1, a and c) are strongly favored in both cases, promoting slippage and expansion in steps of two triplets, A\(_0\) \( \rightarrow \) A\(_2\) \( \rightarrow \) A\(_4\), etc. (Fig. 2b). The CAG loop domains (Fig. 1c), having a 4 °C lower melting temperature (52 versus 56 °C) (Fig. 2a), also show a 4-fold higher rate of slippage and expansion (Figs. 2b and 5). As seen in Fig. 5, slippage rate declines as hairpin structures become more stable with increasing numbers of triplets added, adding support to the idea that slippage initiates in less stable regions of loop structure (4).

By replacing CAG/CTG with GAC/GTC, we find that even-numbered loop domains are still favored by GTC repeats (Fig. 1b) but not by GAC repeats (Fig. 1d). Compared with even-numbered CTG loops (Fig. 1c), the GTC loops are considerably less stable and much more prone to slippage, melting at a 9 °C lower temperature (47 °C, Fig. 3a) and showing a 6-fold higher rate of slippage and expansion (Figs. 3b and 5). For GTC, CAG, and CTC triplets, all of which form even-numbered loops almost exclusively (Fig. 6), a trend of decreasing slippage rate with increasing loop domain melting temperature is indicated (Fig. 4, solid circles).

The GAC triplets, on the other hand, forming odd-numbered as well as even-numbered loops, appear to slip somewhat more rapidly than expected (Fig. 4, open circle). To explain why even-numbered loops are strongly favored by CAG, CTG, and

### Table I

Nearest neighbor doublet values of \( T_m \) and \( \Delta H^\circ \) for melting obtained for the CAG/CTG and GAC/GTC classes of DNA triplet repeats in 0.02 and 1 M Na\(^+\) salt solution

| Doublet       | \( T_m \) 0.02 M Na\(^+\) °C | \( T_m \) 1.0 M Na\(^+\) °C | \( \Delta H^\circ \) kcal/mol |
|---------------|-----------------------------|-----------------------------|-----------------------------|
| CAG/CTG class |                             |                             |                             |
| GC/CG         | 136\(^a\)                   | 139\(^a\)                   | 10.9\(^a\)                  |
| CT/GA         | 55\(^a\)                    | 105\(^a\)                   | 4.6\(^b\)                   |
| CA/GT         | 55\(^b\)                    | 107\(^b\)                   | 4.4\(^b\)                   |
| CT/GT         | 19 ± 3\(^c\)                | 38 ± 6\(^d\)                | 1.5 ± 0.3\(^b\)             |
| CA/GA         | 13 ± 3\(^c\)                | 26 ± 6\(^d\)                | 1.0 ± 0.3\(^b\)             |
| GAC/GTC class |                             |                             |                             |
| GC/GC         | 72\(^a\)                    | 149\(^a\)                   | 5.8\(^b\)                   |
| GT/CA         | 98\(^a\)                    | 108\(^b\)                   | 7.5\(^b\)                   |
| GA/CT         | 86\(^c\)                    | 87\(^b\)                    | 6.9\(^b\)                   |
| GA/GA         | 50 ± 6\(^c\)                | 50 ± 6\(^d\)                | 4.0 ± 0.6\(^b\)             |
| GT/CT         | 40 ± 6\(^d\)                | 40 ± 6\(^d\)                | 3.2 ± 0.6\(^b\)             |

\(^a\) From Ref. 24.

\(^b\) Calculated by the formula, \( \Delta H^\circ \) (kcal/mol) = 0.08 \( T_m \) (°C), based on enthalpy-entropy compensation described by Equation 10.

\(^c\) Calculated from \( T_m \) observed for the two transitions in each case (Fig. 2, a and b, or Fig. 3, a and b), as described in terms of doublets by Equations 7 and 9.

\(^d\) Calculated from doublet \( T_m \) (°C) value in 0.02 M salt, by multiplying this value by 2 or 1, in case of doublet beginning with C or G, respectively; see Discussion.

\(^e\) From Ref. 28.
Starting with physiological temperature. Accordingly, odd-numbered loops in odd-numbered loops (Fig. 1, that the lone C/G bp holding the three-base CTG or CAG bend $T_m$79 °C) in all four duplex melts at a similar high temperature ($T_m$ average (GAC)$_{16}$(GTC)$_4$. GC/CG, CA/GT, and CT/GA; for GAC/GTC repeats, they are three bp doublets. For CAG/CTG repeats, these doublets are so poorly stacked ($T_m$) from opening up to seven unpaired bases. This may be the reason why GTC repeats still prefer GTC triplet repeats, but not by GAC repeats, we examine the base stacking properties of nearest neighbor doublets of base pairs and mispairs formed in each case.

**Stacking of Nearest Neighbor Doublets**—In the p/t duplex formed by complementary primer and template triplet repeats, the stacking of Watson-Crick bp can be described in terms of three bp doublets. For CAG/CTG repeats, these doublet triplets are GC/GC, CA/GT, and CT/GA; for GAC/GTC repeats, they are CG/GC, GA/CT, and GT/CA. As seen in Table I, GC/GC is much stronger than CG/GC in 0.02 M salt, but CA/GT and CT/GA are considerably weaker than GA/CT and GT/CA. The resultant average doublet strength is nearly the same in each case. The average doublet value of $T_m$ being similar explains why the p/t duplex melts at a similar high temperature (~79 °C) in all four cases examined (Figs. 2a and 3a). With p/t duplex being more or less equally stable in all cases, we see that p/t slippage is probably controlled by the less stable t/t domain, whose $T_m$ shows an inverse relationship to slippage rate (Fig. 4).

The hairpin loop t/t domains have lower melting temperatures because they have mispaired bases ($\Delta A$ or $\Delta T$) in place of correct (A/T or T/A) base pairs. The stacking of a correct C/G bp followed by mispair of type $\Delta A$ or $\Delta T$ corresponds to doublet CA/GA or CT/GT, respectively, while the stacking of G/C followed by $\Delta A$ or $\Delta T$ corresponds to GA/CA or GT/CT, respectively. As shown (Table I), CA/GA stacking is very weak in 0.02 M salt ($T_m = 13 °C$), whereas GA/CA stacking is comparatively very strong ($T_m = 50 °C$). This large difference in base stacking helps explain why GAC repeats form much more stable odd-numbered loops than CAG repeats.

**Even-Odd Character of Hairpin Bends Explained by Mispair Stacking Energies**—In loop domains of CAG or CTG repeats, each strongly stacked GC/GC doublet ($T_m = 136 °C$) is accompanied by two weakly stacked doublets of type CA/GA ($T_m = 13 \pm 3 °C$) or GT/CT ($T_m = 19 \pm 3 °C$). Because the latter doublets are so poorly stacked ($T_m$ well below 37 °C), we see that the lone C/G bp holding the three-base CTG or CAG bend in odd-numbered loops (Fig. 1, a and c) is probably unstable at physiological temperature. Accordingly, odd-numbered loops starting with A$_1$ should have seven unpaired bases rather than only three (Fig. 1, a and c). Thus, with more unpaired bases in the bend, A$_1$ becomes less stable than A$_0$, both of which have similar four-base bends firmly held by the strongly stacked GC/GC doublet (Fig. 1, a and c). The enlarged odd-numbered bend with poor stacking is probably the reason why odd-numbered loops are so disfavored for (CTG)$_6$(GAG)$_4$ and (CAG)$_6$(GTC)$_4$, so that slippage occurs primarily in even numbered steps of two triplets (Fig. 2b).

In GAC and GTC loop domains, on the other hand, the corresponding mispaired doublets, GA/CA and GT/CT, have much better stacking as indicated by their high $T_m$ values of 50 ± 6 and 40 ± 6 °C, respectively (Table I). The $T_m$ of GA/CA appears high enough to maintain a stable 3-base bend in odd-numbered loops as shown (Fig. 1d). However, GT/CT being only marginally stable at 37 °C ($T_m = 40 °C$), is less likely to keep such a bend (Fig. 1b) from opening up to seven unpaired bases. This may be the reason why GTC repeats still prefer even-numbered loops, whereas GAC repeats favor odd-numbered loops (Figs. 3b and 6).

Remarkably, GA/CA has $T_m$ and $\Delta H^\circ$ values almost as high as the normal doublets, CA/GT and CT/GA, in 0.02 M salt (Table I). This means that the $\Delta A$ mispair, although not hydrogen-bonded, is strongly stacked in the doublets GA/CA and $\Delta C/\Delta G$, which are equivalent. With $T_m = 50 \pm 6 °C$, well above 37 °C, the stacking appears strong enough to hold the lone G/C bp needed to form a stable 3-base bend, as shown (Fig. 1d) for (GAC)$_6$(GTC)$_4$.

**Dependence on Salt Concentration**—It has been known for some time that normal doublets CA/GT, CT/GA, and GC/GC show similar percentage increases in $T_m$ (°C) as salt concentration is raised (24). For Na$^+$ raised from 0.02 to 1 M, the increase is about 100% in all three cases (Table I), so we expect a 100% increase for CA/GA and CT/GT as well. On the other hand, no increase for GA/CA and GT/CT is expected, since GT/CA, GA/CT, and GC/GC show almost no sensitivity to salt concentration. The resultant $\Delta H^\circ$ values calculated by Equa
tion 10 from $T_m$ in 1 M Na$^+$ (Table I, last column) are consistent with recently published values (28), shown in parentheses.

**Conclusion**—Our analysis of melting profiles and polymerase extensions of self-priming sequences of DNA triplet repeats reveals several novel features of the CAG/CTG triplet class implicated in disease. (a) The intrastrand loops formed by this class of triplet repeats are stabilized largely by the strongly stacked GC/GC bp doublet. (b) The accompanying doublets, CA/GA and CT/GT, containing $\Delta A$ and $\Delta T$ mispairs, respectively, are very weakly stacked in low salt solution. (c) The weak stacking of CA/GA and CT/GT destabilizes the C/G bp that holds three-base bends of odd-numbered loops, making the four-base even-numbered bends much more favorable for both CAG repeats and CTG repeats. (d) The weaker stacking of CA/GA than CT/GT makes CAG loop domains less stable and more slippery than their CTG counterparts, resulting in a 4 °C lower melting temperature and a 4-fold higher rate of repeat expansion by slippage in DNA polymerase reactions.

We also find that GAC/GTC triplet repeats, not yet implicated in disease, have very different base-stacking properties, although forming similar hydrogen-bonded base pairs. (a) Their intrastrand loops are mainly stabilized by the correct doublet, CG/GC, which has considerably weaker base stacking than GC/GC in low salt. (b) The weaker stacking of GC/GC is compensated by stronger stacking in the accompanying doublets with mispairs, GA/CA and GT/CT. (c) The stacking of CA/GA in low salt is surprisingly high, comparable with that of normal CA/GT and CT/GA doublets in correct duplexes of CAG and CTG repeats. (d) The strong GA/CA stacking preferentially stabilizes odd-numbered GAC loops and makes them equivalent to even-numbered CAG loops in stability, although CG/GC is less stable than GC/GC. (e) The GT/CT doublet, being weaker than GA/CA, is unable to stabilize odd-numbered bends sufficiently, so that GTC repeats still primarily form even-numbered loops, which melt 9 °C lower than CTG loops and show a 6-fold higher rate of expansion by slippage.

Several models have been proposed to indicate how hairpin loop formation may contribute to the expansion of trinucleotide repeats with DNA polymerase (4, 26, 27, 29). The data presented here should be helpful in testing such models. Our data reveal that nearest neighbor base stacking has a strong influence on loop stability and slippage rate, leading to repeat expansions. These data were obtained in 0.02 M salt in order to observe the melting of both loop and duplex domains in our self-priming sequences of triplet repeats. Since salt can strongly affect the stacking of base pairs and mispairs, additional data at salt concentrations closer to 0.15 M are needed to make more reliable predictions of DNA triplet-repeat folding and slippage under physiological conditions.

**REFERENCES**

1. Levinson, G., and Gutman, G. A. (1987) Mol. Biol. Evol. 4, 203–221
2. Schlotterer, C., and Tautz, D. (1992) Nucleic Acids Res. 20, 211–215
3. Wells, R. D. (1996) J. Biol. Chem. 271, 23468–23479
4. Petruska, J., Hartenstein, M. J., and Goodman, M. F. (1998) J. Biol. Chem. 273, 5204–5210
5. Wells, R. D., and Warren, S. T. (eds) (1998) Genetic Instabilities and Hereditary Neurological Diseases, Academic Press, Inc., San Diego, CA
6. Jakupcik, J. P., and Wells, R. D. (1999) J. Biol. Chem. 274, 23468–23479
7. Sutherland, G. R., and Richards, R. J. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 3636–3641
8. Davies, S. W., Turmaine, M., Cozenz, B. A., DiFiglia, M., Sharp, A. H., Ross, C. A., Scherzinger, E., Wanker, E. E., Magiarini, L., and Bates, G. P. (1997) Cell 90, 537–548
9. Skinner, P. J., Koshy, B. T., Cummings, C. J., Klement, I. A., Helin, K., Servadio, A., Zoghbi, H. Y., and Orr, H. T. (1997) Nature 389, 971–974
10. Matilla, A., Koshy, B. T., Cummings, C. J., Isobe, T., Orr, H. T., and Zoghbi, H. Y. (1997) Nature 389, 974–978
11. Kahlem, P., Terre, C., Green, H., and Dijan, P. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 14580–14585
12. Cooper, A. J. L., Sheu, K.-F. R., Burke, J. R., Ondoera, O., Strittmatter, W. J., Roses, A. D., and Blass, J. P. (1997) J. Neurochem. 69, 431–434
13. Gilad, G. M., and Varun, L. E. (1985) J. Neurochem. 45, 1522–1526
14. Ohashi, H., Itah, Y., Eirckhichler, P. J., and Takeuchi, Y. (1995) J. Biochem. (Tokyo) 118, 1271–1278
15. Krahe, R., Ashizawa, T., Abbruzzese, C., Reeder, E., Carango, P., Giacanelli, M., Funanage, V. L., and Siciliano, M. J. (1995) Genomics 28, 1–14
16. Davis, B. M., McCurrach, M. E., Taneia, K. L., Singer, R. H., and Houseman, D. E. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 7388–7393
17. Verkerk, A. M. J. H., de Graaf, E., De Boulle, K., Eichler, E. E., Konaacki, D. S., Rey妮ers, E., Manca, A., Poustka, A., Willems, P. J., Nelson, D. L., and Oostra, B. A. (1993) Hum. Mol. Genet. 2, 399–404
18. Duval-Valentin, G., Thuong, N. T., and Helene, C. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 504–508
19. Behn-Kraupa, A., and Doerrler, W. (1994) Hum. Mutat. 3, 19–24
20. Gacy, A. M., Goellner, G., Jurianc, N., Macura, S., and McMurray, C. T. (1995) Cell 81, 533–540
21. Mitas, M., Yu, A., Dill, J., Kamp, T. J., Chambers, E. J., and Haworth, I. S. (1995) Nucleic Acids Res. 23, 1059–1659
22. Petruska, J., Arneheim, N., and Goodson, M. F. (1996) Nucleic Acids Res. 24, 1992–1996
23. Derbyshire, V., Freemont, P. S., Sanderson, M. R., Reese, L., Friedman, J. M., Jyce, C. M., and Steitz, T. A. (1988) Science 240, 199–201
24. Petruska, J., and Goodman, M. F. (1995) J. Biol. Chem. 270, 746–750
25. Mitas, M. (1997) Nucleic Acids Res. 25, 2245–2253
26. Pearson, C. E., Eichler, E. E., Lorenzetti, D., Asharya, S., Kramer, P. R., Kramer, S. F., Nelson, D. L., Zoghbi, H. Y., and Sinden, R. R. (1996) Am. J. Hum. Genet. 59, 244
27. Gellibolian, R., Bacolla, A., and Wells, R. D. (1997) J. Biol. Chem. 272, 16793–16797
28. Peyret, N., Seneviratne, P. A., Allawi, H. T., and SantaLucia, J. J. (1999) Biochemistry 38, 3468–3477
29. Kunkel, T. A., Resnick, M. A., and Gordenin, D. A. (1997) Cell 88, 155–158
Base Stacking and Even/Odd Behavior of Hairpin Loops in DNA Triplet Repeat Slippage and Expansion with DNA Polymerase
Michael J. Hartenstine, Myron F. Goodman and John Petruska

J. Biol. Chem. 2000, 275:18382-18390. doi: 10.1074/jbc.275.24.18382 originally published online June 9, 2000

Access the most updated version of this article at doi: 10.1074/jbc.275.24.18382

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 28 references, 10 of which can be accessed free at http://www.jbc.org/content/275/24/18382.full.html#ref-list-1