An Intramolecular Triplex in the Human γ-Globin 5′-Flanking Region Is Altered by Point Mutations Associated with Hereditary Persistence of Fetal Hemoglobin*

(Received for publication, May 18, 1995, and in revised form, August 1, 1995)

Albino Bacolla, Michael J. Ulricht§, Jacquelyn E. Larson, Timothy J. Ley‡, and Robert D. Wells†

From the Institute of Biosciences and Technology, Texas A&M University, Texas Medical Center, Houston, Texas 77030-3303 and the Division of Bone Marrow Transplantation and Stem Cell Biology, Departments of Medicine and Genetics, Washington University Medical Center, St. Louis, Missouri 63110

The properties of an intramolecular triplex formed in vitro at the 5′-flanking region of the human γ-globin genes were studied by chemical and physical probes. Chemical modifications performed with osmium tetroxide, chloroacetaldehyde, and diethyl pyrocarbonate revealed the presence of non-paired nucleotides on the "coding strand" at positions −209 through −217. These reactivities were induced by negative supercoiling, low pH, and magnesium ions. Downstream point mutations associated with hereditary persistence of fetal hemoglobin (HPFH) altered the extent of the modifications and some of the patterns. Specifically, C →G and C →T significantly decreased the reactivities, whereas the patterns were increased and altered in the T →C. C →T and C →G caused local decreases in reactivity. Modifications at the upstream flanking duplex were modulated by the composition of the vector sequence. In summary, our data indicates the formation of an intramolecular triplex between nucleotides −209 to −217 of the "non-coding strand" and the downstream sequence containing the HPFH mutations. All of the HPFH point mutations altered the structure. More than one sequence alignment is possible for each of the triplexes. In addition, a consequence of some of the point mutations may be to facilitate slippage of the third strand relative to the Watson-Crick duplex.

Human hemoglobin is synthesized from two sets of clustered genes designated the α-cluster (ζ-chain, ω-chain, δ-chain, γ-chain) located on chromosome 16 and the β-cluster (ε-chain, δ-chain, γ-chain, β-chain) on chromosome 11. Expression of the genes follows a developmentally, as well as tissue-specific, regulated program that allows ζ and ε to be transcribed during early embryonic life in placental yolk sac-derived red cells. At subsequent stages of development, globin expression shifts to the α and γ genes in the red cells of hepatic origin. At the time of birth, β-chains are predominantly expressed, and erythropoiesis shifts to the bone marrow (1, 2). This pattern of expression can be altered by mutations affecting any of the transcribed genes (1–3), but a condition that has attracted particular attention is the hereditary persistence of fetal hemoglobin (HPFH) caused by point mutations at the 5′-flanking region of either one of the γ-genes. Single nucleotide changes cause the affected allele to permanently express high levels of γ-chains in adult red cells.

Although the molecular mechanisms responsible for this protracted expression are not known, alterations in the recognition of cis-acting elements by regulatory factors and/or in the supramolecular assembly of chromatin have been invoked (4–7 and references therein). Ulrich et al. suggested that selected mutations in the −200 region destabilize a non-B DNA structure formed during the course of the γ-to-β switching (4). This hypothesis was supported by the finding that some of the mutations abolished an S1 nuclease-hypersensitive site (S1-HSS) located just upstream, which suggested the formation of an intramolecular triplex (I.T.). This occurrence is interesting for several reasons. First, it is becoming increasingly evident that various types of sequence motifs, including simple repeating defined-order-sequences, have the potential of adopting non-B conformations and that these structural transitions occur in biologic systems (8–11). Second, the isolation of DNA-binding proteins specific for pyrimidine-, or purine-rich, motifs are intriguing (12–16 and references cited therein) since these sequences are known to undergo conformational polymorphisms. Finally, diseases inherited by non-Mendelian genetic mechanisms have been recently associated with the expansion of tandemly repeated DNA motifs (17–23). The mechanism(s) through which such aberrant expansions are carried out is unknown, but the propensity of defined-order-sequences to adopt multiple conformations suggests that these properties may be directly involved in the process (24–26).

Intramolecular triplexes have been well characterized in recent years and are known to form at mirror repeat oligopyrimidine-oligopyrimidine tracts under the influence of negative supercoiling and low pH. Sequences of this type, but usually with imperfect mirror repeat symmetries, are often found at regulatory regions in eukaryotic genomes and have been proposed to participate in the regulation of physiological processes such as transcription and recombination. Additionally, some of these sequences have been shown to adopt I.T. structures under appropriate conditions (27–33).

Here we extend the previous studies on the human γ-globin 5′-flanking sequence, which identified the formation of an I.T. based upon S1 nuclease and oligomer binding assays. By ap-

*This work was supported by National Institutes of Health Grants GM52982 (to R. D. W.) and DK49786 (to T. J. L.), National Science Foundation Grant DMB-9103942, and the Robert A. Welch Foundation (to R. D. W.). 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.

§Present address: Dept. of Biology, 2345 Campus Box, Elon College, Elon College, NC 27244.

‡To whom correspondence should be addressed. Tel.: 713-677-7651; Fax: 713-677-7689; E-mail: rwells@bt.tamu.edu.

1The abbreviations used are: HPFH, hereditary persistence of fetal hemoglobin; S1-HSS, S1 nuclease-hypersensitive site; I.T., intramolecular triplex; bp, base pair(s); CAA, chloroacetaldehyde; DEPC, diethyl pyrocarbonate; nt, nucleotide(s).
Triplex DNA and HPFH Mutations in Human \( \gamma \)-Globin Genes

By investigating chemical probe analyses and two-dimensional agarose gel electrophoresis, we now identify the bases associated with the Hoogsteen-pairing third strand and describe the structural alterations introduced by the HPFH point mutations.

**EXPERIMENTAL PROCEDURES**

Plasmid DNA—Plasmid py-200 contains the sequence of the 5'-flanking region of human \( \gamma \)-globin genes from bp –228 to –189 inserted at the HindIII site of pUC9. Plasmids –202G, –202T, –198, –196T, and –195G harbor the same \( \gamma \)-globin fragment with the sequence diverging at the indicated position to reproduce the HPFH point mutations. The mutated inserts were doned at the HindIII site in –202G and –202T, and at Smal in –198C, –196T, and –195G. The cloning was described previously (4). Plasmid py-200 is as py-200, but the \( \gamma \)-globin insert was doned at the Smal site of pUC9. This was performed using synthetic oligodeoxynucleotides and standard procedures (34). Plasmid DNA was amplified in the Escherichia coli strain DH5\( \alpha \) and purified twice through CsCl banding. Na\( ^+ \) ions were exchanged for Cs\( ^+ \) ions by supercoiling, protonation at specific residues, and Mg\( ^2+ \) ions in a concentration-

Preparation of Topoisomers at Defined Superhelical Density—12 \( \mu \)g of plasmid DNA was incubated with various concentrations of ethidium bromide from 0.1 to 4.0 \( \mu \)g/ml for 90 min at 37\( ^\circ \)C in 300 \( \mu \)l of 1% bovine serum albumin, pH 7.6, 50 mM KCl, 50 mM Tris acetate, pH 4.5, 0.1 mM EDTA, 2 mM MgCl\(_2\), 150 mM NaCl.

Data were processed using the SigmaPlot program (Jandel Scientific, San Rafael, CA).

**RESULTS**

Fig. 1 shows the sequences analyzed in this study, the location of the S1-HSS, and a schematic model for the I-T. based on previous data (4). We have now extended these data using chemical probe analyses (OsO\(_4\), CAA, and DEPC) in order to detect perturbations at the bp level (27–29, 37–39).

**Fig. 1. Sequence of the human \( \gamma \)-globin 5'-flanking sequence and model for I-T. formation.** A, the sequences of human \( \gamma \)-globin 5'-flanking regions from bp –228 to –189 were doned in pUC9 as described: single point mutations leading to HPFH are shown in boldface and indicate the bp change on the top (coding) strand as well as the name of the plasmids carrying the respective mutations. The S1 nucleus hypersensitive site is indicated by S1-HSS. A stretch of bp containing two adjacent purine-rich motifs centered on S1-HSS and HPFH is underlined. B, schematic representation of the I-T. structure proposed to be adopted by the –200 region. The structure forms under conditions of negative supercoiling and low pH and is stabilized by Hoogsteen-type hydrogen bonds between the two adjacent purine-rich motifs, represented by a thicker line. The structure leaves unpaired pyrimidine residues on the top strand which become a substrate for S1 nucleus (S1-HSS). The model shows how residues affected by the HPFH point mutations may destabilize the I-T. structure. The 5' terminus on each of the DNA strands is indicated by a filled circle.

Chemical Modifications on the Coding Strand—We examined the reactivities on the coding strand of the wild type sequence (py-200) by using OsO\(_4\), which reacts with the C5-C6 double bond of unpaired thymines (39). Lanes 1 and 2 of Fig. 2A show a typical result. While relaxed (R) py-200, lane 1, did not show any strong reaction, introduction of supercoiling (S, lane 2, \( \tau = 0.137 \)) induced strong cleavages at thymines –209, –210, –212, –215, and –217. The percentage of modifications ranged from 7 to 14 (Fig. 2B). Weaker bands were observed corresponding to thymines –227 and –228 (less than 1%). Modifications were pH dependent, being observed only at pH 4.5. At pH 5.0 or higher (up to 7.5 was tested), no appreciable signals were detected. In addition, reactivities were increased by the addition of up to 10 mM Mg\( ^{2+} \) ions in a concentration-dependent manner (not shown).

CAA forms \( \epsilon \)-etheno adducts with unpaired cytosines, adenes, and, to a lesser extent, guanines (37). Supercoiling-induced cleavages were found at C\( ^{211} \) to C\( ^{213} \), C\( ^{213} \) to C\( ^{214} \), and A\( ^{216} \) (Fig. 2A, lanes 3 and 4). However, these bands accounted only for 0.5–1.5% of the total radioactivity (Fig. 2B). Acid treatment of the samples, before piperidine cleavage, did not improve the signal-to-noise ratio. The sites of modification complemented those detected by OsO\(_4\) and suggested that the 5' end of the single-stranded region was 3' of A\( ^{219} \).

The character of the flanking nt was determined by DEPC, a probe specific for unpaired adenes and guanines (37). As shown in lanes 5 and 6 of Fig. 2A, supercoiling-induced reactivities extended from A\( ^{216} \) to A\( ^{226} \), the strongest band corresponding to A\( ^{216} \) (0.94 ± 0.05% of the total radioactivity). Taken together, these data define two sets of accessible nt: a major site that extends from T\( ^{209} \) to T\( ^{211} \) that we interpret as single-stranded and a minor one, from nt –218 to –228, that we consider being weakly bonded (filled and open circles in Figs. 2A and 4C). These structural transitions are influenced by supercoiling, protonation at specific residues, and Mg\( ^{2+} \) ions.

Reactions with OsO\(_4\), CAA, and DEPC were conducted and...
visualized on the coding strand of plasmids containing point mutations associated with HPFH (Fig. 1A). Relaxed and supercoiled DNA (−σ = 0.137) were treated under the same conditions as pγ−200. The results demonstrated that the modifications occurred at the same residues as the wild type (not shown) but that there were quantitative differences, which are summarized in Fig. 3. Here the columns represent the percentages of the signals from OsO₄ and CAA normalized to the wild type sequence. The major changes were caused by the G⁻²⁰² and T⁻²⁰² mutations, which produced a general reduction in modification; the consequences of the G⁻²⁰² mutation were more severe than those of T⁻²⁰². −198C exhibited a 2-fold increase in modification at T⁻²⁰⁹ and T⁻²¹⁰, whereas −196T and −195G displayed a modest reduction in modification at the middle residues. The normalization at A⁻²¹⁶ with DEPC gave the following values: 0.00 for T⁻²⁰², 0.38 for −202T, 1.19 for −198C, 0.91 for −196T, and 1.32 for −195G.

Chemical Modifications on the Non-coding Strand—Plasmids carrying the wild type sequence or the HPFH point mutations were reacted with OsO₄, CAA, or DEPC, and the sites of cleavage were monitored on the non-coding strand. In response to supercoiling (−σ = 0.137), CAA modified C⁻²⁰⁵ and C⁻²⁰⁶ in the wild type sequence (Fig. 4A, lanes 1 and 2). A second set of reactive nt was present from A⁻²¹⁷ to C⁻²²⁴. Since modifications at these latter positions were also detected on the coding strand (Fig. 4C), they support the interpretation of a locally distorted duplex DNA. However, in contrast to the coding strand, no cleavages were seen between purines −209 and −215. This strand-selective pattern of reactivity, together with the modifications at C⁻²⁰⁵ and C⁻²⁰⁶, is indicative of an I.T. Accordingly, the protected purines constitute the third strand of the structure, whereas C⁻²⁰⁵ and C⁻²⁰⁶ would be located in the loop (Figs. 1B and 7) (40, 41).

The data with DEPC showed the accessibility of A⁻²¹⁷, G⁻²²⁰, and G⁻²²³ (not shown). OsO₄ modified the thymines spanning positions −216 to −226 to a moderate extent (Fig. 4B), confirming the chemical accessibility of the I.T. flanking
patterns of modifications with CAA and DEPC qualitatively identical to that of the wild type. These mutants also showed a considerable reduction in reactivity at T \textsuperscript{216} (Fig. 4, A and B). These results, together with those on the coding strand (Fig. 3), suggest that these two mutations perturb the overall geometry of the triplex structure, but do not change the sequence alignment. In contrast, the CAA-induced cleavages at C \textsuperscript{195} and C \textsuperscript{205} were not observed in the –198C mutant. Instead, signals were detected at C \textsuperscript{204}, G \textsuperscript{203} and G \textsuperscript{202} (denoted by asterisks in Fig. 4, A and C), indicating an alteration in the loop structure. This may be accomplished by a slippage of the purine-rich third strand relative to the Watson-Crick duplex or by multiple adjustments in the interactions among nt that retain their sequence alignment. In either case, it is clear that this mutation has profound consequences on the triplex structure.

The increase in OsO\textsubscript{4} modification at T \textsuperscript{222} and T \textsuperscript{226} in mutants –198C, –196T, and –195G relative to py-200 (Fig. 4B) was quite significant (3–4-fold). However, this behavior is not due to differences in the I.T. structures. A detailed analysis of these reactivities will be given in the last section under “Results.”

Superciling-dependent Titration of OsO\textsubscript{4} Modification—Since the previous experiments were conducted at very high superhelical densities, they left open the question on whether mutations C \textsuperscript{198}, T \textsuperscript{196}, and G \textsuperscript{195} had a significant influence on I.T. formation. To address this issue, OsO\textsubscript{4} was reacted with a set of topoisomers spanning superhelical densities from 0 to –0.137. At each supercoil density, the percentages of cleavage on the coding strand (from T \textsuperscript{209} to T \textsuperscript{217}) were totaled. Fig. 5A shows the results with py-200, –202G, and –202T; a substantial amount of free energy from supercoiling was required to form the I.T. even in the wild type sequence. In fact, at a superhelical density of –0.06, typical of plasmid DNA isolated from E. coli, only about 10% reactivity was detected. The results from –202G and –202T confirm their strong inhibitory effect. Panel B shows the results with –198C and py-200 (as a dotted line). Whereas the increase in modification for –198C may be attributed to the stronger reactivity at T \textsuperscript{209} and T \textsuperscript{210} (Fig. 3), the significant shift in –\pi at 50% modification (0.068 \pm 0.001 versus 0.078 \pm 0.001) indicates that this mutation decreases the amount of free energy required for the duplex to triplex transition. On the other hand, the data for –196T and –195G, reported in panels C and D, respectively, do not show dramatic perturbations.

We also conducted two-dimensional gel electrophoresis to assay for the relaxation associated with the DNA structural transition. Topoisomers of py-200 and all five mutant plasmids were separated on agarose gels in the same buffer solution used for the chemical modification experiments. Chloroquine concentrations of 4 and 30 \mu M were employed in the second dimension, which afforded the resolution of topoisomers up to 23 negative superhelical turns. A transition centered at 15 negative superhelical turns was observed; however, this was attributed to the formation of a non-B DNA structure in the vector. No transitions due to the duplex-triplex conversion were observed. Two explanations are possible: first, the transition may be too small to be detected in this range of topoisomers or, second, the I.T. structure may be too unstable to survive the electrophoretic conditions.

Kinetics of Modifications at A-T Base Pairs Support the I.T. Model—Reactivities to OsO\textsubscript{4} on the non-coding strand and DEPC at the corresponding adenines were interpreted from a dynamic standpoint.

Scheme 1 represents a bp in conventional duplex DNA. As bp opening depends on \( k_1 \) and \( k_{-1} \), the two constants regulate the extent of modification at both residues. However, if Scheme 2
occurs, in which T may interact with a second partner (X) once in an open conformation, the reactivities will depend on k₁ and k₂ in addition to k₃ and k⁻¹. This is expected to selectively increase modification at A, since this residue remains in an open conformation for a longer time than its partner T. If the percentages of cleavage are expressed as a ratio of T/A, values for A-T pairs involved in a type 2 process will be smaller than those progressing through Scheme 1.

Table I shows the results for A-T bp –226, –225, –222, –219, and –216 for py-200 and mutant plasmids. The values varied considerably, from 0.7 to 66, but, with the exception of –216, they were relatively homogeneous at a given locus. The variations observed may be interpreted in terms of modulation in the accessibility to the reactants. In fact, not shown in the schemes are intermediate states in which a given bp may adopt distorted conformations and/or alterations in the stacking interactions with neighbor residues. These changes, which are favored by high levels of supercoiling and flanking non-B DNA structures such as the I.T., not only facilitate bp opening, but also increase the rate of chemical attack on partially unpaired conformations (42).

Locus –216 appears to be a different case. Here the low values, which spanned a greater range (0.7–6.3), were determined by an increased modification at A(open) (A–216) associated with normal or low cleavages at T (Figs. 4B and 6A) and are accordingly accounted for by a type 2 process. Thus, the ratio of T/A modifications is a measure of the relative chemical accessibility and hence the extent to which the A or the T residue has reassociated with another partner.

Reactivities at the I.T. Flanking Sequence Are Affected by the Adjacent Cloning Site—We observed variations in chemical modifications among the mutants that did not correlate with the stability of the respective triplex structures. For example, the OsO₄ modifications at T–225 and T–226 were –3–4-fold higher in –198C, –196T, and –195G relative to py-200 (Fig. 4B), whereas an increase in stability was apparent only for –198C (Fig. 5). Since this behavior reflected the differences in the donating sites for mutants, a comparison of OsO₄, CAA, and DEPC reactivities between py-200 and py-2005 was performed to resolve this issue. Fig. 6A shows the results from DEPC treatment of the coding strand. The tracings represent the signal from the supercoiled plasmids (–σ = 0.137) less the modification found for the relaxed DNA. Quantitation of the peak areas revealed that A–226 and A–225 acquired an ~2-fold increase when subcloned at the Smal site (py-2005), whereas smaller differences were associated with the other residues. Also, the supercoiling-dependent modifications with OsO₄ at T–227 and T–228 (Fig. 6B) revealed that greater signals were associated with the plasmids containing the γ-globin insert at the Smal site. Hence, these data show that the flanking vector sequences are responsible for the reactivities at these positions, rather than the HPFH point mutations.

**DISCUSSION**

These chemical probe analyses on the 5′-flanking region of the human γ-globin genes enable a molecular description of the I.Ts. at a level of detail previously not possible. The duplex to triplex transition characteristic of oligopurine-oligopyrimidine sequences is accomplished by the purine residues simultaneously engaging in Watson-Crick and Hoogsteen hydrogen bonds (27–30). In general, the bound third strand may occupy a parallel, or antiparallel, orientation relative to the purine residues, depending on its sequence composition. Accordingly, a purine-rich third strand will be accommodated in the major groove in an antiparallel orientation, whereas a pyrimidine-rich third strand will occupy the reverse position. Therefore, stable hydrogen bonds may form between G-G, A-A, and A-T in the former case, and G-C and A-T in the latter (30). Low pH is required in order to stabilize a pyrimidine-rich strand containing cytosine residues in this arrangement.²

The structures formed at the 5′-flanking region of the human γ-globin genes, both in the wild type as well as the HPFH point mutations, deviate considerably from this general scheme. In fact, these and previous data (4) indicate that the third strand is purine-rich, yet low pH is required for stabilization. Our results show that the third strand –209AAGAGGATA–217 is hybridized in an antiparallel orientation to the downstream –198GGGAAGGG–203 containing the sites of mutations. Since these two sequences are 9 and 10 bases long, their inter-

² A colon is used to designate the association between two bases by Hoogsteen or reversed Hoogsteen pairs whereas a center dot designates the interaction between bases in a Watson-Crick pairing motif.
action leads to two possible alignments (Fig. 7). In no case can homogeneous G:A, A:A, or A:T Hoogsteen base pairing take place. Rather, mismatches of the G:A, A:G, and G:T type must also be considered. In both of the reported models, the most abundant triplet is C\textsuperscript{205}G:A\textsuperscript{1}, a combination that has recently been observed in other I.Ts. (43, 44). The stabilization induced by the protonated reversed-Hoogsteen-bound adenines agrees well with our observations. The Hoogsteen G:T pair has been described by NMR only in the parallel orientation (45), where T(H3) shares one hydrogen bond with G(N7). Since parallel and antiparallel thymines display a 2-fold symmetry about N3 (30), it is possible that the antiparallel G:T pair maintains this type of hydrogen bond. Antiparallel A:G has not been documented, however, close interactions are possible.

Overall, the paucity of stable C\textsuperscript{205}G:G and T\textsuperscript{206}A:A triplets, together with the short length of the I.T. stem, accounts for the observed requirement of high levels of supercoiling and the inability to detect a duplex to triplex transition by two-dimensional agarose gel electrophoresis. Finally, both models are consistent with the data which locate C\textsuperscript{198} and C\textsuperscript{202} in the loop.

Allofthe HPFH point mutations alter the normal I.T. structure, some slightly (T\textsuperscript{196} and G\textsuperscript{195}), others profoundly (C\textsuperscript{198}, G\textsuperscript{202}, and T\textsuperscript{202}). The destabilizing effects of G\textsuperscript{202} and T\textsuperscript{202}

| DNA    | -226 | -225 | -222 | -219 | -216 |
|--------|------|------|------|------|------|
| p\textsuperscript{g}-200 | 33   | 62   | 11   | 42   | 3.6  |
| -198C  | 17   | 66   | 9    | 23   | 6.2  |
| -196T  | 14   | 63   | 11   | 50   | 0.9  |
| -195G  | 13   | 54   | 11   | 43   | 0.7  |

Fig. 6. Effect of the cloning site on chemical modification at the I.T. flanking sequence. A, plasmids p\textsuperscript{g}-200 and p\textsuperscript{g}-200S contain the wild type \textgamma-globin insert cloned at the HindIII or SmaI site of pUC9, respectively. The DNAs were reacted with DEPC at −10 (R) and 0.137 (S) and processed as described for the coding strand. Quantitations were performed as follows. Pixel values for each line graph were converted to percentage of the total signal (sum of pixel values for each lane). Percentages of the R sample were then subtracted from those of S after R and S were aligned on their highest pixel value. The range of the x axes was adjusted so as to align the two relevant sets of peaks. Areas were calculated by cutting and weighing the peaks. Values (mg × 10\textsuperscript{3}) at selected locations are reported. Since the amount of material in A−216 for both DNAs was identical (±9%), this reinforces the quantitative methodology. B, supercoiling-dependent OsO\textsubscript{4} modification at T\textsuperscript{227} and T\textsuperscript{228}. The percentage of modification at T\textsuperscript{227} and T\textsuperscript{228} was added and expressed as a single (y) value. Interpolation was conducted as explained in the legend to Fig. 5. For clarity, standard deviations were removed.
observed previously (4), are confirmed. Both of these mutations disrupt a GGGCCC motif, a sequence that has been shown to acquire an induced bend upon complexation of Mg$^{2+}$ ions (46). The effect of these nt changes may be that of abolishing this induced bend, which may represent the nucleation step for I.T. formation or that of disrupting critical Hoogsteen hydrogen bonds. Our results favor the first interpretation.

The stabilization mediated by C–198 was suspected from an earlier work (47). However, the previous studies (4, 47) did not anticipate that this mutation may alter the sequence alignment of the triplex by inducing a slippage of the third strand relative to the Watson-Crick duplex.

The results for T–196 and G–199 are unexpected. Here, we find subtle changes in the overall I.T. structures, whereas substantial destabilizations were predicted from former assays (4). It is likely that these discrepancies originate from the experimental conditions used. In fact, we found no modifications at pH 5.0, whereas SI nuclease degradations were previously detected at this pH. Since I.T. formation occurs in the pH range of 4.5–5.0, slight variations are likely to affect the stabilities greatly. Also, the results of the oligomer binding assays may have been influenced by the distortion of the DNA flanking the I.T., as well as by the difference in the cloning site between pBS200 and mutant plasmids (Fig. 6). The two sets of schemes in Fig. 7 predict stable triplexes for C–198, T–196, and G–199 (48, 49); our data do not permit a delineation between these alternatives.

From a physiological standpoint, the combination of low pH and elevated superhelical density required to induce these structures in vitro may raise concerns about their stability in a cellular environment. However, base protonation may occur, and be maintained, in polynucleotides at several pH units above the pK$_a$ of the free base (50–53). Also, divalent metal ions, polynucleotides (54–58), and the aforementioned single-strand-specific binding proteins may cooperate in lowering the activation energy needed for the I.T. transition.

In vivo, the chromatin in the 5′-flanking region of the human $\gamma$-globin genes has been shown to be hypersensitive to DNase I digestion or restriction enzyme deactivation in cells where the $\gamma$-globin genes are actively transcribed (59, 60). This behavior, which is also observed in other systems, is likely to be correlated with the loss of position nucleosomes along the DNA, and the acquisition of new interactions between cis regulatory sequences and cognate transcription factors (61, 62). Indeed, $\gamma$-globin regulatory elements such as the CACCC and CCAAT boxes appear to be selectively occupied in K562 cells, which express these genes (63, 64).

In vivo, no protein complex has been identified to date that interacts with the upstream region that contains the HPFH point mutations. In addition, experiments conducted in transgenic mice have demonstrated that, at least in the case of G–202, a strong correlation exists between this point mutation and the HPFH phenotype (7). Therefore, a macromolecular complex may assemble at T–200; this complex could be involved in $\gamma$-globin gene silencing. A polypeptide that binds and stabilizes the I.T. structure induced at this location in vitro has been found (14). The interaction of this protein with the I.T. might be altered by any of the HPFH point mutations. It remains to be established whether such an interaction reflects the formation of an I.T. complex that operates in vivo to temporally regulate the expression of the human $\gamma$-globin genes.

Acknowledgments—We thank Dr. Richard R. Sinden and Erna Baum for topoisomerase I and Drs. Xiaolian Gao, Richard Bowater, and Vladimir N. Potaman for suggestions.
48. Durland, R. H., Rao, T. S., Revankar, G. R., Tinsley, J. H., Myrick, M. A., Seth, D. M., Rayford, J. Q., Singh, P., and Jayaraman, K. (1994) Nucleic Acids Res. 22, 3233–3240
49. Dittrich, K., Gu, J., Tinder, R., Hogan, M., and Gao, X. (1994) Biochemistry 33, 4111–4120
50. Inman, R. B. (1964) J. Mol. Biol. 9, 624–637
51. Wells, R. D., and Larson, J. E. (1972) J. Biol. Chem. 247, 3405–3409
52. Carbonnaux, C., van der Marel, G. A., van Boom, J. H., Guschlbauer, W., and Fazakerley, G. V. (1991) Biochemistry 30, 5449–5458
53. Leonard, G. A., Booth, E. D., and Brown, T. (1990) Nucleic Acids Res. 18, 5617–5623
54. Kohwi, Y., and Kohwi-Shigematsu, T. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 3781–3785
55. Tung, C.-H., Breslauer, K. J., and Stein, S. (1993) Nucleic Acids Res. 21, 5489–5494
56. Malkov, V. A., Voloshin, O. N., Soifer, V. N., and Frank-Kamenetskii, M. D. (1993) Nucleic Acids Res. 21, 585–591
57. Washbrook, E., and Fox, K. R. (1994) Nucleic Acids Res. 22, 3977–3982
58. Kang, S., Wohlrab, F., and Wells, R. D. (1992) J. Biol. Chem. 267, 1259–1264
59. Gimble, J. M., Max, E. E., and Ley, T. J. (1988) Blood 72, 606–612
60. Bresnick, E. H., and Felsenfeld, G. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 1314–1317
61. Lewin, B. (1994) Cell 79, 397–406
62. Benezra, R., Cantor, C. R., and Axel, R. (1986) Cell 44, 697–704
63. Ikuta, T., and Kan, Y. W. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 10188–10192
64. Reddy, P. M. S., Stamatoyannopoulos, G., Papayannopoulou, T., and Shen, C.-K. J. (1994) J. Biol. Chem. 269, 8287–8295