Nucleic Acids Research, Vol. 20, No. 20 5339–5343

Effects of sequence and length on imino proton exchange and base pair opening kinetics in DNA oligonucleotide duplexes

M.Leijon and A.Gräslund*
Department of Medical Biochemistry and Biophysics, University of Umeå, S-901 87 Umeå, Sweden

Received July 20, 1992; Accepted August 25, 1992

ABSTRACT

The base catalysed imino proton exchange in DNA oligonucleotides of different sequences and lengths was studied by ^{1}H-NMR saturation recovery experiments. The self-complementary sequences studied were GCGCGAATTCGCGC (I), CGCGAATTCGCG (II), GCGAATTCGC (III), and CGCGATCGCG (IV). The evaluation of base pair lifetimes was made after correction for the measured 'absence of added catalyst' effect which was found to be characterized by recovery times of 400–500 ms for the AT base pairs and 250–300 ms for the GC base pairs at 15°C. End effects with rapid exchange is noticeable up to 3 base pairs from either end of the duplexes. The inner hexamer cores GAATTC of sequences I–II show similar base pair lifetime patterns, around 30 ms for the innermost AT, 5–10 ms for the outer AT and 20–50 ms for the GC base pairs at 15°C. The shorter sequences III and particularly IV show much shorter lifetimes in their central AT base pairs (11 ms and 1 ms, respectively).

INTRODUCTION

The DNA double helix is a dynamic structure with motions occurring on many different time scales. One such motion is the transient opening of an individual base pair, which occurs on the millisecond time scale. A series of studies have shown that the kinetics of base pair opening in oligonucleotides may be studied by ^{1}H-NMR, by observing the base catalysed kinetics of exchange of each imino proton (one per base pair) after selective saturation of its resonance (1–3).

The results have shown that the rate of imino proton exchange and opening of each base pair is an individual property of the base pair. The base pair lifetime around 20°C has typically been reported as of the order of a few ms for AT base pairs, and generally somewhat longer for GC base pairs. There is a certain sequence dependence of the base pair lifetimes, which is particularly significant for long A tracts; here anomalously long lifetimes of the order of 50–100 ms are observed for the central AT base pairs (4, 5) probably correlated with a deformed helix structure with a narrow minor groove and bending of the DNA in these sequences (6).

Another aspect concerns the mechanism of base pair opening, whether it is truly an individual base pair event or if duplex dissociation contributes to the observed proton exchange. In a study addressing this problem (7), no measurable salt or polyamine dependence was found for the proton exchange of the octamer d(GGAATTCC)$_2$ under conditions of high catalyst concentration. The interpretation was that at high pH (above 7.0) localized opening events dominate exchange, whereas at a lower pH opening of several base pairs or duplex dissociation may contribute.

The present study concerns sequence and oligonucleotide length effects on the double helix dynamics on the millisecond time scale. We have addressed the question of how far along a helix the effects of a broken chain are sensed. The phenomenon of 'fraying' at the ends of an oligonucleotide has been noticed in passing in most of the previous studies, but no systematic study of how far into the helix this effect is observable has been made. In principle, this distance may also reflect the maximum distance that the opening of one base pair may affect the stability of the surrounding sequence.

We have undertaken a parallel investigation of imino proton exchange and base pair opening kinetics in three self-complementary oligonucleotides of different lengths (10, 12 and 14 base pairs, sequences I to III, Figure 1) with the same central hexamer core (GAATTC, as in the Eco RI endonuclease recognition sequence) and in a second 10-mer with a different core (CGATCG, sequence IV, Figure 1). All sequences have alternating G's and C's added on each side of the hexamer core. Observations on base pair lifetimes of these duplexes give a basis for the discussion of sequence effects on the base pair lifetimes.

MATERIALS AND METHODS

The oligonucleotides (obtained from Symbicom, Umeå) were synthesized by the phosphoamidite method and purified by reverse phase h.p.l.c. and finally through passage (twice) through a NAP-10 column (G-25 material, Pharmacia). The samples were

* To whom correspondence should be addressed

```
                p q r s t
I:    G C G C G A A T T C G C G C
      1 2 3 4 5 6 7

                q r s t
II:     C G C G A A T T C G C G
        1 2 3 4 5 6

                  r s t
III:      G C G A A T T C G C
          1 2 3 4 5

                  r s t
IV:       C G C G A T C G C G
          1 2 3 4 5
```

Figure 1: The different oligonucleotides studied, with the notation used in Table 1 and Figure 6.

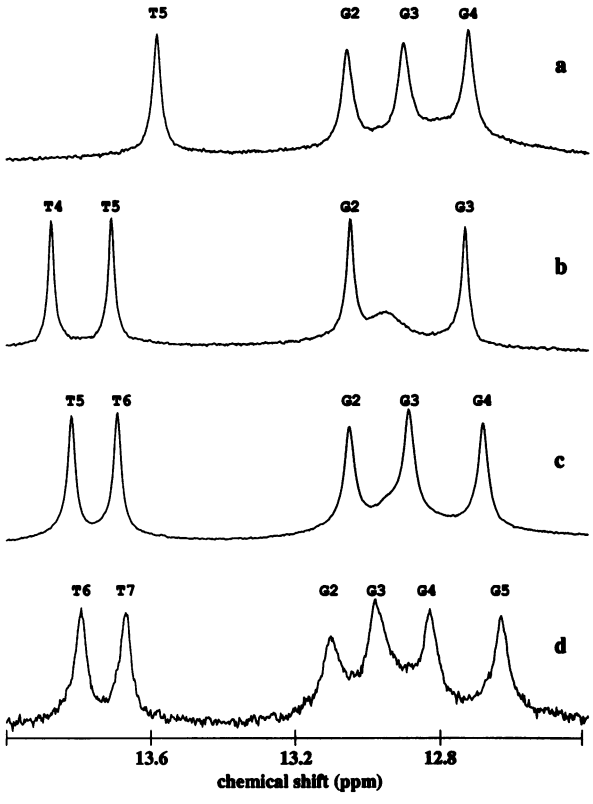


Figure 2: The imino proton NMR spectra of the different oligonucleotides a) IV; b) III; c) II; and d) I (sequences defined in Figure 1) in 150 mM NaCl, 0.5 mM NaEDTA (90% H_2O, 10% D_2O) at 15°C with assignments indicated.

finally dissolved in 150 mM NaCl, 0.5 mM NaEDTA (90% H_20, 10% D_20) and titrated with concentrated ammonia buffer (ammonium chloride and ammonia) at constant pH = 8.8. The concentration of base at each buffer concentration is given by:

$$[\text{base}]^{-1} = (1 + 10^{(pK_a - pH)}) \times [\text{total buffer}]^{-1} \quad (1)$$

where pK_a for the ammonia buffer is 9.25.

^{1}H-NMR experiments were performed on a Bruker AM 500 spectrometer at 15°C. The exchange rates of the imino protons

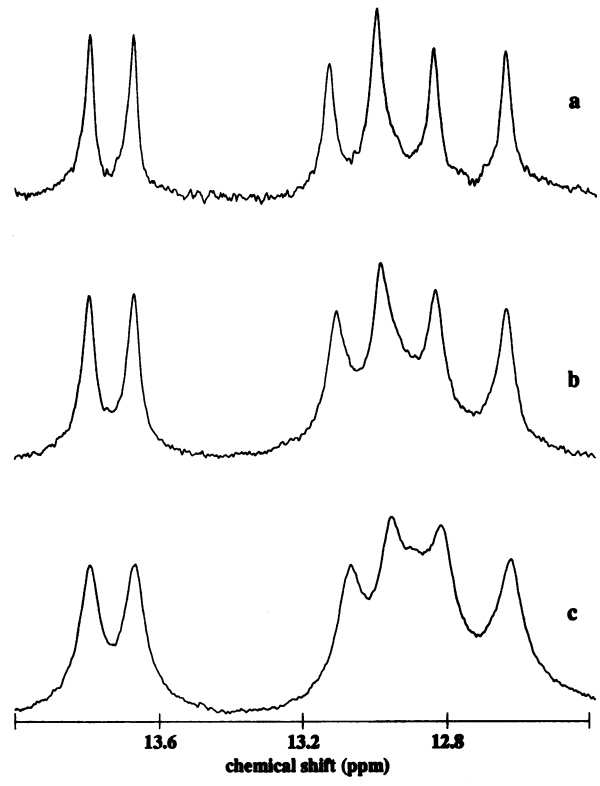


Figure 3: Concentration dependence of the imino proton NMR spectrum of I: a) 1.5 mM; b) 3 mM; and c) 5 mM. Buffer conditions as in Figure 2, 15°C.

were obtained from selective saturation recovery experiments (8, 9, 10) using solvent suppression through the 1331 pulse sequence (11) as previously described (3). The ppm reference was DSS, sodium 3-(trimethyl-silyl)-1-propanesulfonate.

The saturation recovery curves for different values of t (the waiting time between end of saturation and observation pulse) were fitted to an exponential function $I = X_1 + X_2 \cdot \exp(-t/X_3)$, where X_1, X_2 and X_3 were fitted through a non-linear method of least squares.

X_3 is the parameter describing the recovery time T_{rec}. In each experiment series a value of T_{10}, the recovery time in the absence of added catalyst, was measured. T_{10} was generally measured in the presence of a very small amount of buffer (< 10 mM) to stabilize the pH.

RESULTS AND DISCUSSION

Four self-complementary oligonucleotides were studied, with sequences and base pair numbering as shown in Figure 1. Imino proton NMR spectra with assignments are shown in Figure 2. Assignments were made through comparison with previous studies of an identical 12-mer (3) and verified by observation of temperature dependence and one-dimensional NOE effects. The imino proton of the outermost base pair is broadened beyond detection at 15°C and imino proton resonances of the inner base pairs are sequentially broadened from the exterior toward the interior of the helix as the temperature is raised.

For the 14-mer we found a distinct concentration dependence of the imino proton spectra (Figure 3). At high concentrations (*ca.* 5 mM) the resonances are broadened and a new resonance

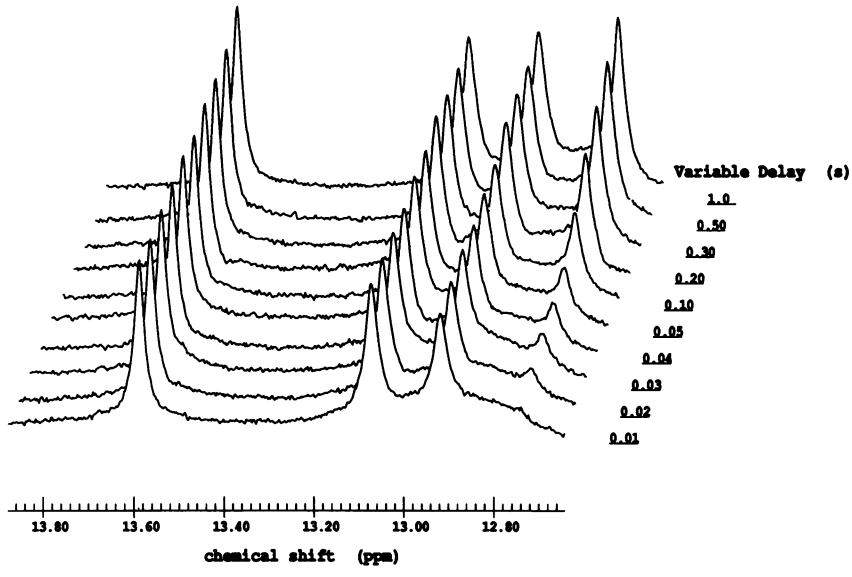


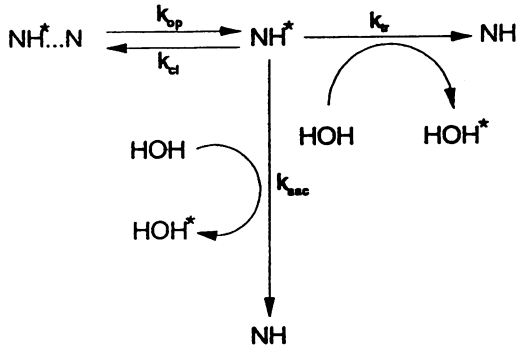
Figure 4: Saturation recovery of the innermost GC base pair in IV in the absence of added exchange catalyst. The variable delay times are indicated.

appears around 12.9 ppm. Attempts to add small amounts of EDTA did not change the spectrum whereas dilution did. Although not further studied, we consider it likely that end-to-end stacking of the oligonucleotides occurs at the higher concentrations and gives rise to the observed effects. The further kinetic studies on the 14-mer were performed with a diluted sample.

The selective longitudinal relaxation time T_1 was measured at 15°C for each imino proton at different concentrations of added base catalyst (ammonia buffer, pH 8.8). Theoretically the following expression is valid:

$$T^{-1}_{rec} = \tau^{-1}_{ex} + T^{-1}_{10} \quad (2)$$

where T^{-1}_{rec} is the measured recovery rate, T^{-1}_{10} is a rate contribution arising from cross-relaxation with neighbouring protons and exchange by internal catalysis (2, 12). τ^{-1}_{ex} is the base catalysed proton exchange rate. The imino proton exchange takes place via transient base pair opening (2, 12, 13):



through base catalysis by added catalyst (rate k_{tr}) or by intrinsic catalysis (rate k_{aac}) by the cyclic nitrogen of the complementary base (2, 12). k_{op} and k_{cl} are the base pair opening and closing rates, respectively. The following expression is valid under conditions of stable structure: $k_{cl} >> k_{op}$. Since $\tau_{ex} = 1/k_{ex}$, $\tau_{op} = 1/k_{op}$, $K_d = k_{op}/k_{cl}$, $T^{-1}_{aac} = K_d k_{aac}$ and $k_{tr} = K_d \alpha k_i [B]$ (14) where α is an accessibility parameter ($0 \leq \alpha \leq 1$), k_i is the proton transfer rate from the isolated nucleotide and [B] is the base catalyst concentration, we obtain

$$\tau_{ex} = \tau_{op} + 1/(K_d \alpha k_i [B] + 1/T_{aac}) \quad (3)$$

When the ammonia base concentration is high, the base catalysed step (k_{tr}) dominates, and the recovery time is given by

$$\tau_{ex} = \tau_{op} + 1/K_d \alpha k_i [B] \quad (4)$$

In the limit of high base catalyst concentration the exchange time ($\tau_{ex} = T_{10}T_{rec}/(T_{10}-T_{rec})$), depends linearly on the inverse base catalyst concentration. An extrapolation to infinite base concentration then gives the base pair lifetime τ_{op}.

Figure 4 shows the recovery of the innermost GC base pair imino proton resonance in CGCGATCGCG after selective saturation. From such experiments on base pair imino protons in all investigated oligonucleotides the recovery times T_{rec} could be obtained by a nonlinear least square fitting of the intensities to an exponential function as described under Materials and Methods.

Figure 5 shows the measured values of τ_{ex}, obtained from the measured values of T_{rec} and T_{10} from eq. (2), as functions of $[B]^{-1}$ towards higher base catalyst concentrations for the oligonucleotides studied. The curves appear fairly linear and were fitted by linear regression to eq. (4).

Figure 6 shows the measured values of T_{10} and the evaluated values of τ_{op} for the base pairs of the studied oligonucleotides. Table 1 summarizes the same values of τ_{op} with corresponding values of the apparent dissociation constant $\alpha K_d = \alpha k_{op}/k_{cl}$.

The measurements were repeated at least twice for every oligonucleotide. The individual lifetimes are reproducible between different experiments within about 50%. For reasonable accuracy it is important to use as concentrated samples as possible, without running into aggregation effects.

The results obtained for T_{10} (Figure 6a) shows that T_{10} generally is around 400–500 ms for AT base pairs and 250–300 ms for GC base pairs at 15°C. These values have contributions

from magnetic dipolar relaxation as well as exchange (in the absence of added catalyst). The relatively short T_{10} values for the central AT base pairs of 10-mer (IV) are consistent with the short lifetime of these base pairs (IV:t, Table 1) which should make the intrinsic exchange process more effective (cf. 2).

There is a general agreement between the pattern of lifetime values found here (Figure 6b and Table 1) at 15°C for the 12-mer (II) and those reported earlier for the same sequence at 25°C using TRIS buffer as catalyst (3, 5). The base pair lifetimes of the AATT base pairs here appear longer than in the previous studies. This may be an effect of the lower temperature in the present study. The lifetime of the innermost GC base pair is here found to be about 25 ms, significantly shorter than previously reported. This should be mainly because a relatively low measured value of T_{10} is included in the present evaluations.

The results of Figure 6b and Table 1 do not include the 2 outermost base pairs, which both have lifetimes shorter than what can be reliably estimated with the present method (< 1 ms). The third base pair from the end (particularly II:q, III:r and IV:r) are similar and have somewhat shorter lifetimes (about 5 ms) and higher αK_d than the GC base pairs further inside in these sequences. The 21 ms determined for I:p could arise because of the extreme length of the 14-mer, but could also be influenced by end-to-end stacking which might contribute to the stability

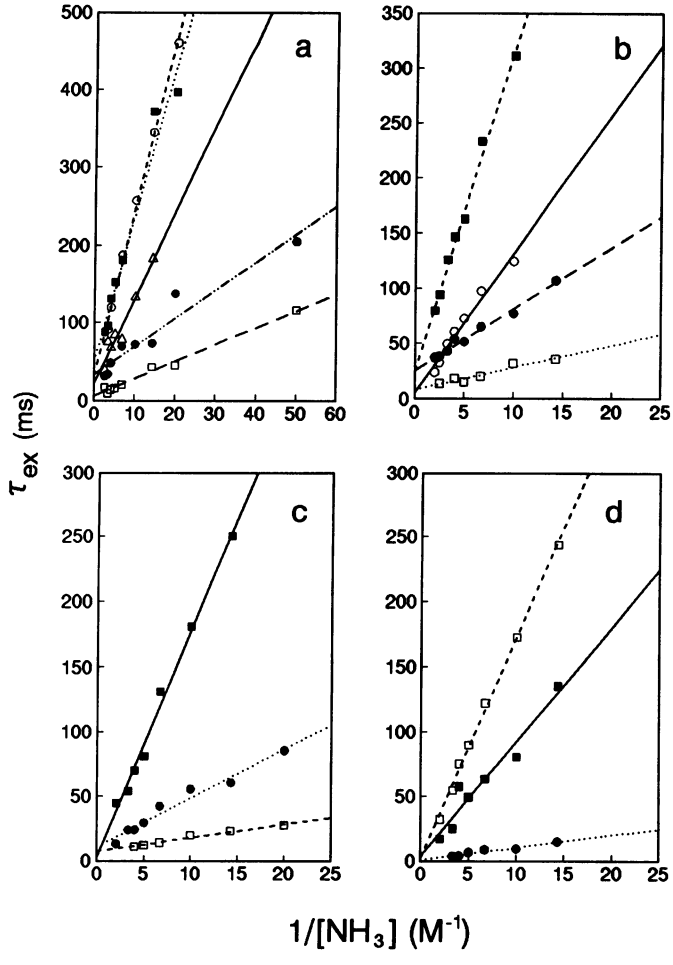


Figure 5: Exchange times τ_{ex} at 15°C as a function of inverse base catalyst concentration for base pairs in a) I: p(△), q(○), r(■), s(□), t(●); b) II: q(○), r(■), s(□), t(●); c) III: r(■), s(□), t(●); d) IV: r(■), s(□), t(●). The labeling is defined in Fig. 1. The samples were dissolved in 150 mM NaCl, 0.5 mM NaEDTA (90% H_2O, 10% D_2O). The catalyst was ammonia buffer at pH 8.8. The straight lines were fitted to eq. (4) through linear regression as described in the text. The base pair lifetimes are obtained as the vertical axis intercepts of these lines.

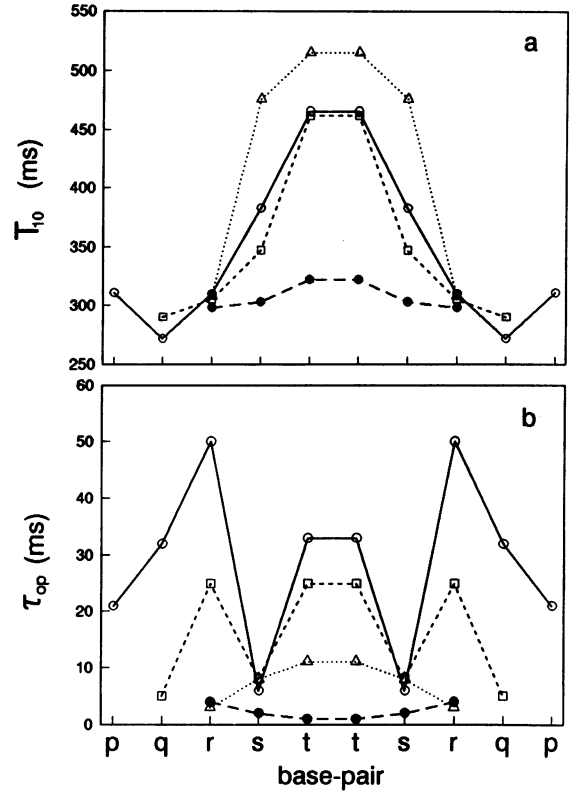


Figure 6: Saturation recovery times in the absence of added catalyst (a), and evaluated base pair lifetimes (b), as functions of base pair position in the different sequences defined in Figure 1: I (○); II (□); III (△); IV (●).

Table 1. Base pair lifetimes and apparent dissociation constants for the different base pairs in the sequences defined in Figure 1 at 15°C. The data is evaluated from the results shown in Figure 5. The value used for k_i when evaluating the apparent dissociation constants, (αK_d), was 2×10^8 (s·M)$^{-1}$ (12). The lifetime results are also presented in Figure 6b.

	base pair lifetime, τ_{op} (ms) base pair					apparent dissociation constant, ($\alpha K_d \times 10^5$) base pair				
sequence	p	q	r	s	t	p	q	r	s	t
I	21	32	50	6	33	0.045	0.023	0.026	0.25	0.13
II		5	25	8	26		0.038	0.017	0.25	0.08
III			3	8	11			0.029	0.5	0.13
IV			4	2	1			0.055	0.029	0.5

of the duplex even after dilution of the sample. The fourth and fifth GC base pairs from the ends (I:q, I:r, II:r) have relatively long and similar lifetimes (25–50 ms), probably not influenced by end effects.

The central -AATT- core shows similar lifetimes in the 12- and 14-mer, about 30 ms for the innermost AT base pairs and 6–8 ms for the outer ones (I:s, I:t, II:s, II:t Table 1). The AT base pairs of the two 10-mers on the other hand show relatively short lifetimes, particularly sequence (IV) which has a lifetime of around 1 ms for the central AT base pairs (IV:t). At the same time the values of αK_d are not changed to the same degree. (Here it should be noticed that αK_d ($= \alpha k_{op}/k_{cl}$) is generally an order of magnitude larger for AT base pairs than for GC base pairs, when the AT and GC basepairs have comparable lifetimes. This means that a GC basepair closes about ten times as fast as an AT basepair with similar lifetime if the accessibility is assumed to be the same.) Similar values of αK_d accompanied by a large decrease in τ_{op} show that a simultanous increase occurs in k_{op} as well as k_{cl}, indicating generally more rapid dynamics in sequence IV compared to I–III. One possible explanation for the different behaviour of IV might be found in the length of the stacked purine sequences, which contain 3 purines in I–III but only 2 in IV. The short lifetime of the innermost GC base pairs, IV:s, in CGCGATCGCG is probably also related to the relatively short lifetime, 1 ms, of the innermost AT base pair, IV:t, in this sequence.

Our results suggest that a longer stretch of stacked purines may contribute to the overall stability of the helix. This view is supported by the observation of anomalously long base pair lifetimes in sequences containing 5 or more successive adenines (4, 5, 15). The end effects on base pair dynamics are generally observable into the third base pair from the end in the investigated 12- and 14-mers. In the 10-mers on the other hand the base pair dynamics in the whole helix seems to be affected by the length of the oligonucleotide, although complete helix dissociation probably does not contribute significantly to the proton exchange (cf. (7)). The short lifetimes of the central AT base pairs in CGCGATCGCG compared to GCGAATTCGC indicate that longer AT-stretches are more stable than short ones, in line with previous results (4, 5, 15). The results further indicate that the dynamics of base pair opening and helix breathing has a significant sequence dependence, and seems to be quite strongly influenced by interactions (mainly stacking?) between purines in neighbouring base pairs.

ACKNOWLEDGEMENTS

This study was supported by grants from the Swedish Natural Science Research Council and the Magn. Bergwall Foundation.

REFERENCES

1. Leroy,J.L., Broseta,D. and Guéron,M. (1985) *J. Mol. Biol.* **184**, 165–178.
2. Guéron,M., Kochoyan,M. and Leroy,J.L. (1987) *Nature* **328**, 89–92.
3. Lycksell,P.O., Gräslund,A., Claesens,F., McLaughlin,W., Larsson,U. and Rigler,R. (1987) *Nucl. Acids Res.* **15**, 9011–9025.
4. Leroy,J.L., Charretier,E., Kochoyan,M. and Guéron,M. (1988) *Biochemistry* **27**, 8894–8898.
5. Moe,J.G. and Russu,I.M. (1990) *Nucl. Acids Res.* **18**, 821–827.
6. Koo,H.S., Wu,H.M. and Crothers,D.M. (1986) *Nature* **320**, 501–506.
7. Braunlin,W.H. and Bloomfield,V.A. (1988) *Biochemistry* **27**, 1184–1191.
8. Johnston,P.D. and Redfield,A.G. (1977) *Nucl. Acids. Res.* **4**, 3599–3615.
9. Patel,D.J., Pardi,A. and Itakura,K. (1982) *Science* **216**, 581–590.
10. Johnston,P.D. and Redfield,A.G. (1981) *Biochemistry* **20**, 3996–4006.
11. Hore,P.J. (1983) *J. Magn. Reson.* **55**, 283–300.
12. Leroy,J.L., Gao,X., Guéron,M. and Patel,D.J. (1991) *Biochemistry* **30**, 5653–5661.
13. Teitelbaum,H. and Englander,S.W. (1975) *J. Mol. Biol.* **92**, 55–78.
14. Eigen,M. (1964) *Angew. Chem. Int. Edn., England* **3**, 1–19.
15. Fritzsche,H., Zdunek,J., Leijon,M. and Gräslund,A. (1992) in preparation.