Sequence-dependent structural features of the DNA double helix have a strong influence on the base pair opening dynamics. Here we report a detailed study of the kinetics of base pair breathing in tracts of GC base pairs in DNA duplexes derived from $^3$H NMR measurements of the imino proton exchange rates upon titration with the exchange catalyst ammonia. In the limit of infinite exchange catalyst concentration, the exchange times of the guanine imino protons of the GC tracts extrapolate to much shorter base pair lifetimes than commonly observed for isolated GC base pairs. The base pair lifetimes in the GC tracts are below 5 ms for almost all of the base pairs. The unusually rapid base pair opening dynamics of GC tracts are in striking contrast to the behavior of AT tracts, where very long base pair lifetimes are observed. The implication of these findings for the structural principles governing spontaneous helix opening as well as the DNA-binding specificity of the cytosine-5-methyltransferases, where flipping of the cytosine base has been observed, are discussed.

Many DNA-binding proteins are highly selective in their recognition of particular DNA sequences. Besides sequence-specific hydrogen bonding and van der Waals interactions, sequence-dependent structure and dynamics of DNA are likely to play an important role in DNA-protein interaction. In addition, the adaptability of a DNA sequence element to structural changes necessary for sequence-specific interaction is important in recognition (1).

Base pair opening is required in many fundamental processes in the cell, for example, transcription and recombination. Recently, base pair opening was found to participate in processes in the cell, for example, transcription and recombination.

**EXPERIMENTAL PROCEDURES**

Sample Preparations and Base-catalyst Titrations—All oligonucleotides were either synthesized by using automated phosphoramidite chemistry on a DNA synthesizer (Applied Biosystems model 394) or purchased from Cybergene Inc. (Sweden). The oligonucleotides were purified by reverse-phase high performance liquid chromatography and desalted by Sephadex G-25 column chromatography. The NMR samples were prepared by dissolving the oligonucleotides in a 3 mM borate buffer at pH 8.8 containing 100 mM NaCl (90% H$_2$O and 10% D$_2$O). The duplex concentrations were in the range 1.1–1.8 mM. Ammonia was added in appropriate amounts from a 6.6 M stock solution at pH 8.8 exchange-time ($t_{\text{ex}}$) data were combined and linearly fit $versus$ the inverse base-concentration ($1/[B]$) via Equation 1 (13).

$$
t_{\text{ex}} = t_{\text{op}} + \frac{1}{k_B [B]} 
$$

(Eq 1)
The base pairs of the duplexes are numbered according to Scheme 1, where numbers from the ends of the decamers, III:T2, IV:T2, and V:T2 are otherwise (24, 25).

The infrared spectra in D2O and H2O of duplex III in the absence of added catalyst and at the ammonia concentration reached at the end point of the titration are shown in Fig. 2. The difference between the two spectra is very small, indicating that no structural alterations occur in course of the titration even at the high duplex concentration used in the IR experiments. The influence of the strong buffer conditions during the titrations was further investigated by performing NOESY experiments on sequence V at an ammonia base concentration of 20 mM and 0.9 M, which corresponds to the titration range. Only minor differences are observed in the chemical shifts and the relative intensity of the cross-peaks in the two spectra (data not shown).

**RESULTS**

**Imino Proton NMR Spectra and IR Spectroscopy**—In the following, GC tracts are synonymous for sequences of GC base pairs with no particular order of guanine and cytosine bases and with a length of at least four base pairs. G tracts are sequences of the type 5'-GnCmG-3'. Isolated GC base pairs will mean at most two consecutive GC base pairs. The studied DNA sequences are shown in Scheme 1. Imino proton spectra of the five DNA duplexes at 15 °C are displayed in Fig. 1, without addition of exchange catalyzing base (left) and in presence of 0.1 M ammonia (right). The assignment of the imino proton resonances is based on two-dimensional 1H NOESY experiments in two-dimensional NOESY experiments in the imino-imino connectivities in two-dimensional NOESY experiments in the imino-imino connectivities in two-dimensional NOESY experiments. The infrared spectra in H2O and D2O were measured with 6 mM oligomer duplex concentration at pH 8.8 and 20 °C as described previously (31).

**Saturation Transfer Experiments**—Saturation transfer experiments utilized a rapid removal of the water magnetization by a 1-1.2-1 ms Gaussian 90°-pulse followed by a pulsed field gradient. This was repeated in a loop 1-10 times. In each loop element the gradient strength and length was randomly varied with 10% around 10 Gauss/cm and 1 ms, respectively, yielding an effective solvent suppression without refocusing effects. Each loop contained a 1-ms gradient recovery delay. The initial loop was followed by a presaturation period of variable length and finally a jump-return observe pulse.2 The decay of the imino proton resonances was fitted to a single exponential function to yield the exchange time and the magnetic spin-lattice relaxation time according to published procedures.

The infrared spectra in H2O and D2O were measured with 6 mM oligomer duplex concentration at pH 8.8 and 20 °C as described previously (24, 25).

The larger broadening observed for III, IV:T2 and III, IV:T3 as compared with V:T2 and V:T3 indicates that the 5'-T2G3' step exerts a destabilizing effect on the ends. This is consistent with the reduced stacking and higher flexibility suggested for this step from the NMR solution structure of sequence III (26). The effect becomes even larger in the dodecamer sequences I and II, where the T2 imino proton resonances have completely disappeared by exchange broadening. This may indicate a cooperative formation of a homogenous G tract type of structure promoted by a flexible 5'-TG-3' step at the G tract ends.

The infrared spectra in D2O and H2O of duplex III in the absence of added catalyst and at the ammonia concentration reached at the end point of the titration are shown in Fig. 2. The difference between the two spectra is very small, indicating that no structural alterations occur in course of the titration even at the high duplex concentration used in the IR experiments. The influence of the strong buffer conditions during the titrations was further investigated by performing NOESY experiments on sequence V at an ammonia base concentration of 20 mM and 0.9 M, which corresponds to the titration range. Only minor differences are observed in the chemical shifts and the relative intensity of the cross-peaks in the two spectra (data not shown).

**Imino Proton Exchange**—The exchange times of the imino protons were obtained from inversion recovery times of the NMR resonances as described previously (9). Addition of an exchange catalyst yields in the limit of infinite catalyst concentration the kinetic parameters for the base pair opening (Equation 1). In Fig. 3, the exchange times of the guanine imino protons of the five GC tracts are displayed as a function of the inverse ammonia concentration at 15 °C. The exchange times display the linear dependence on the inverse base-concentration expected from Eq. 1. The base pair lifetime \( \tau_{\text{op}} \) as well as the apparent dissociation constant \( K_d \), obtained from the linear fits, are given in Table II. Unexpectedly, the base pair lifetimes are below 15 ms for all GC base pairs in the GC tracts. The general pattern is that the lifetimes are much smaller than observed for isolated GC base pair (cf. Table I), and the lifetimes decrease even further when the tract are of the type GmCn. On the average, base pair lifetimes in GC tracts are about 10 times shorter than for isolated GC base pairs (Tables I and II).

Only the averaged lifetimes of the G5 and G6 imino protons of duplex I and the G4 and G5 imino protons of duplex II could be determined due to spectral overlap at high ammonia base concentration (i.e. the last three titration steps). However, at low and intermediate catalyst concentration, it is clear that the exchange properties are very similar for these base pairs (data not shown).

Several sequence-dependent characteristics of the base pair dynamics of the tracts are apparent. In all sequences the outermost GC base pairs display similar behavior with base pair lifetimes around 4 ms and a relatively high base pair dissociation constant in the range 5-14×10^{-7} (Fig. 3, left panel). Notably, from Fig. 3 (A and C), it is seen that the dissociation constant increase with the length of the tract (Table II). In duplex II the central GC step of duplex I has been reversed to a CG step. This leads to an increase of the base pair lifetime and a decrease of the dissociation constant for this base pair as compared with sequence I with roughly a factor of 2 (Fig. 3, A and B; Table II). This is consistent with GmCn-type tracts having unique properties leading to higher base pair dissociation constants (see below). Although the dissociation constants and

\[^{2}M.\ \text{Leijon, manuscript in preparation.}\]
base pair lifetimes of the outermost GC base pairs in the
decamer duplexes are similar (Fig. 3C), the innermost GC base
pairs display different kinetics in the three decamers. The
central base pairs of the alternating tract 5'-GCGC-3' of duplex
IV and 5'-CGCG-3' of duplex V are more stable than any other
base pairs, with dissociation constants \(3.3 \times 10^{-2}\) and \(2.3 \times
10^{-2}\), respectively. These values are close to what is commonly
observed for isolated GC base pairs (12). The central base pair
of sequence III retains the typical properties observed for the
longer G tracts with a high dissociation constant (Fig. 3, C and
D; Table II).

In principle, the increased recovery rates of the imino proton
resonances upon addition of the catalyst could be due to an
increase of the magnetic relaxation rates as well as increased
exchange rates. The former could be the result of aggregation
induced by the high ionic strength present at the high buffer
concentrations necessary to reach near opening-limited ex-
change conditions in the course of the titration. In most studies
of base pair dynamics, it has been implicitly assumed that any
changes in the magnetic relaxation during a catalyst titration
remain so small that they can be neglected. Indeed, this holds
true in most cases (15). However, the unexpected rapid dynam-
ics inferred from the observed exchange behavior of the gua-
ine imino-protons in the G tracts prompted us to investigate
this possibility by carrying out saturation transfer experiments
on sequence V. In this type of experiment, the exchange and the

---

FIG. 1. ¹H NMR imino proton spec-
tra with assignments at 15 °C. No ad-
dition of catalyst (left) and with ammonia
buffer added to yield a base concentration
of 0.1 m (right). The Roman numerals refer
to the numbering in Scheme 1.

FIG. 2. Infrared spectra of duplex
III in D₂O and H₂O (inset). No addition
of catalyst (——) and in the presence of 4
m ammonia buffer at pH 8.8 (-----). The
duplex concentration was 6 mm, and the
spectra were measured at 20 °C.
magnetic relaxation contributions are separated and the exchange is directly measured (23). In Fig. 4, the exchange times derived from the inversion recovery experiments and those derived from the saturation transfer experiments display a close similarity, strongly indicating that changes in magnetic relaxation do not significantly influence the results of the inversion recovery experiments.

**DISCUSSION**

Sequence-dependent structural features of the DNA double helix have a strong influence on the base pair opening dynamics. For instance, the base pair lifetime of the central GC base pairs of the self-complementary duplex d(CCTTTCGAAAGG)$_2$ is 40 ms, while that of the reverse sequence is only 7 ms (Table I). This difference was interpreted as caused by a kink in the center of the helix in the latter case (14). In the DNA dodecamer duplex d(CGCACTGTTGCG)$_2$, the base pair opening rates in the CACA/GTGT motif were 3–8 times higher than in a same dodecamer with the central CG base pair reversed to a GC base pair (16). Despite these differences, GC base pair lifetimes below 5 ms at 15 °C have to our knowledge never been observed in the interior of DNA oligomers where end-fraying effects are negligible. In the present study, base pair lifetimes below 5 ms were observed in almost all of the GC tracts. Structural properties specific for G/GC tracts have been proposed to originate from stacking effects (26, 27), the unmethylated state of the major groove (28, 29), or the pattern of hydrogen-bond donor and acceptor groups in the major groove (30). By comparative studies of different DNA oligomer crystal structures, Dickerson and co-workers (27) concluded that guanine bases prefer to stack flat atop one another, without the helix-following roll observed in A tracts. To maintain hydrogen-
Bonding with the complementary strand, it becomes necessary to break the stack after a certain distance. Hence, a competing situation occurs where the hydrogen-bonding may be compromised for the benefit of optimized stacking. The reason for the difference in stacking properties was suggested to be the presence of a “projecting” N2 amine in guanine. The crystal structure of the GC base pair decamer d(CCGGCCCGG)2 display an unusually deep and wide minor groove and a shallow and accessible major groove. It was implied that shallowness may be a characteristic feature of a major groove devoid of methyl groups. A further manifestation of the accessibility of the major groove in G tract is the unusual groove-backbone interactions observed in the crystal packing of G tract containing oligomers. In fact, A tracts and G tracts exhibit a striking reciprocity with respect to groove dimensions, hydration and base pair dynamics. These features have been summarized in Table III.

It seems probable that one or several of the structural properties listed in Table III are responsible for the very different base pair dynamics observed in A tracts and G tracts. For example, an accessible major groove where the stacking properties of the guanine bases lead to a tendency to transiently break the hydrogen bonding with the complementary strand can give rise to a more rapid base pair dynamics observed in G tracts.

The base pair dissociation constant is higher for G tracts than GC tracts and increases with the length of the G tract. In addition, upon increasing the length of the G tract, the ends of the helix become more labile. A possible explanation is a cooperative formation of a structure in the G tract part with bifurcated hydrogen bonds between the cytosine amino group and the guanine carbonyl group but with poor stacking with the ends. This may lead to a base-pairing shift in the major groove, as has been observed in crystal structures of DNA oligomers with similar sequences, which may stabilize the open state of the base pair leading to a higher base pair dissociation constant.

On the basis of the similarity of the NMR relaxation of the 19F resonances of 5-fluorocytosine at the target site of the cytosine-5-methyltransferase M.HhaI and at a reference position unaffected by protein binding, it was suggested that the protein does not accelerate base pair opening. However, fluorination at the 5-position of uridine leads to an increase of the imino proton exchange by almost a factor of 60, indicating a substantial increase in the base pair opening rate. A similar effect by fluorination of cytosine at the same position could potentially be the dominating contribution to the dynamic behavior in both the absence and presence of the protein. In the present study, we have shown that tracts of four or more GC base pairs exhibit a unique rapid opening dynamics that may contribute to the specificity of the cytosine-5-methyltransferases. It should be noted that it is the guanine base that carries the solvent exchangeable imino proton that is used in deriving the base pair dynamics. In case the fluctuations of the two bases in the GC base pair are independent, i.e., the opening is asymmetric, no information is provided on the fluctuations of the cytosine base. However, no evidence exists that base pair opening is asymmetric in the DNA double helix, although this appears to be the case in PNA-DNA hybrid. Furthermore, the cytosine base is not recognized itself by the methylation transferases. For example, M.HhaI binds with higher affinity to an abasic site than to the cognate sequence.

Interestingly, M.HhaI and M.HacIII cause only quite modest deformation of the DNA helix, whereas adenine-N6-methyltransferase (M.EcoRI) with the recognition sequence 5'-GAAATTC-3' causes a severe 52° bend of the DNA helix. It is possible that this reflects the intrinsic tendency for the G tract base pair to open. On the other hand, the stable, A tract type of recognition sequence of M.EcoRI may require larger deformations of the helix to facilitate the base pair opening.

REFERENCES

1. Steitz, T. A. (1990) Q. Rev. Biophys. 23, 205–280
2. Klimasauskas, S., Kumar, S., Roberts, R. J., and Cheng, X. (1994) Cell 76, 357–369
3. Reinsch, K. M., Chen, L., Verdine, G. L., and Lipscomb, W. N. (1995) Cell 82, 143–153
4. Cheng, X., and Blumenthal, R. M. (1996) Structure 4, 639–645
5. Klimasauskas, S., and Roberts, R. J. (1995) Gene (Amst.) 157, 163–164
6. Yang, A. S., Shen, J. C., Zingg, J. M., Mi, S., and Jones, P. A. (1995) Nucleic Acids Res. 23, 1380–1387
7. Leroy, J. L., Guo, X. L., Guéron, M., and Patel, D. J. (1991) Biochemistry 30, 5653–5661
8. Leroy, J. L., Guo, X. L., Misra, V., Guéron, M., and Patel, D. J. (1992) Biochemistry 31, 1407–1415
9. Lejon, M., Schüttel, U., Nielsen, P. E., and Graslund, A. (1997) J. Mol. Biol. 271, 438–455
10. Nelson, H. C., Finch, J. T., Luis, B. F., and Klug, A. (1987) Nature 330, 221–226
11. Leroy, J. L., Charretier, E., Kouchou, M., and Guéron, M. (1988) Biochemistry 27, 8894–8898
12. Guéron, M., and Leroy, J. L. (1992) in Nucleic Acids and Molecular Biology (Eckstein, F., and Lilley, D. M. J., eds) Vol. 6, pp. 1–22, Springer Verlag, Berlin.
13. Guéron, M., and Leroy, J. L. (1995) Methods Enzymol. 261, 383–413
14. Kouchou, M., Leroy, J. L., and Guéron, M. (1987) J. Mol. Biol. 196, 599–609
15. Lejon, M., and Leroy, J. L. (1997) Biochimie (Paris) 79, 775–779
16. Polt-Stogniew, E., and Russu, I. M. (1994) Biochemistry 33, 11016–11024
17. Guéron, M., Charretier, E., Hagerhorst, J., Kouchou, M., Leroy, J. L., and Moraillon, A. (1990) in Structure and Methods (Sarma, R. H., and Sarma, M. H., eds) Vol. 3, pp. 113–137, Adenine Press, Guildenland, NY
18. Boelens, R., Scheek, R. M., Dijkstra, K., and Kaptein, R. (1985) J. Magn. Reson. 62, 378–386
19. Piotto, M., Saudek, V., and Sklenar, V. (1992) J. Biol. Chem. 267, 1661–1665
20. Plateau, P., and Guéron, M. (1982) J. Am. Chem. Soc. 104, 7310–7311

**TABLE III**

| Properties of A and G tract DNA |
|--------------------------------|
| A tract                        |
| Narrow minor groove            |
| Wide and deep major groove     |
| Methylated major groove        |
| No amino group in the minor groove |
| Hydration spine in the minor groove |
| Very slow base pair dynamics   |
| G tract                        |
| Wide and deep minor groove     |
| Shallow major groove           |
| Unmethylated major groove      |
| Amino group in the minor groove |
| Poor hydration of the minor groove |
| Very rapid base pair dynamics  |
Base Pair Breathing Dynamics of GC Tracts

21. Bauer, C., Freeman, R., Frenkiel, T., Keeler, J., and Shaka, A. J. (1984) J. Magn. Reson. 58, 442–457
22. Geen, H., and Freeman, R. (1992) J. Magn. Reson. 93, 93–141
23. Leijon, M. (1996) J. Magn. Reson. B 112, 181–185
24. Sarkar, M., Dornberger, U., Rozners, E., Fritzsche, H., Strömberg, R., and Graslund, A. (1997) Biochemistry 36, 15463–15471
25. Dornberger, U., Behlke, J., Birch-Hirschfeld, E., and Fritzsche, H. (1997) Nucleic Acids Res. 25, 822–829
26. Dornberger, U., Flemming, J., and Fritzsche, H. (1998) J. Mol. Biol. 284, 1453–1463
27. Dickerson, R. E., Goodsell, D., and Kepka, M. L. (1996) J. Mol. Biol. 256, 168–125
28. Heinemann, U., and Alings, C. (1989) J. Mol. Biol. 210, 369–381
29. Heinemann, U., and Hahn, M. (1992) J. Biol. Chem. 267, 7332–7341
30. Timsit, Y., Vilbois, E., and Moras, D. (1991) Nature 354, 167–179
31. Heinemann, U., Alings, C., and Bansal, M. (1992) EMBO J. 11, 1931–1939
32. Timsit, Y., Westhof, E., Fuchs, R. P., and Moras, D. (1989) Nature 341, 459–462
33. Klimasauskas, S., Szyperski, T., Serva, S., and Wuthrich, K. (1998) EMBO J. 17, 317–324
34. Gmeiner, W. H., Sahasrabudhe, P., and Pon, R. T. (1995) Magn. Reson. Chem. 33, 449–452
35. García, R. A., Bustamante, C. J., and Reich, N. O. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 7618–7622