ACKNOWLEDGMENTS

We thank Mitch Lewis for helpful discussion; Anna Jeitler-Nilsson, Kathy Hehir, and Andrea Jeffrey for help with protein purification; Kathy Hehir and Andrea Jeffrey for protein sequencing; Michael Hecht, Robert Stearman, and Hillary Nelson for mutant plasmids; Eric Suchanek for computer graphics; Dinshaw Patel, David States, Jeff Hoch, David Ruben, Ponzy Lu, and Alfred Redfield for advice regarding NMR techniques; and Anne Lawthers for preparation of the manuscript. The 500-MHz spectra were recorded at the High-Field NMR Resource of the Francis Bitter National Magnet Laboratory (NIH, RR-00995).

REFERENCES

Ackers, G. K. (1970) *Adv. Protein Chem. 24*, 343–446.
Amann, E., Brosius, J., & Ptashne, M. (1983) *Gene 25*, 167–174.
Aue, W. P., Bartholdi, E., & Ernst, R. R. (1976) *J. Chem. Phys. 64*, 2229–2241.
Chadwick, P., Pirrotta, V., Steinberg, R., Hopkins, N., & Ptashne, M. (1971) *Cold Spring Harbor Symp. Quant. Biol. 35*, 283–294.
Feeney, J., Batchelor, J. G., Albrand, J. P., & Roberts, G. C. K. (1979) *J. Magn. Reson. 33*, 519–529.
Hecht, M. H., Nelson, H. C. M., & Sauer, R. T. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 2676–2680.
Johnson, A. D. (1980) Ph.D. Thesis, Harvard University.
Johnson, A. D., Pabo, C. O., & Sauer, R. T. (1980) *Methods Enzymol. 65*, 839–856.
Karplus, M., & McCammon, A. (1983) *Annu. Rev. Biochem. 52*, 263–300.
Lewis, M., Jeffrey, A., Wang, J., Ladner, R., Ptashne, M., & Pabo, C. O. (1983) *Cold Spring Harbor Symp. Quant. Biol. 47*, 435–440.
Maniatis, T., Ptashne, M., Backman, K., Kleid, D., Flashman, S., Jeffrey, A., & Mauer, R. (1975) *Cell (Cambridge, Mass.) 5*, 109–113.
Nelson, H. C. M., Hecht, M. H., & Sauer, R. T. (1983) *Cold Spring Harbor Symp. Quant. Biol. 47*, 441–449.
Ogata, R. T., & Gilbert, W. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 5851–5854.
Pabo, C. O. (1980) Ph.D. Thesis, Harvard University.
Pabo, C. O., & Lewis, M. (1982) *Nature (London) 298*, 443–447.
Pabo, C. O., & Suchanek, E. (1986) *Biochemistry 25*, 5987–5991.
Pabo, C. O., Sauer, R. T., Sturtevant, J., & Ptashne, M. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 1608–1612.
Sauer, R. T., & Anderegg, R. (1978) *Biochemistry 17*, 1092–1100.
Sauer, R. T., Pabo, C. O., Meyer, B. J., Ptashne, M., & Backman, K. C. (1979) *Nature (London) 279*, 396–400.
Sauer, R. T., Hehir, K., Stearman, R. S., Weiss, M. A., Jeitler-Nilsson, A., Suchanek, E. G., & Pabo, C. O. (1986) *Biochemistry 25*, 5992–5998.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.
States, D. J., Dobson, C. M., Karplus, M., & Creighton, T. E. (1984) *J. Mol. Biol. 174*, 411–418.
Thomas, G. J., Jr., Prescott, B., Benevides, J. M., & Weiss, M. A. (1986) *Biochemistry 25*, 6768–6778.
Weiss, M. A., Sauer, R. T., Patel, D. J., & Karplus, M. (1984) *Biochemistry 23*, 5090–5095.
Weiss, M. A., Stearman, R., Jeitler-Nilsson, A., Karplus, M., & Sauer, R. T. (1986) *Biophys. J. 49*, 29–33.
Weiss, M. A., Karplus, M., & Sauer, R. T. (1987) *Biochemistry* (preceding paper in this issue).

Proton Nuclear Magnetic Resonance Studies on Bulge-Containing DNA Oligonucleotides from a Mutational Hot-Spot Sequence†

Sarah A. Woodson and Donald M. Crothers*
Department of Chemistry, Yale University, New Haven, Connecticut 06511
Received June 16, 1986; Revised Manuscript Received September 23, 1986

ABSTRACT: A series of bulge-containing and normal double-helical synthetic oligodeoxyribonucleotides, of sequence corresponding to a frame-shift mutational hot spot in the λ C_I gene, are compared by proton magnetic resonance spectroscopy at 500 MHz. The imino proton resonances of d(GATGGGCAG)·d(CTGCCCATC), d(GATGGGCAG)·d(CTGCCCCATC), and d(GATGGGCAG)·d(CTGACCCATC) are assigned by one-dimensional nuclear Overhauser effect spectroscopy. Nonselective T_1 inversion–recovery experiments are used to determine exchangeable proton lifetimes and to compare helix stability and dynamics of the three duplexes. An extra adenosine flanking the internal G·C base pairs has a strongly localized effect on helix stability, but the destabilizing effect of an extra cytidine in a C tract is delocalized over the entire G·C run. These data lead to the conclusion that the position of the bulge migrates along the run in the fast-exchange limit on the NMR time scale. Rapid migration of the bulge defect in homopolymeric sequences may help rationalize both frame-shift mutagenesis and translational frame shifting. We estimate that the unfavorable free energy of a localized bulge defect is 2.9–3.2 kcal/mol, in good agreement with earlier estimates for RNA helices.

Frame-shift mutations are known to occur predominantly in monotonous runs of base pairs, which can slip one or two positions and still maintain base pairing (Okada et al., 1972; Ames et al., 1973), but the precise role of local DNA structure in mutational events is not yet well understood. Early studies (Lerman, 1963, 1964; Drake, 1964) of mutation genetics led Streisinger to propose a bulged base model for frame-shift

†Supported by Grant GM21966 from the National Institutes of Health.

0006-2960/87/0426-0904$01.50/0

mutation (Streisinger et al., 1966; Okada et al., 1966). According to the model, slippage of one strand of DNA causes one or more bases to loop out. Drug binding stabilizes the lesion, allowing repair processes to seal in the bulge, resulting in either +1 or −1 frame shifts after replication (Mizusawa et al., 1981).

Both DNA sequence and specific drug–DNA interactions play important roles. Calos and Miller (1981) found that in the *lac* i gene, most frame-shift mutations occurred in sequences containing a run of three G·C base pairs flanked by a C·G, with -CGGGG·GCCCC- having the highest frequency. Intercalating drugs may stack with one strand opposite a bulge to stabilize the extra base (Young & Kallenbach, 1981; Lee & Tinoco, 1978), but clearly the geometry is important since some intercalators, such as ethidium bromide, do not cause frame-shift mutations. Nelson and Tinoco (1985) have recently shown that ethidium has an increased affinity for an oligonucleotide containing a bulged cytidine.

Current physical techniques and the availability of synthetic DNA sequences have made it possible to address these questions in a much more defined way. ^{1}H NMR[1] has already been used to study the structure of oligonucleotides containing bulged bases and loops (Patel et al., 1982, 1983; Pardi et al., 1982; Haasnoot et al., 1980), as well as drug binding (Patel & Shapiro, 1985; Patel, 1974; Pardi et al., 1983) and base modifications (Patel et al., 1985; Quignard et al., 1985). We have examined a series of DNA oligomers containing bulges by ^{1}H NMR in order to better understand how local DNA structure can accommodate or even promote helix perturbations that could lead to mutation events.

Our model system is a sequence derived from the "hot spot" for frame-shift mutagenesis in the λ C_1 gene (Skopek & Hutchinson, 1984). The sequences

```
5'dGATGGGCAG     5'dGATGGG-CAG     5'dGATGGG-CAG
  CTACCCGTC         CTACCCCGTC        CTACCCAGTC
```

were synthesized by the phosphoramidite method. Imino proton resonances were assigned by one-dimensional NOE difference spectroscopy, and complete assignment of the nonexchangeable protons is under way by two-dimensional nuclear Overhauser effect spectroscopy (NOESY) and correlated spectroscopy (COSY) experiments. Lifetimes of the imino protons were measured by T_1 inversion–recovery experiments. Comparison of these sequences demonstrates distinct differences between bulges in homogeneous and heterogeneous sequences.

Materials and Methods

Oligonucleotides were synthesized on a 10-μmol scale in an Applied Biosystems 380B DNA synthesizer. Reagents for DNA synthesis were supplied by Applied Biosystems (Foster City, CA). Purification was on a Waters HPLC system with a semipreparative Vydac C_4 peptide column or a Macherey-Nagel Nucleogen C_8 column at 30 or 35 °C. Purity was checked by analytical reverse-phase chromatography and 20% denaturing polyacrylamide gel electrophoresis. Overall yield was 35% for G-rich sequences and 50% for C-rich sequences.

Samples containing 100–150 A_{260} units of duplex were dialyzed extensively against 1 M NaCl, 10 mM phosphate, pH 7.0, and 0.1 mM EDTA, followed by 83.9 mM $NaClO_4$, 10 mM phosphate, and 0.1 mM EDTA, and dissolved in 400 μL of the same, with 10% D_2O. Tris samples were dialyzed against water and then dissolved in 20 mM perdeutero-Tris-HCl, pH 7.0, 0.1 mM EDTA, 80 mM $NaClO_4$, and 10% D_2O. Final sample concentration was 2.0–2.5 mM duplex. For ^{31}P spectra, samples were dialyzed extensively against 1 M NaCl, 10 mM phosphate, pH 7.0, 10 mM EDTA, and 10 mM EGTA and then against 10 mM phosphate, 50 mM NaCl, 10 mM EDTA, and 10 mM EGTA.

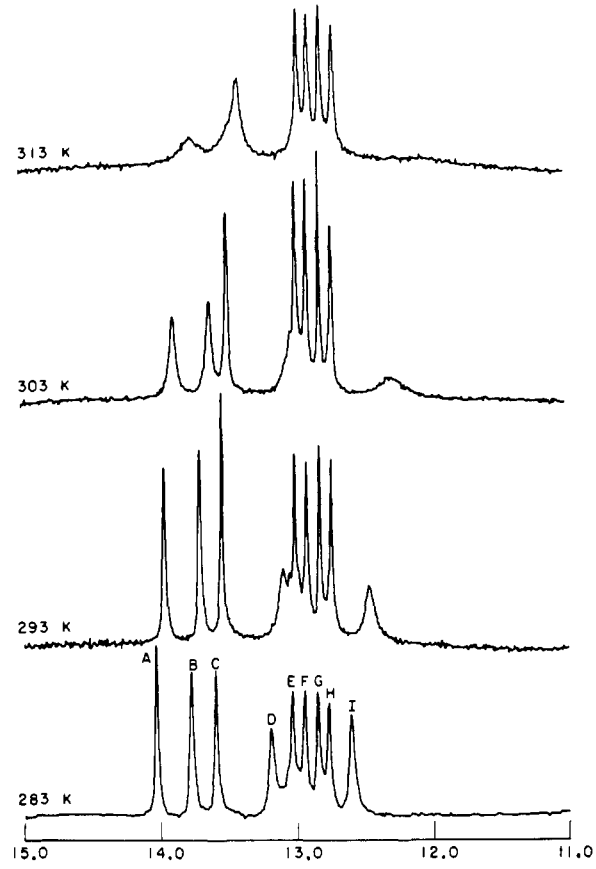


FIGURE 1: Imino proton spectra of perfect 9-mer 5′-d-(GATGGGCAG)·d(CTGCCCATC), at 490 MHz in 10 mM phosphate, pH 7.0, and 100 mM Na^+. Chemical shifts are referenced to internal dioxane at 3.741 ppm relative to DSS. Assignments are shown in Table I.

^{1}H NMR spectra were acquired on the Yale 490 and Bruker WM-500 spectrometers, with the use of twin-pulse observation (Kime & Moore, 1983) and ADA data acquisition (Roth et al., 1980) to suppress the water resonance. Chemical shifts were referenced to internal dioxane set at 3.471 ppm relative to 3-(trimethylsilyl)-1-propanesulfonic acid. One-dimensional NOE difference spectra were acquired in a similar manner, with a 0.3-s irradiation time and 50% saturation of the desired resonance. A Lorentzian-to-Gaussian multiplication was applied to difference FIDs to enhance resolution. Inversion–recovery experiments used a 90–D_1–90–τ–45–D_2–45 pulse sequence with the carrier frequency in the center of the imino region and ADA acquisition with a 1.2-s relaxation delay. FIDs were weighted with an exponential function (3–5 Hz) to improve the signal-to-noise ratio. T_1 values were determined by a linear least-squares program. Error in the relaxation rates was generally within 10%.

Optical melting curves were recorded on a Cary 219 UV–vis spectrometer. Samples were in the same buffer as that for proton NMR experiments, with total strand concentrations ranging from 5 to 100 μM. Data were fit to a two-state model by use of a nonlinear least-squares program. Extinction

[1] Abbreviations: NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; FID, free induction decay; EDTA, ethylenediaminetetraacetic acid; EGTA, [ethylenebis(oxyethylenenitrilo)]tetraacetic acid; DSS, 3-(trimethylsilyl)-1-propanesulfonic acid; HPLC, high-performance liquid chromatography; Tris, tris(hydroxymethyl)aminomethane.

Chart I

```
Perfect 9-mer          1 2 3 4 5 6 7 8 9
                   5'd G A T G G G C A G
                       C T A C C C G T C

C Bulge 9-mer          1 2 3 4 5 6   7 8 9
                   5'd G A T G G G - C A G
                       C T A C C C C G T C

A Bulge 9-mer          1 2 3 4 5 6   7 8 9
                   5'd G A T G G G - C A G
                       C T A C C C A G T C
```

Table I: Chemical Shifts and Assignments of Imino Protons

resonance	base pair	chemical shift (ppm) 9-mer	C bulge	A bulge
A	A·T 7	13.966	14.156	14.229
B	A·T 2	13.739	13.738	13.767
C	T·A 3	13.557	13.562	13.632
D	G·C 9	13.156	13.144	13.128
E	G·C 5	13.002	13.038	13.016
F	G·C 6	12.912	13.011	13.040
G	C·G 7	12.818	12.926	12.907
H	G·C 4	12.735	12.775	12.771
I	G·C 1	12.569	12.599	12.640

coefficients were calculated from values given by Janik (1971).

Results

Perfect 9-mer. The imino region of the proton spectrum of the perfect 9-mer at 490 MHz is shown in Figure 1. At 283 K, all nine resonances are clearly resolved, corresponding to three A·T base pairs between 14.0 and 13.5 ppm and six G·C base pairs between 13.2 and 12.5 ppm. At 293 K, two G·C resonances begin to broaden, and at 303 K, two A·T resonances also broaden and disappear as the temperature is increased. The other G·C resonances finally begin to broaden simultaneously at 323 K. This corresponds exactly to what one would expect for this sequence (Chart I) with a gradual fraying of the helix ends, as previously observed (Pardi et al., 1982).

Hydrogen-bonded imino protons on adjacent base pairs in a DNA duplex are near enough for NOE effects between them to be observed. This fact can be used to assign the imino resonances in a ladderlike fashion, as demonstrated by Roy and Redfield (1981). Some of the NOE difference spectra for the perfect 9-mer are shown along with the reference spectrum in Figure 2. Resonance C is an A·T base pair that broadens at higher temperatures than those at which A and B broaden, so it is assigned to T·A 3. B and C give NOEs to each other, and C and H give NOEs to each other. The other A·T resonance A gives an NOE to G. Continuing in this fashion, one arrives at the assignment shown in Table I. It was not possible to observe NOEs to the terminal G·C resonances even at low temperature, but in several experiments where the signal-to-noise ratio was very high, irradiation of I gave a weak NOE peak at resonance B, so I is assigned to G·C 1 and D to G·C 9. This conclusion remains consistent with later data. Irradition of A·T resonances A, B, and C gave sharp NOE peaks in the aromatic region corresponding to the H-2 protons. A weaker effect was seen to the neighboring A H-2. The spectrum of the analogous 8-mer d-(GATGGCAG)·d(CTGCCATC) is very similar, except that resonance E is absent (data not shown). An independent assignment of this compound confirms the assignment given for the 9-mer.

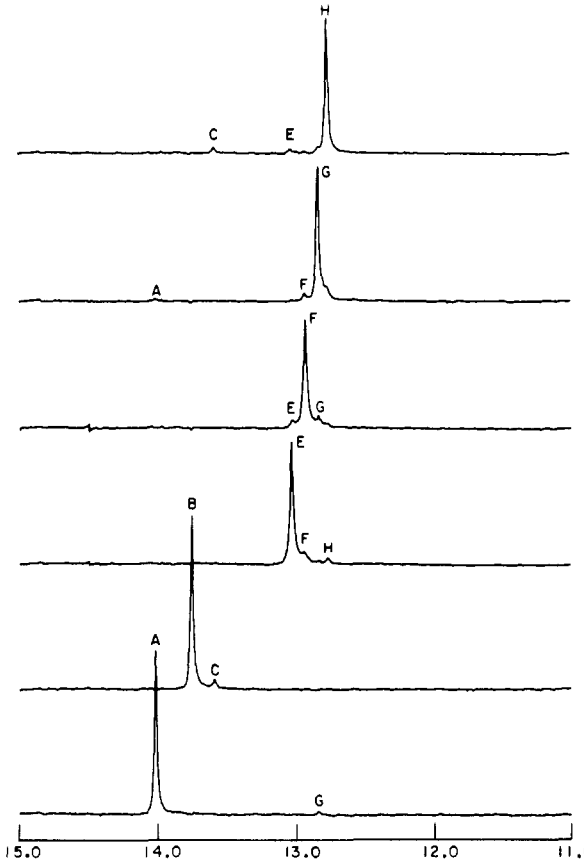


FIGURE 2: 500-MHz NOE difference spectra of perfect 9-mer at 283 K. Labeling of resonances is the same as in Figure 1. Off- and on-resonance FIDs were differenced and weighted with a Gaussian function before transformation.

The T_1 values for the perfect 9-mer in phosphate and Tris buffers are given in Table II. Each measurement was based on 20 τ values, which were fit to the equation

$$I_z = I_0 + e^{-k\tau+b}$$

where I_z is the observed intensity at time τ, I_0 is the equilibrium intensity at long τ, and k is the relaxation rate, or $1/T_{1_{obsd}}$. The observed relaxation of exchangeable protons has both magnetic and exchange components, where

$$1/T_{1_{obsd}} = 1/T_{1_{magn}} + 1/\tau_{exchange}$$

Magnetic contributions to relaxation predominate at low temperatures, but exchange becomes more important at higher temperatures as the helix begins to melt.

The theory of imino proton exchange of oligonucleotides has been described elsewhere (Pardi & Tinoco, 1982; Crothers et al., 1974). The rate-limiting factor for proton exchange depends on the relative rates of helix opening and closing and on the rate of the proton-transfer reaction in which an exposed imino proton is transferred to a base catalyst, generally buffer or OH^- (Crothers et al., 1974). Determination of the rate of helix opening requires extrapolation to high buffer concentration to remove proton transfer as a potential rate-determining factor (Leroy et al., 1985). Since our interest in proton exchange rates lies in comparison of the perfect duplex with bulge-containing molecules in order to identify regions of the molecule destabilized by the bulge, we have not attempted to determine actual helix-opening rates. Instead, we rely on comparison of proton exchange rates at analogous sites in the set of molecules.

Exchange rates for the perfect 9-mer are similar to those previously observed (Pardi et al., 1982; Pardi & Tinoco, 1982).

Table II: T_1 Relaxation Times for Perfect 9-mer (ms)

buffer	temp (K)	base pair 1	2	3	4	5	6	7	8	9
10 mM phosphate, pH 7.0	278	37	202	317	262	240	247	252	177	32
	283	40	230	388	407	377	362	380	182	29
	288	21	116	314	318	322	302	317	96	11
	293	11	66	235	335	326	299	315	54	16
	298	6	37	164	338	328	292	319	29	14
	303		17	76	324	349	277	270	15	
	308		7	36	264	284	188	176	7	
	313			14	82	88	78	68		
	318				22	25	24	21		
20 mM Tris-HCl, pH 7.0	278	27	116	173	163	174	163	163	79	32
	283	16	77	194	194	193	199	198	52	18
	288	7	46	171	230	232	237	229	31	7
	293		25	103	258	258	257	261	18	
	298		17	62	269	275	252	268	11	
	303		4	25	260	264	236	226	4	
	308			11	166	224	179	146		

The terminal G·C resonances have very short T_1 values of roughly 30–40 ms at 278 and 283 K, decreasing rapidly to 10 ms or less by 298 K. The penultimate A·T resonance proton lifetimes are 230 and 182 ms in 10 mM phosphate at 283 K. They are less subject to end fraying and reach a lifetime of 7 ms at 308 K. In phosphate buffer, the lifetimes of the internal resonances are long, between 300 and 400 ms, and increase slightly with temperature until 303 K. Above 303 K, exchange mechanisms begin to predominate over dipolar relaxation, and the apparent T_1 values decrease rapidly to 25 ms at 318 K.

In 20 mM Tris-HCl, pH 7.0, the observed relaxation rates are somewhat shorter, although the overall pattern of temperature dependence remains the same. The terminal resonances are particularly affected, reaching a lower limit of 7 ms at 288 K instead of 298 K in phosphate. The core G·C resonances have lifetimes of 250–275 ms at 293 K.

An Arrhenius plot of the relaxation rates vs. inverse temperature yields an estimate of the activation energy ΔE_a for the overall process leading to proton exchange, where

$$k_{obsd} = 1/T_{1_{magn}} + k_{ex}$$

and at high temperature

$$k_{obsd} \approx k_{ex} = Ae^{-\Delta E_a/RT}$$

Only data points at higher temperature that had a logarithmic temperature dependence were used to determine the nearest fit. In some cases, the contribution of several mechanisms to the observed relaxation rate produced a multiple exponential curve which was difficult to analyze. In these cases, the highest temperature points were fit to a single exponential as nearly as possible. While these activation energies measure the sum of pathways leading to proton exchange, it is useful to compare the values obtained for analogous base pairs in the three duplex molecules.

The activation energies are listed in Table V. For the perfect 9-mer they range from 50 kcal/mol for internal base pairs to 22 kcal/mol for terminal base pairs, which correspond to the limits previously reported for cooperative transitions and single base pair opening (Pardi et al., 1982). Arrhenius plots for base pairs A·T 3, G·C 4, and G·C 7 are shown in Figure 7. All data points are plotted; however, the lines shown were only fitted to the points at high temperature. Semilogarithmic plots are shown for ease of visual comparison, but exponential plots gave better results and were used to obtain the values in Table V.

C-Bulge 9-mer. The imino region of the spectrum of the C bulge containing oligonucleotide d(GATGGGCAG)·d(CTGCCCCATC) is shown in Figure 3. All nine imino

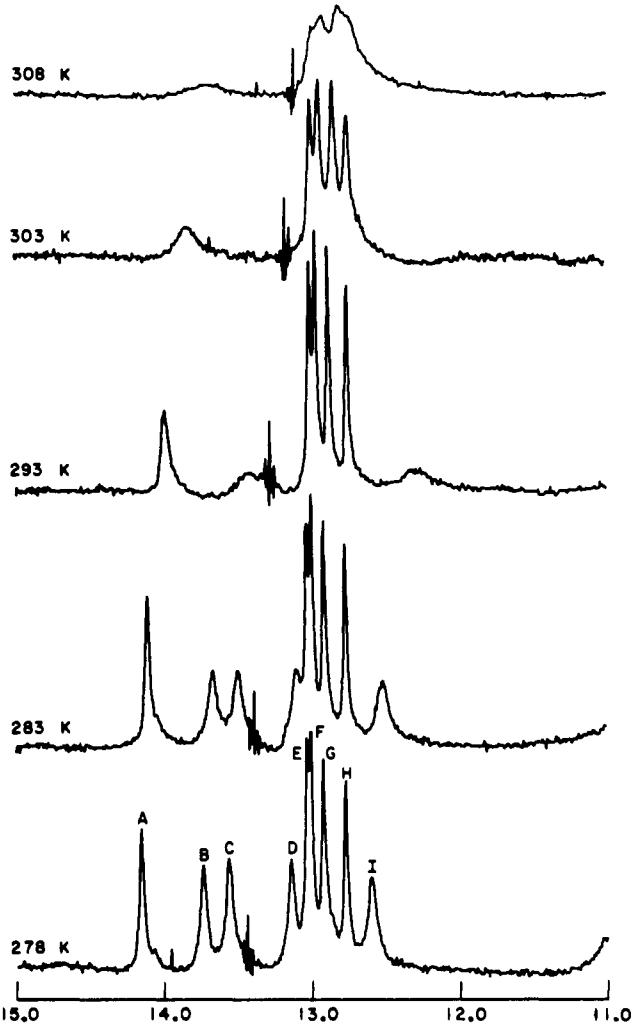


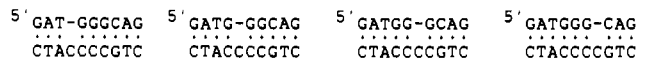
FIGURE 3: Imino proton spectra of C-bulge 9-mer, 5′-d(GATGGGCAG)·d(CTGCCCCATC), at 490 MHz in 10 mM phosphate. Assignments are shown in Table I.

resonances are accounted for and are still adequately resolved. One-dimensional NOE difference spectra show that the order of assignment has not changed from the perfect 9-mer (see Table II). The largest changes in chemical shift occur in base pairs 6–8. Resonance A (A·T 8) shifts from 14.00 to 14.16 ppm. F (G·C 6) and G (G·C 7) also more downfield by about 0.1 ppm. The other resonances change only slightly. E (G·C 5) is unaffected, and H (G·C 4) moves upfield 0.03 ppm. Since the base sequence has changed very little, the chemical shift

Table III: T_1 Relaxation Times of C-Bulge 9-mer (ms) in 10 mM Phosphate, pH 7.0

	base pair								
temp (K)	1	2	3	4	5	6	7	8	9
278	25	66	60	225	255	261	219	123	28
283	14	34	32	243	230	236	249	71	24
288	10	22	17	224	176	187	240	39	
293	6	7	7	130	159	132	140	18	
298				84	105	100	86	17	
303				32	55	47	41		
308				14	16	15	13		

Chart II



differences in base pairs 6–8 could indicate a local change in conformation.

The overall stability of the helix is lowered substantially by the presence of the bulge. The terminal G·C base pairs begin to broaden at 283 K, and the internal base pairs broaden by 308 K, compared to 323 K for the perfect helix. The most striking change is in resonances B (A·T 2) and C (T·A 3), which already begin to broaden at 283 K, while resonance A (A·T 8) is exchanging with the solvent much more slowly and does not disappear until 308 K. Note that C, which is the third base pair from the end, broadens much earlier than A, a penultimate base pair. This seems to indicate that the destabilizing effect of the bulge is felt most strongly by the two A·T base pairs on the left flank of the central G·C run.

This change is reflected even more clearly by the proton relaxation times. The T_1 relaxation times were measured between 278 and 308 K in 10 mM phosphate (Table III) and 20 mM Tris (not shown). As expected, the lifetimes of all the exchangeable protons are shorter than in the perfect 9-mer, due to a shortening of the overall helix lifetime. Again, the core G·C base pairs have roughly equivalent lifetimes until 298 K, where the adjacent base pairs (T·A 3 and A·T 8) are exchanging rapidly. At this point, the outer two imino protons (G·C 4 and C·G 7) seem to be affected by the end fraying and have slightly shorter T_1 values.

The destabilization of A·T 2 and T·A 3 is reflected in their relative T_1 values. While the lifetime of A·T 8 has only slightly decreased from the perfect 9-mer, the lifetimes of the other two A·T resonances (B and C) have sharply decreased and show a much stronger temperature dependence. Moreover, they are nearly equal to each other, despite the fact that one normally expects the imino protons nearer to the ends of the molecule to undergo more rapid exchange.

It is clear that in this molecule the bulge increases imino exchange rates on one end of the helix relative to the other. Besides helix-to-coil transitions and end fraying, one might expect opening of the internal region caused by the bulge to contribute to the observed lifetimes. In addition, these opening pathways may combine, especially at the more A·T-rich left end of the molecule, to produce fraying that runs as far as the bulge defect, exposing base pairs 1–3 to solvent exchange. This effect probably accounts for the more rapid exchange rates at the left end of the molecule compared to the right end.

The strongly increased exchange rates at the left boundary of the G·C tract suggest superficially that the bulge may be localized there. However, the nearly uniform increase in proton exchange rates within the G·C tract argues to the contrary, as discussed in detail below.

As one might expect, the temperature dependences of the relaxation rates are not straightforward in the case of the C-bulge 9-mer and are difficult to analyze quantitatively. In

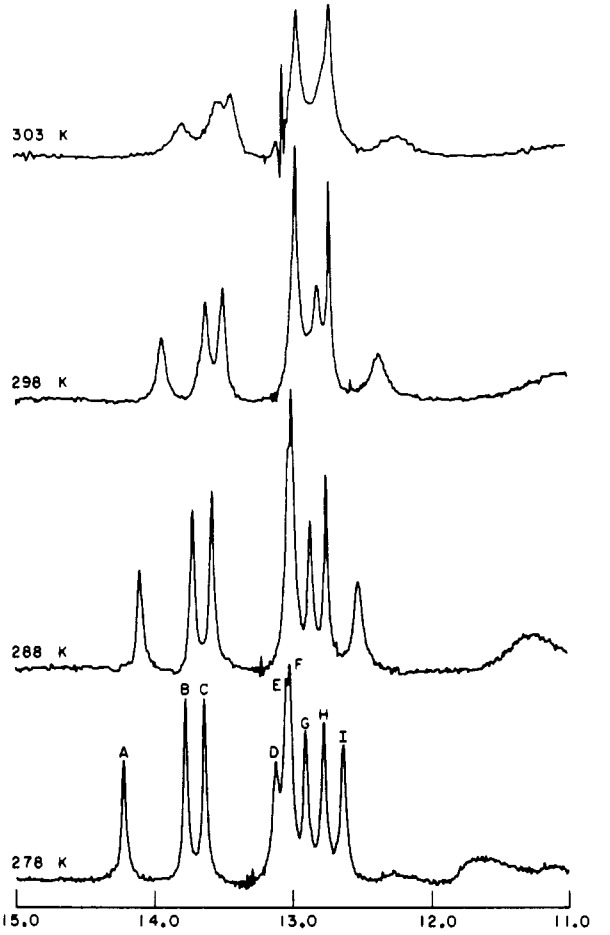


FIGURE 4: Imino proton spectra of A-bulge 9-mer, 5′-d-(GATGGGCAG)·d(CTGACCCATC), at 490 MHz in 10 mM phosphate. Assignments are shown in Table I.

general, the apparent ΔE_a of proton exchange is lower for the bulged molecule (see Table V and Figure 7). For G·C 5–7, ΔE_a is about 35 kcal/mol, higher than the value reported by Early et al. (1977) for terminal base pairs but slightly lower than the value usually given for internal base pairs, about 45–47 kcal/mol. G·C 4 has a much lower activation energy, 26 kcal/mol, which is equal to that of T·A 3. If one considers the left end of the molecule to be fraying more rapidly, it is reasonable that G·C 4 would also be affected.

A-Bulge 9-mer. In the C-bulge oligonucleotide, the position of the cytidine is not fixed but could be anywhere along the G-C run and still allow maximum base pairing (see Chart II). In order to compare these results with a molecule in which the position of the bulge is known, we prepared a third duplex containing an extra adenosine instead of an extra cytidine on the C-tract strand.

The imino spectrum shown in Figure 4 is similar to that of the C-bulge 9-mer. The order of assignment is unchanged except that resonance E is now G·C 6 and F is assigned to G·C 5. Similar to the C-bulge 9-mer, base pairs 6–8 are most

Table IV: T_1 Relaxation Times of A-Bulge 9-mer (ms) in 10 mM Phosphate, pH 7.0

	base pair								
temp (K)	1	2	3	4	5	6	7	8	9
278	49	208	220	224	183	124	160	117	62
283	21	162	210	287	168	69	125	82	
288	20	100	148	293	87	40	86	49	
293	11	52	79	297	88	32	61	24	
298	9	25	43	196	81	32	38	14	
303		12	21	63	56		25	7	
308			2	15	16	16	15		

Table V: Estimated Activation Energies for Proton Exchange (kcal/mol) in 10 mM Phosphate, pH 7.0, 100 mM Na^+, and 83.9 mM ClO_4^-

	base pair								
	1	2	3	4	5	6	7	8	9
perfect	22	28	30	50	50	43	44	26	
C bulge	15	28	26	26	36	35	34	23	
A bulge	16	24	23	49	20	12	16	22	

strongly shifted, which is not surprising in this case since one might expect those base pairs surrounding the bulge site to change the most. Resonance A (A·T 8) moves from 14.0 to 14.23 ppm, an even larger shift than in the C-bulge 9-mer. E (G·C 6) and G (G·C 7) both shift downfield by about 0.1 ppm, while again F (G·C 5) and H (G·C 4) do not change much.

The series of spectra in Figure 4 tells a much different story than that of either of the previous two oligonucleotides. At 278 K, the nine imino resonances are all present, although E and F nearly coincide. As the temperature increases, the terminal and penultimate base pairs broaden as expected but two of the four resonances assigned to the central G·C pairs, E and G, also begin to broaden and disappear, while the other two remain reasonably sharp. Checking the assignment, we see that these indeed correspond to the two base pairs flanking the bulged A. Further examination shows that, between the two penultimate A·T pairs, A·T 8, which is nearer the bulge, broadens earlier than A·T 2, and the terminal base pair G·C 9 seems to broaden before G·C 1, although it unfortunately merges with E. From this simple experiment alone it can be seen that in this case the bulged base does have a localized effect on helix stability.

T_1 relaxation times were measured between 278 and 308 K and are listed in Table IV. It was not possible to get good data on resonance D beyond 278 K, and it was difficult to obtain good data on resonances E and F by these methods due to poor chemical shift resolution. Again, all the lifetimes are somewhat shorter than in the perfect 9-mer. The most stable base pair H (G·C 4), however, has longer lifetimes (290 ms at 283–293 K) than the core resonances have in the C-bulge molecule. On the other hand, E (G·C 6) and G (C·G 7), which flank the bulged adenosine, have comparatively short lifetimes, 69 and 125 ms, respectively, at 283 K. F (G·C 5) is intermediate between E and H.

The activation energies were determined as before. Base pairs 1 and 9 have ΔE_a of 16 and 22 kcal/mol, in the range for single base pair opening. Base pairs 2 and 3 have activation energies that are somewhat higher. As one might expect, G·C 6 and C·G 7, which flank the bulged A, have very low activation energies (12 and 16 kcal/mol) for internal resonances. G·C 5 activation energy is somewhat higher (about 20 kcal/mol), and that for G·C 4 is 49 kcal/mol, in the range usually seen for internal base pairs of oligonucleotides. This conforms to the expectation that the base pairs will melt more easily around a bulge defect. In Figure 7B,C, the relaxation rates for G·C 4 and 7 in each duplex are compared. Note that the slope of G·C 4 is nearly the same in the A-bulge 9-mer as in the perfect 9-mer, whereas the slope of G·C 7 is much less in the A-bulge duplex than in either of the other two.

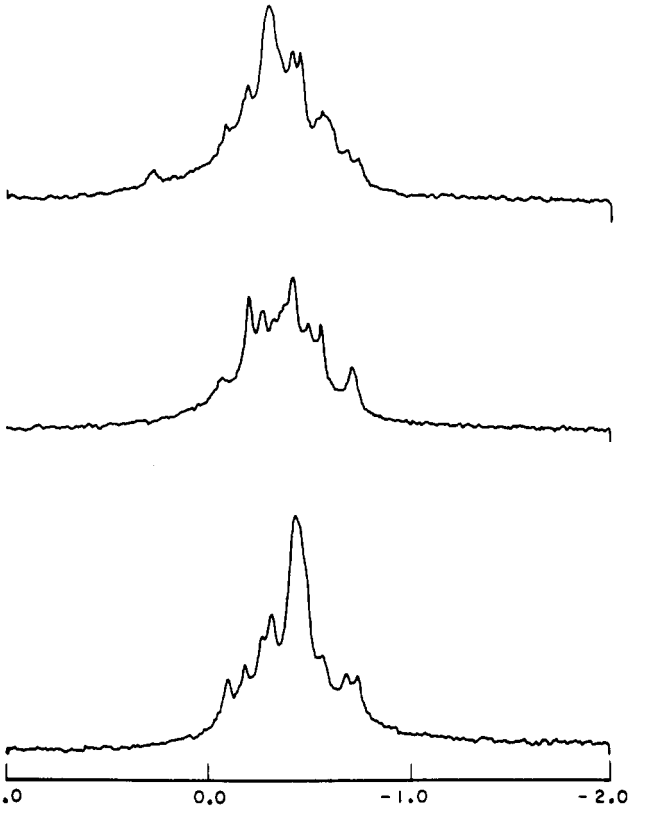


FIGURE 5: Phosphorus spectra at 200 MHz of A-bulge 9-mer (top), C-bulge 9-mer (middle), and perfect 9-mer (bottom). Samples were in 10 mM phosphate, 83.9 mM $NaClO_4$, 10 mM EDTA, and 10 mM EGTA at 293 K. Chemical shifts are relative to 85% phosphoric acid.

^{31}P NMR Spectroscopy. ^{31}P spectra at 200 MHz are shown in Figure 5. The spectra were acquired at 293 K in 10 mM phosphate, 50 mM NaCl, 10 mM EDTA, and 10 mM EGTA. Although the resonances are not well resolved, it is evident that the phosphorus chemical shifts change substantially from one oligonucleotide to the next, which suggests that the backbone conformation is altered significantly to accommodate the extra nucleotide. In the spectrum of the A-bulge 9-mer (top), one resonance has shifted upfield away from the others and may correspond to the bulged adenosine phosphate. It is interesting to note however, that the C-bulge spectrum (middle) does not have a similarly shifted resonance.

Optical Melting Data. The apparent thermodynamic parameters and T_m values for helix melting were obtained from

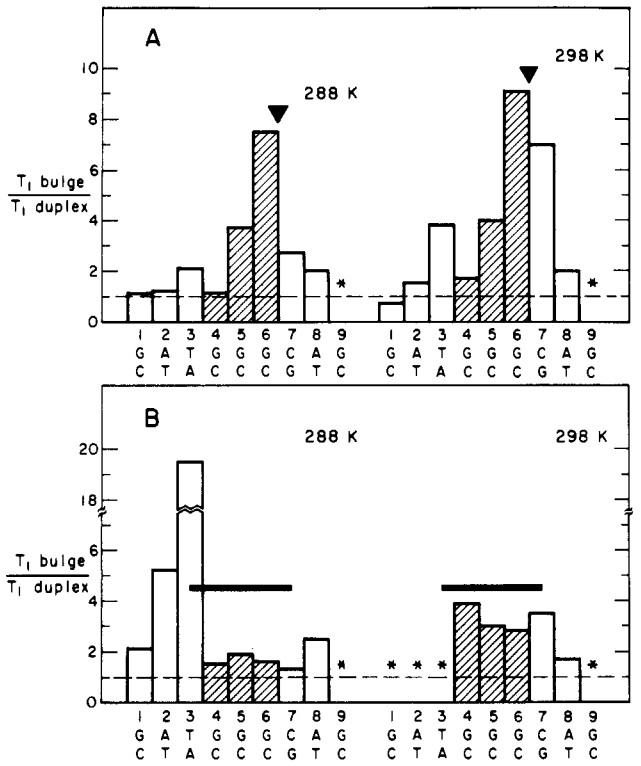


FIGURE 6: Ratios of T_1 values for bulge-containing duplexes to T_1 for the perfect 9-mer at 288 and 298 K, in 10 mM phosphate. The sequence of the perfect duplex is indicated beneath each graph. Asterisks denote resonances where T_1 values are very short. Hatched bars correspond to the G·C run. (A) A-bulge 9-mer. (▼) denotes the position of the bulged adenosine. (B) C-bulge 9-mer. The horizontal bar denotes the range of positions occupied by the bulged cytosine.

optical melting curves for each of the three duplexes. Data were fit to a simple two-state model (Petersheim & Turner, 1983). Values obtained for T_m, $\Delta H°$, and $\Delta S°$ are shown in Table VI; corrections to the two-state model (Gralla & Crothers, 1973) are likely to be substantial for the bulge-containing molecules due to enhanced fraying, so the apparent $\Delta H°$ and $\Delta S°$ values are certainly smaller than the total enthalpy and entropy of helix formation.

Table VI: Apparent Thermodynamic Parameters from Optical Melting Data

	concn (μM)	T_m (°C)	$\Delta H°$ (kcal)	$\Delta S°$ (cal/deg)
perfect	4.16	38.05	−69.96	−197
	11.52	41.53	−68.04	−191
	22.80	43.27	−70.32	−198
	100.20	46.25	−65.79	−185
			−69.43	−196
			± 1.23[a]	± 3[a]
C bulge	4.62	27.37	−55.64	−158
	10.14	29.21	−60.39	−174
	23.00	31.70	−58.60	−168
			−57.67	−165
			± 1.9[a]	± 7[a]
A bulge	4.18	21.53	−54.19	−157
	7.20	23.15	−51.07	−146
	13.90	25.81	−53.78	−155
			−53.02	−153
			± 1.4[a]	± 6[a]

[a] Mean ± SD.

The T_m of the perfect 9-mer is 41 °C at 10 μM strand concentration, with an average enthalpy of helix formation of 69 kcal/mol and an average entropy of 195 cal/(deg·mol). The enthalpy agrees roughly with base stacking energies reported by Freier et al. (1983). As expected, a bulge nucleotide destabilizes the helix considerably; however, the T_m of the A-bulge 9-mer at 23 °C is much lower than that of the C-bulge 9-mer, which is 29 °C.

Discussion

Bulge Migration? A key question in the interpretation of our results is the location of the extra nucleotide in the C-bulge 9-mer. In a homogeneous run an extra base can fit in anywhere, so that all of the conformations in Chart II are possible. The extra base could also migrate along the run and occupy all of the positions with some frequency. Since the imino resonances are narrow and stoichiometric, the molecule must either be in one stable conformation or be in the fast-exchange limit for bulge migration.

The data from the A-bulge 9-mer indicate that one can expect a bulge to exert a local destabilizing effect. As shown in Figure 6A, the ratio of T_1 relaxation times for the bulge-containing and normal duplex molecules shows a pronounced

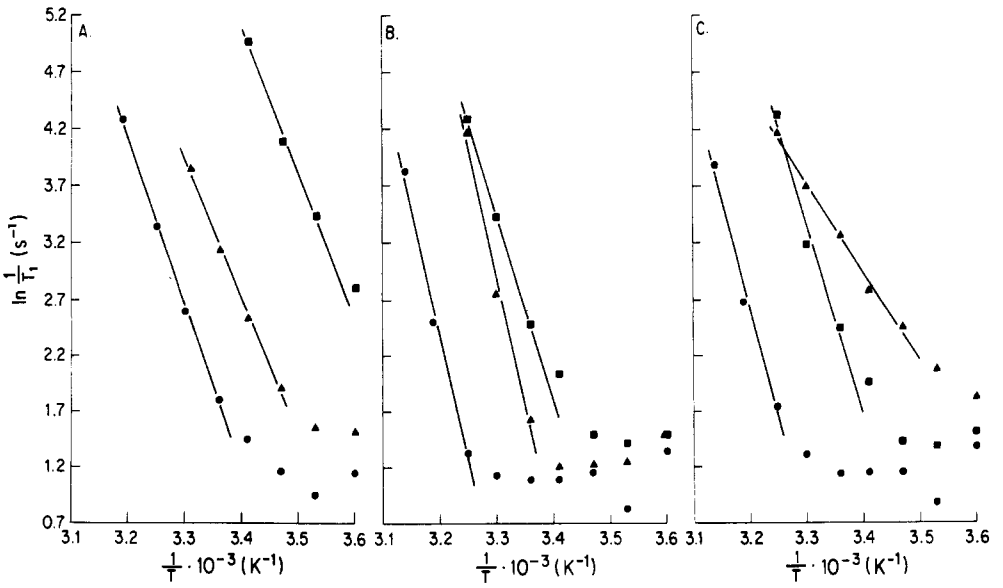


FIGURE 7: Comparison of the logarithm of relaxation rates (s^{-1}) vs. inverse temperature (K^{-1}) for base pairs A·T 3 (A), G·C 4 (B), and G·C 7 (C) in each duplex: (●) perfect 9-mer, (▲) A-bulge 9-mer, and (■) C-bulge 9-mer. Vertical axes are the same. Points at higher temperatures were fit to a straight line as shown, where the slope $\approx \Delta E_a/R$ for proton exchange.

peak for the base pairs immediately 3′ and 5′ to the bulged nucleotide, with a somewhat greater fractional effect on T_1 for the 3′ base pair. Base pair 4, located at the far end of the G·C tract from the bulge between base pairs 6 and 7, shows only a small change in relaxation time due to the bulge, supporting the view that the perturbation effect of the bulge is localized.

In contrast, the C bulge shows a marked absence of a localized perturbation within the G·C tract. The similarity of the relative lifetimes of the four core G·C imino protons (Figure 6B) argues strongly that the observed resonances represent an average of all the positions shown in Chart II. Moreover, the destabilization of the left end of the sequence is readily understood if one takes into account the fact that the A·T base pairs are less able to overcome the destabilizing effect of the bulge and maintain base pairing, whereas the C·G base pair at the right end "contains" the effect of the bulge. The intermediate values for ΔE_a support the idea that the observed resonances are averages (in or near the fast-exchange limit) of conformers with the bulge in different locations. Moreover, the nonlinearity of the temperature dependence suggests that several mechanisms, including bulge migration, are contributing to the observed relaxation times.

Free Energy of the Bulge Defect. The optical melting transitions measured on our series of bulge-containing compounds allow us to estimate the destabilizing free energy of the bulge defect, a parameter analogous to that determined for RNA duplexes some years ago (Fink & Crothers, 1972). We estimate the T_m shift to be $\delta T_m = -16.5$ °C for the A bulge and −10.8 °C for the C bulge. The corresponding free energy change is obtained from

$$\delta(\Delta G^\circ) = -\Delta S^\circ \delta T_m$$

in which ΔS°, assumed to be independent of temperature, is set equal to −196 cal K^{-1}, obtained from analysis of the optical melting curves for the perfect duplex (Table VI). The result is an estimated destabilizing free energy of 3.2 kcal/mol for the A bulge, in close agreement with the estimate of 2.8 kcal/mol for bulged A residues in RNA double helices (Fink & Crothers, 1972).

The C bulge, at 2.1 kcal/mol, is found to be less destabilizing than the A bulge, but one can rationalize most of the difference on the basis of the positional degeneracy of the bulge within the C tract, which increases the entropy and therefore the stability of the bulged helix. Specifically, if the positional degeneracy is 4, the total destabilizing free energy is given by

$$\delta(\Delta G^\circ) = \delta(\Delta G^\circ)[\text{isolated bulge}] - RT \ln 4$$

from which we conclude that the free energy of an isolated C bulge is 2.9 kcal/mol, in close agreement with the result obtained for the isolated A bulge.

Conclusions

From these experiments it is not possible to speculate on the detailed mechanism of the postulated migration of the bulge defect; however, it is interesting to note that the rate of migration is rapid on the NMR time scale and faster than the rate of solvent exchange of the protons involved, even in the presence of Tris buffer. Unfortunately, we cannot directly measure the rate of bulge migration by these experiments, but the fact that no additional broadening of the imino resonances is observed in the C-bulge oligomer places a lower limit on this exchange process. (Imino resonances are already considerably broadened by solvent exchange.) In the fast-exchange limit, the line width $\approx \Delta\omega^2/\omega_{ex}$. Even for a small chemical-shift difference (10–25 Hz) this implies an exchange rate of 100–1000 s^{-1}, about an order of magnitude faster than solvent exchange at temperatures less than 308 K.

Moreover, neither rapid migration of the bulge nor imino proton exchange can be accounted for by a mechanism involving an accelerated helix–coil or melting transition in the less stable bulged duplex. An estimate of helix-opening rates can be obtained from the optical data. For a non-self-complementary duplex DS consisting of strands A and B

$$A + B \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} DS$$

Assuming a value of $\sim 10^6$ M^{-1} s^{-1} for k_{12} (Pörschke & Eigen, 1971; Craig et al., 1971) and a sample concentration of 2 mM, at T_m

$$k_{21} \approx (10^6 \text{ M}^{-1} \text{ s}^{-1})(A_0/2) \approx 1.0 \times 10^3 \text{ s}^{-1}$$

We assume that $\Delta E^\dagger_{21} \approx \Delta H_{helix}$ (Pörschke & Eigen, 1971; Craig et al., 1971) and use the relationship

$$k_{21}(T)/k_{21}(T_m) = e^{(-\Delta E^\dagger_{21}/R)(1/T - 1/T_m)}$$

For the C-bulge 9-mer, $\Delta H = \Delta E^\dagger_{21} = 57.7$ kcal/mol and the T_m is 47 °C at 2.0 mM, so at 288 K, $k_{21} \approx 0.04$ s^{-1}, and at 278 K, $k_{21} \approx 0.001$ s^{-1}. These helix-opening rates are far too slow to account for the observed changes in relaxation rates or the averaging of resonances at lower temperatures. Furthermore, the high activation energy implied by this mechanism would predict a much stronger temperature dependence than observed.

If the rate of bulge migration is faster than solvent exchange, then the imino protons must be protected from attack by solvent molecules during the rapid breaking and reforming of the hydrogen bonds. One can imagine arrangements where one strand is stretched a little and the other is compressed slightly to accommodate the extra base, which remains stacked in the helix. The two strands could shift rapidly with respect to each other, forming and reforming base pairs in different combinations. Two-dimensional NMR studies on the non-exchangeable protons are in progress and should help resolve these questions.

These studies demonstrate that an extra bulged base behaves much differently in a heterogeneous sequence than in a homogeneous G·C run. Rapid isomerization of the bulge site delocalizes its destabilizing effect so that four or more base pairs are weakened, but none are completely disrupted. Furthermore, rapid migration of the bulge in a homopolymer run allows it to be formed at one position and to rapidly exert its frame-shifting influence some distance away, for example, at the site of copying by RNA or DNA polymerase. It is also possible that the bulge defect could migrate rapidly across a codon–anticodon double helix in a homopolymeric mRNA sequence, thus promoting both negative and positive translational frame shifting.

On the other hand, a fixed bulge appears to have a strongly localized effect, at least in these sequences, which is severe at the site of the bulge but is rather abruptly attenuated. Certainly, in molecules of this size, delocalization of the bulge results in less overall helix destabilization. Clearly the idiosyncrasies of local DNA sequences cannot be ignored if phenomena such as frame-shift mutagenesis are to be understood on a molecular level. It is our hope that a continuation of these studies will contribute to the basis of this understanding.

Acknowledgments

Our thanks to S. D. Levene and J. Nadeau for their help with the computer software and to Franklin Hutchinson for

bringing this problem to our attention and communicating results in advance of publication.

Registry No. d(GATGGGCAG)·d(CTGCCCATC), 105784-72-3; d(GATGGGCAG)·d(CTGCCCCATC), 105784-74-5; d(GATGGGCAG)·d(CTGACCCATC), 105784-76-7.

REFERENCES

Ames, B., Lee, F. D., & Durston, W. E. (1973) *Proc. Natl. Acad. Sci. U.S.A. 70*, 782–786.

Calos, M. P., & Miller, J. H. (1981) *J. Mol. Biol. 153*, 39–66.

Craig, M. E., Crothers, D. M., & Doty, P. (1971) *J. Mol. Biol. 62*, 383–401.

Crothers, D. M., Cole, P. E., Hilbers, C. W., & Shulman, R. G. (1974) *J. Mol. Biol. 87*, 63–88.

Drake, J. W. (1964) *J. Cell. Comp. Physiol. 64* (*Suppl. 1*), 19–32.

Early, T. A., Kearns, D. R., Hillen, W., & Wells, R. D. (1977) *Biochemistry 16*, 541–551.

Ferguson, L. R., Denny, W. A., & MacPhee, D. G. (1985) *Mutat. Res. 157*, 29–37.

Fink, T. R., & Crothers, D. M (1972) *J. Mol. Biol. 66*, 1–12.

Freier, S. M., Burger, B. J., Alkema, D., Neilson, T., & Turner, D. H. (1983) *Biochemistry 22*, 6198–6206.

Gralla, J., & Crothers, D. M. (1973) *J. Mol. Biol. 78*, 301–319.

Haasnoot, C. A. G., den Hartog, J. H. J., de Rooij, J. F. M., van Boom, J H., & Altona, C. (1980) *Nucleic Acids Res. 8*, 169–181.

Janik, B. (1971) *Physicochemical Characteristics of Oligonucleotides and Polynucleotides*, Plenum, New York.

Kime, M. J., & Moore, P. B. (1983) *Biochemistry 22*, 2615–2622.

Lee, C.-H., & Tinoco, I., Jr. (1978) *Nature* (*London*) *274*, 609–610.

Lerman, L. S. (1963) *Proc. Natl. Acad. Sci. U.S.A. 49*, 94–102.

Lerman, L. S. (1964) *J. Cell. Comp. Physiol. 64* (*Suppl. 1*), 1–8.

Leroy, J.-L., Broseta, D., & Gueron, M. (1985) *J. Mol. Biol. 184*, 165–178.

Mizusawa, H., Lee, C.-H., Kakefuda, T., McKenny, K., Shimatake, H., & Rosenberg, M. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 6817–6820.

Morden, K. M., Chu, Y. G., Martin, F. H., & Tinoco, I. (1983) *Biochemistry 22*, 5557–5563.

Nelson, J. W., & Tinoco, I. (1985) *Biochemistry 24*, 6416–6421.

Okada, Y., Terzaghi, E., Streisinger, G., Emrich, J., Inouye, M., & Tsugita, A. (1966) *Proc. Natl. Acad. Sci. U.S.A. 56*, 1692–1698.

Okada, Y., Streisinger, G., Owen, J., Newton, J., Tsugita, A., & Inouye, M. (1972) *Nature* (*London*) *236*, 338–341.

Pardi, A., & Tinoco, I. (1982) *Biochemistry 21*, 4686–4693.

Pardi, A., Morden, K. M., Patel, D., & Tinoco, I. (1982) *Biochemistry 21*, 6567–6574.

Pardi, A., Morden, K. M., Patel, D. J., & Tinoco, I. (1983) *Biochemistry 22*, 1107–1113.

Patel, D. J. (1974) *Biochemistry 13*, 2388–2395.

Patel, D. J., & Shapiro, L. (1985) *Biochimie 67*, 887–915.

Patel, D., Kozlowski, S. A., Marky, L. A., Broka, C., Rice, J. A., Itakura, K., & Breslauer, K. J. (1982) *Biochemistry 21*, 445–451.

Patel, D., Kozlowski, S. A., & Bhatt, R. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3908–3912.

Patel, D. J., Shapiro, L., Kozlowski, S. A., Gaffney, B. L., Kuzmich, S., & Jones, R. J. (1985) *Biochimie 67*, 861–886.

Pörschke, D., & Eigen, M. (1971) *J. Mol. Biol. 62*, 361–381.

Quignard, E., Teiole, R., Guy, A., & Fazakerley, G. V. (1985) *Nucleic Acids Res. 13*, 7829–7836.

Roth, K., Kimber, B. J., & Feeney, J. (1980) *J. Magn. Reson. 41*, 302–309.

Roy, S., & Redfield, A. G. (1981) *Nucleic Acids Res. 9*, 7073–7083.

Skopek, T. R., & Hutchinson, F. (1984) *MGG, Mol. Gen. Genet. 195*, 418–423.

Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Terzaghi, E., & Inouye, M. (1966) *Cold Spring Harbor Symp. Quant. Biol. 31*, 77–84.

Young, P. R., & Kallenbach, N. R. (1981) *J. Mol. Biol. 145*, 785–813.