Halpin, R. R., Hegeman, G. D., & Kenyon, G. L. (1981) *Biochemistry 20*, 1525.
Harris, R. K., & Newman, R. H. (1976) *J. Magn. Reson. 24*, 449.
Jeener, J. (1971) Ampere International Summer School, Basko Polje, Yugoslavia.
Krepinsky, J., Carver, J. P., Rajalakshmi, S., Rao, P. M., & Sarma, D. S. R. (1979) *Chem.-Biol. Interact. 27*, 381.
Kuhlmann, K. F., & Grant, D. M. (1968) *J. Am. Chem. Soc. 90*, 7355.
Marker, A. F. H., & Whittingham, C. P. (1967) *J. Exp. Bot. 18*, 732.
Oliver, D. J. (1979) *Plant Physiol. 64*, 1048.
Prestegard, J. H., Koerner, T. A. W., Demov, P. C., & Yu, R. K. (1982) *J. Am. Chem. Soc. 104*, 4993.
Rabson, R., Tolbert, N. E., & Kearney, P. C. (1962) *Arch. Biochem. Biophys. 98*, 154.
Schaefer, J., Stejskal, E. O., & Beard, C. F. (1975) *Plant Physiol. 55*, 1048.
Skokut, T. A., Varner, J. E., Schaefer, J., Stejskal, E. O., & McKay, R. A. (1982) *Plant Physiol. 69*, 314.
Sogn, J. A., Craig, L. C., & Gibbons, W. A. (1974) *J. Am. Chem. Soc. 96*, 4694.
Solomon, I. (1955) *Phys. Rev. 99*, 559.
Thevelein, J. M., den Hollander, J. A., & Shulman, R. G. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 3503.
Traynor, P. L., & Flashman, S. M. (1981) *Theor. Appl. Genet. 60*, 31.
Ts'o, P. O. P., & Kan, L. S. (1979) in *Chromatin Structure and Function* (Nicolini, C. A., Ed.) Part A, p 217, Plenum Press, New York.
Van Divender, J. M., & Hutton, W. C. (1982) *J. Magn. Reson. 48*, 272.
Wang, D., & Burris, R. H. (1963a) *Plant Physiol. 38*, 430.
Wang, D., & Waygood, E. R. (1963b) *Plant Physiol. 37*, 826.
Woo, K. C. (1979) *Plant Physiol. 63*, 783.
Woo, K. C., & Osmond, C. B. (1976) *Aust. J. Plant Physiol. 3*, 771.

Sequence-Specific Recognition of DNA: NMR Studies of the Imino Protons of a Synthetic RNA Polymerase Promoter†

Shan-Ho Chou, David E. Wemmer, Dennis R. Hare, and Brian R. Reid*

ABSTRACT: We have synthesized both strands of a DNA duplex containing the consensus Pribnow promoter sequence TATAATG, flanked by GC base pairs to stabilize the ends of the helix. The stability of this duplex has been studied by using ^1H nuclear magnetic resonance. The imino protons have been assigned by using the sequential nuclear Overhauser effect approach. Exchange rates have been monitored by using selective inversion recovery measurements. The helix is relatively unstable in the center of the AT-rich region even when surrounded by GC base pairs, and there is considerable asymmetry in the melting of the helix.

We have begun a systematic study of the structural and dynamic characteristics of short DNA duplexes containing operator and promoter sequences that function in the regulation of DNA expression (Chou et al., 1983). Promoters are short regulatory DNA sequences, located upstream of structural genes, that function in transcriptional control of DNA by binding RNA polymerase such that RNA synthesis is correctly initiated. The promoter sequence most proximal to the site of initiation is the well-known Pribnow sequence, or "TATA" box, located upstream at about position −10 from the first transcribed residue (Rosenberg & Court, 1979; Siebenlist et al., 1980; Hawley & McClure, 1983). Several different promoter sequences occur naturally, and these variations probably govern the efficiency of transcription at a given gene locus. We have chosen to first examine the prototype or consensus sequence

CG*T*T*A*T*A*A*T*GCG
GC*A*A*T*A*T*T*A*CGC

in which the Pribnow octamer is embedded in two GC pairs at each end of the helix. The boldface letters represent base pairs that are strongly conserved in promoters whereas the italic letters are only moderately conserved (Rosenberg & Court, 1979). The central ten base pairs of this sequence constitute the actual promoter sequence of the *spc* ribosomal protein operon (Post et al., 1978), and the central eight base pairs, including those that are highly conserved, are the specific sequence for a promoter in the SV40 virus (Dhar et al., 1974). The remaining letters thus represent structural support and were introduced into the sequence to reduce the effects of end fraying rather than being in the promoter sequence per se. We will henceforth refer to this sequence as the consensus Pribnow sequence. In this paper, we describe the assignment of the hydrogen-bonded imino proton spectrum and discuss the dynamics of this sequence as probed by imino proton exchange. In the accompanying paper (Wemmer et al., 1984) we assign most of the nonexchangeable protons and discuss the structural features of this dodecamer.

NMR[1] spectroscopy is one of the few methods that can provide both structural and dynamic information on molecules in solution and has seen increasing application in the study of nucleic acids in recent years. Imino proton assignment and dynamic studies on several tRNAs and small DNA sequences have been carried out including two operator DNA sequences (Johnston & Redfield, 1981; Hare & Reid, 1982a,b; Early

† From the Department of Chemistry and the Department of Biochemistry, University of Washington, Seattle, Washington 98195. *Received July 25, 1983.* We gratefully acknowledge NIH Grant R01 GM32302 and instrumentation grants from the Murdock Foundation, the National Institutes of Health (GM2874-01S1), and the National Science Foundation (PCM80-18053).

[1] Abbreviations: DMT, dimethoxytrityl; NMR, nuclear magnetic resonance; DSS, 4,4-dimethyl-4-silapentane-1-sulfonate; NOE, nuclear Overhauser effect; FID, free-induction decay.

0006-2960/84/0423-2257$01.50/0

et al., 1981a,b; Chou et al., 1983; Zuiderweg et al., 1981; Patel et al., 1982c). Methods for almost complete assignment of nonexchangable protons have also been reported (Scheek et al., 1983; Hare et al., 1983) and are described for the Pribnow sequence in the accompanying paper.

The development of solid-phase triester methods for DNA synthesis (Matteucci & Caruthers, 1981; Beaucage & Caruthers, 1981; Chow et al., 1981; Tanaka & Letsinger, 1982) has made possible the synthesis of sufficient quantities of pure 10–20 base pair oligomers for systematic NMR studies of biologically relevant sequences. Once the NMR spectra of such DNA molecules have been assigned and correlated with the structural features of the molecule, then complexes of these DNA sequences with their conjugate proteins can be studied and interpreted.

Materials and Methods

DNA Synthesis. Deoxyadenosine, deoxyguanosine, deoxycytidine, and thymidine were purchased from Biosearch (San Rafael, CA). N-Blocked 5′-DMT deoxynucleosides were prepared according to Ti et al. (1982). 3′-Phosphoramidites were prepared according to Beaucage & Caruthers (1981). ^{31}P NMR was used to verify that the blocked 5′-DMT 3′-phosphoramidites were at least 95% pure. The appropriate blocked nucleoside was coupled to silica gel beads according to Chow et al. (1981). Prior to use in chain extension, the phosphoramidites were activated by the addition of tetrazole amidite activator (Biosearch, San Rafael, CA). The general procedure used for synthesizing and purifying DNA strands has been described previously (Chou et al., 1983; Hare et al., 1983). After deprotection and purification, the d(CGTTATAATGCG) and d(CGCATTATAACG) strands were combined in equal amounts, heated briefly to 70 °C, and allowed to cool slowly to room temperature.

NMR Spectroscopy. The promoter DNA duplex (ca. 10 mg) was dissolved in 0.4 mL of 20 mM sodium phosphate buffer, pH 7.0, and transferred to a 5-mm NMR tube. For spectra in 100 mM NaCl and 5 mM $MgCl_2$, appropriate amounts of concentrated solutions of these salts in H_2O were added. Spectra were obtained on a Bruker WM 500 spectrometer. In H_2O solution, a modified Redfield 21412 pulse sequence was used with the carrier frequency placed near 12 ppm. DSS (ca. 0.1 mM) was added to all samples as an internal chemical shift reference. NOE difference spectra were collected directly as difference FIDs between 0.5 s of on- and off-resonance irradiation, alternating eight scans on-resonance and eight scans off-resonance (Hare & Reid, 1982a,b). Selective inversion of imino resonances was accomplished by using the decoupler frequency source to generate pulses centered in the imino region. By adjusting the length and strength of the pulse, a "long pulse" null at the water resonance frequency could be obtained, while achieving near inversion of the imino proton resonances without significantly perturbing the aromatic proton region. The degree of inversion deteriorates for the imino protons furthest away from the carrier frequency and also decreases as the imino proton exchange rates increase, due to relaxation during the inverting pulse. However, the data were fitted with a three-parameter expression (Sass & Zeissow, 1977), with the initial degree of inversion as one of the parameters, so that reasonably accurate fits are still obtained at all temperatures.

Results and Discussion

The spectrum of the imino and aromatic proton regions of the Pribnow sequence in H_2O solution is shown in Figure 1. In the imino region at 0 °C, there are ten sharp and two somewhat broad resonances, accounting for the expected 12 imino protons of this dodecamer. The two broadened peaks disappear rapidly with increasing temperature, as shown in Figure 2. From the identity of the 10 base pairs remaining at 37 °C (vide infra), these two labile resonances can be assigned to the terminal base pairs whose imino protons exchange rapidly with water due to fraying of the helix. To assign the remaining resonances to particular base pairs, we have used the sequential NOE connectivity technique (Roy & Redfield, 1981; Hare & Reid 1982a,b; Chou et al., 1983; Kan et al., 1982). From the intra-base pair NOE to aromatic CH or amino protons, the base-pair type can be determined, and from the inter-base pair imino–imino NOEs, the two neighbors can be ascertained, leading to complete sequential assignments. In some cases where near degeneracies in chemical shift occur, second-order NOEs can be observed for long saturation times, which reveal next nearest neighbors, and NOEs to and from the adenine C2-H protons of neighboring bases can also aid to the assignment procedure (Chou et al., 1983). In Figure 1 the imino resonances are labeled A–J from the downfield end of the spectrum. The imino–imino NOEs, as shown in Figure 3, establish the nearest-neighbor connectivity sequence of peaks G–E–B–D–J–H. Similarly, peaks I–A–C–F are also sequentially linked (data not shown). Peaks A, C, F, G, E, B, and D show strong C2-H NOEs in the aromatic region, indicating that they are AT base pairs; peaks H, J, and I are GC imino protons. Given the sequence CGTTATAATGCG and the fact that the imino protons of

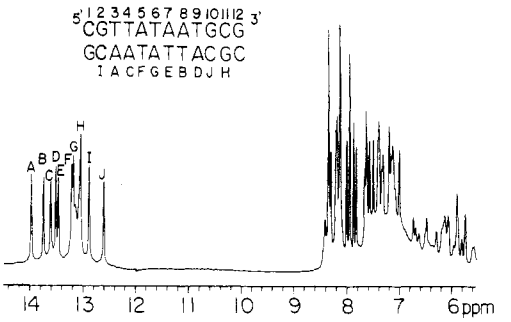

FIGURE 1: Spectrum (500 MHz) of the imino and aromatic proton regions of the consensus Pribnow dodecamer in 90% H_2O/10% D_2O solution in 20 mM phosphate buffer, pH 7.0 at 0 °C. The spectrum is the average of 800 scans with a modified Redfield 214 pulse, with a delay of 1.34 s between pulses. The DNA sequences is shown with the peaks lettered for reference to the text. The assignments of peaks to particular base pairs are shown below the sequence and are derived from NOE measurements discussed in detail in the text.

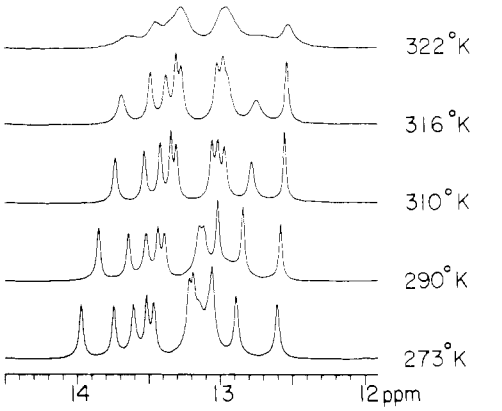

FIGURE 2: Temperature dependence of the imino proton region of the NMR spectrum of d(CGTTATAATGCG). Peaks are referenced to internal DSS at each tempeature. Solution conditions were as in Figure 1, except that the buffer contains 100 mM NaCl and 5 mM $MgCl_2$.

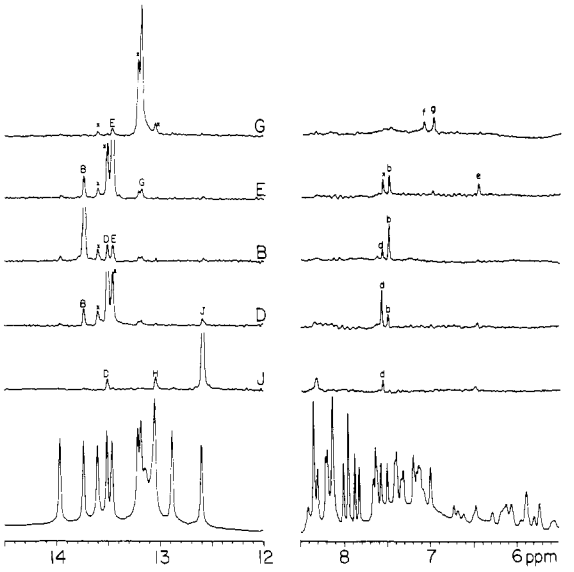

FIGURE 3: Examples of selected NOEs among imino and aromatic protons upon selective saturation of individual imino protons. Large capital letters indicate the irradiated imino proton, small capital letters indicate the neighboring imino protons, and lower case letters indicate the adenine C2-H proton in the same base pair as the correspondingly lettered imino proton; an × indicates "spill over" irradiation of an imino peak neighboring the irradiated resonance. The imino resonances were saturated for 0.5 s before the observation pulse, and all other conditions were as in Figure 1.

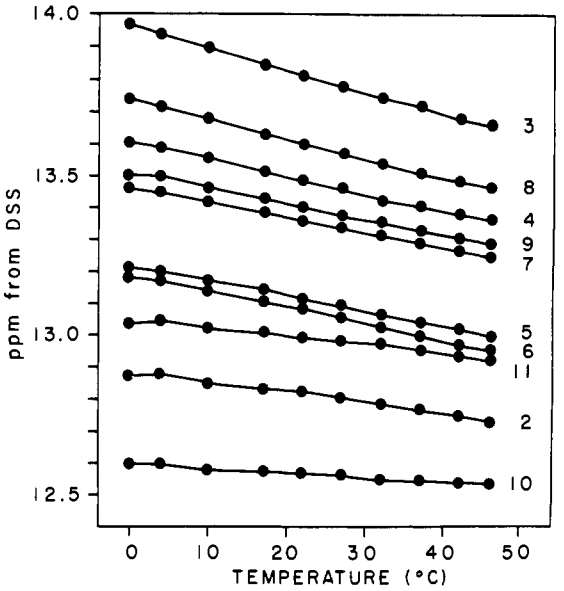

FIGURE 4: Temperature dependence of the chemical shifts of the imino protons of the Pribnow dodecamer. Representative spectra are shown in Figure 2. Identical plots were obtained in the absence of 100 mM NaCl and 5 mM $MgCl_2$.

CG 1 and GC 12 are exchanging too rapidly to show NOEs, the demonstration that GC resonances J and H belong to neighboring base pairs assigns them to GC 10 and CG 11, respectively, and peak I is thus GC 2. Hence, F and G must be neighbors in the helix interior but, because they are extremely close in chemical shift, the direct NOE from G to F cannot easily be observed (see Figure 3). To confirm the F–G nearest-neighbor relationship, we extended the irradiation time in order to observe second-order NOEs. As expected, at longer times peak C shows a weak cross saturation of G, as does E to F. Additional confirmation was obtained by irradiating the two C2-H protons, f and g, which were affected by saturation of peaks F and G. Since they are better resolved than the imino peaks (cf. Figure 3), they could be selectively irradiated, and again, as expected, each C2-H peak cross saturated both of the imino peaks F and G (data not shown). With the F–G connection established, the assignments are complete and are shown below the sequence in Figure 1. This AT-rich sequence shows an unusually wide dispersion in the shifts of the adenosine C2-H protons, probably due to the large ring current shifts induced by the adenine rings. As can be seen in the reference spectrum in Figure 1, the amplitude of the most upfield of the aromatic peaks is significantly attenuated by the decrease in nutation angle for peaks near the water resonance. For this reason, the NOEs to neighboring C2-H protons appear to be as large as the intra-base-pair C2-H NOE, though the latter are actually much larger when corrected for the shape of the observation pulse.

The imino proton spectra at several temperatures are shown in Figure 2, with the temperature dependence of the imino proton chemical shifts shown graphically in Figure 4. There is significant upfield shifting of all AT imino protons with increasing temperature. Similar effects have been observed for the AT protons in the self-complementary sequences CGCGAATTCGCG and CGCGTATACGCG (Patel et al., 1982a,b). The magnitudes of the shifts in the Pribnow sequence are slightly larger over the temperature range 0–45 °C than those in the sequences mentioned above. The sequences studied by Patel et al. could be monitored over a greater temperature range due to greater helix stability and showed continued shifting up to the helix melting temperature. In those sequences very little shifting of the neighboring GC resonances was observed. The relatively GC-rich sequence CCGCACTGATGG has also been studied (Early et al., 1981a), and in this case the AT pairs exhibit shifting with increasing temperature but not to the extent of the other sequences mentioned above; the GC pairs in this restriction fragment show virtually no shift change with temperature, even for the base pairs bordering the AT pairs. In the Pribnow sequence there is significant shifting of GC 2, which borders the seven base-pair AT segment, but less shifting is observed for GC 10 and CG 11 at the other end of the helix. The largest shifts are observed for AT pairs 3 and 8, with the smallest for AT pairs 5 and 6. The cause of these shifts has not been determined, although it has been suggested that propeller twisting of the bases could lead to such shifts (Patel et al., 1982a). This seems unlikely because those AT pairs with the smallest upfield shifts from their intrinsic position show the largest temperature dependence. By comparison with CCGCACTGATGG, it appears that such shifts are larger in AT-rich segments than in isolated AT pairs, and it appears that helix instability and shifting are related.

There have been many measurements of the exchange rates of the imino protons in both DNA and RNA (Johnston & Redfield, 1981; Reid & Hare, 1982; Hurd & Reid, 1980; Pardi et al., 1982; Patel et al., 1982a,b). It has been shown that, in the hydrogen-bonded double-strand form, the exchange of the imino protons with solvent water is slow, and that some sort of helix opening must first occur before exchange can take place (Teitelbaum & Englander, 1975). If the rate of exchange from the open state is rapid compared to helix reclosing, then measurement of the exchange rate reflects the rate of opening of the helix at that base pair. Measurements of exchange rates for individual protons in various RNA and DNA helices have shown that the exchange at any particular base is not necessarily correlated with that of its immediate neighbors, indicating that the opening event often appears to be a very local perturbation of the structure. Such effects might be expected to be sequence dependent, although it has

Table I: Observed Inversion Recovery Rates (s^{-1})

temp (°C)	base-pair number[a] 2	3	4	5	6	7	8	9	10	11
0	8.1 (0)	5.7 (0)	5.7 (0)	6.2 (0)	5.6 (0)	5.7 (0)	5.7 (0)	5.2 (0)	7.8 (0)	12.8 (0)
17	7.3 (2)	5.1 (1)	4.0 (0)	4.9 (0)	5.2 (1)	3.8 (0)	3.4 (0)	3.6 (0)	4.4 (0)	7.0 (0)
22	9.1 (4)	6.7 (3)	4.6 (1)	6.1 (2)	5.7 (2)	4.2 (1)	3.9 (0)	3.7 (0)	4.4 (0)	7.9 (2)
27	14.5 (10)	10.1 (7)	6.5 (3)	7.1 (4)	7.0 (4)	5.5 (2)	4.4 (1)	4.2 (1)	4.3 (0)	11.2 (6)
32	26.3 (22)	16.7 (14)	9.1 (6)	9.7 (7)	10.6 (8)	7.1 (4)	6.4 (4)	5.4 (3)	4.9 (1)	18.5 (14)
37	45.4 (42)	28.6 (26)	14.9 (12)	13.5 (11)	14.1 (12)	11.1 (9)	8.5 (6)	7.9 (6)	7.2 (4)	29.4 (25)
42	91.7 (88)	53.8 (52)	29.4 (27)	24.4 (22)	26.3 (24)	20.0 (17)	15.6 (13)	14.3 (12)	13.0 (9)	51.3 (47)

[a] The numbers in parentheses are the approximate helix opening rates obtained by subtracting the estimated magnetic contribution to the relaxation.

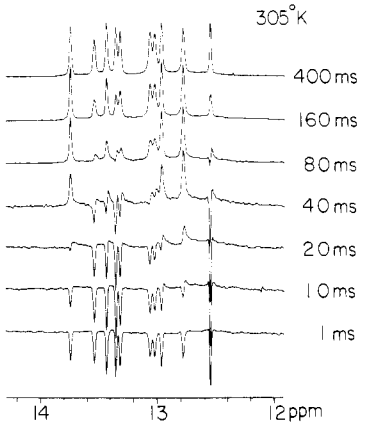

FIGURE 5: Representative selective inversion recovery spectra of d(CGTTATAATGCG). The time listed on each spectrum is the delay between the inverting and probing pulses. The recycle delay between a probing pulse and the next inverting pulse was 1.04 s except above 310 K, where it could be reduced to 0.59 s due to rapid recovery.

been suggested that AT and GC pairs undergo opening with characteristic rates that are different from each other, but are relatively independent of the sequence in which they occur (Early et al., 1981a,b).

At low temperatures, the experimentally observed rate of recovery from a saturation or inversion pulse is primarily determined by magnetic relaxation rather than chemical exchange. Since the GC imino protons have a larger number of close neighboring protons, they recover faster in the spin-diffusion limit than do AT pairs (the GC recovery rate is predicted to exceed the AT rate by a factor of 2 in the spin-diffusion limit; Early et al., 1981a).

An example of the recovery data for the Pribnow sequence after selective inversion of the imino protons is presented in Figure 5, and the measured recovery rates at several temperatures are given in Table I. The imino protons of base pairs 1 and 12 exchange too rapidly for analysis, but it is obvious from Table I that the relaxation rates at 0 °C for the three GC pairs in positions 2, 10, and 11 are 1.5–2 times faster than those for the AT pairs. Furthermore, the seven consecutive AT pairs in positions 3–9 have recovery rates of 5.5 ± 0.5 s^{-1}, which are identical within experimental error. That these rates reflect relatively little exchange contribution and are predominantly magnetic cross-relaxation rates can be seen from the fact that they all *decrease* upon raising the temperature to 17 °C. From the expected viscosity and temperature effects on the rotational correlation time of a rigid isotropic molecule, the relaxation rate should drop by 30–40% at 17 °C, and this is indeed observed for most imino protons (see Table I). It should be pointed out that the error in the experimental measurement is approximately 1 s^{-1} so that low exchange rates in the 1–2-s^{-1} range cannot be accurately determined in the presence of a background magnetic relaxation

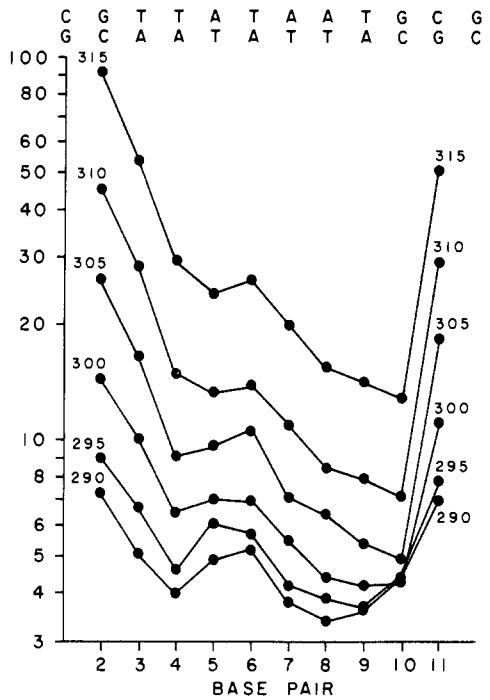

FIGURE 6: Observed inversion recovery rates of the imino protons of d(CGTTATAATGCG) as a function of a position in the helix at several temperatures. The asymmetry in the melting of the helix and the relatively unstable segment at base pairs 5 and 6 are clearly visible. Rates are in reciprocal seconds, and the absolute temperature is shown at the left and right ends of each curve.

rate of 4–5 s^{-1} at 17 °C. By 27–32 °C, the magnetic relaxation rate should drop to about 50% of its 0 °C value due to faster rotational diffusion; i.e., one expects a magnetic contribution of approximately 3 s^{-1} for AT pairs and 4–5 s^{-1} for GC pairs. Despite the uncertainty in this value, it is nevertheless apparent that, although TA 3, TA 6, and TA 9 have the same magnetic relaxation rates at 0 °C, by 27 °C TA 3 is echanging at 6–7 s^{-1} whereas TA 9 is exchanging no faster than 1–2 s^{-1}. Similarly, at 32 °C TA 3 (13–14 s^{-1}) and TA 6 (7–8 s^{-1}) are exchanging faster than TA 9 (2–3 s^{-1}). Thus, the internal AT pairs do not open at equal rates, and there is marked asymmetry in the base-pair dynamics with the TTA end of the promoter being more labile than the ATG end. At a higher level of detail, the central TA 6 seems to be consistently more labile than its neighbors.

The overall melting temperature (which in the NMR case reflects helix opening at greater than 200 s^{-1}) is relatively low as might be expected for an AT-rich sequences (see Figure 2). The addition of 100 mM NaCl and 5 mM $MgCl_2$ to the original 20 mM phosphate buffer increases the melting temperature of the helix by only 3–4 °C and makes no qualitative change in either the shift or exchange data. The observed recovery rates ($k_{mag} + k_{exch}$) as a function of position in the

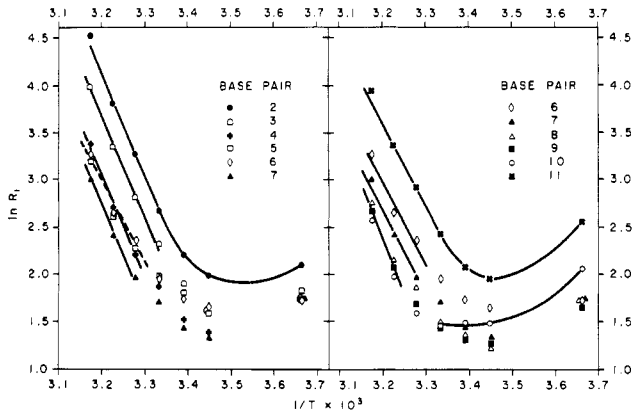

FIGURE 7: Arrhenius plots of the observed inversion recovery data for the consensus Pribnow sequence. The activation energies described in the text were derived by replotting the data after correction for the magnetic contribution to the relaxation at each temperature, which was calculated as described in the text.

helix are plotted in Figure 6 at several temperatures.

The temperature dependence of the experimental recovery rate for each base pair is shown in Figure 7 in the form of an Arrhenius plot. Clearly, the raw data must be corrected for the contribution from magnetic relaxation before the activation energy of the open-limited exchange process can be determined (Hurd & Reid, 1980; Early et al., 1981a,b; Patel et al., 1983a). The magnetic contributions at each temperature were calculated from the purely magnetic relaxation rates at 0 °C according to the Stokes–Einstein effect of viscosity and temperature on the correlation time and the linear dependence of the magnetic relaxation rate on the correlation time (spin-diffusion limit). Values of 64, 58, 53, 48, 44, and 41% of the 0 °C rates were used at 17, 22, 27, 32, 37, and 42 °C, respectively.

The uncertainty in estimating the magnetic contribution to the overall recovery rate leaves only a small range of useful temperatures at which exchange is dominant, thus making the activation energy determinations relatively inaccurate. Over the range of 295–310 K, the corrected plots for the imino protons are linear and yield activation energies of about 28 kcal/mol for base pairs 2, 4, 7, 8, 9, and 10. For base pairs 3 and 11, E_a is about 26 kcal/mol, and for base pairs 5 and 6, it is about 23 kcal/mol. In this temperature range, Patel et al. (1982a) have shown that the exchange rates are open limited at neutral pH, and the activation energy should be that of the opening process alone. The activation energies we have determined are somewhat higher than the ca. 15-kcal values reported by others (Early et al., 1981a,b; Patel et al., 1982c; Pardi et al., 1982). In this connection, we note that the slope of our uncorrected recovery rates corresponds to an activation energy close to 15 kcal. Furthermore, reprocessing the raw data from these earlier reports according to our own simple correction for magnetic relaxation at each temperature yields approximately 28 kcal of activation energy for AT pairs, i.e., values close to our own. Pardi et al. (1982) and Patel et al. (1982c) did not correct for magnetic relaxation whereas Early t al. (1981a,b) used a more complicated correction based on T_1 to T_2 ratios. The steep temperature dependence of their T_2 values at lower temperatures, perhaps due to partial aggregation, led to negligible estimates of magnetic relaxation rates above 20 °C. Regardless of the absolute value of the activation energy, it is still apparent that AT 5 and TA 6 are less stable than their surrounding AT pairs; i.e., they exchange more rapidly.

Since the activation energy is similar for the base pairs at each end of the helix, the asymmetry in the helix melting cannot be explained by sequential fraying from an unstable end because this would result in higher activation energies toward the interior of the helix. Patel has also observed unusually rapid exchange and low exchange activation energy for the unstable TATA subsequence in non-Pribnow sequences (Patel et al., 1983a) and also in a symmetrized Pribnow subsequence (Patel et al., 1983b). With the 2-fold symmetry imposed on this sequence (CGATTATAATCG), asymmetry in the helix melting is obviously not possible. The differences in stability toward the ends of the Pribnow helix do not seem to arise simply from differences in activation energy between the termini of the duplex and, hence, must arise through differences in the preexponential factors. These, unfortunately, are not well understood at present, and a physical explanation of the effect is not possible. As the number of sequences studied in detail increases, the basis of these effects should become more evident. The recovery rates at low temperatures are not interpretable since they are a complex function of the rates of magnetic relaxation, opening of the helix, and exchange from the open state. Since it is thought that RNA polymerase binds to promoters with concomitant separation of the two DNA strands, it is not surprising that these DNA sequences show relatively low stability. The possibility that the directionality of polymerase transcription is related to the asymmetric melting behavior, as well as the relationship between helix stability and polymerase affinity, must await studies on point mutations with known biological effects. Such studies are under way.

Registry No. d(CGTTATAATGCG)·d(CGCATTATAACG), 89144-59-2.

References

Beaucage, S. L., & Caruthers, M. H. (1981) *Tetrahedron Lett. 22*, 1859.

Chou, S.-H., Hare, D. R., Wemmer, D. E., & Reid, B. R. (1983) *Biochemistry 22*, 3037.

Chow, F., Kempe, T., & Palm, G. (1981) *Nucleic Acids Res. 9*, 2807.

Dhar, R., Weissman, S. M., Zain, B. S., Pan, J., & Lewis, A. M. (1974) *Nucleic Acids Res. 1*, 595.

Early, T. A., Kearns, D. R., Hillen, W., & Wells, R. D. (1981a) *Biochemistry 20*, 3756.

Early, T. A., Kearns, D. R., Hillen, W., & Wells, R. D. (1981b) *Biochemistry 20*, 3764.

Hare, D. R., & Reid, B. R. (1982a) *Biochemistry 21*, 1835.

Hare, D. R., & Reid, B. R. (1982b) *Biochemistry 21*, 5129.

Hare, D. R., Wemmer, D. E., Chou, S.-H., Drobny, G. D., & Reid, B. R. (1983) *J. Mol. Biol. 171*, 319.

Hawley, D. K., & McClure, W. R. (1983) *Nucleic Acids Res. 11*, 2237.

Hurd, R. E., & Reid, B. R. (1980) *J. Mol. Biol. 142*, 181.

Johnston, P. D., & Redfield, A. G. (1981) *Biochemistry 20*, 3996.

Kan, L.-S., Cheng, D. M., Jayaraman, K., Letzinger, E. E., Miller, P. S., & T'so, P. O. P. (1982) *Biochemistry 21*, 6723.

Matteucci, M. D., & Caruthers, M. H. (1981) *J. Am. Chem. Soc. 103*, 3186.

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

Patel, D. J., Kozlowski, S. A., Marky, L. A., Broka, C., Rice, J. A., Itakura, K., & Breslauer, K. J. (1982a) *Biochemistry 21*, 428.

Patel, D. J., Kozlowski, S. A., Marky, L. A., Broka, C., Rice, J. A., Dallas, J., Itakura, K., & Breslauer, K. J. (1982b) *Biochemistry 21*, 437.

Patel, D. J., Pardi, A., & Itakura, K. (1982c) *Science (Washington, D.C.) 216*, 581.
Patel, D. J., Ikuta, S., Kozlowski, S., & Itakura, K. (1983a) *Proc. Natl. Acad. Sci. U.S.A. 80*, 2184.
Patel, D. J., Kozlowski, S. A., Ikuta, J. S., Itakura, K., Bhatt, R., & Hare, D. R. (1983b) *Cold Spring Harbor Symp. Quant. Biol. 47*, 197.
Post, L. E., Arfsten, A. E., Reusser, F., & Nomura, M. (1978) *Cell (Cambridge, Mass.) 15*, 215.
Reid, B. R., & Hare, D. R. (1982) *Ciba Found. Symp. 93*, 208.
Rosenberg, M., & Court, D. (1979) *Annu. Rev. Genet. 13*, 319.
Roy, S., & Redfield, A. G. (1981) *Nucleic Acids Res. 9*, 7073.
Sass, M., & Ziessow, D. (1977) *J. Magn. Reson. 25*, 263.
Scheek, R. M., Russo, N., Boelens, R., Kaptein, R., & van Boom, J. H. (1983) *J. Am. Chem. Soc. 105*, 2914.
Siebenlist, U., Simpson, R. B., & Gilbert W. (1980) *Cell (Cambridge, Mass.) 20*, 269.
Tanaka, T., & Letsinger, R. L. (1982) *Nucleic Acids Res. 10*, 3249.
Wemmer, D. E., Chou, S.-H., Hare, D. R., & Reid, B. R. (1984) *Biochemistry* (following paper in this issue).
Zuiderweg, E. R. P., Scheek, R. Mm., Veeneman, G., van Boom, J. H., Kaptein, R. Ruterjans, R., & Beyreuther, K. (1981) *Nucleic Acids Res. 9*, 6553.

Sequence-Specific Recognition of DNA: Assignment of Nonexchangeable Proton Resonances in the Consensus Pribnow Promoter DNA Sequence by Two-Dimensional NMR†

David E. Wemmer, Shan-Ho Chou, Dennis R. Hare, and Brian R. Reid*

ABSTRACT: The resonances of most of the nonexchangeable protons of both + and − strands of the consensus Pribnow dodecamer d(CGTTATAATGCG) have been assigned by two-dimensional nuclear magnetic resonance methods. Application of the two-dimensional nuclear Overhauser effect (NOESY) sequential connectivity method, combined with two-dimensional autocorrelated (COSY) spectra to reveal scalar-coupled protons, results in assignment of virtually all of the base and sugar protons, except the sugar C5 protons which are inadequately resolved. Analysis of the nuclear Overhauser data indicates that the helix assumes a fairly uniform B form conformation.

The recent improvements in both DNA synthesis and NMR[1] methods have allowed studies of nucleic acid structure in solution to be undertaken with the possibility of revealing detailed informtion on the sequence dependence of structure and dynamics. The first sequences that we have decided to examine by these methods are those involved in the regulation of the expression of DNA. Proteins interact with these sequences, i.e., promoters and operators, in a highly sequence-specific manner to either enhance or abolish the expression of specific gene loci in the adjacent DNA. Crystallographic studies have been carried out on several specific DNA binding proteins (Anderson et al., 1981; Pabo & Lewis, 1981; McKay & Steitz, 1981) and, in the case of the cro repressor from phage λ, have given rise to the first detailed model for the interaction of a regulatory protein with its cognate DNA (Ohlendorf et al., 1982). However, since no regulatory DNA sequences and only a few DNA binding proteins have been crystallized, it will certainly be of continuing interest to examine these nucleic acids and proteins in solution, and NMR is perhaps the only method available that can provide reasonably high-resolution structural information as well as dynamic information about such structures in solution. NMR evidence has shown that the conformation of DNA observed in crystals need not be preserved in solution (Reid et al., 1983a), thus providing additional impetus for the examination of solution structures of nucleic acids. In the present work, we report the proton assignments for, and preliminary structural analysis of, the consensus Pribnow promoter sequence d(CGTTATAATGCG). As described in the preceding paper (Chou et al., 1983), this sequence, or a close analogue, is involved in the binding of RNA polymerase prior to transcribing the DNA into the mRNA copy from which protein synthesis takes place (Rosenberg & Court, 1979; Siebenlist et al., 1980; Hawley & McClure, 1983).

The development of two-dimensional NMR methods has made possible the systematic assignment of resonances in many moderately large biopolymers. The application of 2D NMR to proteins has been described in detail by Wuthrich and co-workers (Wuthrich et al., 1982; Billeter et al., 1982; Wagner & Wuthrich, 1982), and more recently, these techniques have been applied to nucleic acids (Feigon et al., 1982; Pardi et al., 1983; Scheek et al., 1983; Hare et al., 1983). Other workers have used one-dimensional methods to achieve partial assignments of DNA spectra (Reid et al., 1983a,b; Sanderson et al., 1983; Kan et al., 1982). In 2D NMR, use is made of both through-bond and through-space couplings to identify groups of protons in bases and sugars of the nucleic acid and to establish which groups are neighbors in the primary sequence. Once the assignments of the resonances to particular protons have been made, one can then proceed to the study of DNA structure and DNA–ligand interactions. This is

† From the Department of Chemistry and the Department of Biochemistry, University of Washington, Seattle, Washington 98195. *Received July 25, 1983.* We gratefully acknowledge NIH Grant R01 GM32302 and instrumentation grants from the Murdock Foundation, the National Science Foundation (PCM80-18053), and the National Institutes of Health (GM2874-01S1).

[1] Abbreviations: NMR, nuclear magnetic resonance; 2D NMR, two-dimensional NMR; NOESY, two-dimensional nuclear Overhauser effect; COSY, two-dimensional autocorrelated; DSS, 4,4-dimethyl-4-silapentane-1-sulfonate.

0006-2960/84/0423-2262$01.50/0