Vol. 164, No. 1, 1989 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
October 16, 1989 Pages 88-93

OBSERVATION OF ARGINYL-DEOXYOLIGONUCLEOTIDE INTERACTIONS IN *Taq*I ENDONUCLEASE BY DETECTION OF SPECIFIC ^{1}H NMR SIGNALS FROM 140kD [Nη1, Nη2, ^{15}N Arg]*Taq*I/OLIGOMER COMPLEXES

John Glushka[a], Francis Barany[b], and David Cowburn[a]

[a]The Rockefeller University, 1230 York Ave., N.Y., N.Y.,10021

[b]Dept. of Microbiology, Hearst Microbiology Research Center, Cornell University Medical College, N.Y., N.Y., 10021

Received August 14, 1989

Proton and nitrogen signals of the guanidinium amines in [Nη1, Nη2 ^{15}N Arg]*Taq*I endonuclease were observed using isotope filtered experiments and proton detected $^{1}H\{^{15}N\}$ heterocorrelated two dimensional NMR spectroscopy. These rapidly exchanging protons could be detected in the free enzyme only at pH 4.5; at pH 8.5, no signals were measured after extensive signal averaging. Addition of deoxyribonucleotide oligomers resulted in the appearance of two groups of signals at about 6.8 and 7.5 ppm. Since these signals are independent of the presence of cognate sequence or Mg^{2+}, it is assumed they represent nonspecific arginyl-DNA interactions. This labeling/NMR approach provides a new method for investigating the role of arginine in protein-DNA interactions.

Restriction endonucleases are specialized DNA binding proteins that recognize and cleave only their canonical sequence (1). DNA sequence recognition, as well as non-specific DNA binding, can be mediated by ionic contacts to the phosphate backbone (2,3), hydrogen bonds to the major groove (4), and numerous polar and hydrophobic interactions across a large protein/DNA surface. Crystallographic studies of an *Eco*RI-DNA complex suggest that sequence specificity is mediated by 12 hydrogen bonds. For each subunit, Arg^{200} forms two hydrogen bonds with guanine while Arg^{145} forms two hydrogen bonds with adjacent adenine residues (4). In contrast, for example, the crystal structure of the *trp* repressor/operator complex suggests specificity is mediated via multiple direct hydrogen bonds to the DNA phosphate backbone, as well as solvent mediated bonds to the bases (5). *Taq*I endonuclease recognizes its canonical sequence, TCGA, and upon addition of Mg^{2+} cofactor cleaves both strands prior to dissociating (6,7). Based on 'star' site cleavage preferences, one of us has proposed that the sequence discrimination is mediated by eight hydrogen bonds formed between *Taq*I and the cognate nucleotides within the major groove (8). In addition, ethylation inhibition of *Taq*I cleavage suggests that contacts are made to each phosphate within the canonical sequence (9). Although the residues involved in sequence recognition and DNA binding by *Taq*I are not known, arginine and lysine are suitable candidates as hydrogen bond donors. As part of an investigation to determine the functional regions and sequence specificity of *Taq*I endonuclease, NMR studies of arginyl residues were initiated.

The detection of the signals from arginyl guanidinium groups is challenging, owing to the exchange broadening of these protons (10), as well as the line broadening from the short transverse relaxation times of the high molecular weight protein-DNA complex. ^{15}N isotope labelling and use of isotope-filtered experiments (11) can select the signals of interest out of the background of amide and aromatic protons, and differentiate between the protein and the oligodeoxynucleotides.

0006-291X/89 $1.50

Materials and Methods

Sample Preparation. [Nη1, Nη2 ^{15}N Arg]*Taq*I endonuclease was purified from a strain of *E. Coli*, RR1[pFBLT88], which overproduces the endonuclease to 30% of soluble protein (12). RR1[pFBLT88] cells were induced by overnight growth in 2 liters 0.2 mM phosphate-MOPS medium containing 0.01% w/v [Nη1,Nη2 ^{15}N]Arg (13) , 0.01% w/v of the other 19 (unlabeled) amino acids, and 50 μg/ml ampicillin. Cells were harvested and sonicated, and *Taq*I was purified on a phosphocellulose column and concentrated by ammonium sulfate precipitation as described (12). Analysis of purified endonuclease by 0.1% SDS-12% polyacrylamide gel electrophoresis showed a single band migrating at 33 kDa, in agreement with the revised nucleotide sequence of *Taq*I endonuclease (14). The endonuclease elutes from a sizing column as a tetramer and is active both as a tetramer and dimer. The protein solutions used for NMR studies were 12 mg/ml (approximately 370 μM) dialyzed against either 20 mM sodium acetate, 1 mM EDTA, and 0.01% sodium azide at pH 4.5, or 20 mM Tris-SO_4 buffer and 0.5 mM EDTA at pH 8.5. NMR samples were made from 0.4 ml of these solutions plus 0.04 ml of D_2O. Oligodeoxynucleotides were a generous gift from Ira Schildkraut of New England Biolabs® Inc. (15). For the preparation of DNA-*Taq*I complexes, 0.4 ml of protein solution were added to 1 mg, 0.13 μmoles, of DNA linker, to give a molar ratio of 0.29 for protein tetramer to double strand oligodeoxynucleotide. Mg^{2+} was introduced by adding μl amounts of a 100 mM solution of $MgSO_4$ to the NMR sample.

NMR. All spectra were run on a GE GN500 spectrometer at 500 MHz, using 5mm probes provided by Cryomagnetics, Inc. or GE-NMRI. The one dimensional spectra were obtained with an ^{15}N filtered one pulse sequence (16, 17). Typically a Redfield 214 ^{1}H selective pulse was used in order to avoid excitation of the water signal (18). The short transverse relaxation time of the complex required shortening the 1/2*J* delay to 2.5 ms, and restricted the use of additional delays for ^{15}N proton refocussing and decoupling. Most spectra were acquired as antiphase signals coupled by ^1J(N-H), though in some cases, in-phase absorptive peaks could be obtained with an isotope filtered 1-1 refocussed sequence (19). In this case the delay was set to 1.5 ms. Typically, signals for the pH 4.5 sample were obtained with 1000-4000 scans. The pH 8.5 complex required 20,000 to 60,000 scans for adequate signal to noise, and so two-dimensional heterocorrelated spectra were impractical. One-dimensional data was processed with line broadening of 20 Hz. Because of the line widths (approximately 100 Hz at half height) and the low signal to noise ratio, spectra obtained with the non-refocussed pulse sequence could not be be readily phased to reveal distinct antiphase doublets. Therefore, for analysis, the spectra were phase adjusted to provide the lineshapes displayed in Figure 2, or a magnitude calculation was performed to estimate integrated intensities. The refocussed spectra obtained with the 1-1 pulse sequence were phased normally to give absorptive peaks (Figure 1). The two-dimensional ^{1}H{^{15}N} heterocorrelated spectrum (Figure 3) was obtained using the pulse sequence of Bax et al (20). A 214 ^{1}H selective pulse was used as the primary proton pulse, and the phase cycling provided zero quantum signals in t_1. The 2048 × 64 matrix was processed with a gaussian in t_2 and exponential line broadening and zero filling in t_1 to a final size of 2048 × 256. The spectrum is displayed in the absolute value mode.

Results and Discussion

At pH 4.5, the ^{15}N filtered one-dimensional spectrum (Figure 1b) shows a broad signal at 6.7 ppm, corresponding to typical arginyl guanidinium amine proton chemical shifts of 6.63 ppm at pH 3 (21). The ^1J(N-H) coupling is obscured due to the overlapping broad peaks. Upon cooling to 5°C, the signal splits into two peaks at 7.0 ppm and 6.5 ppm (data not shown), corresponding to the protons on each Nη amino group (22, 23). Although it is not possible to accurately estimate the integrated intensities of the arginyl resonances, the values obtained (10 - 30% of the theoretical value) indicate that most of the guanidinium amino protons are represented by the observed signal.

The 2D ^{1}H{^{15}N} heterocorrelated spectrum (Figure 3) shows a peak at 6.7 ppm from TSP for the ^{1}H signal and 70 ppm from NH_3 for the ^{15}N signal. The nitrogen chemical shift is in agreement with the expected value of 74 ppm that was observed for the guanidinium Nη1 and Nη2 nitrogens at pH less than 12 and temperatures above 4°C, in DMSO/H_2O mixtures (24, 25, 26). The 4 ppm difference is most likely due to the solvent, an effect which is also observed for amide nitrogens (27). Processing for increased resolution gives some fine structure and peak dimensions that roughly corresponds to the 90 Hz one bond coupling. Individual arginyl residues were not, however, identifiable.

At pH 8.5, no signals are observed in the filtered one-dimensional spectrum after extensive signal averaging, as might be expected from the rapid exchange rate of these protons in the free amino acid. There are no reported signals from arginyl residues at this pH in other proteins, although signals in *E. Coli* Adenylate Kinase were observed at pH 8 (28), and tentatively identified in ferrodoxin at pH 5.7 (29). Arginyl residues in *Taq*I involved in internal hydrogen bonding are not observable under these conditions.

When *Xho*I linker, d(CCGCTCGAGCGG) containing the TCGA canonical binding site was added, two groups of signals could be observed (Figure 2a) in the phase adjusted non-refocussed spectrum (see Materials and Methods). The upfield signal at 6.8 ppm is similar in chemical shift to that obtained with free arginine and also from the ^{15}N labeled *Taq*I at pH 4.5. The downfield signal at 7.5 ppm has approximately 30% of the upfield signal intensity. The combined intensity of these signals is 20 to 30% of the signal at pH 4.5. Variation in peak intensity due to proton exchange and other factors, coupled with the low signal to noise and phase distortions make it difficult to interpret these values in terms of number of residues participating.

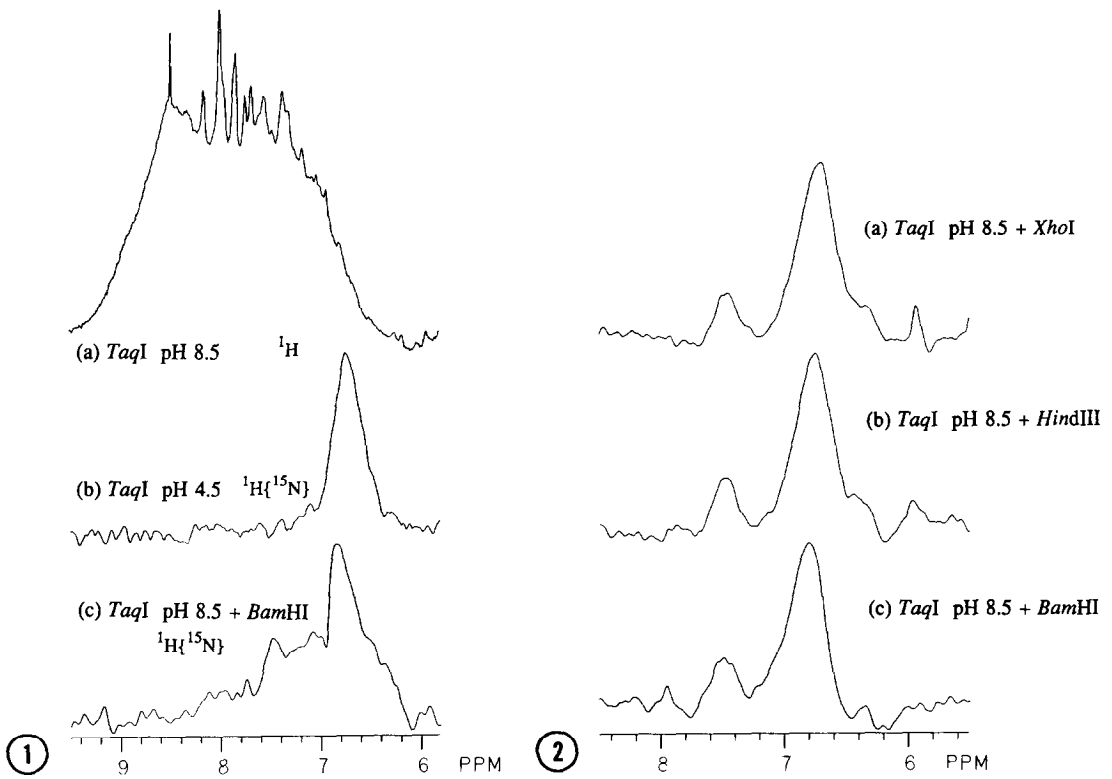

Figure 1: Downfield region of ^{1}H NMR spectra of [Nη1, Nη2 ^{15}N Arg]*Taq*I endonuclease, 0.4 mM in 90%H_2O/10%D_2O at 25°C. The pH and presence of oligodeoxynucleotides is indicated. Spectrum (a) was acquired using a 1-1 refocussed pulse sequence with 1024 scans and 3 Hz exponential line broadening applied; (b) represents 8000 scans of a similar sequence with ^{15}N isotope filtering; 20 Hz exponential line broadening was used. The ^{15}N filtered spectrum (c) represents 29,000 scans with line broadening of 30 Hz. The spectra are not scaled.

Figure 2: Downfield region of ^{15}N filtered ^{1}H NMR spectra of [Nη1, Nη2 ^{15}N Arg]*Taq*I endonuclease, 0.4 mM in 90%H_2O/10%D_2O. Samples were at pH 8.5 with the linker indicated. The pulse sequence used a Redfield 214 selective pulse and had no refocussing delay; the resonances were phase adjusted for presentation purposes only, and do not represent true gaussian shaped peaks. (see Materials and Methods)

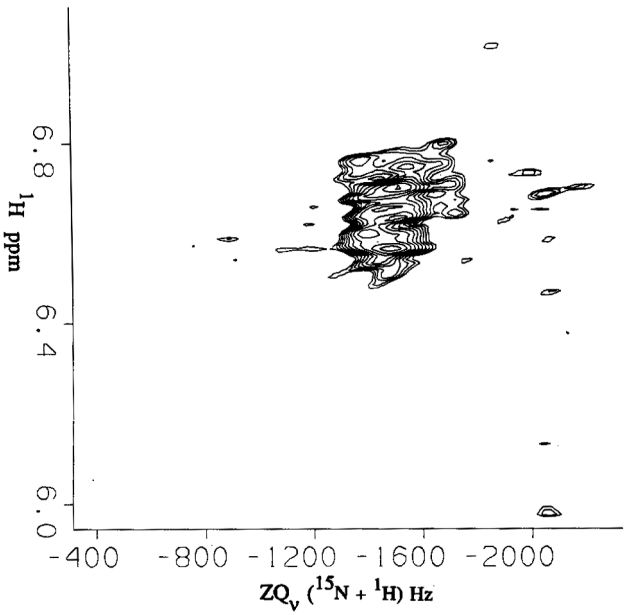

Figure 3: $^1H\{^{15}N\}$ HMP spectrum of [Nη1, Nη2 ^{15}N Arg]*Taq*I endonuclease, 0.4 mM, pH 4.5 in 90% H_2O/10%D_2O taken at 50°C. A 4096 × 64 matrix was collected, with spectral width of 8000 Hz in t_2 and 6666.7 Hz in t_1. A gaussian function was applied to t_2 and exponential line broadening was applied to t_1, followed by zero filling to 256 points. The spectrum is displayed in absolute value mode. The ^{15}N frequency corresponds to 70 ppm relative to liquid NH_3.

Similar results were obtained from addition of *Hin*dIII (Figure 1b) and *Bam*HI linker (Figure 1c), with sequences d(CCCAGCTTGGG) and d(CGCGGATCCGCG), respectively. That is, the same two peaks were observed at 6.8 ppm and 7.5 ppm, in approximately the same intensity ratio. The refocussed, decoupled spectrum of *Taq*I plus *Bam*HI linker shows the same main features with additional intensity in the region of 6.8 to 7.5 ppm (Figure 1c).

The appearance of signals in the complex at this pH demonstrates that solvent exchange rates of the guanidinium protons have been significantly lowered, by restricted access to the solvent and/or hydrogen bonding. The chemical shift distribution may thus arise from heterogeneity of the environments of separate residues, or may arise from the distinct $^1H\eta11$, $^1H\eta12$, $^1H\eta21$ and $^1H\eta22$ protons in conformationally restricted arginyl side chains. It has been observed in arginyl containing peptides that the guanidinium protons shifted downfield, when hydrogen bonded to carboxylate groups (30, 31). It has also been observed that the interaction between guanidinium groups of *Staphylococcus* nuclease and thymidine 3'5'-diphosphate, as well as between methylguanidinium ions and phosphate ions, favours two hydrogen bonds involving the Nε and one Nη amino group,(32), an arrangement that would restrict rotation about the Nε-Cζ bond, and possibly lead to chemical shift dispersion as seen at low temperature. Preliminary temperature studies were therefore initiated with the *Taq*I-oligomer complexes (data not shown). Between 25 and 40°C, the general peak pattern does not change significantly. However, on cooling to 5°C the signal at 6.8 ppm differentiates into two peaks as was observed for the *Taq*I sample at pH 4, and most likely represents freely rotating arginyl groups at 25°C. The signal at 7.5 ppm, however is no longer clearly discernible above the noise at the lower temperature. It is therefore not yet

established whether this peak also represents an averaged resonance of the protons on both Nη1 and Nη2 groups in a different environment, or represents one Nη amino group in a hydrogen bonded residue.

These results suggest that the peaks observed are a result of nonspecific interactions between arginyl side chains and the phosphate or sugar groups of the oligodeoxynucleotides. This conclusion was further supported by addition of $MgSO_4$ to some samples of *Taq*I-DNA complex. This was expected to initiate cleavage, which would be indicated by a loss of signals or chemical shift changes, if specific hydrogen bonds were broken after dissociating. No loss of signal was observed for samples up to 3 mM concentration of Mg^{2+}. At higher concentrations of the cation, signal intensity was reduced eventually to base noise levels, but this was accompanied by sample precipitation. However, we cannot rule out the possiblity that some signal results from a specific interaction in the major groove, where solvent accessiblity to arginyl side chain protons is reduced when the endonuclease binds either canonical or noncanonical DNA.

From an NMR technical view, the data presented confirm that meaningful information can be obtained from protein-DNA complexes of 140 kDa, where the J(N-H) value is on the order of the linewidth (33). With the benefit of less protein aggregation, greater sensitivity and other improvements, further investigation will determine to what extent the observed arginyl signals derive from direct protein-DNA contacts, or from changes in the conformation of the protein.

Acknowledgments

Supported by grants from NIH(GM-36596,D.C.; GM-41337-01,F.B.) and NSF(DMB-8714352,F.B.). NMR equipment was purchased with funds from NSF, NIH and the Keck Foundation.

References

1. Smith, H.O. (1979) Science 205, 455-466.
2. Lu, A.-L., Jack, W.E., and Modrich, P. (1981) J. Biol. Chem. 256, 13200-13206.
3. Terry, B.J., Jack, W.E., and Modrich, P. (1985) J. Biol. Chem. 260, 13130-13137.
4. McClarin, J. A., Frederick, C. A., Wang, B., Greene, P., Boyer, H. W., Grable, J., and Rosenberg, J. M. (1986) Science 234, 1526-1541.
5. Otwinowski, Z., Schevitz, R.W., Zhang, R.-G., Lawson, C.L., Joachimiak, A., Marmorstein, R.Q., Luisi, B.F., and Sigler, P.B. (1988) Nature 335, 321-329.
6. Sato, S., Hutchinson, C.A., and Harris, J.I. (1977) Proc. Natl. Acad. Sci. USA 74, 542-546.
7. Barany, F. (1987) Gene 56, 13-27.
8. Barany, F. (1988) Gene 65, 149-165.
9. Zebala, J. and Barany, F., unpublished results.
10. Wüthrich, K. (1986) NMR of Proteins and Nucleic Acids, Wiley, New York.
11. Griffey, R. K. and Redfield, A. G. (1987) Quart. Rev. Biophysics 19, 51-82.
12. Barany, F. (1988) Gene 65, 167-177.
13. Cambridge Isotope Laboratories, Woburn, MA.
14. Barany, F., unpublished results.
15. New England BioLabs Inc. 32 Tozer Rd., Beverly, MA., 01915-9990.
16. Freeman, R., Mareci, T. H., and Morris, G. A. (1981) J. Magn. Res. 42, 341-345.
17. Bendall, M. R., Pegg, D. T., Doddrell, D. M., and Field, J. (1981) J. Am. Chem. Soc. 103, 934-936.
18. Redfield, A. G., Kunz, S., and Hurd, T. (1975) J. Magn. Res. 19, 114-117.
19. Sklenar, V. and Bax, A. (1987) J. Magn. Res. 74, 469-479.
20. Bax, A., Griffey, R. H., and Hawkins, B. L. (1983) J. Magn. Res. 55, 301.
21. Bundi, A. and Wüthrich, K. (1979) Biopolymers 18, 285-298.
22. Klevan, L. and Crothers, D.M. (1979) Biopolym. 18, 1029-1044.
23. Smith, R.J., Williams, D.H., and James, K. (1989) J.Chem.Soc., Chem. Commun., 682-683.
24. Blomberg, F., Maurer, W., and Ruterjans, H. (1976) Proc. Natl. Acad. Sci. USA 73, 1409-1413.
25. Kanamori, K., Cain, A. H., and Roberts, J. D. (1978) J. Am. Chem. Soc. 100(16), 4979-4981.
26. Kanamori, K. and Roberts, J. D. (1983) J. Am. Chem. Soc. 105, 4698-4701.

27. Kricheldorf, H. R. (1981) Org. Magn. Res. 15, 162-177.
28. Glushka, J., Barzu, O., Cowburn, D., unpublished results.
29. Wand, A.J., DiStefano, D.l., Feng, Y., Roder, H., and Englander, S.W. (1989) Biochemistry 28, 186-194, Amer. Chem. Soc..
30. Lancelot, G., Mayer, R., and Hélène, C. (1979) J.Am.Chem.Soc. 101, 1569-1576.
31. Mayer, R. and Lancelot, G. (1981) J.Am.Chem.Soc. 103, 4738-4742.
32. Cotton, F.A., Day, V.W., Hazen, Jr. E.E., and Larsen, S. (1973) J.Amer.Chem.Soc. 95, 4834-4840.
33. Live, D. H., Kojiro, C. L., Cowburn, D., and Markley, J. L. (1985) J. Am. Chem. Soc. 107, 3043-3045.