Article

Affinity NMR: Decoding DNA Binding

Robert C. Anderson, Mengfen Lin, and Michael J. Shapiro

J. Comb. Chem., **1999**, 1 (1), 69-72• DOI: 10.1021/cc980004o • Publication Date (Web): 01 December 1998

Affinity NMR: Decoding DNA Binding

Robert C. Anderson,[†] Mengfen Lin,[‡] and Michael J. Shapiro*,[‡]

Department of Analytics, and Metabolic and Cardiovascular Diseases, Preclinical Research, Novartis Pharmaceuticals Corporation, 556 Morris Avenue, Summit, New Jersey 07901

Received June 30, 1998

We have shown that affinity NMR can be used to edit a NMR spectrum so that ligands that have affinity to DNA can be observed in the presence of other nonbinding molecules. Diffusion encoded spectroscopy (DECODES) can be used to identify the binding ligands. We were able to identify Hoechst 33342 as binding to the Drew−Dickerson dodecamer d(CGCGAATTCGCG)$_2$ in the presence of the nonbinding molecules adenine, adenosine, and thiamine. Affinity NMR appears to be readily applicable to DNA systems for the following reasons. (1) The relaxation rate of the DNA oligonucleotides is favorable, thus the signal intensity loss due to relaxation is not severe. (2) A comparison of the patterns of the DNA cross-peaks upon binding in the two-dimensional total correlation spectroscopy and correlation spectroscopy spectrum are easily performed, and the ligand signals in the two-dimensional DECODES spectrum can be readily identified. (3) The aromatic part of the DNA spectrum is devoid of 2D cross-peaks in these correlation spectra, greatly facilitating the interpretation of the bound ligand in the DECODES spectrum.

Introduction

The discovery and development of small-molecule, sequence-specific DNA ligands will likely have a major impact on the treatment of human diseases. Given sufficient affinity to disrupt DNA−protein interactions at repressor or effector binding sites within gene promoter regions, these ligands could modulate gene transcription in a therapeutically beneficial manner. Major groove triple-helix formation with oligonucleotides,[1] strand invasion with peptide nucleic acids,[2] minor groove ligands such as distamycin A,[3] calicheamicin-derived oligosaccharides,[4,5] and polyamides derived from *N*-methylimidazole and *N*-methylpyrrole amino acids[6] have initially demonstrated many favorable properties for such a therapeutic approach. With the overall concept of selective small-molecule transcription factors validated, attention will increasingly be turned to the discovery and optimization of such ligands.

Despite the amazing progress in de novo design of polyamide ligands,[6] this area of research may benefit from the application of combinatorial library approaches currently being applied to more traditional endeavors.[7] Such an approach may be immediately applicable to the polyamide series since this series is currently being prepared on resin.[8] Screening mixture libraries for high-affinity ligands may, however, be quite difficult. A reporter gene assay would be amenable to high-capacity screening but, being a cell-based assay, would be plagued by false positives and acquisition of definitive ligand binding structure−activity information would be restricted. In considering the problem, we speculated that our recently disclosed NMR diffusion spectroscopy (affinity NMR) approach could provide a solution to the problem.[9] This methodology, which relies on pulse field gradient (PFG) NMR to spatially encode molecules in solution, enables structure determination of the individual components due to differences in translational diffusion coefficients.[10]

The foundation of the approach is based upon the principle that translational diffusion in solution is size-dependent. We reasoned that the diffusion coefficient of a small molecule would appear altered on a time-averaged basis by complexation with a larger partner in solution and that the diffusion coefficients of the complexing ligand(s) would be significantly different from the noncomplexing compounds. This concept is reminiscent of separation of compounds by affinity chromatography. Using affinity NMR, molecules can be edited from the spectrum based upon their diffusion coefficients, and thus the bound ligands will be spectroscopically separated from the unbound molecules.

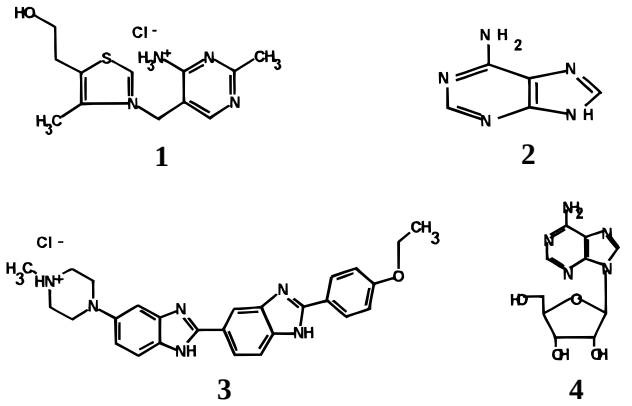
Furthermore, since the translational diffusion coefficient value is an intrinsic property of a molecule as a whole, it can be used to distinguish resonances arising from different molecules.[10] Diffusion encoded spectroscopy (DECODES), which involves (PFG) NMR combined with total correlation spectroscopy (TOCSY) allows the structure of the bound substrate to be identified.[11]

Results and Discussions

To test our approach, we selected the Drew−Dickerson dodecamer d(CGCGAATTCGCG)$_2$[12] (DDD) as a test DNA fragment[13] and Hoechst 33342 (**3**) as the binding ligand. A 1:1 mixture of DDD and **3** was obtained by stepwise addition of aliquots of **3** into a solution of DDD. The formation of a 1:1 complex was monitored by following the changes in the thiamine methyl resonances of the oligonucleotide.[14] Two-dimensional TOCSY and correlation spectroscopy (COSY) experiments were used to assign the resonances of **3**, which were observed at 8.10, 8.03, 7.89, 7.69, 7.36, and 7.14 ppm. The signals of the ethyloxy group from the 1D

[†] Metabolic and Cardiovascular Diseases.
[‡] Department of Analytics.

10.1021/cc980004o CCC: $18.00
Published on Web 12/01/1998



1 **2** **3** **4**

^{1}H and TOCSY spectra of **3** were observed at 1.51 and 4.25 ppm.

The mixture of nonbinding test molecules consisted of adenine (**1**), adenosine (**2**), and thiamine (**4**). One equivalent of each ligand was added into the Hoechst−DDD mixture. The PFG conditions under which the signals of the mixture of potential ligands **1**−**4** in the absence of the DDD are not observed were determined from a one-dimensional PFG NMR experiment. Then, one-dimensional PFG NMR experiments were repeated on the entire mixture, which included d(CGCGAATTCGCG)$_2$, **3**, in addition to the three nonbinding ligands.

In this experiment the PFG NMR spectra were collected with both a strong gradient pulse, under which the signals of the ligand are not observed in the absence of DNA, and also with a weaker gradient pulse. The resulting spectrum is presented in Figure 1.

As shown in Figure 1, the sharp resonances of the three nonbinding ligands are not observed in the PFG NMR spectrum using a strong gradient strength while the resonances of **3** remain observable. Comparison of the two PFG NMR spectra collected with weak and strong gradients shows that the rate of change of signal intensity arising from **3** was approximately the same as that of the d(CGCGAATTCGCG)$_2$, indicating that these compounds were bound to each other.

More detailed information of the ligand involved in binding can be obtained with the two-dimensional affinity NMR experiments. A DECODES−TOCSY experiment was conducted using the strong gradient pulse. The resulting spectrum is shown in Figure 2 along with the TOCSY spectrum obtained from the mixture. A comparison of the normal TOCSY spectrum with the DECODES TOCSY spectrum clearly shows a decrease in number of signals. In the DECODES spectrum, the cross-peak signals from **3** in the aromatic region at $\delta = 7.14$, 7.36, and 7.89 can still be clearly observed. All of the signals of the nonbinding ligands are missing throughout the spectrum.

The adenosine cross-peaks at $\delta = 5.95$, 4.69, and 4.19 observed in the regular TOCSY spectrum are absent in the DECODES spectrum. In addition, the cross-peaks of thiamine at 3.08, 3.78 and 2.46, 2.45 ppm have also disappeared in the DECODES spectrum. The absence of adenine, which has no cross-peaks in the TOCSY spectrum (signals at $\delta =$ 8.13 and 8.22 in the normal 1D NMR spectrum), can be deduced from inspection of the 1D PFG spectrum. Combining the results from the 1D PFG NMR spectrum with the DECODES result, we can identify the ligand that has high affinity to DDD as **3**. Interestingly additional signals arising from impurities in the purchased DNA sample were also found to be missing in the affinity-edited spectrum, indicating these molecules are small and also do not bind to the DNA.

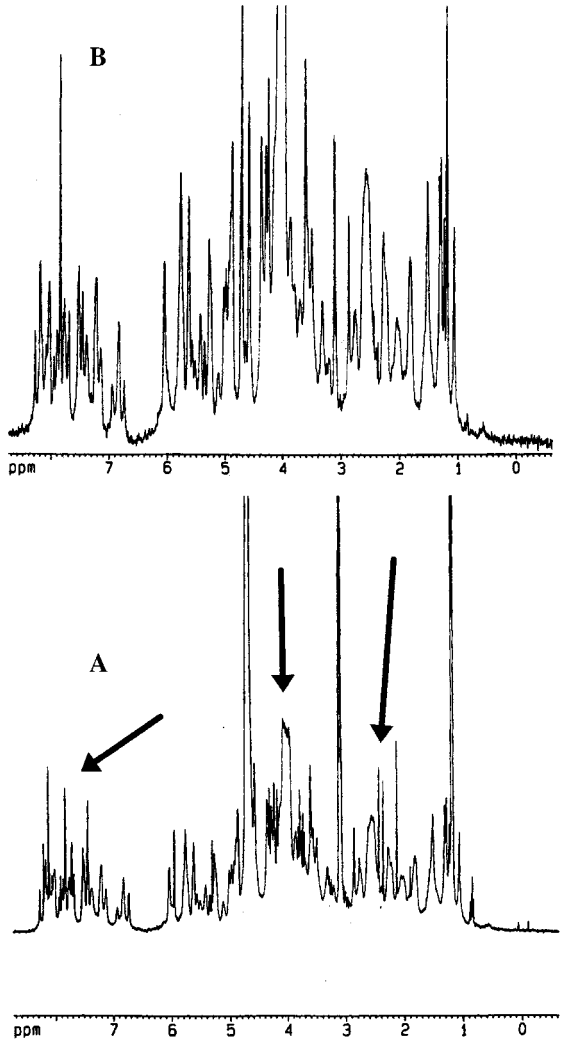


Figure 1. Comparison of the two PFG NMR spectra collected with (A) weak and (B) strong gradients for mixtures of DDD with compounds **2**−**5**. Arrows indicate where peaks have disappeared from the spectrum.

The ligands with high affinity to DNA can also be identified with a two-dimensional DECODES−COSY experiment collected using strong gradient pulses. A DECODES−COSY spectrum is shown in Figure 3. A comparison made between a COSY spectrum and a DECODES−COSY spectrum of the mixture demonstrates that the cross-peaks of the thiamine ethylene group at 3.08 and 3.87 ppm are absent in the DECODES−COSY spectrum. The cross-peaks of the adenosine at 3.81, 3.72 and 4.18, 4.67, and 5.94 ppm were also not observed in the DECODES−COSY spectrum. The cross-peaks arising from 3, however, were observed at 7.77−7.14, 8.01−7.37, and 8.10−7.88 ppm in addition to cross-peaks arising from DDD.

In conclusion, we have shown that affinity NMR can be used to edit the NMR spectrum so that ligands that have

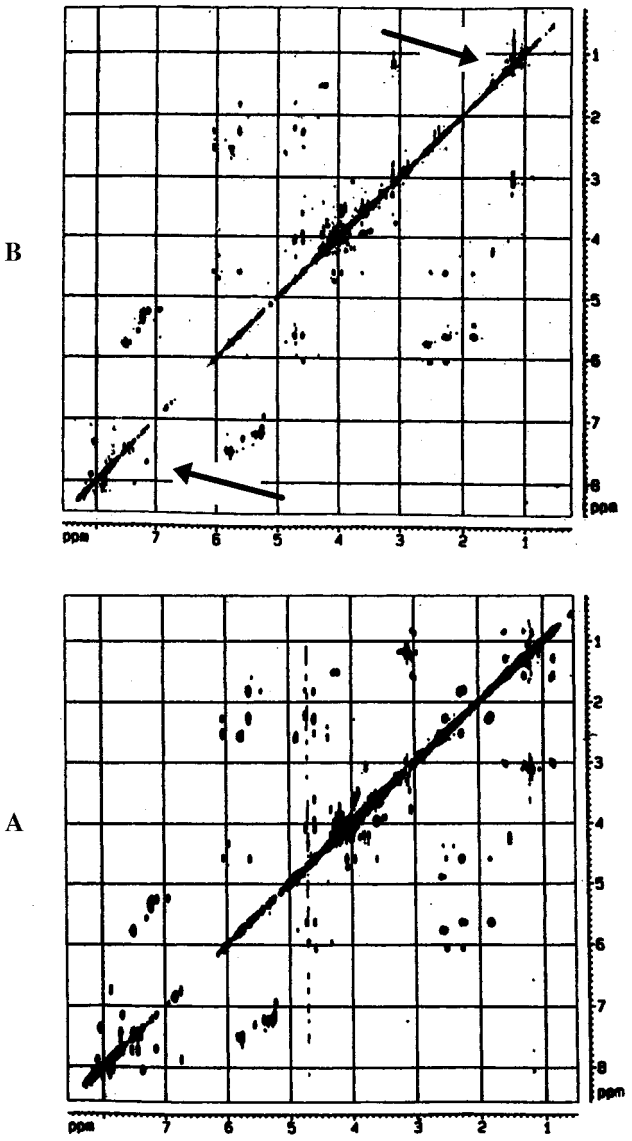


Figure 2. Comparison of a normal TOCSY spectrum (A) with the (B) DECODES TOCSY spectrum. Arrows indicate where peaks have disappeared from the spectrum.

affinity to DNA can be observed and that DECODES can be used to identify the ligands that are bound. Affinity NMR appears to be readily applicable to DNA systems. (1) The relaxation rate of the DNA oligonucleotides is favorable, thus the signal intensity loss due to the relaxation is not severe. As a result, a longer gradient pulse can be used to eliminate the unbound ligands, which is useful especially when the maximum gradient strength is small due to hardware limitations. (2) A comparison of the patterns of the DNA cross-peaks upon binding in the 2D TOCSY and COSY spectrum is easily performed, and the ligand signals in the two-dimensional DECODES spectrum can be readily identified. (3) The aromatic part of the DNA spectrum is devoid of cross-peaks, greatly facilitating the interpretation of the bound ligand in the DECODES spectrum.

Experimental Section

Materials. Purified d(CGCGAATTCGCG)$_2$ was purchased from National Biosciences, Inc., and was used without further purification. The DNA sample was prepared by dissolving 1 mg of the dodecamer in 0.5 mL of phosphate buffer containing 5 mM of phosphate and 10 mM of NaCl, pH 7. The sample was then lyophilized and resuspended in D_2O (99.9%). The final concentration of the DNA sample used in the NMR experiments was 0.5 mM. Hoechst 33428 was obtained from Sigma and was used without further purification. Thiamine, adenosine, and adenine were obtained from Aldrich.

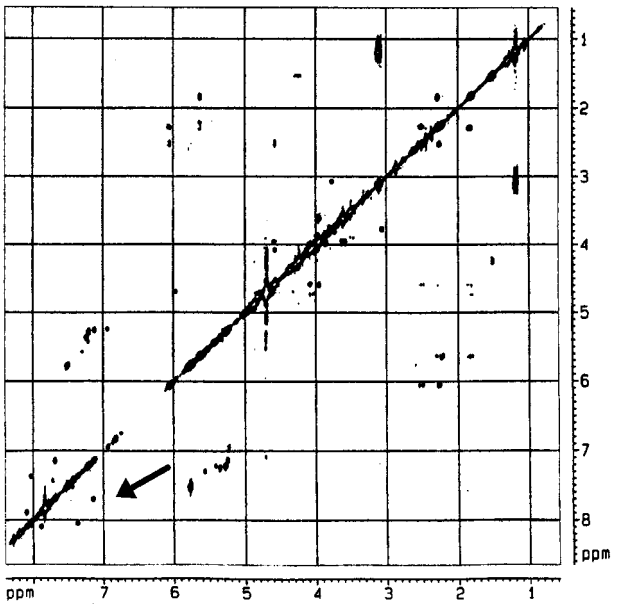


Figure 3. DECODES−COSY spectrum obtained with strong gradients. The arrow indicates the cross-peaks arising from compound **3**.

NMR Experiments. All NMR experiments were performed on a Bruker DMX-500 NMR spectrometer equipped with an Acustar gradient accessory. One-dimensional proton spectra were acquired into 32K data point over a spectrum width of 6K Hz. One-dimensional pulsed field gradient (PFG) NMR experiments were performed using the longitudinal eddy current delay pulse sequence.[10a] Two-dimensional experiments TOCSY and COSY were collected into the 4K complex for 512 t_1 increments. Two-dimensional DECODES spectra were collected with a ledmlevtp sequence using a gradient pulse strength of 40 G/cm.[9] 4K data points were collected at f2 dimension for the DECODES spectrum, and 512 increments were collected in the f1 dimension. The diffusion edited COSY spectrum was obtained using the bipolar bpp-COSY sequence.[10a] A total of 512 f1 increments and 4K f2 data points were collected for the DECODES−COSY spectrum. The gradient strengths used were the same as that used in the DECODES experiment.

References and Notes

(1) Beal, P. A.; Dervan, P. B. *Science* **1991**, *251*, 1360−1363.
(2) Hanvey, J. C.; Peffer, N. J.; Bisi, J. E.; Thomson, S. A.; Cadilla, R.; Josey, J. A.; Ricca, D. J.; Hassman, C. F.; Bonham, M. A.; Au, K. G.; Carter, S. G.; Bruckenstein, D. A.; Boyd, A. L.; Noble, S. A.; Babiss, L. E. *Science* **1992**, *258*, 1481−1485.
(3) Zimmer, C.; Wahnert, U. *Prog. Biophys. Mol. Biol*. **1986**, *47*, 32−112.
(4) Ho, S. N.; Boyer, S. H.; Schreiber, S. L.; Danishefsky, S. J.; Crabtree, G. R. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 9203−9207.
(5) Liu, C.; Smith, B. M.; Ajito, K.; Komatsu, H.; Gomez-Paloma, L.; Li, T.; Theodorakis, E. A.; Nicolaou, K. C.; Vogt, P. K. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 940−944.

(6) Gottesfeld, J. M.; Neely, L.; Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Nature* **1997**, *387*, 202−205.
(7) Ellman, J.; Stoddard, B.; Wells, J. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 2779−2782.
(8) Swalley, S. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc*. **1997**, *119*, 6953−6961.
(9) Lin, M.; Shapiro, M. J.; Wareing, J. R. *J. Am. Chem. Soc*. **1997**, *119*, 5249.
(10) (a) Morris, K. F.; Johnson, C. S., Jr. *J. Am. Chem. Soc*. **1992**, *114*, 3139−3141. (b) Morris, K. F.; Johnson, C. S., Jr. *J. Am. Chem. Soc*. **1993**, *115*, 4291−4299. (c) Morris, K. F.; Stilbs, P.; Johnson, C. S., Jr. *Anal. Chem*. **1994**, *66*, 211−215. (d) Gibbs S. J.; Johnson, C. S., Jr., *J. Magn. Reson.* **1991**, *93*, 395. (e) Lin, M.; Jayawickrama, D. A.; Rose, R. A.; Delvisio, J. A.; Larive, C. K. *Anal. Chim. Acta* **1995**, *307*, 449. (f) Stilbs, P.; Paulsen, K.; Griffiths, P. C. *J. Chem. Phys.* **1996**, *100*, 8180. (g) Lin, M.; Shapiro, M. J.; Wareing, J. R. *J. Org. Chem.* **1997**, *62*, 8930.
(11) Shapiro, M. J.; Lin, M. *J. Org. Chem.* **1996**, *61*, 7617.
(12) Drew, H. R.; Dickerson, R. E. *J. Mol. Biol*. **1981**, *151*, 535−556.
(13) Parkinson, J. A.; Barber, J.; Douglas, K. T.; Rosamond, J.; Sharples, D. *Biochemistry* **1990**, *29*, 10181−10190.
(14) Sriam, M.; van der Marel, G. A.; Roelen, H. L. P. F.; van Boom, J. H.; Wang, A. H.-J. *Embo J.* **1992**, *11*, 225−232.

CC980004O