ASSIGNMENTS OF THE ^{1}H NMR SPECTRUM OF A CONSENSUS DNA-BINDING PEPTIDE FROM THE HMG-I PROTEIN

Jeremy N. S. Evans§‡*, Mark S. Nissen§ and Raymond Reeves§
Departments of Biochemistry/Biophysics§ and Chemistry‡,
Washington State University, Pullman, WA 99164-4660. U. S. A.

INTRODUCTION

The HMG-I subfamily [1-3] of high mobility group (HMG[1]) chromatin proteins [4] consists of DNA-binding proteins that preferentially bind to stretches of A•T-rich sequence both *in vitro* [5-8] and *in vivo* [9]. Recently, members of the HMG-I family have been suggested to bind *in vitro* to the narrow minor groove of A•T-DNA by means of an 11 amino acid peptide binding domain (BD) which, because of its predicted structure is called the "A•T-hook motif" [10]. The HMG-I proteins are specific substrates for the cell cycle regulating enzyme(s) $p34^{cdc2/cdc28}$ kinase (also known as histone H1 kinase) both *in vivo* and *in vitro* [11-13]. The sites of phosphorylation by cdc2 kinase are the threonine residues at the amino terminal ends of the A•T-hook motifs and such modifications have been demonstrated to reduce markedly the affinity of binding of the HMG-I proteins to DNA [12,13].

HMG-I proteins are also of considerable biological interest because they are expressed at elevated levels in actively proliferating cells and have been observed to be a characteristic feature of undifferentiated [1,7] or neoplastically transformed cellular phenotypes [14,15]. High HMG-I levels have been found to be a consistent feature of rat and mouse malignant cells [14-17] and have been suggested to be protein markers for both neoplastic transformation [15] and metastatic potential [18]. The HMG-I proteins have also been implicated in control of DNA replication [19,20] and the regulation of gene transcription [8,21,22]. It is known from their primary sequences that the HMG-I proteins have the overall structure typical of Ptashne-type transcriptional activator proteins possessing both a DNA binding domain(s) and a highly acidic COOH terminus [23]. *In vitro* HMG-I-like proteins have been demonstrated to increase transcription of isolated ribosomal genes [21] and to alter the conformation and stability of A•T-rich regions of DNA [24] properties often associated with DNA-binding gene regulatory proteins.

As an initial attempt to determine in molecular detail the interaction of the BD peptide with A•T rich DNA, we have examined a synthetic 13 residue BD peptide by NMR spectroscopy. In this paper we report the assignments of the resonances for the peptide at 295 K and pH 3.4, and provide preliminary evidence on the peptide structure.

[1] Abbreviations: HMG, high mobility group; BD, binding domain; DQF, double quantum filtered; COSY, correlated spectroscopy; NOESY, nuclear Overhauser effect spectroscopy; ROESY, rotating-frame Overhauser effect spectroscopy; TOCSY, total correlated spectroscopy.

MATERIALS AND METHODS

Peptide Synthesis. The 13mer BD peptide (VPTPKRPRGRPKK) was synthesized by solid-phase synthesis (on a Departmental Applied Biosystems model 431A peptide synthesizer), and purified by reverse-phase HPLC on a Vydak C_4 column (1 × 25 cm) using a water (containing 0.1% trifluoroacetic acid)–acetonitrile gradient under standard conditions. The 13mer BD peptide eluted at 15% acetonitrile (R_t = 13 mins @ 1.5 mL min^{-1}).

NMR Spectroscopy. High field Fourier transform (FT) NMR studies were performed on a Varian VXR-500S (11.75 T, 500MHz ^{1}H) NMR spectrometer. Deuterium was used for locking the field. ^{1}H NMR chemical shifts were referenced externally to samples of similar dielectric constant containing sodium 3-(trimethylsilyl) propanoate-2,2,3,3-d_4 (TSP) in D_2O buffer (δ_H = 0.00 ppm). Sample temperature was maintained with a Varian variable temperature control unit, using gaseous nitrogen (from boil-off liquid nitrogen) cooled using an FTS XR-85 cryo-cooler. The majority of samples of peptide were maintained at 295 K, except where the amide exchange rates were being measured, when the sample was maintained at 277 K. Data was downloaded to either a Silicon Graphics 4D25TG or a 4D70GT workstation, and converted from Varian format to FELIX format using the VNMR2FELIX conversion program (a gift from Darrell Davies, University of Utah). The output from this was processed using FELIX (Hare Research Ltd.). All 2D data was obtained using the hypercomplex phase sensitive method [25] and processed as 2K × 2K complex data sets with baseline correction and sine-bell squared weighting functions in both dimensions.

DQF-COSY were recorded with the pulse sequence t_0–90°–t_1–90°–δ–90°–t_2, where t_1 is the evolution time, t_2 is the acquisition time, and δ is a fixed delay of 3 μs [26]. TOCSY spectra were recorded with the pulse sequence t_0–90°–t_1–SL_x–(MLEV-17)-SL_x–t_2, where SL_x was a 4 ms trim pulse along the x axis [27]. The MLEV-17 spin-locking pulse sequence was repeated to give a mixing time of 40 ms. ROESY spectra were recorded with the pulse sequence t_0–90°–t_1–90°–SL_x(30°)–90°–t_2, where SL_x(30°) is a small spin-lock pulse repeated to give a mixing time of 200 ms [28]. NOESY spectra were recorded with the pulse

TABLE 1 Sequential Assignments of HMG-I 13mer Binding Domain Peptide, pH 3.4, 295K

Residue	NH	αCH	βCH	Others	
V1	–	4.23	2.36	1.03, 1.14	
P2	–	4.58	1.91, 2.37	2.08	3.67, 3.80
T3	8.44	4.60	4.16	1.31	
P4	–	4.33	1.91, 2.35	2.04	3.73, 3.90
K5	8.10	4.23	1.76, 1.87	1.43	3.03, 3.25
R6	8.35	4.68	1.77, 1.87	1.71	3.05, 3.25, 7.55
P7	–	4.46	1.93, 2.34	2.05	3.68, 3.87
R8	8.54	4.34	1.73, 1.87	1.44, 1.50	3.05, 3.25
G9	8.41	3.92, 4.04			
R10	8.28	4.68	1.76, 1.88	1.73	3.25
P11	–	4.45	1.94, 2.34	2.05	3.66, 3.84
K12	8.44	4.32	1.79, 1.84	1.51	3.05
K13	7.24	3.27	1.74, 1.89	1.49	3.05

sequence t_0–90°–t_1–90°–τ_m–90°–t_2 [29,30] with mixing times, τ_m, of 100, 300, 400, and 600 ms. Solvent suppression was achieved by presaturation of the H_2O resonance for all the 2D experiments.

Sample Preparation. The BD peptide was dried with successive cycles of lyophilization and re-hydration with either H_2O or D_2O and then dissolved in one of the following: Buffer (i) 25 mM potassium phosphate, 0.01% (w/v) NaN_3, in 10% v/v D_2O/H_2O, pH 3.4; or Buffer (ii) 25 mM potassium phosphate, 0.01% (w/v) NaN_3, in 99% v/v D_2O/H_2O, pH 3.4. pH titrations were carried out by careful addition of small quantities of HCl or NaOH and measurement with a 4 mm pH electrode (Ingold Co.).

RESULTS

The 13mer peptide was studied by NMR at 277 and 295 K and at a variety of pH values. The linewidths of the 1D 500 MHz NMR spectra were not sensitive to concentration in the range 0.1-20 mM, indicating negligible aggregation of the peptide [31]. The use of 2-dimensional double-quantum filtered correlated spectroscopy (DQF-COSY) and total correlated spectroscopy (TOCSY) has enabled us to assign resonances to amino acid spin types, although distinctions

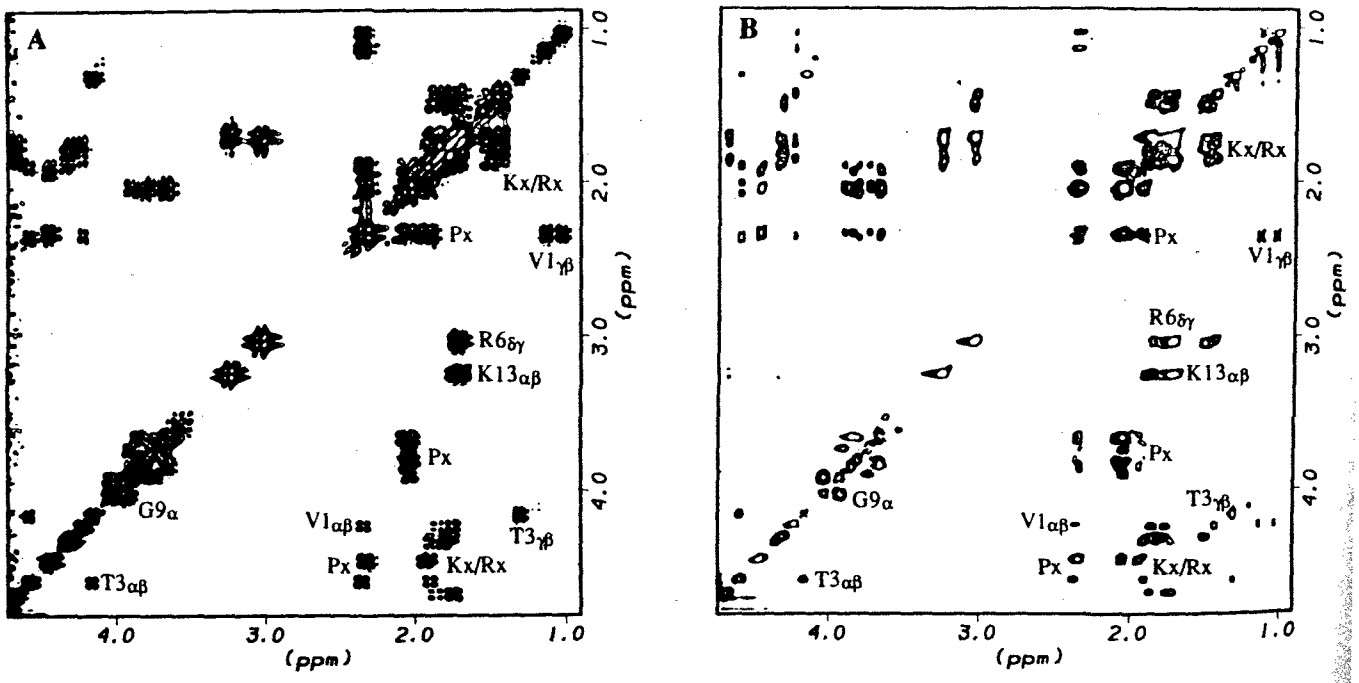

Figure 1 500 MHz ^1H DQF-COSY (A) and TOCSY (B) NMR spectra of the aliphatic resonances of the 13mer BD peptide (20 mM) in 90%H_2O/10%D_2O phosphate (25 mM) buffer, pH 3.4, 295K. Cross-peak assignments are indicated.

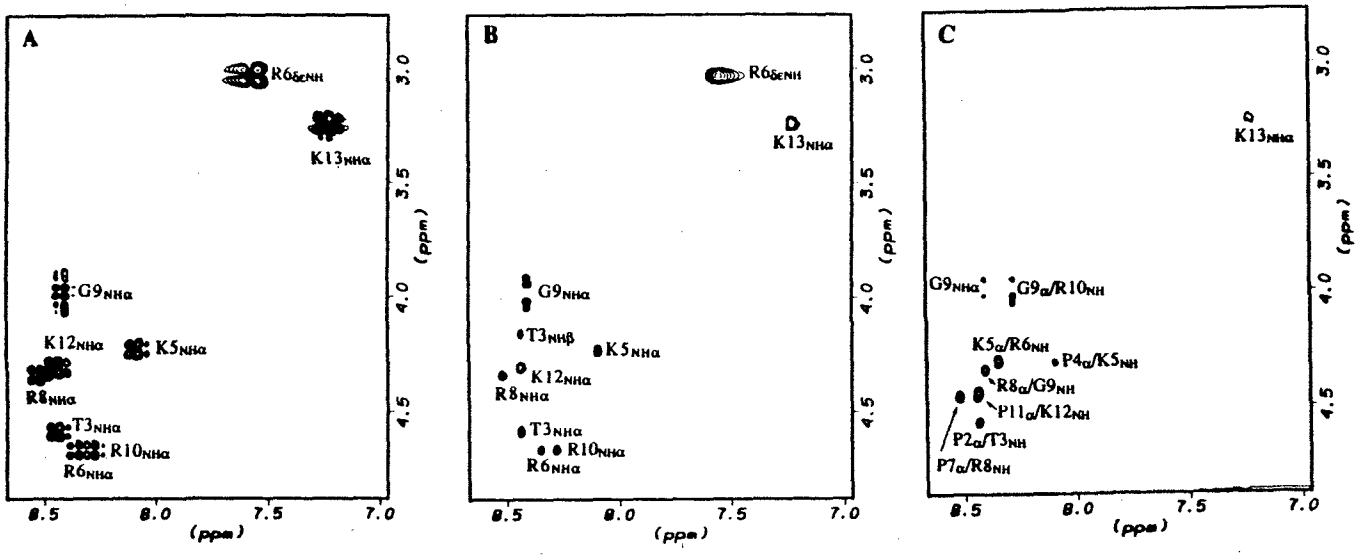

Figure 2 500 MHz ^{1}H DQF-COSY (A) and TOCSY (B) and 300 ms NOESY (C) NMR spectra of the amide resonances of the 13mer BD peptide (20 mM) in 90%H_2O/10%D_2O phosphate (25 mM) buffer, pH 3.4, 295K. Cross-peak assignments are indicated, with the subscript x used to denote unspecified residue numbers in crowded regions of the spectrum (sequential assignments are indicated in Table 1).

between lysine and arginine are not possible with these experiments alone. Figure 1 shows the DQF-COSY and TOCSY spectra of the aliphatic region of the 13mer BD peptide. Those cross-peaks which have been assigned are indicated on the figure. Sequential assignments were made using the $d_{\alpha N}(i, i+1)$ NOE connectivities from the NOESY spectrum (see Figure 2), and they are listed in Table 1. All four prolines were found to have *trans* conformations, on thebasis of the observed $d_{N\alpha}(i\text{-}1, i)$ NOE connectivities, where i=proline. Further confirmation was provided by examining the key NOE connectivities suggested by Wüthrich [32] using the ROESY experiment, in which the $d_{\alpha\delta}(i, i+1)$ NOE connectivities characteristic of a *trans*-proline were detectable for all four prolines. The strong $d_{\alpha\alpha}$ NOE expected for the *cis*-proline conformation was not detected. Finally, multiple cross-peaks for the V1 or T3 methyl resonances were not detected, and these would be expected for peptides containing mixed proline conformers [33], since *cis-trans* proline isomerization is known to be slow on the NMR timescale [34,35]. It should be noted that the absence of $d_{NN}(i, i+1)$ and $d_{\alpha N}(i, i+3)$ NOE connectivities is to be expected for a peptide containing a proline every 3-5 residues. The lack of $d_{\alpha\beta}(i, i+3)$ cross peaks in the ROESY or NOESY spectra suggests that the peptide side chains are relatively disordered under these conditions. However, since there is a high degree of symmetry in the molecule, there is also the possibility that those $d_{\alpha\beta}$ cross peaks for one residue may involve interresidue NOEs.

Preliminary examination of the amide exchange rates are consistent with selected amide resonances that are slowly exchanging. This was carried out by dissolving protonated BD peptide in deuterium oxide at 277 K, pH 3.4, and recording TOCSY spectra as a function of time. Also the persistence of amide resonances has been examined by obtaining TOCSY spectra at 295 K as a function of pH (2.2, 3.4, 5.5, and 7.1). The rapidly exchanging amides exchange with $t_{1/2}$ = < 5 mins, whereas the slowly exchanging amides exchange with $t_{1/2}$ = > 6 hours. Interestingly, the slowly exchanging amide resonances also do not titrate over the pH range 2.2 to 7.1, with the exception of the K5 NH, and all appear to be located along one side of the peptide molecule.

DISCUSSION

A model for how the consensus binding domain peptide from the HMG-I protein binds to the minor groove of A•T rich DNA has been proposed by this laboratory [10] on the basis of molecular modelling. In order to test this model, we have initiated NMR studies of a 13mer BD peptide in solution. While no $d_{NN}(i, i+1)$ and $d_{\alpha N}(i, i+3)$ NOE connectivities are detected for the BD peptide, this is to be expected for a peptide containing a proline every 3-5 residues. Furthermore while the amino acid side-chains would appear to be disordered, the detection of $d_{\alpha\delta}(i, i+1)$ NOE connectivities suggests that all the prolines adopt the *trans*-conformation. This implies that the backbone is conformationally restrained in the regions containing the prolines, and together with the existence of slowly exchanging amide protons on one side of the peptide these data suggest an asymmetric crescent structure. Interestingly, the amides which exchange slowly are consistent with the model of Reeves & Nissen [10], in which the amides lying along the inside of the crescent of the peptide include T3, K5, R6, R8, R10, and R(or K)12 (adopting the residue numbering of

the 13mer peptide). These amides might be expected to be less accessible to solvent, and potentially may participate in hydrogen-bonds with the carbonyl groups from (i, i + 2) residues. At this stage, there is no additional corroborating evidence from medium-range or long-range NOEs either to support or refute these speculations. However, in the presence of A•T rich DNA, the conformational lability of the peptide backbone and side-chains would be expected to be reduced significantly, and longer range NOEs detectable.

ACKNOWLEDGEMENTS

We should like to acknowledge Gerhard Munske for the synthesis of the 13mer peptide, Wendy Shuttleworth for help with some sample preparations, Darrell Davies (University of Utah) for the VNMR2FELIX conversion software. Supported in part by an American Cancer Society Institutional Grant IGR-119L (JNSE), NIH Grant R01 GM46352 (RR), and the WSU NMR Center is supported by NIH grant RR 0631401, NSF grant CHE 9115282 and Battelle Pacific Northwest Laboratories Contract No.12-097718-A-L2.

REFERENCES

[1] Lund, T., Holtlund, J., Fredriksen, M. & Laland, S. (1983) *FEBS Lett. 152*, 163-167.
[2] Johnson, K., Lehn, D., Elton, T., Barr, P. & Reeves, R. (1988) *J. Biol. Chem. 263*, 18338-42.
[3] Johnson, K., Lehn, D. & Reeves, R. (1989) *Mol. Cell. Biol. 9*, 2114-23.
[4] Goodwin, G. & Bustin, M. (1988) in Kahl, G. *Architecture of Eukaryotic Genes*, 187-205, VCH Weinheim, Germany.
[5] Strauss, F. & Varshavsky, A. (1984) *Cell 37*, 889-901.
[6] Solomon, M., Strauss, F. & Varshavsky, A. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 1276-1280.
[7] Elton, T., Nissen, M. & Reeves, R. (1987) *Biochem. Biophys. Res. Commun. 143*, 260-265.
[8] Reeves, R., Elton, T., Nissen, M., Lehn, D. & Johnson, K. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 6531-6535.
[9] Disney, J., Johnson, K., Magnuson, N., Sylvester, S. & Reeves, R. (1989) *J Cell Biol. 109*, 1975-82.
[10] Reeves, R. & Nissen, M. (1990) *J. Biol. Chem. 265*, 8573-8582.
[11] Lund, T. & Laland, S. (1990) *Biochem. Biophys. Res. Commun. 171*, 342-7.
[12] Reeves, R., Langan, T. & Nissen, M. (1991) *Proc. Natl. Acad. Sci. U S A 88*, 1671-5.
[13] Nissen, M., Langan, T. & Reeves, R. (1991) *J. Biol. Chem. 266*, 19945-19952.
[14] Giancotti, V., Pani-D'Andrea, P., Berlingieri, M. T., Di Fiore, P. P., Fusco, A., Veccio, G., Philip, R., Crane-Robinson, C., Nicolas, R. H., Wright, C. A., & Goodwin, G. H. (1987) *EMBO J. 6*, 1981-1987.
[15] Giancotti, V., Buratti, E., Perissin, L., Zorzet, S., Balmain, A., Portella, G., Fusco, A., & Goodwin, G. H. (1989) *Exp. Cell Res. 184*, 538-45.
[16] Elton, T. & Reeves, R. (1986) *Anal. Biochem.157*, 53-62.
[17] Vartiainen, E., Palvimo, J., Mahonen, A., Linnala-Kankkunen, A. & Maenpaa, P. (1988) *FEBS Lett 228*, 45-48.
[18] Bussemakers, M., van de Ven, W., Debruyne, F. & Schalken, J. (1991) *Cancer Res. 51*, 606-611.
[19] Grummt, F., Holst, A., Muller, F., Wegner, F., Schwender, S., Luksza, H., Zastow, G., & Klavinius, A. (1988) *Cancer Cells 6*, 463-466.
[20] Wegner, M., Zastrow, G., Klavinius, A., Schwender, S., Muller, F., Luksza, J., Hoppe, J., Wienberg, J. & Grummt, F. (1989) *Nucleic Acids Res. 17*, 9909-9932.
[21] Yang-Yen, H. & Rothblum, L. (1988) *Mol. Cell. Bio 8*, 3406-3414.
[22] Eckner, R. & Birnstiel, M. (1989) *Nucleic Acids Res. 17*, 5947-59.
[23] Ptashne, M. (1988) *Nature 335*, 683-689.
[24] Lehn, D., Elton, T., Johnson, K. & Reeves, R. (198 *Biochem. Int. 16*, 963-971.
[25] States, D. J., Haberkorn, R. A., & Ruben, D. J. (198 *J. Magn. Reson. 48*, 286-293.
[26] Rance, M., Sørensen, O.W., Bodenhausen, G., Wagne G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479-485.
[27] Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355-360.
[28] Kessler, H., Griesinger, C., Kerssebaum, R., Wagner, R., & Ernst, R. R. (1987) *J. Am. Chem. Soc. 109*, 607-609.
[29] Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. I (1979) *J. Chem. Phys. 71*, 4546-4553.
[30] Bodenhausen, G., Kogler, H., & Ernst, R. R. (1984 *J. Magn. Reson. 58*, 370.
[31] Dyson, H. J., & Wright, P. E. (1991) *Annu. Rev. Biophys. Biophys. Chem. 20*, 519-538.
[32] Wüthrich, K., Billeter, M., & Braun, W. (1984) *J. Mol. Biol. 180*, 715-740.
[33] Mayo, K. M., Parra-Diaz, D., McCarthy, J. B., & Chelberg, M. (1991) *Biochemistry 30*, 8251-8267.
[34] Grathwohl, C. & Wüthrich, K. (1976) *Biopolymers 15*, 2025-2041.
[35] Grathwohl, C. & Wüthrich, K. (1981) *Biopolymers 20*, 2623-2633.