# Homo- and Heteronuclear Two-Dimensional NMR Studies of the Globular Domain of Histone H1: Sequential Assignment and Secondary Structure<sup>†</sup>

Corinne Cerf,\*,‡ Guy Lippens,‡ Serge Muyldermans,§ Alain Segers,§ V. Ramakrishnan, Shoshana J. Wodak,‡ Klaas Hallenga, and Lode Wyns§

Unité de Conformation des Macromolécules Biologiques (UCMB), CP 160/16, Université Libre de Bruxelles, Avenue P. Héger, B-1050 Brussels, Belgium, Instituut voor Molekulaire Biologie, Vrije Universiteit Brussel, Paardenstraat 65, B-1640 Sint-Genesius-Rode, Belgium, Biology Department, Brookhaven National Laboratory, Upton, New York 11973, and Corvas Int. N. V. at UCMB and Corvas Int. N. V., Plateaustraat 22, B-9000 Ghent, Belgium

Received July 26, 1993®

ABSTRACT: A recombinant 75 amino acid polypeptide corresponding to the globular domain of the chicken histone H1 (GH1) has been studied by  $^{1}$ H homonuclear and  $^{1}$ H $^{-15}$ N heteronuclear 2D NMR spectroscopy. Sequential assignment of the backbone and  $\beta$ -proton resonances has enabled us to determine the secondary structure of GH1. It was found to consist of three helical regions (T7-S17, L25-Y37, E40-K56) and probably a  $\beta$ -hairpin (L59-L73). This structure is similar to the structure of the globular domain of histone H5 (GH5) obtained both by NMR spectroscopy [Zarbock et al. (1986) *Proc. Natl. Acad. Sci. U.S.A.* 83, 7628-7632; Clore et al. (1987) *EMBO J.* 6, 1833-1842] and by X-ray crystallography [Ramakrishnan et al. (1993) *Nature* 362, 219-223]. The  $\beta$ -hairpin as suggested for GH1 is also present in the X-ray structure of GH5 but has not been reported for the NMR structure of GH5.

The linker histone H1 is a component of the nucleosome, and is responsible for the folding and maintenance of chromatin into the 30-nm filament (Thoma et al., 1979). It is also thought to play a direct role in the regulation of DNA transcription and replication, and in cell proliferation (Croston et al., 1991; Zlatanova, 1990; Kamakaka & Thomas, 1990; Sun et al., 1989). The histones H1 constitute a family with many variants, differing in their affinity for chromatin (Kumar & Walker, 1980). Several variants are simultaneously present in a single cell. For example, the nucleated erythrocytes of birds contain both H1 and H5, the latter being an extreme variant of H1.

H1 and its variants have a multidomain structure (Bradbury et al., 1975). There is a central, trypsin-resistant, globular domain (GH1), I flanked by extended N- and C-terminal domains with a high content of lysines and arginines and thus very positively charged. GH1 is essential for the binding of H1 to the nucleosome. It is the part of H1 which is responsible for the specific positioning of the protein on the nucleosome, and it confers the same nucleosome protection from micro-

- \* To whom correspondence should be addressed.
- <sup>‡</sup> Université Libre de Bruxelles.
- § Vrije Universiteit Brussel.
- Brookhaven National Laboratory.
- <sup>⊥</sup> Corvas Int. N. V.
- Abstract published in Advance ACS Abstracts, October 1, 1993.

coccal nuclease digestion as complete H1 (Allan et al., 1980; Buckle et al., 1992). The isolated globular domains of different H1 variants also show different affinities for chromatin (Thoma et al., 1983).

Natural GH1, obtained by trypsin digestion of complete chicken histone H1, turned out to be too heterogeneous to allow successful NMR studies. The breakthrough was achieved by working with a 75-residue recombinant polypeptide having the sequence of a particular GH1 variant. As such, a more homogeneous preparation could be obtained compared to the natural fragment. This recombinant GH1 was studied by means of <sup>1</sup>H homonuclear and <sup>1</sup>H-<sup>15</sup>N heteronuclear 2D NMR spectroscopy. Comparison between analyzable parts of the natural GH1 spectra and the corresponding parts of the recombinant GH1 spectra permitted us to verify that both polypeptides possess the same fold (Segers, 1991). Here we present the assignment of <sup>1</sup>H and <sup>15</sup>N backbone resonances (up to  $\beta$ -protons) and the secondary structure of GH1. These data are discussed in the light of GH5 structures obtained previously by NMR (Clore et al., 1987) and more recently by X-ray crystallography (Ramakrishnan et al., 1993).

## MATERIALS AND METHODS

Preparation of Natural GH1. Natural chicken erythrocyte GH1 was prepared and purified on a TSK SP-5PW Ultropac (LKB) ion-exchange column according to Segers et al. (1991).

Preparation of Recombinant GH1. A GH1 gene was synthesized on the basis of the amino acid sequence of the globular domain of the 11L H1 variant from chicken erythrocyte (methionine residue +40-113 fragment). It is noticeable that this sequence is shared by globular domains of other chicken H1 variants (36-109 segment of .10 H1, 37-110 segment of 11R H1) (Coles et al., 1987). The gene was cloned in a pET 3c-based vector (Studier et al., 1990) and expressed in the Escherichia coli BL21(DE3) cells. Cells were grown until late-logarithmic phase in LB medium containing ampicillin (100  $\mu$ g/mL), or in minimal medium

<sup>&</sup>lt;sup>†</sup> C.C. is a Research Assistant of the National Fund for Scientific Research (Belgium). V.R. is supported by NIH Grant GM42796 and by the Office of Health and Environmental Research of the U.S. Department of Energy. Part of this work was financially supported by the National Fund for Scientific Research (Belgium). This paper presents research results of the Belgian Programme on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming.

<sup>&</sup>lt;sup>1</sup> Abbreviations: GH1, globular domain of histone H1; GH5, globular domain of histone H5; 1D, one-dimensional; 2D, two-dimensional; NMR, nuclear magnetic resonance; DQF-COSY, double-quantum-filtered J-correlated spectroscopy; DQSY, double-quantum spectroscopy; TOC-SY, total correlation spectroscopy; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; HSQC, heteronuclear single-quantum coherence; HMQC, heteronuclear multiple-quantum coherence; SCUBA, stimulated cross-peaks under bleached alphas.

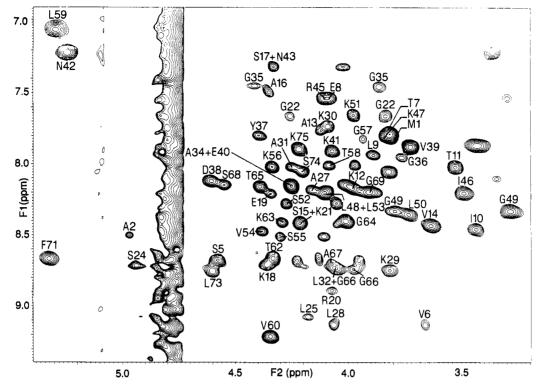


FIGURE 1: Fingerprint region of the TOCSY spectrum of GH1 in  $H_2O$  at 295 K, pH 5.8, with 80-ms spin-locking time. HN-H $\alpha$  cross-peaks are labeled.

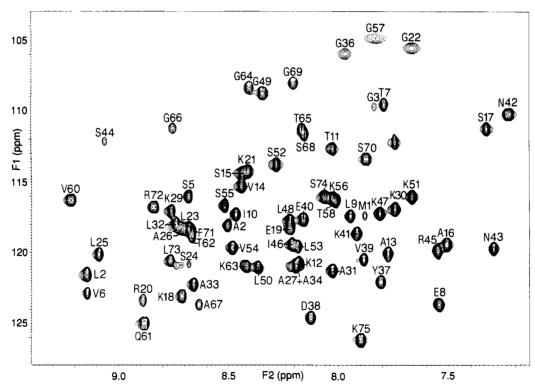


FIGURE 2: Part of the HSQC spectrum of uniformly <sup>15</sup>N-labeled GH1 in H<sub>2</sub>O at 295 K, pH 5.8. Backbone NH-HN correlation peaks are labeled.

containing  $^{15}$ NH<sub>4</sub>Cl (Aldrich) complemented with vitamin B1 (5  $\mu$ g/mL) for the  $^{15}$ N-labeled protein (Sambrook et al., 1989). Protein induction was achieved by addition of isopropyl  $\beta$ -D-galactopyranoside (1 mM). After 3-h induction for GH1 or overnight for  $^{15}$ N-labeled GH1, the cells were harvested by centrifugation and resuspended in 100 mM Tris-phosphoric acid (pH 8.9)/1 mM phenylmethanesulfonyl fluoride. The cells were disrupted by sonication, and GH1 was solubilized by bringing the NaCl concentration above 300 mM. The

cellular debris was removed by low-speed centrifugation; 380 mg of (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>/mL was added to the extract. After centrifugation, GH1 remained in the supernatant which was dialized overnight against 100 mM Tris-phosphoric acid (pH 8.9)/250 mM NaCl and applied on S-Sepharose FF (Pharmacia). After the unbound material passed, the protein was eluted with a linear NaCl gradient from 200 mM until 1 M. GH1 was desalted over a G25 column [replaced by a PD10 column (Pharmacia) for the <sup>15</sup>N-labeled protein] equilibrated

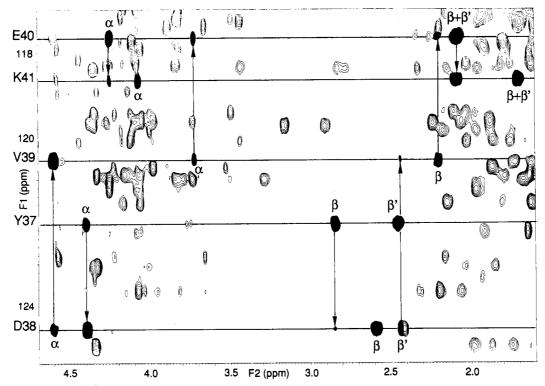


FIGURE 3: Part of the HSQC-NOESY spectrum of uniformly <sup>15</sup>N-labeled GH1 in H<sub>2</sub>O at 295 K, pH 5.8, showing the sequential assignment of the Y37-K41 segment.

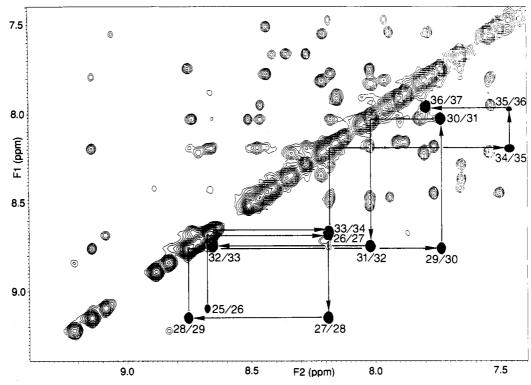


FIGURE 4: Amide region of the NOESY spectrum of GH1 in H<sub>2</sub>O at 295 K, pH 5.8, with 150-ms mixing time. The d<sub>NN</sub>(i,i+1) connectivities corresponding to helix II are highlighted.

with  $0.1 \times S$  buffer (10 × S buffer = 43 g of Na<sub>2</sub>SO<sub>4</sub>, 48.4 g of NaH<sub>2</sub>PO<sub>4</sub>·H<sub>2</sub>O, and 7 g of Na<sub>2</sub>HPO<sub>4</sub>/L, pH 5.8). The protein was lyophilized and brought into 7 × S buffer for NMR experiments. It gave one single band on a sodium dodecyl sulfate-polyacrylamide gel. The quality and the proper folding of GH1, very sensitive to the conditions of pH and ionic strength, were checked by UV absorbance, circular dichroism, and fluorescence (Smerdon & Isenberg, 1976).

NMR Spectroscopy. All NMR experiments were performed at 295 K on a Varian Unity 600-MHz spectrometer. Quadrature detection in  $F_1$  and suppression of axial peaks were achieved according to the TPPI-States method (Marion et al., 1989). We recorded homonuclear <sup>1</sup>H spectra of a 1 mM GH1 sample dissolved successively in 95% H<sub>2</sub>O/5% D<sub>2</sub>O and in 100% D<sub>2</sub>O, and heteronuclear <sup>1</sup>H-<sup>15</sup>N correlation spectra of a 2 mM <sup>15</sup>N-labeled GH1 sample dissolved in 95%

| Table I: 15N and 1H Resonance Assignments of GH1 in H <sub>2</sub> O at 295 K, pH 5.8 <sup>a</sup> |              |      |      |      |      |      |         |                  |      |      |      |      |      |
|--|--------------|------|------|------|------|------|---------|------------------|------|------|------|------|------|
| residue  | 15 <b>NH</b> | HN   | Нα   | Ηα′  | Нβ   | Hβ′  | residue | <sup>15</sup> NH | HN   | Ηα   | Ηα′  | Нβ   | Hβ′  |
| M1   | 117.4        | 7.88 | 3.82 |      |      |      | V39     | 120.6            | 7.88 | 3.72 |      | 2.20 |      |
| A2   | 118.1        | 8.50 | 4.97 |      | 1.80 |      | E40     | 117.7            | 8.16 | 4.25 |      | 2.09 | 2.09 |
| G3   | 109.8        | 7.83 | 3.95 | 3.81 |      |      | K41     | 118.7            | 7.92 | 4.07 |      | 1.70 | 1.46 |
| P4   | $133.1^{b}$  |      | 4.50 |      | 2.25 | 1.75 | N42     | 110.3            | 7.23 | 5.25 |      | 3.35 | 2.65 |
| S5   | 116.1        | 8.68 | 4.59 |      | 4.21 | 3.96 | N43     | 119.8            | 7.30 | 4.35 |      | 3.14 | 2.90 |
| V6   | 122.9        | 9.14 | 3.66 |      | 2.17 |      | S44     | 112.2            | 9.06 | 4.18 |      | 4.00 | 4.00 |
| <b>T7</b>  | 109.6        | 7.79 | 3.81 |      | 4.16 |      | R45     | 119.9            | 7.55 | 4.11 |      | 2.17 | 1.97 |
| E8   | 123.7        | 7.54 | 4.07 |      | 2.20 | 2.06 | I46     | 119.5            | 8.21 | 3.49 |      | 2.10 |      |
| L9   | 117.4        | 7.94 | 3.88 |      | 1.69 | 1.10 | K47     | 117.3            | 7.81 | 3.84 |      | 1.94 | 1.94 |
| I10  | 117.4        | 8.46 | 3.44 |      | 1.69 |      | L48     | 117.8            | 8.22 | 4.11 |      | 1.80 | 1.80 |
| T11  | 112.7        | 8.02 | 3.53 |      | 4.23 |      | G49     | 108.8            | 8.34 | 3.80 | 3.29 |      |      |
| K12  | 120.8        | 8.18 | 3.98 |      | 1.86 | 1.86 | L50     | 121.1            | 8.36 | 3.74 |      | 1.68 | 1.15 |
| A13  | 120.1        | 7.77 | 4.12 |      | 1.32 |      | K51     | 116.1            | 7.66 | 3.98 |      | 1.98 | 1.98 |
| V14  | 115.3        | 8.44 | 3.64 |      | 2.02 |      | S52     | 113.8            | 8.28 | 4.27 |      | 4.05 | 4.05 |
| S15  | 114.5        | 8.44 | 4.22 |      | 4.05 | 4.00 | L53     | 119.7            | 8.18 | 4.09 |      | 2.05 | 2.05 |
| A16  | 119.5        | 7.51 | 4.35 |      | 1.49 |      | V54     | 119.7            | 8.48 | 4.37 |      | 2.23 |      |
| S17  | 111.3        | 7.33 | 4.33 |      | 4.01 | 4.01 | S55     | 116.7            | 8.51 | 4.30 |      | 4.11 | 4.11 |
| K18  | 123.1        | 8.71 | 4.36 |      | 2.03 | 1.85 | K56     | 116.1            | 8.03 | 4.34 |      | 1.98 | 1.67 |
| E19  | 118.4        | 8.21 | 4.35 |      | 2.14 | 1.96 | G57     | 104.9            | 7.83 | 4.35 | 3.94 |      |      |
| R20  | 123.4        | 8.89 | 4.07 |      | 1.93 | 1.93 | T58     | 116.4            | 8.01 | 4.09 |      | 3.97 |      |
| K21  | 114.3        | 8.41 | 4.22 |      | 1.82 | 1.82 | L59     | 115.0            | 7.06 | 5.29 |      | 1.45 | 1.39 |
| G22  | 105.6        | 7.66 | 4.23 | 3.84 |      |      | V60     | 116.3            | 9.21 | 4.34 |      | 1.87 |      |
| L23  | 118.4        | 8.71 | 4.44 |      | 1.62 | 1.62 | Q61     | 125.0            | 8.88 | 4.83 |      | 1.74 | 1.63 |
| S24  | 120.9        | 8.68 | 4.93 |      | 4.43 | 4.43 | T62     | 119.0            | 8.66 | 4.33 |      | 4.13 |      |
| L25  | 120.1        | 9.09 | 4.17 |      | 2.03 | 1.75 | K63     | 121.0            | 8.41 | 4.30 |      | 1.84 | 1.79 |
| A26  | 118.3        | 8.69 | 4.08 |      | 1.47 |      | G64     | 108.4            | 8.40 | 4.01 | 4.01 |      |      |
| A27  | 121.0        | 8.19 | 4.16 |      | 1.64 |      | T65     | 111.3            | 8.16 | 4.39 |      | 4.25 |      |
| L28  | 121.6        | 9.14 | 4.06 |      | 2.15 | 1.63 | G66     | 111.3            | 8.75 | 4.04 | 3.97 |      |      |
| K29  | 117.1        | 8.76 | 3.82 |      | 2.03 | 1.92 | A67     | 123.7            | 8.63 | 4.07 |      | 1.46 |      |
| K30  | 117.0        | 7.74 | 4.09 |      | 1.97 | 1.97 | S68     | 111.8            | 8.15 | 4.54 |      | 4.01 | 4.01 |
| A31  | 121.3        | 8.03 | 4.26 |      | 1.54 |      | G69     | 108.1            | 8.20 | 3.91 | 3.91 |      |      |
| L32  | 118.0        | 8.73 | 4.06 |      | 1.91 | 1.30 | S70     | 113.4            | 7.87 | 4.75 |      | 3.42 | 3.42 |
| A33  | 122.3        | 8.65 | 4.86 |      | 1.65 |      | F71     | 118.4            | 8.68 | 5.34 |      | 2.83 | 2.77 |
| A34  | 121.1        | 8.19 | 4.25 |      | 1.58 |      | R72     | 116.8            | 8.83 | 4.88 |      | 1.72 | 1.63 |
| G35  | 101.6        | 7.47 | 4.42 | 3.86 |      |      | L73     | 120.7            | 8.76 | 4.60 |      | 1.77 | 1.63 |
| G36  | 106.0        | 7.97 | 4.42 | 3.76 |      |      | S74     | 116.1            | 8.06 | 4.21 |      | 3.82 | 3.82 |
| Y37  | 122.1        | 7.80 | 4.40 |      | 2.85 | 2.45 | K75     | 126.2            | 7.90 | 4.21 |      | 1.86 | 1.75 |
| D38  | 124.6        | 8.12 | 4.61 |      | 2.59 | 2.44 |         |                  |      |      |      |      |      |

<sup>a</sup> Chemical shifts (ppm) refer to the HSQC-NOESY spectrum with 150-ms mixing time. <sup>1</sup>H frequencies were measured relative to an internal reference of sodium trimethylsilylpropionate (0 ppm), and <sup>15</sup>N frequencies were measured relative to a fictive <sup>15</sup>N reference calculated by the method of Live et al. (1984). <sup>b</sup> Determined from the direct <sup>15</sup>N 1D spectrum.

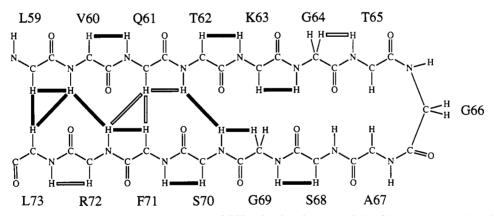


FIGURE 5: Schematic representation of the L59-L73 segment of GH1, showing the sequential and long-range backbone NOE connectivities indicative of the presence of a  $\beta$ -hairpin. Medium to strong NOEs are represented by black bars. The open boxes indicate that peaks could not be seen due to overlap or the presence of the water line.

 $\rm H_2O/5\%~D_2O$ . The former included DQF-COSY, DQSY (Braunschweiler et al., 1983), TOCSY with a spin-locking time of 40 and 80 ms (Bax & Davis, 1985), and NOESY with a mixing time of 150 ms. The latter included HSQC, HSQC-TOCSY with a spin-locking time of 60 ms, and HSQC-NOESY with a mixing time of 150 ms (Bodenhausen & Ruben, 1985; Norwood et al., 1990; Bax et al., 1990). Water suppression was achieved by presaturation, and the SCUBA method for recovery of saturated Hα resonances was used for the homonuclear spectra (Brown et al., 1988).  $^{15}$ N decoupling was accomplished by means of a WALTZ-16 sequence (Shaka

et al., 1983). Homonuclear spectra were typically acquired with 2048 points in  $F_2$  (4096 for the DQF-COSY) and 512 complex points in  $F_1$  (1024 for the DQSY). For the heteronuclear spectra, it was typically 2048 points in  $F_2$  and 256 complex points in  $F_1$ . In each case, zero-filling was applied to obtain a number of points equal to the next power of 2. Data were processed using the Varian software.

#### RESULTS

Sequential Assignment. The <sup>1</sup>H spectra of GH1 display a large amount of resonance overlap due to the amino acid



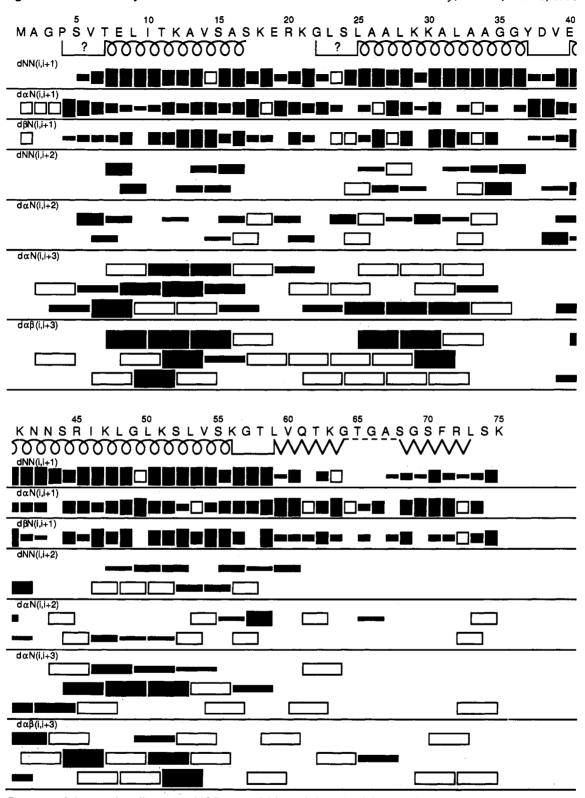


FIGURE 6: Summary of short- and medium-range NOE connectivities and elements of the secondary structure of GH1. NOE intensities are classified into strong, medium, or weak (intensity 3, 2, or 1), and the mean value between homonuclear (NOESY) and heteronuclear (HSQC-NOESY) peak intensities is taken in the case of connectivities involving an amide proton. The resulting intensities are represented by the thickness of the bars. The open boxes indicate that peaks could not be seen due to overlap or the presence of the water line. The secondary structure elements represented,  $\alpha$ -helices (helical lines), extended structures (broken lines), and turns (brackets), were determined on the basis of the sequential and medium-range NOE connectivities, as discussed in the text.

composition, which is rather unfavorable for an NMR study: two-thirds of the residues belong to only five amino acid types. Moreover, only two aromatic residues are present. The assignment thus required a combination of homonuclear and heteronuclear techniques even though the protein is small.

The assignment strategy consisted of identifying first all possible spin systems by the classical comparison of homonuclear DQF-COSY and TOCSY spectra. The TOCSY fingerprint region of GH1 is shown in Figure 1. The DQSY technique was also used to identify the glycine residues by

FIGURE 7: Comparison of the secondary structures of GH5 and GH1. (1) Secondary structure of GH5 deduced from NOE connectivities studied by NMR (Clore et al., 1987). (2) Secondary structure of GH5 deduced from the H-bonding scheme studied by X-ray crystallography (Ramakrishnan et al., 1993). (3) Proposed secondary structure of GH1 (this work). The secondary structure elements represented,  $\alpha$ -helices (helical lines), extended structures (broken lines), and turns (brackets), are discussed in the text.

their characteristic peak at  $HN - (H\alpha + H\alpha')$  frequencies, and the arginine residues by the peak at  $HN\epsilon - (H\delta + H\delta')$  frequencies. Moreover, both aromatic residues and one asparagine were identified, taking into account NOE connectivities between  $H\beta$  and  $H\delta$  ( $HN\delta$  for the asparagine). About 40 spin systems could be identified by this method.

To complete the sequential assignment, <sup>15</sup>N-<sup>1</sup>H heteronuclear spectroscopy was needed. A comparison between the NOESY spectrum of the unlabeled sample and a 15Ndecoupled NOESY spectrum of the labeled spectrum showed that GH1 possesses the same conformation in both samples. For the heteronuclear spectra, the HSOC technique was preferred over the HMQC technique because the former has no proton J couplings in the  $F_1$  dimension (15N) and leads thus to spectra with a better resolution in  $F_1$  (Bax et al., 1990; Norwood et al., 1990; Molinari et al., 1992). The resonances of GH1 appeared to be very well separated on the HSQC spectrum, in particular in  $F_1$  (Figure 2). The HSQC-TOCSY and HSQC-NOESY spectra were significantly better spread and resolved than the corresponding homonuclear TOCSY and NOESY spectra. The strategy consisted of labeling all previously identified spin systems in the heteronuclear spectra, before proceeding with the usual sequential walk by aligning intraresidual  $NH_i - (H\alpha, H\beta)_i$  peaks with interresidual  $NH_{i+1}$  $-(H\alpha, H\beta)_i$  peaks on the HSQC-NOESY spectrum (Figure 3). When a segment contained a sufficient number of identified spin systems, it could be positioned uniquely in the primary sequence of the protein. The segments could be linked together, leading to the sequential assignment of all backbone and  $\beta$ -proton resonances after proline-4. The first three residues before proline-4 were assigned by elimination. The resulting resonance assignments are listed in Table I. The assignment of the side-chain protons is now in progress and will be published subsequently, together with the tertiary structure of GH1.

Secondary Structure. Short- and medium-range NOEs provide a basis for identification of secondary structure elements (Wüthrich, 1986). Helices are characterized by stretches of strong  $d_{\rm NN}(i,i+1)$ , weak  $d_{\alpha \rm N}(i,i+1)$ , and medium or weak  $d_{\rm NN}(i,i+2)$ ,  $d_{\alpha \rm N}(i,i+2)$ , and  $d_{\alpha \rm B}(i,i+3)$  connectivities. The connectivities between any proton and an amide proton can be extracted from homonuclear NOESY but also from heteronuclear HSQC-NOESY, at the  $F_1$  frequency of the amide nitrogen. The relation between the intensity of the peak and the interproton distance is conserved (Molinari et

al., 1992). We carried out a qualitative analysis by classifying NOE intensities into strong, medium, or weak (intensity 3, 2, or 1), and then taking the mean value between homonuclear and heteronuclear peak intensities in the case of connectivities involving an amide proton. For GH1, three helices were identified in this way for residues T7–S17 (helix I), L25–Y37 (helix II), and E40–K56 (helix III). The amide region of the NOESY spectrum is shown in Figure 4. The  $d_{\rm NN}(i,i+1)$  connectivities corresponding to helix II are highlighted.

An extended structure is characterized by consecutive strong  $d_{\alpha N}(i,i+1)$  and weak or inobservable  $d_{NN}(i,i+1)$  connectivities. Using the same procedure as above, extended fragments were found for residues L59–Q61 and S68–R72. Searching for long-range connectivities involving these segments, we observed a pattern of NOEs that could correspond to a  $\beta$ -hairpin from residue L59 to residue L73 (Figure 5).

A turnlike structure involving four residues is characterized by a strong  $d_{\rm NN}(3,4)$  connectivity in combination with a  $d_{\alpha N}(2,4)$  connectivity. Such a pattern is present for P4–T7, G22–L25, Y37–E40, and K56–L59 segments. The attribution of the first two segments to a turnlike structure is a matter of debate. They could also constitute the beginning of helices I and II. However, the presence of the Y37–E40 and K56–L59 turns is more convincing. All observed short- and mediumrange NOEs are summarized in Figure 6.

#### DISCUSSION

While the precise limits of the secondary structure elements and their relative positioning must await detailed structural calculations, it is clear from the pattern of short- and mediumrange NOEs that the secondary structure of GH1 in solution is dominated by three helices. A high content in helical structure was predicted from theoretical considerations (van Helden, 1982). Some evidence also emerged from circular dichroism experiments (Pepe et al., 1990). Moreover, it is reflected by the fact that nearly all  $\alpha$ -protons resonate upfield relative to water (see Figure 1), which is an indication for a helical conformation (Wishart et al., 1991).

A comparison was undertaken between the secondary structures of GH1 and GH5, the globular domain of histone H5, an extreme variant of H1 (Figure 7). These homologous domains display 29 amino acids identities over 75 residues, amounting to 39%. The identities are unevenly distributed among the sequence: 14 in the N-terminal 56-residue part

(containing the 3 helices), representing 25%, and 15 in the C-terminal 19-residue part (containing the  $\beta$ -hairpin), representing 79%. The tertiary structure of GH5 has been determined first by NMR (Zarbock et al., 1986; Clore et al., 1987) and more recently by X-ray crystallography (Ramakrishnan et al., 1993). As far as the secondary structure is concerned, both GH5 structures possess three helices (I, II, and III) at roughly the same position (only the limits differ slightly), but the X-ray structure contains in addition a  $\beta$ -hairpin at the C-terminus and a short three-residue  $\beta$ -strand just before helix II. In GH1, the three helices and the suggested β-hairpin are perfectly aligned with the GH5 X-ray structure, but there is no evidence for a  $\beta$ -strand before helix II. Instead, the corresponding segment seems rather to adopt a turnlike conformation. The conformation of the C-terminal part of GH1 and GH5 is expected to be the same in similar conditions, as evidenced by the very strong sequence homology (79%) in this part of the molecule. If the  $\beta$ -hairpin is effectively present in GH1, the secondary structure of GH1 would thus be closer to the X-ray GH5 structure than to the NMR GH5 structure. suggesting that the difference between the GH5 NMR and X-ray structures is not due to the solution versus crystal state but rather to differences in physicochemical conditions. The β-hairpin present in GH5 at near-physiological pH (the protein was crystallized at pH 8) could be disrupted at the low pH (3.7) at which the NMR experiments on GH5 were carried out, as suggested by Ramakrishnan et al. (1993). The presence of the  $\beta$ -hairpin in the GH1 solution structure determined here could be attributed to the use of an intermediate pH (5.8) and the presence of high salt concentrations. In another respect, it is interesting to note the strong structural homology between GH1 and GH5 in the region of the three helices, displaying only 25% sequence identity. This fact could reflect a high evolutionary pressure to select certain structural

Given the high degree of similarity between GH1 and GH5 secondary structures, it is reasonable to assume a similar overall fold for both proteins. However, we must await the precise tertiary structure determination of GH1 to reveal the possible subtle differences between GH1 and GH5, which could explain functional differences such as their affinities for chromatin (Thoma et al., 1983).

#### **ACKNOWLEDGMENT**

We thank Prof. C. Houssier for his help with circular dichroism, Claude Maerschalk for his help with the <sup>15</sup>N referencing of heteronuclear NMR spectra, and Maria Vanderveken for excellent technical assistance in protein purification.

### REFERENCES

Allan, J., Hartman, P. G., Crane-Robinson, C., & Aviles, F. X. (1980) *Nature 288*, 675-679.

Bax, A., & Davis, D. G. (1985) J. Magn. Reson. 65, 355-360.

Bax, A., Ikura, M., Kay, L. E., Torchia, D. A., & Tschudin, R. (1990) J. Magn. Reson. 86, 304-318.

Bodenhausen, G., & Ruben, D. J. (1985) Chem. Phys. Lett. 69, 185-189.

Bradbury, E. M., Cary, P. D., Chapman, G. E., Crane-Robinson,
C., Danby, S. E., Boublik, M., Palau, J., & Aviles, F. (1975)
Eur. J. Biochem. 52, 605-613.

Braunschweiler, L., Bodenhausen, G., & Ernst, R. R. (1983) Mol. Phys. 48, 535-560.

Brown, S. C., Weber, P. L., & Mueller, L. (1988) J. Magn. Reson. 77, 166-169.

Buckle, R. S., Maman, J. D., & Allan, J. (1992) J. Mol. Biol. 223, 651-659.

Clore, G. M., Gronenborn, A. M., Nilges, M., Sukumaran, D. K., & Zarbock, J. (1987) EMBO J. 6, 1833-1842.

Coles, L. S., Robins, A. J., Madley, L. K., & Wells, J. R. E. (1987) J. Biol. Chem. 262, 9656-9663.

Croston, G. E., Kerrigan, L. A., Lira, L. M., Marshak, D. R., & Kadonaga, J. (1991) Science 251, 643-649.

Kamakaka, R. T., & Thomas, J. (1990) EMBO J. 9, 3997-4006.
Kumar, N. M., & Walker, I. O. (1980) Nucleic Acids Res. 8, 3535-3551.

Live, D. H., Davis, D. G., Agosta, W. C., & Cowburn, D. (1984)
J. Am. Chem. Soc. 106, 1939-1941.

Marion, D., Ikura, M., Tschudin, R., & Bax, A. (1989) J. Magn. Reson. 85, 393-399.

Molinari, H., Esposito, G., Consonni, R., Pegna, M., & Zetta, L. (1992) J. Biomol. NMR 2, 289-299.

Norwood, T. J., Boyd, J., Heritage, J. E., Soffe, N., & Campbell, I. D. (1990) J. Magn. Reson. 87, 488-501.

Pepe, I., Catasti, P., Nizzari, M., & Nicolini, C. (1990) Biochim. Biophys. Acta 1041, 14-21.

Ramakrishnan, V., Finch, J. T., Graziano, V., Lee, P. L., & Sweet, R. M. (1993) *Nature 362*, 219-223.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Segers, A. (1991) Ph.D. Thesis, Vrije Universiteit Brussel.

Segers, A., Muyldermans, S., & Wyns, L. (1991) J. Biol. Chem. 266, 1502-1508.

Shaka, A. J., Keeler, J., Frenkiel, T., & Freeman, R. (1983) J. Magn. Reson. 52, 335-338.

Smerdon, M. J., & Isenberg, I. (1976) Biochemistry 15, 4233–4242.

Studier, F. W., Rosenberg, A. H., Dunn, J. J., & Dubendorff, J. W. (1990) Methods Enzymol. 185, 61-89.

Sun, J.-M., Wiaderkiewicz, R., & Ruiz-Carrillo, A. (1989) Science 245, 68-71.

Thoma, F., Koller, T., & Klug, A. (1979) J. Cell Biol. 83, 403-427.

Thoma, F., Losa, R., & Koller, T. (1983) J. Mol. Biol. 167, 619-640.

van Helden, P. D. (1982) J. Theor. Biol. 96, 327-336.

Wishart, D. S., Sykes, B. D., & Richards, F. M. (1991) J. Mol. Biol. 222, 311-333.

Wüthrich, K. (1986) NMR of Proteins and Nucleic Acids, pp 162-175, Wiley, New York.

Zarbock, J., Clore, G. M., & Gronenborn, A. M. (1986) Proc. Natl. Acad. Sci. U.S.A. 83, 7628-7632.

Zlatanova, J. (1990) Trends Biochem. Sci. 15, 273-276.