Protein Science (2000), *9*:627–636. Cambridge University Press. Printed in the USA.

A helix-turn motif in the C-terminal domain of histone H1

ROGER VILA,[1] IMMA PONTE,[1] M. ANGELES JIMÉNEZ,[2] MANUEL RICO,[2] AND PEDRO SUAU[1]

[1]Departamento de Bioquímica y Biología Molecular, Facultad de Ciencias, Universidad Autónoma de Barcelona, 08193 Bellaterra, Barcelona, Spain
[2]Instituto de Estructura de la Materia, CSIC, Serrano, 119 28006 Madrid, Spain

(RECEIVED October 13, 1999; FINAL REVISION January 5, 2000; ACCEPTED January 28, 2000)

Abstract

The structural study of peptides belonging to the terminal domains of histone H1 can be considered as a step toward the understanding of the function of H1 in chromatin. The conformational properties of the peptide Ac-EPKRSVAFKKT KKEVKKVATPKK (CH-1), which belongs to the C-terminal domain of histone H1° (residues 99–121) and is adjacent to the central globular domain of the protein, were examined by means of ^{1}H-NMR and circular dichroism. In aqueous solution, CH-1 behaved as a mainly unstructured peptide, although turn-like conformations in rapid equilibrium with the unfolded state could be present. Addition of trifluoroethanol resulted in a substantial increase of the helical content. The helical limits, as indicated by $(i, i + 3)$ nuclear Overhauser effect (NOE) cross correlations and significant up-field conformational shifts of the C^{α} protons, span from Pro100 to Val116, with Glu99 and Ala117 as N- and C-caps. A structure calculation performed on the basis of distance constraints derived from NOE cross peaks in 90% trifluoroethanol confirmed the helical structure of this region. The helical region has a marked amphipathic character, due to the location of all positively charged residues on one face of the helix and all the hydrophobic residues on the opposite face. The peptide has a TPKK motif at the C-terminus, following the α-helical region. The observed NOE connectivities suggest that the TPKK sequence adopts a type (I) β-turn conformation, a σ-turn conformation or a combination of both, in fast equilibrium with unfolded states. Sequences of the kind (S/T)P(K/R)(K/R) have been proposed as DNA binding motifs. The CH-1 peptide, thus, combines a positively charged amphipathic helix and a turn as potential DNA-binding motifs.

Keywords: C-terminal domain; circular dichroism; helix-turn motif; histone H1; nuclear magnetic resonance

Histone H1 has a role in the stabilization of both the nucleosome and chromatin higher-order structure. H1 linker histones have a characteristic three domain structure (Hartman et al., 1977). The central globular domain consists of a three-helix bundle with a β-hairpin at the C-terminus (Clore et al., 1987; Cerf et al., 1993; Ramakrishnan et al., 1993) that is similar to the winged-helix motif found in some sequence-specific DNA-binding proteins. The amino-terminal and carboxy-terminal tail-like domains are highly basic.

H1 plays a key role in the folding of the nucleosomal arrays into the 30 nm chromatin fiber. However, experiments with sperm nuclei and cell-free extracts from *Xenopus* eggs show that H1 is not essential for the assembly of morphologically normal nuclei capable of DNA replication (Dasso et al., 1994). Likewise, knockout experiments in *Tethrahymena thermophila* show that linker histones are not essential for cell survival, although the chromosome structure is less condensed (Shen et al., 1995).

Reprint requests to: Dr. P. Suau, Departamento de Bioquímica y Biología Molecular, Facultad de Ciencias, Universidad Autónoma de Barcelona, 08193 Bellaterra, Barcelona, Spain; e-mail: Pere.Suau@uab.es.

Experiments in vivo indicate that H1 does not function as a global transcriptional repressor, but instead participates in complexes that either activate or repress specific genes (Zlatanova & Van Holde, 1992; Khochbin & Wolffe, 1994; Wolffe et al., 1997). Regulated expression of H1 during *Xenopus* development has a specific role in the differential expression of oocyte and somatic 5S rRNA genes (Bouvet et al., 1994). In *Tethrahymena*, H1 does not have a major effect on global transcription, but can act as either a positive or negative gene-specific regulator of transcription in vivo (Shen & Gorovsky, 1996). Previous work has clearly established that the globular domain of H1 is sufficient to direct specific gene repression in early *Xenopus* embryos (Vermaak et al., 1998). Other gene-specific effects, such as gene activation of the mouse mammary tumor virus (Lee & Archer, 1998) or the activation or repression of specific genes in *Tethrahymena* (Dou et al., 1999), however, are regulated by phosphorylation localized to the tail-like domains.

The study of the detailed structure of the terminal domains may provide insight into the binding of H1 in chromatin and thus contribute to the understanding of H1 function. H1 terminal domains have little or no structure in solution. The C-terminal domain,

however, might acquire a substantial proportion of α-helical structure upon interaction with DNA (Clark et al., 1988; Hill et al., 1989). The C-terminal domain is rich in lysine, alanine, and proline and binds to the linker DNA. It is required for chromatin condensation and is responsible for the ordered aggregation of DNA giving rise to the "psi-DNA" spectrum in circular dichroism (CD) (Morán et al., 1985). The low degree of defined secondary structure in the C-terminal domain in aqueous solution may be due to the electrostatic repulsion between the positively charged lysine side chains. Trifluoroethanol and other agents known to stabilize secondary structure induce variable amounts of helical structure in the C-terminal domain of H1 subtypes (Clark et al., 1988; Hill et al., 1989).

Here, we study the conformational properties of a peptide belonging to the C-terminal domain of the H1 subtype H1° by high-resolution NMR and CD. The peptide is adjacent to the globular domain and contains the longest proline-free fragment in the C-terminal domain of H1°. It binds to DNA as shown by gel retardation assays and CD (R. Vila & P. Suau, unpubl. obs.). We show that the peptide acquires a high amount of helical structure in aqueous TFE solutions. The helical region presents a marked amphipathic character, with all positively charged residues concentrated on one face of the helix and all the hydrophobic residues, together with a Glu residue, on the other. The last four residues of the peptide, TPKK, adopt a turn conformation. Sequences of the kind (S/T)P(K/R)(K/R) have been proposed as DNA binding motifs (Suzuki, 1989; Suzuki et al., 1993). The peptide thus combines a positively charged amphipathic helix and a turn as potential DNA-binding motifs.

Results

CD analysis

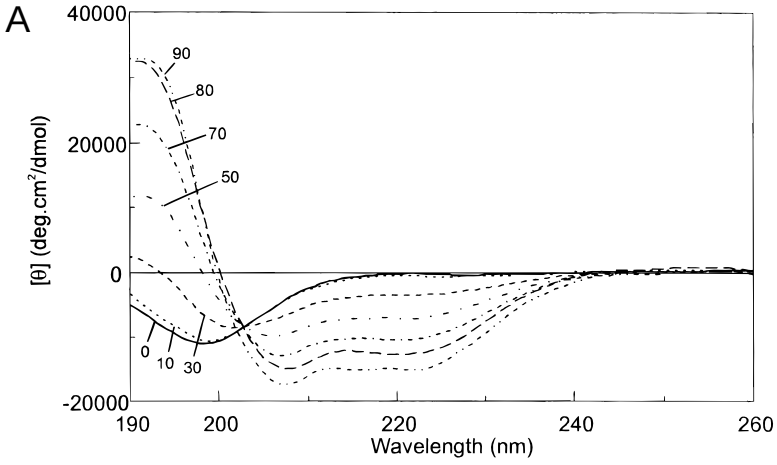
The CD spectrum of the peptide in H_2O, pH 3.5, and 5 °C is dominated by the contribution of the random coil. No sign of the characteristic double minimum at 208 and 222 nm, and the maximum at 190 nm of the α-helix was observed. The mean residue molar ellipticity at 222 nm ($[\theta]_{222}$), taken as diagnostic of helix formation, was negligible in water. However, the small positive peak at ~215 nm, characteristic of the random coil, was not observed, suggesting that a very small amount of structure could be present in water (Fig. 1). Addition of TFE, which is known to stabilize peptide secondary structure, resulted in an increase in the negative ellipticity at 222 nm (Fig. 1). The helical content of the peptide as a function of TFE concentration was estimated by the method of Chen et al. (1974) (Fig. 1). In 50 and 90% TFE solution, the helical populations were estimated to be 21 and 45%, respectively.

NMR analysis

The NMR spectra were recorded in aqueous solution and in 50% and 90% TFE. The presence of helical conformations was established on the basis of the following criteria: (1) the presence of stretches of nonsequential $\alpha N(i,i+3)$, $\alpha N(i,i+4)$, $\alpha\beta(i,i+3)$, as well as side-chain–side-chain and side-chain–main-chain $i,i+3$ and $i,i+4$ NOE connectivities; (2) strong sequential *NN* NOE connectivities, concomitantly with weakened $\alpha N(i,i+1)$ connectivities; and (3) significant up-field shifting of the C^{α} resonances relative to the random coil values.

Figure 2 shows selected regions of the two-dimensional (2D) NOE spectra of the peptide in 90% TFE, pH 3.5, 25 °C, where NOE correlations corresponding to medium-range interactions are indicated. Figure 3 and Tables 1 and 2 summarize all relevant NOE data for the peptide in water and in 50 and 90% TFE solution. The figure also shows the plot of the conformational shifts of the $C^{\alpha}H$ protons, $\Delta\delta = \delta_{\text{observed}} - \delta_{\text{random coil}}$. The peptide does not show a significant helical population in water. However, the presence of abundant medium NOE correlations between sequential amide protons (Fig. 3), which are only close enough in folded structures, suggests that turn-like conformations in rapid equilibrium with the unfolded state could be present.

Analysis of the peptide in 50% TFE solution reveals a significant helical population, as shown by the presence of nonsequential



B

% TFE	$[\theta]_{222}$	% helix
90	-15181,9	45
80	-12625,4	37
70	-10462,6	31
50	-7218	21
30	-3626,8	11
10	-625,7	2
0	-290,2	1

Fig. 1. TFE-dependent conformational transition of the CH-1 peptide measured by CD. **A:** Far-ultraviolet CD spectra in the presence of various concentrations of TFE in phosphate buffer 5 mM, pH 3.5 at 5 °C. The numbers refer to the TFE concentration in percentage by volume. **B:** Variation of the mean residue molar ellipticity ($[\theta]$, deg cm^2 dmol^{-1}) at 222 nm with added TFE. The percentage of helical structure, calculated as described in Materials and methods, is also indicated.

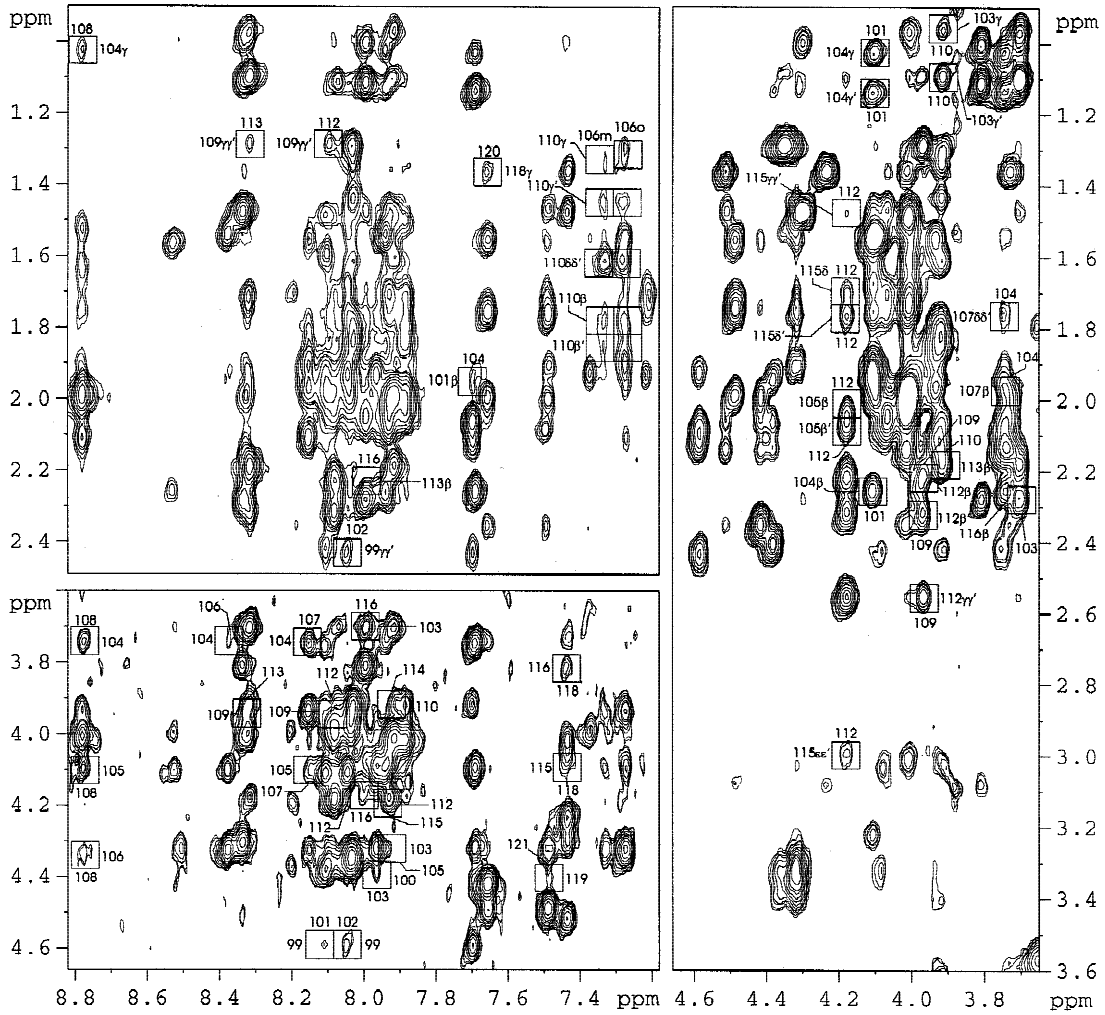


Fig. 2. Selected regions of the 2D NOE spectrum of CH-1. The peptide was 2.2 mM in 90% TFE, pH 3.5, 25 °C, mixing time 150 ms. Medium-range NOE connectivities are boxed. **A:** NOE correlations of aliphatic protons (bottom, $C^{\alpha}H$) with amide or aromatic resonances. **B:** NOE correlations of $C^{\alpha}H$ protons with other aliphatic protons.

NOEs, the intensity of the *NN* NOE cross peaks, several side-chain–side-chain and side-chain–main-chain NOE connectivities and the upfield conformational shifts, $C^{\alpha}H\Delta\delta$. According to the chemical shift criterion, a residue is considered to form part of a helix whenever its $\Delta\delta$ value is negative. In 90% TFE, the proportion of helical structure is considerably enhanced, as revealed by the increase in number and magnitude of the above α-helix NMR diagnostic parameters.

For the determination of the helical limits, we used a criterion based both on the conformational shifts and on the presence of $\alpha N(i,i+3)$, $\alpha\beta(i,i+3)$ NOE cross correlations. In 90% TFE solution, the helix spans from Pro100 to Val116, with Glu99 and Ala117 as N- and C-caps. The helical limits have been confirmed by structure calculations based on the distance constraints arising from NOE cross correlations in 90% TFE. The presence of $(i,i+3)$ cross correlations involving Glu99 (Hγ) and Arg (NH) suggests that Glu99 may act as the N-cap. A Pro residue is found in the next position. This location is consistent with the strong preference of Pro for position N1 (the first residue after the N-cap), when Pro is in the first turn of the helix (Richardson & Richardson, 1988).

The peptide bond Glu99–Pro100 is predominantly in a *trans* conformation as indicated by strong connectivities between Pro H$\delta\delta'$ and Glu Hα (Hinck et al., 1993). A very weak connectivity between the Pro Hα and the Thr Hα was observed, indicating the presence of a small population of Pro in the *cis* conformation.

In 50% TFE solution, there is a significant helical population in the segment Lys101–Lys111, with an average $\Delta\delta$ of -0.12 ppm. The average $\Delta\delta$ for residues Glu112 to Ala117 is close to zero; however, there are several $(i,i+3)$ NOEs in this region, indicating that the helix might be less stable than in 90% TFE or distorted. In 90% TFE, the segment Glu112–Ala117 shows an average $\Delta\delta$ value of -0.29 ppm, equivalent to that observed in the rest of the helix.

The helix populations in 50 and 90% TFE, as calculated from the average conformational shifts and according to the different helical limits, were 39 and 72%, respectively. With the purpose of comparison with the values obtained by CD, these percentages were corrected by the total length of the peptide. Values of 19% (50% TFE) and 53% (90% TFE) were obtained. These values are in satisfactory agreement with those estimated from the CD measurements.

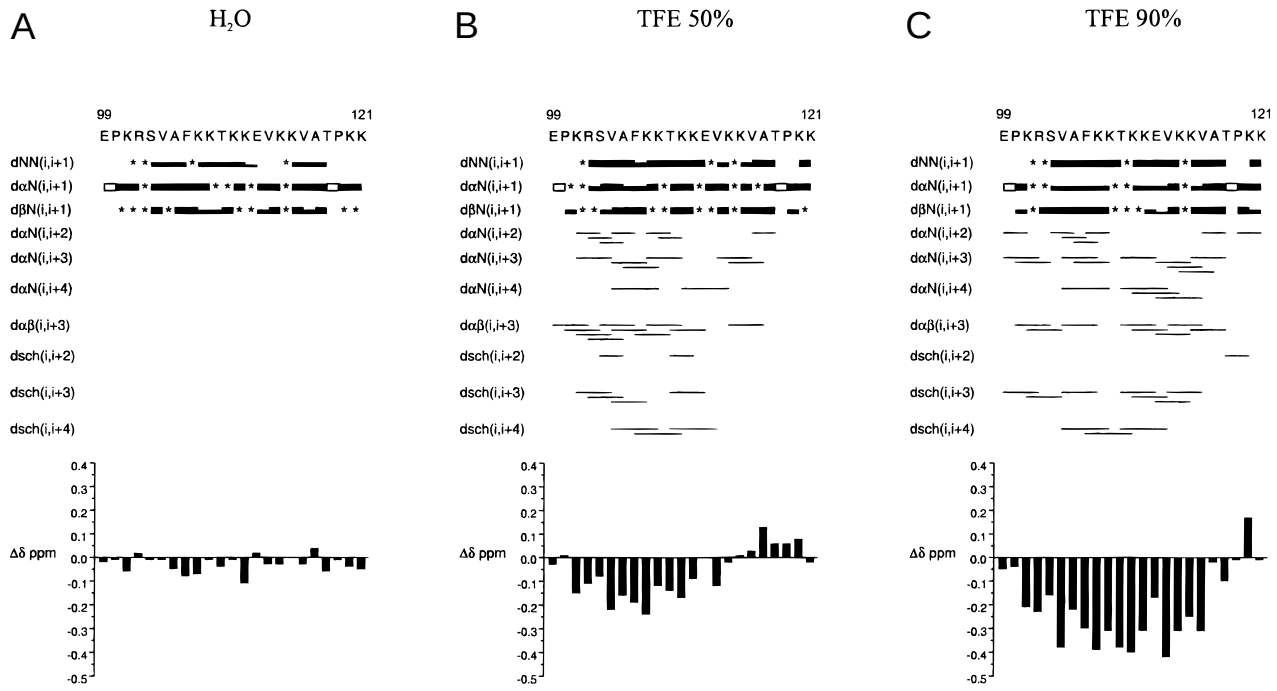


Fig. 3. Summary of the NOE connectivities of CH-1. The results (**A**) in aqueous solution, (**B**) in 50% TFE solution, and (**C**) in 90% TFE solution are presented (25 °C, pH 3.5). The thickness of the lines reflects the intensity of the sequential NOE connectivities, i.e., weak, medium, and strong. An asterisk (*) indicates an unobserved NOE connectivity due to signal overlapping, closeness to the diagonal or overlapping with the solvent signal. An open box indicates a $d\alpha\delta(i, i + 1)$ NOE connectivity, where $i + 1$ is proline. dsch indicates NOE connectivities involving side chains. The conformational shifts with respect to random coil values, $C^{\alpha}H\ \Delta\delta$, are shown as a function of the sequence number (bottom).

Connectivities $\alpha N(i, i + 4)$, which distinguish α helices from 3_{10} helices, were only observed between Val104 and Val116, while, according to the $\alpha N(i, i + 3)$ connectivities and the conformational shifts, the helical structure could extend up to Pro100. This may indicate that helix N-terminal sequence PKRS is in a 3_{10} helical conformation. Such a possibility is supported by structure calculations.

It is interesting to note the series of side-chain–side-chain medium cross peaks between the ring protons of Phe106 and the methylene protons of Lys110 (Fig. 2; Table 1). Side chains in peptides in solution are often quite mobile, so that NOE cross peaks involving side chains are difficult to detect (Muñoz et al., 1995). The high number of connectivities between the aromatic protons of Phe106 and the aliphatic protons of Lys110, therefore, indicates the existence of an interaction between the side chains of these two residues efficient enough to keep them close and less mobile (Fig. 4B). This conclusion is further supported by distinct chemical shifts of the diastereotopic protons $H^{\epsilon 2}$ and $H^{\epsilon 3}$ of Lys110.

Structure of the N-terminal TPKK motif

The peptide has a TPKK sequence in the C-terminal end. It has been pointed out (Suzuki et al., 1993) that a high proportion of (S/T)PXX sequences in crystal structures fold into a type (I) β-turn, which is stabilized by a hydrogen bond between the CO of residue (i) and the NH of residue $i + 3$. Many such β-turns present an additional hydrogen bond between the side-chain oxygen of (Ser/Thr) and either the NH of the $i + 2$ residue (the σ-turn type) or the NH of the $i + 3$ residue (the τ-turn type). The peptide bond Thr118–Pro119 is predominantly in a *trans* conformation as indicated by the presence of strong connectivities between Pro $H\delta\delta'$ and Thr $H\alpha$ (Hinck et al., 1993) (Fig. 3). A connectivity between Pro $H\alpha$ and Thr $H\alpha$, characteristic of the *cis* conformation, was not observed.

In H_2O, the sequence TPKK does not show NOE connectivities characteristic of turns or other folded structures. In 50% TFE, NOE connectivities characteristic of turn conformations are observed: there is a strong connectivity NH$(i + 2)$-NH$(i + 3)$ characteristic of β-turns and a connectivity NH$(i + 2)$-Pro$H\beta$ characteristic of β- or σ-turns. In 90% TFE, other connectivities appear, NH$(i + 2)$-Pro$H\gamma$, NH$(i + 2)$-Proγ' and NH$(i + 2)$-Proδ, also characteristic of β- or σ-turns. In 90% TFE, a connectivity NH$(i + 2)$-Thr$C\gamma$H3, which could indicate the presence of a hydrogen bond between the OH group of Thr and the backbone amide of Lys $i + 2$, is also observed (Table 2).

The connectivities observed between the NH of residue $i + 2$ and the β, γ, and δ protons of Pro are those expected for Type(I) β-turns. The σ-turn can coexist with a Type(I) β-turn to give a $\beta\sigma$-turn with two hydrogen bonds. It is not, however, possible to decide whether in CH-1 the turn conformation of the TPKK sequence is stabilized simultaneously by two hydrogen bonds or whether the β- and σ-turns are in rapid interconversion. Evidence for the presence of a σ-turn in 90% DMSO solution was previously obtained for motifs starting with Ser in peptides SPRKSPRK and GSPKKSPRK, but not for the peptide TPRK (Suzuki et al., 1993). Our results suggest that the σ-turn conformation is also

Table 1. *Summary of side-chain to side-chain and side-chain to main-chain NOE connectivities other than those expected for regular helices ($\alpha\beta_{i,i+3}$) found for the CH-1 peptide in 90% TFE* [a]

	Residue 1	Residue 2	Intensity
	γGlu99	NHArg102	Medium
	αLys101	γH_3Val104	Strong
*	αLys101	γ′H_3Val104	Strong
*	βLys101	NHVal104	Medium
+	δLys101	NHVal104	Weak
+	βArg102	NHVal104	Medium
+	γArg102	NHVal104	Weak
+	γArg102	NHAla105	Weak
+	βSer103	NHAla105	Weak
*	αVal104	δLys107	Medium
*	γH_3Val104	NHLys108	Medium
	oPhe106	βLys110	Medium
	oPhe106	β′Lys110	Medium
	oPhe106	γLys110	Medium
	oPhe106	γ′Lys110	Medium
*	oPhe106	δLys110	Medium
	oPhe106	ϵLys110	Medium
	oPhe106	ϵ′Lys110	Medium
	mPhe106	βLys110	Medium
	mPhe106	β′Lys110	Medium
	mPhe106	γLys110	Medium
	mPhe106	γ′Lys110	Medium
*	mPhe106	δLys110	Medium
	mPhe106	ϵLys110	Medium
	mPhe106	ϵ′Lys110	Medium
+	βThr109	NHLys111	Weak
*	αThr109	γGlu112	Strong
*	γH_3Thr109	NHGlu112	Medium
*	γH_3Thr109	NHVal113	Medium
	αLys110	γVal113	Strong
	αLys110	γ′Val113	Strong
	αGlu112	γLys115	Weak
	αGlu112	δLys115	Strong
	αGlu112	δ′Lys115	Medium
	αGlu112	αGlu112	Medium
	βVal113	NHVal116	Weak
	γH_3Thr118	NHLys120	Medium
	γPro119	NHLys120	Weak
	γ′Pro119	NHLys120	Weak
	δPro119	NHLys120	Weak

[a] *, NOE cross peaks that are also observed in 50% TFE; +, NOE cross peaks observed only in 50% TFE.

Table 2. *NOE connectivities observed for the sequence TPKK in 50% and 90% TFE solution* [a]

	Residue 1	Residue 2	50% TFE	90% TFE
σ	αThr118	NHLys120	*	*
σ	βThr118	NHLys120	—	—
σ	γH_3Thr118	NHLys120	*	Medium
β/σ	βPro119	NHLys120	Medium	Medium
β/σ	β′Pro119	NHLys120	*	*
β/σ	γPro119	NHLys120	*	Weak
β/σ	γ′Pro119	NHLys120	*	Weak
β/σ	δPro119	NHLys120	*	Weak
β/σ	δ′Pro119	NHLys120	*	—
β	αPro119	NHLys121	—	Very weak
β	NHLys120	NHLys121	Strong	Strong

[a] σ, NOE cross peaks expected for σ-turns; β, NOE cross peaks expected for β-turns; β/σ, NOE cross peaks expected for either β- or σ-turns; *, unobserved NOE connectivity due to signal overlapping.

accessible to sequence motifs starting with Thr. The turn structure formed by the sequence TPKK must have limited rotational freedom about the backbone bonds connecting it to the rest of the peptide, as indicated by the presence of strong or medium connectivities between the amide proton of Thr118 and βAla117, αVal116, and αLys115 (Fig. 3).

Structure calculations

Structure calculations were performed on the basis of the NOE cross correlations observed in 90% TFE, since in these conditions the population of folded structures is higher.

A set of 77 distance constraints, composed of 22 sequential and 55 medium-range constraints, was used to calculate the three-dimensional (3D) structures of the peptide. A number of 50 structures were generated for the peptide with the distance geometry program DIANA (Güntert & Wüthrich, 1991). The 15 structures best fitting the experimental data were chosen. The global RMS deviation of the backbone atoms for this set of structures, excluding the first and the last residues, was 0.161 ± 0.061 nm. The maximum NOE violation was 0.034 nm. A superposition of the backbones of the 15 selected structures is shown in Figure 4A. Figure 4B shows one of the calculated structures. The region spanning from Pro100 to Ala117 adopts a well-defined helical structure. The region from Val104 to Ala117 is α-helical, while the residues from Pro100 to Ser103 are in a 3_{10} helical conformation.

The helix bends slightly around the sequence KKT (107–109). This effect arises from the strong interaction between the side chains of Phe106 and Lys110 (Fig. 4B).

The C-terminal sequence, TPKK, adopts a turn conformation. Some of the calculated structures show the hydrogen bond typical of the β-turn, while others have the σ-type hydrogen bond. In the structure with the third lowest free energy, both hydrogen bonds are present in a βσ-turn conformation. A superposition of the backbone of the best 25 structures is shown in Figure 5A. The TPKK sequence in βσ-turn conformation is shown in more detail in Figure 5B.

Discussion

We have shown that the CH-1 peptide in TFE solution presents a significant population of well-defined structure. The longest segment, 100–116, adopts a helical structure. The first four residues of the helical region, PKRS, most probably adopt a 3_{10} helical structure, while the rest of the residues are structured as an α-helix. The peptide is essentially a random coil in H_2O, although the observation of $dNN(i, i+1)$ connectivities is indicative of the presence of a series of folded structures (Dyson et al., 1988), which in TFE solution would be in equilibrium with a significant population of regular helical structures. It is interesting to note that the N-terminus of the helical region, from Pro100 to Lys111, has a higher pro-

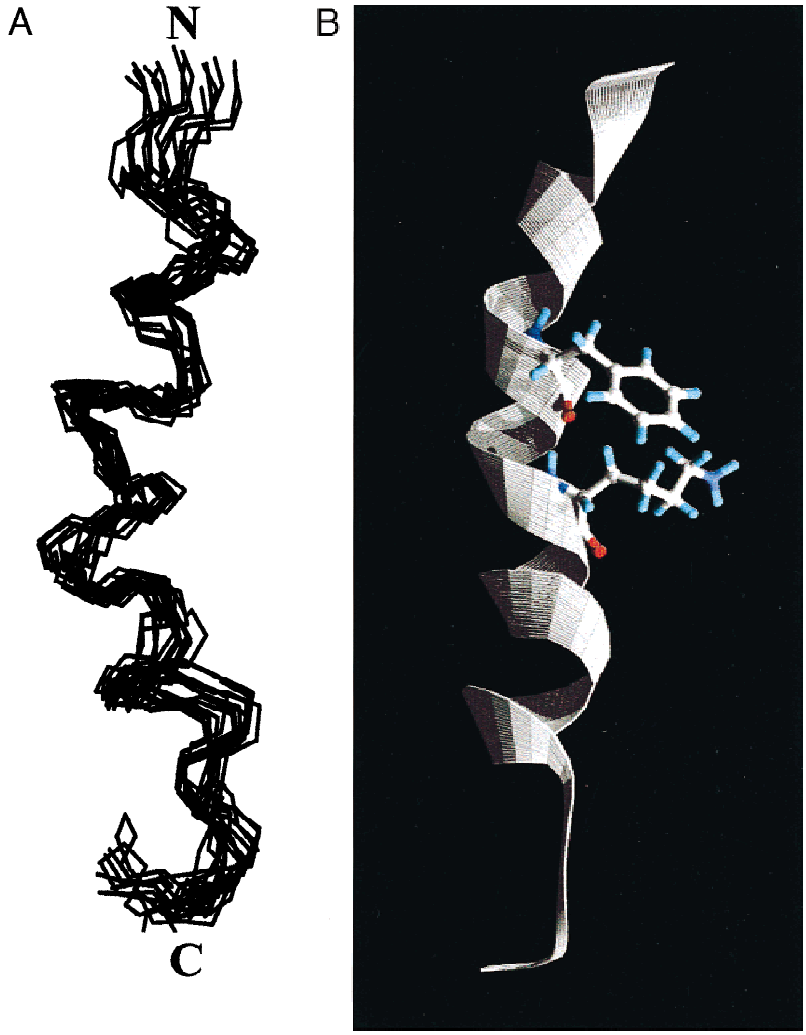


Fig. 4. Structure of the CH-1 peptide. **A:** Superposition of the best 15 structures of the CH-1 peptide calculated by distance geometry methods on the basis of the distance constraints derived from observed NOE cross correlations in 90% TFE solution. **B:** Illustration of the interaction between the Phe106 and Lys110 side chains.

pensity to become helical than the C-terminus, from Glu112 to Ala117. This is manifested by the magnitudes of the negative conformational shifts in 50 and 90% TFE solution. While in 50% TFE, the N-terminus is already largely helical, the TFE has to be increased up to 90% to observe a significant helical population in the C-terminus (Fig. 3). The lack of defined secondary structure when CH-1 is studied in absence of TFE is correctly predicted by the helix prediction program AGADIR, which has been parameterized on the basis of peptide conformations in aqueous solution (Muñoz & Serrano, 1994). When the method of Chou and Fasman (1974), which is based on protein data, is used, the peptide is predicted to be α-helical.

In the middle of the peptide, there is a TKKE sequence. (T/S) XX(E/Q) sequences can adopt the capping box conformation (Harper & Rose, 1993). In CH-1, the TKKE motif is, however, in α-helical conformation, according to the observed NOEs and the negative conformational shifts. The α-helical structure of the sequence TKKE is definitely confirmed by distance geometry structural calculations on the basis of distance constraints derived from NOE cross peaks. This result favors the idea that TXXE sequences are not necessarily stop signals. Jiménez et al. (1994) showed that a capping box signal could be bypassed by favorable interactions between the side chain of a capping box residue and a residue located before it. In CH-1, the series of $(i, i + 4)$ side-chain–side-chain cross peaks between the ring protons of Phe106 and the methylene protons of Lys110, the first Lys in the TKKE sequence (Fig. 2; Table 1), indicates a strong interaction between the side chains of these two residues that very likely contributes to the stability of the α-helical structure (Muñoz et al., 1995).

The peptide ends with a TPKK motif. It has been previously shown that sequences of the kind (S/T)P(K/R)(K/R) can adopt turn conformations in solution in peptides containing one or two such repeats (Suzuki et al., 1993). We have shown that the TPKK motif also adopts a turn conformation when in continuity with a native H1 sequence. NOE connectivities, both for the classical type (I) β-turn and for the σ-turn, are observed. β- and σ-turns could be in dynamic interconversion or combine in a $\beta\sigma$-turn with two hydrogen bonds (Fig. 6). NOE connectivities characteristic of a σ-turn were not previously observed for the peptide TPRK in 90% DMSO solution. They were, however, present in peptides

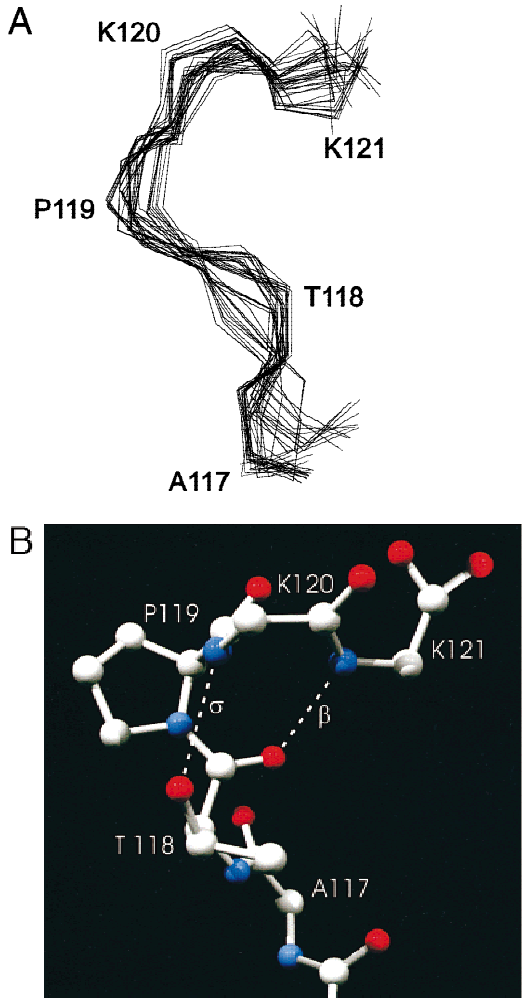


Fig. 5. Turn structure adopted by the C-terminal end sequence TPKK. **A:** Superposition of the best 25 calculated structures of the TPKK motif. **B:** Illustration of one of the 25 computed structures. Broken lines indicate possible hydrogen bonds; β-type (β) and σ-type (σ).

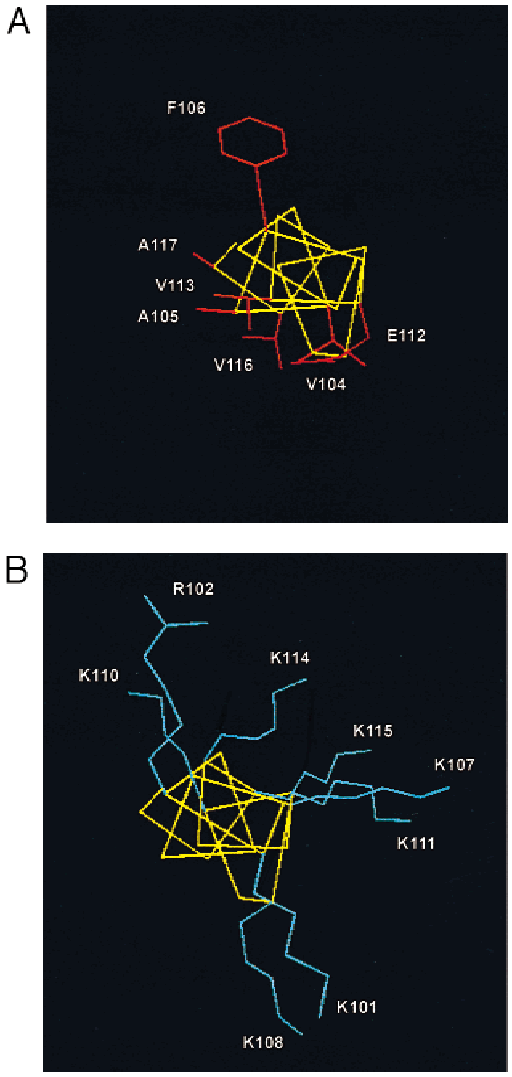


Fig. 6. End view down the helix axis of one of the calculated structures of CH-1. The helical region from Lys101 to Val117 is represented. It shows the amphipathic character of the helix, with the positively charged residues on one face of the helix and the hydrophobic residues and Glu112 on the other face.

starting with Ser and containing two repeats of the motif (Suzuki et al., 1993). Our results suggest that TPXX sequences can also adopt the σ-turn conformation, which therefore would not be exclusive to sequences starting with Ser. This is interesting, as TPXX sequences are frequently found in H1 subtypes. Models for the binding of (S/T)PXX motifs to the DNA minor groove, in either β- or σ-turn conformation have been reported (Suzuki, 1989; Reeves & Nissen, 1990; Suzuki et al., 1993).

The effects of TFE on polypeptide chain conformation are diverse. In CH-1, both a helix and a turn are stabilized by TFE. It has been now well established that for helical structures, TFE reveals the conformational biases of the primary sequence. Possible mechanisms by which TFE affects polypeptide structure include enhancement of internal hydrogen bonding, the disruption of water structure and preferential solvation of certain groups of the polypeptide chain. In previous studies with entire H1 molecules and model peptides, repulsion between positively charged residues was assumed to counteract the tendency toward the α-helical structure (Walters & Kaiser, 1985; Clark et al., 1988; Hill et al., 1989; Johnson et al., 1994). Organic solvents as TFE decrease the dielectric constant of the medium. This effect should increase solvation of the Lys positive charges by counterions and could lead to a more efficient neutralization of the peptide charge (Clark et al., 1988). Charge neutralization may not, though, be sufficient to fully structure C-terminal peptides. The amphipathic nature of CH-1 suggests that in condensed chromatin the hydrophobic face of the peptide could participate in protein–protein interactions. In water mixtures, the CF_3 substituent is hydrophobic. It has been suggested that TFE could cooperatively associate with the hydrophobic surface in amphipathic helices (Buck, 1998), thus mimicking a dehydrated environment.

The helical region has a marked amphipathic character, with all basic residues on one face of the helix and all the hydrophobic ones, together with Glu112, on the opposite face (Fig. 6). The amphipathic character of the peptide may influence its conformational dynamics upon interaction with the DNA, as it has been shown that DNA-induced α-helix formation in short synthetic pep-

tides composed of Ala and Lys requires that the peptide can form amphipathic helices (Johnson et al., 1994). In the case of CH-1, difference CD could not be used to investigate the structure of the peptide bound to DNA, as changes in DNA structure did not allow to attribute with certainty the spectral changes to the peptide.

The concentration of the basic residues on one face of the helix defines a potential DNA-binding site. This kind of arrangement could contribute to the clustering of the positive charges over a short region of the DNA and, since the peptide is adjacent to the globular domain, it could affect the conformation of the initial part of the linker DNA at the exit of the nucleosome. This could have important consequences in those cases where gene regulation is mediated by the tail domains (Lee & Archer, 1998; Dou et al., 1999).

It has been proposed that the C-terminal domain of H1 may contain α-helical elements (Clark et al., 1988; Hill et al., 1989) and β-turns (Erard et al., 1990) when bound to the DNA. Our results with CH-1, where both structural elements are present, support the view that both kinds of secondary structure coexist in the C-terminal domain. It is to be noted that α-helices generally bind to the DNA major groove, while the (S/T)P(K/R)(K/R) motifs are thought to bind to the minor groove (Suzuki, 1989; Erard et al., 1990). This suggests that the H1 C-terminal domain could bind to both the major and minor grooves of the DNA.

Materials and methods

Peptide synthesis

The peptide Ac-EPKRSVAFKKTKKEVKKVATPKK (CH-1) was synthesized by standard methods (Neosystem Laboratoire, Strasbourg, France). Peptide homogeneity was determined by high-performance liquid chromatography on Nucleosil C18. The peptide composition was confirmed by amino acid analysis and the molecular mass was checked by mass spectrometry. The sequence of the peptide corresponds to residues 99 to 121 at the C-terminus of histone H1°. The sequence is common to the mouse and the rat and presents two conservative substitutions in humans, at positions 102 (Arg $\rightarrow$ Lys) and 113 (Val $\rightarrow$ Ile). The peptide was acetylated to remove the dipole destabilization effect.

CD spectroscopy

Samples for CD spectroscopy were 1.23×10^{-4} M in the peptide and 5 mM in sodium phosphate buffer, pH 3.5. Samples in aqueous and mixed solvent with different ratios (v/v) of trifluoroethanol/H_2O were prepared. Spectra were obtained on a Jasco J-715 CD spectrometer in 1 mm cells at 5 °C. The results are expressed as mean residue molar ellipticities $[\theta]$. The helical content was estimated from the ellipticity value at 222 nm ($[\theta]_{222}$) according to the empirical equation of Chen et al. (1974):

$$\% \text{ helical content} = 100([\theta]_{222}/-39{,}500 \times (1 - 2.57/n)),$$

where n is the number of peptide bonds in the helix. The helical length was determined from the NMR data.

^{1}H NMR spectroscopy

Samples were routinely prepared as ~2.2 mM solutions of the peptide in 90% H_2O/10% ^2H_2O, 5 mM phosphate buffer, 70 mM NaCl. The pH was adjusted to 3.5 with minimal amounts of HCl or NaOH in water. Sodium 3-trimethylsilyl (2,2,3,3-$^2H_4$) propio-

Table 3. *Chemical shifts of the CH-1 peptide in 90% TFE solution, pH 3.5, 25 °C*

	NH	CαH	CβH	Cβ'H	CγH	Cγ'H	CδH	Cδ'H	Others
Glu99	7,70	4,59	2,08	2,11	2,44	2,44	—	—	
Pro100	—	4,38	1,95	2,40	2,06	2,14	3,76	3,92	
Lys101	8,11	4,11	1,94	2,01	1,49	1,49	1,75	1,75	$H_2$3,02
Arg102	8,05	4,11	1,93	1,93	1,68	1,76	3,22	3,22	NϵH7,22
Ser103	7,97	4,31	4,08	4,08	—	—	—	—	
Val104	7,70	3,74	2,26	—	CγH$_3$1,03	Cγ'H$_3$1,14	—	—	
Ala105	7,94	4,10	CβH$_3$1,54	—	—	—	—	—	
Phe106	7,28	4,33	3,31	3,34	—	—	—	—	2,6H7,28 3,5H7,33 4H7,29
Lys107	8,15	3,93	1,97	2,10	1,55	1,55	1,78	1,78	$H_2$3,03 NϵH6,42
Lys108	8,78	4,01	1,99	1,99	1,53	1,63	1,75	1,75	CϵH$_2$2,98 NϵH6,42
Thr109	8,04	3,97	4,35	—	CγH$_3$1,30	—	—	—	
Lys110	8,03	3,92	1,79	1,84	1,35	1,44	1,62	1,62	CϵH2,84 Cϵ'H2,92 NϵH6,42
Lys111	7,89	4,01	1,99	2,03	1,50	1,50	1,73	1,73	CϵH$_2$3,00 NϵH6,42
Glu112	8,08	4,18	2,22	2,32	2,55	2,55	—	—	
Val113	8,32	3,70	2,19	—	CγH$_3$0,98	Cγ'H$_3$1,10	—	—	
Lys114	7,92	4,01	2,00	2,03	1,51	1,51	1,71	1,71	CϵH$_2$3,00 NϵH6,42
Lys115	7,93	4,07	2,05	2,07	1,47	1,47	1,72	1,72	CϵH$_2$2,98 NϵH6,42
Val116	7,99	3,81	2,28	—	CγH$_3$1,01	Cγ'H$_3$1,12	—	—	
Ala117	8,34	4,30	CβH$_3$1,48	—	—	—	—	—	
Thr118	7,44	4,51	4,23	—	CγH$_3$1,37	—	—	—	
Pro119	—	4,41	1,99	2,35	2,06	2,14	3,72	4,02	
Lys120	7,66	4,49	1,76	1,99	1,56	1,56	1,72	1,72	CϵH$_2$3,07 NϵH6,42
Lys121	7,49	4,31	1,77	1,91	1,47	1,47	1,72	1,72	CϵH$_2$3,02 NϵH6,42

nate (TPS) was used as internal standard. Spectra were also obtained in the presence of either 50 or 90% deuterated TFE.

Spectra were recorded in a Bruker AMX-600 spectrometer. All 2D spectra were recorded in the phase-sensitive mode using time-proportional phase incrementation (Marion & Wüthrich, 1983) with presaturation of the water signal. Correlation spectroscopy (Aue et al., 1976) and NOE spectroscopy (NOESY) (Kumar et al., 1980) spectra were obtained using standard phase-cycling sequences. Short mixing times of 150–200 ms were used in the NOESY experiments to avoid spin diffusion. Spectra in aqueous solution were obtained at 5 °C, while spectra in TFE solution were obtained at 25 °C. Total correlation spectroscopy (Bax & Davis, 1985) spectra were acquired using the standard MLEV16 spin-lock sequence with a mixing time of 80 ms. The phase shift was optimized for every spectrum.

The assignments of the ^{1}H-NMR spectra of the peptide in aqueous solution or in the presence of different concentrations of trifluoroethanol were performed by standard 2D sequence-specific methods (Wüthrich et al., 1984; Wüthrich, 1986). The chemical shift assignments of the CH-1 peptide in 90% TFE solution are shown in Table 3.

Quantification of the helix populations was carried out on the basis of the well-established up-field shifts of the $C^{\alpha}H$ δ-values upon helix formation, according to Jiménez et al. (1993). The average helical population per residue was obtained by dividing the average conformational shift, $\Delta\delta = \sum(\delta_i^{obs} - \delta_i^{RC})/n$, by the shift corresponding to 100% helix formation. Random coil values δ^{RC} were those given by Wishart et al. (1995). The random coil values used for Glu99 and Thr118 were those given for amino acids followed by Pro. A value of -0.39 ppm was used as the shift for 100% helix formation (Wishart et al., 1991). The helical length n was determined on the basis of NOE cross peaks and conformational shifts, and confirmed by structure calculations.

Structure calculations

Calculations of peptide structures were carried out with the program DIANA (Güntert & Wüthrich, 1991). Distance constraints were derived from the 150 ms NOESY spectrum acquired in 90% TFE at 25 °C, pH 3.5. The intensities of the observed NOEs were evaluated in a qualitative way and translated into upper limit distant constraints according to the following criteria: strong NOEs were set to distances lower than 0.3 nm; medium, lower that 0.35 nm, and weak, lower than 0.45 nm. Pseudo atom corrections were set to the sum of the van der Waals radii. The ψ angles were constrained to the range $-180°$ to $0°$.

Acknowledgments

This work was supported by the Ministerio de Educación y Cultura, (DGICYT, PB95-0608), the Generalitat de Catalunya (1998SGR 00072), and the EU (CEE B104-97-2086).

References

Aue WP, Bertholdi E, Ernst RR. 1976. Two-dimensional spectroscopy. Applications to magnetic resonance. *J Chem Phys 64*:2229–2246.

Bax A, Davis DG. 1985. Practical aspects of two-dimensional spectroscopy. *J Magn Reson 65*:355–360.

Bouvet P, Dimitrov S, Wolffe AP. 1994. Specific regulation of *Xenopus* chromosomal 5S rRNA gene transcription in vivo by histone H1. *Genes Dev 8*:1147–1159.

Buck M. 1998. Trifluoroethanol and colleagues: Cosolvents come of age. Recent studies with peptides and proteins. *Q Rev Biophys 31*:297–355.

Cerf C, Lippens G, Muyldermans S, Sergers A, Ramakrishnan V, Wodak SJ, Hallenga K, Wyns L. 1993. Homo- and heteronuclear two-dimensional NMR studies of the globular domain of histone H1: Sequential assignment and secondary structure. *Biochemistry 32*:11345–11351.

Chen YH, Yang JT, Chau KH. 1974. Determination of the helix and β-form proteins in aqueous solution by circular dichroism. *Biochemistry 13*:3350–3359.

Chou PY, Fasman GD. 1974. Prediction of protein conformation. *Biochemistry 13*:222–245.

Clark DJ, Hill CS, Martin SR, Thomas JO. 1988. α-Helix in the carboxy-terminal domain of histones H1 and H5. *EMBO J 7*:69–75.

Clore GM, Gronenborn AM, Nilges M, Sukumaran DK, Zarbock J. 1987. The polypeptide fold of the globular domain of histone H5 in solution. A study using nuclear magnetic resonance, distance geometry and restrained molecular dynamics. *EMBO J 6*:1833–1842.

Dasso M, Dimitrov S, Wolffe AP. 1994. Nuclear assembly is independent of linker histones. *Proc Natl Acad Sci USA 91*:12477–12481.

Dou Y, Mizzen CA, Abrams M, Allis CD, Gorovsky MA. 1999. Phosphorylation of linker histone H1 regulates gene expression in vivo by mimicking H1 removal. *Mol Cell 4*:641–647.

Dyson J, Rance M, Houghten RA, Wright PE, Lerner RA. 1988. Folding of immunogenic peptide fragments of proteins in water solution. II. The nascent helix. *J Mol Biol 201*:201–217.

Erard M, Lakhdar-Ghazal F, Amalric F. 1990. Repeat peptide motifs which contain β-turns and modulate DNA condensation in chromatin. *Eur J Biochem 191*:19–26.

Güntert P, Wüthrich K. 1991. Improved efficiency of protein structure calculations from NMR using the program DIANA with redundant dihedral angle constraints. *J Biomol NMR 1*:447–556.

Harper ET, Rose GD. 1993. Helix stop signals in proteins and peptides: The capping box. *Biochemistry 32*:7605–7609.

Hartman PG, Chapman GE, Moss T, Bradbury EM. 1977. Studies on the role and mode of operation of the very-lysine-rich histone H1 in eukaryotic chromatin: The three structural regions of the histone H1 molecule. *Eur J Biochem 77*:45–51.

Hill CS, Martin SR, Thomas JO. 1989. A stable α-helical element in the carboxy-terminal domain of free and chromatin-bound histone H1 from sea urchin sperm. *EMBO J 8*:2591–2599.

Hinck AP, Eberhardt ES, Markley JL. 1993. NMR strategy for determining Xaa-Pro peptide bond configurations in proteins: mutants of staphylococcal nuclease with altered configuration at proline-117. *Biochemistry 32*:11810–11818.

Jiménez MA, Bruix M, González C, Blanco FJ, Nieto JL, Herranz J, Rico M. 1993. CD and ^{1}H-NMR studies on the conformational properties of peptide fragments from the C-terminal domain of thermolysin. *Eur J Biochem 211*:569–581.

Jiménez MA, Muñoz V, Rico M, Serrano L. 1994. Helix stop and start signals in peptides and proteins. The capping box does not necessarily prevent helix elongation. *J Mol Biol 242*:487–496.

Johnson NP, Lindstrom J, Baase WA, von Hippel PH. 1994. Double-stranded DNA templates can induce α-helical conformation in peptides containing lysine and alanine: Functional implication for leucine zipper and helix-loop-helix transcription factors. *Proc Natl Acad Sci USA 91*:4840–4844.

Khochbin S, Wolffe AP. 1994. Developmentally regulated expression of linker-histone variants in vertebrates. *Eur J Biochem 225*:501–510.

Kumar A, Ernst RR, Wüthrich K. 1980. A two-dimensional nuclear overhauser enhancement (2D-NOE) experimental for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. *Biochem Biophys Res Commun 95*:1–6.

Lee HL, Archer TK. 1998. Prolonged glucocorticoid exposure dephosphorylates histone H1 and inactivates the MMTV promoter. *EMBO J 17*:1454–1466.

Marion D, Wüthrich K. 1983. Application of phase-sensitive two-dimensional correlated spectroscopy (COSY) for measurement of proton-proton spin-spin coupling constants. *Biochem Biophys Res Commun 113*:967–974.

Morán F, Montero F, Azorín F, Suau P. 1985. Condensation of DNA by the C-terminal domain of histone H1. A circular dichroism study. *Biophys Chem 22*:125–129.

Muñoz V, Serrano L. 1994. Elucidating the folding problem of helical peptides using empirical parameters. *Nat Struct Biol 1*:399–409.

Muñoz V, Serrano L, Jiménez MA, Rico M. 1995. Structural analysis of peptides encompassing all α-helices of three α/β parallel proteins: Che-Y, Flavodoxin and P21-Ras: Implications for α-helix stability and the folding of α/β parallel proteins. *J Mol Biol 247*:648–669.

Ramakrishnan V, Finch JT, Graziano V, Lee PL, Sweet RM. 1993. Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. *Nature 362*:219–223.

Reeves R, Nissen MS. 1990. The A.T-DNA binding domain of mammalian high mobility group I chromosomal proteins. *J Biol Chem 265*:8573–8582.

Richardson JS, Richardson DC. 1988. Amino acid preferences for specific locations of the ends of α helices. *Science 240*:1648–1652.

Shen X, Gorovsky MA. 1996. Linker histone H1 regulates specific gene expression but not global transcription in vivo. *Cell 86*:475–483.

Shen X, Yu L, Weir JW, Gorovsky MA. 1995. Linker histones are not essential and affect chromatin condensation in vivo. *Cell 82*:47–56.

Suzuki M. 1989. SPKK, a new nucleic acid-binding unit of protein found in histone. *EMBO J 8*:797–804.

Suzuki M, Gerstein M, Johnson T. 1993. An NMR study on DNA-binding SPKK motif and a model for its interaction with DNA. *Protein Eng 6*:565–574.

Vermaak D, Steinbach OC, Dimitrov S, Rupp RAW, Wolffe AP. 1998. The globular domain of histone H1 is sufficient to direct specific gene repression in early *Xenopus* embryos. *Curr Biol 8*:533–536.

Walters L, Kaiser ET. 1985. Design of DNA-binding peptides: Stabilization of α-helical structure by DNA. *J Am Chem Soc 107*:6422–6424.

Wishart DS, Brigam CG, Holm A, Hodges RS, Sykes BD. 1995. ^{1}H, ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigations of the nearest-neighbor effects. *J Biomol NMR 5*:67–81.

Wishart DS, Sykes BD, Richards FM. 1991. Relationship between nuclear magnetic resonance chemical shift and secondary structure. *J Mol Biol 222*:311–333.

Wolffe AP, Khochbin PS, Dimitrov S. 1997. What do linker histones do in chromatin? *BioEssays 19*:249–255.

Wüthrich K. 1986. *NMR of protein and nucleic acids*. New York: John Wiley and Sons.

Wüthrich K, Billeter M, Braun W. 1984. Polypeptide secondary structure determination by nuclear magnetic resonance observation of short proton-proton distances. *J Mol Biol 180*:715–740.

Zlatanova J, Van Holde K. 1992. Histone H1 and transcription: Still an enigma? *J Cell Sci 103*:889–895.