OBJECTIUS

OBJECTIUS

Aquesta tesi s'emmarca dins l'objectiu general de caracteritzar l'estructura dels dominis N- i C-terminals de la histona H1. De forma més concreta, els objectius són els següents:

1- Estudiar l'estructura induïda per TFE de pèptids dels dominis N- i C-terminals de varis subtipus de la histona H1 mitjançant DC, FTIR i RMN-2D de protó.

2- Estudiar l'estructura de pèptids dels dominis N- i C-terminals de la histona H1 quan formen part de complexos amb DNA de doble cadena mitjançant FTIR.

L'aproximació i els mètodes pels que hem optat vénen determinats per les següents peculiaritats de l'objecte d'estudi:

Els dominis terminals de la histona H1 es troben desestructurats en solució aquosa, però es creu s'estructuren per a realitzar la seva funció d'unió al DNA internucleosomal i de condensació de la cromatina. Donada la complexitat del sistema en el que la histona H1 realitza la seva funció, difícilment simplificable i d'estructura encara no ben coneguda, és gairebé impossible abordar directament l'estudi detallat de l'estructura de la histona H1 formant part de la cromatina mitjançant les tècniques disponibles actualment.

Els dominis terminals de la histona H1 són altament repetitius en la seva seqüència, amb una gran riquesa en residus lisina, alanina i prolina. Per exemple, de 97 residus de què consta el domini C-terminal de la H1º de ratolí, 40 són de lisina, 16 d'alanina, 12 de prolina, 9 de valina, 7 de serina i 5 de treonina. Aquest fet dificulta l'anàlisi per RMN, tant per la dificultat de realitzar les assignacions, com pels solapaments en les ressonàncies, com pel trencament de l'assignació seqüencial degut a les prolines. Tot això provoca una limitació pràctica de la llargada dels pèptids estudiats.

L'elevada basicitat dels dominis N- i C-terminals, i de la majoria dels seus pèptids, provoca la formació d'un complex insoluble en presència de DNA. Això impossibilita l'ús del dicroïsme circular i la RMN per a analitzar l'estructura d'aquests dominis o pèptids units al DNA.

CAPÍTOL I

CD AND NMR ANALYSIS OF A C-TERMINAL DOMAIN PEPTIDE OF HISTONE H1º

INTRODUCTION

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).

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 and Van Holde, 1992; Khochbin and Wolffe, 1994; Wolffe et al., 1997b). 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 and 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 and Archer, 1998) or the activation or repression of specific genes in *Tethrahymena* (Dou et al., 1999), are, however, 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 might, however, 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 (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 (Vila and Suau, unpublished observations). 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.

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 HPLC 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 → Lys) and 113 (Val → Ile). The peptide was acetylated to remove the dipole destabilization effect.

Circular dichroism spectroscopy

Samples for circular dichroism 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.

^1H NMR spectroscopy

Samples were routinely prepared as approximately 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-$^2H_4$) propionate (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 2-dimensional spectra were recorded in the phase-sensitive mode using time-proportional phase incrementation (Marion and Wüthrich, 1983) with presaturation of the water signal.

COSY (Aue et al., 1976) and NOESY (Kumar et al., 1980) spectra were obtained using standard phase-cycling sequences. Short mixing times of 150-200 milliseconds 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. TOCSY (Bax and Davis, 1985) spectra were acquired using the standard MLEV16 spin-lock sequence with a mixing time of 80 milliseconds. 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 two-dimensional 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 1.

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, $\Delta\delta=\Sigma(\delta_i^{obs}-\delta_i^{RC})/n$, 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 p.p.m. was used as the shift for 100% helix formation (Wishart et al.,1991). The helical length, n, was determined on the basis of NOE crosspeaks and conformational shifts, and confirmed by structure calculations.

Structure calculations

Calculations of peptide structures were carried out with the program DIANA (Güntert et al., 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. ϕ angles were constrained to the range -180° to 0°.

RESULTS

Circular dichroism 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 nm 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 approximately 215 nm, characteristic of the random coil, was not observed, suggesting that a very small amount of structure could be present in water (Figure 1). Addition of TFE, which is known to stabilize peptide secondary structure, resulted in an increase in the negative ellipticity at 222 nm (Figure 1). The helical content of the peptide as a function of TFE concentration was estimated by the method of Chen et al. (1974) (Figure 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 non-sequential $\alpha N(i, i+3)$, $\alpha N(i, i+4)$, $\alpha\beta(i, i+3)$, 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 $\alpha N(i, i+1)$ NOE connectivities; (3) significant up-field shifting of the $C^{\alpha}H$ resonances relative to the random coil values.

Figure 2 shows selected regions of the two dimensional 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 II and III 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, $\Delta\delta = \delta_{observed} - \delta_{random\ coil}$. The peptide does not show a significant helical population in water. However, the presence of abundant medium NOE correlations between sequential amide protons (Figure 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 non-sequential NOEs, the intensity of the NN NOE crosspeaks, 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 ($H\gamma$) 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 and 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 $H\alpha$ (Hinck et al., 1993). A very weak connectivity between the Pro $H\alpha$ and the Thr $H\alpha$ 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 - 0.12ppm. 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 - 0.29ppm, 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.

Connectivities αN(i, i+4), which distinguish α helices from 3_{10} helices, were only observed between Val104 and Val116, while, according to the αN(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 (Figure 2, Table 2). 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 (Figure 4B). This conclusion is further supported by distinct chemical shifts of the diastereotopic protons $H^{\varepsilon 2}$ and $H^{\varepsilon 3}$ 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 oxigen 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 Hδδ' and Thr Hα (Hinck et al., 1993) (Figure 3). A connectivity between Pro Hα and Thr Hα, 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)-ProHβ characteristic of β or σ turns. In 90% TFE, other connectivities appear, NH(i+2)-ProHγ, NH(i+2)-Proγ' and

NH(i+2)-Proδ, also characteristic of β or σ turns. In 90% TFE, a connectivity NH(i+2)-ThrCγH3, that 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 3).

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 βσ-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 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 (Figure 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 structures of the peptide. A number of 50 structures were generated for the peptide with the distance geometry program DIANA (Güntert and Wüthrich, 1991). The 15 structures best fitting the experimental data were chosen. The global mean square root 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 Lys 110 (Figure 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 propensity 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% in order to observe a significant helical population in the C-terminus (Figure 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 and 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 and 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 crosspeaks. 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 crosspeaks between the ring protons of Phe106 and the methylene protons of Lys110, the first Lys in the TKKE sequence (Figure 2, Table 2), 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 βσ-turn with two hydrogen bonds (Figure 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 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 and 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 towards the α-helical structure (Walters and 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 amphypathic 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 (Figure 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 peptides composed of Ala and Lys requires that the peptide can form amphipathic helices (Johnson et al., 1994). In the case of CH-1, difference circular dichroism 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 and 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.

	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	CεH$_2$3,02	NεH6,42	
Arg102	8,05	4,11	1,93	1,93	1,68	1,76	3,22	3,22	NεH7,22		
Ser103	7,97	4,31	4,08	4,08	…..	…..	…..	…..			
Val104	7,70	3,74	2,26	…..	CγH$_3$1,03	Cγ'H$_3$1,14	…..	…..			
Ala105	7,94	4,10	CβH$_3$1,54	…..	…..	…..	…..	…..			
Phe106	7,28	4,33	3,31	3,34	…..	…..	…..	…..	2,6H7,28	3,5H7,33	4H7,29
Lys107	8,15	3,93	1,97	2,10	1,55	1,55	1,78	1,78	CεH$_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	

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

	Residue 1	Residue 2	Intensity
	γGlu99	NHArg102	Medium
	αLys101	γH_3Val 104	Strong
*	αLys101	γ'H_3Val 104	Strong
*	βLys101	NHVal104	Medium
+	δLys101	NHVal104	Weak
+	βArg102	NHVal104	Medium
+	γArg102	NHVal104	Weak
+	γArg102	NHAla105	Weak
+	βSer103	NHAla105	Weak
*	αVal 104	δLys107	Medium
*	γH_3Val 104	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	εLys115	Medium
	βVal113	NHVal116	Weak
	γH_3Thr118	NHLys120	Medium
	γPro119	NHLys120	Weak
	γ'Pro119	NHLys120	Weak
	δPro119	NHLys120	Weak

Table 2. 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.

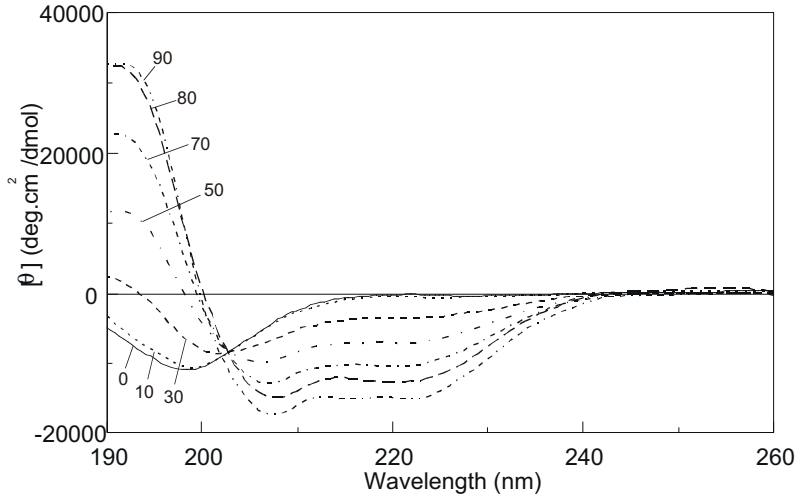
(*) NOE crosspeaks that are also observed in 50% TFE.

(+) NOE crosspeaks observed only in 50% TFE.

	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

Table 3. NOE connectivities observed for the sequence TPKK in 50% and 90% TFE solution. σ, NOE crosspeaks expected for σ-turns; β, NOE cross-peaks expected for β-turns; β/σ, NOE cross-peaks expected for either β- or σ-turns. An asterisk (*) indicates an unobserved NOE connectivity due to signal overlapping.

a



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

Figure 1. TFE-dependent conformational transition of the CH-1 peptide measured by CD. (a) Far-UV 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.

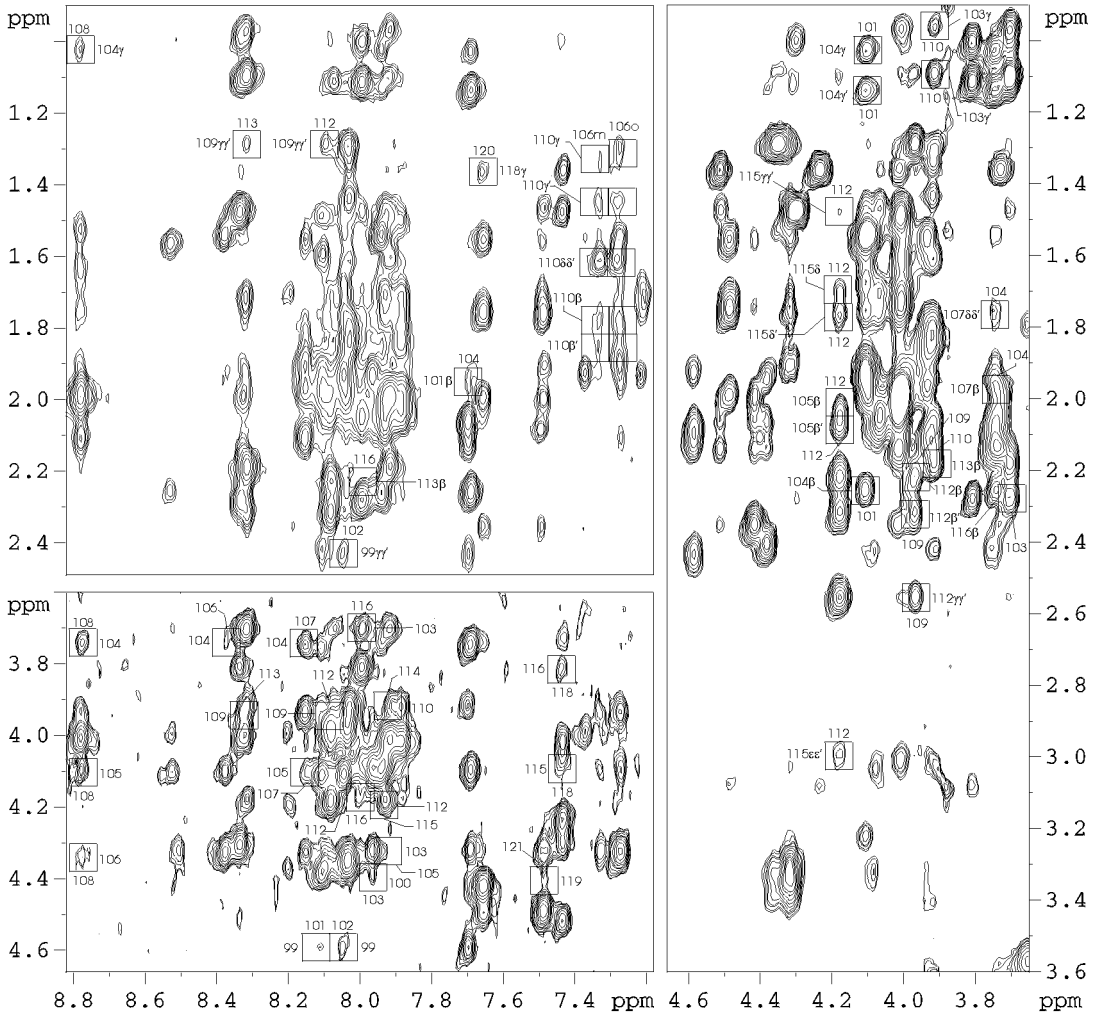


Figure 2. Selected regions of the 2-dimensional NOE spectrum of CH-1. The peptide was 2.2 mM in 90% TFE, pH 3.5, 25 °C, mixing time 150 ms. Medium-range NOE connectivities are boxed. Left, NOE correlations of aliphatic protons (bottom, $C^{\alpha}H$) with amide or aromatic resonances. Right, NOE correlations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

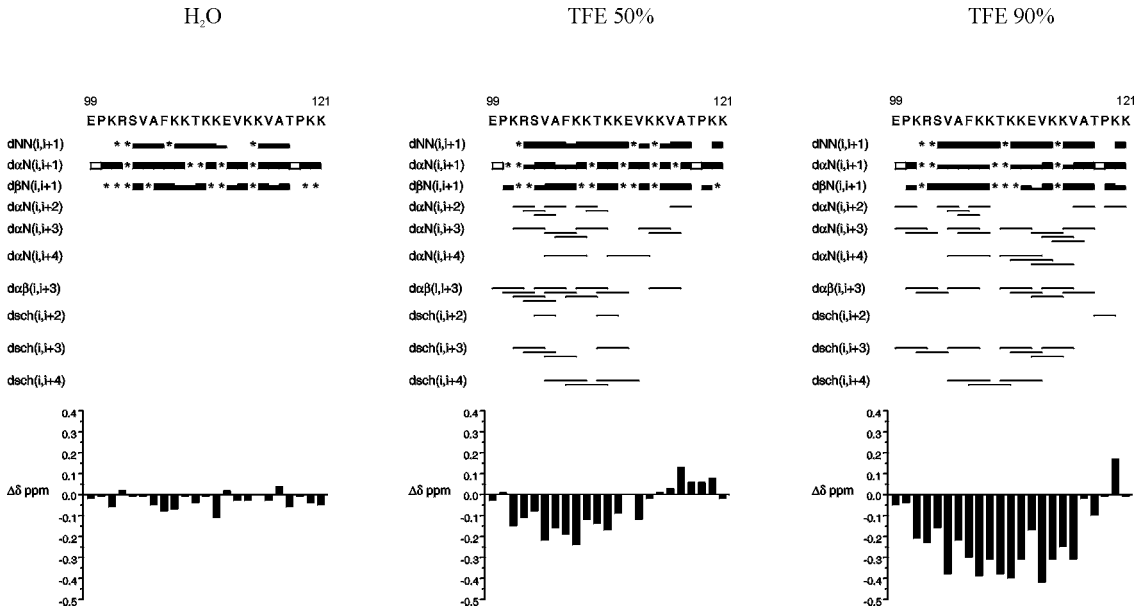


Figure 3. Summary of the NOE connectivities of CH-1. The results in aqueous solution (left), in 50% TFE solution (center) and in 90% TFE solution (right) are presented (25 °C, pH 3.5). The thickness of the lines reflects the intensity of the sequential NOE connectivities, i.e., weak, medium and strong. 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, $C^{\alpha}H$ Dd, are shown as a function of the sequence number (bottom).

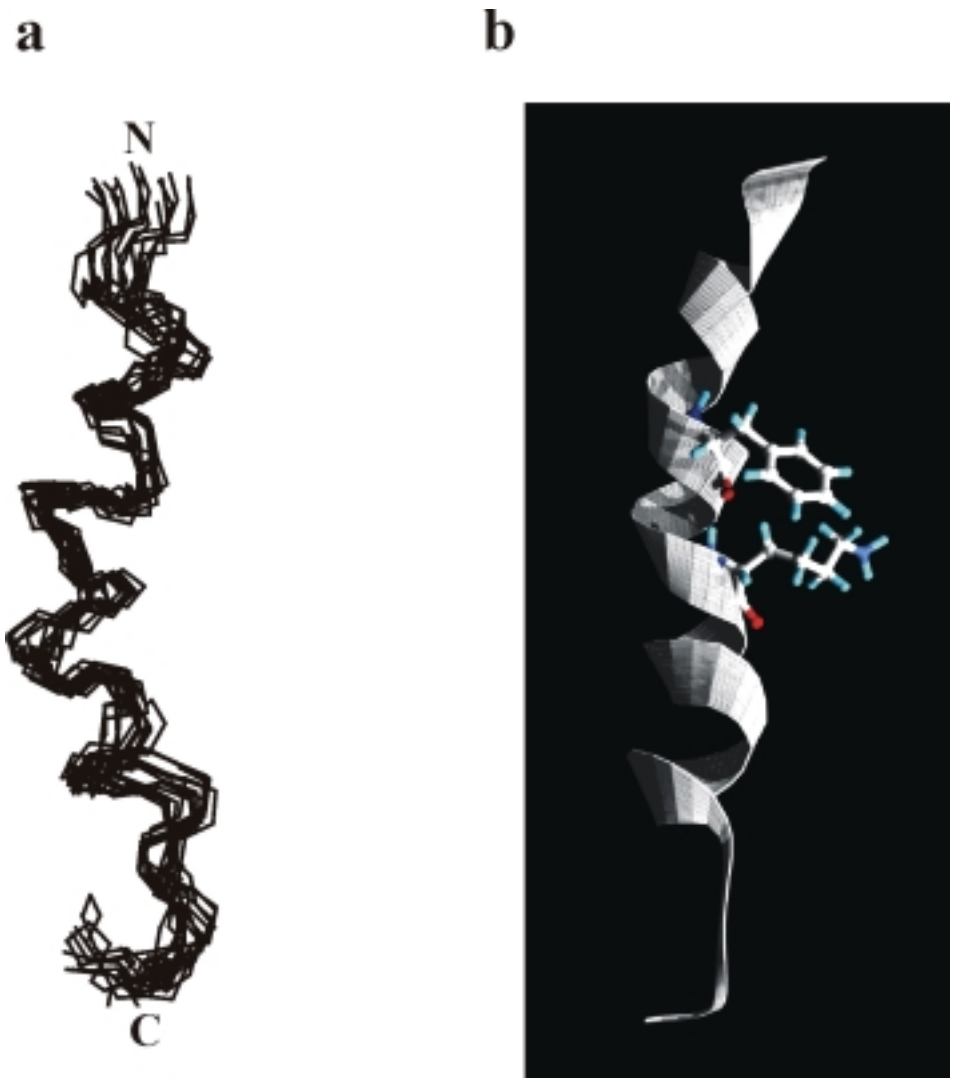


Figure 4. Structure of the CH-1 peptide. (A) Superposition of the best 15 structures of the CH-1 peptide calculated by distance geometry methods on the basis of the distance constraints derived from observed NOE cross-correlations in 90% TFE soluion. (B) Illustration of the interaction between the Phe106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

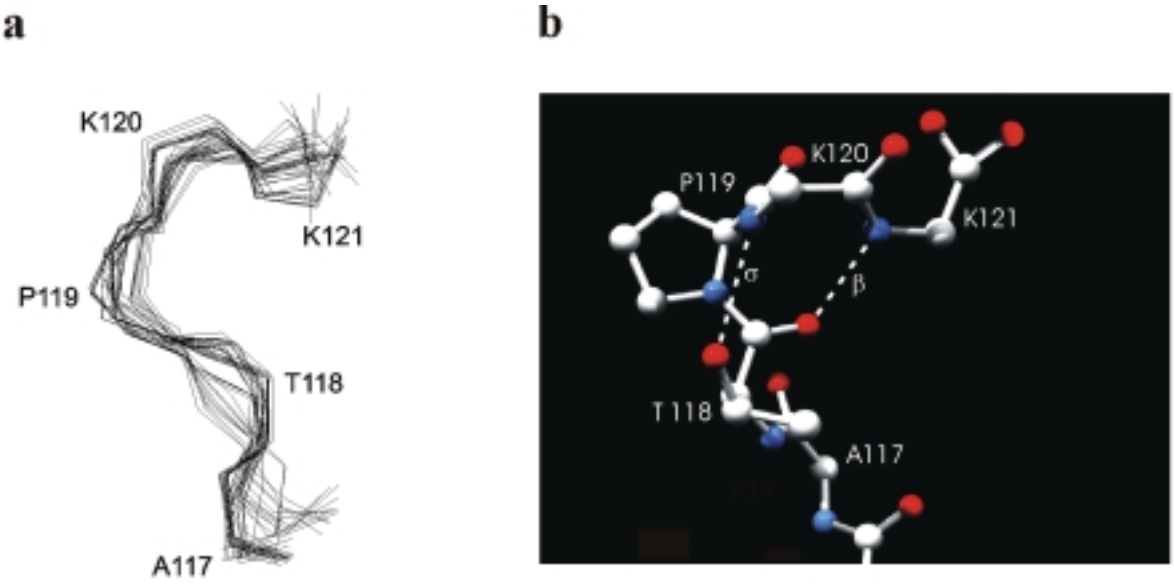


Figure 5. Turn structure adopted by the C-terminal end sequence TPKK. (A) Superposition of the best 25 calculated structures of the TPKK motif. (B) Illustration of one of the 25 computed structures. Broken lines indicate possible hydrogen bonds; β-type (β) and σ-type (σ).

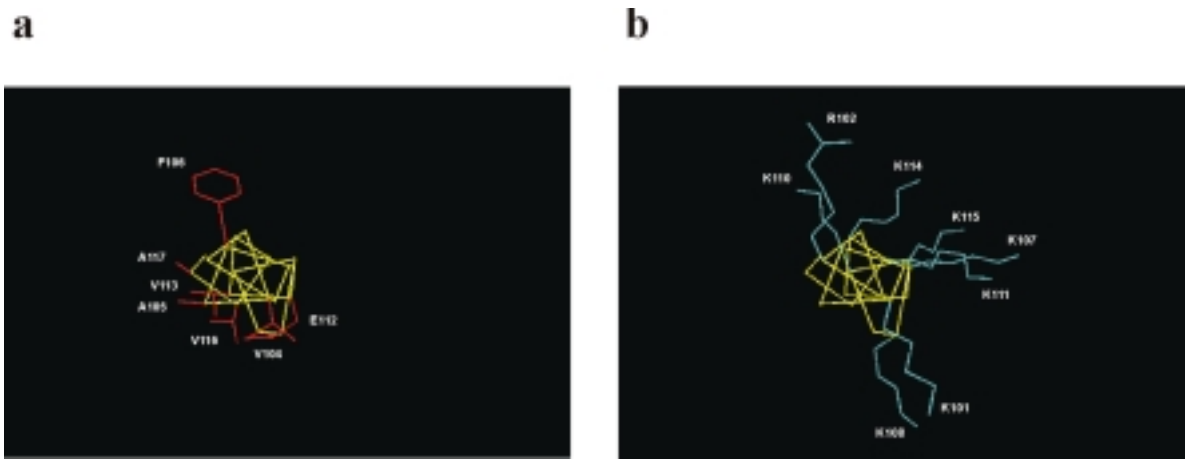


Figure 6. End view down the helix axis of one of the calculated structures of CH-1. The helical region from Lys101 to Val117 is represented. It shows the amphipathic character of the helix, with the positively charged residues on one face of the helix and the hydrophobic residues and Glu112 on the other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

CAPÍTOL II

FTIR ANALYSIS OF A C-TERMINAL DOMAIN PEPTIDE OF HISTONE H1º

INTRODUCTION

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 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 and Van Holde, 1992; Khochbin and Wolffe, 1994; Bouvet et al., 1994; Shen and Gorovsky, 1996; Wolffe et al., 1997b). Previous work has 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 and Archer, 1998) or the activation or repression of specific genes in *Tetrahymena* (Dou et al., 1999) are regulated by phosphorylation localized to the tail-like domains.

The study of the structure of the terminal domains when bound to the DNA may contribute to the understanding of H1 function. H1 terminal domains have little or no structure in solution. The C-terminal domain, however, acquires a substantial proportion of α-helical structure in the presence of helical inducers such as TFE and $HClO_4$ (Clark et al., 1988; Hill et al., 1989; Vila et al., 2000). 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 (Morán et al, 1985).

We previously studied the conformational properties of a peptide (CH-1) belonging to the C-terminal domain of the H1 subtype H1° by high resolution ^{1}H-NMR (Vila et al., 2000). The peptide is adjacent to the globular domain and contains the longest proline-free fragment in the C-terminal domain. In aqueous solution, CH-1 behaves as a

mainly unstructured peptide, although turn-like conformations in rapid equilibrium with the unfolded state may be present. Addition of trifluoroethanol (TFE) resulted in a substantial increase of the helical content. 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 on the opposite face. This kind of arrangement could affect the conformation of the initial part of the linker DNA at the entry/exit point of the nucleosome. The last four residues of the peptide, TPKK, adopt a turn conformation in TFE. Sequences of this kind have been proposed as DNA binding motifs (Suzuki, 1989; Suzuki et al., 1993).

We have now studied the interaction of the CH-1 peptide with the DNA by IR spectroscopy to identify features of the low-resolution structure of the bound peptide. IR spectroscopy is particularly well suited to the study of the complexes of DNA with basic peptides since it is not affected by turbidity. The important questions are whether the interaction with the DNA induces the folding of the peptide and, if so, whether the structure of the bound peptide is similar to the structure in TFE solution. Our results indicate that on binding to non-specific DNA targets the peptide acquires a substantial proportion of helical structure, while the TPKK motif at the C-terminal end of the peptide folds in a stable turn conformation.

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→Lys) and 113 (Val→Ile). The peptide was acetylated to remove the dipole destabilization effect.

Infrared Measurements

Peptide samples were measured at 4.6 mg/ml. Peptide-DNA complexes contained 6.7 mg/ml of DNA and the appropriate amount of peptide. Samples were measured in 10 mM HEPES, pD 7.0, plus either 10 mM or 70 mM NaCl. Prior to sample preparation, the trifluoroacetate counterions were replaced with chlorine ions by five lyophilisations of the peptide in 10 mM HCl. To transfer the samples to D_2O buffer, the aqueous solution was evaporated in a Speed-vac (Savant) evaporator and then reconstituted in D_2O.

The spectra were recorded in a Nicolet Magna II 550 spectrometer equipped with a MCT detector using a demountable liquid cell (Harrick Scientific, Ossining, NY) with calcium fluoride windows and 50 μm spacers. Typically, 1000 scans for each spectrum, background and sample, were collected and the spectra were obtained with a nominal resolution of 2 cm^{-1}, at 22 ºC.

Data treatment and band decomposition of the original amide I' have been described previously (Arrondo and Goñi, 1999). Briefly, for each component, four parameters were considered, band position, band height, band width, and band shape. The number and position of component bands was obtained through deconvolution. Initial heights were set at 90% of those of the original spectrum for the bands in the wings and for the most intense component and at 70% of the original intensity for the other bands. In decomposing the amide I' band gaussian components were used. The baseline was removed before fitting. The curve-fitting procedure was accomplished in two steps: (i)

The band position was fixed, allowing width and heights to approach final values, and (ii) band positions were left to change. Band decomposition was performed using CURVEFIT running under SpectraCalc (Galactic Inc., Salem, NH). The fitting result was evaluated visually by overlapping the reconstituted overall curve on the original spectrum and by examining the residual obtained by subtracting the fitting from the original curve. The procedure gave differences of less than 1% in band areas after the artificial spectra were submitted to the curve fitting procedure. The frequency positions of the band centers were independently evaluated by second derivative procedures; they were always very close to the positions found by deconvolution.

Spectra of peptide-DNA complexes

Complexes of the CH-1 peptide with mouse DNA, alternating poly[dA-dT]·poly[dA-dT], and homopolymeric poly[dA]·poly[dT] were prepared. The DNA contribution to the spectra was subtracted using a DNA sample of the same concentration. The symmetric component of the phosphate vibration at 1087 cm^{-1} was used to check the accuracy of the subtraction, for this band shows no overlap with other vibrations, it is not substantially affected by the interaction of the peptide and its intensity is proportional to the DNA concentration. The DNA spectrum was weighted so as to cancel the absorption at 1087 cm^{-1} in the difference spectra.

RESULTS

IR spectroscopy in aqueous solution and in TFE

The amide I' band of the CH-1 peptide in aqueous (D_2O) and in TFE solution and in complexes with DNA are shown in Figure 1. The number and position of the component bands were obtained by Fourier deconvolution and used for the curve-fitting of the original envelope by an iterative process as described previously (Arrondo and Goñi, 1999) (Figures 2-4). Values corresponding to band position and percentage area are given in Table 1.

In aqueous solution, the amide I' shows a major band (representing 38% of the total amide I' intensity) at 1643 cm^{-1}, which is typical of the random coil (Byler and Susi, 1986). The other main component (25% of the total amide I' intensity), which appears at 1662 cm^{-1}, was assigned to turns. In aqueous solution, the NMR did not show a significant helical population (Vila et al., 2000). However the presence of abundant medium intensity NOE correlations between sequential amide protons, which are only close enough in folded structures, suggested that turn-like conformations in rapid equilibrium with the unfolded state could be present. The turns at 1662 cm^{-1} observed in IR should, therefore, be short-lived and in rapid exchange with open states. Minor components at 1652 cm^{-1} (5%) and 1637 cm^{-1} (6%) were attributed to α-helix and 3_{10} helix, respectively, as discussed below (Figure 2A). A band is present near 1625 cm^{-1} (16%). This component also contributed to the spectra in TFE solution (7%) and in the complexes with the DNA (6-14%). This component was first described in homopolypeptides (Susi, 1969). It has been assigned to peptides in extended configurations involved in intermolecular hydrogen bonding, such as in monomer-monomer interactions. It is also seen in irreversibly aggregated proteins (Alvarez et al., 1987; Surewicz et al., 1990; Jackson et al., 1991; Arrondo and Goñi, 1999). In human low density lipoproteins it was assigned to "low frequency" β-sheets (Goormaghtigh et al., 1989) or to β-structure less accessible to the external solution (Herzyk et al., 1987). In this context, this component could be tentatively assigned to a fraction of aggregated peptide.

In 90% TFE solution (Figure 2B), the random coil vibration at 1643 cm^{-1} was only a minor component (3%), while the α-helix band at 1654 cm^{-1} became the main component, with 51% of the total amide I' band. In TFE, the band at 1637 cm^{-1}, that was present as a minor component in water solution, was observed with increased

intensity (14% of the amide I' intensity). Bands at this wavenumber have been attributed to 3_{10} helix by several authors (Dwivedi et al., 1984; Holloway and Mantsch, 1989; Prestrelski et al., 1991; Miick et al., 1992). This band position is also found in β-sheet structures (Byler and Susi, 1986), but in this structure must be accompanied by a high frequency component at 1675 cm^{-1} (Arrondo and Goñi, 1999), that was not observed in our case. Therefore, in CH-1, we assigned the band at 1637 cm^{-1} to 3_{10} helical structure. The assignment is supported by structure calculations based on NOE constraints observed in 90% TFE, conditions where NMR shows no sign of β-structure (Vila et al., 2000).

In TFE, a component at 1670 cm^{-1}, representing about 20% of the amide I' band, substitutes the component at 1662 cm^{-1}. Bands around 1670 cm^{-1} are also attributed to turns. The main contribution to this band probably arises from the TPKK motif at the C-terminal end of the peptide. The motif has no defined structure in aqueous solution, but in 90% TFE solution it adopts a type (I) β-turn, a σ-turn or a combination of both. In these conditions, the rest of the peptide is basically helical as shown both by IR and NMR. The band at 1662 cm^{-1}, found in aqueous solution, would thus correspond to less stable turn-like conformations in rapid equilibrium with disordered states, while the band at 1670 cm^{-1} would arise from the more stable TPKK motif.

IR spectroscopy of the peptide bound to the DNA

We have studied the structure of the peptide bound to several types of DNA: mouse DNA, alternating poly[dA-dT]·poly[dA-dT], and homopolymeric poly[dA]·poly[dT] (Figure 3). The DNA contribution to the spectra of peptide-DNA complexes was subtracted using a DNA sample of the same concentration. The asymmetric component of the phosphate vibration at 1087 cm^{-1} was used to check the accuracy of the subtraction as described in Materials and Methods. Spectra of complexes of different peptide/DNA ratio (**r**) were also recorded (Figure 4). The spectra were always basically identical regardless of the **r** value of the complexes, indicating that the amide I' region was not significantly affected by DNA spectral changes.

The band at 1670 cm^{-1}, which in TFE solution was attributed to the C-terminal TPKK motif in a turn conformation, was still observed in the complexes (Figure 3). It amounts to about 23% of the amide I' band. As the turn is not stable in aqueous solution, its presence in the peptide-DNA complexes indicates that the turn structure is induced by the interaction with the DNA.

In complexes with the DNA, the α-helix band splits into two components at 1657 cm^{-1} and 1647 cm^{-1}, each representing about 20% of the total amide I' intensity. Both components are in the spectral region of α-helical structures. The possibility that the faces of the amphipathic helical region of the peptide are exposed to different environments in the complexes, giving rise to different vibrations, will be discussed. A band at 1637 cm^{-1}, representing about 20% of the amide I' intensity, was also present. It has been attributed to 3_{10} helix, in accordance with the assignment in the peptide in TFE solution. This component was present with low intensity (6%) in the free peptide in water solution. Its significant increase in the complexes indicates that this kind of helical structure is induced upon interaction with the DNA.

The percentages and positions of the different components of the amide I' band in the spectra of the complexes with alternating poly[dA-dT]· poly[dA-dT] and homopolymeric poly[dA]·poly[dT] are very similar to those obtained with mouse DNA (Figure 3).

DISCUSSION

We have shown that double-stranded DNA can promote helix and turn formation in a peptide belonging to the C-terminal domain of histone H1 that has a very low amount of secondary structure in solution. The C-terminal domain is presumably involved in the organization of the linker DNA in chromatin. The CH-1 peptide is adjacent to the globular domain and could affect the conformation of the initial part of the linker DNA at the entry/exit of the nucleosome. This could have important consequences in those cases where gene regulation is mediated by the tail domains (Lee and Archer, 1998; Dou et al., 1999).

In aqueous solution, the IR spectrum of the free peptide is dominated by the random coil band at 1643 cm^{-1} and by a band at 1662 cm^{-1}, which is usually attributed to turns. Previous NMR studies of the peptide in water solution showed abundant NN(i, i+1) NOE connectivities of medium intensity, indicating that in these conditions the conformational ensemble of the peptide included a substantial population of folded conformations. Medium and strong βN(i, i+1) connectivities, which are commonly observed in type I (or type III) turns and in helical structures (Wüthrich et al., 1984) were also present (Vila et al., 2000). Long range order was, however, absent. We conclude that the band at 1662 cm^{-1} originates in short-lived turn-like conformations, that interconvert rapidly with the extended chain random coil form. Such an ensemble of secondary structures of low stability was referred to by Dyson et al. (Dyson et al., 1988) as *nascent helix*, since they are readily stabilized in an ordered helical conformation by long range interactions or in the presence of TFE. The assignment of the 1662 cm^{-1} band is further supported by the disappearance of this component in TFE solution and in the complexes with the DNA, conditions that both stabilize a significant amount of helical structure.

In TFE solution and in the DNA complexes, the 1662 cm^{-1} band shifts to 1670 cm^{-1}, indicating a change in the turn pattern (Arrondo and Goñi, 1999). We attribute this band to the TPKK motif at the C-terminal end of the peptide. In the presence of secondary structure inducers (such as TFE and DMSO), (T/S)P(K/R)(K/R) sequences adopt turn conformations –that could be a type (I) β-turn, a σ-turn or a $\beta\sigma$-turn with two hydrogen bonds (Suzuki, 1989; Suzuki et al., 1993; Vila et al., 2000). Although it is not easy to attribute a particular component of the IR spectrum to a given region of a peptide; in this case, the intense band at 1670 cm^{-1} can hardly arise from sequences other than the TPKK motif since in 90% TFE the rest of the peptide predominantly adopts a helical

conformation (Table 1). The important observation is that the band at 1670 cm^{-1} is present in the complexes with the DNA, and with a similar intensity as in the free peptide in TFE solution, indicating that the interaction with the DNA stabilizes the turn conformation of the TPKK motif. In summary, the band at 1662 cm^{-1} observed in water solution would arise from transient turn structures that would be stabilized into helical structure with long-range order in water-TFE mixtures or by interaction with the DNA. The band at 1670 cm^{-1} would arise from the well defined, and presumably more stable, TPKK motif. The folding of the TPKK motif upon interaction with the DNA confirms that this sequence indeed behaves in H1 as a DNA binding motif as previously proposed (Suzuki, 1989; Suzuki et al., 1993).

In 90% TFE solution, the CH-1 peptide becomes helical between Pro100 and Val116 (Vila et al., 2000). In these conditions, the amide I' band is dominated by the α-helical component (1654 cm^{-1}), which represents 51% of the total intensity. A component at 1637 cm^{-1}, representing about the 14% of the amide I', is also observed. This band can be assigned to 3_{10} helical structure. Structural calculations based on NMR constraints indicate that the 3_{10} helix is highly likely for the peptide sequence PKRS (Vila et al., 2000). Bands at 1637 cm^{-1} have been attributed to 3_{10} helical structure. Short alanine-based peptides have a maximum at 1637 cm^{-1}, and may form 3_{10} helices rather than α-helices in aqueous solution (Miick et al., 1992). Polyaminoisobutyric acid films in D_2O gave an amide I' band at 1640 cm^{-1} (Dwivedi et al., 1984). In α-lactoalbumin, an amide I' band at 1639 cm^{-1} was also assigned to 3_{10} helix (Prestrelski et al., 1991). The contributions of α-helix and 3_{10} helix estimated by IR agree well with the values estimated by NMR.

The component at 1637 cm^{-1} is also present in the complexes, with an intensity (17-20%) slightly higher than in TFE, suggesting that the 3_{10} helical structure is also induced upon interaction with the DNA. In the complexes, the contribution of the α-helix is split into two components at 1657 cm^{-1} and 1647 cm^{-1}, which together represent about 40% of the total amide I' intensity. The interaction with the DNA would thus induce a considerable amount of α-helical structure as compared with the small amount that may be present in aqueous solution. The presence of two components in the spectral region of the α-helix may be caused by a distortion of the helical region upon interaction with the DNA, which could be favored by the extreme amphipathic character of the helical region of the peptide, with all the basic residues on one side of the helix

and all the hydrophobic residues on the other (Vila et al., 2000). It should also be noted that the vibrations of α-helical structures can be sensitive to their environment (Haris and Chapman, 1995). The turbidity of the solutions of the peptide-DNA complexes, at the concentrations used in the IR experiments, suggests that the hydrophobic faces of the peptide bound to different DNA molecules could be in contact. In the complexes, the charged face of the helix bound to the DNA and the hydrophobic face may, therefore, be exposed to very different environments. Bands around 1647 cm^{-1} have been assigned to α-helical structures. For instance, in magainins, a family of positively charged antimicrobial amphipathic peptides, which are reminiscent of the CH-1 peptide, an amide I' band centered at 1647 cm^{-1} was assigned to α-helix in the presence of negatively charged liposomes (Jackson et al., 1992). Bands at 1656 cm^{-1} and 1646 cm^{-1} were also reported for the chaperonine GroEL and assigned to different types of α-helix (Galán et al., 2001).

Taken together, the contributions of the peptide at 1657 cm^{-1}, 1647 cm^{-1} and 1637 cm^{-1} in the complexes with the DNA, which have been attributed to helical structure, are equivalent to the amount of helical structure of the peptide in TFE (Table 1). In addition, the turn conformation of the TPKK motif is also induced upon interaction with the DNA. It thus appears that in this peptide, TFE reveals the conformational propensities of the DNA-bound peptide with high fidelity.

Our results, showing the presence of inducible helical and turn elements in CH-1, support the view that both kinds of structural elements behave as DNA binding motifs in the C-terminal domain of histone H1.

band position (cm^{-1})	band area (%)				
	D_2O	90% TFE	CH1 / Mouse DNA	CH1 / p[dA-dT]·p[dA-dT]	CH1 / p[dA]·p[dT]
1624-26	16	7	6	14	13
1637-38	6	14	20	21	17
1642-43	38	3	0	0	0
1647	0	0	20	19	19
1652	5	0	0	0	0
1654	0	51	0	0	0
1657	0	0	23	20	21
1662	25	0	0	0	0
1670	0	20	24	20	26

Table 1: Band position (cm^{-1}) and percentage area (%) corresponding to the components obtained after curve fitting of the amide I' band of the CH-1 peptide in aqueous (D_2O) and TFE solution and in the complexes with the DNA.

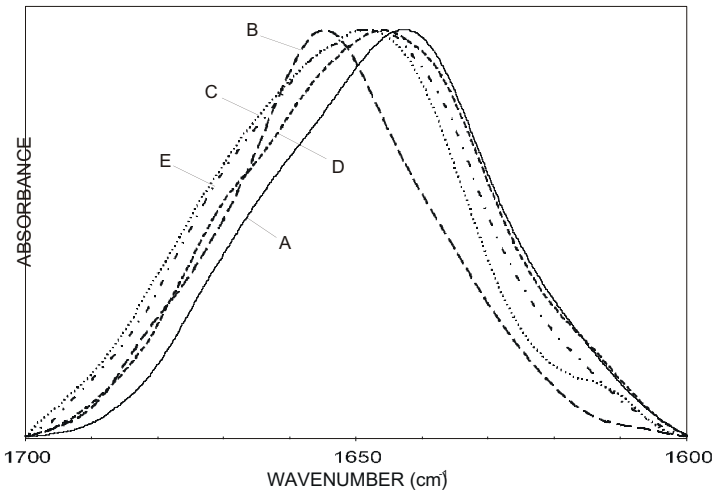


Figure 1. The amide I' band of the IR spectra of the CH-1 peptide at 22 ºC in 10 mM HEPES, 10 mM NaCl and D_2O, pD 7 and 4.6 mg/ml peptide concentration. (A) Free peptide in aqueous solution, (B) Free peptide in 90% TFE solution, (C) Spectrum of the peptide bound to mouse DNA, **r** (peptide/DNA ratio) = 0.7 (w/w), (D) Spectrum of the peptide bound to poly[dA-dT]· poly[dA-dT], **r** = 0.7, (E) Spectrum of the complex of the peptide bound to poly[dA]· poly[dT], **r** = 0.7. The DNA contribution to the spectra of the complexes was subtracted as described in Materials and Methods.

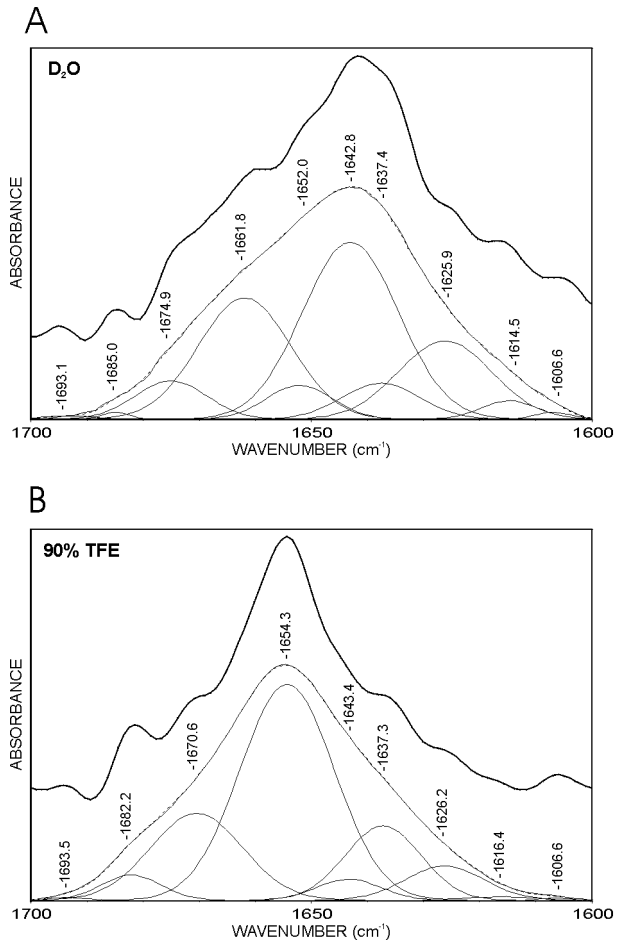


Figure 2. Amide I' band decomposition of the free CH-1 peptide spectra, showing the component bands, the envelope and, in dashed lines, the reconstitution of the amide I' band from the components. (A) Peptide in D_2O buffer. (B) Peptide in 90% TFE solution. The dashed and continuous lines are virtually superimposed, because of the goodness of the fit. The Fourier deconvolution of the absorption spectra, which was used to fix the position of the component bands, is included.

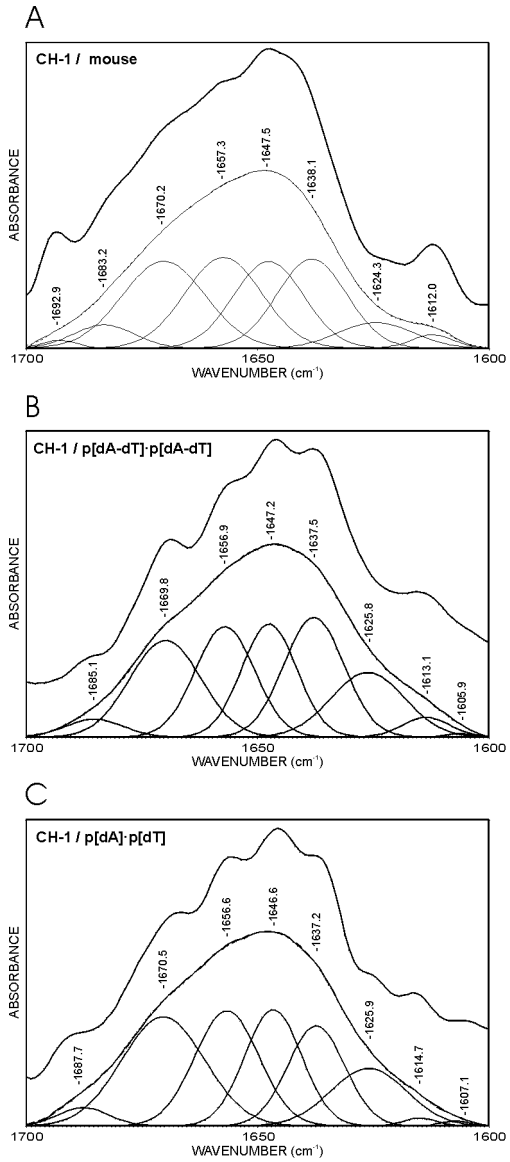


Figure 3. Amide I' band decomposition of the spectra of the CH-1 peptide bound to DNA. (A) Mouse DNA, (B) poly[dA-dT]·poly[dA-dT], (C) poly[dA]·poly[dT]. The reconstitution of the absorption spectra from the component bands is indicated by dashed lines. The Fourier deconvolution of the absorption spectra, that was used to fix the position of the component bands is included. The DNA contribution to the spectra of the complexes was subtracted as described in Materials and Methods. r (peptide/DNA ratio) = 0.7 (w/w).

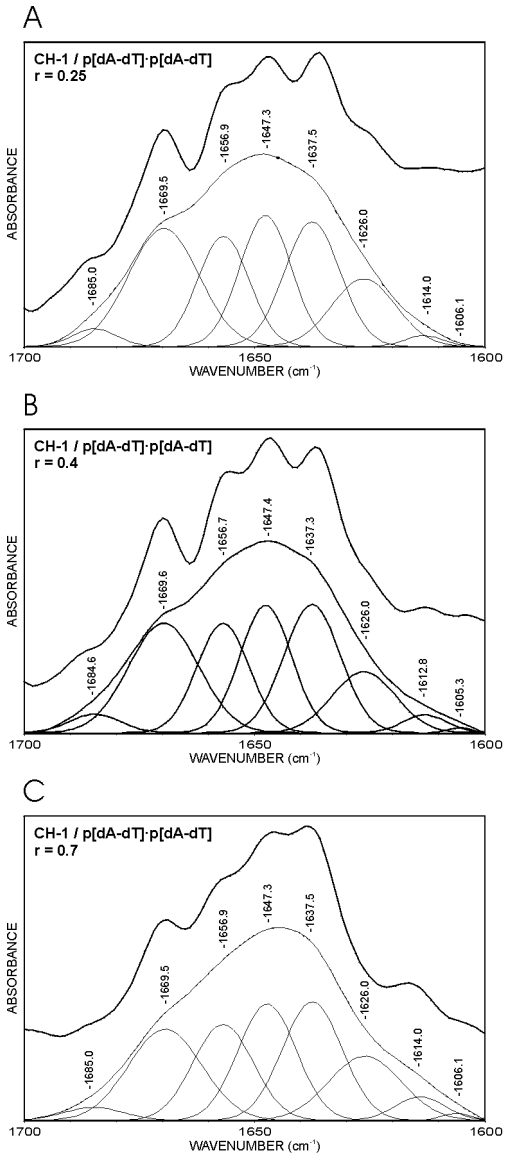


Figure 4. Amide I' band decomposition of the spectra of the CH-1 peptide in complexes with poly[dA-dT]·poly[dA-dT] of different r (peptide/DNA ratio) value (w/w). (A) r = 0.25, (B) r = 0.4, (C) r = 0.7. The reconstitution of the spectra from the component bands is indicated by dashed lines. The Fourier deconvolution of the absorption spectra, that was used to fix the position of the component bands is included. The DNA contribution to the spectra of the complexes was subtracted as described in Materials and Methods.

CAPÍTOL III

CD AND NMR ANALYSIS OF A N-TERMINAL DOMAIN PEPTIDE OF HISTONE H1e

INTRODUCTION

The linker histone H1 has a central role in stabilizing both the nucleosome and higher-order chromatin structure. H1 has a three-domain structure: a central globular domain flanked by highly basic amino- and carboxy-terminal tail-like domains (Hartman et al., 1977). The structure of the globular domain has been determined by X-ray diffraction (Ramakrishnan et al., 1993) and nuclear magnetic resonance (NMR) (Clore et al., 1987; Cerf et al., 1993). It contains a three-helix bundle, which resembles the winged-helix motif found in some sequence-specific DNA-binding proteins.

The N- and C-terminal domains are mainly unstructured in solution. However, the C-terminal domain acquires a substantial amount of α-helical structure in the presence of secondary-structure inducers such as trifluoroethanol (TFE) or $NaClO_4$ (Clark et al., 1988; Hill et al., 1989). A turn and a helix-turn motif in the C-terminal domain have also been characterized by high resolution NMR in the presence of inducers (Suzuki et al., 1993; Vila et al., 2000). These results led to the hypothesis that the C-terminal domain acquires α-helical structure when binding to DNA.

Histone H1 is described as a general transcriptional repressor, since it contributes to chromatin condensation, which limits the access of the transcriptional machinery to DNA. However, recent studies show that linker histone participates in complexes that can either activate or repress specific genes (Zlatanova and Van Holde, 1992; Khochbin and Wolffe, 1994; Wolffe et al., 1997b). These regulatory functions of histone H1 are attributed in some cases to the globular domain and in others to the tail-like domains (Bouvet et al., 1994; Shen and Gorovsky, 1996; Lee and Archer, 1998; Vermaak et al., 1998; Dou et al., 1999). Understanding the structural properties of the H1 terminal domains may clarify its function in chromatin.

The N-terminal domains of H1 histones have two distinct subregions (Böhm and Mitchel, 1985): the distal part is rich in alanine and proline and highly hydrophobic; its function is unknown. The region adjacent to the globular domain is highly basic, and may be involved in the location and anchoring of the globular domain (Allan et al., 1986).

Here we study the conformational properties of NE-1 peptide, which comprises the positively charged region of the N-terminal domain of linker histone subtype H1e. This is the most abundant subtype in many mammalian somatic cells. We show that this region has little structure in aqueous solution, but that it acquires a high degree of α-helical structure in the presence of TFE. In 90% TFE, the peptide is organized in two amphipathic α-helices, separated by a Gly-Gly motif, which allows great freedom to the orientation of the two α-helices. This feature may allow the tracking of the DNA phosphate backbone by the α-helices or simultaneous binding to two non-consecutive DNA segments.

MATERIALS AND METHODS

Peptide synthesis

The peptide Ac-EKTPVKKKARKAAGGAKRKTSG-NH_2 (NE-1) was synthesized by standard methods (DiverDrugs, Barcelona, Spain). Peptide homogeneity was determined by HPLC on Kromasil C8 . 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 15 to 36 at the N-terminus of mouse histone H1e. It presents two substitutions in rat, at positions 19 (Val → Ile) and 34 (Thr → Ala), and three in humans, at positions 26 (Ala → Ser), 29 (Gly → Ala) and 34 (Thr → Ala). The peptide was acetylated and amidated to remove the dipole destabilization effect.

Circular dichroism spectroscopy

Samples for circular dichroism spectroscopy were 2.85×10^{-5} M in the peptide and 5 mM in sodium phosphate buffer, 10mM NaCl, 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 20 °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):

$$\%\text{helical content} = 100([\theta]_{222}/-39500 \times (1-2.57/n)),$$

where n is the number of peptide bonds.

^1H NMR spectroscopy

Samples were routinely prepared as approximately 2.6 mM solutions of the peptide in H_2O, 5 mM phosphate buffer, 10 mM NaCl and the presence of either 50% or 90% deuterated TFE. The pH was adjusted to 3.5 with minimal amounts of HCl or NaOH in water. Spectra were obtained at 25 °C. Sodium 3-trimethylsilyl (2,2,3,3-$^2H_4$) propionate (TPS) was used as internal standard.

Spectra were recorded in a Bruker AMX-600 spectrometer. All 2-dimensional spectra were recorded in the phase-sensitive mode using time-proportional phase

incrementation (Marion and Wüthrich, 1983) with presaturation of the water signal. COSY (Aue et al., 1976) and NOESY (Kumar et al., 1980) spectra were obtained using standard phase-cycling sequences. Mixing times of 150 milliseconds were used in the NOESY experiments to avoid spin diffusion. TOCSY (Bax and Davis, 1985) spectra were acquired using the standard MLEV16 spin-lock sequence with a mixing time of 80 milliseconds. The phase-shift was optimized for every spectrum.

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

The helix populations were quantified on the basis of the 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 calculated by dividing the average conformational shift, $\Delta\delta=\Sigma(\delta_i^{obs}-\delta_i^{RC})/n$, by the shift corresponding to 100% helix formation. Random coil values, δ^{RC}, were those given by Wüthich (1986), except for Thr17, which was the that given for amino acids followed by Pro (Wishart et al., 1995). A value of -0.39 p.p.m. was used as the shift for 100% helix formation (Wishart et al., 1991). The helical length, n, was determined on the basis of NOE crosspeaks, couplings $^{3}J_{C\alpha H\text{-}NH}$ smaller than 5.0 Hz and conformational shifts, and confirmed by structure calculations.

Structure calculations

Peptide structures were calculated with the program DYANA (Güntert et al., 1997). 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 qualitatively 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. ϕ angles for those residues with $^{3}J_{C\alpha H\text{-}NH} < 5.0$ Hz (T3, K6, R10, A12, A13, K17 and K19) were restricted to the range -90º to -30º. The ϕ angles of the rest of the residues except for Gly were constrained to the range -180° to 0°.

RESULTS

Circular dichroism analysis

We have studied the peptide Ac-EKTPVKKKARKAAGGAKRKTSG-NH_2 (NE-1), which corresponds to the positively charged region of the N-terminal domain, immediately adjacent to globular domain, of mouse histone H1e (residues 15 to 36). The first 14 residues of the protein, mainly alanine and proline, are not positively charged.

The CD spectrum of the NE-1 peptide in H_2O, was dominated by the random coil component. Neither the double minimum at 208 nm and 222 nm nor the maximum at 190 nm of the α-helix were observed (Fig. 1A). The mean residue molar ellipticity at 222 nm ($[\theta]_{222}$), taken as diagnostic of helix formation, was negligible in water. However, the absence of the small positive peak at ~215 nm, a feature characteristic of the random coil, suggested that a small amount of structure was present.

Addition of TFE, which stabilizes peptide secondary structure, increased the α-helical content, as shown by the increase in the negative ellipticity at 222 nm and the change in the shape of the spectrum. The transition between the random coil and the helical structure seems to be a two-state equilibrium, as suggested by the presence of an isodichroic point. The helical content of NHe-1, estimated by the method of Chen *et al.* (1974), was 18% and 44% in 50% and 90% TFE, respectively (Fig. 1B).

NMR analysis

The NMR spectra were recorded in 50% and 90% TFE, pH 3.5, 25ºC. Chemical shift assignments of the NE-1 peptide in 90% TFE are shown in Table 1. Figure 2 shows selected regions of the two-dimensional NOE spectra of the peptide in 90% TFE, where NOE correlations corresponding to medium-range interactions are indicated. Figure 3 and Table 2 summarize all relevant NOE data for the peptide in 50% and 90% TFE. The figure also shows the plot of the conformational shifts of the $C^{\alpha}H$ protons, $\Delta\delta = \delta_{observed} - \delta_{random\ coil}$.

The presence of helical conformations was established on the basis of the following criteria: (1) the presence of stretches of non-sequential $\alpha N(i, i+3)$, $\alpha N(i, i+4)$, $\alpha\beta(i, i+3)$, as well as side-chain-side-chain and side-chain-main-chain (i, i+3) and (i, i+4) NOE connectivities; (2) strong sequential $N N$ NOE connectivities, concomitant with

weakened $\alpha N(i, i+1)$ NOE connectivities; (3) significant up-field shift of the $C^{\alpha}H$ resonances relative to the random coil values; (4) $^{3}J_{C\alpha H\text{-}NH}$ couplings smaller than 5.0 Hz.

In 50% TFE, the NE-1 peptide showed a low α-helical population. Thr17 may occupy the first position in the helix, with Pro18 in the statistically favored N+1 position (Richardson and Richardson, 1988). In these conditions, the C-terminal helical limit could not be determined accurately.

In 90% TFE, the helical population increased considerably, and spanned almost the entire peptide, as shown by the presence of numerous (*i, i*+3) and (*i, i*+4) NOE connectivities, the reinforcement of sequential *NN* NOEs, the up-field shift of the $C^{\alpha}H$ resonances and the $^{3}J_{C\alpha H\text{-}NH} < 5.0$ Hz for residues T17, K20, R24, A26, A27, K31 and K33. All these features are interrupted around the Gly28-Gly29 doublet. This indicates that NE-1 is organized in two α-helical elements separated by a Gly-Gly motif. The N-terminal α-helix (Helix N-I) begins at Thr17, as this residue presents a $^{3}J_{C\alpha H\text{-}NH} < 5.0$ Hz and it is the first involved in an $\alpha N(i, i+3)$ NOE connectivity. This helical element spans to Ala27, as shown by the abundance of (*i, i*+3) and (*i, i*+4) NOEs, the negative $C^{\alpha}H$ $\Delta\delta$ and the $^{3}J_{C\alpha H\text{-}NH}$ couplings < 5.0 Hz for residues T17, K20, R24, A26 and A27. The fractional helicity of this region was 54% according to the $C^{\alpha}H$ $\Delta\delta$.

The C-terminal α-helix (Helix N-II) is shorter and somewhat less stable than the N-terminal one, as deduced from its absence in 50% TFE and the slightly lower helical population in 90% TFE (44%). According to the presence of (*i, i*+3) and (*i, i*+4) NOE connectivities, the helix spans from Ala30 to Thr34, with Gly29, which is involved in an $\alpha N(i, i+4)$ NOE crosspeak, as N-Cap. This region contains two residues, Lys31 and Lys33, with $^{3}J_{C\alpha H\text{-}NH} < 5.0$ Hz. The helix may include Ser35, because this residue is involved in an $\alpha N(i, i+3)$ NOE cross-correlation.

Structure calculations

The structure calculations were performed on the basis of the NOE cross-correlations observed in 90% TFE. A set of 89 distance constraints, composed of 21 sequential and 68 medium-range constraints, was used to calculate the three-dimensional structures of the peptide. A number of structures were generated with the torsion angle dynamics program DYANA (Güntert *et al.*, 1997). The 22 best converged

structures were chosen. The global RMS deviation of the backbone atoms, excluding the first and the last residues, for this set of structures was 0.298±0.089 nm and the maximum NOE violation was 0.066 nm. Figure 4 shows the superposition of the backbones of the 22 selected structures. The region spanning from Thr17 to Ala27 adopts a well-defined helical structure, as does the region from Ala30 to Thr34. In some structures Ser35 is also included in the C-terminal helix. The calculated structures are highly flexible about the Gly28-Gly29 motif, which results in a wide range of values for the angle defined by the two helical axes. This obliges us to show the set of calculated structures fitted either for the N-terminal helical region (Fig. 4a) or for the C-terminal one (Fig. 4b).

DISCUSSION

Histone H1 terminal domains have little structure in aqueous solution. Secondary structure inducers promote the formation of turns and helices in the C-terminal domain (Clark et al., 1988; Hill et al., 1989; Suzuki et al., 1993; Vila et al., 2000). However, very little is known so far on the N-terminal domain functional and structural properties.

We have studied the structure of the NE-1 peptide, which corresponds to the charged region of the N-terminal domain of histone H1e, by CD and ^{1}H-NMR. In aqueous solution, NE-1 behaved as mainly unstructured peptide. Addition of TFE resulted in a substantial increase of the helical content. In 90% TFE, the peptide is structured in two α-helices spanning from Thr17 to Ala27 (Helix N-I) and from Gly29 to Thr34 (Helix N-II), with fractional helicities of 54% and 44%, respectively. These results, calculated from the up-field shifts of the $C^{\alpha}H$ δ values, are in good agreement with the average value of 44% obtained by CD for the entire peptide. The presence of an $\alpha N(i, i+3)$ NOE crosspeak involving Ser35 suggests that this residue may also be included in Helix N-II, as in some of the calculated structures. The lack of defined secondary structure when NE-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 and Serrano, 1994). When the method of Chou and Fasman (1974), which is based on protein data, is used, the two α-helical elements are correctly predicted.

Helix N-II surpasses the limit between the N-terminal and the globular domains defined by trypsin cutting at Lys33. This suggests that the structural limit between the domains is the helix-disruptive motif Pro37-Pro38, which is immediately adjacent to Helix I of the globular domain. Fig. 5 shows a model structure of the N-terminal peptide of H1e connected to the NMR structure of the globular domain of chicken histone H1 (Cerf et al., 1993).

TFE stabilizes helical and turn structures (Buck, 1998), revealing the conformational biases of the primary sequence. Although TFE has been extensively used, the mechanism by which it affects the polypeptide structure is not fully understood. Enhancement of hydrogen bonding, disruption of the water structure and preferential solvation of certain groups of the polypeptide chain have been suggested as possible explanations. It has also been argued that TFE could associate with the

hydrophobic surface of amphipathic helices and mimic a dehydrated environment (Buck, 1998). The electrostatic repulsion between the positively charged residues in the amphipathic helices could explain the unusually high TFE concentration necessary to stabilize histone H1 helices (Walters and Kaiser, 1985; Clark et al., 1988; Hill et al., 1989; Johnson et al., 1994; Vila et al., 2000).

Both helical elements in the N-terminal domain of H1e have a marked amphipathic character, with all the basic residues on one face of the helices and the apolar residues on the other. The positively charged amphipathic α-helices have been described as DNA binding motifs. Our results suggest that the positively charged region of the N-terminal domain adopts α-helical structure upon binding to the DNA, as proposed for some of the sequences of the C-terminal domain (Clark et al., 1988; Hill et al., 1989; Vila et al., 2000). Moreover, the proximity of this region to the globular domain and its high positive charge density support the view that it is involved in the location and anchoring of the globular domain to the nucleosome (Allan et al., 1986)

Both N-terminal domain helical elements contain a triplet of basic residues. Basic triplets are not frequent in somatic H1 subtypes, but they are more common in non sequence-specific DNA-binding proteins with high affinity for the DNA, such as sperm linker histones and protamines (Subirana, 1990). The triplets of basic residues probably confer a high DNA-binding affinity to the proximal region of the N-terminal domain. The inducible character of the helical elements of the terminal domains may prevent the highly charged N-terminal region from strongly binding to DNA before the globular domain is correctly positioned on the nuclesosome.

The two α-helical elements in the N-terminal domain are separated by a Gly-Gly motif. The Gly-Gly doublet behaves as a flexible linker between the helical elements. The Gly-Gly doublet is conserved at equivalent positions in many vertebrate H1 subtypes (Sullivan et al., 2000). A Gly-Gly motif is found in the porins of *E.coli* and other bacterial species (Jeanteur et al., 1991). The motif confers flexibility to the third loop of the protein, which appears to have functional implications (Van Gelder et al., 1997). Another example of structural flexibility associated with a Gly sequence is that of the calmodulin-related TCH2 protein from *Arabidopsis*. Its overall structure consists of two globular domains separated by a flexible linker region that contains a $(Gly)_4$ motif (Khan et al., 1997).

In the N-terminal domain of H1e, the wide range of possible orientations of the two α-helices allowed by the Gly-Gly motif could facilitate the tracking of the phosphate backbone by the helical elements or the simultaneous binding of two non consecutive DNA segments.

	NH	CαH	CβH	Cβ'H	CγH	Cγ'H	CδH	Cδ'H	Others
Glu15	7.92	4.30	2.02	2.11	2.43	2.43	---	---	
Lys16	8.08	4.47	1.86	1.99	1.47	1.56	1.76	1.76	$C^{\varepsilon}H_2$3.03
Thr17	7.78	4.60	4.39	---	$C^{\gamma}H_3$1.32	---	---	---	
Pro18	---	4.40	1.99	2.42	2.09	2.13	3.79	3.86	
Val19	7.25	3.78	2.13	---	$C^{\gamma}H_3$1.01	$C^{\gamma'}H_3$1.07	---	---	
Lys20	7.59	4.10	1.94	1.94	1.48	1.62	1.78	1.78	$C^{\varepsilon}H_2$3.01
Lys21	7.90	4.09	1.94	1.94	1.48	1.64	1.75	1.75	$C^{\varepsilon}H_2$3.00
Lys22	7.92	4.02	1.99	1.99	1.50	1.64	1.76	1.76	$C^{\varepsilon}H_2$3.00
Ala23	8.27	4.14	1.55	---	---	---	---	---	
Arg24	8.05	4.11	2.02	2.02	1.74	1.89	3.23	3.25	$N^{\varepsilon}H$7.22
Lys25	8.07	4.14	2.01	2.01	1.55	1.67	1.77	1.77	$C^{\varepsilon}H_2$3.01
Ala26	8.41	4.19	1.56	---	---	---	---	---	
Ala27	8.29	4.21	1.56	---	---	---	---	---	
Gly28	8.20	3.94/3.94	---	---	---	---	---	---	---
Cly29	8.20	3.94/3.94	---	---	---	---	---	---	---
Ala30	8.06	4.20	1.56	---	---	---	---	---	
Lys31	8.21	4.09	2.01	2.01	1.52	1.52	1.75	1.75	$C^{\varepsilon}H_2$3.00
Arg32	8.09	4.14	2.01	2.01	1.71	1.82	3.23	3.23	$N^{\varepsilon}H$7.07
Lys33	8.27	4.20	2.00	2.00	1.56	1.56	1.72	1.72	$C^{\varepsilon}H_2$3.00
Thr34	7.87	4.37	4.44	---	$C^{\gamma}H_3$1.35	---	---	---	
Ser35	7.80	4.54	4.02	4.05	---	---	---	---	
Gly36	7.89	3.93/4.03	---	---	---	---	---	---	

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

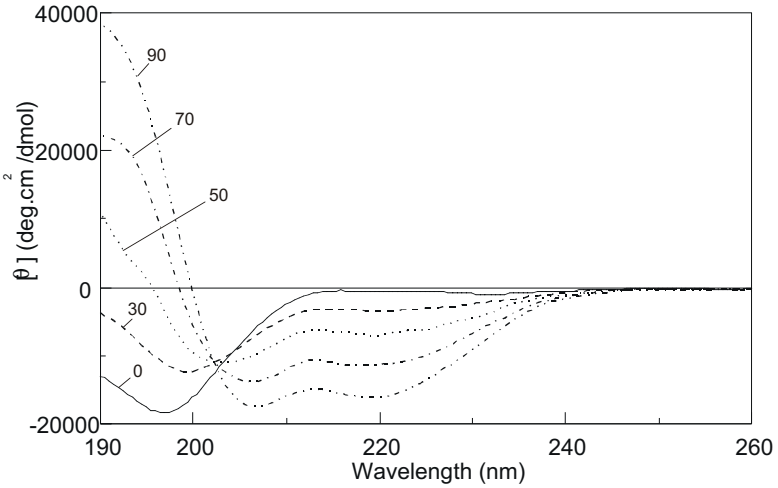
	Residue 1	Residue 2	Intensity
	β'Glu15	NHThr17	Weak
*	$\gamma\gamma'$Glu15	NHThr17	Weak
	αLys16	βPro18	Medium
	αLys16	γH_3Val19	Medium
	αThr17	$\gamma' H_3$Val19	Weak
	αThr17	γLys20	Weak
	αThr17	γ'Lys20	Medium
	αThr17	$\delta\delta'$Lys20	Medium
+	βThr17	NHVal19	Strong
	βThr17	γH_3Val19	Medium
	βThr17	$\gamma' H_3$Val19	Strong
*	γH_3Thr17	NHVal19	Medium
*	γH_3Thr17	NHLys20	Medium
	γH_3Thr17	$\varepsilon\varepsilon'$Lys20	Medium
	αPro18	γLys21	Medium
	αPro18	γ'Lys21	Medium
	αPro18	$\delta\delta'$Lys21	Strong
*	αVal19	γLys22	Medium
	αVal19	γ'Lys22	Medium
*	αVal19	$\delta\delta'$Lys22	Strong
	αVal19	$\varepsilon\varepsilon'$Lys22	Medium
	$\gamma' H_3$Val19	NHLys22	Weak
*	$\gamma' H_3$Val19	NHAla23	Weak
	αLys21	$\delta\delta'$Arg24	Medium
	δArg24	βH_3Ala27	Weak
	δ'Arg24	βH_3Ala27	Weak
	$\alpha\alpha'$Gly29	γArg32	Medium
	$\alpha\alpha'$Gly29	γ'Arg32	Medium
	$\alpha\alpha'$Gly29	$\delta\delta'$Arg32	Strong
	$\alpha\alpha'$Gly29	$\gamma\gamma'$Lys33	Strong
	αAla30	γH_3Thr34	Medium
	αLys31	γH_3Thr34	Medium

Table 2. 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 NHe-1 peptide in 90% TFE.

(*) NOE crosspeaks that are also observed in 50% TFE.

(+) NOE crosspeaks observed only in 50% TFE.

a



b

% TFE	$[\theta]_{222nm}$	% helix
90	-15324	44
70	-10706	31
50	-6378	18
30	-3174	9
0	-366	1

Figure 1. TFE-dependent conformational transition of the NHe-1 peptide measured by CD. (a) Far-UV CD spectra in the presence of various concentrations of TFE in 10mM NaCl, phosphate buffer 5 mM, pH 3.5 at 20 °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.

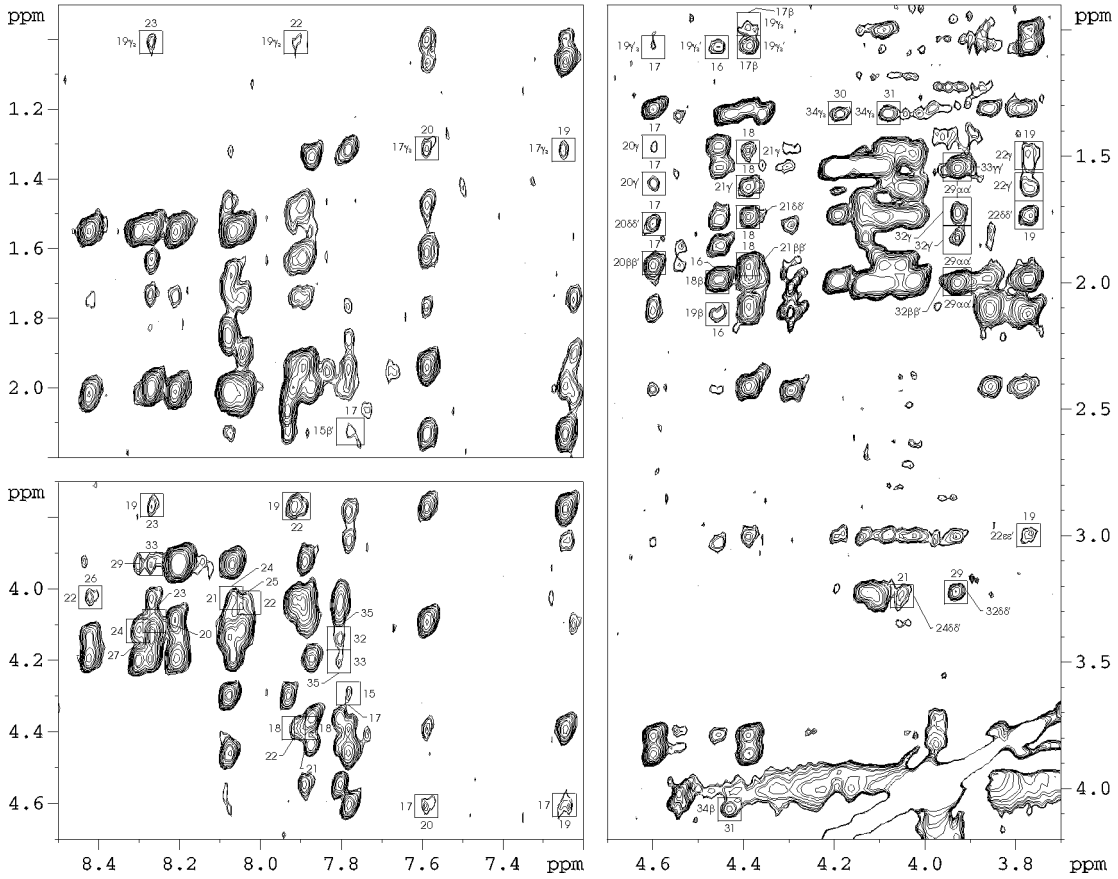


Figure 2. Selected regions of the 2-dimensional NOE spectrum of NHe-1. The peptide was 2.6 mM in 90% TFE, pH 3.5, 25 °C, mixing time 150 ms. Medium-range NOE connectivities are boxed. Left, NOE correlations of aliphatic protons (bottom, $C^{\alpha}H$) with amide or aromatic resonances. Right, NOE correlations of $C^{\alpha}H$ protons with other aliphatic protons.

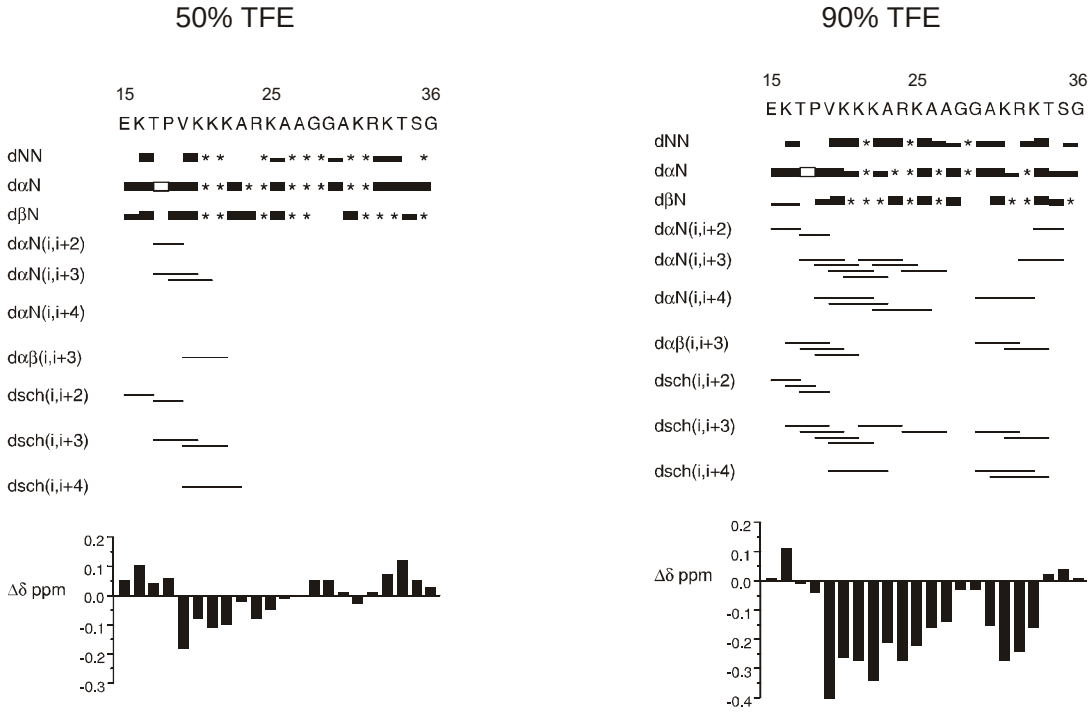


Figure 3. Summary of the NOE connectivities of NHe-1. The results in 50% TFE solution (left) and in 90% TFE solution (right) are presented (25 °C, pH 3.5). The thickness of the lines reflects the intensity of the sequential NOE connectivities, i.e., weak, medium and strong. An asterisk (*) indicates an unobserved NOE connectivity due to signal overlapping, closeness to the diagonal or overlapping with the solvent signal. An open box indicates a dαδ(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, $C^{\alpha}H$ Δδ, are shown as a function of the sequence number (bottom).

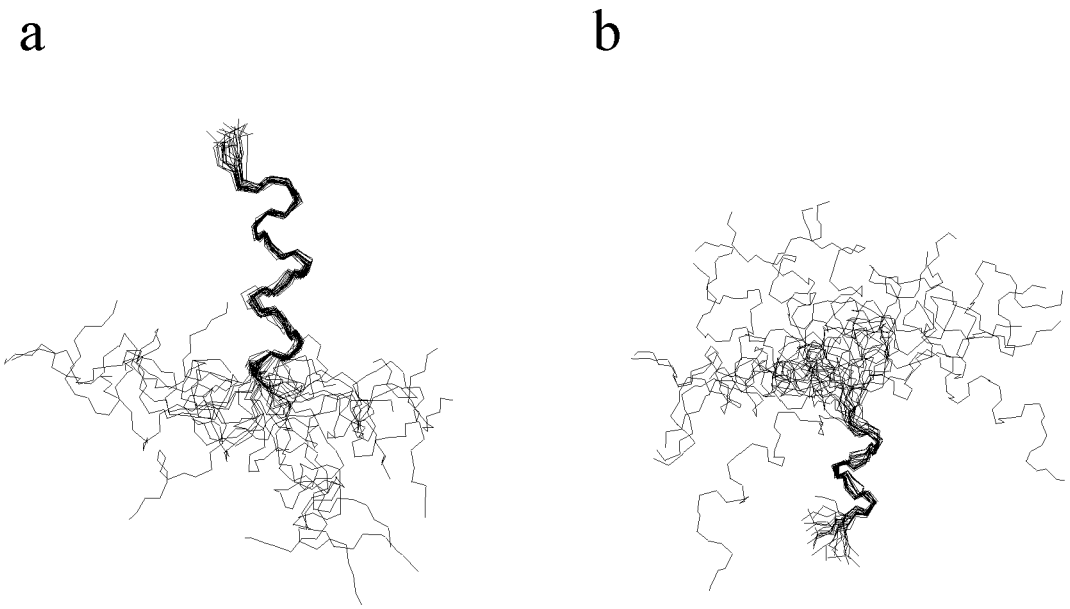


Figure 4. Superposition of the backbone of the NHe-1 peptide best 22 converged structures calculated by the torsion angle dynamics program DYANA (Güntert et al., 1997) on the basis of the distance constraints derived from observed NOE cross-correlations in 90% TFE solution. (A) The 22 structures fitted at the N-terminal helix. (B) The 22 structures fitted at the C-terminal helix.

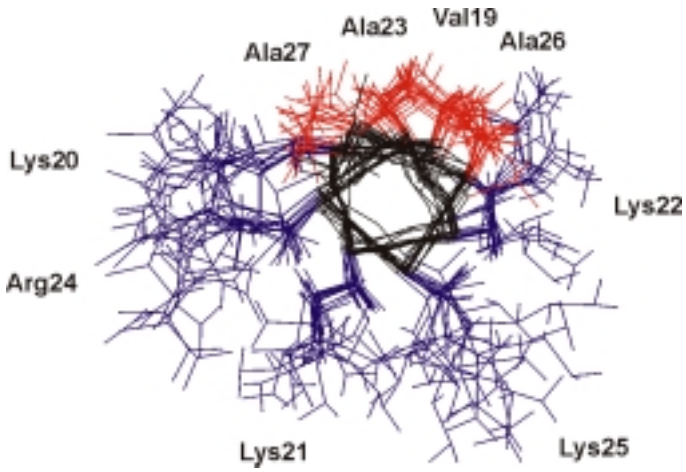


Figure 5. End view down the helix axis of the best 22 converged structures of NHe-1. The helical region from Val19 to Ala27 is represented. It shows the amphipathic character of the helix, with the positively charged residues on one face of the helix and the hydrophobic residues on the other face.

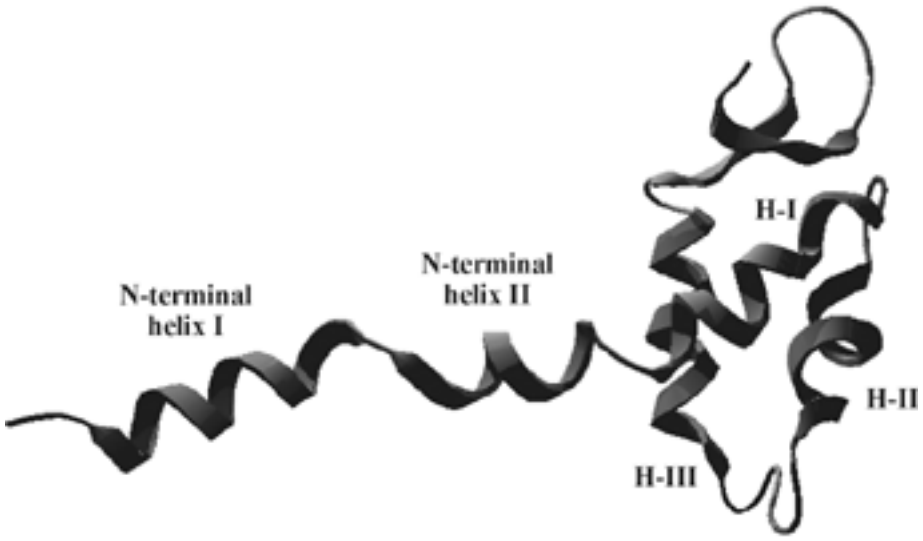


Figure 6. Model structure of the connection of the N-terminal and globular domains of histone H1e. The NMR structure of the folded N-terminal domain NHe-1 peptide is represented connected to the NMR structure of the globular domain of chicken histone H1 globular domain.