J. Mol. Biol. (1995) **245,** 297–308

JMB

Elucidating the Folding Problem of Helical Peptides using Empirical Parameters. III. Temperature and pH Dependence

Victor Muñoz and Luis Serrano

EMBL, Meyerhofstrasse 1
Heidelberg D-69117
Germany

Explaining the helical behaviour of amino acid sequences without tertiary interactions, in aqueous solution, could be considered one of the first steps to solve the protein folding problem in a rational way. In the accompanying paper the information about the conformational behaviour of helical peptides in solution, as well as the studies on α-helix stability in proteins has been utilised to derive a database of energy interactions. This database, when implemented in an algorithm based on the helix-coil transition theory (AGADIR), correctly calculates the average helical behaviour in solution of 423 peptides analysed by circular dichroism. The majority of these peptides have been studied at low temperatures (0 to 10°C), and neutral pH. However, *in vivo*, proteins fold at higher temperatures and in some cases low or high pH values. To understand protein folding it is necessary to calculate the helical behaviour of linear peptides under very different temperature and pH experimental conditions. We have included the temperature and pH effects on the helical behaviour of peptides by means of generally accepted assumptions and simplifications. The inclusion of these terms allow us to calculate the helical behaviour of polyalanine-based peptides, as well as of complex natural sequences, under different experimental conditions.

Keywords: α-helix stability; protein folding; thermodynamics; calorimetry

Introduction

Knowledge of the energetics of protein stability is central to our understanding of the folding process of protein molecules. The rational understanding of the energetics of the formation of the different secondary structure elements in a protein is a first step in that direction. One of the simplest and more studied secondary structure elements are α-helices. In a previous work (Muñoz & Serrano, 1994) and in the accompanying paper (Muñoz & Serrano, 1995), we assigned different ΔG values to the different interactions that occur in α-helices in order to determine the helical behaviour in solution of small peptides. These values were mainly deduced from the analysis of proteins and polyalanine-based peptides, as well as of peptides derived from natural sequences, analysed at low ionic strength (<100 mM, pH between 5 and 7, and temperature between 0 and 10°C). However, systems in thermodynamic equilibrium are obviously dependent on the environmental conditions. In aqueous solution (with volume and pressure constant), the three major environmental factors that could affect the energetics of the helix-coil transition of peptides are temperature, pH and ionic strength.

To understand the effect of temperature on the stability of helical peptides it is first necessary to determine which are the factors that contribute to α-helix stability, their enthalpic and entropic components as well as the changes in heat capacity upon unfolding. So far the models based on the helix-coil theory have used very simple approaches to incorporate the temperature-dependence of peptide helicity, in which only the elongation parameter *s* is temperature-dependent (Scholtz *et al*., 1991*c*), and no change in heat capacity is considered.

The effect of pH on α-helix stability has been experimentally studied in polyalanine-based peptides (Scholtz *et al*., 1993). There are two main energy terms affected by pH: interactions of the charged side-chains with the helix macrodipole and the $i,i+3$ and $i,i+4$ interactions between charged side-chains. Classical theories seem to account for the experimentally observed effects (Scholtz *et al*., 1993), although

Abbreviations used: TFA, trifluoroacetic acid; EDT, ethane dithiol.

0022–2836/95/030297–12 $08.00/0

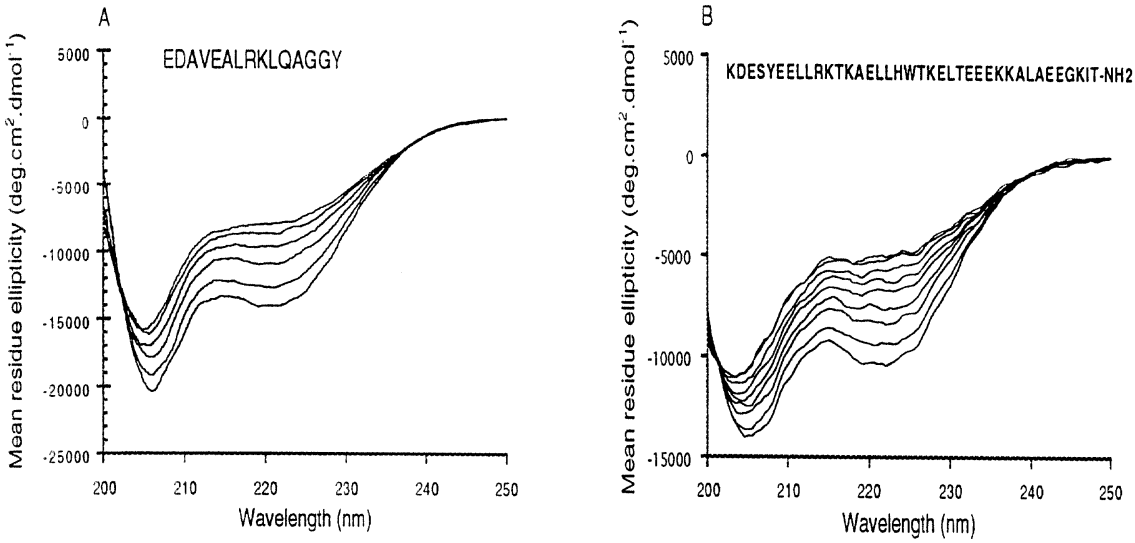

Figure 1. Changes in the far-UV CD spectra of peptides derived from wild-type or mutant protein sequences. All the peptides were analysed in 2.5 mM phosphate buffer pH 7.0. at 5°C. A, Mutant variant of a peptide corresponding to the α-helix 2 of CheY (see the accompanying paper, Muñoz & Serrano, 1995). B, Peptide corresponding to a region of titin (residues 1977 to 2014; Musco *et al*., unpublished results). The temperature increases from 0°C in steps of 10 deg.C. The curves with the minimum ellipticity at 222 nm correspond to the lower temperature.

an estimation of the model and parameters able to describe the effect in general terms has not yet been given.

Ionic strength affects the electrostatic, as well as the hydrophobic, interactions (Scholtz *et al*., 1991*b*). Screening of ionic interactions at high concentrations of salt has been experimentally described and it is indicated that the interactions of charged residues with the helix macrodipole are less affected than those between charged side-chains (Lockhart & Kim, 1993). However, there are no extensive experimental studies addressing the effect of the ionic strength on hydrophobic interactions.

Here, we report a modification of the algorithm AGADIR (Muñoz & Serrano, 1994, 1995), which includes a general model in the framework of classical physicochemical theories for the pH and temperature dependence of helical formation in linear peptides. Calculations obtained with this improved version of AGADIR are compared with the experimental data.

Results

Experimental temperature dependence of the average helical content in peptides corresponding to protein fragments

There are very few examples of peptides derived from proteins that have a high helical content (>25%) in aqueous solution at 0°C and have been denatured by temperature (Shin *et al*., 1993*a*,*b*; Yumoto *et al*., 1993). To test if our theoretical approximation is able to describe the temperature-dependence behaviour of peptide sequences having different types of side-chain to side-chain interactions (i.e. hydrophobic interactions between aliphatic or aromatic groups), we need more examples. For this reason we have analysed the temperature denaturation of two peptides derived from proteins that exhibit a high helical content (CheY, Muñoz & Serrano, 1995, and a titin fragment, Musco *et al*., unpublished results). The changes in the far-UV spectra with the temperature of these two different peptides, corresponding to protein fragments, are shown in Figure 1. There is a decrease of the minimum at 222 nm, as well as a displacement of the second minimum towards random-coil values (~196 nm). The negative ellipticity at 222 nm does not disappear since, as previously indicated, the ellipticity for the random coil decreases with temperature (Scholtz *et al*., 1991*a*; and see Materials and Methods). There is also an isodichroic point at 202 nm, indicative of the presence of a two-state transition (α-helix to random coil).

Comparison between the calculated and experimental temperature dependence of the helical content of monomeric peptides

Using the simplest model in which ΔH and in theoretical procedures (throughout) ΔS are temperature-independent (see equation (5)), the calculated values for polyalanine-based peptides are not very cooperative and there is not a good agreement with the experimental values (Figure 2A to E). If we consider that ΔH and ΔS are temperature-dependent (see equations (6) and (7)), and ΔCp is constant (equations (8) and (9)), the agreement with the experimental values is very good, within the limits of the experimental error (Figure 2A to E).

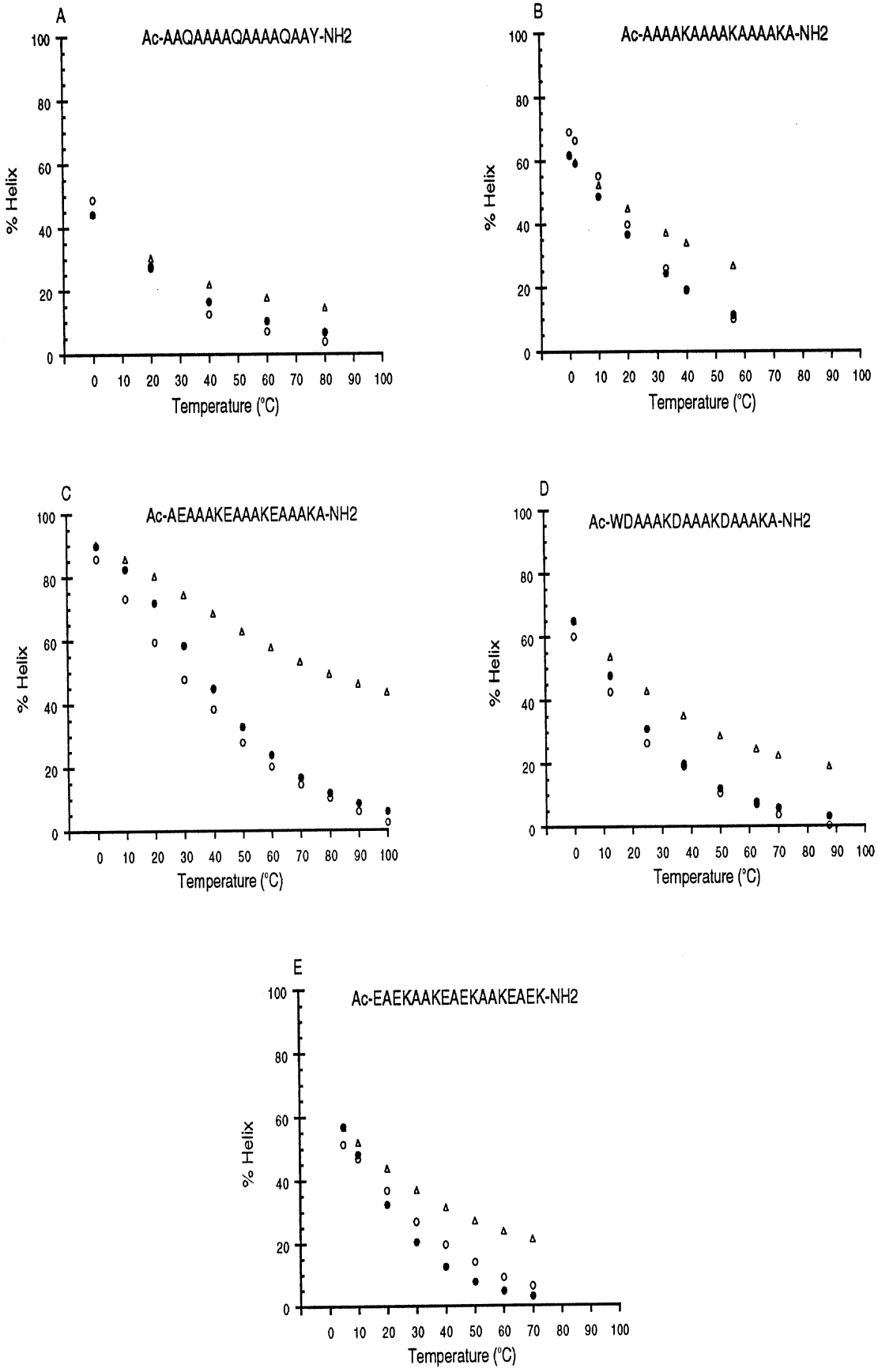

Figure 2. Comparison between the helical temperature-dependence of designed peptides and the calculated values. A, Analysed by Scholtz *et al*. (1991*b*). B, Analysed by Marqusee *et al*. (1989). C, Analysed by Merutka & Stellwagen (1990). D, Analysed by Merutka & Stellwagen (1991). E, Analysed by Zhou *et al*. (1993). (○) Experimental values; (●) calculated values considering the existence of constant heat capacities in the terms ΔG_{pol} and ΔS_{confor} (see Materials and Methods). (△) Calculated values assuming the very simple model shown in equation (5), in which the heat capacities are not considered.

To check whether equivalent results are obtained with sequences containing different types of side-chain to side-chain interactions, including hydrophobic interactions, we have analysed peptides corresponding to protein fragments (Shin *et al*., 1993*a*,*b*; Musco *et al*., unpublished results; Muñoz & Serrano, 1994; Yumoto *et al*., 1993), or designed sequences (Bradley *et al*., 1990). Calculation of the helical content of these peptides using equations (8) and (9) results in temperature denaturations that are too cooperative (with the exception of the titin fragment: Figure 3A to E). This indicates that in the case of peptides having hydrophobic interactions, between aliphatic or/and aromatic side-chains, we need an extra free energy term not present in the polyalanine-based peptides. This free energy term should be more favourable at higher temperatures to decrease the cooperativity of the unfolding process. The most obvious candidate for this free energy term is the energy coming out from the hydrophobic interactions between side-chains. Introduction of a new constant heat capacity increment ($\Delta C_{\text{Phydroph}}$; equation (10)), to take into account these hydrophobic interactions, results in the opposite effect for all the peptides (data not shown). This is not surprising since the calorimetric analysis of proteins indicates that the heat capacity change due to the burial of hydrophobic surface is highly temperature-dependent (Wintrode *et al*., 1994). Considering a linear dependence of $\Delta C_{\text{Phydroph}}$ with temperature (equations (18) and (19)) results in a very good agreement with the experimental values (Figure 3A to D). The inclusion of this extra term does not affect the calculated values for the polyalanine-based peptides since they do not contain hydrophobic interactions between aliphatic or/and aromatic side-chains.

pH-dependence of the average helical content

In Figure 4 we compare the pH-dependence of the average helical content of peptides having different electrostatic interactions, with the calculated values. The change in the repulsion electrostatic energy of two positive charges with pH is illustrated by Figure 4A (Marqusee *et al*., 1989). In Figure 4B we show the changes in helical content of two peptides, one of them having a His residue close to the N terminus and the other to the C terminus (Armstrong & Baldwin, 1994). The behaviour of these two peptides is completely different due to a favourable interaction with the helix dipole when the His^+ residue is close to the C terminus and *vice versa*. Figure 4C illustrates a more complex case: the pH-dependence of peptides having attractive $i,i+4$ electrostatic interactions between Asp and Arg residues placed at different positions in the peptides, as well as attractive or repulsive interactions with the helix macrodipole depending on the location of these residues (Huyghues-Despointes *et al*., 1993). This results in an increase in the helical content of some peptides upon raising the pH, and the opposite in others. The changes with the pH of the interaction energy between an aromatic residue and a His residue, as well as that of the His and the helix macrodipole are illustrated by Figure 4D (Armstrong *et al*., 1993). A more complicated case is that of a peptide designed (Forood *et al*., 1993) to study the capping box interaction (Dasgupta & Bell, 1993; Harper & Rose, 1993). In this case there are several possible electrostatic interactions including a capping box, a non-protected N terminus, an $i,i+3$ favourable attraction and an Arg at the C terminus (Figure 4E). Finally we show a peptide that contains several different types of residues as well as almost every type of electrostatic interaction (Bradley *et al*., 1990). In all the cases there is a very good agreement between the calculated and experimental values, within the limits of the experimental error. In the majority of the cases the calculated titration curves show pK values quite similar to those observed in the experimental titrations. However, some cases show a slight shift at a certain pH, between the calculated and experimental data. This is evident in Figure 4E, in which the calculated helical content increases above pH 9, while the experimental content does it above pH 8. Also in Figure 4F there is some disagreement between the calculated and experimental data around pH 6. The implementation of the pH-dependence in AGADIR allows us also to determine the pH-induced helical changes at a residue level. Figure 5 shows the very good agreement between the calculated values at a residue level and the ΔC_α upfield chemical shifts determined by NMR in a peptide analysed at two different pH values.

Discussion

Experimental conditions influence the helical behaviour of peptides as it generally occurs for systems in thermodynamic equilibrium. In peptides in solution two major factors that influence the helical content of a peptide are the temperature and the pH. Another important complex factor is the ionic strength. This factor affects electrostatic interactions and those between hydrophobic residues. There is no appropriate theory nor experimental data to account for the ionic strength influence on the hydrophobic interactions. Besides, the theoretical framework available to study the ionic strength effect on the electrostatic interactions (Debye-Huckel theory) seems to work properly only for ionic strengths less than 100 μM. The majority of the peptides in our database were analysed at low ionic strength (<0.1 M). This is close to the physiological values and in this range most of the peptides do not show large changes in their helical content. All these arguments led us to the decision of not implementing this parameter in our algorithm, until more experimental data are available.

To implement the temperature and pH dependence, we have used the physical chemistry of systems in aqueous solution and quite simple theoretical approximations. Following Occam's razor, we tried to describe these effects using the simplest possible approximations that account for the

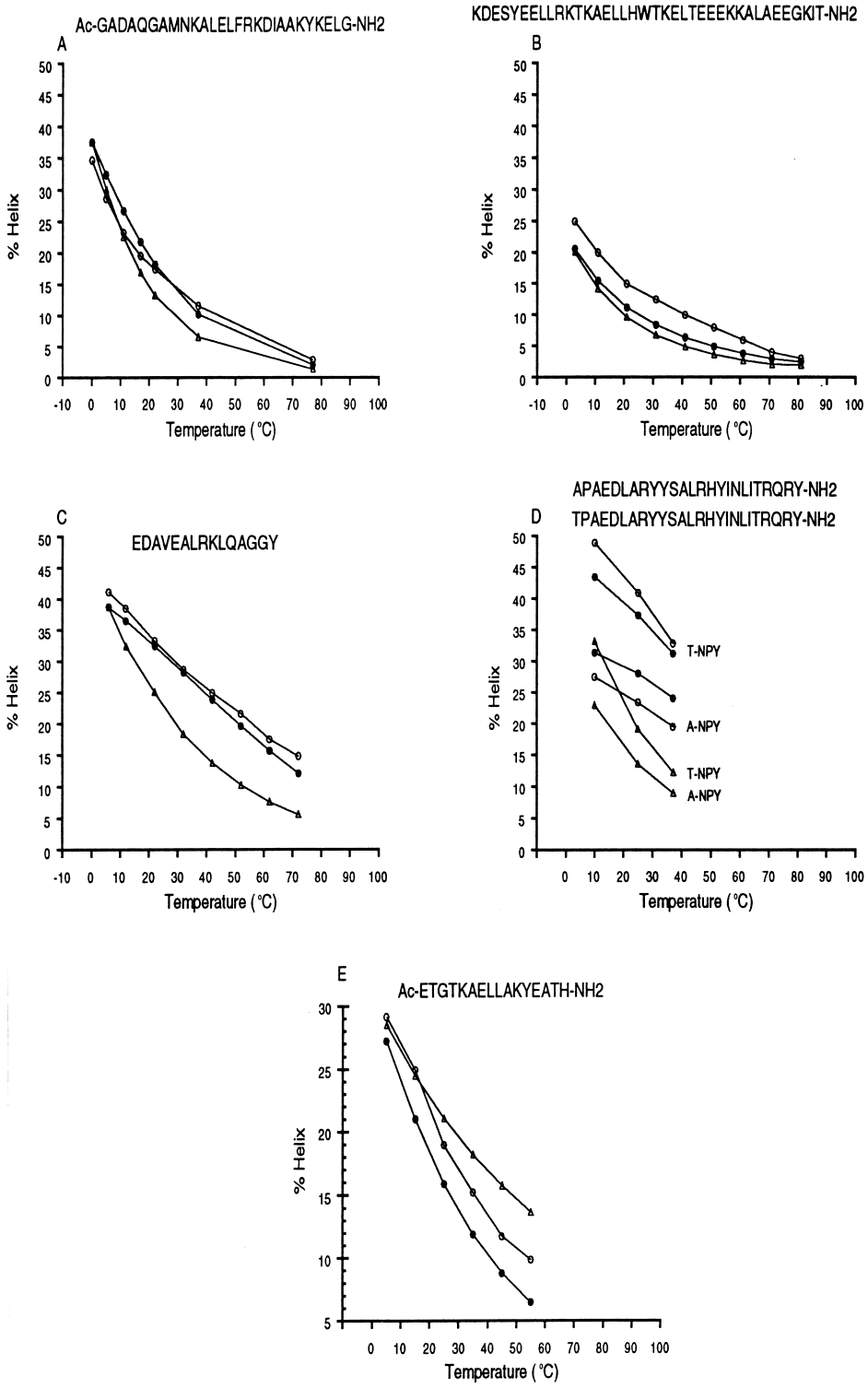

Figure 3. Comparison between the temperature-dependence of the helical behaviour of peptides corresponding to wild-type or mutant sequences derived from proteins, or designed peptides, and the calculated values. A, Peptide corresponding to the α-helix H of myoglobin analysed by Shin *et al*. (1993*a,b*). B, Peptide corresponding to a single amino acid type region of titin (residues 1977 to 2014; Musco *et al*., unpublished results). C, Mutant variant of a peptide corresponding to the α-helix 2 of CheY (see the accompanying paper, Muñoz & Serrano, 1995). D, Wild-type and Ala to Thr variant of the α-helix of neuropeptide Y (Yumoto *et al*., 1993). F, Designed peptide analysed by Bradley *et al*. (1990). The experimental values are shown as open circles. The calculated values considering the terms indicated in equations (8) to (10) are shown as filled circles. The calculated values assuming that there is not a $\Delta C_{\text{Phydrop}}$ term (equation (9)), are shown as open triangles.

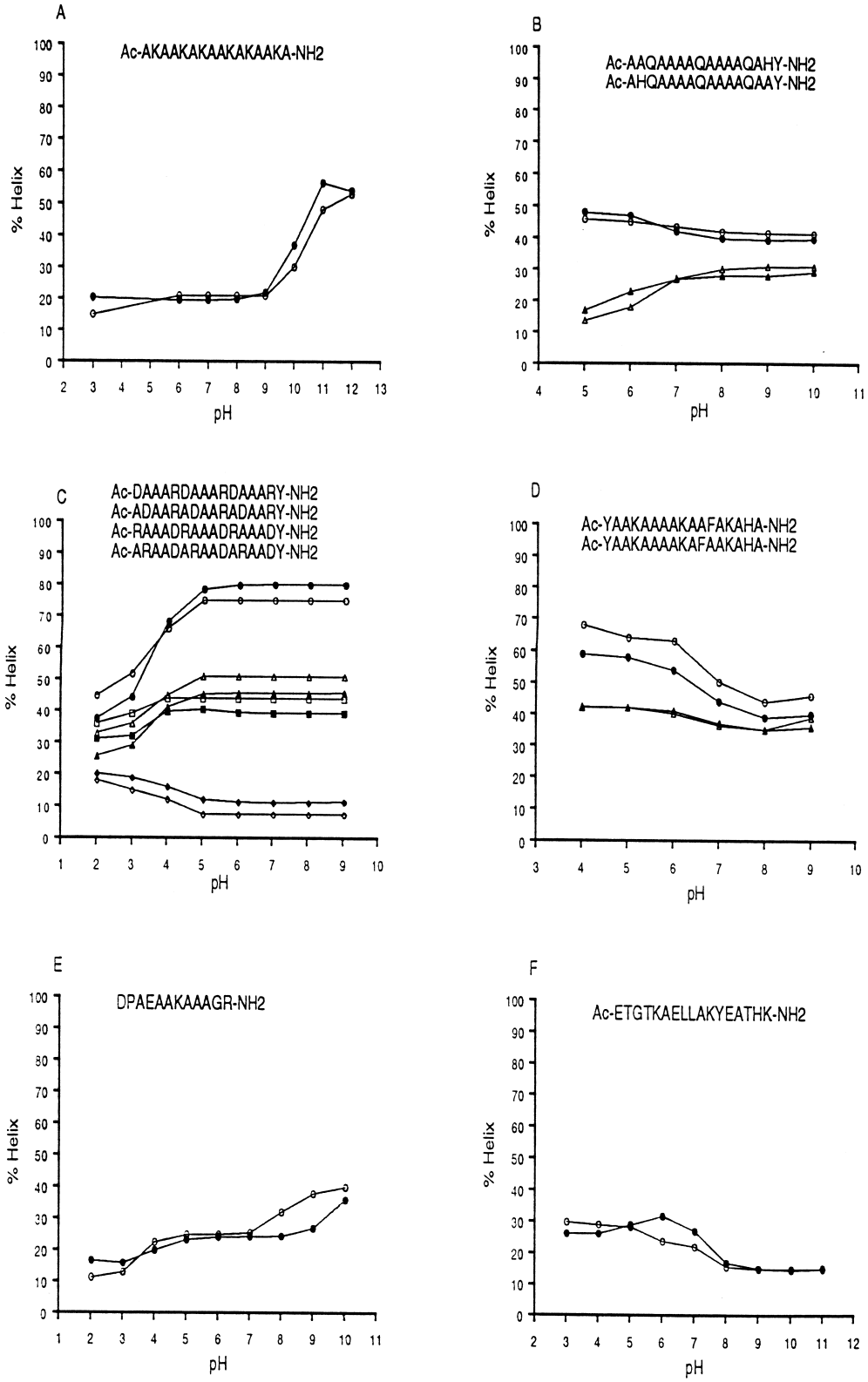

Figure 4. Comparison between the helical content of designed peptides at different pH values and the predicted values for helical content. Open circles represent the experimental values calculated as indicated in Materials and Methods. Filled symbols represent the calculated values. A, Analysed by Marqusee *et al*. (1989). B, Analysed by Armstrong & Baldwin (1993). C, Analysed by Huyghues-Despointes *et al*. (1993). D, Analysed by Armstrong *et al*. (1993). E, Analysed by Forood *et al*. (1993). F, Analysed by Bradley *et al*. (1990).

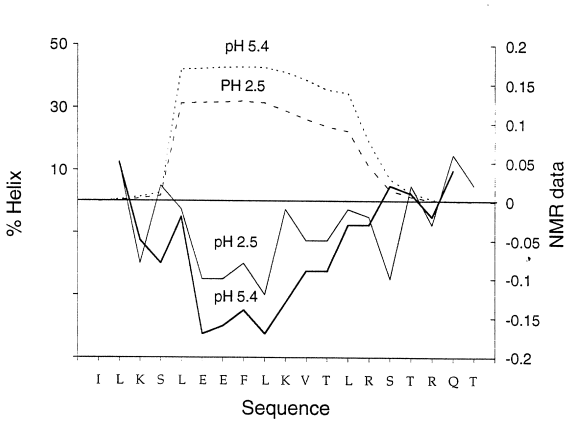

Figure 5. Diagrams showing the agreement between the calculated helicity per residue (broken lines) and the upfield chemical shifts of the C^{α} protons, with respect to random coil values (continuous line). The peptide was analysed at 2 different pH values, which are indicated in the Figure.

experimental observations. Whenever possible we have eliminated those contributions that we could reasonably think were not crucial for the correct description of the helical system (i.e. van der Waals interactions), and we have kept those that have been empirically described and were necessary for the correct description of the systems under study (i.e. heat capacities).

Temperature dependence

The simplest approximation used by other authors (Scholtz *et al*., 1991*a*), in which the nucleation factor, the hydrogen bond and the dielectric constant are assumed to be temperature independent, does not calculate well the experimental data analysed here. The difference between the calculated and experimental values could be because no heat capacity increment is considered in this approximation. In those cases where a change of state occurs accompanied by a change in solvation of different groups, it is mandatory to consider a heat capacity change (ΔCp) in order to analyse the temperature dependence of the process. The formation of an α-helix from a random-coil peptide also involves a change in solvation of different chemical groups. In principle, the total polar and apolar surface buried by each amino acid residue should be different and dependent on its side-chain (Lee *et al*., 1994). All of this, together with the fact that the change in heat capacity is very often a function of temperature (Privalov & Gill, 1988; Wintrode *et al*., 1994; Viguera *et al*., 1994), makes the theoretical analysis of the α-helix temperature dependence complicated. However, the inclusion of a unique heat capacity change (ΔCp), for the 20 amino acids (this term may be sconsidered as the average ΔCp for the different amino acids), in the framework of the modified helix-coil transition theory published previously (Muñoz & Serrano, 1994, 1995), renders an adequate description of the temperature dependence of polyalanine-based peptides. This change in heat capacity has been considered temperature-independent in a theoretical work on α-helix unfolding thermodynamics (Ooi & Obatake, 1991). Those peptides, derived from protein fragments, having side-chain to side-chain hydrophobic interactions, show a more complex behaviour upon changes in temperature than the polyalanine-based peptides. This is not surprising, since two bulky hydrophobic side-chains at $i,i+3$ and $i,i+4$ positions on an α-helix bury some hydrophobic surface and therefore an additional heat capacity change occurs. In proteins, the heat capacity increment corresponding to the solvent exposure of hydrophobic groups is highly dependent on temperature (Wintrode *et al*., 1994), and indeed we find that the introduction of a linear temperature-dependence of $\Delta C_{\mathrm{Phydrof}}$ significantly improves the calculated values.

The final values for the two heat capacity increments considered here are not far away from the values obtained from the calorimetric analysis of proteins and model compounds. The value of 1.5 cal K^{-1} mol^{-1} for ΔCp, in the unfolding direction, is very close to that determined empirically, from the thermodynamic data for protein unfolding and hydration of model compounds, for α-helices (2 cal K^{-1} mol^{-1}; Ooi & Obatake, 1991). In the case of $\Delta C_{\mathrm{Phydrop}}$ (unfolding reaction), the value of $+0.008 - 0.00005(t - 273)$ kcal K^{-1} mol^{-1} for aliphatic-aliphatic and aromatic-aliphatic interactions and of $+0.004 - 0.000025(t - 273)$ kcal K^{-1} mol^{-1} for aromatic-aromatic interactions, is very close to those determined from the calorimetric analysis of protein denaturation (0.0067 kcal K^{-1} mol^{-1} per CH_3 group; Murphy & Gill, 1991) or 0.013 kcal K^{-1} mol^{-1} per CH_3 group (Yang *et al*., 1992); 60% of these values has been reported for aromatic residues (Makhatadze & Privalov, 1990). Although these values do not need to be equal, they could be approximately similar, since in a helical peptide the hydrophobic surface buried by two mobile hydrophobic side-chains is not, on average, much higher than that of a completely buried methyl group.

The model considered here is highly simplified and the energy contributions have been assumed to be either purely enthalpic or entropic. If these terms are more complex, including entropic and enthalpic factors, then it would be necessary to determine their relative contribution. Moreover, we have considered that the heat capacity increments are similar for all the amino acids as well as for all the aliphatic or aromatic side-chain interactions, which is clearly not correct. We have analysed here the helical dependence on temperature of four peptides derived from proteins, with helical content between 25 and 45%. Two of these peptides have been studied by us in this paper (CheY2 and the titin fragment). There are other examples in the literature with less than 20% helical content at 0°C. However, these peptides cannot be used to calibrate the heat capacities considered here since we are missing a large part of the transition

region. Unfortunately it is difficult to find many peptides derived from protein fragments that possess a significant amount of helical population (>25%). The analysis of polyalanine-based peptides, with specific hydrophobic interactions, could allow us to see whether there is need for a more complex model in which additional terms are considered.

pH dependence

The dependence of the helical content of short polypeptide chains, without tertiary interactions, upon changes in the pH is the other main point addressed in this work. The pH value influences the ionization degree of charged groups and consequently affects the putative interactions between them. In a short polypeptide chain there are several groups that can be charged; the end groups when they are not protected and the side-chains of Asp, Glu, Cis, Tyr, Lys, Arg and His. In this work we have considered only those that are charged at pH values close to pH 7.0 (Asp, Glu, Lys, Arg and His). The effect of the pH on the ionisation of these groups is described by the classical equations shown above (equations (14) and (15)). However, the presence of other ionisable groups affects the ionisation state of the group being considered and *vice versa*, making necessary a further calculation of its pK_a taking into account the electrostatic environment. Calculation of pK_a values of charged residues in proteins is today an important field of study. Several approximations involving statistical mechanics have been developed (Yang & Honig, 1993; Antosiewicz *et al*., 1994). These calculations have large errors (around 0.9 unit of pK_a) and the computer time that would be required to do them for all the possible helical conformations in a polypeptide makes it prohibitive. These reasons prompted us to simplify the calculation of the pK_a values of charged residues. The main simplification consists of including only the effects arising from the interactions with the helix macrodipole. Despite the simplifications introduced, the implementation of the theory seems successful in describing the pH-induced effects in the helicity of all the experimentally studied peptides. Some of these peptides are very complicated and have several different types of interactions. More important, in this case there has not been any refinement of the parameters involved and therefore all the peptides shown here could be considered a blind test. It is possible that in peptides with a large number of charges the calculation of the pK of charged groups accumulates severe errors, since we are not taking into account side-chain to side-chain electrostatic interactions, but there is no straightforward way to overcome the problem.

Conclusions

The work presented here indicates that classical physicochemical theories introduced in a simple statistical mechanics framework with an empirical parametrization may, in principle, describe the helical dependence on temperature and pH of monomeric peptides in solution. This allows us to calculate, at different pH values and temperatures, the helical content at a residue level of short linear polypeptides that are monomeric.

Materials and Methods

Experimental procedures

Peptide synthesis

The solid-phase synthesis of the peptides was performed on an Abimed AMS422 multiple peptide synthesizer using Fmoc chemistry and PyBOP activation at a 0.025 mmol scale. After synthesis was completed, protecting groups were removed and the peptide chains were cleaved from the resin with a mixture of 10 ml of TFA, 0.75 g of phenol, 0.25 ml of EDT, 0.5 ml of thioanisole and 0.5 ml of water for three hours. The peptides were purified on a Vydac C-18 reverse phase column (20 mm × 250 mm, 0.01 mm particle size) at a flow-rate of 10 ml / min. Solvent A was water containing 0.1% (v / v) TFA and solvent B was 70% acetonitrile, 0.1% TFA in water. Peptide homogeneity (>98%) was determined by HPLC using an acetonitrile gradient of 0.7% / min. The peptide composition was confirmed by amino acid analysis and the molecular mass was checked by matrix-assisted laser desorption time-of-flight mass spectrometry.

Peptide concentration

The concentration of the different peptides was determined by amino acid analysis, or UV absorbance using the method of Gill & von Hippel (1989). The error is around 10%.

Circular dichroism analysis

CD spectra were recorded on a Jasco-710 instrument at a temperature of 5°C. The peptides (roughly 1.5 mg) were dissolved in 1 ml of 2.5 mM sodium phosphate buffer (pH 7.0), unless otherwise indicated. To check for concentration dependence of the CD spectra, different dilutions of the peptides (10 to 750 mM), using cuvettes with different pathlengths (0.1 mm to 0.5 cm), were scanned. CD spectra in the range 190 to 250 nm were obtained using the continuous scan option (100 nm / min scan speed), with a one second response time and taking points every 0.1 nm. For every sample we took 30 scans and the experiment was repeated three times on different days. The ellipticity was calibrated using D-10-camphorsulphonic acid. The thermal denaturation curves have been obtained from several discrete measurements at different temperatures, as described above.

Determination of the helical percentage from the circular dichroism spectra, at different temperatures

To estimate the helical population of the different peptides at 0°C, (%$Helix_{0C}$), we used the mean residue ellipticity at 222 nm, taking into account the peptide length (Chen *et al*., 1974):

$$\%Helix_{0C} = 100\ Ellipticity\ /\ (39{,}500\ (1 - 2.57\ /\ n), \quad (1)$$

where n is the number of residues in the peptide. It has been reported that the mean residue ellipticity at 222 nm

decreases in a random coil peptide with temperature following the equation:

$$Ellipticity_{coil} = Ellipticity_{0Ccoil} - 45\ t, \quad (2)$$

where $Ellipticity_{OCcoil}$ is the ellipticity of a random coil peptide at 0°C (400 deg cm^{-1} dmol^{-1}), $Ellipticity_{Ccoil}$ is the ellipticity of a random coil peptide at temperature t, and t is temperature in degrees Celsius (Scholtz *et al*., 1991*c*). On the other hand, the ellipticity of the helical conformation increases with temperature following the equation:

$$Ellipticity_{helix} = Ellipticity_{OChelix} + 100\ t, \quad (3)$$

where $Ellipticity_{OChelix}$ is the ellipticity of a peptide having 100% α-helical conformation at 0°C, $Ellipticity_{helix}$ is the ellipticity of a peptide having 100% α-helical conformation at temperature t, and t is temperature in degrees Celsius (Scholtz *et al*., 1991*c*). To calculate the helical average percentage of a particular peptide at a certain temperature we used:

$$\%Helix = 100 / (1 + ((Ellipticity - Ellipticity_{helix}) / (Ellipticity_{coil} - Ellipticity))), \quad (4)$$

where *Ellipticity* is the mean residue ellipticity of the peptide at a particular temperature.

Theoretical procedures

Helical dependence on temperature

The difference in free energy of a particular helical conformation, with respect to the random coil, might be described by the following classical equation:

$$\Delta G_{Hel} = \Delta H - t\Delta S, \quad (5)$$

where ΔH is the difference in enthalpy between the helix and the random coil states, ΔS is the difference in entropy between both states and t is the temperature considered.

In classical helix-coil transition formalism, where side-chain to side-chain interactions, capping effects and interactions with the helix macrodipole are not considered, the ΔH term arises mainly, but not exclusively, from the main-chain to main-chain $i,i+4$ hydrogen-bonding network and ΔS reflects the entropic cost of fixing the amino acid residues in helical dihedral angles. This has been the theoretical basis of the approach used by Baldwin and collaborators to check the validity of the helix-coil transition on the temperature-induced unfolding transition of helical peptides of different lengths (Scholtz *et al*., 1991*c*). On the other hand, it is known that any process involving a change of state in aqueous solution, as happens for the helix-coil transition, also involves a change in solvation of different groups. In particular, in an α-helix certain groups of the polypeptide chain are less exposed to the solvent than in the random-coil state. Amide and carbonyl groups from the backbone plus the C^{α} and C^{β} atoms of all the residues (with the exception of Gly for the C^{β}), are partly buried in the helix and solvent-exposed in the random coil. The result of the solvation change is a change in the heat capacity (Cp) between both states (ΔCp). The existence of ΔCp implicates that ΔH and ΔS are functions of temperature, as is expressed in equations (6) and (7):

$$\Delta H = \Delta H^{ref} + \Delta C\text{p}\ (t - t^{ref}) \quad (6)$$

$$\Delta S = \Delta S^{ref} + \Delta C\text{p}\ \ln(t/t^{ref}), \quad (7)$$

where ΔH^{ref} is the difference in enthalpy between both states at the isoenthalpic temperature (t^{ref}). ΔS^{ref} is the difference in entropy between both states at the isoentropic temperature (t^{ref}), and ΔCp is the change in heat capacity between both states.

Besides, ΔCp might also be temperature-dependent, making the analysis more complex. There are several calorimetric analyses on the unfolding transition of different proteins describing the heat capacity changes of ΔCp with the temperature (Privalov & Gill, 1988; Livingstone *et al*., 1991; Wintrode *et al*., 1994; Viguera *et al*., 1994). Typically, ΔCp shows a non-linear dependence on temperature for those proteins analysed, but it might be approximated to a constant term when the temperature does not exceed 75°C (Wintrode *et al*., 1994). Definition of ΔCp as a constant term facilitates considerably the theoretical analysis of protein thermal denaturation and several studies have been carried out in that direction (Murphy & Gill, 1991; Ooi & Obatake, 1991; Yang *et al*., 1992). In our case, since helical peptides are mostly unfolded at 75°C and following the idea of using the simplest possible assumptions, we have defined ΔCp as a constant term not dependent upon temperature changes. The initial value used in our parameterization has been extracted from the theoretical analysis made by Ooi & Obatake (1991), specifically addressed to the thermal unfolding of α-helices (2 cal K^{-1} mol^{-1}). Once a value of ΔCp is assigned, it is straightforward to introduce this term into our algorithm (equations (8) and (9)), since the term previously described ΔG_{Hbond} (Muñoz & Serrano, 1994, 1995), corresponds mainly to enthalpic interactions arising from hydrogen-bond formation and van der Waals interactions, $\Delta G_{Hbond} \sim \Delta H_{Hbond}$. The term ΔG_{Int}, representing the entropic cost of fixing the residues in helical dihedral angles (Muñoz & Serrano, 1994, 1995), might be considered as arising uniquely from the loss of conformational entropy, $\Delta G_{Int} = -t\Delta S_{confor}$:

$$\Delta G_{HBond} = \Delta H^{ref}_{HBond} + \Delta C\text{p}\ (t - t^{ref})) \quad (8)$$

$$\Delta G_{Int} = -t\Delta S_{confor} = -t(\Delta S^{ref}_{confor} + \Delta C\text{p}\ \ln(t/t^{ref})). \quad (9)$$

t^{ref} is the reference temperature that in our case is 0°C, because the initial parameters have been calculated and refined for this temperature (Muñoz & Serrano, 1994, 1995). ΔG^{ref}_{HBond} is the difference in the summation of the net free energy contribution of all the main-chain hydrogen bonds within the helical segment at t^{ref}. ΔS^{ref}_{confor} is the cost in conformational entropy of fixing all the amino acid residues forming the helical segment at t^{ref} and ΔCp is the change in heat capacity as defined above.

There are several other energy contributions that need to be taken into account in order to achieve a correct description of the helical formation process (Muñoz & Serrano, 1994, 1995; see equation (10) in Muñoz & Serrano, 1995). ΔG_{nonH} refers to the contribution to the helical segment stability of the flanking residues (N-cap and C-cap). This contribution is particularly important in those cases involving side-chains that could make hydrogen bonds (Ser, Asn, Asp and Thr at the N-cap, His, Lys and Arg at the C-cap, and Glu and Gln at position $N+3$ in the capping-box motif; Dasgupta & Bell, 1993; Harper & Rose, 1993). In these cases we have considered a similar temperature dependence as that defined for ΔG_{HBond}. This dependence stands also for the hydrogen-bond component due to the protection of the peptide ends (acetylation and succinylation of the N terminus and amidation of the C terminus; Muñoz & Serrano, 1994, 1995). The temperature-dependencies for other possible cases (i.e. possible hydrogen bond between a Gln and Asp at positions i and $i+4$; Huyghues-Despointes *et al*., 1993) are neglected, since their contribution to the total energy is in principle smaller and there is not a clear division between its entropic and enthalpic components.

An important contribution to the final stability of the helical segment is ΔG_{SD}, which accounts for all the interactions between the side-chains in the helical segment. Interactions between side-chains might be of different kinds: hydrophobic interactions, electrostatic interactions and others. The hydrophobic interactions arise from the fact that two side-chains bury some hydrophobic surface upon interacting with each other. This means that every different hydrophobic interaction should have, in principle, a different value for $\Delta Cp_{hydroph}$. Since the free energy values for the majority of the hydrophobic interactions are very similar (Muñoz & Serrano, 1995), and the energy of the interaction should be proportional to the area buried, then the change in hydrophobic heat capacity was considered constant and equal for all hydrophobic interactions. The only exception to this is the distinction between interactions involving aliphatic side-chains or aromatic side-chains. This is so because different heat capacity increments due to aliphatic or aromatic groups have been experimentally determined in proteins. For aliphatic residues $\Delta Cp_{hydrop} = 13.4$ cal mol^{-1} K^{-1} (Yang *et al*., 1992), or 7 cal mol^{-1} K^{-1} (Murphy & Gill, 1991). For aromatic groups, Makhatadze & Privalov (1990) have estimated that the heat capacity increment is 60% of that of aliphatic groups. The difference in free energy in the hydrophobic interactions arises fundamentally from the differences in the entropic term and therefore the formulation of the temperature-dependence is immediate:

$$\Delta G_{hydroph} = -t\Delta S_{hydroph} = -t(\Delta S^{ref}_{hydroph} + \Delta Cp_{hydroph} \ln(t/t^{ref})), \quad (10)$$

where $\Delta G_{hydroph}$ is the difference in free energy for the hydrophobic interaction. $\Delta S_{hydroph}$ reflects the gain in entropy due to the hydrophobic interaction. t is the temperature and t^{ref} is the reference temperature (0°C). The experimental values described above have been used to parameterise the changes in heat capacity for both aliphatic and aromatic interactions (see refinement of heat capacities).

The interactions between charged side-chains at positions $i,i+3$ and $i,i+4$ are grouped with the interactions of charged side-chains and/or unprotected ends with the helix macrodipole, and they are considered electrostatic interactions. These interactions normally do not involve large changes in solvation of the participating groups (they are long-range interactions), but they are affected by temperature. Electrostatic interactions are inversely proportional to the dielectric constant of the medium. This constant shows a well-known exponential decrease with temperature, with the consequence that at higher temperatures both the attractions and repulsions are stronger. The effect of temperature on the dielectric constant is expressed in the following equation:

$$\epsilon = 88.1(e^{-0.004314\Delta t}), \quad (11)$$

where ϵ is the dielectric constant at a given temperature t. The term Δt refers to the difference between the temperature t and t^{ref} (0°C). The attractive or repulsive electrostatic interactions will then vary according to the following equation:

$$\Delta G_t = \Delta G_{0C}(e^{0.004314\Delta t}), \quad (12)$$

where ΔG_{0C} is the free energy of interaction at 0°C, and ΔG_t is the free energy of interaction at the temperature t.

Other interactions between side-chains of residues in positions $i,i+3$ and $i,i+4$ are assumed not to be temperature-dependent. This clearly is an oversimplification, but since there is no clear way to determine their entropic and enthalpic components, there are no experimental data about their temperature-dependence and their contribution to the final energy is normally small, we decided to neglect them.

Helical dependence on the pH

The average helical content of a peptide containing ionisable groups depends on the electrostatic interaction between these groups, as well as on the interaction of these groups with the helix macrodipole (Scholtz *et al*., 1993). These interactions obviously change with the degree of ionisation and consequently with the pH of the solution.

We can contemplate two different cases depending on whether the interactions are repulsive or attractive. In the first case we considered that the repulsion energy disappears completely when the charged groups become neutral. However, in the second case the situation is more complicated. For the $i,i+3$, and $i,i+4$ attractive interactions, there is experimental evidence that when one of the charged groups becomes neutral, there is still an attractive interaction probably due to the formation of a hydrogen bond (Scholtz *et al*., 1993). We have considered that in these cases 60% of the attractive interaction disappear when one of the two charged groups becomes neutral. The electrostatic interactions ($i,i+1$) are considered to be null when one of the charged groups becomes neutral, since their side-chains cannot hydrogen-bond to each other. From these premises to calculate the changes in the electrostatic free energy of interaction with the pH, we need to determine only the pK_a of the different ionisable groups on the peptide.

The intrinsic pK_a of a single ionisable group in a polypeptide chain changes from its standard value depending on the electrostatic environment:

$$pK_a = pK^0_a - (\gamma\Delta\Delta G^{env}/2.3\,Kt), \quad (13)$$

where pK^0_a is the pK_a of the group being considered when there are no other charged groups, γ is 1 or -1, depending on whether the group is basic or acidic, $\Delta\Delta G^{env}$ is the free energy associated with charging the group in the polypeptide compared with that when charging it in a model compound. This last term depends on the solvation of the group, the interaction with permanent dipoles and the interaction with other charged groups.

In principle, the titration curve of each ionisable group in the protein can be obtained from statistical mechanics considering all the possible states (Yang & Honig, 1993). However, the complexity of this approach increases exponentially with the number of titrable groups, and it is impractical when it reaches several tens of residues. There are no clear-cut experiments about the mutual effect on the pK_a values of two groups placed at different distances on an α-helix, so we decided not to include this interaction effect on the pK_a of charged groups. This clearly introduces an error in the estimation of the pK_a, but we assume that this should not be large except in very special cases in which several charged groups are located in proximity to the group whose pK_a is being calculated. In a peptide in solution the charged groups should be very accessible so we can discard also the solvent exposure contribution to $\Delta\Delta G^{env}$. This means that we are considering only the effect of the helix macrodipole on the pK_a of a charged group.

The term $\Delta\Delta G^{env}$ for Asp, Glu, Lys, Arg and His residues, located at different distances from the helix dipole, was determined in the accompanying paper (Muñoz & Serrano, 1995). Then, we need to know the standard pK_a values of the groups being considered. We have used the following values, obtained from titration of peptides without

secondary structure or other charged groups: Asp, 3.86; Glu, 4.25; Arg, 12.48; Lys, 10.53; His, 6.50; N-terminal, 9.6; C-terminal, 2.2. Once the pK_a of the group is known, it is straightforward to calculate the ionisation degree of the group using the classical equations:

$$I_{\text{acidic}} = 1/(1 + (10^{\text{pH}}/10^{\text{p}K_a})) \quad (14)$$

$$I_{\text{basic}} = 1/(1 + (10^{\text{p}K_a}/10^{\text{pH}})), \quad (15)$$

where I_{acidic} is the ionisation degree for an acidic group and I_{basic} is that for a basic group. The dependence of the free energy of interaction between two charged groups, with the pH, is indicated in equation (16). That between a charged group and the helix dipole, is indicated in equation (17):

$$\Delta G_{\text{int}}^{\text{pH}} = \Delta G_{\text{int}}^{\text{pHref}} I^a I^b \quad (16)$$

$$\Delta G_{\text{int}}^{\text{pH}} = \Delta G_{\text{int}}^{\text{pHref}} I \quad (17)$$

where $\Delta G_{\text{int}}^{\text{pHref}}$ is the interaction energy at pH 7.0.

Parameter refinement of the heat capacity increments

The values for the different heat capacities were refined by looking at the temperature denaturation of several polyalanine-based peptides (Marqusee *et al*., 1989; Scholtz *et al*., 1991*b*; Merutka & Stellwagen, 1990, 1991; Zhou *et al*., 1993), as well as some peptides derived from proteins (Shin *et al*., 1993*a,b*; Musco *et al*., unpublished results; Muñoz & Serrrano, 1994; Yumoto *et al*., 1993). In our algorithm the helix-coil transition is defined in the direction of helix formation (folding reaction). Besides, the reference temperature in our case is 0°C and not the isoenthalpic and isoentropic temperatures. The last ones are very similar for proteins (Yang *et al*., 1992) and of the order of 380 K. This results in our reference being on the other side of the transition. A final value of $-1.5\ \text{cal K}^{-1}\ \text{mol}^{-1}$ (folding direction), was obtained for ΔCp after refinement of the initial value given by Ooi & Obatake (1991).

There is experimental evidence that the term $\Delta C_{\text{Phydrop}}$ is not constant in protein unfolding (Wintrode *et al*., 1994). Initially we assumed it to be constant (see Materials and Methods, Theoretical procedures), but during the refinement it became obvious for the necessity of considering it as a linear function of temperature, see equation (18). For the interaction between aliphatic residues or between aliphatic (Leu, Ile, Val and Met) and aromatic residues (Tyr, Phe and Trp), the heat capacity increment changes according to:

$$\Delta C_{\text{Phydrop}} = -8 + 0.05(t - 273), \quad (18)$$

where the term $\Delta C_{\text{Phydrop}}$ is expressed in $\text{cal K}^{-1}\ \text{mol}^{-1}$.

References

Antosiewicz, J., McCammon, A. & Gilson, M. K. (1994). Prediction of pH-dependent properties of proteins. *J. Mol. Biol.* **238**, 415–436.

Armstrong, K. M. & Baldwin, R. L. (1994). Charged histidine affects α-helix stability at all positions in the helix by interacting with the backbone charges. *Proc. Nat. Acad. Sci., U.S.A.* **90**, 11337–11340.

Armstrong, K. M., Fairman, R. & Baldwin, R. L. (1993). The (*i*,*i* + 4) Phe-His interaction studied in an alanine-based α-helix. *J. Mol. Biol.* **230**, 284–291.

Bradley, E. K., Thomason, J. F., Cohen, F. E., Kosen, P. A. K. & Kuntz, I. D. (1990). Studies of synthetic peptides using circular dichroism and nuclear magnetic resonance. *J. Mol. Biol.* **215**, 607–622.

Chen, C. C., Zhu, Y., King, J. A. & Evans, L. (1992). A molecular thermodynamic approach to predict the secondary structure of homopolypeptides in aqueous systems. *Biopolymers*, **32**, 1375–1392.

Chen, Y. H., Yang, J. T. & Chau, K. H. (1974). Determination of the helix and β-form of proteins in aqueous solution by circular dichroism. *Biochemistry*, **13**, 3350–3359.

Dasgupta, S. & Bell, J. (1993). Design of helix ends. *Int. J. Pept. Res.* **41**, 499–511.

Forood, B., Feliciano, E. J. & Nambiar, K. P. (1993). Stabilization of α-helical structures in short peptides *via* end capping. *Proc. Nat. Acad. Sci., U.S.A.* **90**, 838–842.

Gill, S. C. & von Hippel, P. H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182**, 319–326.

Harper, E. T. & Rose, G. D. (1993). Helix stop signals in proteins and peptides: the capping box. *Biochemistry*, **32**, 7605–7609.

Huyghues-Despointes, B. M. P., Scholtz, J. M. & Baldwin, R. L. (1993). Helical peptides with three pairs of Asp-Arg and Glu-Arg residues in different orientations and spacings. *Protein Sci.* **2**, 80–85.

Lee, K. H., Freire, E. & Amzel, L. M. (1994). Estimation of changes in side chain configurational entropy in binding and folding: general methods and application to helix formation. *Proteins: Struct. Funct. Genet.* In the press.

Livingstone, J. R., Spolar, R. S. & Record, T. M. (1991). Contribution to the thermodynamics of protein folding from the reduction in water-accessible nonpolar surface area. *Biochemistry*, **30**, 4237–4244.

Lockhart, D. J. & Kim, P. S. (1993). Electrostatic screening of charge and dipole interactions with the helix backbone. *Science*, **260**, 198–202.

Makhatadze, G. I. & Privalov, P. L. (1990). Heat capacity of proteins. I. Partial molar heat capacity of individual amino acid residues in aqueous solution. *J. Mol. Biol.* **213**, 375–384.

Marqusee, S., Robbins, V. H. & Baldwin, R. L. (1989). Unusually stable helix formation in short alanine-based peptides. *Proc. Nat. Acad. Sci., U.S.A.* **86**, 5286–5290.

Merutka, G. & Stellwagen, E. (1990). Positional independence and additivity of amino acid replacements on helix stability in monomeric peptides. *Biochemistry*, **29**, 894–898.

Merutka, G. & Stellwagen, E. (1991). Effect of amino acid ion pairs on peptide helicity. *Biochemistry*, **30**, 1591–1594.

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

Muñoz, V. & Serrano, L. (1995). Elucidating the folding problem of helical peptides using empirical parameters. II. Helix macrodipole effects and rational modification of the helical content of natural peptides. *J. Mol. Biol.* **245**, 275–297.

Murphy, K. P. & Gill, S. J. (1991). Solid model compounds and the thermodynamics of protein unfolding. *J. Mol. Biol.* **222**, 699–709.

Ooi, T. & Obatake, M. (1991). Prediction of the thermodynamics of protein unfolding: the helix-coil transition of poly(l-alanine). *Proc. Nat. Acad. Sci., U.S.A.* **88**, 2859–2863.

Presta, L. G. & Rose, G. D. (1988). Helix signals in proteins. *Science*, **240**, 1632–1641.

Privalov, P. L. & Gill, S. J. (1988). Stability of protein structure and hydrophobic interactions. *Advan. Protein Chem.* **39**, 191–234.

Scholtz, J. M., Marqusee, S., Baldwin, R. L., York, E. J., Stewart, J. M., Santoro, M. & Bolen, D. W. (1991*a*). Calorimetric determination of the enthalpy change for the α-helix to coil transition of an alanine peptide in water. *Proc. Nat. Acad. Sci., U.S.A.* **88**, 2854–2858.

Scholtz, J. M., York, E. J., Stewart, J. M. & Baldwin, R. L. (1991*b*). A neutral water-soluble, α-helical peptide: the effect of ionic strength on the helix-coil equilibrium. *J. Amer. Chem. Soc.* **113**, 5102–5104.

Scholtz, M., Qian, H., York, E. J., Stewart, J. M. & Baldwin, R. L. (1991*c*). Parameters of helix-coil theory for alanine-based peptides of varying chain lengths in water. *Biopolymers*, **31**, 1463–????

Scholtz, J. M., Quian, H., Robbins, V. H. & Baldwin, R. L. (1993). The energetics of ion-pair and hydrogen-bonding interactions in a helical peptide. *Biochemistry*, **32**, 9668–9676.

Shin, H. C., Merutka, G., Waltho, J. P., Tennant, L. L., Dyson, H. J. & Wright, P. E. (1993*a*). Peptide models of protein folding initiation sites. 3. The G-H helical hairpin of myoglobin. *Biochemistry*, **32**, 6356–6364.

Shin, H. C., Merutka, G., Waltho, J. P., Wright, P. E. & Dyson H. J. (1993*b*). Peptide models of protein folding initiation sites. 2. The G-H turn region of myoglobin act as a helix stop signal. *Biochemistry*, **32**, 6348–6355.

Viguera, A. R., Martinez, J. C., Filimonov, V., Mateo, P. & Serrano, L. (1994). Thermodynamic and kinetic analysis of the SH3 domain of spectrin. *Biochemistry*, **33**, 2142–2150.

Wintrode, P. L., Makhatadze, G. I. & Privalov, P. L. (1994). Thermodynamics of ubiquitin unfolding. *Proteins: Struct. Funct. Genet.* **18**, 246–253.

Yang, A. & Honig, B. (1993). On the pH dependence of protein instability. *J. Mol. Biol.* **231**, 459–474.

Yang, A., Sharp, K. A. & Honig, B. (1992). Analysis of the heat capacity dependence of protein folding. *J. Mol. Biol.* **227**, 889–900.

Yumoto, N., Murase, S., Hattori, T., Yamamoto, H., Tatsu, Y. & Yoshikawa, S. (1993). Stabilization of α-helix in C-terminal fragments of neuropeptide Y. *Biochem. Biophys. Res. Commun.* **196**, 1490–1495.

Zhou, N. E., Kay, C. M., Sykes, B. D. & Hodges, R. S. (1993). A single-stranded amphipatic α-helix in aqueous solution: design, structural characterisation, and its application for determining α-helical propensities of amino acids. *Biochemistry*, **32**, 6190–6197.

Edited by A. R. Fersht

(Received 28 June 1994; accepted 15 September 1994)