Contributions of the ionization states of acidic residues to the stability of the coiled coil domain of matrilin-1

Sonja A. Dames[a], Richard A. Kammerer[b], Detlef Moskau[c], Jürgen Engel[b], Andrei T. Alexandrescu[a,*]

[a] *Department of Structural Biology, Biozentrum, University of Basel, Basel CH-4056, Switzerland*
[b] *Department of Biophysical Chemistry, Biozentrum, University of Basel, Basel CH-4056, Switzerland*
[c] *Bruker AG, Industriestrasse 26, Faellanden CH-8117, Switzerland*

Received 19 November 1998; received in revised form 25 January 1999

Abstract The pK_a values of eight glutamic acid residues in the homotrimeric coiled coil domain of chicken matrilin-1 have been determined from 2D H(CA)CO NMR spectra recorded as a function of the solution pH. The pK_a values span a range between 4.0 and 4.7, close to or above those for glutamic acid residues in unstructured polypeptides. These results suggest only small favorable contributions to the stability of the coiled coil from the ionization of its acidic residues.

Key words: Cartilage matrix protein; Ion pair; Salt bridge; Electrostatic; Stability; Specificity; Ionization constant

1. Introduction

The coiled coil is an oligomeric protein folding motif consisting of two to five amphipatic α-helices associated into a left-handed superhelix [1–3]. The amino acid sequences of coiled coils are arranged in distinctive $(abcdefg)_n$ heptad repeats [2–4]. Positions a and d are preferentially occupied by non-polar residues which form a hydrophobic core. The remaining positions are usually hydrophilic. Sites e and g are often occupied by charged residues. While the aliphatic portions of these residues participate to some extent to the hydrophobic core, their charged side chain groups have the potential to form interhelical i+5′ ion pairs (e.g. between residue g on chain 1, and residue e′ on chain 2) [5,6,9]. Additional ion pairs within the α-helix monomers, can form between oppositely charged residues separated by an i+3 or i+4 sequence spacing [6,7]. Electrostatic interactions have the potential to modulate the stability, as well as the oligomerization state specificity of coiled coils [6,8–11].

The C-terminal domain of matrilin-1 (also named cartilage matrix protein) forms a stable ($T_m > 100$°C), disulfide-linked, homotrimeric coiled coil [12,13]. The amino acid sequence of the coiled coil domain from the chicken protein (CMPcc) is shown in Fig. 1. Three pairs of oppositely charged residues (Lys-16-Glu-19, Glu-29-Lys-33, Lys-33-Glu-36) have an i+3 or i+4 sequence spacing and have the potential to participate in intrahelical ion pairs [6,7]. A possible interhelical i+5′ ion pair could form between Arg-34 in a g position, and Glu-39 in the e′ position of the following heptad [5,6]. The solution structure of CMPcc has recently been solved by NMR [14]. The α-helix monomers of the coiled coil run from Ser-9 to Ile-43 and are capped at their N-termini by a ring of three interchain disulfide bonds (formed between Cys-5 and Cys-7′). The N-terminal residues upstream of Cys-5 are disordered and extend away from the long axis of the coiled coil [14]. The α-helix monomers are associated through an extensive hydrophobic interface formed by the 'knobs-into-holes' packing of a and d layers from neighboring polypeptide chains. The backbone, together with the side chains of the hydrophobic core residues are precisely defined in the NMR structures. The side chains of charged residues which are all surface exposed, however, are relatively poorly defined. Using the criterion of a distance shorter than 4.0 Å between the side chain nitrogen and oxygen atoms of oppositely charged residues [15], we examined the ensemble of 20 lowest energy NMR structures for possible ion pairs. Based on the NMR structures, if any ion pairs exist in CMPcc, they are only fractionally populated. The pairs Lys-16-Glu-19 and Glu-29-Lys-33 meet the 4.0 Å cut-off in 10 and 20% of the structures, respectively. An interchain Arg-34-Glu-39 contact is present in 20% of the structures. When all of the remaining oppositely charged residues in the protein are considered, none comes within ion pairing distance in more than 10% of the NMR structures. Although this analysis suggests that the propensity for ion pair formation in CMPcc is weak, the lack of uniquely defined ion pairs could be a consequence of the low structural precision for surface side chains in the NMR structures [14].

More direct information on electrostatic interactions and on their contributions to protein stability can be obtained from NMR measurements of residue-specific pK_a values. The pK_a value of an ionizable group in a protein may be shifted from its intrinsic value by its environment. Factors which contribute to pK_a shifts include attractive and repulsive electrostatic interactions with other charged groups, hydrogen bonding interactions and solvent exposure. By thermodynamic linkage it is possible to relate the pK_a shift to the contribution from the ionization of a titratable group to the change in Gibbs free energy upon folding [9,11,16,17]. For example, a negatively charged group that makes a favorable contribution to the protein stability will have its pK_a shifted to a lower value in the folded compared to the unfolded state. Conversely, a negatively charged group that destabilizes the folded state will have a raised pK_a in the folded compared to the unfolded state [9,16]. Here we report NMR-derived pK_a values for the eight glutamates in CMPcc, at 50°C and at a physiological ionic strength of 150 mM NaCl, and assess the contributions made by the ionization of these residues to the stability of the coiled coil.

*Corresponding author. Fax: (41) (61) 267 2109.
E-mail: alexandrescu@ubaclu.unibas.ch

PII: S0014-5793(99)00186-6

2. Materials and methods

Uniformly ^{13}C-labelled recombinant chicken CMPcc was expressed as a 6×His-tagged protein in *Escherichia coli* and purified by affinity chromatography on Ni^{2+}-Sepharose [13]. The 6×His-tag was removed by thrombin cleavage. As in previous studies [13,14], the recombinant CMPcc protein used in this work contains the four extraneous N-terminal residues GSHM. These residues are disordered based on NMR data [13,14] and are not included in the numbering scheme of this paper.

NMR spectra were obtained on a Bruker Avance spectrometer operating at 600 MHz. A temperature of 50°C was used for all measurements to allow comparisons with previous work [13,14]. All samples were dissolved in D_2O. The pH dependence of Glu Hγ and Cδ resonances were measured for samples containing 0.3 mM CMPcc trimer and 150 mM NaCl, by using a 2D H(CA)CO version of the 3D HCACO experiment [18], modified to optimize the detection of coupling between adjacent ^{13}Cβ/γ and ^{13}CO side chain resonances [19]. Frequency discrimination in ω1 was achieved with the TPPI method. Hard, off resonance phase modulated squared pulses (square-128) and sin(x)/x-shaped pulses were used for the excitation of ^{1}H, ^{13}Cβ/γ and ^{13}CO nuclei, respectively. The ^{1}H dimension was acquired with a spectral width of 4810 Hz digitized into 512 complex points. Spectral widths for the ^{13}CO dimension were 1210 Hz (64 complex points), 1810 Hz (128 complex points) or 2716 Hz (256 complex points). In some experiments the non-titrating Asn Cγ and Gln Cδ resonances were folded into the spectrum in order to optimize the resolution of the Glu Cδ resonances. The titration of the CMPcc Arg-34 Cζ resonance was monitored using 1D ^{13}C experiments recorded with 16 384 complex points and spectral widths of 30 303 Hz. The pK_a value of ~40 mM *N*-acetyl-glutamic acid α-methyl ester (Ac-GluOMe: Bachem, Switzerland) was determined from the titration of the molecule's downfield Glu Hγ resonance (2.245 ppm at pH 7.2) in 1D ^{1}H spectra recorded as a function of pH.

Sample pH adjustments were made with 0.5–1.0 μl aliquots of 1 M DCl and NaOD stock solutions. Samples, stock solutions, pH calibration buffers and the pH electrode immersed in KCl, were kept in a 50°C constant temperature bath between NMR experiments. The sample pH was measured before and after each NMR experiment, following equilibration to 50°C for 5–10 min. In the pH range 1.8–7.4, the differences between the two readings were in the order of 0.05 pH U and the pH value was taken as the average of the two measurements. For the two most basic pH points, the differences between the two measurements were 0.2 pH U and the second reading was considered more accurate. All reported pH values are uncorrected for the deuterium isotope effect on pK_a (~0.4 U), which is approximately equal and opposite in sign to the deuterium isotope effect on the pH electrode [20]. Note that all of the pK_a values considered in this work are for D_2O solutions, making it unnecessary to correct for the isotope effect in comparisons between the pK_a values of folded CMPcc and models of the unfolded state of the protein.

^{1}H chemical shifts were referenced to internal DSS. ^{13}C chemical shifts were referenced indirectly assuming a ^{13}C/^{1}H frequency ratio of 0.251449530 [21]. The four parameters pK_a (apparent ionization constant), δ_{low} (low pH chemical shift plateau), δ_{high} (high pH plateau), and *n* (Hill coefficient) were determined from non-linear least squares fits of the chemical shift (δ) data as a function of pH, to the modified Henderson-Hasselbach equation [20,22]:

$$\delta = \delta_{low} - \frac{\delta_{low} - \delta_{high}}{1 + 10^{n(pK_a - pH)}} \quad (1)$$

The midpoints of the transitions were independent of their slopes. Thus the pK_a's obtained from the ^{13}C chemical shift data, differed by less than 0.06 pH U between fits where *n* was treated as a free variable and fits in which the value of *n* was fixed to unity.

3. Results

Fig. 2 shows a 2D H(CA)CO spectrum of CMPcc at pH 5.6. The 2D H(CA)CO experiment correlates backbone and side chain carbonyl and carboxyl carbons with immediately adjacent protons [18,19]. For the present work, the experiment was optimized for the detection of side chain correlations

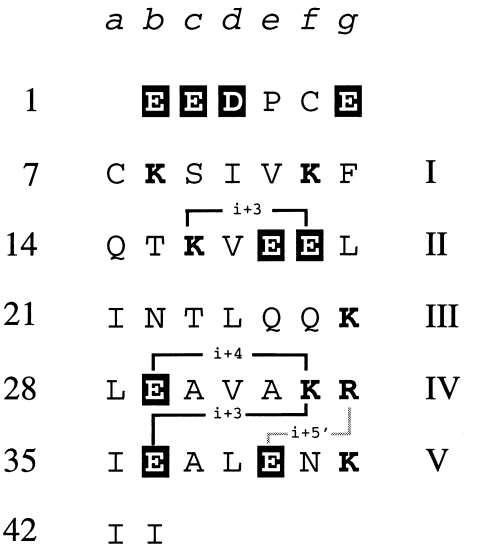

Fig. 1. The sequence of the coiled coil domain of chicken matrilin-1 (residues 451–492). The positions of the five heptad repeats are indicated by roman numerals to the right of the figure. Basic residues are shown in bold, acidic residues in white letters. Possible intrachain ion pairs (based on an i+3 or i+4 sequence spacing) include Lys-16-Glu-19, Glu-29-Lys-33, Lys-33-Glu-36. Arg-34-Glu-39′ could potentially form an i+5′ (g-e′) interchain ion pair.

(Hβ-Cγ of Asp and Asn: Hγ-Cδ of Glu and Gln) by placing the carrier frequency for the ^{13}Cβ/γ nuclei at 38 ppm. As the excitation profile is not completely selective, most of the protein's backbone Hα-C′ correlations are also observed in the spectrum (Fig. 2A), albeit at a lower intensity than when the carrier frequency is placed in the middle of the Cα region (56 ppm). Due to the large chemical shift differences between the Hα and Hβ/Hγ resonances (Fig. 2A), the backbone correlations do not interfere with the side chain correlations. Indeed, the backbone Hα-C′ correlations serve as useful additional probes of the pH titrations of ionizable groups.

The α-helix monomers of CMPcc are magnetically equivalent, being related by a 3-fold symmetry axis running along the length of the parallel homotrimeric coiled coil. Consequently, NMR spectra of CMPcc contain a single set of signals for each amino acid in the protein sequence [13]. The sequence of CMPcc contains 14 residues with side chain carboxyl or carbonyl groups. It was possible to identify all of these groups except for Asp-3 (Hβ1 2.88, Hβ2 2.62 ppm), which occurs in a disordered N-terminal segment of the protein preceding the coiled coil. The Hβ protons of Asn residues and the Hγ protons of Glu and Gln residues were previously assigned based on 3D ^{1}H-^{15}N TOCSY-HSQC, 3D HNHB and 3D HCCH-TOCSY spectra, recorded on CMP samples at pH 5.6–6.0 and a temperature of 50°C [13]. A 2D HCA(CO) spectrum (not shown) enabled us to connect the Asn Hβ resonances and the Gln Hγ resonances with the corresponding, previously assigned Cβ and Cγ resonances. The Cγ resonances of the eight Glu residues, however, span a very narrow range of chemical shifts between 36.6 and 37.9 ppm [13] and were completely unresolved in the HCA(CO) spectrum. Nevertheless, each of the eight Glu residues has at least one unique Hγ chemical shift, which made it possible to link the previously assigned Hγ resonances with the corresponding resolved Cδ resonances in the 2D H(CA)CO spectrum (Fig. 2B).

On lowering the pH from 6 to 2, the Glu Cδ resonances

shift upfield as a group by about 4 ppm, while the Hγ resonances shift downfield by about 0.3 ppm. The pH dependence of the Hγ and Cδ resonances is similar for all Glu residues indicating similar ionization constants. In the 2D H(CA)CO experiment each side chain carboxyl group is defined by two (Hγ,Cδ) and in some cases three (Hγ,Hγ′,Cδ) chemical shifts. Thus in spite of similar ionization constants and chemical shifts, it was possible to completely characterize the pH titrations for seven of the eight Glu residues (Fig. 3Table 1). Glu-39 gives a single weak Hγ-Cδ cross-peak (Fig. 2B), which merges with the downfield Hγ-Cδ cross-peak of Glu-19 at pH values below 4.5. Consequently, we were unable to determine the pK_a of Glu-39 based on the pH titration of its Hγ-Cδ cross-peak. The ionization constant of Glu-39 is of particular interest, as a mutagenesis study of a peptide corresponding to

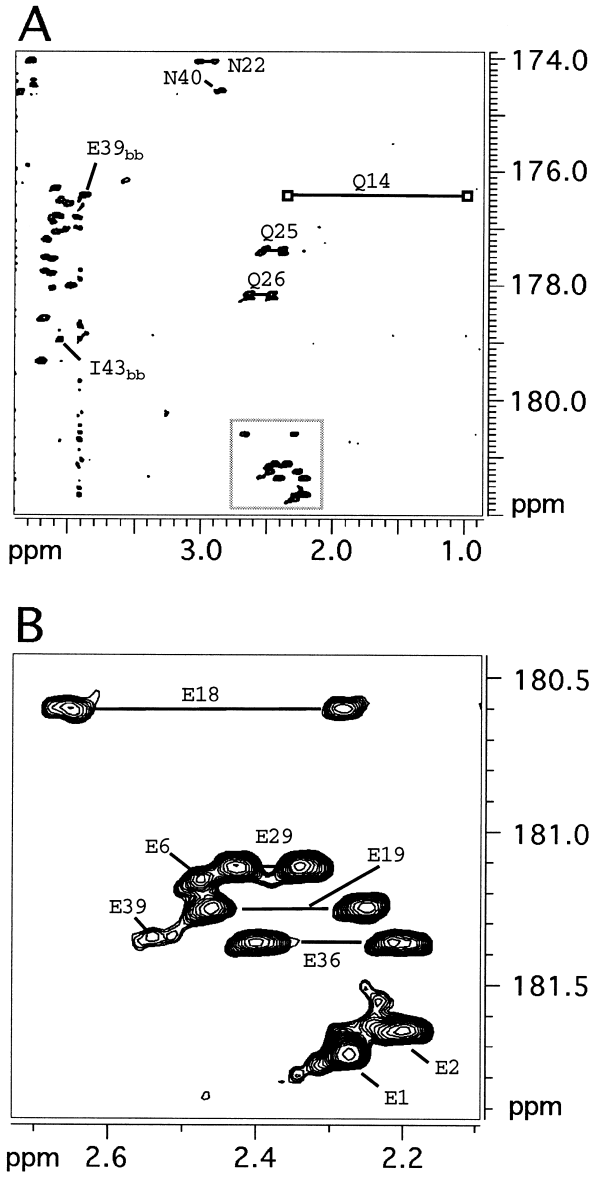

Fig. 2. 2D H(CA)CO spectrum of CMPcc in 150 mM NaCl, at pH 5.6 and 50°C. (A) Portion of the spectrum showing backbone and side chain correlations. The spin system for Gln-14 is only visible at lower contours. (B) Expansion of the region corresponding to the Hγ-Cδ correlations of Glu residues, indicated by the gray box in (A).

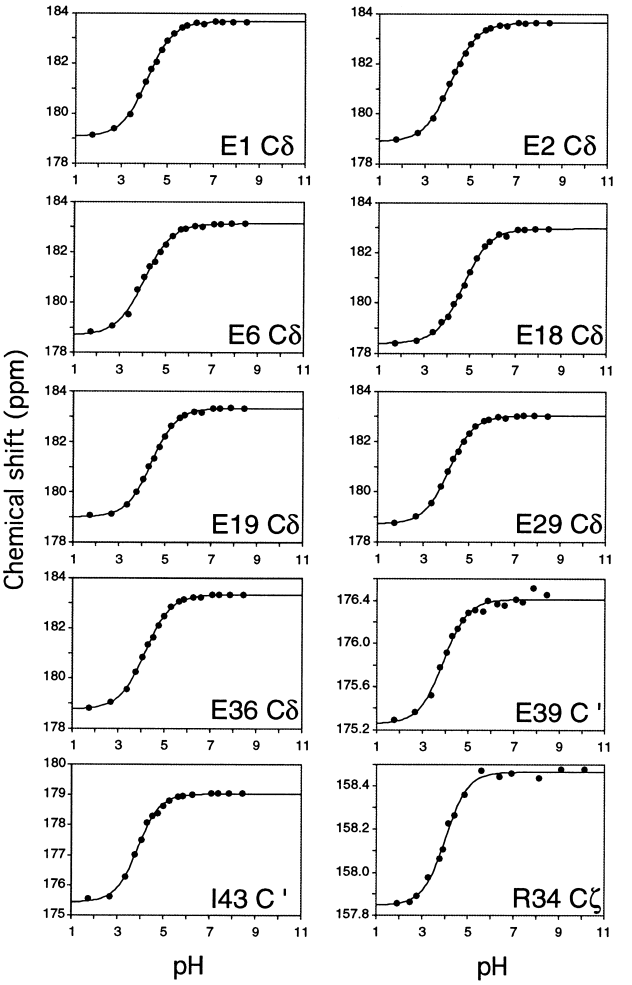

Fig. 3. Titration curves for ^{13}C resonances in CMPcc. The curves represent non-linear least squares fits of the pH titration data to Eq. 1.

residues 8–43 of the human CMPcc sequence (h-CMP-C36) suggests that Arg-34 could be involved in an electrostatic interaction with Glu-39′ in the neighboring chain [6].

We thus examined the possibility of obtaining the pK_a of Glu-39 from the titration of its backbone Hα and C′ chemical shifts. The Hα proton of Glu-39 resonates at 3.85 ppm [13]. The only other residue with a similar Hα proton chemical shift is Pro-4. A single cross-peak with an Hα value of 3.86 is observed in the 2D H(CA)CO spectrum at pH 5.6 (Fig. 2A). The cross-peak shows an upfield ^{13}C shift of 1.2 ppm and a downfield ^{1}H shift of 0.1 ppm with decreasing pH: consistent with an assignment to the Hα-C′ correlation of Glu-39. A least squares fit of the Glu-39 C′ chemical shift as a function of pH gives a pK_a of 3.9. This pK_a value is lower than those of the other seven Glu residues in the protein (Table 1). A pK_a of 3.9, however, could also be consistent with the titration of the C-terminus [23]. Glu-39 is separated by four positions in the sequence from the C-terminal residue Ile-43. The $^{3}J_{HNH\alpha}$ coupling constants of residues Asn-40–Ile-43 are in the range between 5 and 7 Hz, indicative of conformational averaging. Nevertheless, the four C-terminal residues exhibit dNN(i,i+2), dαN(i,i+3) and dαN(i,i+4) NOE's, indicative of a predominantly α-helical conformation [13]. In an α-helix, residues separated by four positions in the se-

Table 1
pH titration parameters for CMPcc in 150 mM NaCl, at a temperature of 50°C[a]

Resonance		pK_a	δ_{low} (ppm)	δ_{high} (ppm)	n
Glu-1	Cδ	4.14 ± 0.02	179.09 ± 0.04	183.67 ± 0.02	0.79 ± 0.02
	Hγ	4.16 ± 0.05	2.52 ± 0.01	2.264 ± 0.003	0.56 ± 0.03
Glu-2	Cδ	4.12 ± 0.02	178.89 ± 0.04	183.65 ± 0.02	0.75 ± 0.02
	Hγ	4.03 ± 0.08	2.50 ± 0.01	2.197 ± 0.005	0.54 ± 0.05
Glu-6	Cδ	4.08 ± 0.04	178.69 ± 0.10	183.14 ± 0.04	0.73 ± 0.04
	Hγ	3.80 ± 0.08	2.71 ± 0.01	2.485 ± 0.003	0.57 ± 0.05
Glu-18	Cδ	4.71 ± 0.02	178.39 ± 0.05	182.97 ± 0.03	0.76 ± 0.03
	Hγ	4.85 ± 0.06	2.855 ± 0.005	2.636 ± 0.004	0.57 ± 0.04
	Hγ′	4.87 ± 0.04	2.58 ± 0.01	2.245 ± 0.005	0.71 ± 0.05
Glu-19	Cδ	4.42 ± 0.02	179.00 ± 0.04	183.33 ± 0.02	0.80 ± 0.02
	Hγ	4.43 ± 0.03	2.78 ± 0.01	2.456 ± 0.003	0.72 ± 0.03
	Hγ′	4.46 ± 0.03	2.60 ± 0.01	2.236 ± 0.003	0.70 ± 0.04
Glu-29	Cδ	4.12 ± 0.01	178.72 ± 0.03	183.03 ± 0.02	0.80 ± 0.02
	Hγ	4.06 ± 0.04	2.671 ± 0.005	2.443 ± 0.002	0.80 ± 0.05
	Hγ′	4.10 ± 0.03	2.636 ± 0.004	2.348 ± 0.002	0.76 ± 0.03
Glu-36	Cδ	4.19 ± 0.02	178.76 ± 0.04	183.33 ± 0.02	0.79 ± 0.02
	Hγ	4.16 ± 0.03	2.57 ± 0.01	2.208 ± 0.003	0.74 ± 0.04
	Hγ′	4.04 ± 0.04	2.630 ± 0.005	2.414 ± 0.002	0.75 ± 0.04
Glu-39	C′	3.91 ± 0.07	175.26 ± 0.04	176.41 ± 0.02	0.82 ± 0.08
	Hα	3.68 ± 0.08	3.948 ± 0.005	3.869 ± 0.002	1.25 ± 0.24
C-term	C′	3.91 ± 0.03	175.44 ± 0.06	179.02 ± 0.03	0.91 ± 0.05
	Hα	3.88 ± 0.02	4.254 ± 0.003	4.007 ± 0.001	0.85 ± 0.03
Arg-34	Cζ[b]	4.02 ± 0.05	157.85 ± 0.02	158.46 ± 0.01	0.89 ± 0.09

[a]The titration parameters summarized in this table provide NMR assignment information for the respective resonances over the entire pH range investigated in this work. The pK_a values of Glu residues, however, are more accurately reflected by the ^{13}C data because the magnitudes of pH-dependent chemical shift changes are greater for ^{13}C than for ^{1}H resonances.
[b]From 1D ^{13}C spectra recorded as a function of pH.

quence are in close spatial proximity. Indeed in the NMR structures of CMPcc the average distance between the C′ carbons of Glu-39 and Ile-43 is 5.7 Å, a value comparable to the average distance between the backbone C′ and the side chain Cδ carbons of Glu-39 ($<d>$ = 4.9 Å). The C-terminal residue Ile-43 has an Hα chemical shift of 4.05. While there are three other residues with similar Hα chemical shifts in the protein, the Hα-C′ cross-peak of Ile-43 can be easily distinguished based on its downfield C′ chemical shift and based on the magnitude of its chemical shift changes as a function of pH (Δδ C′ = 3.6 ppm, Δδ Hα = 0.25 ppm). Ile-43 titrates with a pK_a of 3.9, a value identical to that obtained for the C′ resonance of Glu-39. Consequently we could not distinguish if the titration of the Glu-39 C′-Hα cross-peak is a function of the ionization of the Glu-39 side chain carboxyl, the C-terminus or both groups.

Table 2
pK_a and standard $\Delta\Delta G_{titr}$ values for glutamic acid residues in CMPcc, at a temperature of 50°C[a]

Residue	pK_a	$\Delta\Delta G_{titr}$ (kcal/mol)[b]
Glu-1	4.14	0.0
Glu-2	4.12	0.0
Glu-6	4.08	−0.1
Glu-18	4.71	0.8
Glu-19	4.42	0.4
Glu-29	4.12	0.0
Glu-36	4.19	0.1
Glu-39[c]	4.02	−0.2

[a]Experimental uncertainties in the reported pK_a and $\Delta\Delta G_{titr}$ values are estimated to be ± 0.05 pH U and ± 0.1 kcal/mol, respectively.
[b]Contributions from the ionization of acidic residues to the change in free energy of folding per CMPcc monomer, calculated assuming a value of 4.15 ± 0.05 for the pK_a values of glutamates in unfolded CMPcc. Alternatively, if the pK_a values of glutamates in unfolded CMPcc are modeled according to the pK_a of AcGluOMe (4.33 ± 0.01, at 50°C in a D_2O solution containing 150 mM NaCl), the $\Delta\Delta G_{titr}$ values are shifted by −0.3 kcal/mol.
[c]From the apparent pK_a of the Arg-34 Cζ resonance. All other pK_a values are from the titration curves of the respective Glu Cδ resonances (Table 1).

Assuming that Arg-34 is involved in an electrostatic interaction with Glu-39′, a change in the ionization state of Glu-39 would be expected to induce perturbations in the chemical shifts of Arg-34 [24]. The Cζ carbon in the guanidinium group of Arg-34 should be particularly sensitive to pH-dependent changes in an interaction with the carboxyl group of Glu-39′. Arginine guanidinium carbons (Cζ carbons) resonate with highly unique chemical shifts of about 160 ppm [25]. The closest ^{13}C resonances in proteins are Tyr C4 ring carbons (~157 ppm), however, CMPcc contains no tyrosines. The next closest ^{13}C resonances are backbone carbonyls (starting ~10 ppm downfield) and carbons in aromatic rings (~20 ppm upfield). The sequence of CMPcc contains only one arginine residue (Fig. 1). A single resonance is observed in the region of the ^{13}C spectrum of CMPcc between 142 and 168 ppm (Fig. 4). The resonance is a sharp singlet in ^{13}C spectra acquired without ^{1}H decoupling (not shown), consistent with a quarternary carbon. Based on its distinct chemical shift we assign this resonance to the guanidinium (Cζ) carbon of Arg-34. The resonance titrates over a range of 0.6 ppm, with an apparent pK_a of 4.0 (Fig. 3). This value is much lower than the typical pK_a of ~12 for the ionization of arginine [23] and is within the experimental error of that obtained for the Glu-39 C′ resonance (Table 1). The Cζ atom of Arg-34 has an average distance of 6.0 Å to the Cδ atom of Glu-39′ from a neighboring chain in the trimer. By contrast, an interaction between Arg-34 and the C-terminus is extremely unlikely. The closest intra or interchain distance between the Ile-43 C′ and Arg-34 Cζ atoms is 13.7 Å. Indeed, the chemical shift of the Cγ carbon of Asn-40, which is closer to the C-terminus ($<d>$ = 8.7 Å), is invariant to changes in pH. Based on the titration of the Arg-34 Cζ resonance with an apparent

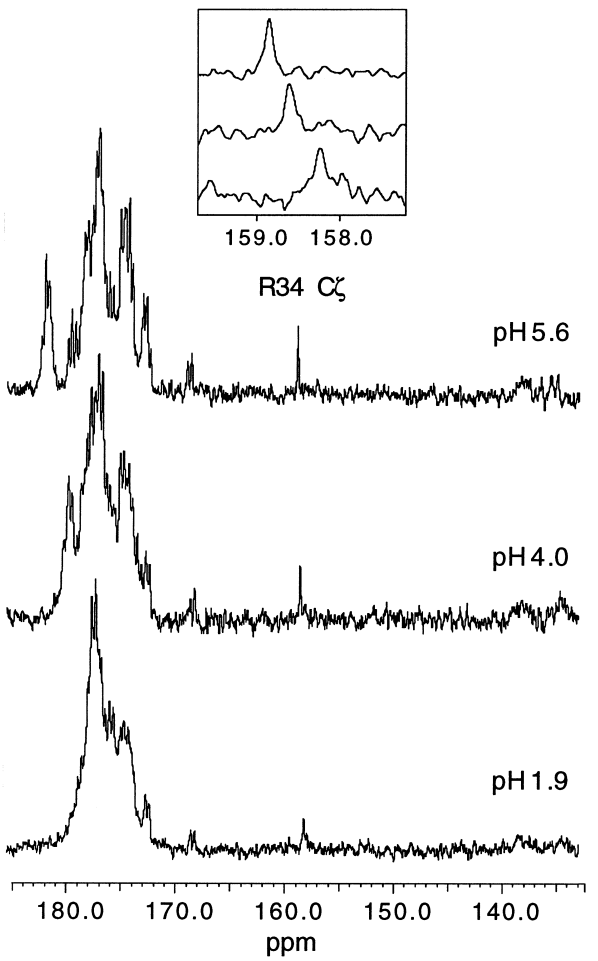

Fig. 4. Selected 1D ^{13}C spectra of CMPcc showing the titration of the Arg-34 Cζ resonance (expansion in inset). The group of resonances that titrates from ~183 ppm at high pH to ~178 ppm at low pH are the Cδ carboxyl carbons of Glu residues. The non-titrating resonances between ~170 and ~180 ppm are from backbone and side chain carbonyls.

pK_a of 4.0 (Fig. 4), we conclude that Arg-34 interacts weakly with Glu-39′ in the neighboring chain and that the apparent pK_a of the Arg-34 Cζ resonance reflects the ionization constant for the side chain carboxyl group of Glu-39′.

4. Discussion

The energetic contributions to protein stability from the titration of individual acidic groups can be expressed as [9,11]:

$$\Delta\Delta G_{titr} = -2.303RT*(pK_{a,U} - pK_{a,F}) \quad (2)$$

where p$K_{a,F}$ and p$K_{a,U}$ are the ionization constants in the folded and unfolded forms of the protein, respectively [9,11]. The Glu moiety in the blocked model compound AcGluOMe titrates with a pK_a of 4.33 under the solution conditions used to study CMPcc. This value is about 0.2 pH U higher than the pK_a values of Glu residues in CMPcc which are not expected to be involved in electrostatic interactions: for example, Glu-1 and Glu-2 which are in the unstructured [13,14] N-terminal segment of the protein. It has been suggested that the pK_a values for acidic residues in unfolded proteins are slightly lower than those in isolated blocked amino acids or short peptides [26], reflecting differences between the dielectric constants of an unfolded polypeptide and water. Thus the pK_a values for Glu-20 and Glu-22 in the unfolded form of GCN4-p1 at 25°C and 150 mM NaCl are 4.20 ± 0.05 and 4.13 ± 0.05, respectively [9]. These values closely match those obtained in the present study for Glu-1 and Glu-2, two residues which are in an unstructured segment of CMPcc. We thus conclude that a pK_a value of ~4.15 ± 0.05 is more akin to that of a Glu residue in an unfolded protein, than the value of 4.33 obtained from the AcGluOMe model compound. $\Delta\Delta G_{titr}$ values for CMPcc, calculated from Eq. 2 with the assumption of an unfolded state p$K_{a,U}$ of 4.15 ± 0.05, are listed in Table 2. The n values which pertain to the cooperativities of the ionization reactions are summarized in Table 1. NMR pH titration curves almost invariably give $n = 1$ [20]. An n value lower than unity implies negative cooperativity [20,23]. The n values obtained from the ^{13}C data for CMPcc are usually larger than those obtained from the ^{1}H data, even when the values pertain to the same residue (e.g. Glu-1 in Table 1). The only clear distinction between the ^{1}H and ^{13}C data is that the ^{1}H resonances span a smaller effective range of chemical shifts as a function of pH. In this regard, it is worth noting that the n value obtained for AcGluOMe is also lower than unity (0.92 ± 0.02), even though the molecule contains a single ionizable group. These observations suggest that n values determined from NMR are subject to systematic errors which are not reflected in the precision of the fits of the pH titration data to Eq. 1. Nevertheless, the n values in Table 1 are close to unity and indicate that the ionization of charged groups related by the 3-fold symmetry of CMPcc is non-cooperative, or slightly anticooperative. As such, the $\Delta\Delta G_{titr}$ values in Table 2 pertain to differences in free energies per mole of CMPcc monomers. To obtain the corresponding values expressed per mole of CMPcc trimer, the $\Delta\Delta G_{titr}$ values in Table 2 should be multiplied by a factor of three.

Results on the roles of electrostatic interactions in coiled coils do not easily lend themselves to generalization [6,9–11,27–29]. The strengths of charge-charge interactions depend on the effective dielectric constants [23]. Thus salt screening of electrostatic charges, specific ion binding and the degree of solvent exposure of charged residues, can have complex effects on electrostatic interactions [27,28]. With the exception of the present work, NMR determination of site-specific pK_a values in coiled coils has been limited to residues Glu-20 and Glu-22 of the GCN4-p1 dimer [9]. Although both glutamates form interhelical salt bridges in the X-ray structure of GCN4-p1, neither makes a favorable contribution to the stability of the protein compared to the corresponding neutral glutamic acid forms [9]. By contrast, mutagenesis and protein design experiments have often suggested key roles for electrostatic interactions [6,10,28,29]. The energetic consequences of the addition or deletion of charged residues, however, can extend beyond changes in electrostatic interactions [9,16,28]. A peptide corresponding to residues 8–43 of human CMPcc (h-CMP-C36) illustrates this point. Substitution of the residue corresponding to Glu-39 in CMPcc to a glutamine, results in a h-CMP-C36 peptide that forms a tetramer at neutral pH but reverts to a trimeric state below pH 3.0 or above pH 11.5 [6]. By contrast, the wild-type h-CMP-C36 peptide maintains a trimeric oligomerization state throughout the entire pH range studied [6]. The latter observation suggests that new electrostatic interac-

tions are formed in the tetramer [6], which are absent in the trimer.

The growing number of coiled coil structures, together with studies on model peptides, provide a data base for the types of electrostatic interactions that can occur in coiled coils. At present, however, it would seem extremely difficult to predict electrostatic interactions from the amino acid sequence. Thus potential ion pairs are not always formed in coiled coil structures [11] and even when detected in high resolution X-ray structures, their presence is not necessarily concomitant with a favorable contribution to the stability [9]. In CMPcc, Glu-18 and Glu-19 have raised pK_a values implying that the charged forms of these residues destabilize the native state by 0.8 ± 0.1 and 0.4 ± 0.1 kcal/mol, respectively. The raised pK_a values may be due to the high density of six negative charges that results when the two Glu residues adjacent in the sequence, are brought into spatial proximity by the parallel coiled coil trimer structure. Consistent with some previous studies on coiled coils [9,11], the most pronounced electrostatic contributions to the stability of CMPcc are unfavorable. Of the four possible ion pairs predicted from the amino acid sequence of CMPcc (Fig. 1), the only one supported by the pK_a data is between Arg-34 and Glu-39′ on a neighboring chain. The ionization of Glu-39 makes a modest favorable contribution to the free energy of folding of -0.2 ± 0.1 kcal/mol of CMPcc monomer, or -0.6 ± 0.3 kcal/mol of CMPcc trimer. By the way of comparison, the stability of CMPcc to unfolding at pH 6.2, where all of the acidic groups in the protein are charged, is estimated to be -7.0 kcal/mol of CMPcc trimer [13]. The pK_a values of Glu-29 and Glu-36 are unperturbed, in spite of the potential for either residue to form an intrachain ion pair with Lys-33. The pK_a of Glu-19 is raised, presumably because repulsive interactions predominate over a potentially attractive interchain electrostatic interaction with Lys-16.

To summarize, the present work suggests that with the exception of the amelioration of unfavorable electrostatic terms for Glu-18 and Glu-19, neutralization of negative charges accounts for only small effects on the stability of the CMPcc coiled coil. The pK_a values for acidic residues and hence the contributions of negative charges to the stability of the coiled coil, are not easily anticipated from the protein's amino acid sequence. An assessment of the role of charge interactions in determining the oligomerization state specificity presents an even more formidable challenge, as the electrostatic properties of two distinct conformations (e.g. dimer, trimer) need to be considered.

Acknowledgements: This work was supported by grants from the Swiss National Science Foundation to J.E. and A.T.A.

References

[1] Crick, F.H.C. (1953) Acta Cryst. 6, 689–697.
[2] Cohen, C. and Parry, D.A.D. (1990) Proteins 7, 1–15.
[3] Lupas, A. (1996) Trends Biochem. Sci. 21, 375–382.
[4] Hodges, R.S., Sodek, J., Smillie, L.B. and Jurasek, L. (1972) Quant. Biol. 37, 299–310.
[5] McLachlan, A.D. and Stewart, M. (1975) J. Mol. Biol. 98, 293–304.
[6] Beck, K., Gambee, J.E., Kamawal, A. and Bächinger, H.P. (1997) EMBO J. 16, 3767–3777.
[7] Marqusee, S. and Baldwin, R.L. (1987) Proc. Natl. Acad. Sci. USA 84, 8898–8902.
[8] Graddis, T., Myszka, D. and Chaiken, I. (1993) Biochemistry 32, 12664–12671.
[9] Lumb, K.J. and Kim, P.S. (1995) Science 268, 436–439.
[10] Lavigne, P., Sönnichsen, F.D., Kay, C.M. and Hodges, R.S. (1996) Science 271, 1136–1137.
[11] Lumb, K.J. and Kim, P.S. (1996) Science 271, 1137–1138.
[12] Haudenschild, D.R., Tondravi, M.M., Hofer, U., Chen, Q. and Goetinck, P.F. (1995) J. Biol. Chem. 270, 23150–23154.
[13] Wiltscheck, R., Kammerer, R.A., Dames, S.A., Schulthess, T., Blommers, M.J.J., Engel, J. and Alexandrescu, A.T. (1997) Protein Sci. 6, 1734–1745.
[14] Dames, S.A., Kammerer, R.A., Wiltscheck, R., Engel, J. and Alexandrescu, A.T. (1998) Nat. Struct. Biol. 5, 687–691.
[15] Barlow, D.J. and Thornton, J.M. (1983) J. Mol. Biol. 168, 867–885.
[16] Yang, A.-S. and Honig, B. (1992) Curr. Opin. Struct. Biol. 2, 40–45.
[17] Tanford, C. (1961). Physical Chemistry of Macromolecules, Wiley, New York.
[18] Grzesiek, S. and Bax, A. (1993) J. Magn. Reson. Ser. B 102, 103–106.
[19] Yamazaki, T., Nicholson, L.K., Torchia, D.A., Wingfield, P., Stahl, S.J., Kaufman, J.D., Eyermann, C.J., Hodge, C.K., Lam, P.Y.S., Ru, Y., Jadhav, P.K., Chang, C.-H. and Weber, P.C. (1994) J. Am. Chem. Soc. 116, 10791–10792.
[20] Markley, J.L. (1975) Accounts Chem. Res. 8, 70–80.
[21] Wishart, D.S., Bigam, C.G., Yao, J., Abilgaard, F., Dyson, H.J., Oldfield, E. and Markley, J.L. (1995) J. Biomol. NMR 6, 135–140.
[22] Jeng, M.-F. and Dyson, H.J. (1996) Biochemistry 35, 1–6.
[23] Creighton, E.E. (1993). Proteins: Structures and Molecular Principles, 2nd edn., W.H. Freeman and Company, New York.
[24] Alexandrescu, A.T., Evans, P.A., Pitkeathly, M. and Dobson, C.M. (1993) Biochemistry 32, 1707–1718.
[25] Wishart, D.S., Bigam, C.G., Holm, A., Hodges, R.S. and Sykes, B.D. (1995) J. Biomol. NMR 5, 67–81.
[26] Tan, Y.-J., Oliveberg, M., Davis, B. and Fersht, A.R. (1995) J. Mol. Biol. 254, 980–992.
[27] Kohn, W.D., Kay, C.M. and Hodges, R.S. (1997) J. Mol. Biol. 267, 1039–1052.
[28] Eckert, D.M., Malashkevich, V.N. and Kim, P.S. (1998) J. Mol. Biol. 284, 859–865.
[29] Zhou, N.E., Kay, C.M. and Hodges, R.S. (1994) J. Mol. Biol. 237, 500–512.