Research Article

Journal of PeptideScience

Received: 31 January 2011 Revised: 18 March 2011 Accepted: 24 March 2011 Published online in Wiley Online Library: 6 June 2011

(wileyonlinelibrary.com) DOI 10.1002/psc.1377

Cooperativity network of Trp-cage miniproteins: probing salt-bridges

Petra Rovó,[a] Viktor Farkas,[b] Orsolya Hegyi,[c] Orsolya Szolomájer-Csikós,[c] Gábor K. Tóth[c] and András Perczel[a,b]*

Trp-cage miniprotein was used to investigate the role of a salt-bridge (Asp^{9} – Arg^{16}) in protein formation, by mutating residues at both sides, we mapped its contribution to overall stability and its role in folding mechanism. We found that both of the above side-chains are also part of a dense interaction network composed of electrostatic, H-bonding, hydrophobic, etc. components. To elucidate the fold stabilizing effects, we compared and contrasted electronic circular dichroism and NMR data of miniproteins equipped with a salt-bridge with those of the salt-bridge deleted mutants. Data were acquired both in neutral and in acidic aqueous solutions to decipher the pH dependency of both fully and partially charged partners. Our results indicate that the folding of Trp-cage miniproteins is more complex than a simple two-state process as we detected an intermediate state that differs significantly from the native fold. The intermediate formation is related to the salt-bridge stabilization; in the miniprotein variants equipped with salt-bridge the population of the intermediate state at acidic pH is significantly higher than it is for the salt-bridge deleted mutants. In this molecular framework Arg^{16} stabilizes more than Asp^{9} does, because of its higher degree of 3D-fold cooperation. In conclusion, the Xxx^{9} ↔Yyy^{16} salt-bridge is not an isolated entity of this fold; rather it is an integrated part of a complex interaction network.

Supporting information may be found in the online version of this article

Keywords: Trp-cage; salt-bridge; H-bonding network; NMR structure

Introduction

The thermodynamic stability of proteins has long been of interest to biochemists [1]. As native protein folds are only 5–10 kcal/mol more stable than their denaturated states [2], no type of intermolecular force can be neglected when accounting for folding. There is considerable evidence that hydrophobic interactions among non-polar residues dominate the protein folding [3–5] and *van der Waals* [6] and H-bonding interactions [7,8] have a lower contribution to the overall thermodynamic stability. The role of electrostatic interactions (such as salt-bridges and dipoles) is more ambiguous as electrostatic effects are highly variable, sometimes favorable but sometimes not [7]. However, salt-bridges are important as they contribute to protein folding, structure, residual flexibility and function. (i) Salt-bridges may increase the specificity of folding by reducing the number of ways of molecular packing [9]. (ii) Solvent exposed charges help to solubilize proteins in aqueous environments [10]. (iii) Charged residues could create a suitable local electrostatic environment for specific protein–protein and protein–ligand interactions [11–15].

In spite of their high importance the direct experimental detection of salt-bridge formation is a challenge. While there are several methods for measuring H-bonding interactions directly – such as H/D exchange experiments [16–19] and heteronuclear NOE [20,21] by NMR, or by measuring IR C═O shifts [22–24] – the presence of a salt-bridge can only be noticed indirectly by point-mutations (mutational approach), or by changing the pH or ionic strength of the medium (pK_a approach) [25]. Applying the above concepts on salt-bridges, model peptides may shed some light on the mechanistic details of electrostatic interactions. An understanding of these interactions is required as the biological relevance of protein–protein and protein–ligand recognition is becoming more apparent.

The Trp-cage miniprotein, TC5b (NLYIQ WLKDG GPSSG RPPPS), is only a 20-residue long polypeptide, notable for its protein-like 3D-fold in quasi physiological conditions [26]. In its 3D structure an α-helix, a 3_{10}-helix and a polyproline II helix shield the central hydrophobic Trp^{6} residue. Unlike most small protein inhibitors and toxins, the Trp-cage fold lacks disulfide bridges. In their absence this structure is stabilized through a series of hydrophobic interactions, all organized around Trp^{6}, completed by a salt-bridge formed between residues 9 and 16 (e.g. Asp^{9} ↔Arg^{16}). The stable and well-defined time average structure of the Trp-cage scaffold is the consequence of a series of truncations and mutations carried out on the parent protein exendin-4 (Ex-4) [27]. One of the key structure stabilizing modifications was the salt-bridge insertion

* *Correspondence to: András Perczel, Laboratory of Structural Chemistry and Biology, Institute of Chemistry, Eötvös University, H-1117 Budapest, Pázmány Péter stny 1/A, Hungary. E-mail: perczel@chem.elte.hu*

a *Laboratory of Structural Chemistry and Biology, Institute of Chemistry, Eötvös University, H-1117 Budapest, Pázmány Péter stny 1/A, Hungary*

b *Protein Modeling Group HAS-ELTE, Institute of Chemistry, Eötvös University, H-1538 Budapest, P.O. Box 32, Hungary*

c *Department of Medical Chemistry, Faculty of General Medicine, University of Szeged, H-6720 Szeged, Dóm tér 8., Hungary*

Journal of PeptideScience

Table 1. Sequence and chemical characterization of the examined TC5b variants

Name of TC5b variants	Sequence 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	M_w/Da Calc.	Meas.	HPLC t_{ret}/min
D9E	N	L	Y	I	Q	W	L	K	**E**	G	G	P	S	S	G	**R**	P	P	P	S	2183.47	2183.5	8.56
D9Aad_R16K	N	L	Y	I	Q	W	L	K	**Aad**[a]	G	G	P	S	S	G	**K**	P	P	P	S	2169.46	2171.0	7.61
R16hR	N	L	Y	I	Q	W	L	K	**D**	G	G	P	S	S	G	**hR**[a]	P	P	P	S	2183.44	2184.5	8.59
D9N	N	L	Y	I	Q	W	L	K	**N**	G	G	P	S	S	G	**R**	P	P	P	S	2168.46	2169.6	7.17
R16A	N	L	Y	I	Q	W	L	K	**D**	G	G	P	S	S	G	**A**	P	P	P	S	2084.31	2084.8	6.45
D9N_R16A	N	L	Y	I	Q	W	L	K	**N**	G	G	P	S	S	G	**A**	P	P	P	S	2083.33	2083.4	7.07
D9S	N	L	Y	I	Q	W	L	K	**S**	G	G	P	S	S	G	**R**	P	P	P	S	2141.41	2141.4	8.50

[a] Aad stands for L-aminoadipic acid; hR stands for L-homoarginine.

by Neidigh *et al*. [26], where the N9D and A16R mutations made the structure stable enough to be folded as a single time average structure more than 95% in water at 280 K.

The detailed structural analysis of the model miniprotein TC5b and its derivatives could provide valuable insights into the folding mechanism of proteins [28]. For this reason, the investigation of the structure stabilizing interactions such as the hydrophobic effect between Trp^6 and Pro^{12}, Pro^{18}; the $\pi-\pi$ stacking between Tyr^3 and Trp^6; and the salt-bridge between Asp^9 and Arg^{16} is essential [29]. The latter salt-bridge was of high interest in recent computational [30–33] and experimental studies [28,31,34,35].

In this work we present selected mutants of the $Xxx^9 \leftrightarrow Yyy^{16}$ salt-bridge, all within the TC5b scaffold to decipher the true nature of the salt-bridges in a molecular-context dependent manner. The sequences designed, synthesized and studied are presented in Table 1 and a schematic overview of the salt-bridge enforced TC5b scaffold is depicted in Figure 1. The two salt-bridge forming residues are located in two non-adjacent secondary structural elements: Asp^9 is the C-terminal residue of the α-helix, while Arg^{16} is the first residue of PPII. Furthermore, this salt-bridge frames and stabilizes the central 3_{10}-helix by incorporating Pro^{12}, the key residue for preventing water penetration into the core part of this foldamer [32].

To monitor the structure stabilizing factor of the above salt-bridge and its role in the folding mechanism, TC5b and several of its variants were studied at two different pHs (~3.0 and ~7.0), over a wide range of temperatures relevant for proteins ($5 \leq T \leq 85\ ^\circ C$). Secondary structural changes recorded by electronic circular dichroism (ECD) data were analyzed in conjunction and quantified by the deconvolution program convex constraint analysis plus (CCA+) [36–38]. To locate structural differences at an atomic level in the fully folded states, high resolution $^1H-^1H$ NMR studies were performed ($T = 280$ K) both at neutral ($6.5 \leq pH \leq 7.1$) and acidic ($2.8 \leq pH \leq 3.2$) pH conditions. Using the above models and tools, our aim was to evaluate the role of salt-bridge(s) on protein folding and stabilization in a molecular-context dependent manner.

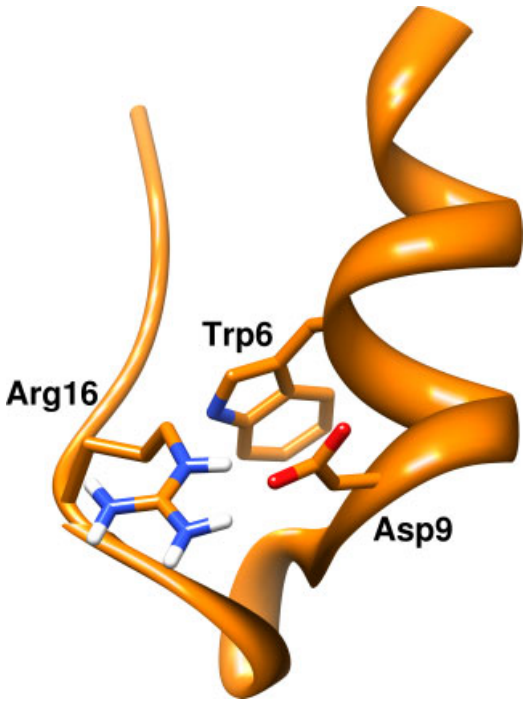

Figure 1. The Asp^9-Arg^{16} salt-bridge of TC5b also holding together the three secondary structural elements, the α-helix (L^2-D^9), the 3^{10}-helix (P^{12}-S^{14}) and the PPII-helix (R^{16}-P^{19}) jointly forming a protein-like 3D-scaffold. This figure is available in colour online at wileyonlinelibrary.com/journal/jpepsci.

Materials and Methods

Peptide Synthesis, Purification and Analysis Peptides

Peptides D9E, D9S, R16A and D9N_R16A were synthesized using the standard Boc solid-phase peptide synthesis method manually. Merrifield resins were used as the solid support and the side-chain protecting groups used were as follows: Boc-Ser(Bzl), Boc-Arg(Tos), Boc-Asp(OcHex), Boc-Glu(OcHex), Boc-Tyr(2BrZ), and Boc-Lys(2ClZ). Couplings were performed with DCC and HOBt. Amino acid incorporation was monitored by the quantitative ninhydrin test. The completed peptide resins were treated with liquid HF/dimethyl sulphide/p-cresol/p-thiocresol (86 : 6 : 4 : 2, vol/vol) at 0 °C for 45 min. HF was removed and the resulting free peptides were solubilized in 10% aqueous acetic acid, filtered and lyophilized.

Peptides D9N, D9Aad_R16K and R16hR were synthesized by Fmoc chemistry using Wang resins as the solid support. The following side-chain protections were used: Fmoc-Tyr(But), Fmoc-Lys(Boc), Fmoc-Ser(But), Fmoc-Aad(OBut) and Fmoc-hArg(Pmc). Couplings were performed with DCC and HOBt. The completed peptide resins were treated with TFA/DTT/phenol/H_2O (87 : 3 : 5 : 5) and the resulting free peptides were solubilized in 10% aqueous acetic acid, filtered and lyophilized. The crude peptides were puri-

fied by reverse-phase HPLC on a Phenomenex Jupiter C-18 column (21.2 mm × 250 mm), using a gradient of water/acetonitrile (solvent system was the following: 0.1% TFA in water and 0.1% TFA, 80% acetonitrile in water, the gradient was: 0 → 40% B in 80 min, flow 3 ml/min and detection completed at 220 nm). Collected fractions were lyophilized and their identity confirmed using a Finnigan TSQ 7000 tandem quadrupole mass spectrometer equipped with an electrospray ion source. The analytical data of the peptides are summarized in Table 1.

ECD Data and Ensemble Deconvolution by CCA+

ECD spectra were recorded on a Jasco J810 dichrograph in cells with a path-length of 1.0 and 10 mm. Typical spectral accumulation parameters were: a scan rate of 50 nm/min with a 1 nm bandwidth, and a 0.2 nm step resolution over the wavelength range 185–260 nm with four scans averaged for each spectrum at temperatures ranging from 5 to 85 °C. The temperature at the cell was controlled by a Peltier-type heating system. The spectrophotometer was equilibrated at each temperature for 5 min before acquisition. Solvent reference spectra were used as baselines which were automatically subtracted from the peptide spectra. The raw ellipticity data were converted into mean residue molar ellipticity units ($[\Theta]_{MR}$, deg × cm^2/dmol). All spectra were analyzed using the latest release of the CCA+ deconvolution program [36–38].

NMR Data Acquisition Procession and Structure Determination

All NMR experiments were performed with a Bruker DRX 500 MHz spectrometer on peptide samples of ~1 to 2 mM concentration at 6.5 < pH < 7.1 and 2.8 < pH < 3.5, with 5–8% D_2O. 4,4-dimethyl-4-silapentane-1-sulphonic acid (DSS) was used as the internal proton reference standard set to 0 ppm for all conditions. Proton NMR chemical shift assignments were achieved using standard procedures which required the recording of ^1H–^1H TOCSY and ^1H–^1H NOESY spectra and were completed using the program Sparky [39]. TOCSY measurements were taken with 65 ms < spinlock time < 80 ms, and 100 ms < mixing time < 300 ms was used for the NOESY spectra. The CSDs for all peptides under neutral conditions were calculated using the reference random coil shifts determined by Bundi and Wuthrich [40] and modified for the Trp-cage by Fesinmeyer *et al*. [41] and for the acidic spectra the δ_{rc} defined by Schwarzinger *et al*. [42] were used. The NOE intensities were converted into distance constraints and the size category, and allowed ranges were adjusted manually during the refinement procedure. Assignments are listed in Supporting information. All structures were calculated with the program CNS Solve 1.1 [43]. The 10 best structures were selected out of 50 and their qualities checked using the validation server of Protein Data Bank [44]. All structures reported here were prepared using Chimera [45].

Results and Discussion

Secondary Structure Content Analysis by ECD Spectroscopy

Far and near UV ECD spectral analysis was completed by recording the ECD spectra of the eight miniprotein variants analyzed here, both at neutral and acidic pH and over a temperature range of 80 °C (Figure 2A). In total, 272 ECD (= 2 pH × 8 mutants × 17 temperature) spectra were recorded and analyzed together using the latest version of CCA+ algorithm [36–38]. The spectral deconvolution resulted in three pure ECD components as depicted in Figure 2B. The first component is similar to a U-type ECD spectrum but less intense (labeled as U′ pure ECD curve) and was assigned to a molten globule or intermediate state (I). The second pure ECD component curve with an $n \rightarrow \pi^*$ transition (~222 nm) and an exciton couplet (194 and 209 nm) resembles that of an α- and 3_{10}-helical component (C-type ECD curve) and represents the folded form (F) of the molecular scaffold. The third pure ECD curve of type U can be associated with the unfolded or atypical ensemble of backbone structures (U). For all mutants, the increase in temperature causes the decrease of the α- and/or 3_{10}-helical component and the increase of the unordered component, while the U′ component has a pH dependent maximum at ambient temperatures [Figure 2C; for details see Figure S2 (Supporting information)]. The decrease in the helical content is likely to be associated with a shift in the Trp-cage equilibrium toward an unordered state, rather than the shortening of the α-helical segment with respect to the native conformation.

At acidic pH the relative ratio of the U′ ECD component increases significantly and reaches a maximum at around 45–55 °C. Thus, the molecular scaffold of the TC5b variants favors an alternative, mostly unstructured intermediate (I) state over the Trp-cage fold (F) at higher temperatures. This could imply that under acidic pH the unfolding of these miniproteins is not a simple two-state unfolding mechanism as suggested by Neidigh *et al*. [26] and Streicher *et al*. [46,47] but rather it is more complex process as proposed by Mok *et al*. [48] and Ahmed *et al* [49]. Our data show that the α-helical segment (and the 3_{10}-helix) melts first into a turn-rich molten globule conformation and this intermediate state denaturates further at higher temperatures. The assumed folding route resembles a nucleation–condensation (NC) mechanism in which the protein first forms the native state tertiary contacts, while the secondary structure is still solvated. This NC route has been suggested by Zhou *et al*. and Juraszek *et al*. based on replica-exchange MD simulations of Trp-cage in explicit solvent [31,32,50]. The formation of the intermediate state can be detected as well at neutral pH, although its contribution is much lower (for details see Figure S2).

Interestingly, there is no significant difference between the melting behavior of the mutants with or without a salt-bridge as monitored by ECD, which reflects a globally similar temperature (and pH) induced unfolding process of the TC5b scaffold. Only the R16A mutants (R16A and R16A_D9N) show higher U′-type contribution (more intermediate fold) even at neutral pH with respect to the other variants. The removal of the R^{16} side-chain prohibits the salt-bridge formation as well as the efficient burial of the central indole ring, thus the formation of a molten globule state is preferable even at low (5 °C) temperatures.

The ECD melts were subjected to thermodynamic studies based on the monitoring of the unfolding process at 222 nm wavelength (Figure S3). The observed ECD melting curves are much broader than usual protein-melting curves, implying marginal cooperativity during the folding process. The CCA+ deconvolution and the latter NMR studies also confirmed that unfolding of the TC5b scaffold is not a pure two-state process. Thus the ECD melting curves should fit with (at least) a three-state model where besides the folded (F) and unfolded (U) states, an intermediate state (I) populates the thermal unfolding; i.e. two different sets of enthalpy of transition (ΔH_U), heat capacity change (ΔC_p) and melting temperature (T_m) should be derived, which would be six variables fitted onto 17 measured data points.

Journal of PeptideScience

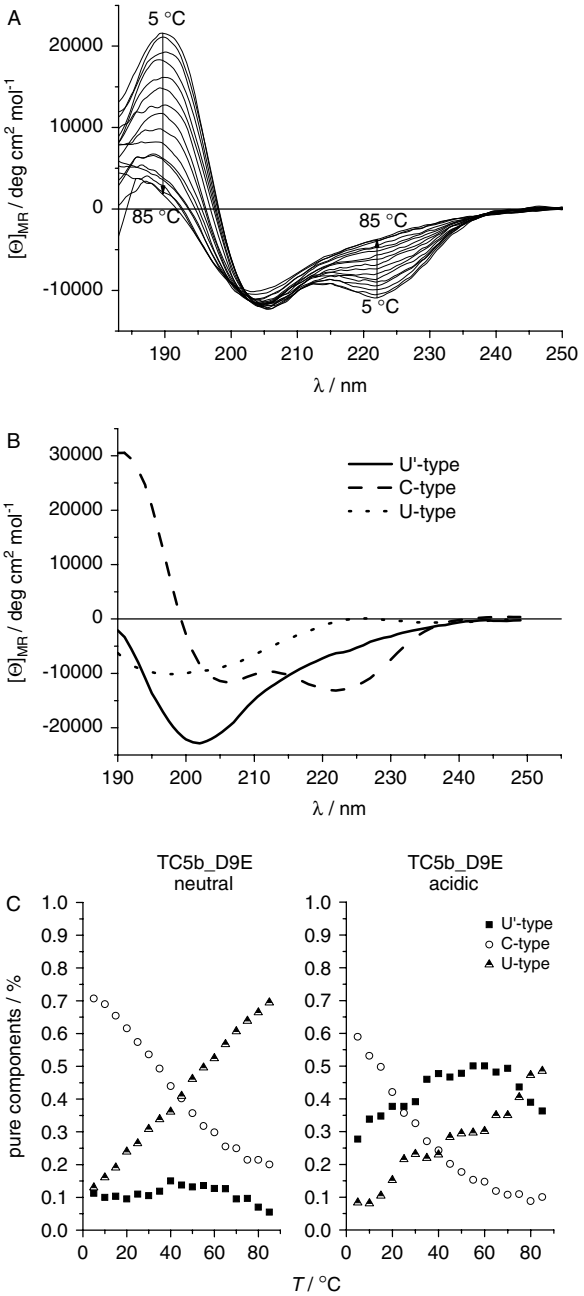

Figure 2. (A) Far UV ECD spectra of TC5b_D9E from 5 to 85 °C at 5 °C increments showing an isodichroic point at ~205 nm. (B) The three pure component ECD curves resulted in from the CCA+ deconvolution: a U′-type (solid line); a C-type (dashed line) and a U-type (dotted line) pure ECD curve. (C) The relative contribution (%) changes of the pure component ECD curves over the temperature of TC5b_D9E assigned as: molten globule or intermediate state, U′-type (black squares); folded, C-type (open circle) and unfolded, U-type structures (semi-filled triangle) both at neutral and acidic pHs. (For similar ECD data of the other Trp-cage mutants see Supporting information).

The determination of the folded and unfolded baselines is also uncertain since it is difficult to estimate this information simply from the low-temperature and the high-temperature part of the ECD melting curves. The ambiguous baselines and the high number of fitted parameters made the thermodynamic studies reasonably doubtful. In Supporting information, however, we present the thermal stabilities (ΔG_U^{280K}, ΔG_U^{300K}) and melting temperatures derived from a two-state unfolding model (Table S1, see Supporting information).

In conclusion, the comprehensive temperature-dependent ECD analysis revealed that a typical TC5b scaffold melts in a more complicated manner than described by a two-state unfolding model. The subtle but important changes warranted a closer look and thus NMR structure elucidation was carried out at neutral and acidic pHs to reveal atomic resolution structural and folding information.

Structure Elucidation and Analysis

The 3D structure ensembles of all salt-bridge variants were determined based on the ^{1}H–^{1}H NOE distances (Table 2).

The D9E mutant seems to have the most well-defined NMR structure as the ensemble was calculated from as many as 415 restraints, of which 67 were long-range NOEs ($i \rightarrow i+n, n > 4$). The other three salt-bridge variants, TC5b, D9Aad_R16K and R16hR, had a moderate number of NOE restraints, thus somewhat larger RMSDs were obtained. For comparisons of the fingerprint regions of TC5b, D9E, D9Aad_R16K and R16hR see Figure S4. Upon lowering the pH, the number of NOEs decreased significantly. Most of the mutants with salt-bridge deletion can only be characterized with structure ensembles with high RMSD, this was also apparent from their few numbers of long-range NOE peaks. However, even at acidic pH these weakly stabilized molecular scaffolds – except that of D9N_R16A – have characteristic NOEs (e.g. cross-peaks of W^6–P^{12}, W^6–R/A^{16}, W^6–P^{17} and Y^3–P^{19}) representative of the original 3D structure. Thus, the Trp-cage folded structure (F) itself is somewhat preserved, but the overall fold becomes highly dynamicand the system is heading toward an intermediate state. Structures at acidic pH are of exceptional interest, as they could provide information about the molten globule form of the Trp-cage scaffold as seen from the CCA+ deconvolution studies.

Under acidic conditions, the NMR spectra of the salt-bridge equipped Trp-cage scaffolds (TC5b, D9E, D9Aad and R16hR) present a remarkable number of additional cross-peaks besides the signals associated with the major conformer (Figure 3). The 'acidic' spectra of the salt-bridge deleted mutants (D9S, D9N and R16A but except D9N_R16A) contain no or only a few extra NOE cross-peaks. These peaks emerged mostly in the fingerprint region of the spectra and are likely to belong to a second (intermediate) conformational state seen in the ECD deconvolution studies. In the cases of TC5b, R16hR and D9Aad, the peaks of the minor conformer were too ambiguous to be assigned, but for D9E an approximate assignment was completed (Figure 3). A second network of ^{1}H-resonances corresponded to the L^2, Y^3, I^4, Q^5, L^7, G^{10} and S^{14} residues, namely to those of the α-helical segment, and to that of the 3_{10}-helix. Contrary to expectations, Xaa–Pro peptide bond isomerization was not observed. The intense cross-peaks between the Hδ resonances of the prolines and the Hα resonance of the preceding residues (such as $G^{11}H\alpha1$–$P^{12}H\delta1$, Hδ2; $R^{16}H\alpha$–$P^{17}H\delta$ 1, Hδ2; $P^{17}H\alpha$–$P^{18}H\delta$# and $P^{18}H\alpha$–$P^{19}H\delta1$, Hδ2) are characteristic indicators of *trans* peptide bond conformation. No alternative cross-peaks between the i^{th} Hα and the $(i+1)^{th}$ proline Hα were observed which would be a sign of *cis* peptide bond conformation. However, a more reliable indicator of *cis* Xaa–Pro bond conformation would be the chemical shift difference of the $\alpha^{13}C\beta$ and $\delta^{13}C\gamma$ resonances. In the absence of ^{13}C labeled samples such information is not available for these miniproteins.

The observations regarding the minor conformer could be rationalized with a stable intermediate structure in which the

Table 2. NMR restraints and structure characterization of TC5b salt-bridge (Xxx[9] ↔Yyy[16]) mutants as measured at neutral and acidic pHs

	Exp. conditions		NOEs: intraresidual, sequential and long-range							RMSD[a] (Å)	
TC5b variants	pH	*T*/K	Σ	$i \rightarrow i$	$i \rightarrow i+1$	$i \rightarrow i+2$	$i \rightarrow i+3$	$i \rightarrow i+4$	$i \rightarrow i+n$	backbone	all atom
TC5b	6.9	282	270	73	65	14	50	18	50	0.26	1.96
TC5b	3.2	282	174	110	42	0	4	0	18	2.83	3.68
D9E	6.3	282	415	170	106	25	31	18	67	0.05	1.78
D9E	3.2	280	209	125	52	3	12	3	14	0.73	2.22
R16hR	6.7	280	274	139	58	9	17	15	36	0.27	0.89
R16hR	2.9	280	248	131	64	9	16	7	21	0.77	2.46
D9Aad_R16K	6.7	280	302	152	62	12	25	9	42	1.43	1.83
D9Aad_R16K	2.9	280	211	116	50	7	15	4	19	1.22	2.49
D9S	6.7	280	220	135	42	4	14	8	17	1.37	2.75
D9S	2.8	280	168	111	39	6	3	0	9	1.74	2.69
D9N	6.5	282	245	127	61	3	19	10	25	0.41	2.11
D9N	3.2	282	244	131	68	3	12	8	23	0.48	2.05
R16A	6.9	280	251	143	68	0	14	2	24	0.62	1.63
R16A	2.9	280	156	102	41	1	4	0	8	n.d.	n.d.
D9N_R16A	6.9	280	32	n.d.[b]	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
D9N_R16A	3.5	280	155	123	29	0	2	0	1	n.d.	n.d.

[a] Over the accepted 10 best structures pair-wise RMSD were calculated for residues 3–19 for all backbone atoms (Cα, C, N), as well as for all atoms including all Hs.
[b] n.d. no data available.

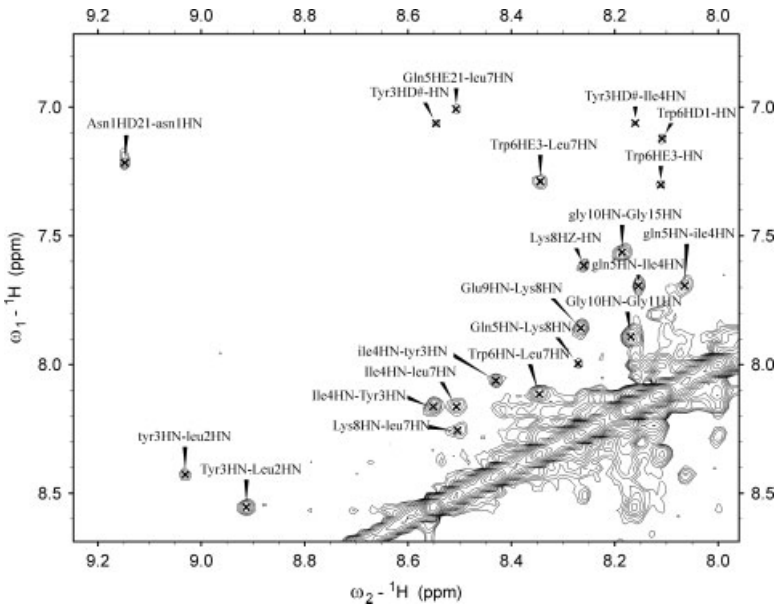

Figure 3. The HN–HN part of the NOESY spectrum of TC5b_D9E, pH 3.2 at 280 K. Major resonances are labeled with majuscule and minor resonances are labeled with minuscule.

hydrophobic core (P^{12}, P^{17}, P^{18} and P^{19}) around W^6 still exists, but the α- and the 3_{10}-helical regions have alternative fold(s). Interestingly, these minor peaks have large chemical shift dispersions with, in some cases, even larger chemical shift deviation (CSD) than that of the major conformer, but the CSD pattern of the minor peaks differs significantly from that of the Trp-cage F-state, which suggests an alternative fold. Although most of the minor peaks could be assigned, no structure calculation was performed on the minor conformer(s) because of the insufficient number of NOE cross-peaks. If we assume that under neutral pH the exchange rate between different states is fast on the NMR time-scale, then solely from the NMR spectra, the presence of an intermediate state cannot be excluded. However, the ECD deconvolution studies predict a considerable percent of intermediate state only for R16A and D9N_R16A while for the other variants they predict a maximum of 10% (for details see Figure S2).

Atomic Resolution Structure of Salt-bridge Variants

TC5b_D9E

The starting model is the previously described TC5b_D9E [28]. In this variant a more favorable contact between the hydrophobic

Journal of PeptideScience

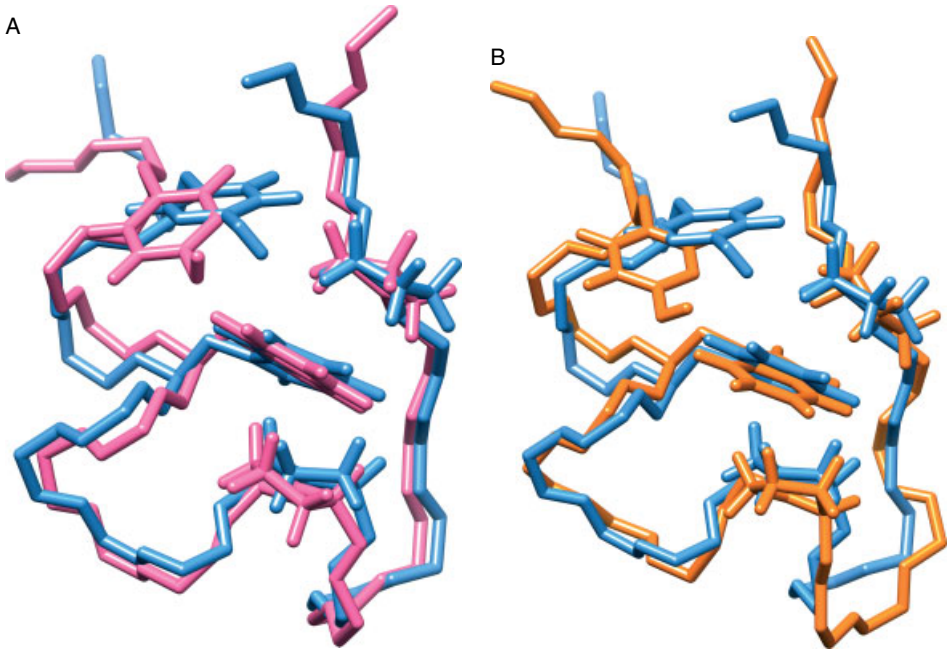

Figure 4. An overlay of the representative structures of TC5b_D9Aad_R16K (pink) and TC5b_R16hR (orange) with TC5b_D9E (blue). Note the high scaffold similarity. (All atoms are displayed for Y^3, W^6, P^{12} and P^{18} key residues, whereas only N, Cα and C′ atoms are shown for the remaining ones).

backside of the side-chain of the E^9 and the hydrophobic core was observed, which therefore resulted in a more compact 3D-structure (Table 2). The elongation of the original aspartate side-chain by a methylene group allows a higher flexibility for the salt-bridge, thus E^9 takes part more easily in a time average manner both in the salt-bridge formation with R^{16} and in the QxxxE interaction [51]. Besides these favorable interactions, the longer side-chain of E^9 (with respect to D^9) interacts favorably with the hydrophobic core. The Cγ atom of E^9 is just 2.86 Å away from the Cδ of W^6; and the overall elongated salt-bridge allows an alternative position for the PPII segment (R^{16}–P^{19}), so that the core becomes tighter as the PPII region shields the indole ring more efficiently. Although, the calculated ensemble at $T = 282$ K displayed a more compact 3D-structure with respect to TC5b, the ECD melting studies suggest that it is a less thermostable foldamer.

TC5b_D9Aad_R16K

With the aim of stabilizing the original molecular scaffold by using a better tuned salt-bridge mutant, D^9 and R^{16} were replaced by aminoadipic acid (Aad) and lysine residues respectively, in order to maintain the overall length of the salt-bridge. The structure was solved using many more NOE restraints (a total of 302) than detected for R16hR (274), but fewer than observed for D9E (415). The calculated structure ensemble has backbone and all atom RMSD values of 1.83 and 1.43 Å, respectively (Table 2). The double mutation has a marginal effect on the overall Trp-cage structure. In comparison with the D9E variant, the most obvious difference is related to the position of the Y^3 side-chain with respect to the indole ring, as it turns slightly 'away' from the hydrophobic core (Figure 4). Despite the scaffolds similarity to TC5b and TC5b_D9E, this variant is less thermostable than D9E.

TC5b_R16hR

The NMR structure of TC5b_R16hR was solved using a total of 274 distance restraints, of which 139 were interresidual; the calculated ensemble has backbone and all atom RMSD values of 0.27 and 0.89 Å, respectively (Figure 4). This R^{16} side-chain elongated salt-bridge variant has an extensive cross-peak network around the 'arm' of hR^{16} with residues W^6, S^{13}, S^{14}, G^{15} and P^{17}. This implies that the hR^{16} side-chain covers the hydrophobic core while shielding the 3_{10}-helical part, resulting in a compact molecular packing. The homoarginine side-chain is long enough to reach the backbone of the α-helix as a NOE cross-peak emerges between $hR^{16}H\varepsilon\#$ and D^9HN, which is a direct indicator of the salt-bridge formation. However, despite the detectable salt-bridge the overall hydrophobic core is slightly destabilized, in accordance with the ECD melting data.

In conclusion, most of these scaffolds display the features characteristic of the Trp-cage structure, namely an α-helical segment from L^2 to K^8, a short 3_{10}-helix between residue P^{12}–S^{14} and a PPII structural motif (X^{16}–P^{19}) at the C-terminus of the molecule. The central W^6 is packed against Y^3, L^7, G^{11}, P^{12} and P^{18} residues as they form the hydrophobic core. The differences between the high resolution NMR structures are the following: (i) orientation and position of residues 9 and 16: the only NOE-based direct information for the $Xxx^9 \leftrightarrow Yyy^{16}$ salt-bridge is between Asp^9 and $hArg^{16}$ in TC5b_R16hR, while for the other variants only indirect evidence is seen; (ii) the QxxxY interaction: directly measured for D9E and D9Aad_R16K variants (apparent from the NOEs between $Q^5H\varepsilon22$–$E^9H\beta\#$, $E^9H\gamma\#$ and $Q^5H\varepsilon22$–$Aad^9H\beta\#$, $Aad^9H\gamma\#$); (iii) the tightness of the hydrophobic core: the most compact core is formed in D9E followed by D9Aad_R16K and R16hR mutants, signaled by the different number of NOE cross-peaks surrounding the indole ring; and (iv) the $\pi-\pi$ stacking interaction: the salt-bridge fine-tuning affects the position of the phenol ring with respect to the indole ring so that the phenol ring of D9E is closer to the indole ring while the phenol rings of D9Aad_R16K and R16hR are somewhat further away from W^6, however the tilt-angles of the aromatic systems remain unchanged.

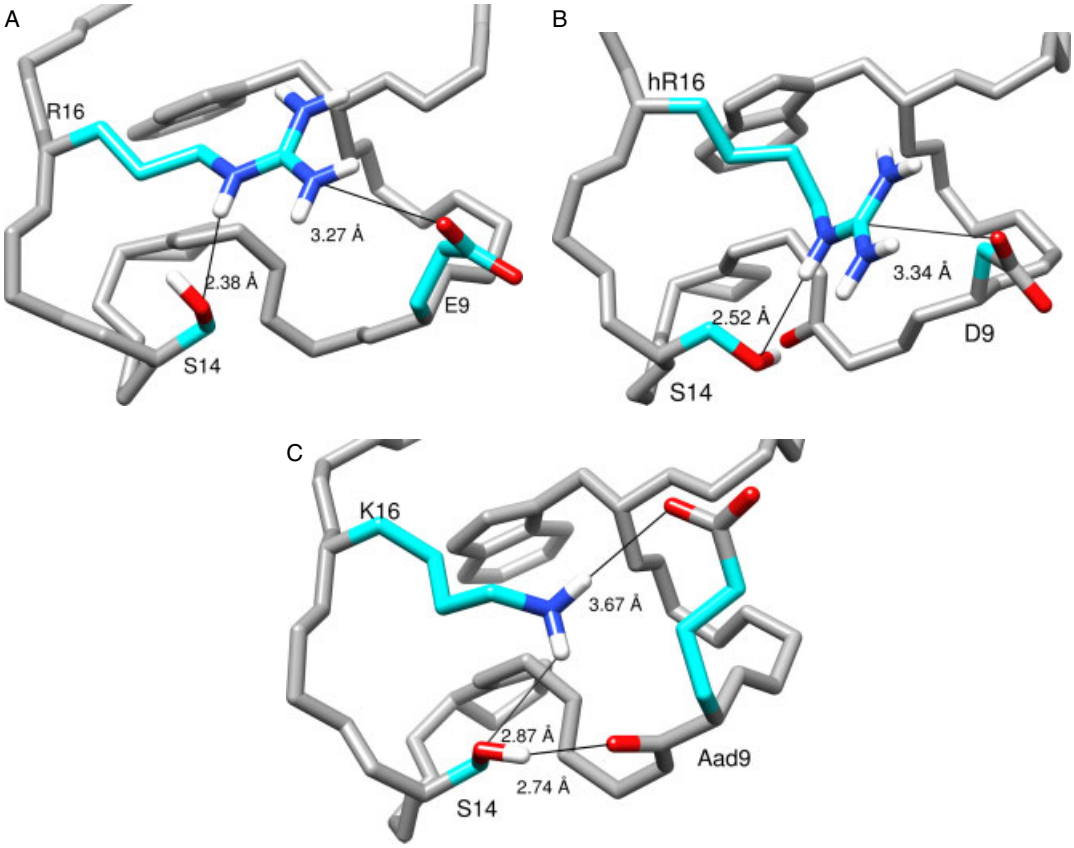

Figure 5. The H-bonding network of the salt-bridge and the neighboring residue(s) anchoring residues 9, 16 and 14. (A) TC5b_D9E: $R^{16}H\varepsilon$ points toward the side-chain oxygen of S^{14}, while the carboxyl group of E^{9} and the guanidinium group of R^{16} form the salt-bridge. (B) TC5b_R16hR: $hR^{16}H\zeta$ points toward the $S^{14}O\gamma$, while the carboxyl group of D^{9} and the guanidinium group of hR^{16} form the salt-bridge. (C) TC5b_D9Aad_R16K: $K^{16}H\zeta\#$ points toward the $S^{14}O\gamma$ and $S^{14}H\gamma$ points toward the carbonyl of Aad, while the carboxyl group of Aad^{9} and the amino group of K^{16} form the salt-bridge. This figure is available in colour online at wileyonlinelibrary.com/journal/jpepsci.

H-bonding Network

The H-bonding network of the salt-bridging residues was scrutinized for the above salt-bridge variants. Representative structures from the NMR ensembles of three selected models in Figure 5 show the complex nature of the H-bonding networks. For clarity, just one structure is depicted for all salt-bridge mutants, the ensemble representation gives a similar H-bonding network as can be seen in Figure S5. Although, the NOE-based structural ensembles do not incorporate direct experimental data on the H-bonds anchoring the salt-bridge, it was possible to evaluate the structural details due to a series of side-chain positioning NOEs. The RMSDs calculated for the backbone and for the side-chain heavy atoms of residues 9, 14 and 16 are 0.33 Å for D9E, 0.65 Å for R16hR and 0.58 Å for D9Aad.

Barua *et al*. [34] have reported that in TC10b the R^{16} $NH\gamma$ proton interacts with the $S^{14}O\gamma$ through an H-bond, which enhances the H-bonding ability of $S^{14}H\gamma$ toward the carbonyl oxygen of D^{9} or G [11]. This H-bonding network stabilizes the 3_{10}-helix with respect to the α-helix (*via* D^{9}) and the PPII structure (*via* R^{16}). For TC5b_D9E the arrangement of E^{9}, S^{14} and R^{16} residues is optimal, both for salt-bridge and H-bond formation (Figure 5A). The distances between the charged groups and between the H-bonding R^{16} $NH\varepsilon$ and $S^{14}O\gamma$ atoms are 3.27 and 2.38 Å, respectively. The position of $S^{14}H\gamma$ is not detectable with ^{1}H NMR as the side-chain hydroxyl proton of S^{14} is solvent exposed. The interaction network of residues 9, 14 and 16 of the R16hR mutant is similar to that of D9E (Figure 5B). In a representative model the salt-bridge separation, measured by the distance between the $hR^{16}C\eta$ and $D^{9}O\delta1$ atoms, is 3.34 Å, while the $hR^{16}H\zeta \rightarrow S^{14}O\gamma$ H-bond distance is 2.52 Å. The orientation of $S^{14}H\gamma$ is ideal for a H-bond toward the carbonyl oxygen of G^{10}. In TC5b_D9Aad, two new H-bonds were detected: one between $K^{16}H\zeta\#$ and $S^{14}O\gamma$ and the other between the $S^{14}H\gamma$ and the backbone carbonyl group of Aad^{9} (Figure 5C). This specific H-bond/salt-bridge network of the charged amino group therefore stabilizes the nearby 3_{10}-helix as well as the overall fold of the molecule.

The formation of a cation–π interaction could be another fold stabilizing effect associated with R^{16}, however in this molecular framework the orientation of the R^{16} side-chain differs from the optimal cation–π geometry where the positively charged guanidino group faces the electron-rich indole ring. In Trp-cages the R^{16} side-chain instead covers the electron-poor nodal plane of the indole ring.

Destabilized Mutants of Salt-bridge Deletion

Because of the elimination of one of the salt-bridge forming partners, a significant structural destabilization is expected to occur. Indeed, the total number of NOE cross-peaks decreased accordingly (Table 2). However, the remaining sets of NOEs of D9S, D9N and R16A are consistent with the Trp-cage formation and the calculated ensembles are similar to those in which the salt-bridge is operative. On the other hand, the neutral spectra of the

Journal of PeptideScience

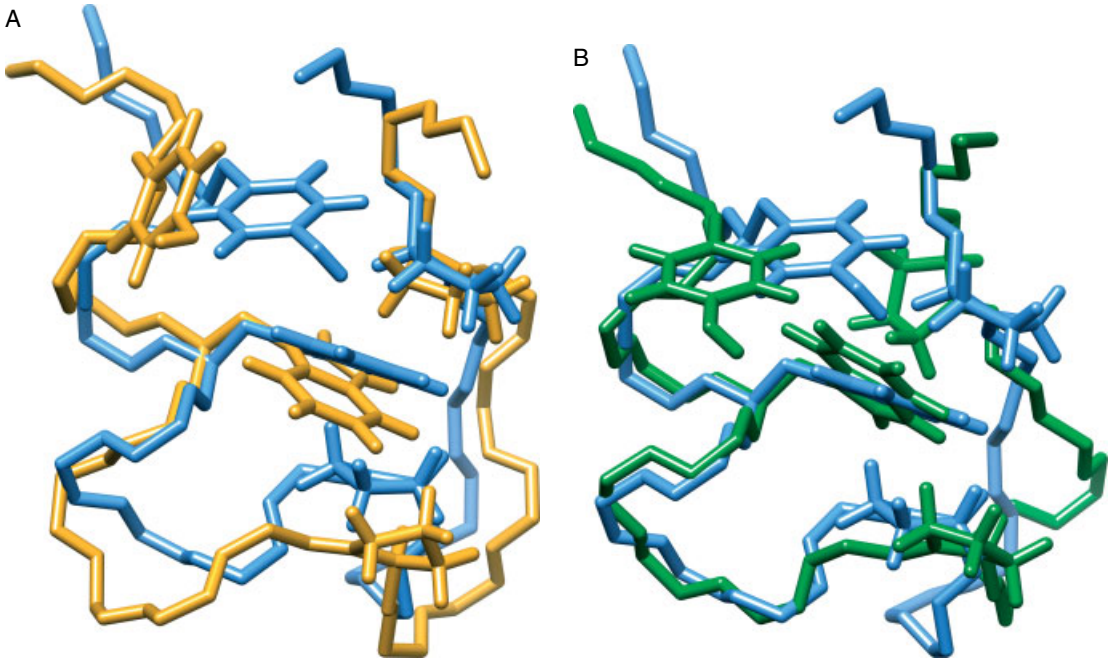

Figure 6. (A) An overlay of the representative structures of TC5b_D9S (gold) and TC5b_D9E (blue). The D9S mutation eliminates the parent salt-bridge, thus increases the flexibility of the 3_{10}-region. (B) An overlay of the representative structures of TC5b_R16A (green) and TC5b_D9E (blue). Although, the R16A mutation eliminates the salt-bridge, it does not affect the 3_{10}-helix formation. All atoms are displayed for Y^3, W^6, P^{12} and P^{18} whereas only N, Cα and C′ atoms are shown for the remaining residues.

double mutant D9N_R16A could only partly be assigned because of the sparseness of chemical resonances in the fingerprint region; therefore no structure elucidation was possible. The extraordinarily low number of cross-peaks can be explained by a medium slow exchange between folded (F), unfolded (U) and probably between intermediate (I) states, where the exchanging peaks are broadened. Thus, unlike the single mutant, the D9N_R16A double mutation affects the time-scale of unfolding significantly and not just the overall stability. However, at acidic pH (3.5) the NMR signals of D9N_R16A sharpened so the NOESY spectrum was assignable, but because of lack of long-range NOEs structure determination was not completed.

In the NOESY spectra of D9S, D9N and R16A the important long-range NOEs characteristic of the Trp-cage scaffold were assigned (e.g. those connecting residues Y^3-P^{19}, W^6-R/A^{16}, W^6-P^{18} and W^6-P^{19}), while no NOEs were found to connect neither residue 9 to 14, nor 13 to 16. Although S^9 and N^9 residues in D9S and D9N have high mobility and thus their side-chains have no interresidual NOE cross-peak, the R^{16} side-chain is partly fixed: it takes part in the core formation as usual by interacting with the indole ring of W^6 as observed from the $W^6H\delta1-R^{16}H\beta\#$ and $W^6H\delta1-R^{16}H\gamma\#$ NOE cross-peaks, respectively. The NOESY spectrum of D9N_R16A reports on a highly dynamic structure ensemble in which all the folded, unfolded and intermediate states are present. The folded state is moderately populated since the splitting of $G^{11}H\alpha$ resonances – which is a diagnostic site of Trp-cage formation [34] – is reduced from 2.46 ppm (TC5b) to 0.56 ppm ($\Delta G^{11}H\alpha$ of D9N and R16A are 1.21 and 0.79 ppm, respectively); the only long-range NOE was observed between $W^6H\delta1$ and $A^{16}H\beta\#$. According to the ECD deconvolution studies the intermediate U′ state is the major species for D9N_R16A at acidic pH and indeed in the fingerprint region of the NOESY spectrum several new sets of spin-systems emerged similar to those seen for the salt-bridge equipped Trp-cage variants. Again, there was no evidence of *cis* Xaa–Pro peptide bond formation.

In conclusion, in D9S and D9N both the hydrophobic core and the α-helix are formed, while the 3_{10}-helix presents an elevated flexibility and thus remains less defined. The scaffold loosened at one end by the lack of the salt-bridge $Xxx^9 \leftrightarrow Yyy^{16}$ which affects the other side of the foldamer which is apparent from the lack of a $\pi-\pi$ stacking interaction between the side-chains of the Y^3 and W^6 residues (Figure 6A). Although the chemical shift deviations of R16A indicate a loosely defined structure, the NOE-based 3D-structure is rather compact (Figure 6B), as over 25 long-range and well-distributed NOEs ensure the overall scaffold. The calculated structure ensembles show all the necessary secondary structural elements of the Trp-cage fold and only the PPII subunit is a bit extended as Ala^{16} presumably favors the PPII conformation.

pH Induced Destabilization

Acidification destroys the fold for all variants as can be seen in Figure S6, although some sign of hydrophobic core formation is still detectable. The structure determination from the spectra recorded at acidic pH was completed based on the comparison of the NOESY spectra recorded at neutral pH. Although, the peak dispersion was lower as almost all resonances shifted toward their random coil values and notably fewer interresidual NOE cross-peaks could be assigned. In spectra of TC5b, D9E, D9Aad and R16R the assignments were ambiguous because of the presence of peaks from the intermediate conformational state; therefore the calculated structures have higher uncertainty. Similarly, as in the neutral spectra of the salt-bridge deletion mutants, the main long-range NOEs were present indicating the preserved Trp-cage fold, and the reduced number of NOE cross-peaks from $G^{10}-S^{14}$ sequence indicates the destabilized 3_{10}-helix (Figure S6). The $\pi-\pi$ stacking interaction of Y^3 and W^6 is also reduced or diminished

at acidic pH which is an indirect consequence of the α-helix destabilization by the weakened H-bonds of the QxxxY interaction and of the α-helix.

The salt-bridge deleted TC5b variants retained some pH dependency as well as seen from the ECD analysis and the calculated ensembles changed accordingly. No structure was determined from the acidic R16A and D9N_R16A spectra because of the low number of long-range NOE cross-peaks. The acidic structure of D9S was solved using only 18 interresidual NOEs, thus the calculated ensemble has high all atom RMSD (2.69 Å). D9S at low pH has a highly flexible molten globule like structure in which the hydrophobic core is still partially formed as shown by the NOE cross-peaks between W^6-P^{12}, W^6-P^{18} and Y^3-P^{19}. Interestingly, D9N preserved much of the TC5b scaffold (Figure S6E) at acidic pH and both the α-helix and the hydrophobic core are formed (112 interresidual NOEs based the structure determination).

Conclusions

In summary, we found that the salt-bridge of the Trp-cage scaffold is not an isolated structure stabilizing element but rather an integrated part of a dense interaction network. Our observations strongly suggest that a salt-bridge in any protein should be evaluated in the context of its surrounding residues, as the interaction between the charged groups is at the same magnitude as the background interaction of the side-chain atoms.

In the cases of the examined mutants, the observed stability tendencies can be understood by considering four specific, but coupled interactions (i) electrostatic interaction between the charged groups; (ii) helix-stabilizing QxxxY interaction (iii) 3_{10}-helix stabilizing H-bond between residue 16 and S^{14} and (iv) hydrophobic interaction of the $-(CH_2)_3-$arm of R^{16} with the indole ring of W^6. On the basis of herein reported mutation studies it seems that the interaction network of the arm of residue 16 is of higher importance than the one operative between the negatively charged D^9/E^9 and the positively charged guanidino group. The elimination of one of the salt-bridge forming partners (e.g. D9S, D9N) is less drastic than the elimination of the hydrophobic arm of R^{16} (e.g. R16A).

The deconvolution of ECD spectra and the analysis of the acidic NMR spectra revealed that the melting scenario of the Trp-cage is not a simple two-state mechanism (F $\rightarrow$ U) but rather a more complex process with at least one intermediate state (F $\rightarrow$ I $\rightarrow$ U). The intermediate conformation differs significantly from the native fold: the magnitude and the pattern of minor conformer shifts suggest that the intermediate state is no longer helical but has rather a turn-rich conformation.

Acknowledgements

We are indebted to our colleague Imre Jákli for his crucial contribution to the CCA+ deconvolution studies. This work was supported by grants from ICGEB (CRP/HUN08-03), the Hungarian Scientific Research Fund (OTKA K72973, NK67800 and NI-68466) and TÁMOP-4.2.1.B-09/1/KMR.

Supporting information

Supporting information may be found in the online version of this article.

References

1 Anfinsen C. Principles that govern the folding of protein chains. *Science* 1973; **181**: 223–230.
2 Pace C. The stability of globular proteins. *CRC Crit. Rev. Biochem.* 1975; **3**: 1–43.
3 Kauzmann W. Some factors in the interpretation of protein denaturation. *Adv. Protein Chem.* 1959; **14**: 1–63.
4 Dill K. Dominant forces in protein folding. *Biochemistry* 1990; **29**: 7133–7155.
5 Stigter D, Dill K. Charge effects on folded and unfolded proteins. *Biochemistry* 1990; **29**: 1262–1271.
6 Israelachvili J. Van der Waals forces in biological systems. *Q. Rev. Biophys.* 1973; **6**: 341–387.
7 Rose G, Wolfenden R. Hydrogen bonding, hydrophobicity, packing, and protein folding. *Annu. Rev. Biophys. Biomol. Struct.* 1993; **22**: 381–415.
8 Wolfenden R. Waterlogged molecules. *Science* 1983; **222**: 1087–1093.
9 Wimley W, Gawrisch K, Creamer T, White S. Direct measurement of salt-bridge solvation energies using a peptide model system: implications for protein stability. *Proc. Natl. Acad. Sci. U.S.A.* 1996; **93**: 2985–2990.
10 Arakawa T, Timasheff S. Theory of protein solubility. *Methods Enzymol.* 1985; **114**: 49–77.
11 Harris B, Lau F, Fujii N, Guy R, Lim W. Role of electrostatic interactions in PDZ domain ligand recognition. *Biochemistry* 2003; **42**: 2797–2805.
12 Rodgers K, Sligar S. Mapping electrostatic interactions in macromolecular associations. *J. Mol. Biol.* 1991; **221**: 1453–1460.
13 Schreiber G, Shaul Y, Gottschalk K. Electrostatic design of protein-protein association rates. *Methods Mol. Biol.* 2006; **340**: 235–249.
14 Shaul Y, Schreiber G. Exploring the charge space of protein-protein association: a proteomic study. *Proteins* 2005; **60**: 341–352.
15 McCoy A, Chandana Epa V, Colman P. Electrostatic complementarity at protein/protein interfaces. *J. Mol. Biol.* 1997; **268**: 570–584.
16 Clarke J, Itzhaki L, Fersht A. Hydrogen exchange at equilibrium: a short cut for analysing protein-folding pathways? *Trends. Biochem. Sci.* 1997; **22**: 284–287.
17 Clarke J, Itzhaki L. Hydrogen exchange and protein folding. *Curr. Opin. Struct. Biol.* 1998; **8**: 112–118.
18 Englander S. Protein folding intermediates and pathways studied by hydrogen exchange. *Annu. Rev. Biophys. Biomol. Struct.* 2000; **29**: 213–238.
19 Englander S, Mayne L. Protein folding studied using hydrogen-exchange labeling and two-dimensional NMR. *Annu. Rev. Biophys. Biomol. Struct.* 1992; **21**: 243–265.
20 Perczel A, Hollósi M, Sándor P, Fasman G. The evaluation of type I and type II beta-turn mixtures. Circular dichroism, NMR and molecular dynamics studies. *Int. J. Pept. Protein Res.* 1993; **41**: 223–236.
21 Renner C, Schleicher M, Moroder L, Holak T. Practical aspects of the 2D 15N-[1H]-NOE experiment. *J. Biomol. NMR* 2002; **23**: 23–33.
22 Vanderkooi J, Dashnau J, Zelent B. Temperature excursion infrared (TEIR) spectroscopy used to study hydrogen bonding between water and biomolecules. *Biochim. Biophys. Acta* 2005; **1749**: 214–233.
23 Hollosi M, Perczel A, Fasman G. Cooperativity of carbohydrate moiety orientation and beta-turn stability is determined by intramolecular hydrogen bonds in protected glycopeptide models. *Biopolymers* 1990; **29**: 1549–1564.
24 Gloaguen E, Pollet R, Piuzzi F, Tardivel B, Mons M. Gas phase folding of an (Ala)4 neutral peptide chain: spectroscopic evidence for the formation of a beta-hairpin H-bonding pattern. *Phys. Chem. Chem. Phys.* 2009; **11**: 11385–11388.
25 Anderson D, Becktel W, Dahlquist F. pH-induced denaturation of proteins: a single salt bridge contributes 3–5 kcal/mol to the free energy of folding of T4 lysozyme. *Biochemistry* 1990; **29**: 2403–2408.
26 Neidigh J, Fesinmeyer R, Andersen N. Designing a 20-residue protein. *Nat. Struct. Biol.* 2002; **9**: 425–430.
27 Neidigh J, Fesinmeyer R, Prickett K. Andersen N. Exendin-4 and glucagon-like-peptide-1: NMR structural comparisons in the solution and micelle-associated states. *Biochemistry* 2001; **40**: 13188–13200.
28 Hudáky P, Stráner P, Farkas V, Váradi G, Tóth G, Perczel A. Cooperation between a salt bridge and the hydrophobic core triggers fold stabilization in a Trp-cage miniprotein. *Biochemistry* 2008; **47**: 1007–1016.

29 Hatfield M, Palermo N, Csontos J, Murphy R, Lovas S. Quantum chemical quantification of weakly polar interaction energies in the TC5b miniprotein. *J. Phys. Chem. B.* 2008; **112**: 3503–3508.
30 Snow C, Zagrovic B, Pande V. The Trp cage: folding kinetics and unfolded state topology via molecular dynamics simulations. *J. Am. Chem. Soc.* 2002; **124**: 14548–14549.
31 Zhou R. Trp-cage: folding free energy landscape in explicit water. *Proc. Natl. Acad. Sci. U.S.A.* 2003; **100**: 13280–13285.
32 Juraszek J, Bolhuis P. Sampling the multiple folding mechanisms of Trp-cage in explicit solvent. *Proc. Natl. Acad. Sci. U.S.A.* 2006; **103**: 15859–15864.
33 Hu Z, Tang Y, Wang H, Zhang X, Lei M. Dynamics and cooperativity of Trp-cage folding. *Arch. Biochem. Biophys.* 2008; **475**: 140–147.
34 Barua B, Lin J. Williams V, Kummler P, Neidigh J, Andersen N. The Trp-cage: optimizing the stability of a globular miniprotein. *Protein Eng. Des. Sel.* 2008; **21**: 171–185.
35 Shi X, Parks J. Fluorescence lifetime probe of biomolecular conformations. *J. Am. Soc. Mass. Spectrom.* 2010; **21**: 707–718.
36 Jákli I, Perczel A. The inherent flexibility of peptides and protein fragments quantitized by CD in conjunction with CCA+. *J. Pept. Sci.* 2009; **15**: 738–752.
37 Perczel A, Hollósi M, Tusnády G, Fasman G. Convex constraint analysis: a natural deconvolution of circular dichroism curves of proteins. *Protein Eng.* 1991; **4**: 669–679.
38 Perczel A, Park K, Fasman, G. Analysis of the circular dichroism spectrum of proteins using the convex constraint algorithm: a practical guide. *Anal. Biochem.* 1992; **203**: 83–93.
39 Goddard T, Kneller D. *SPARKY 3*, University of California: San Franscisco, 1999.
40 Bundi A, Wuthrich K. ^{1}H-NMR parameters of the common amino acid residues measured in aqueous solutions of the linear tetrapeptides H-Gly-Gly-X-L-Ala-OH. *Biopolymers* 1979; **18**: 285–297.
41 Fesinmeyer R, Hudson F, Andersen N. Enhanced hairpin stability through loop design: the case of the protein G B1 domain hairpin. *J. Am. Chem. Soc.* 2004; **126**: 7238–7243.
42 Schwarzinger S, Kroon G, Foss T, Chung J, Wright P, Dyson H. Sequence-dependent correction of random coil NMR chemical shifts. *J. Am. Chem. Soc.* 2001; **123**: 2970–2978.
43 Brünger A, Adams P, Clore G, DeLano W, Gros P, Grosse-Kunstleve R, Jiang J, Kuszewski J, Nilges M, Pannu N, Read R, Rice L, Simonson T, Warren G. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* 1998; **54**: 905–921.
44 Berman H, Westbrook J, Feng Z, Gilliland G, Bhat T, Weissig H, Shindyalov I, Bourne P. The Protein Data Bank. *Nucl. Acids Res.* 2000; **28**: 235–242.
45 Pettersen E, Goddard T, Huang C, Couch G, Greenblatt D, Meng E, Ferrin T. UCSF Chimera – a visualization system for exploratory research and analysis. *J. Comput. Chem.* 2004; **25**: 1605–1612.
46 Streicher W, Makhatadze G. Unfolding thermodynamics of Trp-cage, a 20 residue miniprotein, studied by differential scanning calorimetry and circular dichroism spectroscopy. *Biochemistry* 2007; **46**: 2876–2880.
47 Wafer L, Streicher W, Makhatadze G. Thermodynamics of the Trp-cage miniprotein unfolding in urea. *Proteins* 2010; **78**: 1376–1381.
48 Mok K, Kuhn L, Goez M, Day I, Lin J, Andersen N, Hore P. A pre-existing hydrophobic collapse in the unfolded state of an ultrafast folding protein. *Nature* 2007; **447**: 106–109.
49 Ahmed Z, Beta I, Mikhonin A, Asher S. UV-resonance raman thermal unfolding study of Trp-cage shows that it is not a simple two-state miniprotein. *J. Am. Chem. Soc.* 2005; **127**: 10943–10950.
50 Juraszek J, Bolhuis P. Rate constant and reaction coordinate of Trp-cage folding in explicit water. *Biophys. J.* 2008; **95**: 4246–4257.
51 Huyghues-Despointes B, Klingler T, Baldwin R. Measuring the strength of side-chain hydrogen bonds in peptide helices: the Gln.Asp (i, i + 4) interaction. *Biochemistry* 1995; **34**: 13267–13271.

