Cooperation between a Salt Bridge and the Hydrophobic Core Triggers Fold Stabilization in a Trp-Cage Miniprotein†

Péter Hudáky,‡ Pál Stráner,‡ Viktor Farkas,§ Györgyi Váradi,∥ Gábor Tóth,∥ and András Perczel*,‡,§

Laboratory of Structural Chemistry and Biology, Institute of Chemistry, Eötvös Loránd University, Pázmány Péter sétány 1/A, H-1117 Budapest, Hungary, Protein Modelling Group MTA-ELTE, Institute of Chemistry, Eötvös Loránd University, Pázmány Péter sétány 1/A, H-1117 Budapest, Hungary, and Department of Medical Chemistry, Faculty of General Medicine, University of Szeged, Dóm tér 8, H-6720 Szeged, Hungary

Received July 11, 2007; Revised Manuscript Received October 18, 2007

ABSTRACT: Miniproteins are adequate models to study various protein-structure modifying effects such as temperature, pH, point mutation(s), H-bonds, salt bridges, molecular packing, etc. Tc5b, a 20-residue Trp-cage protein is one of the smallest of such models with a stable 3D fold (Neidigh J. W. et al. (2002) *Nat. Struct. Biol. 9*, 425−430). However, Tc5b exhibits considerable heat-sensitivity and is only stable at relatively low temperatures. Here we report a systematic investigation of structural factors influencing the stability of Tc5b by solving its solution structure in different environments, varying temperature, and pH. The key interactions identified are the hydrophobic stacking of the aromatic rings of Tyr3 and Trp6 and the salt bridge formed between Asp9 and Arg18. To verify the importance of these interactions, selected variants (mutated, glycosylated and truncated) of Tc5b were designed, prepared, and investigated by NMR. Indeed, elimination of either of the key interactions highly destabilizes the structure. These observations enabled us to design a new variant, Tc6b, differing only by a methylene group from Tc5b, in which both key interactions are optimized simultaneously. Tc6b exhibits enhanced heat stability and adopts a stable fold at physiological temperature.

The atomic level description of protein folding, providing characteristic properties of sequence-dependent optimum conditions of the process, is of long-standing interest (*1*−*4*). The determination of the minimal requirements of a polypeptide chain to fold up into a compact 3D structure, composed of characteristic secondary structural elements, is under investigation (*5*, *6*). Typically, short linear oligopeptides do not adopt a well-defined and unique overall fold. Thus, the folding ability of specific peptide sequences, composed of 20−40 amino acid residues only, called as miniproteins (*7*) have been intensively studied (*8*−*14*). There is a contest for finding the miniprotein composed of the shortest primary sequence showing at the same time the most characteristics of globular proteins. Several different fold types were recognized as miniproteins, such as betanova (*15*), WW (*16*), and Trp-cage (*17*). To date, perhaps the best candidate is a Trp-cage miniprotein that contains only 20 amino acid residues and has in many respects globular protein-like structural features, without the structure-stabilizing effect of disulfide bridge(s). As expected, the 3D-fold of these "fragile" molecules is very sensitive to most environmental and structure-modifying effects (*18*). Although a uniformly accepted set of criteria defining the self-folding principles of miniproteins has yet to be created (*19*, *20*), there are several important features that are generally accepted: (i) existence of and contacts between multiple regular secondary structural elements; (ii) well-defined side-chain orientations due to 3D-packing; (iii) significant chemical shift dispersion of the constituent atoms; (iv) a major fraction of amide protons protected from H/D exchange; (v) folding cooperativity.

Neidigh et al. have designed a 20 residue-long sequence (NLYIQWLKDG GPSSGRPPPS) (*17*) called Tc5b that spontaneously folds into a "Trp-cage" 3D structure containing both well-defined secondary structure elements and tertiary structure contacts typical of larger proteins. The solution state structure reported by Neidigh et al. (1L2Y) (*17*) reveals a compact hydrophobic core (Tyr3 and Trp6) against which three prolines, Pro12, Pro18, and Pro19, plus a glycine (Gly11), are packed. Its structure is typically described by an α-helix spanning residues 1 through 9, followed by a central region composed of three β-turns (residues 10 to 14), and a polyproline II structure at the C-terminus (*17*). Although not detected directly in solution, molecular dynamics-based structure analysis revealed that the salt bridge between the side-chain carboxylic group of Asp9 and the guanidino group of Arg16 could provide a major force in stabilizing the overall fold (*21*). Salt bridges are known to have fold orienting effects in proteins (*22*, *23*).

Although TC5b may adopt a protein-like 3D-fold having a single dominant conformer well represented by an average

† This work was supported by grants from the Hungarian Scientific Research Fund (OTKA T047186, TS49812, NI68466), MediChem-Bats2, ICGEB (Hun04-03), and by the Hungarian Academy of Sciences (TKI ProtModGroup).

* To whom correspondence should be adressed. Tel.: +36-1-209-0555/1653; fax: +36-1-372-2620; e-mail: perczel@chem.elte.hu.

‡ Laboratory of Structural Chemistry and Biology, Eötvös Loránd University.

§ Protein Modelling Group MTA-ELTE, Eötvös Loránd University.

∥ University of Szeged.

10.1021/bi701371x CCC: $40.75
Published on Web 12/28/2007

Table 1: Selected Analytical Parameters of Miniprotein Variants[a]

Code	Description	Sequence	MW (Calculated) [Da]	MW (Measured) [Da]	HPLC retention time [min]
TC5b	Original molecule	NLYIQWLK**D**GGPSSGRPPPS	2169.44	2169.7	7.29
TC6a	Salt bridge-deficient	NLYIQWLK**N**GGPSSGRPPPS	2168.46	2169.6	7.17
TC6as	Galactosylated	NLYIQWLK**N**GGPSSGRPPPS	2329.66	2330.9	8.85
TC6b	Salt bridge-optimized	NLYIQWLK**E**GGPSSGRPPPS	2183.47	2183.5	8.56
TC5b-5	N-terminally truncated	WLK**D**GGPSSGRPPPS	1536.71	1537.2	7.05

[a] The crossed N denotes galactosylated asparagine (N(Gal)). The underlined amino acid at position 9 indicates the change of the peptide sequence.

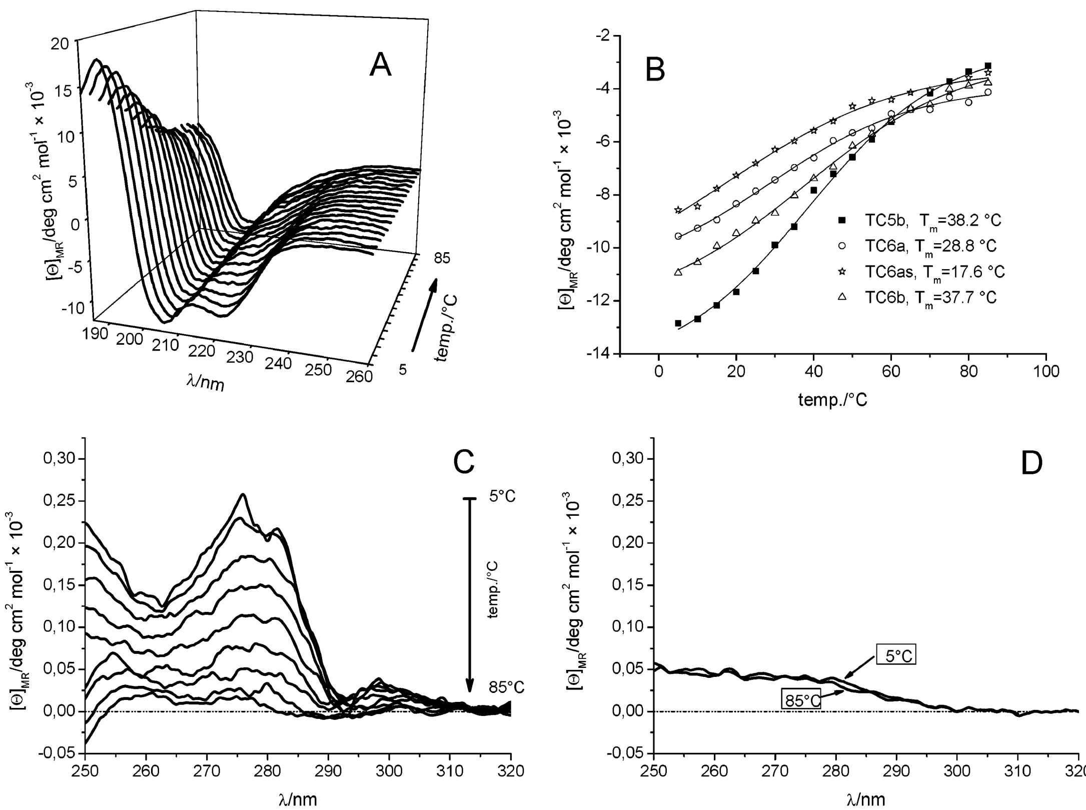

FIGURE 1: (A) Temperature-dependent far-UV CD spectra of Tc6as, a typical variant of a Trp-cage miniprotein as a function of the temperature (5 °C to 85 °C). (B) Thermal denaturation curves of Tc5b, Tc6a, Tc6as, and Tc6b by using CD spectral information detected at 222 nm. (C) Temperature-dependent near-UV CD spectra of Tc6b (D) and of Tc5b-5. (The arrow on panel C indicates the increasing temperature from 5 °C to 85 °C.)

structure, the stability of the molecule is relatively low, requiring data collection at 282 K, a temperature remarkably below the physiological one. We have revisited the original model and engineered several variants to find a mutant stable at room temperature, with the aim to determine the structural effects related to thermosensitivity. We introduced several types of modifications related to Asp9, a key residue with multiple roles: it is (i) the residue terminating the first secondary structure element (the α-helix), (ii) the amino acid also forming the first type I β-turn, and (iii) the moiety forming the salt bridge with Arg16. Our variations, summarized in Table 1, include mutations (TC6a and TC6b), specific side-chain modifications (glycosylation at position 9, TC6as) as well as the truncation of the N-terminus of the molecule (TC5b-5). The glycosylated model, a prototype of the post-translationally modified proteins, incorporates Asn(Gal) which has the potential of stabilizing the 3D-fold of proteins (*24*−*27*) and having a role in specifying the activity of several enzymes (*28*).

The versatile nature of CD and NMR spectroscopy is useful to study solution-phase structural properties of these molecules. Dynamic nature of these systems can also be assessed by these spectroscopic methods (*29*, *30*). Experimental data collection in conjunction with theoretical calculations (e.g., molecular dynamics) has proved fruitful in invesitigating structure−stability properties of miniproteins (*31*−*34*) Fine-tuning the overall fold of TC5b was possible by assessing the effect of the temperature and the pH for most mutants.

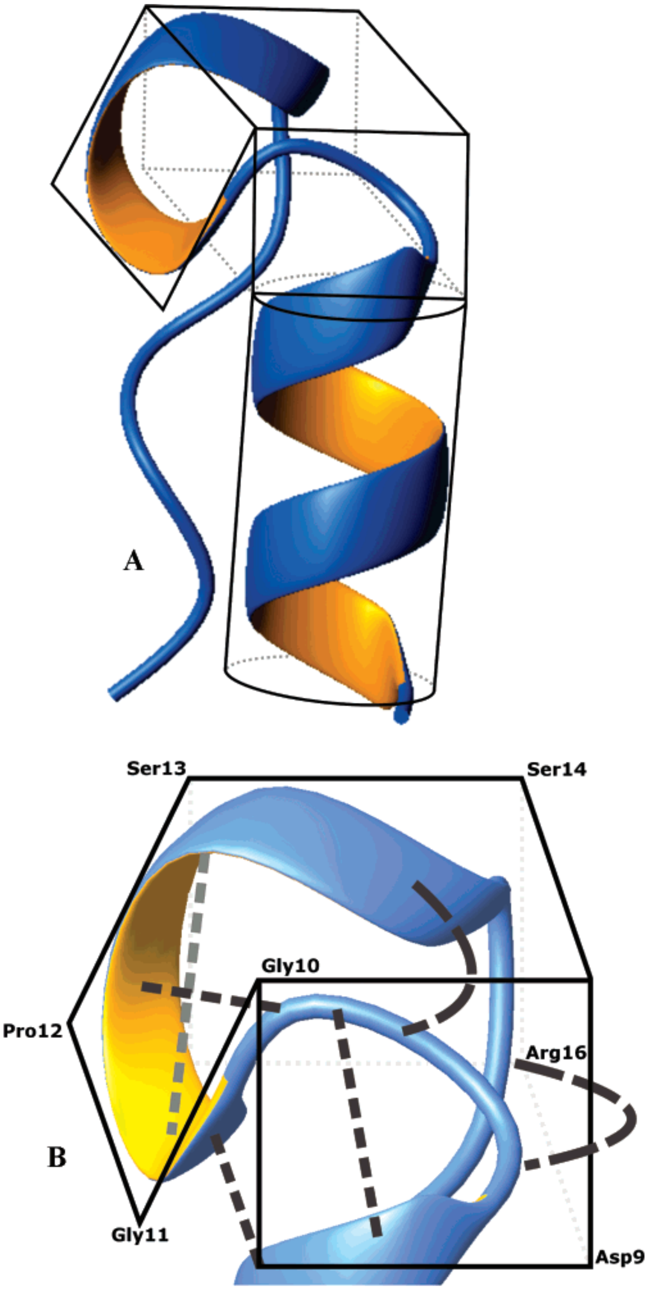

FIGURE 2: Average solution structure of TC5b (A) with the enlarged central region, Asp9-Ser14 (B), containing as many as three different β-turns.

MATERIALS AND METHODS

Peptide Synthesis, Purification and Analysis. Peptides were synthesized by solid-phase technique utilizing Fmoc chemistry. Wang resin (0.79 mmol/g; 0.25 mmol; 0.32 g) was used as a solid support. The side-chain protecting groups were as follows: Asn(Trt), Asp(Obu^t), Arg(Pbf), Glu(Obu^t), Gln(Trt), Lys(Boc), Ser(Bu^t), and Tyr(Bu^t). After the incorporation of the C-terminal serine by applying DCC/HOBt coupling in the presence of 4-dimethylaminopyridine, peptides were synthesized according to the standard DCC or DCC/HOBt protocol with a threefold excess of reagents. A glycosylated asparagine residue was incorporated as Fmoc-Asn(Gal)-OH and was prepared according to the following procedure (*35*). Fmoc-Asp-Obu^t and equivalent amounts of HBTU and HOAt were dissolved in DMF. After the mixture was stirred for 2 min, 1 equiv of 1-deoxy-1-aminogalactose and 1.5 equiv of *N*,*N*-diisopropylethylamine were added and stirred for 4 h. The formed Fmoc-Asn(Gal)-Obu^t was isolated by precipitating with ice-cold diethyl ether. Then TFA was added, stirred for 30 min, and evaporated. The resultant Fmoc-Asn(Gal)-OH was washed with ice-cold diethyl ether, dried, and built to the peptide using DCC/HOBt activation. Peptides were detached from the resin with a 95:5 (v:v) TFA/water mixture containing 3% (m/v) dithiothreitol at room temperature for 2.5 h. The resin was removed, and the peptides were precipitated by the addition of ice-cold diethyl ether. The precipitated peptides were collected by filtration and dried. The crude peptides were purified by semipreparative RP-HPLC using a solvent system of (A) 0.1% TFA and (B) 80% MeCN, 0.1% TFA on a Phenomenex Jupiter (Torrance, CA) C18 10 μm column (15 $\times$ 250 mm). The HPLC apparatus was made by Knauer (Berlin, Germany). Absorbance was detected at 220 nm. The peptides were eluted first with 10% solvent B for 3 min, followed by a linear gradient from 10 to 35% B over 50 min (flow rate: 3.5 mL/min). The purity of the peptides was evaluated by analytical RP-HPLC on a 4.6 $\times$ 250 mm Vydac or Lichrosorb C18 column using an HP 1050 HPLC system made by Hewlett-Packard (Palo Alto, CA). The integrity of the products was verified by mass spectrometry using a Finnigan TSQ 7000 tandem quadrupol mass spectrometer equipped with electrospray ion source. The analytical data of the peptides are summarized in Table 1.

CD Data and Ensemble Deconvolution by CCA+. Far-UV and near-UV CD measurements were performed on a Jasco J810 dichrograph in 0.1 and 1.0 cm quartz cells, respectively. The temperature at the cell was controlled by a Peltier-type heating system. For each measurement, the sample in the cell was allowed to equilibrate for 5 min at the adjusted temperature prior to data acquisition. All experiments reported in this paper were measured in a range of 5 °C to 85 °C by increasing the temperature by increments of 5 °C. Each spectrum was obtained by averaging a total of four scans. The solvent reference spectra were used as baselines, which were automatically subtracted from the peptide CD spectra. CD band intensities were expressed in mean residue ellipticity ($[\Theta]_{MR}$, deg $\times$ cm^2/dmol). A total of 85 spectra were collected and comprehensively analyzed by using the latest release of the CCA+ deconvolution program (*36*−*38*). Spectral analyses have resulted in three pure component CD curves with the appropriate conformational weights (Figure 1).

NMR Data Acquisition Procession and Structure Determination. NMR experiments were recorded on a Bruker DRX500 spectrometer at two pH values (low pH ∼3.1 and neutral pH ∼6.6 and at two different temperatures (low T = 282 K and ambient T = 300 K) (Table S2).

NMR samples contained ∼2 mM protein dissolved in 90% H_2O/10% D_2O. No buffer was applied. All NMR data were referenced to H_2O. Resonance assignment of protons was obtained from $^1H-^1H$ TOCSY and $^1H-^1H$ NOESY experiments. Typical spectral widths were 10 ppm. TOCSY measurements were typically taken with 58 to 72 ms spinlock time. Unambiguous backbone as well as side-chain proton assignment of the residues could be achieved for all residues.

To maintain conditions for comparison of the resonances, the same processing protocol was used for all spectra measured at all conditions (282 and 300 K and at 3.1 and

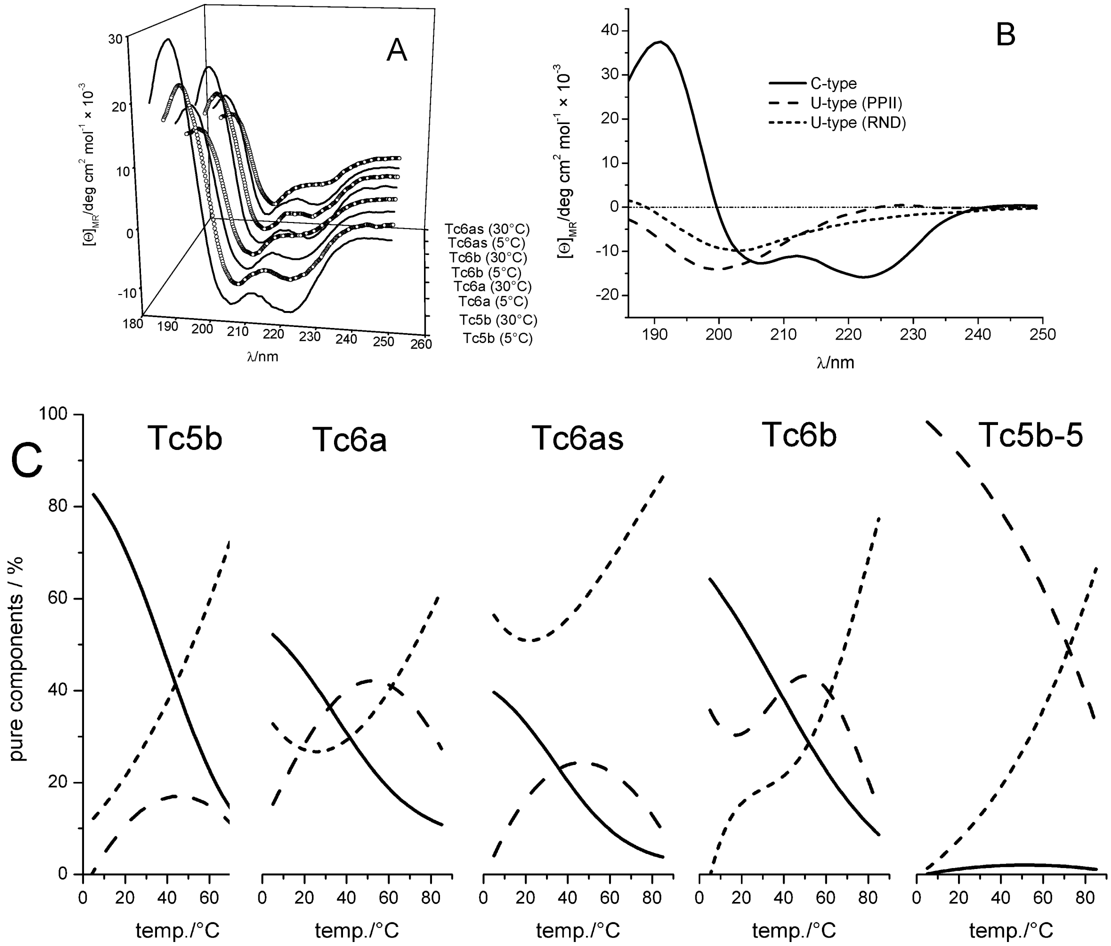

FIGURE 3: (A) Far-UV CD spectra of the five miniproteins recorded at 5 °C and 30 °C. (B) The three pure components resulting from deconvolution. (C) Relative contribution (%) of the pure component CD curves (C-type (—), U-type (PPII) (− − −), and U-type (RND) (---) in the experimental spectra recorded as the function of the temperature for all five miniproteins (curves fitted by using the deconvolution results in relative weights).

6.6 pH values). All data were processed by using the program nmrPipe (*39*). Spectra were apodized with a 90° shifted cosine-squared function and zero-filled before Fourier transformation. Polynomial baseline correction was applied. Resonance assignments were completed with the program Xeasy (*40*). All structures were calculated with the program Xplor (*41*). The 20 best structures were selected out of 100. The quality of the structures was evaluated by Procheck-NMR (*42*). All figures depicting molecular folds were created with the program MolMol (*43*).

RESULTS AND DISCUSSION

Lowering thermal motion stabilizes "softly" folded macromolecules. The structure of TC5b was solved previously by Neidigh et al. (*17*) at 280 K, which consists of an α-helix (Leu2-Lys8), a type I (Lys8-Asp9), a type II′ (Gly11-Pro12), and again a type I (Pro12-Ser13) β-turn followed by a linker (Ser14-Arg16), and completed by a PPII (Pro17-Pro19) type secondary structure element (Figure 2).

In order to locate structural differences as a function of the temperature, CD and NMR experiments were conducted at both low and room temperatures, at neutral as well as at acidic pH values.

1. Secondary Structure Content Analysis of the Model Systems by CD Spectroscopy. Far- and near-UV CD spectra analysis was completed by recording the CD spectra of all five model systems over a wide temperature range (5 to 85 °C) in water (Figure 3A). The ensemble of a total of 85 (= 5 × 17) CD curves were studied simultaneously, by using the recently fine-tuned version of the Convex Constraint Analysis Plus (CCA+) program (*36*−*38*). The spectral deconvolution have resulted in three different pure CD curves (Figure 3B). The first one shows spectral features typical of a C-type (*44*, *45*), the second one resembles a U-type CD curve. The third pure CD curve, a shifted U-type pure component spectrum, is often associated with unordered or atypical backbone folds (*46*). It is common to associate a C-type CD curve with α- or 3_{10}-helices as well as with type I β-turn(s). The second type of pure CD curve, with a small but significant positive maximum at 225 nm and a large negative maximum around 205 nm is an indicator of polyproline II conformation and of type II′ β-turns. (Figure 3B) For all five miniproteins, the increase of temperature correlates with the increase of relative ratio of the third component CD curve. The latter one has a large negative maximum at around 200 nm, and it is commonly associated with the presence of an unstructured backbone fold (Figure 3C). Except for Tc6a the ratio of unfolding increases and reaches 70% or higher at about 85 °C, a temperature at which some residual helix or type I β-turn as well as some PPII structure content can still be traced. Except for Tc5b-5 (Figure 1D), the decreasing intensity at ∼275 nm of the near

UV CD spectra also supports the gradual unfolding of the Trp minicage (Figure 1C). In the case of Tc5b-5, the lack of the backside stabilizing effect provided by Tyr3 for Trp6 makes the miniprotein fold vulnerable resulting in the complete unfolding of the first half of the molecule. This is clearly reported by the absence of the near-UV CD bend (Figure 1D) and by the close to zero helix content of the far-UV CD spectra (Figure 3C). However, the latter molecule exhibits relatively large PPII content, especially at low temperature, which could be initiated by the second half (-[16]-Arg-Pro-Pro-[19]Pro-) of the model system in the absence of α-helix. The former type of secondary structure element rapidly decreases with the increase of temperature (Figure 1C), Thus, as expected, the PPII content is also lost at higher temperature.

Unlike for TC6a and TC6as, the thermal denaturation curves of Tc6b and Tc5b (Figure 1B) reflect a similar level of structure stability signaling that their 3D-fold "melts" only at above 38 °C. However, the 15−20% lower contribution of the RND type pure CD curve in Tc6b with respect to Tc5b signals that they must exhibit a somewhat different fold which can hardly be traced and explained at a "global" level depicted by CD spectroscopy. In order to reveal the atomic details and the subtle differences of the above Trp-cage miniproteins, a detailed NMR structure elucidation was conducted, varying pH and temperature.

2. Structure and Characteristics of TC5b. 2.1. TC5b at Low Temperature (T = 282K) and Neutral pH (6.9). As the first step of this study, we revisited the structural properties of TC5b using conditions similar to those applied by Neidigh et al. (*17*). The chemical shifts obtained are shifted a little, $\Delta\delta \sim 0.14 \pm 0.22$ ppm on average, most likely due to minor temperature, pH, and buffer differences. The overall 3D structure of the miniprotein is very similar to what was determined previously. The original structure was solved using 169 distance restraints (NOEs), of which 115 were assigned to be interresidual. We were able to collect a higher number of distance restraints (260), of which 197 are interresidual. As expected, the higher number of NOEs gave rise to a somewhat lower backbone rmsd value; namely, 0.64 Å compared to 0.81 Å. The "all heavy atom" rmsd values also decreased from 1.52 Å to 1.20 Å. (rmsd values of the "original structure" were determined by using structures calculated with the deposited restraint set.) (Figure 4).

Minor structural differences between the original and our structure are as follows: (i) the first amino acid residue at the N-terminus of the α-helix has higher structural flexibility (there is only one distance restraint in our structure between residue 1 and 4), and (ii) the phenyl ring of Tyr3 in our structure is slightly rotated to have a more planar orientation with respect to the indole ring of Trp6.

2.2. TC5b at Ambient Temperature (T = 300 K) and at Neutral pH (6.52). The increase in temperature was reported to destabilize the overall 3D-fold of TC5b, now tested at 300 K. The total number of NOEs at higher temperature decreased (260→152). While the total number of intraresidual NOEs remained roughly the same (63 at 282 K and 62 at 300 K) the number of interresidual NOEs decreased significantly (197→90). The protein at room temperature has backbone and heavy atom rmsd values of 1.16 Å and 1.78 Å, respectively.

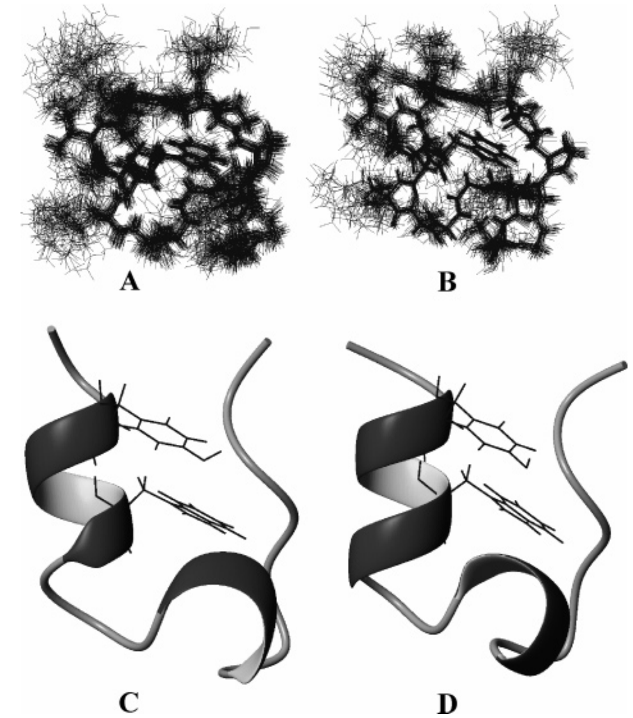

FIGURE 4: Structure ensemble of TC5b using the original data set of restraints (A), and the one currently determined (B), plus the corresponding two representative structures.

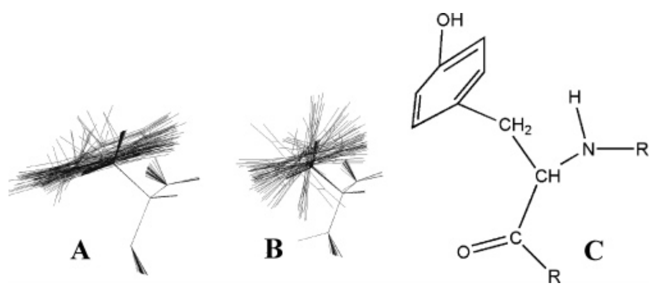

FIGURE 5: Superposition of the aromatic ring of Tyr3 at low (A) and room temperature (B). (The chemical structure of the residue is depicted (C)). Note that a higher fluctuation of the aromatic ring is associated with the increase in temperature.

The most pronounced difference between the low- and the ambient temperature structures is related to Tyr3 (Figure 5). The phenyl ring of Tyr3 is involved not only in the hydrophobic core formation but also in connecting the N− and C-termini of the molecule. As many as four NOEs were detected at 282 K and none at 300 K. In conclusion, the overall fold of TC5b at higher temperature is loosened, and destabilization is located near Tyr3 because of increased thermal motion. As depicted by CD (Figure 3C), the increase of temperature monotonically decreases the α-helical content and increases the unstructured ratio within the miniprotein.

2.3. TC5b at Low Temperature (T = 282 K) and Acidic pH. Lowering the pH reduces (*47*) the exchange of the amide protons, making some resonances narrower, thus tempting the NMR spectroscopist to carry out experiments at lower pH. Although for many proteins it was proven that their structures remain essentially the same at pH ≈3 as at neutral pH (*48*), there are examples where partial unfolding occurs in acidic conditions (*49*). This seems to be the case for Tc5b. At pH = 3.2 the structure is more mobile, indicated by the lower number of NOEs (67 interresidual). Thus, the backbone

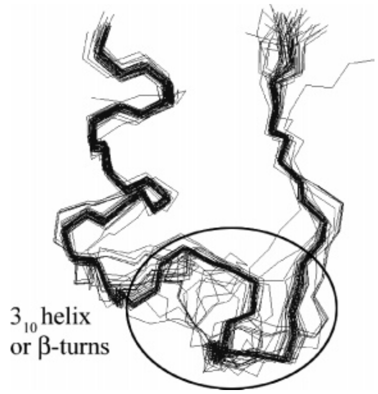

FIGURE 6: Superimposition of the backbone structures of Tc5b determined at pH = 3.17 and at lower temperature (T = 282 K). Protonation of the side chain of Asp9 increases the flexibility of the central β-turn region of the miniprotein by eliminating the Asp9-Arg16 salt bridge.

rmsd value increases from 0.64 Å up to 1.2 Å, and the "all heavy atom" rmsd values increase from 1.20 Å to 1.75 Å.

In contrast to the room-temperature structure determined at neutral pH, the position of the side chain of Tyr3 remains well defined at acidic conditions. Both its position and relative orientation with respect to Trp6 is the same as determined for the neutral-pH structure at low temperature (T = 282K). Furthermore, there are two crucial long-range NOEs connecting Tyr3 and Pro18, indicating spatial connection of the α-helix and the PPII region. However, the stability of the central region with β-turns is destabilized at pH = 3.2 (Figure 6). Since the side chain pK_a of Asp9 is around 3.7, the increased flexibility at pH = 3 is due to its protonation, weakening the Asp9-Arg16 salt bridge. This is signaled by the elevated local rmsd values of Asp9, namely 1.49 Å for the backbone and 2.22 Å for all heavy atoms.

2.4. TC5b at Ambient Temperature (T = 300 K) and at Acidic pH (3.12). At room temperature (T = 300 K) under acidic conditions (pH = 3.12), the overall fold is weakened significantly since only 99 NOEs (34 interresidual) could be measured. The extremely low number of detected distance restraints indicates that the overall fold of TC5b is very flexible, experiencing a multitude of different subconformers in a time average manner. As a consequence, NOEs can no longer be assigned neither between Tyr3 and Pro18 nor between Tyr3 and Pro19, indicating the loss of close contact motion between the residual α-helical part and the PPII secondary structure elements.

3. Structure and Characteristics of TC6a at Neutral pH (6.50) and Low Temperature (T = 282 K). The structure of TC6a was solved by using a moderate number of interresidual NOEs (a total of 62), resulting in a structure ensemble in which the "all heavy atom" and backbone rmsd values are 1.77 Å and 1.24 Å, respectively. The comprehensive structure analysis of TC5b has revealed that the salt bridge of Asp9-Arg16 is crucial in holding the structure of the miniprotein together. The mutation of Asp to Asn, carrying out a "negative-design", makes the formation of the salt bridge impossible even at neutral pH. Interestingly, mutation at position 9 also affects the orientation of the side chain of Tyr3. A somewhat altered location of the aromatic ring with respect to that of TC5b is depicted. In TC6a the aromatic ring of Tyr3 is turned outward from the core and it is placed beside, rather than above, the indole ring of Trp6. The tilt angle between the planes of Tyr3 and Trp6 is around 75°. Thus, the relative orientation of the two aromatic systems becomes almost perpendicular, destabilizing the overall fold and, in particular, the C-terminus of the α-helix (Table S1). Furthermore, although the Asp9-Arg16 salt bridge is eliminated, the central β-turn-containing segment of the molecule maintains its unique fold. The biased structural behavior of the polyproline II region can be explained by the increased mobility of Arg16 in the absence of the salt bridge, which has a role in orienting the PPII secondary structure.

The NMR structure elucidation indicates that Tc6a has a less compact structure than TC5b, which is in perfect agreement with the global structural data obtained from CD curve decomposition. The initially lower helical content drops with the increase of temperature most likely because of the misplacement of Tyr3. The larger relative ratio of PPII remains around 35% while the RND contribution increases from 30 to 65% as the temperature increases from 40 to 85 °C. (Figure 3C)

4. Structure and Characteristics of TC5b-5 at Neutral pH and Low Temperature. We have pointed out so far that Tyr3 seems crucial in stabilizing Trp6, the chief moiety of the hydrophobic center. A variant, TC5b-5 was synthesized with the omission of the first five amino acid residues of the N-terminus (Table 1). As the near-UV CD spectra show no changes with increasing temperature **(**Figure 1D), we conclude that the hydrophobic core does not exist for TC5b-5 even at low temperature (T = 277 K). However, CD data analysis indicates a high content of PPII conformation at lower temperatures, which vanishes at elevated temperatures. This can be possible by the intrinsic structural preferences of the core PPII segment, Pro17-Pro19, promoting PPII conformation even in the absence of the interacting α-helix. Although NMR spectra were recorded for TC5b-5, they lack most of the long-range NOEs consistent with a conformation with high PPII content. This latter property enables, however, to quantitate the PPII content of the longer variants.

5. Structure and Characteristics of TC6as at Neutral pH (6.58) and Low Temperature (T = 282 K). Glycosylation of proteins has a widespread biological importance including dipole−dipole structure stabilization. We intended to test whether a carbohydrate moiety has the capacity of forming a suitable dipole−dipole interaction sufficient for substantial stabilization of the Trp-cage fold. The spectral analysis of TC6as reveals a moderate number of interresidual NOEs (89), resulting in a less defined structure ensemble where the "all heavy atom" and backbone rmsd values were 2.11 Å and 0.96 Å, respectively. (Unambiguous assignment of the carbohydrate moiety could not be achieved; thus, its relative orientation with respect to the protein remains unclear.) The main structural features of TC6as are similar to that of the "acidic" TC5b. CD spectra deconvolution (Figure 3C) shows that apart from the highly flexible central region, the first eight amino acid residues (40% of the total 20) forming the α-helix seem intact at low temperature, as found also by NMR (Table S1). The elevation of the temperature unscrews the α-helix as well as the PPII structure (Figure 3C) of the molecule, resulting in an almost totally unfolded polypeptide chain at 85 °C. In conclusion, the 3D structure of the central region of TC6as is blurred; thus, we can conclude that the attached N-galactosyl moiety cannot sufficiently stabilize the overall fold of the miniprotein.

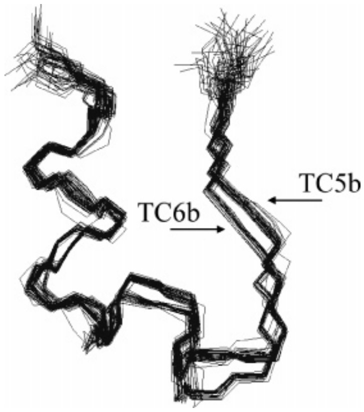

FIGURE 7: The backbone trace of TC6b and TC5b demonstrating a subtle but significant fold alteration.

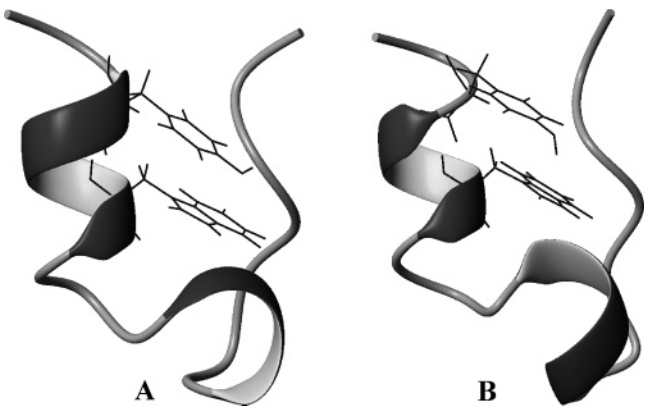

FIGURE 8: Representative conformer of TC6b at 282 K (A) and at 300 K (B). It is apparent that the two structures are similar.

6. Structure and Characteristics of TC6b. 6.1. TC6b at Low Temperature (T = 282 K) and Neutral pH (6.61). It was shown that for the stability of the present Trp-cage miniprotein, the formation of the Asp9-Arg16 salt bridge is essential in addition to a compact hydrophobic core formed by the aromatic side chains of Trp6 and Tyr3 residues. Molecular modeling and structure analysis suggest that the structure organizing effect of the salt bridge can be improved by "lengthening" the side chain of residue 9. The insertion of methylene group, mutation of Asp9 to Glu (Table 1), could allow a greater conformational flexibility and perhaps fine-tune the interaction. If so, this might also have a long-range effect, and the interaction of the prolines with the aromatic rings could be improved by forming an even tighter hydrophobic core to further stabilize the overall 3D-fold of the miniprotein. The structural elucidation of TC6b (T = 282 K, pH = 6.61) has revealed that, according to our expectations, its molecular fold is more compact than that of TC5b. This is indicated by the larger number of interresidual NOEs (251) assigned with respect to those of the 197 in TC5b. When the coherence of the structural ensembles is analyzed, the rmsd values are slightly lower for TC6b than for TC5b: the backbone rmsd value is 0.62 Å (0.64 Å for TC5b) and "all heavy atom" rmsd value is 1.14 Å (1.20 Å for TC5b).

Both regions that are crucial for the stability of the overall structure are well determined. Namely the Tyr3 and the Trp6 interaction as well as the central β-turn region of the fold are as stable as any other part of the miniprotein. However, a uniform stability of the secondary structure elements was also noticed for TC5b. Therefore, the extra stability measured by the increased number of interresidual NOEs should come from elsewhere. The atomic level analysis has revealed a subtle but significant reorientation of the side chain of Tyr3: the angle between the two aromatic moieties is around 36° in TC5b and 28° in TC6b. The Tyr3-Trp6 interaction is reshuffled, resulting in a tighter hydrophobic core for TC6b. This conformational change is coupled with a slight shift of the polyproline II region made possible by the optimized arm of the salt bridge. The lengthening of the side chain has two consequences. First, the side chain of residue 9 in TC6b (Glu) is longer than that of TC5b (Asp) which participate also in the stabilization of the backside of the hydrophobic cage. The C^{γ} atom of Glu9 is just 2.86 Å away from the C^{δ} of Trp6. This type of hydrophobic interaction is present at the same time when the COO^{-} group of the same residue anchors to the guanidino group of Arg16. Note that in TC5b the side chain of Asp9 turns outward from the cage, avoiding any hydrophobic contacts with Trp6. Second, the overall salt bridge is elongated, allowing an alternative position for the PPII segment. This provides more space for the phenyl ring of Tyr3 to be positioned more in-plane relative to the indole ring of Trp6. The rearrangement of hydrophobic contacts is shown by the different NOE patterns of TC5b and TC6b at neutral pH and T = 282 K. The NOE pattern analysis of Trp6 shows that it is in the spatial vicinity of residues 16−18, as well as those of 12, 14 and 19, respectively. More importantly, the side chain of Tyr3 has NOEs to Pro19 in TC6b, unlike in TC5b. The rearrangement of the polyproline II region (Figure 7) thus contributes to the different orientations of the phenyl ring of Tyr3.

6.2. TC6b at Ambient Temperature (T = 300 K) and Neutral pH (6.61). Thermal destabilization also decreases the number of interresidual NOEs for TC6b: at T = 300 K the number of interresidual NOEs is 149. The backbone rmsd value of TC6b at 300 K (0.83 Å) is lower than that of TC5b at the same conditions (1.16 Å) and only just slightly higher than that of TC5b at 282 K (0.64 Å). From the point of view of structure refinement the structure of TC6b at 300 K is as good as that of TC5b at 282 K (Figure 8) because of the favorable NMR-signal dispersion and the high quality of the restraint set.

As a response to the increased thermal motion, no significant increase in structure mobility or local unwinding of structural elements is observed for TC6b at 300 K. Most structural features are similar to those of the parent "low" temperature structure. Both the side-chain packing of Tyr3 and the stability of the central region remains similar and well determined (Table S1). Even at elevated temperatures the angle between the aromatic planes, that is the alignment

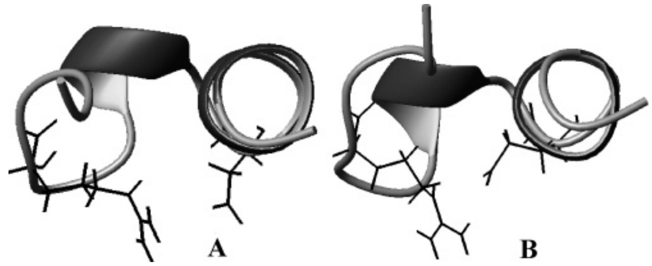

FIGURE 9: Structures of TC5b (A) and TC6b (B) with respect to the crucial salt bridge formed between residues 9 and 16.

Table 2: Selected Structural Refinement Parameters of Tc5b and Its Variants

model	T (K)	pH	no. of interresidual NOEs	all NOEs	backbone rmsd	all heavy atom rmsd	Ramachandran map analysis (Procheck)			
							core	allow	generously allowed	disallowed
1L2Y	282	7.0	115	169	0.81	1.52	–	–	–	–
TC5b	282	6.90	197	260	0.64	1.20	78.6	15.9	3.4	2.0
		3.17	67	148	1.2	1.75	81.8	9.1	9.1	0.0
	300	6.52	90	152	1.16	1.78	62.7	12.7	14.5	10.0
		3.12	34	99	1.64	2.40	43.6	40.0	14.5	1.8
TC6a	282	6.5	62	150	1.24	1.77	81.8	18.2	0.0	0.0
TC6as	282	6.58	89	187	0.96	2.11	85.2	14.1	0.4	0.4
TC6b	282	6.61	251	324	0.62	1.14	76.5	19.7	3.8	0.0
	300	6.61	149	210	0.83	1.42	76.6	15.6	6.3	1.5

of Tyr3 with respect to Trp6, is preserved at approximately 28°. Similarly to that of the "low" temperature structure, the role of the side chain of Glu9 at higher temperatures (T = 300 K and above) is twofold: (i) the methylene group participates in the formation of the backside of the hydrophobic cage, and (ii) the carboxylic end forms the salt bridge with Arg16. The attractive force operative between the negatively charged acidic group of Glu9 and the positively charged guanidino group of Arg16 provides an optimal distance between the two sides of the molecule, becoming a key element for the increased compactness and for the improved heat resistance (Figure 9). However, two minor structural changes can also be identified at elevated temperature. First, Trp6 is slightly shifted toward the "other facade" of the molecule as the temperature increases. At higher temperature (300 K) the distance between the indole ring and the side-chain atoms of Pro19 decreases, indicating a shift between the two sides of the molecule. This occurs in parallel with the slight increase of the distance between Tyr3 and Pro19, signaling a gentle rearrangement of the hydrophobic core. Second, the first residue of the miniprotein only has a limited number of NOE contacts within the α-helix. Thus, the N-terminal "end" of the higher temperature structure seems more flexible, which is perfectly supported by the CD analysis (Figure 3C). In conclusion, apart from the minor conformational differences described above, the main structural features of TC6b are nearly identical at low and higher temperatures, which is clearly not the case for TC5b. Detailed structural data are presented in Table 2 and in Supporting Information tables. All experimental constraints are fulfilled.

Analyzing the ratios of pure CD curves related to Tc6b as a function of the temperature shows (Figure 3C) that the helix/turn content decreases much more slowly than in the case of Tc5b. Additionally, the PPII content of Tc6b is markedly higher, indicating a more stable polyproline II core region than found in TC5b. Finally, the random coil content of Tc6b is significantly lower than it is in Tc5b. This latter feature may be explained by a more stable central region, also identified by NMR and by a more rigid PPII C terminus. In conclusion, both CD and NMR data support that TC6b is more compact and more thermoresistant than the parent TC5b.

CONCLUSION

It has been previously shown (*17*) that the Trp-cage miniprotein TC5b adopts a protein-like 3D-fold at 280 K. The presented comprehensive structural analysis has demonstrated the extreme sensitivity of its structure toward both pH and temperature changes. By deciphering the key effect in the stabilization of the Trp-cage fold, our goal was to design a more compact and thermoresistant mutant. To do so, we have revisited the folding properties of the original molecule and designed, synthesized, and determined the 3D structure of different mutants and variants of the Trp-cage miniprotein. We have uncovered important molecular features of this miniprotein, strongly contributing to its overall stability, namely:

(i) the significance of the aromatic residues Tyr3 and Trp6 in forming the core of the protein,

(ii) the importance of the relative spatial orientation of the aromatic rings of both Tyr3 and Trp6, showing that an improved arrangement is achieved if the two rings are above each other and have a smaller tilt angle,

(iii) the impact of the salt bridge formed between residues 9 (Asp or Glu) and 16 (Arg),

(iv) the favorable contact between the hydrophobic backside of the side chain of Glu9 and the hydrophobic core.

Deciphering the interaction network led us to the rational design of TC6b with the introduction of a single methylene group at a selected position. The fine-tuning of the length of the crucial salt bridge between residue 9 and 16 have resulted in a clearly more compact and thermostable mutant. In the latter molecule even at higher temperature, minor rearrangements were observed only involving the polyproline II segment, the α-helix of the N-terminus, and the hydrophobic core of the protein. Unlike for TC5b, CD data show that even at higher temperature (85 °C) a noticeable fraction (>30%) of both the α-helix and the PPII secondary structures remains. In summary, we have shown the existence of multiple secondary structure elements (e.g., α-helix, PPII) as well as their interaction in this rationally modified mutant. We have found that most of the side chains in TC6b adopt a unique conformation, especially those of the core region, e.g., Tyr3 and Trp6. We have improved some of the folding properties of the miniprotein, now clearly demonstrating a greater compactness and an improved thermal resistance, features typical of globular proteins.

ACKNOWLEDGMENT

The authors appreciate the help of Suzanne K Lau. Dedicated to Professor Imre G. Csizmadia on the occasion of his 75th birthday.

SUPPORTING INFORMATION AVAILABLE

Average backbone parameters, PROCHECK-NMR results, and rmsd values of different TC5b mutants at various conditions. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Dobson, C. M., and Karplus, M. (1999) The fundamentals of protein folding: bringing together theory and experiment, *Curr. Opin. Struct. Biol. 9*, 92−101.
2. Markossian, K. A., and Kurganov, B. I. (2004) Protein folding, misfolding and aggreration. Formation of inclusion bodies and aggresomes, *Biochemistry (Moscow) 69*, 9, 971−984.
3. Dobson, C. M. (2004) Experimental investigation of protein folding and misfolding, *Methods 34*, 1, 4−14.
4. Swope, W. C., Pitera, J. W., and Suits, F. (2004) Describing protein folding kinetics by molecular dynamics simulations. 1. Theory, *J. Phys. Chem. B 108*, 21, 6571−6581.
5. Pitera, J. W., and Swope, W. (2003) Understanding folding and design: Replica-exchange simulations of "Trp-cage" miniproteins, *Proc. Nat. Acad. Sci. U.S.A. 100* (13), 7587−7592.
6. Barua, B., Andersen, N. H. (2002) Determinants of miniprotein stability: Can anything replace a buried H-bonded Trp sidechain? *Lett. Pept. Sci. 8* (3−5), 221−226.
7. Vita, C., Vizzavona, J., Drakopoulou, E., Zinn-Justin, S., Gilquin, B., and Ménez, A. (1998) Novel Miniproteins Engineered by the Transfer of Active Sites to Small Natural Scaffolds, *Biopolymers 47* (1), 93−100.
8. Bunagan, M. R., Yang, X., Saven, J. G., and Gai, F. (2006) Ultrafast folding of a computationally designed trp-cage mutant: Trp 2-cage, *J. Phys Chem. B 110* (8), 3759−3763.
9. Qiu, L., Pabit, S. A., Roitberg, A. E., and Hagen, S. J. (2002) Smaller and faster: The 20-residue Trp-cage protein folds in 4 μs, *J. Am. Chem. Soc. 124* (44), 12952−12953.
10. Chowdhury, S., Lee, M. C., Xiong, G., and Duan, Y. (2003) Ab initio folding simulation of the Trp-cage mini-protein approaches NMR resolution, *J. Mol. Biol. 327* (3), 711−717.
11. Naduthambi, D., Zondlo, N. J. (2006) Stereoelectronic tuning of the structure and stability of the trp cage miniprotein. *J. Am. Chem. Soc. 128* (38), 12430−1.
12. Ágoston. V., Cemazar, M., Kaján, L., and Pongor, S. (2005) Graph-representation of oxidative folding pathways. *BMC Bioinformatics 6* (1), 19.
13. Gáspári, Z., Vlahovicek, K., and Pongor, S. (2005) Efficient recognition of folds in protein 3D structures by the improved PRIDE algorithm. *Bioinformatics 21* (15), 3322−3323.
14. Patthy, L (1993) Modular design of proteases of coagulation, fibrinolysis, and complement activation: Implications for protein engineering and structure - function studies, *Methods Enzymol. 222*, 10−21.
15. Kortemme, T., Ramírez-Alvarado, M., and Serrano, L. (1998) Design of a 20-amino acid, three-stranded β-sheet, *Protein Sci. 281* (5374), 253−256.
16. Bork, P., and Sudol, M. (1994) The WW domain: A signalling site in dystrophin? *Trends Biochem. Sci. 19* (12), 531−533.
17. Neidigh, J. W., Fesinmeyer, R. M., and Andersen, N. H. (2002) Designing a 20-residue protein, *Nat. Struct. Biol. 9* (6), 425−430.
18. Rhee, Y. M., Pande, V. S. (2003) Multiplexed-replica exchange molecular dynamics method for protein folding simulation, *Biophys. J. 84* (2), 775−786.
19. Neuweiler, H., Doose, S., and Sauer, M. (2005) A microscopic view of miniprotein folding: Enhanced folding efficiency through formation of an intermediate, *Proc. Nat. Acad. Sci. U.S.A. 102* (46), 16650−16655.
20. Ahmed, Z., Beta, I. A., Mikhonin, A. V., and Asher, S. A. (2005) UV-resonance Raman thermal unfolding study of Trp-cage shows that it is not a simple two-state miniprotein, *J. Am. Chem. Soc. 127* (31), 10943−10950.
21. Juraszek, J., Bolhuis, P. G. (2006) Sampling the multiple folding mechanisms of Trp-cage in explicit solvent, *Proc. Nat. Acad. Sci. U.S.A. 103* (43) 15859−15864.
22. Bosshard, H. R., Marti, D. N., and Jelesarov, I. (2004) Protein stabilization by salt bridges: Concepts, experimental approaches and clarification of some misunderstandings, *J. Mol. Recognit. 17* (1), 1−16.
23. Kumar, S., and Nussinov, R. (2002) Close-range electrostatic interactions in proteins, *ChemBioChem 3* (7), 604−617.
24. Solá, R. J., and Griebenow, K. (2006) Chemical glycosylation: New insights on the interrelation between protein structural mobility, thermodynamic stability, and catalysis, *FEBS Lett. 580* (6), 1685−1690.
25. Csonka, G. I., Schubert, G. A., Perczel, A., Sosa, C. P., and Csizmadia, I. G. (2002) Ab initio conformational space study of model compounds of O-glycosides of serine diamide, *Chemistry - A Eur. J. 8* (20), 4718−4733.
26. Láng, E., Hargittai, B., Majer, Z., Perczel, A., Mák, M., Kajtár-Peredy, M., Radics, L., Fasman, G. D., and Hollósi, M. (1996) *Protein Pept. Lett. 3* (1), 9−16.
27. Perczel, A., Kollát, E., Hollósi, M., and Fasman, G. D. (1993) Synthesis and conformational-analysis of n-glycopeptides. 2. Cd, molecular-dynamics, and nmr spectroscopic studies on linear n-glycopeptides, *Biopolymers 33* (4), 665−685.
28. Ghirlando, R., Lund, J., Goodall, M., and Jefferis, R. (1999) Glycosylation of human IgG-Fc: influences on structure revealed by differential scanning micro-calorimetry, *Immunol. Lett. 68*, 47−52.
29. Lu, S., Deng, P., Liu, X., Luo, J., Han, R., Gu, X., Liangi, S., Wangparallel, X., Liparallel, F., Lozanov, V., Patthy, A., and Pongor, S. (1999) Solution structure of the major α-amylase inhibitor of the crop plant amaranth, *J. Biol. Chem. 274* (29), 20473−20478.
30. Jackson, P. J., McNulty, J. C., Yang, Y.-K., Thompson, D. A., Chai, B., Gantz, I., Barsh, G. S., and Millhauser, G. L. (2002) Design, pharmacology, and NMR structure of a minimized cystine knot with agouti-related protein activity, *Biochemistry 41* (24), 7565−7572.
31. Ding, F., Buldyrev, S. V., and Dokholyan, N. V. (2005) Folding Trp-cage to NMR resolution native structure using a coarse-grained protein model, *Biophys. J. 88* (1), 147−155.
32. Zhou, R. (2003) Trp-cage: Folding free energy landscape in explicit water, *Proc. Nat. Acad. Sci. U.S.A. 100* (23), 13280−13285.
33. Nikiforovich, G. V., Andersen, N. H., Fesinmeyer, R. M., and Frieden, C. (2003) Possible locally driven folding pathways of TC5b, a 20-residue protein, *Proteins: Struct. Funct. Genet. 52* (2), 292−302.
34. Seshasayee, A. S. N. (2005) High-temperature unfolding of a Trp-cage mini-protein: A molecular dynamics simulation study, *Theor. Biol. Med. Modelling 2*, 7−11.
35. Ötvös, L., Jr., Ürge, L., Hollósi, M., Wroblewski, K., Graczyk, G., Fasman, G. D., and Thurin, J. (1990) Automated solid-phase synthesis of glycopeptides. Incorporation of unprotected mono- and disaccharide units of N-glycoprotein antennae into T cell epitopic peptides, *Tetrahedron Lett. 31* (41), 5889−5892.
36. Perczel, A., Hollósi, M., Tusnády, G., and Fasman, G. D. (1991) Convex constraint analysis - a natural deconvolution of circular-dichroism curves of proteins, *Protein Eng. 4*, 669−679.
37. Perczel, A., Park, K., and Fasman, G. D. (1992) Analysis of the circular-dichroism spectrum of proteins using the convex constraint algorithm - a practical guide, *Anal. Biochem. 203*, 83−93.
38. http://www.chem.elte.hu/departments/protnmr/cca/.
39. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR 6*, 277−293.
40. Bartels, C., Xia, T.-H., Billeter, M., Güntert, P., and Wüthrich, K. (1995) The program XEASY for computer-supported NMR spectral analysis of biological macromolecules, *J. Biomol. NMR 5*, 1−10.
41. Nilges, M., Clore, G. M., and Gronenborg, A. M. (1988) Determination of three dimensional structures of proteins from interproton distance data by chemical simulated annealing from a random array of atoms, *FEBS Lett. 239*, 129−136.
42. Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR, *J. Biomol. NMR 8* (4), 477−486.
43. Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures, *J. Mol. Graph. 14* (1), 51−55.
44. Woody, R.W. (1974) Studies of theoretical circular dichroism of polypeptides: Contributions of β-turns, In *Peptides, Polypeptides and Proteins* (Blout, E. R., Bovey, F. A., Goodman, M., and Lotan, N., Eds.) pp 338−360, John Wiley & Sons, New York.

45. Woody, R. W. (1985) Circular dichroism of peptides, In *Peptides* (Hruby, V. J., Ed.) pp 15−113, Vol. 7, Academic Press, New York.
46. Hollósi, M., Ötvös, L., Jr., Ürge, L., Kajtár, J., Perczel, A., Laczkó, I., Vadász, Z., and Fasman, G. D. (1993) Ca^{2+}-induced conformational transitions of phosphorylated peptides, *Biopolymers 33*, 497−510.
47. Hu, Hao, Clarkson, Michael W., Hermans, Jan, and Lee, Andrew L. (2003) Increased Rigidity of Eglin c at Acidic pH: Evidence from NMR Spin Relaxation and MD Simulations, *Biochemistry 42* (47), 13856−13868.
48. Kasimova, M. R., Kristensen, S. M., Howe, P. W. A., Christensen, T., Matthiesen, F., Petersen, J., Sørensen, H. H., and Led, J. J. (2002) NMR studies of the backbone flexibility and structure of human growth hormone: A comparison of high and low pH conformations, *J. Mol. Biol. 318* (3), 679−695.
49. Kahmann, J. D., O'Brien, R., Werner, J. M., Heinegard, D., Ladbury, J. E., Campbell, I. D., and Day, A. J. (2000) Localization and characterization of the hyaluronan-binding site on the Link module from human TSG-6, *Structure 8* (7), 763−774.

BI701371X