CHEMISTRY
A EUROPEAN JOURNAL

DOI: 10.1002/chem.200500458

A Designed Well-Folded Monomeric Four-Helix Bundle Protein Prepared by Fmoc Solid-Phase Peptide Synthesis and Native Chemical Ligation**

Gunnar T. Dolphin*[a]

Abstract: The design and total chemical synthesis of a monomeric native-like four-helix bundle protein is presented. The designed protein, GTD-Lig, consists of 90 amino acids and is based on the dimeric structure of the de novo designed helix-loop-helix GTD-43. GTD-Lig was prepared by the native chemical ligation strategy and the fragments (45 residues long) were synthesized by applying standard fluorenylmethoxycarbonyl (Fmoc) chemistry. The required peptide–thioester fragment was prepared by anchoring the free γ-carboxy group of Fmoc-Glu-allyl to the solid phase. After chain elongation the allyl moiety was orthogonally removed and the resulting carboxy group was functionalized with a glycine–thioester followed by standard trifluoroacetic acid (TFA) cleavage to produce the unprotected peptide–thioester. The structure of the synthetic protein was examined by far- and near-UV circular dichroism (CD), sedimentation equilibrium ultracentrifugation, and NMR and fluorescence spectroscopy. The spectroscopic methods show a highly helical and native-like monomeric protein consistent with the design. Heat-induced unfolding was studied by tryptophan absorbance and far-UV CD. The thermal unfolding of GTD-Lig occurs in two steps; a cooperative transition from the native state to an intermediate state and thereafter by noncooperative melting to the unfolded state. The intermediate exhibits the properties of a molten globule such as a retained native secondary structure and a compact hydrophobic core. The thermodynamics of GuHCl-induced unfolding were evaluated by far-UV CD monitoring and the unfolding exhibited a cooperative transition that is well-fitted by a two-state mechanism from the native to the unfolded state. GTD-Lig clearly shows the characteristics of a native protein with a well-defined structure and typical unfolding transitions. The design and synthesis presented herein is of general applicability for the construction of large monomeric proteins.

Keywords: helical structures · native chemical ligation · peptides · protein design · protein folding

Introduction

The de novo design and total synthesis of monomeric globular proteins with linear sequences in excess of 50 amino acid residues is a considerable challenge. Adding to this challenge is the design of proteins that fold into well-defined tertiary structures. In de novo design research, a synthetic approach is required as it offers fast access to the designed protein for structure and stability analysis. So far, only two monomeric globular proteins have been synthesized that have a native-like structure.[1,2] These are three-helix bundles and are constructed of 65 and 71 residues. Other designed monomeric proteins with well-defined folds consist of short amino acid sequences,[3–5] which is a challenge in itself. Recent advances in computer-assisted design[6] have yielded well-folded structures[7–9] as well as monomeric proteins of more than 90 residues that are prepared by protein expression.[10,11] For the preparation of larger monomeric structures, protein designers have employed methods that involve protein assembly on templates.[12–15] A number of designed well-folded proteins have been reported that have a total of more than 50 residues. They are constructed though of aggregates of secondary-structure-containing building-blocks such as α-helices,[8,16] β-strands,[17] and the helix-loop-helix motif.[18,19] This strategy for designing oligomeric aggregates is interesting as it offers a simple way to prepare tertiary structures. However, there is a need to prepare monomeric globular proteins for folding analysis and for the engineering of unique binding sites as this is a common motif in nature.

[a] Dr. G. T. Dolphin
LEDSS 5, Université Joseph Fourier
BP 53, 38041 Grenoble Cedex 9 (France)
Fax: (+33) 476-514-946
E-mail: gunnar.dolphin@ujf-grenoble.fr
gunnar_dolphin@hotmail.com

[**] Fmoc = fluorenylmethoxycarbonyl.

Supporting information for this article is available on the WWW under http://www.chemeurj.org/ or from the author.

De novo protein design offers the ultimate test for our understanding of protein folding and structure, which has practical applications in the engineering of, for example, biosensors,[20] novel catalysts,[21] immunogenic folds,[22] pharmaceutical applications,[23] and cofactor-linked proteins.[24,9] Previously, the majority of de novo designed proteins were best described as molten globules due to elusive design principles. Today much has been learned from de novo design which now enables the construction of peptides and proteins with well-defined structures.[25–27] It is important to consider the fact that shape-complementarity between the hydrophobic interfaces of amphiphilic α-helices and β-sheets is not enough to achieve well-defined tertiary structures and that the incorporation of conformational constraints is necessary.[28,29]

Previously the author has reported with collaborators on the de novo design and synthesis of GTD-43, a 43-residue peptide that dimerizes to form a four-helix bundle.[29,30] The related and more soluble GTD-C has also been reported previously.[31] These structures have been shown to fold into well-defined four-helix bundles by circular dichroism (CD), sedimentation equilibrium ultracentrifugation, and NMR spectroscopy. The well-folded nature of the structures was attributed to the designed aromatic ensemble in the central hydrophobic core, designed to form a conformational constraint and restrict the dynamics of the hydrophobic residues in the folded structure.[30] In addition, GTD-C showed that there was increased specificity other than the contribution from the aromatic ensemble, which was assigned to the interhelical salt bridge between His17 and Asp27 located at the hydrophobic core-loop boundary.[31]

The de novo design and structural characterization of GTD-Lig, a monomeric four-helix bundle protein, is reported herein. The design is based on the folded dimeric structure of GTD-43. The sequence of GTD-43 was redesigned in order to connect all four helices together through a rearrangement of the loops. This new monomeric four-helix bundle is folded from a single polypeptide chain containing 90 amino acid residues. In the work described herein the modifications required to transform a well-folded dimeric four-helix bundle protein into the corresponding monomeric four-helix bundle were investigated.

Also presented in this report is the total synthesis of the designed monomeric globular protein, prepared through the use of the native chemical ligation strategy (NCL).[32]

A major advance in protein chemistry has been the development of chemical ligation methods by which two unprotected peptide fragments are chemoselectively ligated together. These methods have drastically increased the size of synthetic proteins and allowed their redesign.[33,34]

Up to now, the syntheses of the required C-terminal peptide thioesters have been performed mainly by the Boc-protection strategy (Boc; *tert*-butyloxycarbonyl) of solid-phase peptide synthesis (SPPS). The disadvantage of the Boc strategy is the repeated use of strong acid for α-amino deprotection and final HF cleavage. Other disadvantages are encountered when the peptide segment is modified with HF sensitive glycosidic linkages or fluorescent probes. For these reasons the Fmoc-protection strategy (Fmoc; fluorenylmethoxycarbonyl) is preferred for its milder conditions. However there is one problem to overcome with the Fmoc strategy, which is the susceptibility of the thioester linkage towards nucleophilic piperidine which is used for Fmoc removal. In recent years, several strategies have been reported that overcome this problem.[35] The published methods however either require extensive optimization, special agents/linkers, or are reported to cause undesirable side-reactions.

A simple alternative to previously published methods of peptide-thioester preparation with Fmoc SPPS is the side-chain anchoring of an orthogonally protected amino acid on to the resin. Tulla-Puche and Barany have reported on the synthesis of a peptide-thioester by side-chain anchoring of Fmoc-Asp-Allyl for on-resin ligation.[36] Here, as the thioester is cleaved from the resin and purified before ligation, Fmoc-Glu-Allyl is used in combination with a peptide-amide resin that forms a glutamine upon cleavage to avoid possible side-reactions such as ester migration.[37] The allyl is an orthogonal protecting group on the α-carboxy group, which is simply and selectively removed after completion of the sequential solid-phase peptide synthesis. Glycine-benzyl thioester is then coupled to the resulting free α-carboxy group using standard coupling reagents;[38] thereafter, the desired peptide-thioester is released from the resin through standard TFA-cleavage procedures.

The monomeric four-helix bundle protein presented herein has been synthesized in high yield with the aid of NCL and studied by a number of biophysical analytical methods to determine its structure and stability. The results show that the structure is highly helical and that it folds into a native-like monomeric structure. These results have implications for the de novo design and total synthesis of large monomeric proteins with applications, for example, in the research of membrane channels, metalloproteins, and the binding of cofactors. The preparation of GTD-Lig, which contains 90 residues, demonstrates that the presented strategy for peptide-thioester synthesis in combination with native chemical ligation provides an easy access to large peptides/proteins. The results also show that the redesign of de novo designed and native proteins is possible as the rearrangement of the loops in the four-helix bundle had little effect on the tertiary fold.

Results

The design of GTD-Lig: The monomeric four-helix bundle protein GTD-Lig was designed based on GTD-43 and GTD-C. These 43-residue polypeptides fold into helix-loop-helix motifs that dimerize to form antiparallel four-helix bundles in aqueous solution. The de novo design procedure of GTD-43 and GTD-C, which has previously been described in detail, involves the design of an aromatic ensemble, Phe10 and Trp13 in helix 1 and Phe34 in helix 2, that induces unique conformational stability.[30] The conformational

stability of GTD-43 is characterized by a well-dispersed ^{1}H NMR spectrum that exhibits slow exchange on the NMR timescale, slow amide proton exchange rates in part of the sequence, and a narrow temperature range for thermal denaturation.

For the design of the monomeric four-helix bundle, it was necessary for the loops to undergo rearrangement to avoid a loop stretching the length of the helical structure. To do this, the loop between helices 1 and 2 was removed and two new loops between helices 1 and 1′ and between helices 2′ and 2 were designed (Figure 1). The loop between helices 1′ and 2′ was unchanged. The new loops were designed to have the same sequence as in the original helix-loop-helix, namely, Gly-Thr-Gly-Pro.

The site for native chemical ligation was designed to be in the region of the loop in the center of the amino acid sequence. This has the advantage of producing ligating fragments of similar length and the introduction of the cysteine necessary for ligation has helix-destabilizing properties. Cysteine replaces the threonine in the loop as this leads to only small changes in the structural features of the loop, which now becomes Gly-Cys-Gly-Pro. The two peptide fragments to be ligated together are designated GTD-(1,1′)-SBzl and GTD-(2′,2), where (1,1′) and (2′,2) indicate the individual helices in the four-helix bundle and SBzl is the C-terminal benzyl thioester (Figure 2).

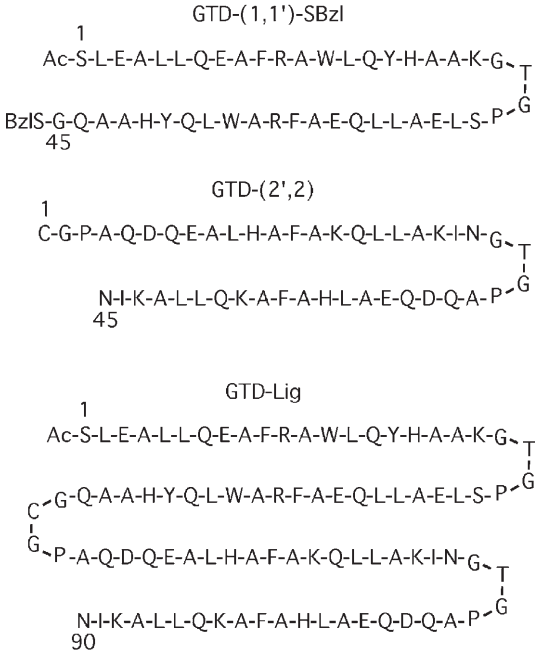


Figure 2. Amino acid sequences of the fragments GTD-(1,1′)-SBzl and GTD-(2′,2), and the ligated protein GTD-Lig. A one-letter code is used for the amino acid residues. The helices are connected by the loop sequences GTGP and GCGP. BzlS is the benzyl thioester.

Synthesis of GTD-Lig: The required peptide fragment with a C-terminal glycine-thioester [GTD-(1,1′)-SBzl] was designed to be synthesized by the Fmoc strategy. This was possible and easily achieved by using the α-carboxy orthogonal-protected Fmoc-Glu-α-Allyl in a two-dimensional solid-phase synthesis (Scheme 1). This amino acid is coupled to the solid-phase resin through its side-chain and allows, first, the Fmoc group to be used for sequential peptide synthesis. Secondly, after SPPS the allyl group can be selectively removed to give the α-carboxy group, which is then functionalized with the appropriate amino acid-thioester.

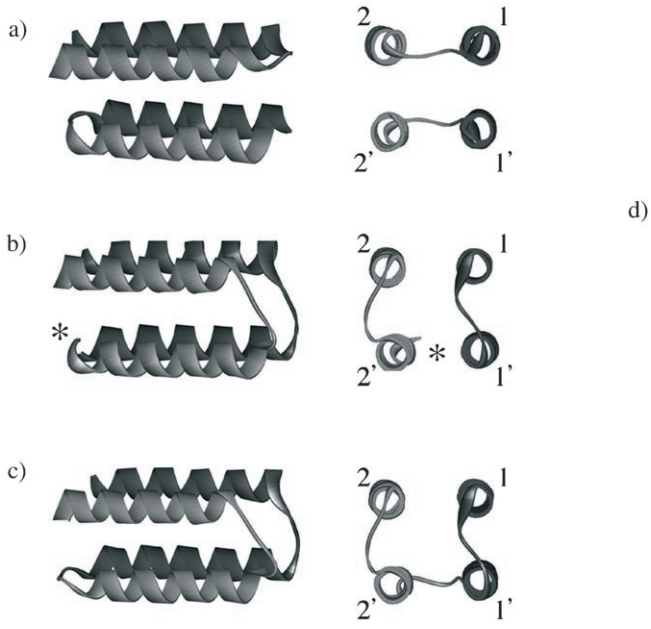
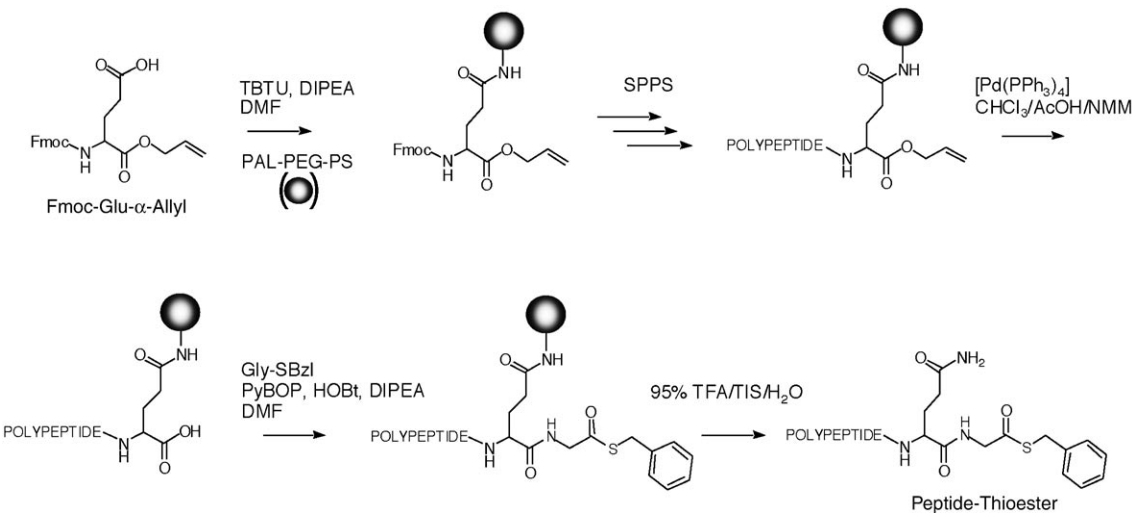


Figure 1. Schematic representation of antiparallel and dimeric GTD-43 and its redesign to monomeric GTD-Lig. a) Dimer of GTD-43. The helices are denoted 1 and 2 in the top monomer and 1′ and 2′ in the bottom monomer. b) Intermediate fragments GTD-(1,1′)-SBzl and GTD-(2′,2). The * indicates the location of the ligation reaction. c) Modeled representation of GTD-Lig with the individual helices having the same relative orientation to one another as those in the dimeric structure of GTD-43. d) Schematic representation of GTD-Lig showing the hydrophobic amino acids phenylalanine and tryptophan. The two aromatic ensembles originate from the designed aromatic ensemble of GTD-43.

To synthesize the 45-residue peptide benzyl thioester, GTD-(1,1′)-SBzl, amino acids were coupled by TBTU/DIPEA activation starting with the first residue Fmoc-Glu-α-Allyl. Thereafter the allyl group was selectively removed by treatment with palladium(0) for 3 h at room temperature. Glycine-benzyl thioester was then coupled to the resulting free α-carboxy moiety by PyBop/HOBt/DIPEA activation. The desired



Scheme 1. Fmoc strategy for the preparation of peptide–thioesters starting from Fmoc-Glu-α-Allyl. TBTU = 2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium tetrafluoroborate, DIPEA = diisopropylethylamine, PyBOP = benzotriazole-1-yl-oxy-tris(pyrrolidino)phosphonium hexafluorophosphate, HOBt = 1-hydroxybenzotriazide, TIS = triisopropylsilane.

peptide thioester was released from the resin through normal TFA cleavage procedures. In this synthesis a peptide amide linker was chosen that transforms the attached side-chain of the glutamic acid from a carboxylic acid into an amide, and the resulting cleaved peptide has a glutamine residue at this position. The GTD-(2′,2) fragment was prepared by the standard Fmoc-SPPS procedure.

Native chemical ligation: Native chemical ligation of the two synthetic peptide fragments was performed using standard methodology[39] and was complete within 22 h. The ligation reaction was followed by analytical HPLC (Figure 3). After 1.5 h the formation of the reactive phenyl thioester GTD-(1,1′)-SPh was observed as well as a small amount of ligation product. After 7 h the ligation reaction was about $^2/_3$ complete and additional GTD-(2′,2) was added to the reaction mixture at this point. After completion of the ligation reaction, GTD-Lig was readily purified in good yield and identified by electrospray mass spectroscopy: found: 9836.5 Da; calculated: 9837.1 Da.

Biophysical characterization of GTD-Lig: The pH dependence of GTD-Lig was studied with far-UV CD in the range of pH 2.1–9.8. The protein was found to precipitate near its calculated pI of 7. At a pH below 5 and above 8, GTD-Lig is fully soluble and the CD spectrum is independent of pH (data not shown). To analyze and find the optimum conditions in which GTD-Lig folds into the desired well-defined four-helix bundle protein, the pH and temperature were varied and the structure was studied by ^{1}H NMR spectroscopy. At pH 3 the ^{1}H NMR spectrum of GTD-Lig is well dispersed and has narrow line-widths relative to the spectrum at pH 4.5, which shows less dispersion and broad line-widths (Figure 4). Well-dispersed spectra with narrow line-widths are characteristics that distinguish between a native-like structure and the molten globular state. The broad lines that

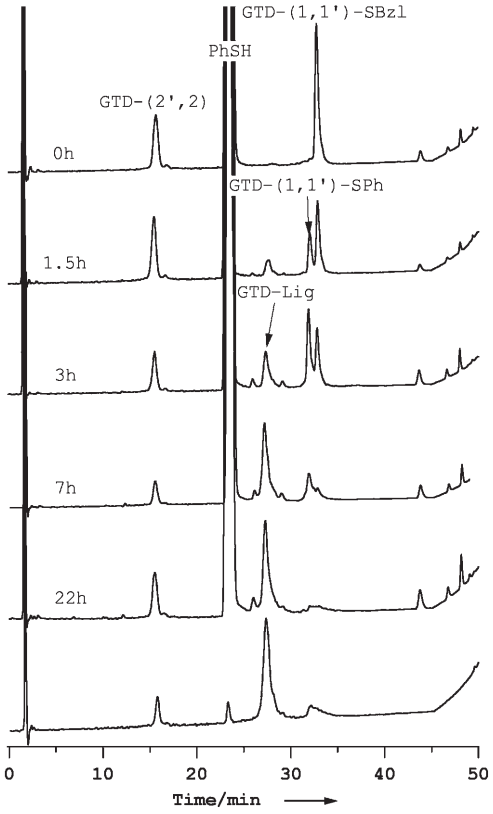


Figure 3. The ligation between GTD-(1,1′)-SBzl and GTD-(2′,2) was monitored by analytical HPLC (chromatograms at 230 nm are shown). After 22 h, the formation of the product GTD-Lig was essentially complete. The bottom chromatogram shows the reaction mixture after removal of excess thiophenol by diethyl ether extraction.

are found in the spectrum at pH 4.5 indicate the formation of a molten globular structure. The molten globular structure is most probably induced by nonspecific aggregation as the pH approaches the protein's pI.

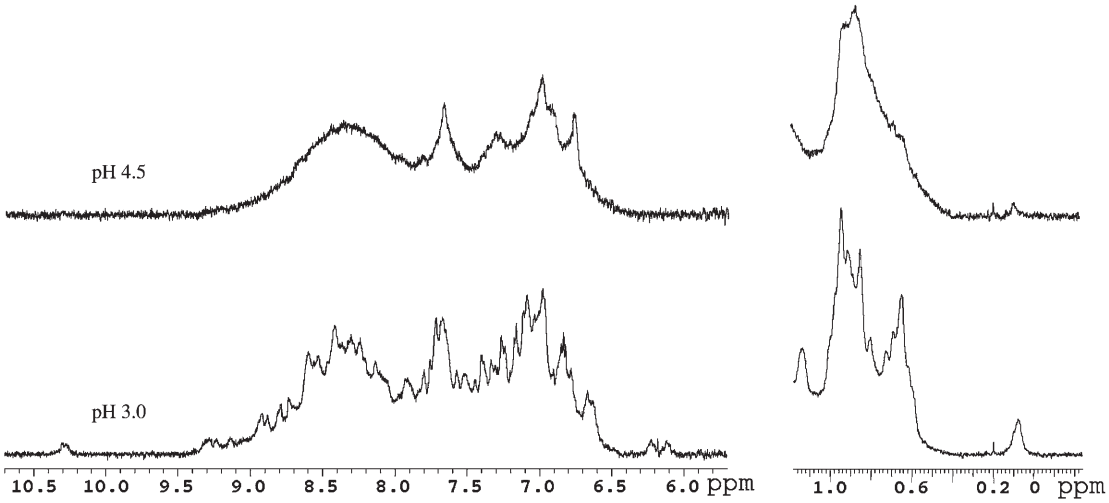


Figure 4. pH dependence of the aromatic-amide and methyl region of the ^{1}H NMR spectrum of GTD-Lig at a concentration of 0.5 mM and 5 °C.

The temperature dependence of the ^{1}H NMR spectrum of GTD-Lig was studied at pH 3, at which pH the protein shows native-like behavior (Figure 5). With increasing temperature the chemical-shift dispersion decreases and the line-widths become broader; this shows that at pH 3 and below 20 °C GTD-Lig is in slow exchange on the NMR timescale.[40]

The CD spectra in the far-UV region (Figure 6) show that the protein in aqueous solution at pH 3 at different temperatures has the characteristics of an α-helical protein, with minima at 222 and 208 nm and a maximum at 195 nm. The protein structure is the most helical below 20 °C, with $[\theta]_{222} = -20500$ deg cm^2 dmol^{-1}, which is the same, within experimental error, as that reported for both GTD-43 and GTD-C. Thus the modifications introduced into the sequences of the parent proteins to create a monomeric structure with rearranged loops apparently do not affect the secondary structure. The CD spectrum of GTD-Lig was also found to be independent of concentration in the measured range of 2–500 μM at pH 3 and 21 °C. The following structure and conformational studies of GTD-Lig were all performed within this concentration range and at pH 3.

The two tryptophan residues in the aromatic ensemble of the hydrophobic core are convenient probes for studying the structure and stability of GTD-Lig. The tryptophan fluorescence maximum is at 329 nm in aqueous solution at pH 3 and 20 °C; this value is blue-shifted relative to that of a fully exposed tryptophan, which usually has a maximum at about 350 nm (see the Supplementary Information). The significantly lower value shows that the tryptophan molecules are partly buried in the hydrophobic core of the folded structure.[41] Note also that the fluorescence maximum is red-shifted to 337 nm at 80 °C. This indicates that the tryptophan residues are still buried but to a lesser extent.

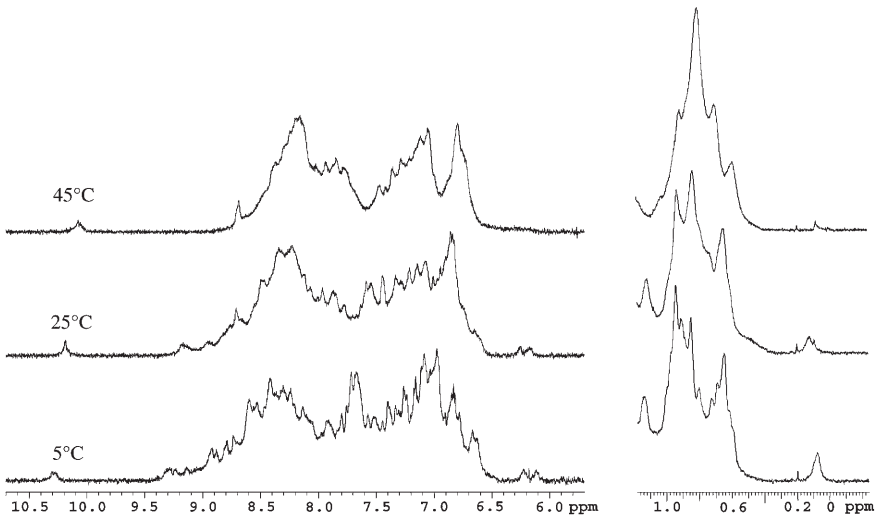


Figure 5. The aromatic-amide and methyl region of the ^{1}H NMR spectrum of GTD-Lig as a function of temperature at pH 3.0. Same sample as in Figure 4.

Sedimentation equilibrium ultracentrifugation: To verify that GTD-Lig is indeed monomeric and to probe for concentration dependence, sedimentation equilibrium ultracentrifugation of 6 and 30 μM aqueous solutions (salt-free, pH 3) of the protein was performed; the same conditions as those used for stability and structural measurements. The equilibrium curves were measured at five speeds between 15000 and 35000 rpm. Analysis was performed at each concentration

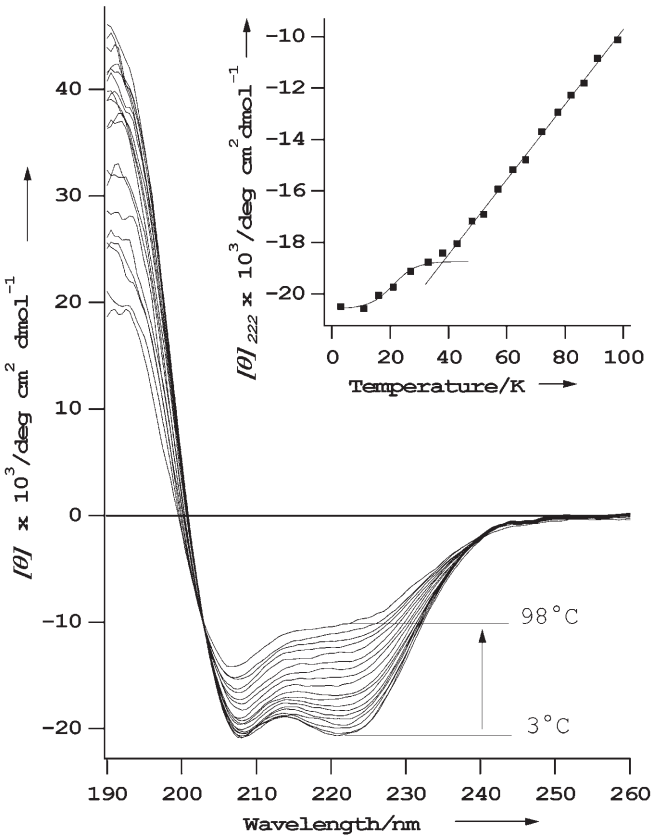


Figure 6. Temperature dependence of the far-UV CD spectra. Temperatures range from 3°C (bottom trace) to 98°C (top trace at 222 nm). (Inset): the mean residue ellipticity at 222 nm as a function of temperature. The solid lines fitted to the data are only to aid the eye and are described in the Discussion section.

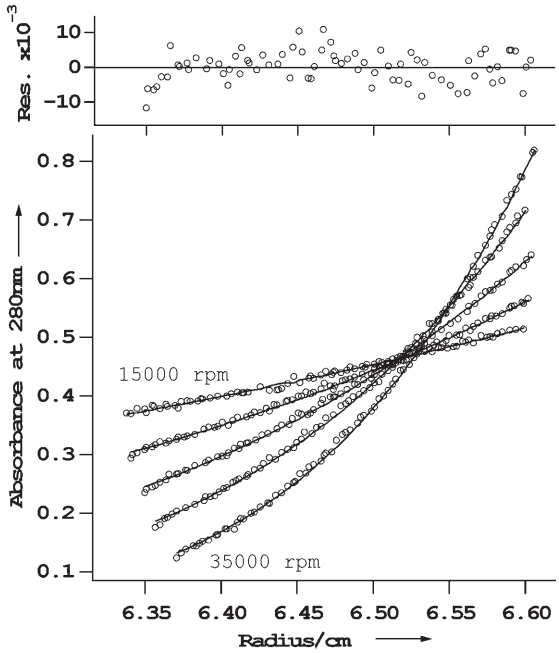


Figure 7. Equilibrium ultracentrifugation analysis of GTD-Lig (30 μM) in aqueous solution at pH 3.0. Data are from measurements taken at 15000, 20000, 25000, 30000, and 35000 rpm and globally fitted to a single species of a monomer with M_W fixed at 9837.1 Da. The second virial coefficient was estimated to be 5.4×10^{-5} mol L g^{-2}. The solid lines represent the calculated fit. The residuals of the fit of the 25000 rpm data are shown above.

by simultaneously fitting the data measured at different rotor speeds to the equation for a single species of a monomer. For the 6 and 30 μM solutions, the apparent molecular weights, $M_{W,app}$, were determined to be 9100 and 6500 Da, respectively. The value of $M_{W,app}$ determined for the low-concentration sample is close to the calculated molecular weight, $M_{W,calcd}$, which is 9837 Da; this indicates that the protein GTD-Lig, as designed, is a monomer at pH 3.0. As can be seen, when the data are fitted individually at each concentration, $M_{W,app}$ becomes less than $M_{W,calcd}$ and decreases with increasing protein concentration. This indicates that the solutions of GTD-Lig are nonideal.[42] Nonideality can be described in terms of the second virial coefficient B, which was determined from the equilibrium curves obtained from the 30 μM solution in which the nonideality is most obvious. Data were fitted with M_W set equal to $M_{W,calcd}$ (9837 Da), which gave a value of B of 5.4×10^{-5} mol L g^{-2} (Figure 7). The residuals of the 25000 rpm data-fit are shown and are representative of the goodness-of-fit as the residuals are random and centered around zero.

The conformational stability of GTD-Lig: The appearance of the ^{1}H NMR spectrum shows that GTD-Lig has a well-defined native-like structure at pH 3 and at temperatures below 20°C (Figure 4 and Figure 5). To complement this NMR thermal study, the thermal unfolding of GTD-Lig was studied by monitoring the tryptophan absorbance and circular dichroism in both the far- and near-UV regions at pH 3. Tryptophan is a suitable chromophore for studying transitions of the tertiary structure, whereas far-UV CD is suitable for studying changes to the secondary structure. Figure 8 shows the temperature dependence of the tryptophan absorbance which decreases sharply between 20 and 40°C. Above 40°C there is a less pronounced linear decrease in the absorbance. The unfolding transition is highly cooperative and reversible: When the temperature was lowered, the value of the tryptophan absorbance returned to its original level.

Monitoring the thermal unfolding by far-UV CD (Figure 6) shows that the negative values of the mean residue ellipticities at 208 and 222 nm decrease with increasing temperature, which indicates a decrease of helical content. There is a transition between 10 and 40°C and above 40°C the thermal unfolding is diffuse. Thus the thermal unfolding monitored by far-UV CD corresponds to the temperature dependence of the tryptophan absorbance. The helical structure of GTD-Lig is remarkably stable to temperature as the far-UV CD spectrum at 98°C retains its characteristic α-helical minima at 208 and 222 nm and the maxima at 195 nm. The value of $[\theta]_{222}$ at 98°C is about half the low-temperature

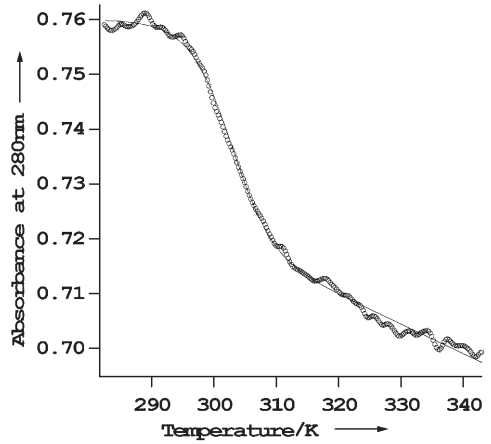


Figure 8. Heat-induced unfolding of GTD-Lig. The thermal unfolding was followed by monitoring the absorbance at 280 nm using a peptide concentration of 25 μM at pH 3.0. The solid line shows the nonlinear least-squares best fit of the experimental results to Equation (2) (see Experimental Section).

value. Usually natural proteins are completely denatured at high temperatures.

The near-UV CD spectra of GTD-Lig at different temperatures are shown in Figure 9. The spectrum of GTD-Lig at pH 3 and 5 °C is characterized by a strong broad negative band with minima at 283 and 290 nm and a broad maximum

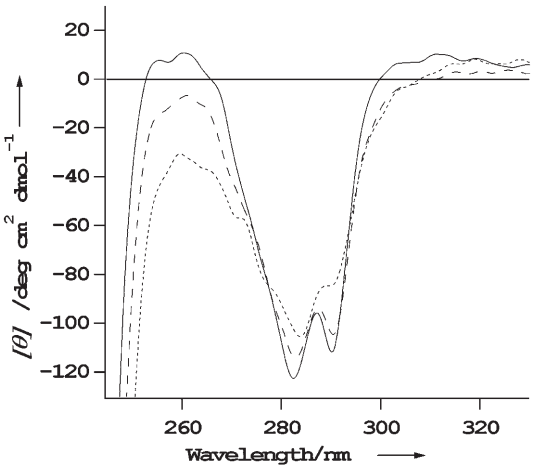


Figure 9. Effect of temperature on the near-UV CD spectra of GTD-Lig at a peptide concentration of 100 μM in aqueous solution at pH 3. Solid line: 5 °C; dashed line: 43 °C; dotted line: 98 °C.

at 260 nm. The spectrum is typical of tryptophan in an ordered environment and from the 7 nm vibronic spacing between the two minima, they are identified as the tryptophan ^1L_b vibronic bands.[43] Upon heating the spectrum loses some of its vibronic structure as both the minima at 283 and 290 nm decrease. The small change in the mean residue ellipticities upon heating indicates that the tryptophan residues are still in a hydrophobic environment at high temperatures.

Binding studies using the hydrophobic fluorescent dye ANS (1-anilinonaphthalene-8-sulfonic acid) are commonly performed to discriminate molten globules from native proteins. ANS has a strong affinity for the hydrophobic core found in molten globules and its binding to it is accompanied by a strongly enhanced fluorescence intensity and a blue shift of the fluorescence maximum wavelength.[44] In the study of a ten-fold excess of ANS (20 μM) and its interaction with GTD-Lig (2 μM), only an eight-fold increase in the fluorescence intensity is seen and the fluorescence maximum wavelength, λ_{max}, is blue-shifted to 480 nm (Figure 10).

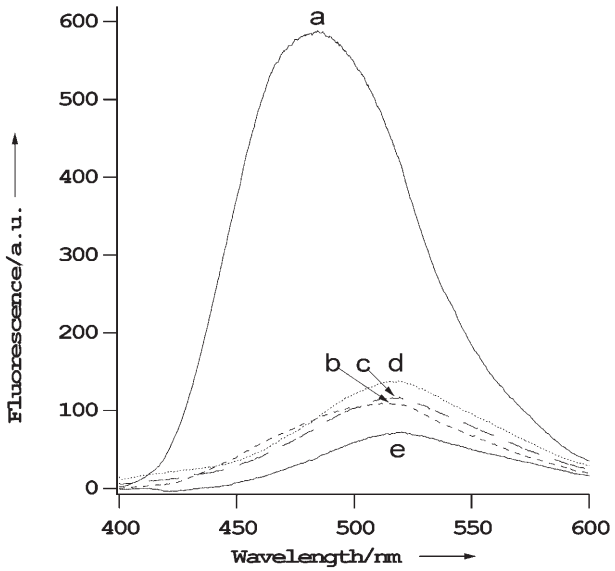


Figure 10. Study of the interaction of ANS (20 μM) and GTD-Lig (2 μM) in aqueous solution at pH 3 by fluorescence spectroscopy: a) No GuHCl added; b) 1 M GuHCl; c) 4 M GuHCl; d) 7 M GuHCl; e) ANS alone.

When studying the same mixture with 1, 4, and 7 M GuHCl, the fluorescence intensity and λ_{max} values are almost the same as those of ANS alone. As GTD-Lig has a native structure at 1 M GuHCl (see Figure 11), the observed affinity of ANS for GTD-Lig at 0 M GuHCl is probably due to electrostatic interactions between the negative charge of ANS and the large positive charge of the protein at pH 3,[45] which are reduced at a high ionic concentration of GuHCl.

Heat denaturation: The NMR, CD, and UV thermal studies indicate that in the low-temperature range, below 40 °C, a transition occurs that disrupts the well-defined tertiary structure and that an intermediate state is formed. With further heating this intermediate state shows characteristics of a molten globule, such as diffuse thermal unfolding. Also, at 40 °C, the values of the CD signals in the far- and near-UV region are still considerable, indicating a highly helical compact structure typical of molten globules.

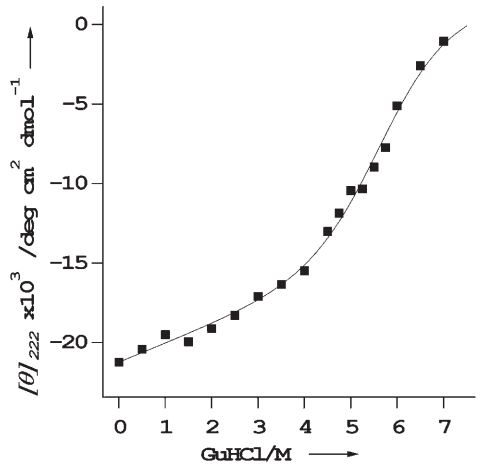


Figure 11. GuHCl-induced unfolding of GTD-Lig monitored by the mean residue ellipticity at 222 nm and 20 °C. The solid line is the nonlinear least-squares best fit of the experimental results to Equation (5).

By regarding the temperature dependence of the tryptophan absorbance in Figure 8 as a two-state transition between the native structure (N) and the intermediate molten globule (I) [see Equation (1) in the Experimental Section], the Gibbs free energy of thermal unfolding to the I state, $\Delta G_{NI}(T)$, was calculated. A two-state transition between the native and intermediate states is defined as the UV data show that thermal unfolding is reversible. The temperature at the midpoint of the transition, $T_{m,NI}$, was determined to be 29 °C when the linear change above 40 °C was defined as a post-transitional baseline. The UV absorbance data at 280 nm are well-fitted to the two-state model according to Equation (2). The value of ΔC_P, the difference in heat capacity between the two states, is estimated[46] to be 14 kcal mol^{-1} K^{-1} res^{-1} and the derived parameters $T_{m,NI}$ and $\Delta H_{m,NI}$ were used to calculate $\Delta G_{NI}(10\,°C)$ using Equation (3) (see Experimental Section). The value of $\Delta H_{m,NI}$ was determined to be 54 ± 1.2 kcal mol^{-1} and $\Delta G_{NI}(10\,°C)$ was calculated to be 2.6 kcal mol^{-1}.

Denaturation with GuHCl: Chemical denaturation allowed the stability properties of GTD-Lig to be further characterized. Denaturation was followed by monitoring the mean residue ellipticity at 222 nm, $[\theta]_{222}$, as a function of GuHCl concentration at 20 °C and pH 3 (Figure 11). Unlike heat denaturation, GuHCl denaturation completely unfolds the protein. The denaturation first gives a pre-transitional baseline from 0 to 4 M GuHCl and thereafter a sigmoidal curve with a sharp decrease between 4 and 7 M GuHCl is observed. The midpoint of the transition is 5.5 M. Assuming that the denaturation may be described by a two-state model that involves only a native structure (N) and an unfolded random coil (U), nonlinear least-squares fitting of the experimental results to Equation (5) (see Experimental Section) gives the Gibbs free energy of unfolding in aqueous solution, ΔG_{H_2O}. The constant m_G in Equation (5) is also determined and is a measure of the cooperativity of the denaturation. The values of ΔG_{H_2O} and m_G are 4.6 ± 0.5 kcal mol^{-1} and 0.8 kcal mol^{-1} M^{-1}, respectively.

Discussion

The design, total synthesis, and structural and thermodynamic characterization of a monomeric protein, GTD-Lig, is discussed.

Structure: The structure of GTD-Lig has been demonstrated to fold into the designed monomeric four-helix bundle protein by NMR, far- and near-UV CD spectroscopy, sedimentation equilibrium ultracentrifugation, and fluorescence spectroscopy. GTD-Lig exhibits strong CD signals in the far- and near-UV regions, and the value of the mean residue ellipticity at 222 nm is $-20\,500$ deg cm^2 dmol^{-1}, which is the same, within experimental error, as that of the original dimeric bundles.[29–31] Note that the ^{1}H NMR spectra of GTD-Lig and GTD-43 are nearly identical, which indicates that the structures have the same fold. The sedimentation equilibrium study shows that GTD-Lig is a monomer in the concentration range measured. As the experiment was carried out under the same conditions as those used for structure and stability studies (salt-free, pH 3.0), the observed nonideality is mainly the result of charge–charge repulsion between the proteins, which alters the protein concentration distribution.[47]

Native-like structure: Proteins with well-defined native structures are discriminated from molten globule structures by the use of spectroscopic probes.[48] A tightly packed native protein has, for example, a well-dispersed ^{1}H NMR spectrum with narrow line-widths, a well-defined thermal melting point, strong CD signals in the far- and near-UV regions, and a low affinity to the hydrophobic fluorescent dye ANS. GTD-Lig fulfils these criteria as it has a well-dispersed ^{1}H NMR spectrum with narrow line-widths at pH 3.0 and 5 °C. The NH chemical shifts cover a wide range with signals shifted downfield to $\delta = 9.3$ ppm and the methyl region has a width of 0.9 ppm. GTD-Lig is comparable to the much larger naturally occurring four-helix bundle IL-4 (133 residues), which exhibits similar chemical-shift dispersion with NH chemical shifts as far downfield as $\delta = 9.3$ ppm and a methyl range of 0.94 ppm.[49] The methyl range for molten globules is less than 0.3 ppm. At pH 4.5 the ^{1}H NMR spectrum exhibits broadened resonances and poor resolution, which is typical of molten globule structures. Between pH 5 and 8 the protein is near its isoelectric point and in this pH range the protein has low solubility. The broad lines in the NMR spectrum at pH 4.5 may be due to aggregation. The temperature dependence of the ^{1}H NMR spectrum shows that the resolution is dramatically better at lower temperatures than at elevated temperatures. This shows that the structure is in slow conformational exchange on the NMR timescale and that the conformational exchange processes are accelerated at elevated temperatures,

which is demonstrated by the broadening of the resonances. The hydrophobic fluorescent dye ANS is found to bind weakly to GTD-Lig. This binding is expected to result from an electrostatic attraction[45] as ANS has a negative charge and GTD-Lig has a large positive charge at pH 3, at which it displays a native-like fold and at which pH the study was performed. This was confirmed by increasing the ionic strength of the solution, which resulted in a return of the ANS fluorescence to intensity and λ_{max} values of the unbound dye. Further evidence of a well-developed structure comes from the observation of a strong near-UV CD spectrum and the blue shift of the tryptophan fluorescence maximum to 329 nm. This indicates that the tryptophan residues are buried in the hydrophobic core.

Design: Rearranging the loops in the original de novo designed dimeric structure of GTD-43 has obviously not altered the well-folded characteristics of the tertiary structure. One of the factors contributing to the similar folding in the monomeric and dimeric structures is the choice of loop. The two new loops connecting helix 1 to helix 1′ and helix 2′ to helix 2 were given the sequence Gly-Thr-Gly-Pro, the same loop sequence as in the GTD-43 monomeric structure. The other main factor that contributes to the conformational stability of the folded structure is the hydrophobic core. The hydrophobic core is constructed of aromatic residues, which, through their reduced rotational freedom as well as aromatic interactions, are designed to restrict the dynamics between different conformations. The aromatic ensemble in GTD-Lig is constructed of Phe10 and Trp13 in helix 1, Phe34 and Trp37 in helix 1′, Phe58 in helix 2′, and Phe81 in helix 2 (Figure 1). It seems that the hydrophobic interactions, together with the reduced dynamics of the aromatics, allow large structural changes such as the ones demonstrated here with rearrangement of loops that convert the dimeric tertiary structure into a monomeric structure.

Heat denaturation: The thermal unfolding of GTD-Lig, followed by monitoring the tryptophan absorbance (Figure 8), apparently occurs in two steps as demonstrated by the presence of, first, a sigmoidal transition with a transition midpoint, T_m, of 29 °C. Secondly, there is an uncooperative melting of the protein at higher temperatures. After, the first transition an intermediate state (I) is populated. The N–I transition is cooperative as the observed change in the tryptophan absorbance is reversible. Similar conclusions can be inferred from the thermal unfolding monitored by far-UV CD. Though less obvious, it also shows a two-step unfolding process which seems to correspond to the temperature dependence of the tryptophan absorbance (Figure 6, inset; the low-temperature data is fitted to Equation (2) with the slope of the pre- and post-unfolding baselines set to zero, while the high-temperature data is linear). The observed isodichroic point at 203 nm in the CD study indicates a two-state transition.[50] This shows that the secondary and tertiary structures follow each other, indicating a cooperative unfolding to the intermediate state. In noncooperative unfolding secondary and tertiary structures unfold at different temperatures.[51]

GTD-Lig exhibits a well-defined structure at 10 °C and the value of $\Delta G_{NI}(10\,^{\circ}C)$ is calculated to be 2.6 kcal mol^{-1}, which is apparently enough of an energy difference for the folding of a native-like structure from the intermediate found at 40 °C. This temperature-induced intermediate shows structural characteristics typical of a molten globule, such as uncooperative melting and a retained helical structure. Further evidence of a molten globule structure is provided by the following points: The magnitudes of ΔG_{NI}(10 °C) and $\Delta H_{m,NI}$ correspond to partial unfolding, the ^{1}H NMR spectrum at 45 °C exhibits decreased chemical-shift dispersion and broad lines characteristic of a molten globule, and finally the study of the tryptophan residues at 40 °C by near-UV CD and fluorescence spectroscopy (see the Supporting Information) shows that the transition from the native to the intermediate state does not lead to greater exposure of the tryptophan residues to the solvent, which is typical for a molten globule structure.

GuHCl denaturation: GuHCl-induced unfolding of GTD-Lig is well-approximated by a two-state transition. The N–U transition occurs between 4 and 7 M GuHCl (Figure 11). The protein is moderately stable with a $\Delta G^{o}_{H_2O}$ value of 4.5 kcal mol^{-1}. This value is apparently enough for a well-folded structure to form. For a comparison with other monomeric four-helix bundle proteins of similar size, the binding domain of the kinas FRAP contains 95 residues and is stabilized by 6.27 kcal mol^{-1} with a value of m_G = 2.6 kcal mol^{-1} M^{-1} at 10 °C.[52] The corresponding values for the immunity proteins Im 7 (85 residues) and Im 9 (87 residues) are $\Delta G^{o}_{H_2O}$ = 5.6 kcal mol^{-1} and m_G = 1.1 kcal mol^{-1} M^{-1}, and $\Delta G^{o}_{H_2O}$ = 7.4 kcal mol^{-1} and m_G = 1.1 kcal mol^{-1} M^{-1}, respectively, at 10 °C.[53] So GTD-Lig is at the lower end of the stability range for monomeric four-helix bundle proteins of similar size. This could be due to a number of factors such as the size of the hydrophobic core and the number of electrostatic interactions. The stability of GTD-Lig was determined at pH 3. Therefore the number of stabilizing electrostatic interactions is expected to be low as a result of the protonation of aspartic and glutamic acids. The value of m_G is a measure of cooperativity and the value found for GTD-Lig is 0.8 kcal mol^{-1} M^{-1}, which is at the lower end of the range observed for natural proteins of similar size. The concentration of GuHCl at the midpoint of the denaturation, $[GuHCl]_{1/2}$, is 5.5 M. This high stability observed towards GuHCl is common for small proteins with low m values.[46]

It is difficult to compare the stabilities of monomeric and oligomeric proteins. Ferreira and co-workers have studied the dissociation and unfolding of the dimeric protein triosephosphate isomerase as well as a designed helix-loop-helix dimer (four-helix bundle) and found that both proteins are stabilized mainly through the association of the monomers.[18,54] Comparison of the stabilities of the monomeric GTD-Lig ($\Delta G^{o}_{H_2O}$ = 4.6 kcal mol^{-1}) and the dimeric GTD-43 ($\Delta G^{o}_{H_2O}$ = 10 kcal mol^{-1})[30] shows the dimeric structure to

have an increased stability, which could partly be a result of the association between the two monomers.

Conclusions

GTD-Lig clearly shows characteristics of a native protein with a well-defined structure and typical unfolding transitions such as a cooperative GuHCl denaturation and the formation of a heat-induced intermediate (I) state. Such I states are believed to be involved in the folding of natural proteins.[55,56] The de novo design of large monomeric proteins is another step towards improving our understanding of the forces and interactions that stabilize protein structures. The total synthetic approach utilizing native chemical ligation and standard Fmoc chemistry offers facile access to long polypeptides for protein-folding analysis as well as novel protein folds. The design of an aggregating four-helix bundle monomer that forms a tetramer may be possible in the future. By designing electrostatic and hydrophobic interactions, such a tetramer would create a quaternary structure with 16 helices. This would indeed test our understanding of protein folding and association.

Experimental Section

Peptide synthesis: Peptides were synthesized on a Pioneer automated peptide synthesizer (Applied Biosystems) using general fluorenylmethoxycarbonyl (Fmoc) procedures. The synthesis of GTD-(2′,2) was performed on a PAC-PEG linked Fmoc-Asn(Trt) polystyrene polymer (Applied Biosystems) with a substitution level of 0.18 mmol g^{-1}. For the synthesis of GTD-(1,1′)-SBzl a Fmoc-PAL-PEG polystyrene polymer (Applied Biosystems) with a substitution level of 0.17 mmol g^{-1} was used. The side-chains of the amino acids (Calbiochem-Novabiochem AG) were protected with piperidine-stable groups: *tert*-butyl (Ser, Thr, and Tyr), *tert*-butyl ester (Asp and Glu), *tert*-butoxycarbonyl (Lys and Trp), trityl (Asn, Gln, and His), and 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Arg). The Fmoc protecting groups were removed from the amino termini by treatment with 20% piperidine in DMF (v/v) for 8 min. A fourfold excess of amino acid was used in each coupling and amino acids were preactivated with a mixture of TBTU (Alexis Biochemicals)/DIPEA (1:2) in DMF. A standard amino acid coupling time of 60 minutes was used except in the cases of Asn, Arg, and Gln (120 min).

Synthesis of GTD-(1,1′)-SBzl: Fmoc-Glu-α-Allyl (Applied Biosystems) was coupled to the resin by its side-chain using the standard procedures described above. After completion of the sequence the N-terminus of the peptide was capped with acetic anhydride (0.3 M) in DMF. The C-terminal allyloxycarbonyl group was selectively removed by treatment with tetrakis(triphenylphosphine)palladium(0) [$Pd(PPh_3)_4$] (3 equiv) in a mixture of trichloromethane, acetic acid, and morpholine (17:2:1, v/v; 20 mL per gram of polymer) for 3 h at room temperature. To remove the Pd^0 the resin was washed sequentially with diethyldithiocarbamic acid (20 mM) in DMF and DIPEA (30 mM) in DMF, and finally with DMF. The C-terminal thioester of GTD-(1,1′) was prepared by coupling Gly-SBzl (4 equiv) with PyBop (4 equiv), HOBt (4 equiv), and DIPEA (8 equiv) in DMF (6 mL per g of polymer) for 1 h. Gly-SBzl was synthesized as previously described.[38]

GTD-(1,1′)-SBzl was deprotected and cleaved from the resin by stirring with a mixture of TFA/TIS/H_2O (95:2.5:2.5, v/v; 15 mL per g of resin) for 2 h at room temperature. After removing the resin by filtration the peptide was precipitated and washed three times with cold diethyl ether. GTD-(1,1′)-SBzl was purified by reversed-phase HPLC on a semipreparative Kromasil C-8 column eluting isocratically with 36.5% isopropanol and 0.1% TFA at a flow rate of 10 mL min^{-1} (retention time = 30 min). The identity of the peptide was verified by electrospray mass spectrometry with a VG Analytical ZabSpec instrument. The calculated weight for GTD-(1,1′)-SBzl is 5189.7 Da; found 5188.9 Da.

Synthesis of GTD-(2′,2): After SPPS, GTD-(2′,2) was deprotected and cleaved from the resin using a mixture of TFA/H_2O/1,2-ethanedithiol/TIS (95:2.5:2.5:1, v/v). Work up, purification, and identification was performed as for GTD-(1,1′)-SBzl. For HPLC purification GTD-(2′,2) had a retention time of 22 min, eluting isocratically with 34% isopropanol and 0.1% TFA at a flow rate of 10 mL min^{-1}. The calculated weight for GTD-(2′,2) is 4770.5 Da; found 4769.8 Da.

Native chemical ligation: The ligation of the two unprotected synthetic peptide fragments was performed using a standard methodology.[32] Briefly, the GTD-(1,1′)-thioester (4.9 mg, 0.94 μmol) and the C-terminal segment GTD-(2′,2) (6.3 mg, 1.3 μmol) were dissolved in GuHCl (6 M) and sodium phosphate (100 mM, pH 7.5) to give a final concentration of 1 mM of the thioester. Thiophenol (3%, v/v) was added and the mixture was stirred for 22 h and monitored by analytical C-8 reversed-phase HPLC. Chromatographic separations were achieved by using a 30.7–55% AcCN + 0.1% TFA linear gradient at 0.57% min^{-1} and a flow rate of 1.5 mL min^{-1}. After the ligation, the product was purified by first removing the thiophenol by extraction with diethyl ether followed by semipreparative C-8 reversed-phase HPLC and lyophilization to afford 6 mg [65% yield based on the limiting GTD-(1,1′)-thioester] of the full-length protein GTD-Lig. The product was characterized by analytical C-8 reversed-phase HPLC and ES-MS: calculated 9837.1 Da; found 9836.5 Da.

General: An equivalent of DTT was used in the structure and stability experiments, unless otherwise noted, to reduce eventual oxidized cysteine sulfhydryl groups.

NMR spectroscopy: 500 MHz ^1H NMR spectra were recorded with a Varian Unity 500 NMR spectrometer equipped with a matrix shim system from Resonance Research Inc. using a 90° pulse of 11.5 μs and a sweep width of 8000 Hz. Solvent suppression was accomplished by weak pre-irradiation of the water resonance for 1.5 s. The NMR spectra obtained are an average of 64 transients. Samples contained 0.5 mM peptide in H_2O/D_2O (90:10) with [D_{10}]DTT (5 mM) to keep cysteines reduced. The temperature dependence of the ^1H NMR spectrum of GTD-Lig was determined at pH 3 and the pH dependence was performed at 5 °C.

Circular dichroism spectroscopy: Circular dichroism (CD) spectra were acquired by signal averaging on a Jasco J-714 spectropolarimeter and a baseline was recorded separately and subtracted. The instrument was routinely calibrated using [D]-10-(+)-camphorsulfonic acid. The spectra were processed and presented using the IGOR software from Wavemetrics Inc. Far-UV spectra were recorded from 260 to 185 nm using quartz cells, 1 nm bandwidths, and 0.25 s response times. Near-UV CD spectra were recorded from 350 to 240 nm using a 1.0 cm quartz cell, 1 nm bandwidths, and 0.5 s response times. The far- and near-UV spectra obtained are the averages of 10 scans and ellipticities are reported as mean residue ellipticities. Stock peptide concentrations were determined spectrophotometrically in GuHCl (6.0 M) using an extinction coefficient at 280 nm of 13 512 $cm^{-1} M^{-1}$.

The far-UV CD pH profile of GTD-Lig was recorded using an aqueous peptide solution (42 μM) in a 0.1 cm quartz cell at 21 °C. The pH was adjusted with HCl and NaOH (0.1 M).

The temperature dependence of the mean residue ellipticity was determined in the far- and near-UV regions. For the far-UV CD temperature profile, a jacketed quartz cell with a 0.05 cm path length was used and the spectra of GTD-Lig were recorded using an aqueous peptide solution (53 μM) at pH 3. The near-UV CD temperature profile was recorded using an aqueous peptide solution (100 μM) at pH 3. The temperature was varied in the range from 3 to 98 °C, controlled by a Lauda-T water bath. The sample cell temperature was raised in 4–6 °C increments with 10-minute delays between spectra to obtain thermal equilibrium.

A solution of GTD-Lig (20 μM) was denatured by the addition of 0–7 M GuHCl at 20 °C. The samples were prepared from a stock solution of the peptide (100 μM) by the addition of aliquots from a stock solution of

GuHCl (7.68 M) followed by dilution with water to the desired volume; the pH was set to 3.0 in all solutions. The samples were equilibrated for 30 minutes before analysis. The concentration of the stock solution of GuHCl was determined from its refractive index. Spectra were recorded from 260 to 210 nm using a 0.05 cm quartz cell, 1 nm bandwidth, and 2 s response time.

The CD spectrum of GTD-Lig was recorded as a function of concentration at pH 3 and 21 °C. The aqueous concentrations studied and respective cell lengths were: 1.9 μM, 0.5 cm; 9.6 μM, 0.5 cm; 96 μM, 0.05 cm; 480 μM, 0.01 cm.

Sedimentation equilibrium ultracentrifugation: The aggregation state of GTD-Lig was determined by equilibrium sedimentation ultracentrifugation at 20 °C using a Beckman Coulter Optima XL-I analytical ultracentrifuge equipped with an An-50 Ti Rotor and six sector cells. Rotor speeds of 15000–35000 rpm were used with steps of 5000 rpm. The equilibrium at each speed was reached after 20 h. The sedimentation was monitored by measuring the absorbance at 280 nm. Aqueous solutions of GTD-Lig (6 and 30 μM) were analyzed at pH 3.0. The SedNterp software was used to calculate the solvent density (0.99823 g mL^{-1}) and the partial specific volume (0.7364 mL g^{-1}) of the protein. The sedimentation properties were analyzed using the self-association model in the Beckman data analysis software package. Global data fits were obtained by fitting five sets of data simultaneously to the sedimentation equilibrium equation. The quality of the fits was assessed by comparing the randomness of the residuals, the magnitude of the computed variance, and by checking the physical reality of offsets and association constants.

Fluorescence measurements: The temperature dependence of the tryptophan fluorescence was studied using a Spex 1680 Fluorolog τ2 spectrometer using a 0.3 cm reduced cell with a sample volume of 60 μL. The bandwidth for excitation was 0.5 nm and for emission 4 nm, and the response time was 2 s. The changes in fluorescence intensity and wavelength maximum were studied using an aqueous peptide solution (4 μM) at pH 3 with excitation at 295 nm. The emission spectra were recorded at 1 nm intervals between 300 and 400 nm.

The fluorescence spectra of ANS (1-anilinonaphthalene-8-sulfonate) binding were recorded with an F-4500 spectrofluorimeter (Hitachi High-Technologies Corp.) using a 1.0 cm quartz cell at 21 °C. An excitation wavelength of 370 nm was used and the emission was monitored between 400 and 600 nm. Excitation and emission slits of 5 and 10 nm, respectively, were used. The baseline was recorded separately and subtracted from the spectra. The fluorescence of ANS (20 μM) in the presence of GTD-Lig (2 μM) at pH 3.0 was recorded and compared with solutions containing GuHCl and ANS alone.

UV spectroscopy: The temperature dependence of the tryptophan absorbance at 280 nm was monitored using a Cary 100 instrument with a Cary temperature controller in a temperature range of 8–70 °C. An aqueous solution of GTD-Lig (25 μM) at pH 3 in a 1 cm quartz cell, 1 nm bandwidths, and 5 s response times was used. The temperature was increased at 0.25 °C min^{-1} and data was collected at each 0.25 °C.

Data analysis: The thermal unfolding curve derived from the temperature dependence of the tryptophan absorbance was analyzed for the two-state transition between native (N) and partly unfolded or intermediate (I) structures, as described by Equation (1), where $K_{NI}=[I]/[N]$, the equilibrium constant between the native and populated intermediate states. The free energy of unfolding of GTD-Lig to the I state was determined from the thermal unfolding. The unfolding to the completely unfolded state (U) could not be analyzed as this transition is an undefined melting process. The observed absorbance, A_{obs}, can be described in terms of the native, A_N, and unfolded baselines, A_I, and the fraction of native protein, f_N, by Equation (2).

$$N \overset{K_{NI}}{\Leftrightarrow} I \overset{K_{IU}}{\Leftrightarrow} U \tag{1}$$

$$A_{obs} = f_N(A_N - A_I) + A_I \tag{2}$$

The folded and unfolded (intermediate) baselines were approximated from the linear temperature functions, $A_N = A^{\circ}_N + m_N T$ and $A_I = A^{\circ}_I + m_I T$, where A°_N and A°_I are the extrapolated absorbances of the baselines at 0 K and m_N and m_I are the slopes of the baselines. The fraction of native protein, f_N, is expressed as $f_N = 1-\{[I]/([I]+[N])\}$, and combining this with $K_{NI}=[I]/[N]$, $\Delta G^{\circ} = -RT \times \ln K_{NI}$, and the Gibbs–Helmholtz equation [Eq. (3)], the fraction of native protein, f_N, can be expressed in terms of thermodynamic parameters by Equation (4), where ΔH_m is the enthalpy of unfolding at the transition midpoint temperature, T_m, ΔC_p is the difference in heat capacity between the two states, and R is the gas constant.

$$\Delta G(T) = \Delta H_m(1 - T/T_m) + \Delta C_p[T - T_m - T\ln(T/T_m)] \tag{3}$$

$$f_N = \frac{1}{1 + \exp\{\Delta H_m/R(T/T_m - 1/T) + \Delta C_p/RT[T_m - T + T\ln(T/T_m)]\}} \tag{4}$$

For the N–I transition the thermodynamic parameters were determined from nonlinear least-squares fitting to the thermal denaturation curve with Equations (2) and (4); this gives $\Delta H_{m,NI}$, $T_{m,NI}$, $\Delta C_{p,NI}$, A°_N, A°_I, m_N, and m_I. In the initial fitting process, all the thermodynamic parameters were allowed to vary. As $\Delta H_{m,NI}$ and $\Delta C_{p,NI}$ are strongly linked and the $\Delta C_{p,NI}$ value obtained was difficult to define, a $\Delta C_{p,NI}$ contribution of 14 cal mol^{-1} K^{-1} per amino acid was used to give an estimate, which was fixed for the data analysis.[46] The value of 14 cal mol^{-1} K^{-1} per amino acid is an average value for globular protein N–U transitions and has been used here for the transition to the I state, as the corresponding value for the I–U transition, $\Delta C_{p,IU}$, is expected to be small. The values obtained for $\Delta H_{m,NI}$ and $T_{m,NI}$ were used to calculate the free energy of unfolding of the protein to the I state, $\Delta G^{\circ}_{NI}(T)$, from Equation (3).

The experimental results of GuHCl-induced unfolding, as monitored by CD at 222 nm (Figure 11), were fitted to a two-state model using Equation (5) described by Santoro and Bolen,[57] where θ°_N and θ°_U are the extrapolated signals of the pre- and post-transitional baselines to zero denaturant concentration [D] and m_N and m_U are the slopes of the baselines. $\Delta G^{\circ}_{H_2O}$ is the free energy of unfolding extrapolated to aqueous solution and m_G is the slope. Nonlinear least-squares fitting of the experimental results to Equation (5) gives the free energy of unfolding in aqueous solution. All nonlinear least-squares fittings were performed with the Igor Pro (WaveMatric) program.

$$\theta_{obs} = \frac{(\theta^{\circ}_N + m_N[D]) + (\theta^{\circ}_U + m_U[D])\exp\{-(\Delta G^{\circ}_{H_2O} + m_G[D])/RT\}}{1 + \exp\{-(\Delta G^{\circ}_{H_2O} + m_G[D])/RT\}} \tag{5}$$

Acknowledgements

I would like to thank L. Baltzer for fruitful discussions and support, M. Lundqvist for assistance with the equilibrium sedimentation ultracentrifugation and B. G. Karlsson for the kind gift of [D_{10}]DTT.

[1] J. W. Bryson, J. R. Desjarlais, T. M. Handel, W. F. DeGrado, *Protein Sci.* **1998**, *7*, 1404–1414.
[2] J. S. Johansson, B. R. Gibney, J. J. Skalicky, A. J. Wand, P. L. Dutton, *J. Am. Chem. Soc.* **1998**, *120*, 3881–3886.
[3] V. Sieber, G. R. Moe, *Biochemistry* **1996**, *35*, 181–188.
[4] M. D. Struthers, R. P. Cheng, B. Imperiali, *Science* **1996**, *271*, 342–345.
[5] T. Kortemme, M. Ramirez-Alvarado, L. Serrano, *Science* **1998**, *281*, 253–256.
[6] C. M. Kraemer-Pecore, A. M. Wollacott, J. R. Desjarlais, *Curr. Opin. Chem. Biol.* **2001**, *5*, 690–695.
[7] B. I. Dahiyat, S. L. Mayo, **1997**, *Science 278*, 82–87.
[8] P. B. Harbury, J. J. Plecs, B. Tidor, T. Alber, P. S. Kimt, *Science* **1998**, *282*, 1462–1467.
[9] F. V. Cochran, S. P. Wu, W. Wang, V. Nanda, J. G. Saven, M. J. Therien, W. F. DeGrado, *J. Am. Chem. Soc.* **2005**, *127*, 1346–1347.
[10] J. R. Calhoun, H. Kono, S. Lahr, W. Wang, W. F. DeGrado, J. G. Saven, *J. Mol. Biol.* **2003**, *334*, 1101–1115.

[11] B. Kuhlman, G. Dantas, G. C. Ireton, G. Varani, B. L. Stoddard, D. Baker, *Science* **2003**, *302*, 1364–1368.
[12] M. Mutter, G. G. Tuchscherer, C. Miller, K. H. Altmann, R. I. Carey, D. F. Wyss, A. M. Labhardt, J. E. Rivier, *J. Am. Chem. Soc.* **1992**, *114*, 1463–1470.
[13] S. Futaki, M. Aoki, T. Ishikawa, F. Kondo, T. Asahara, M. Niwa, Y. Nakaya, T. Yagami, K. Kitagawa, *Bioorg. Med. Chem.* **1999**, *7*, 187–192.
[14] R. Schnepf, P. Horth, E. Bill, K. Wieghardt, P. Hildebrandt, W. Haehnel, *J. Am. Chem. Soc.* **2001**, *123*, 2186–2195.
[15] J. Brask, J. M. Dideriksen, J. Nielsen, K. J. Jensen, *Org. Biomol. Chem.* **2003**, *1*, 2247–2252.
[16] D. A. Lindhout, J. R. Litowski, P. Mercier, R. S. Hodges, B. D. Sykes, *Biopolymers* **2004**, *75*, 367–375.
[17] K. H. Mayo, E. Ilyina, *Protein Sci.* **1998**, *7*, 358–368.
[18] A. Chapeaurouge, J. S. Johansson, S. T. Ferreira, *J. Biol. Chem.* **2002**, *277*, 16478–16483.
[19] G. Ghirlanda, J. D. Lear, N. L. Ogihara, D. Eisenberg, W. F. DeGrado, *J. Mol. Biol.* **2002**, *319*, 243–253.
[20] K. Enander, G. T. Dolphin, B. Liedberg, I. Lundström, L. Baltzer, *Chem. Eur. J.* **2004**, *10*, 2375–2385.
[21] K. Johnsson, R. K. Allemann, H. Widmer, S. A. Benner, *Nature* **1993**, *365*, 530–532.
[22] P. T. P. Kaumaya, K. D. Berndt, D. B. Heidorn, J. Trewhella, F. J. Kezdy, E. Goldberg, *Biochemistry* **1990**, *29*, 13–23.
[23] Y. B. Yu, *Adv. Drug Delivery Rev.* **2002**, *54*, 1113–1129.
[24] F. Rabanal, W. F. DeGrado, P. L. Dutton, *J. Am. Chem. Soc.* **1996**, *118*, 473–474.
[25] W. F. DeGrado, C. M. Summa, V. Pavone, F. Nastri, A. Lombardi, *Annu. Rev. Biochem.* **1999**, *68*, 779–819.
[26] R. B. Hill, D. P. Raleigh, A. Lombardi, W. F. DeGrado, *Acc. Chem. Res.* **2000**, *33*, 745–754.
[27] L. Baltzer, H. Nilsson, J. Nilsson, *Chem. Rev.* **2001**, *101*, 3153–3163.
[28] T. M. Handel, S. A. Williams, W. F. DeGrado, *Science* **1993**, *261*, 879–885.
[29] G. T. Dolphin, L. Brive, G. Johansson, L. Baltzer, *J. Am. Chem. Soc.* **1996**, *118*, 11297–11298.
[30] L. Brive, G. T. Dolphin, L. Baltzer, *J. Am. Chem. Soc.* **1997**, *119*, 8598–8607.
[31] G. T. Dolphin, L. Baltzer, *Folding & Design* **1997**, *2*, 319–330.
[32] P. E. Dawson, T. W. Muir, I. Clarklewis, S. B. H. Kent, *Science* **1994**, *266*, 776–779.
[33] P. E. Dawson, S. B. H. Kent, *Annu. Rev. Biochem.* **2000**, *69*, 923–960.
[34] G. G. Kochendoerfer, S.-Y. Chen, F. Mao, S. Cressman, S. Traviglia, H. Shao, C. L. Hunter, D. W. Low, E. N. Cagle, M. Carnevali, V. Gueriguian, P. J. Keogh, H. Porter, S. M. Stratton, M. C. Wiedeke, J. Wilken, J. Tang, J. J. Levy, L. P. Miranda, M. M. Crnogorac, S. Kalbag, P. Botti, J. Schindler-Horvat, L. Savatski, J. W. Adamson, A. Kung, S. B. H. Kent, J. A. Bradburne, *Science* **2003**, *299*, 884–887.
[35] For a recent review, see: C. M. Gross, D. Lelievre, C. K. Woodward, G. Barany, *J. Pept. Res.* **2005**, *57*, 395–410.
[36] J. Tulla-Puche, G. Barany, *J. Org. Chem.* **2004**, *69*, 4101–4107.
[37] M. Villain, H. Gaertner, P. Botti, *Eur. J. Org. Chem.* **2003**, 3267–3272.
[38] J. Alsina, T. S. Yokum, F. Albericio, G. Barany, *J. Org. Chem.* **1999**, *64*, 8761–8769.
[39] P. E. Dawson, M. J. Churchill, M. R. Ghadiri, S. B. H. Kent, *J. Am. Chem. Soc.* **1997**, *119*, 4325–4329.
[40] K. Wüthrich, *NMR of Proteins And Nucleic Acids*, Wiley, New York, **1986**.
[41] Y. Lakowicz, *Principles of Fluorescence Spectroscopy*, Plenum Press, New York, **1983**.
[42] T. M. Laue, *Methods Enzymol.* **1995**, *259*, 427–452.
[43] E. H. Strickland, *CRC Crit. Rev. Bioeng.* **1974**, *2*, 113–175.
[44] G. V. Semisotnov, N. A. Rodionova, O. I. Razgulyaev, V. N. Uverskii, A. F. Gripas, R. I. Gil′manshin, *Biopolymers* **1991**, *31*, 119–128.
[45] V. Ali, K. Prakash, S. Kulkarni, A. Ahmad, K. P. Madhusudan, V. Bhakuni, *Biochemistry* **1999**, *38*, 13635–13642.
[46] J. K. Myers, C. N. Pace, J. M. Scholtz, *Protein Sci.* **1995**, *4*, 2138–2148.
[47] C. Tanford, *Adv. Protein Chem.* **1968**, *21*, 121–282.
[48] D. A. Dolgikh, L. V. Abaturov, I. A. Bolotina, E. V. Brazhnikov, V. E. Bychkova, R. I. Gil′manshin, Y. O. Lebedev, G. V. Semisotnov, E. I. Tiktopulo, O. B. Ptitsyn, *Eur. Biophys. J.* **1985**, *13*, 109–121.
[49] R. Powers, D. S. Garrett, C. J. March, E. A. Frieden, A. M. Gronenborn, G. M. Clore, *Biochemistry* **1992**, *31*, 4334–4346.
[50] N. Greenfield, G. D. Fasman, *Biochemistry* **1969**, *8*, 4108–4115.
[51] V. De Filippis, P. Polverino de Laureto, N. Toniutti, A. Fontana, *Biochemistry* **1996**, *35*, 11503–11511.
[52] N. J. Marianayagam, F. Khan, L. Male, S. E. Jackson, *J. Am. Chem. Soc.* **2002**, *124*, 9744–9750.
[53] S. A. Gorski, A. P. Capaldi, C. Kleanthous, S. E. Radford, *J. Mol. Biol.* **2001**, *312*, 849–863.
[54] A. W. M. Rietveld, S. T. Ferreira, *Biochemistry* **1998**, *37*, 933–937.
[55] K. Kuwajima, *Proteins* **1989**, *6*, 87–103.
[56] O. B. Ptitsyn, R. H. Pain, G. V. Semisotnov, E. Zerovnik, O. I. Razgulyaev, *FEBS Lett.* **1990**, *262*, 20–24.
[57] M. M. Santoro, D. W. Bolen, *Biochemistry* **1988**, *27*, 8063–8068.

Received: April 22, 2005
Revised: August 30, 2005
Published online: November 11, 2005

