

# Compactness of Protein Folds Alters Disulfide-Bond Reducibility by Three Orders of Magnitude: A Comprehensive Kinetic Case Study on the Reduction of Differently Sized Tryptophan Cage Model Proteins

Dániel Horváth,<sup>[a]</sup> Nóra Taricska,<sup>[a]</sup> Ernő Keszei,<sup>[b]</sup> Pál Stráner,<sup>[a]</sup> Viktor Farkas,<sup>[a]</sup> Gábor K. Tóth,<sup>[c]</sup> and András Perczel<sup>\*[a]</sup>

A new approach to monitor disulfide-bond reduction in the vicinity of aromatic cluster(s) has been derived by using the near-UV range ( $\lambda = 266\text{--}293\text{ nm}$ ) of electronic circular dichroism (ECD) spectra. By combining the results from NMR and ECD spectroscopy, the 3D fold characteristics and associated reduction rate constants ( $k$ ) of E19\_SS, which is a highly thermostable, disulfide-bond reinforced 39-amino acid long exenatide mimetic, and its N-terminally truncated derivatives have been determined under different experimental conditions.

Single disulfide bond reduction of the E19\_SS model (with an 18-fold excess of tris(2-carboxyethyl)phosphine, pH 7, 37 °C) takes hours, which is 20–30 times longer than that expected, and thus, would not reach completion by applying commonly used reduction protocols. It is found that structural, steric, and electrostatic factors influence the reduction rate, resulting in orders of magnitude differences in reduction half-lives ( $900 > t_{1/2} > 1\text{ min}$ ) even for structurally similar, well-folded derivatives of a small model protein.

## Introduction

Structural disulfide (SS) bonds, which are stable in the harsh oxidative extracellular environment, maintain the native fold of proteins by fixing and protecting them from thermal fluctuation induced by elevated internal dynamics. The SS bond formation is perhaps the most fundamental post-translational modification that stabilizes the 3D fold of globular proteins. The absence of regulated SS formation leads to diseases including diabetes,<sup>[1]</sup> cancer,<sup>[2]</sup> neurodegenerative conditions,<sup>[3]</sup> and cardiovascular diseases.<sup>[4]</sup> Non-native SS bond pairing evokes backbone misfolding, which jeopardizes both function and bioactivity, although some proteins may present alterna-

tive SS states and still achieve similarly well-folded forms.<sup>[5]</sup> In protein evolution, the presence of SS bonds shows a significant correlation with the complexity of the organism.<sup>[6]</sup> Approximately 50% of all cysteine residues found in proteins form SS bonds,<sup>[7]</sup> and thus, these cysteine residues become the most conserved among all amino acids, despite being added late to the genetic code during protein evolution.<sup>[8]</sup> Due to the unique pairing pattern of cysteine residues, SS bonds stabilize the 3D fold of proteins unambiguously.<sup>[9]</sup>

Contrary to structural disulfides, redox-active disulfides are highly dynamic, and their formation is reversible. The redox potential of the surrounding environment controls the regulation and cellular localization of these proteins.<sup>[10]</sup> Intramolecular formation of these redox-active disulfides is common for oxidoreductases (thioredoxin<sup>[11]</sup> or glutaredoxin<sup>[12]</sup> family) and allosteric disulfides,<sup>[13–15]</sup> whereas an intermolecular SS linkage results in glutathionylated<sup>[16]</sup> or cysteinylated<sup>[17]</sup> small molecule–protein adducts. The redox potential and stability of the SS bond is highly dependent on several factors, such as the  $pK_a$  of the thiols (the standard  $pK_a$  is 8.5, but this can range from 3.5 to 12.8, depending on the local environment),<sup>[18]</sup> the strain introduced by the SS bond of the protein structure, and the entropic cost of SS bond formation.<sup>[19,20]</sup> The Cys residues of an SS bond are typically distant in the primary sequence; 49% of the SS-bond-forming cysteine residues are more than 25 residues apart from each other.<sup>[21]</sup> The SS-bond formation is thermodynamically more favorable if the cysteine residues are placed in spatial vicinity by the native fold itself before oxidation,<sup>[22]</sup> otherwise—in the absence of chaperones assisting folding<sup>[23]</sup>—the protein precipitates. Adjacent cysteine residues oxi-

[a] D. Horváth, N. Taricska, Dr. P. Stráner, Dr. V. Farkas, Prof. Dr. A. Perczel  
Laboratory of Structural Chemistry and Biology and  
MTA-ELTE Protein Modeling Research Group at the Institute of Chemistry  
Eötvös Loránd University  
112, P.O. Box 32, 1518 Budapest (Hungary)  
E-mail: perczel@chem.elte.hu

[b] Prof. Dr. E. Keszei  
Chemical Kinetics Laboratory, Institute of Chemistry  
Eötvös Loránd University  
112, P.O. Box 32, 1518 Budapest (Hungary)

[c] Prof. Dr. G. K. Tóth  
Department of Medical Chemistry  
Faculty of General Medicine, University of Szeged  
Szeged Dóm tér 8, H-6720 Szeged (Hungary)

Supporting information and the ORCID identification numbers for the authors of this article can be found under <https://doi.org/10.1002/cbic.201900470>.

© 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

dized to a SS bond are rare, although examples can be found among enzymes, receptors, and toxins.<sup>[24,25]</sup> The SS bond or SS bond pattern in prokaryotic proteins is formulated by ribosomal mRNA translation, followed by oxidation and post-translational modifications catalyzed by various enzymes located in the periplasm (DsbC, DsbG, DsbD)<sup>[26]</sup> or cytoplasm (DsbA, DsbB).<sup>[27,28]</sup> In eukaryotic species, this process is performed in specific cell organelles, such as the mitochondria (Mia40, ERV1), endoplasmic reticulum (PDI, ERO1, Erv2), and chloroplasts (PSI, PSII, LTO1, LQY1, CYO1).<sup>[29]</sup>

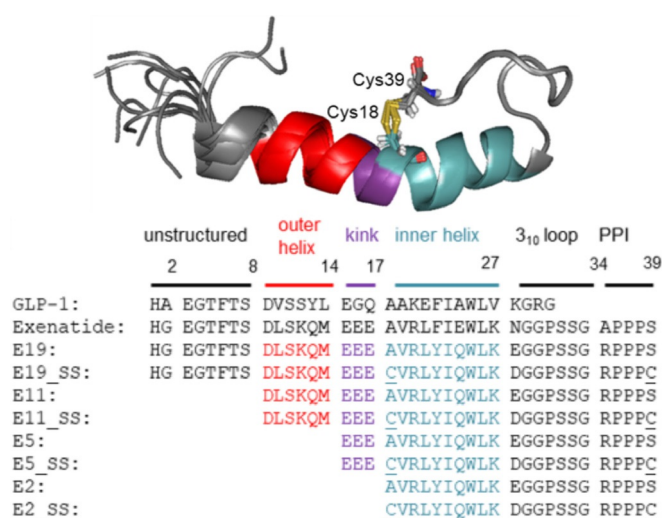
The SS bonds form the core of hundreds of proteins of known 3D structures. Hydrophobic and/or aromatic residues (e.g., Trp and Tyr) may condense around the SS bond and form the network of key interactions that determine the 3D structure of a large number of different proteins.<sup>[21]</sup> In at least 50% of protein families, this type of interaction is invariant. In dozens of proteins (e.g., tick anticoagulant peptide,<sup>[30]</sup> phospholipase A2<sup>[31]</sup>) the SS unit(s) are reinforced by associated aromatic–aromatic interactions<sup>[32,33]</sup> and vice versa. For instance, 92Tyr of RNase-A effectively shields the solvent-exposed nearby SS bond (40Cys–95Cys) from reducing agents (RAs), and thus, helps to maintain the native fold of the protein.<sup>[34]</sup>

If SS bonds are reduced, the thiol groups of the free cysteine residues often adopt an ensemble of local conformers that also loosen the compactness of neighboring residues. In the era of manufacturing recombinant proteins (e.g., insulin), the SS bond cyclized peptides (e.g., vasopressin, oxytocin, desmopressin, octreotide)<sup>[35]</sup> and human monoclonal IgG antibodies,<sup>[36]</sup> which are produced on a large scale by the biopharmaceutical industry, it is vital to have reliable and fully tested methods for SS bond reduction.

In addition to  $\beta$ -mercaptoethanol or 1,4-dithio-D-threitol (DTT), more recently tris(2-carboxyethyl)phosphine (TCEP) has become commonly used as a RA of SS bonds because it is chemically more stable, nonvolatile, odorless, and it reduces SS bonds more effectively, even at low pH.<sup>[37,38]</sup> TCEP is claimed to selectively and completely reduce water-soluble alkyl disulfides over a wide pH range within a few minutes (<5 min).<sup>[39]</sup> Some protocols recommend using 1–100 molar equivalents of TCEP relative to protein concentration.<sup>[40,41]</sup> The reduction time and appropriate temperature greatly depend on the nature of the protein, but, generally, elevated temperature and/or TCEP<sup>[42]</sup> concentration and longer times make the reduction more complete, but these conditions also initiate a multitude of side reactions, which are poorly described, to date.

Exendin-4<sup>[43]</sup> or exenatide<sup>[44]</sup> (synthetic name), which has been used in clinical practice since 2005, is an incretin mimetic<sup>[45]</sup> glucagon-like peptide-1 (GLP-1) analogue, which is a 39-residue peptide with complex physiological actions<sup>[46]</sup> in multiple organs, used in the treatment of type 2 diabetes mellitus.<sup>[47]</sup> Exenatide acts as an agonist of the GLP-1 receptor<sup>[48]</sup> (GLP-1R). Its amphipathic helix binds to the extracellular domain of the GLP-1R, the mainly unstructured N terminus activates the receptor,<sup>[49]</sup> and the structure-stabilizing Trp-cage<sup>[50,51]</sup> fold is not directly involved in interactions to GLP-1R.<sup>[52,53]</sup> We have synthesized and studied the 3D fold of several dozens of Trp-cage folds, including analogues of exenatide, such as E19,<sup>[54,55]</sup>

which is a 39 amino acid protein of comparable bioactivity, but improved water solubility. As a “natural tool” for enhancing the compactness of the 3D fold, we introduced two solvent-exposed Cys residues into E19, making E19\_SS (Figure 1), and



**Figure 1.** A) Structure ensemble of E19\_SS, and B) amino acid sequences of GLP-1; exenatide; parent E19 and its truncated derivatives E11, E5, and E2; their SS analogues E19\_SS, E11\_SS, E5\_SS, and E2\_SS; and their reduced 2SH analogues E19\_2SH, E11\_2SH, E5\_2SH, and E2\_2SH. The position of the SS bridge is highlighted by stick representation and underlined as C. The sequences of E19 is divided into six major parts: 1) 2–8 unstructured N terminus, 2) 9–14 outer helix, 3) 15–17 kink region, 4) 18–27 inner helix, 5) 28–34 3<sub>10</sub> helix, and 6) 35–39 polyproline region. This apportionment of the sequence coincides with the truncation of the peptides.

a loop from residues 18 to 39 in E19\_A18C\_S39C (E19\_2SH). E19\_2SH oxidized to E19\_SS spontaneously with atmospheric O<sub>2</sub> dissolved in water at room temperature. The SS bond of E19\_SS extends the hydrophobic core of the native Trp fold in the spatial proximity of 22Tyr, which is surrounded by explicit negative charges (15Glu, 16Glu, 17Glu). Although E19\_SS is small in size, (MW: 4334.9 g mol<sup>-1</sup>), its reduction takes several hours to reach equilibrium with >10 molar excess of TCEP in water at room temperature. Because the SS bond reduction time turned out to be significantly longer than that expected based on literature data and common laboratory practice, we launched a comparative study, including three designed and truncated analogues of E19\_SS, namely, E11\_SS, E5\_SS, and E2\_SS. Notably, the model systems thus created (Figure 1) oxidize spontaneously and rapidly adopt the Trp-cage 3D fold.<sup>[56]</sup> Moreover, the “loop size” created by the SS bond, in other words, the number of residues between the two reacting cysteine residues, is 20 amino acids long, which is close to the average value ( $\approx 17$ ) observed in thousands of proteins.<sup>[57]</sup>

E11\_SS was designed by removing the “HGEFTFTS” tail, which was the unstructured GLP-1R-activating N-terminal eight residues of E19\_SS. Shortening by an additional six residues removes the outer helical part of E19\_SS, namely, the “HGEFTFTS-DLSKQM” subunit,<sup>[56]</sup> affording E5\_SS. Although 14 residues shorter than that of E19\_SS, E5\_SS still adopts a compact Trp-cage fold and comprises the entire interface for bind-

ing to GLP-1R.<sup>[52]</sup> Finally, in E2\_SS, the entire N terminus preceding 18Cys of E19\_SS was omitted, namely, “HGEGTFTS-DLSKQ-EEE” was cleaved, to give a folded protein with a fully exposed SS bond at its surface; this is considered to be a construct ready for a rapid SS bond reduction (Figure 1).

Herein, we discuss the structure and properties of both the oxidized and reduced forms of the four model proteins of different  $\alpha$ -helical lengths, in comparison with the parent (Cys-free) miniproteins, and the kinetics of reduction. We introduce spectroscopic approaches that make the monitoring of the reduction progress fast and easy. The effect of the compactness of the protein fold, the accessibility, and the local explicit charges of the SS bond and the reagent type on reduction rate and the mechanism are also explained herein.

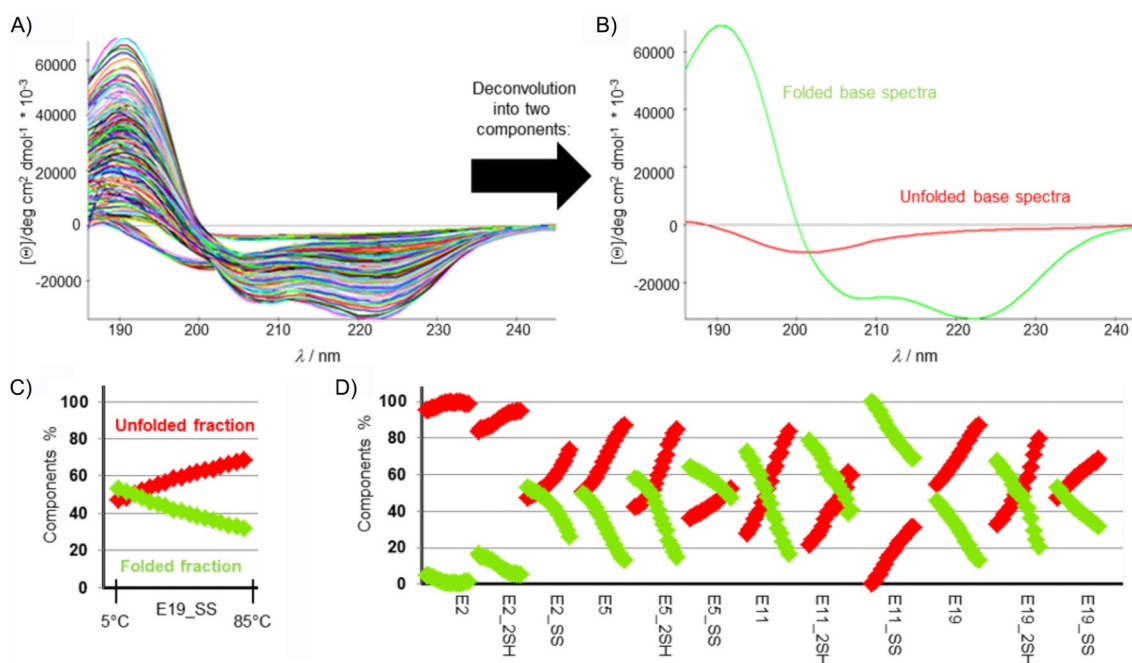
## Results and Discussion

### Three-dimensional fold characterized by far-UV electronic circular dichroism (FUV-ECD) spectra

Circular dichroism (CD) spectroscopy is increasingly recognized as very sensitive indicator of protein conformation,<sup>[58,59]</sup> relying on a plethora of electronic transitions. The FUV-ECD spectra of Trp-cage proteins (e.g., Exenatide, E19, E19\_SS) are typically the weighted sums of the C- (folded and highly helical) and U-type (unfolded) base curves (Figure 2A), as assigned and verified by means of NMR spectroscopy.<sup>[60,61]</sup> As the temperature increases, the shape of the FUV-ECD spectra changes: those of the parent proteins—E2, E5, E11, and E19—acquire more and more U-type characteristics, as they unfold gradually. The tem-

perature-dependent FUV-ECD spectra for all four SS bond enforced model peptides were recorded between 5 and 85 °C (in steps of 5 °C, resulting in 17 spectra for each protein; Figure S1 in the Supporting Information). Aside from E2\_SS, the SS-bond-containing mutants have similar FUV-ECD spectra to that of their parent proteins at low temperatures. On the other hand, because the SS bond makes the 3D folds of SS variants more rigid, they preserve their C-type characteristics better and delay unfolding, even at higher temperatures. Once the SS bond is reduced (see below for details), the spectral properties of the SH variants revert to those of the parent proteins. Their 3D-scaffold compactness decreases as the temperature increases; this is less apparent in the case of E2 and E2\_2SH because they both already present an ensemble of dynamic backbone structures at 5 °C.

Ensemble deconvolution<sup>[62,63]</sup> of the 204 (12×17) ECD spectra,  $f(\lambda, T)$ , made the quantitative analysis of the relative abundance of secondary structural elements belonging to each peptide in each state possible because the pure ECD curves were successfully assigned.<sup>[60,64–66]</sup> The results in Figure 2B indicate that 1) the SS bond stabilizes the less folded protein scaffolds more effectively, for example, whereas the difference at 4 °C between the E2 and E2\_SS folded fraction is 48%, the same difference between E5 and E5\_SS is 14%; in the case of E11 and E11\_SS, it is only 28%, and for E19 and E19\_SS it is 7% (Figure 2D). 2) The ratio of the folded, helical components increases upon going from E2 to E5 and E11; however, the compact  $\alpha$ -helical content of E19\_SS, E19\_2SH, and E19 is lower than those of E11\_SS, E11\_2SH, and E11 because the unfolded eight-residue-long N-terminal part elevates the overall



**Figure 2.** A) Temperature-dependent FUV-ECD spectra (204 in total) of the four primal peptides (E2, E5, E11, and E19) and their four reduced (\_2SH) and four oxidized (\_SS) variants. B) The two pure ECD curves were derived from the ensemble analysis of the 204 ECD spectra by using CCA+. Pure component 1 (red) represents that of the unfolded/U-type, whereas component 2 (green) represents the folded/C-type backbone structure. C) The associated relative proportions [%] of the two pure components at each measured temperature are given for E19\_SS as an example, D) as well for each 204 spectra starting from E2 (at 5 °C) up to E19\_SS (at 85 °C).

backbone dynamics, which destabilizes the compact 3D fold. 3) All four reduced proteins (E2\_2SH, E5\_2SH, E11\_2SH, E19\_2SH) have a higher helix content ( $\approx 7\text{--}15\%$ ) than that of the parent proteins over the entire temperature range. 4) The 3D folds stabilized by SS bonds are less sensitive to temperature (Figure 2 C, D).

### Three-dimensional folds of proteins determined and characterized by NMR spectroscopy

The ensemble of the temperature-dependent FUV-ECD spectra confirms that SS bonds preserve the fold of the model proteins and increase thermostability. NMR spectroscopy analysis at  $15^\circ\text{C}$  allowed further characterization of the 3D structures of each variant. Fold, chemical shift, and secondary chemical shift (SCS) information<sup>[67]</sup> were derived from the appropriate 2D homonuclear NMR spectroscopy experiments ( $^1\text{H}, ^1\text{H}$  COSY,  $^1\text{H}, ^1\text{H}$  TOCSY, and  $^1\text{H}, ^1\text{H}$  NOESY) at  $T=15^\circ\text{C}$ ; the ensemble of the ten lowest energy structures were analyzed. This comprehensive analysis conducted at  $15^\circ\text{C}$  provided the following useful structural descriptors: the root-mean-square deviation (RMSD) of the 3D structures, the average chemical-shift deviation (CSD) of backbone  $\text{H}\alpha$  protons per residue ( $[\sum\text{CSD}_{(i)}^{\text{H}\alpha}]/i$ ) in the helical segment and the compactness of the Trp-cage core by SCS sum of selected protons:  $\text{CSD}_{\text{cage}}$  (Table 1).

A comparison of the helices of different lengths is more straightforward if the helical segment is divided into three parts: 1) the outer  $\alpha$ -helix, 2) the kink region in the vicinity of the SS bond, and 3) the inner  $\alpha$ -helix (Figure 3).

The outer helix compactness seems to be affected by the length of the  $\alpha$ -helix. Interestingly, this part of E11 variants is slightly more structured, in spite of being the terminal part

Table 1. Selected measures characterizing the degree of folding of the model protein ( $T=15^\circ\text{C}$ and $\text{pH } 7$ ).					
		Degree of the fold by FUV-ECD [%] <sup>[a,b]</sup>	Backbone RMSD [ $\text{\AA}$ ] <sup>[c]</sup>	$\text{CSD}_{\text{cage}}$	$[\sum\text{CSD}_{(i)}^{\text{H}\alpha}]/i$
E19	_SS	51.6	0.7	11.7	0.5
	_SH	65.8	0.7	11.1	0.6
	parent	43.7	1.2	10.9	0.5
E11	_SS	96.0	0.3	11.5	0.5
	_SH	76.0	0.3	11.0	0.6
	parent	70.0	1.5	10.9	0.5
E5	_SS	63.4	0.2	11.4	0.5
	_SH	42.8	0.1	10.4	0.5
	parent	51.5	1.6	10.3	0.4
E2	SS	51.8	0.1	11.3	0.5
	_SH	14.8	0.3	9.6	0.4
	parent	3.7	1.5	3.8	0.3

[a]  $T=15^\circ\text{C}$ ,  $c_{(\text{protein})}=20\text{--}30\ \mu\text{M}$  at  $\text{pH}\approx 7$  (typical conditions applied for CD measurements). [b] Calculated % from the joint deconvolution (CCA+) of 204  $T$ -dependent FUV-ECD spectra (Figure 2). [c]  $T=15^\circ\text{C}$ ,  $c_{(\text{protein})}=0.8\text{--}1.8\ \text{mM}$  at  $\text{pH}\approx 7$  (typical conditions applied for  $^1\text{H}$  NMR spectroscopy measurements). [d] RMSD of all backbone atoms of the ten best structures. [e] Xf-cage values<sup>[51,55]</sup> were used to correlate the fold of the protein. The following "H" atoms were involved in calculations: W25H $\epsilon$ 1, L26H $\alpha$ , G30H $\alpha$ 2, P31H $\beta$ 2, R35H $\alpha$ , P37H $\alpha$ , P37H $\beta$ 2, P38H $\delta$ 1, and P38H $\delta$ 2. [f] The average CSD of backbone  $\text{H}\alpha$  protons per residue.

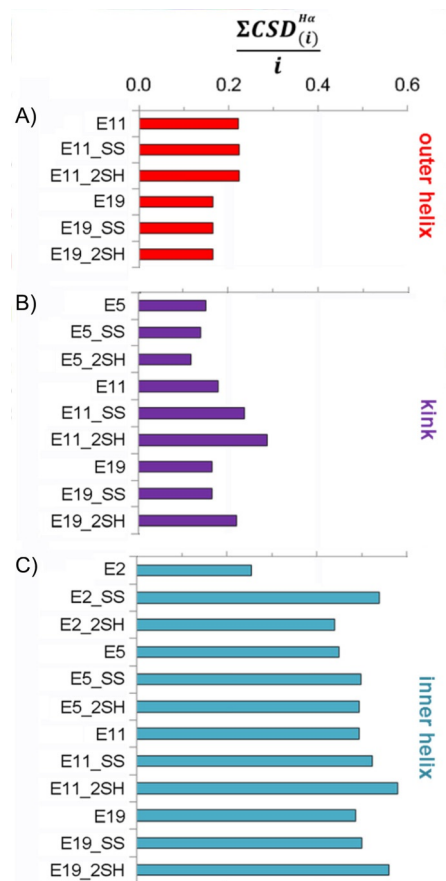


Figure 3. The average CSD of backbone  $\text{H}\alpha$  protons per residue ( $[\sum\text{CSD}_{(i)}^{\text{H}\alpha}]/i$ ) in the three different helical regions: A) outer helix ( $i=6$ ), B) kink region ( $i=3$ ), and C) inner helix ( $i=10$ ). Higher residual values imply a more structured  $\alpha$ -helix.

(usually flexible and unstructured), as opposed to the outer helices of the E19 variants, where this helical segment is flanked (Figure 3A). The above tendency is also true for the kink region, but here the presence and state (\_SS or \_2SH) of the SS bond are also differentiated (Figure 3B). These distant helical parts have generally lower  $[\sum\text{CSD}_{(i)}^{\text{H}\alpha}]/i$  values than those of the inner helix. The compactness of the inner helices is similar (expect for E2). Interestingly, reduced longer polypeptides show slightly increased  $[\sum\text{CSD}_{(i)}^{\text{H}\alpha}]/i$  values, which may be the indicative of ring tension in the SS bond cyclized variants in these systems (Figure 3C).

$^1\text{H}$  NMR spectroscopy studies also confirm that all model proteins, except E2, have a common, compact, and folded Trp-cage core structure at  $T=15^\circ\text{C}$  (Tables 1 and S1 and Figure S2), regardless of the differently structured tails attached to them (Table 1). E2 is predominantly unfolded, even at low temperature ( $15^\circ\text{C}$ ), but because the SS bond joins together the N and C termini of E2\_SS, the hydrophobic core folds properly. Interestingly, even E2\_2SH forms a more compact Trp cage than that of E2. In agreement with data reported in the literature, cysteine residues promote and stabilize  $\alpha$ -helices, if located at their N termini.<sup>[68]</sup> The cage values of the longer reduced peptides are close to that of their oxidized counterparts (Table 1).

NMR spectroscopy data reveal that a longer  $\alpha$ -helix results in a more structured Trp cage, in all cases studied.

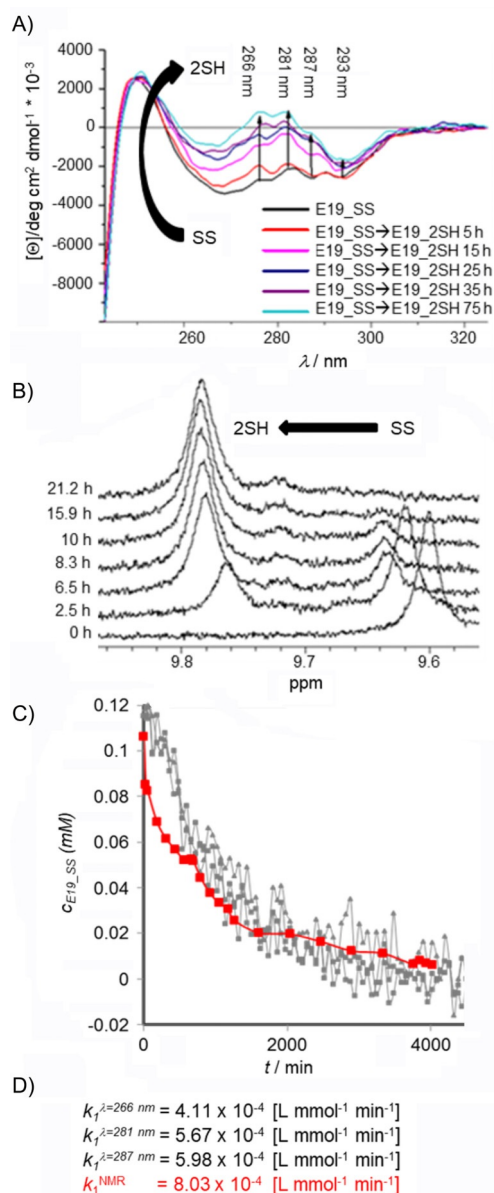
In general, it seems that the core of the reduced (SH<sup>-</sup>) proteins is almost as well folded as those that are SS bonded. The following 3D fold compactness order was established:  $CSD_{cage}^{SS} > CSD_{cage}^{2SH} > CSD_{cage}^{parent}$ , but the differences are small, aside from those of E2 ( $CSD_{cage}^{E2} = 3.8$ )  $\rightarrow$  E2\_SS ( $CSD_{cage}^{E2\_SS} = 11.3$ ).

### Oxidized and reduced states defined by near-UV (NUV) ECD data

As shown above, reduction does not have a dramatic effect on the tertiary structure content of the model systems at room temperature; thus, to detect reduction, NUV-ECD spectra (instead of FUV) had to be used. The interpretation of the changes to the observed chiroptical properties of the Trp/Tyr/SS  $\rightarrow$  Trp/Tyr/2SH (Figure 4A) complex chromophore system is less straightforward because the assignment of “pure” NUV-ECD spectra has not yet been completed. The conformation-dependent fine structure of Tyr/Trp chromophores<sup>[60]</sup> ( $260 \leq \lambda \leq 320$  nm) comprises the <sup>1</sup>Lb of Tyr ( $\lambda \approx 276$  nm, with a shoulder at  $\lambda \approx 287$  nm), <sup>1</sup>Lb of Trp ( $\lambda \approx 281$  and  $\approx 293$  nm), and <sup>1</sup>La of Trp transitions appearing as superimposed broad bands. In addition, the SS bond may also contribute in form of a relatively weak but broad band with a maximum at  $\lambda \approx 260$ –270 nm. For the current proteins with SS bonds, a larger negative band was recorded (Figure S3). The bands of Trp and Tyr in the SS-bond-containing proteins shifted to the negative ellipticity range, which was not observed in the case of the parent proteins (E2, E5, E11, E19),<sup>[60]</sup> for which the bands of these amino acids were detected in the positive range (except the Trp band at  $\lambda \approx 293$  nm). The reduction kinetics of E19\_SS  $\rightarrow$  E19\_2SH were monitored over time as the band intensities at  $\lambda \approx 266$ , 281, 287, and 295 nm increased from larger negative to smaller negative and/or positive values, similar to those of the parent proteins (Figures 4A and S3). We were encouraged to use NUV-ECD spectral changes to monitor SS to SH reduction in proteins if embedded in a suitable molecular environment such as that of the Trp cage motif.

Due to the acidic nature of TCEP, to avoid any pH shift, a phosphate buffer (50 mM, pH 7) was used to sustain near-physiological pH. The groups of Han<sup>[39]</sup> and Whitesides<sup>[37]</sup> described the chemical instability of TCEP above pH 7 in 300–400 mM phosphate buffer. They found that the autoxidation of TCEP depended on how the reagent was stored (open air or in capped vials), whether the solution is stirred, and on the elapsed time ( $24 \leq t \leq 72$  h) during storage. However, herein, we have monitored TCEP stability by means of <sup>31</sup>P NMR spectroscopy in 50 mM phosphate buffer, and found no significant spectral changes connected to TCEP oxidation or degradation at room temperature over 14 days.

Reduction of the E19\_SS protein was followed by recording NUV-ECD spectra ( $\approx 0.113$  mM E19\_SS, pH 7, 15 °C, 18-fold excess of TCEP, cell length = 10 mm) at four different wavelengths (266, 281, 287, and 293 nm). Thus, by following band intensity changes of selected (one or more) <sup>1</sup>Lb transitions of Tyr or Trp, we could monitor the redox state of the SS/SH



**Figure 4.** A) NUV-ECD spectral changes measured for the reduction of E19\_SS ( $\approx 0.113$  mM E19\_SS, pH 7, 15 °C, 18-fold excess of TCEP) at four different wavelengths ( $\lambda = 266$ , 281, 287, and 293 nm). No spectral changes were observed after about 55 h (3300 min). B) <sup>1</sup>H NMR spectra of E19\_SS  $\rightarrow$  E19\_2SH reduction ( $c \approx 0.115$  mM, pH 7, 15 °C, 18-fold excess of TCEP) in water. The chemical shift of the indole H $\epsilon$ 1 of Trp25 was used to monitor reduction: H $\epsilon$ 1 shifted upfield from  $\delta = 9.60$  (SS) to 9.78 ppm (2SH during the reduction). Reaching steady state, the integral ratio of E19\_2SH and E19\_SS was found to be 92 to 8%. C) Concentration change of E19\_SS [mM] measured during reduction by different approaches plotted as a function of time. D) The calculated rate constants (see the discussion of modeling reduction kinetics).

groups and determine the “end point” as a steady state. Thus, if a suitable aromatic residue (Tyr, Trp, Phe) is coupled to the SS bond as a chromophore, it enables its reduction/oxidation state to be monitored, even if the molecular system shows no coupled backbone conformational changes ( $CSD_{cage}^{E19\_SS} = 11.66$ ;  $CSD_{cage}^{E19\_2SH} = 11.07$ ). The measured absorbance was converted into concentration by using Equation (1):

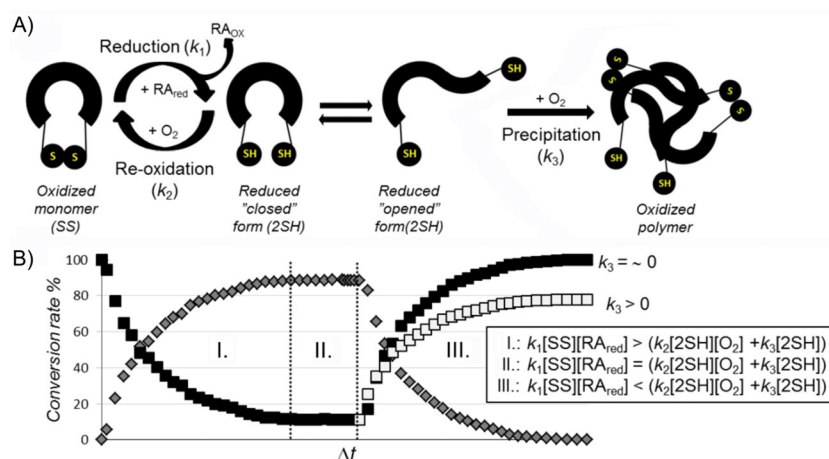
$$c(t) = \frac{A_{\infty} - A}{A_{\infty} - A_0} [SS]_0 \quad (1)$$

Steady state was reached conclusively after about 55 h. We determined the rate constant,  $k_1$ , at each wavelength by parameter estimation to be  $k_1^{\lambda=266\text{ nm}} = 4.11 \times 10^{-4} \text{ Lmmol}^{-1} \text{ min}^{-1}$ ,  $k_1^{\lambda=281\text{ nm}} = 5.67 \times 10^{-4} \text{ Lmmol}^{-1} \text{ min}^{-1}$ , and  $k_1^{\lambda=287\text{ nm}} = 5.98 \times 10^{-4} \text{ Lmmol}^{-1} \text{ min}^{-1}$  (Figures 4C and S4). The deviations of the fitted model from the measured data at  $\lambda = 293 \text{ nm}$  were remarkably large; therefore, parameter estimation was not performed on this dataset. NUV-ECD monitoring enables one to observe the clean and clear changes in the spectra, but it does not make it possible to extract the absolute value of the concentration,  $[SS]_{\text{ocr}}$  at the end point of the reaction. Based only on the intensity of the molar ellipticity, it cannot be decided if reduction is fully completed or not. To ascertain the absolute values of the concentrations in the redox system, reduction was repeated under the same conditions in NMR tubes with a diameter ( $\varnothing$ ) of 5 mm ( $\approx 0.113 \text{ mm}$  E19\_SS, pH 7, 15 °C, 18-fold TCEP) by recording  $^1\text{H}$  NMR resonances (Figure 4B). By using both SS and SH state integrals of the signals at selected resonance frequencies (e.g.,  $\text{He}^{\text{TPP}}$ ),  $^1\text{H}$  NMR spectroscopy driven quantitative analysis of the reduction was performed (Figure 3E) and the rate constant was determined to be  $k_1^{\text{NMR}} = 8.03 \times 10^{-4} \text{ Lmmol}^{-1} \text{ min}^{-1}$ . Although 18-fold excess of TCEP was used,  $^1\text{H}$  NMR spectroscopy data showed that, at steady state,  $\partial[\text{E19\_SS}]/\partial t \approx 0$  and  $\partial[\text{E19\_2SH}]/\partial t \approx 0$ , reduction was incomplete and about 8% of E19\_SS remained oxidized. A comparison of the calculated reaction rates of the two methods (NMR and CD spectroscopy) shows that not only are the orders of magnitudes the same, but the values are also quite similar. The reduction rate of NUV-ECD measured at  $\lambda = 287 \text{ nm}$  is closest to that of  $k_1^{\text{NMR}}$  (Figure 4C). Monitoring the intensity of the molar ellipticity by NUV is a fast and efficient method to define the end of the reaction. It also provides an approximate

value of the reduction rate if the conversion is close to completion. Based on  $^1\text{H}$  NMR spectroscopy integrals, it is possible to determine the rate of the conversion and obtain evidence for the reversibility of the redox system. Taking into account incomplete conversion, despite the presence of the 18-fold excess of RA, the role of dissolved oxygen and reoxidation should also be included in the kinetic mechanism.

### Concept of the reversible redox system

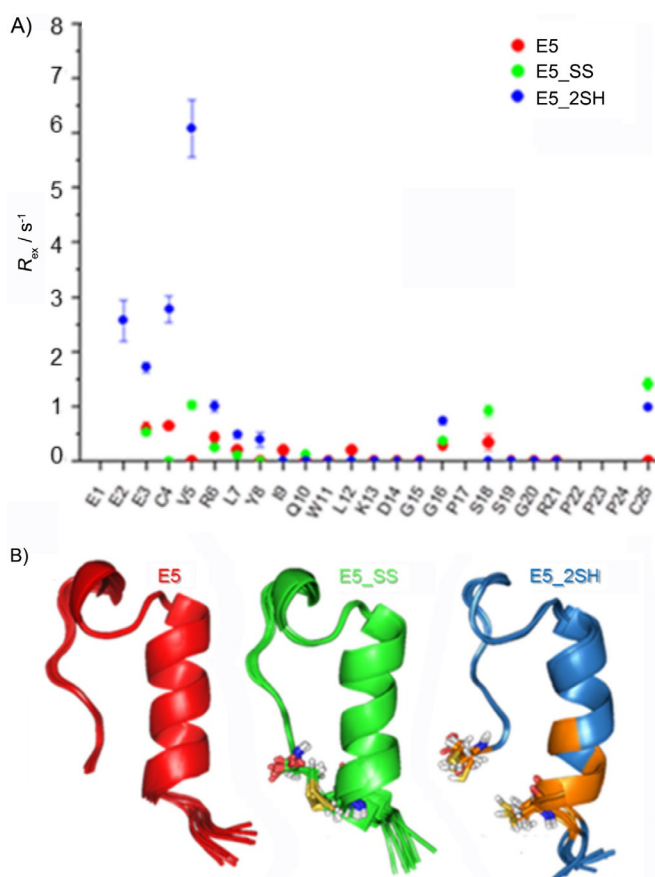
Physiological solutions contain dissolved  $\text{O}_2$  from the air, and thus, Cys-SH groups of any protein might oxidize spontaneously to form the SS bond(s). The apparent rate constant depends on several microequilibrium constants, which are explicitly not elaborated herein.<sup>[69]</sup> However, it certainly depends on the width of the conformational space of the reduced molecular fold. Furthermore, the concentration of dissolved  $\text{O}_2$  (and thus,  $T$  and  $p$ ), the diffusion rate of TCEP, and the protein concentration are all rate-influencing factors. Because our model protein forms a coupled reaction cycle, once E19\_2SH is reduced by excess TCEP, E19\_SS will be instantaneously reoxidized by dissolved  $\text{O}_2$  (Figure 5). Before exploring the mechanism of these redox-cycle-related electron-transfer processes, it should be noted that, at a macroscopic level, these coupled cycles remain hidden, as steady state ( $\sum \partial \xi / \partial t = 0$ ) is reached. Reduction concludes in a "normal way" if all dissolved  $\text{O}_2$  is consumed; however, if the concentration of the RA declines faster than that of  $\text{O}_2$ , then oxidation will dominate the process and spectral properties will change accordingly (Figure 5). It is hard to a priori predict the end point of the latter process because, unlike the oxidized fold of a protein, the reduced one could have a multitude of backbone conformers in exchange at various timescales (e.g.,  $\mu\text{s}$  to  $\text{ms}$ ). Among these 3D folds of the reduced state, the "closed-SH" forms (Figure S2II), in which both the C and N termini are close to each other, lead only to



**Figure 5.** A) A schematic illustration of the redox cycle. The oxidized state (SS) in the presence of RA (e.g., TCEP) becomes reduced (2SH), in which open and closed conformers are present in equilibrium. In the case of the closed conformer, in which the -SH groups are closely fixed to each other, intramolecular re-oxidation can occur in the presence of  $\text{O}_2$ , whereas the open conformer is more likely to aggregate due to intermolecular interactions. B) Three stages of the theoretical redox setups provide the state at which reduction dominates the overall process (I), a steady state (II), and a state (III) in which excess dissolved  $\text{O}_2$  and the absence of RA lead to oxidation back to the reduced state. The black square denotes the relative concentration of the oxidized form; gray diamonds represent the reduced form. If precipitation occurs ( $k_3 > 0$ ), then at the end point of the redox cycle the soluble protein concentration has decreased relative to the initial one.

intramolecular reoxidation. However, if “open-SH” backbone forms become highly populated (e.g., as is the case of E2\_2SH; Figure S2/XIII), then intermolecular oxidation will be more prevalent, giving rise to oligo- and polymer formation (see below).

Capturing internal backbone dynamics occurring on the timescale of micro- to milliseconds was successfully attempted by means of Carr–Purcell–Meiboom–Gill (CPMG) NMR spectroscopy.<sup>[70]</sup> Herein, we present the characterization of E5, E5\_SS, and E5\_2SH as examples. We found that only the backbone NH groups of Glu3, Cys4, Val5, Arg6, Tyr8, and Cys25 of E5\_2SH partake in such slow motion. Considering the fact that all of these NH groups are close to both Cys residues (Figure 6), the CPMG data suggest that either E5\_2SH presents alternative backbone structures, which interconvert at a slow exchange rate, or, due to incomplete reduction, the remaining oxidized form (1–8%, see a discussion of the conversion rate below) constantly interconverts with the reduced form. The minor amount of coexisting oxidized form (E5\_SS) could contribute to the stabilization of the dominant backbone fold of E5\_2SH. The conformational equilibrium between the oxidized and reduced states seems to be the most likely explanation for the above-described slow exchange; however, both scenarios of motion can occur in a concerted way.



**Figure 6.** A) CPMG-determined NH  $R_{ex}$  values of E5 (red), E5\_SS (green), and E5\_2SH (blue), and B) their backbone structures, with the key Cys residues highlighted. Slow exchange was measured for backbone NH groups of E3, C4, V5, R6, Y8, and C25 of E5\_2SH only. Notably, residues that give  $R_{ex}$  are in the proximity of the Cys residues colored orange.

## Modeling of the SS bond reduction kinetics

The SS bond reduction by TCEP is a bimolecular nucleophilic substitution ( $S_N2$ ) reaction.<sup>[37]</sup> Thus, both the concentration of the oxidized form of the protein [ $SS$ ] and that of the RA contribute to the rate of the reduction. In an ideal case, we should consider only nucleophilic attack of the RA ( $k_1$ ), but, as we explained previously, in practice, we also have to take into account back oxidation ( $k_2$ ), which takes place simultaneously, and, in some cases, depending on the size and shape of the protein, precipitation ( $k_3$ ; Figure 5). The mechanism of reduction, therefore, can be described by Equations (2), (3), and (4):



By fitting this model to the concentration–time functions determined by means of NMR spectroscopy,  $k_1$ ,  $k_2$ , and  $k_3$  can be determined, and half-lives can be calculated. We focused on the determination of the reduction rate constants  $k_1$ ; therefore, sampling was more frequent in the reduction phase (stage I; Figure 5B). Based on parameter estimation (see the Experimental Section),  $k_2$  and  $k_3$  are very often either negligibly small, or, due to a lack of sufficient data, cannot be confidently estimated. Obtaining key kinetic parameters allowed us to describe and compare the reduction kinetics of the SS-containing mini-proteins under various experimental conditions. Some protocols reported in the literature apply extreme conditions, such as high temperature (e.g., 50–80 °C), to obtain short reduction times; this is clearly unsuitable for maintaining the integrity of the protein, or > 20-fold molar excess of reagent. By performing the reduction of E19\_SS (0.8 mM) under such conditions (60 °C with 18-fold TCEP excess), the reaction seemed almost instantaneous ( $t_{1/2} < 5$  min), but the sample became opalescent and side reactions (e.g., precipitation) were instantly detected. Similarly to most globular proteins, the conformational ensemble of E19\_2SH at 60 °C is distinctly different from that of 15 °C; thus presenting many more unfolded states. The folded fraction of E19\_2SH is 64% at 15 °C, whereas it is 41% at 60 °C, according to FUV-ECD analysis. Instead of intramolecular reoxidation, undesirable intermolecular reoxidation might occur between particles. (Reducing E19\_SS for 120 min, followed by centrifugation gave practically zero soluble protein concentration.) In general, reduction and reoxidation at higher  $T$  (e.g.,  $\geq 60$  °C) is expected to be less effective, and accompanied by multiple side reactions, such as  $\beta$ -elimination<sup>[71]</sup> (which already occurs at a lower  $T$ ),<sup>[72,73]</sup> racemization,<sup>[74,75]</sup> and aggregation. In principle, the reduction rate can be enhanced at lower  $T$  by increasing the TCEP molar ratio (15–20-fold molar excess); however, this also triggers obscure unwanted processes (Figure S5). Experiments were repeated at different temperatures (15, 25, and 37 °C) with 0.8 mM protein and 18-fold excess of TCEP (Table 2 and Figure S6). The Arrhenius equation allows the activation energy ( $E_a$ ) of the redox reaction to be derived, resulting in a value of about 44.3 kJ mol<sup>-1</sup>. For comparison, the activa-

**Table 2.** Kinetic parameters of temperature-dependent E19\_SS reduction with 0.8 mM protein and 18-fold excess of TCEP. For detailed results of parameter estimation, see Figure S6.

<i>T</i> [°C]	Elapsed time to steady state [h]	Conversion rate [%]	<i>k</i> <sub>1</sub> [L mmol <sup>-1</sup> min <sup>-1</sup> ]	<i>t</i> <sub>1/2</sub> [min]	Relative standard deviation of <i>k</i> <sub>1</sub> [%]
15	≈ 15	92	3.05 × 10 <sup>-4</sup>	181	3.27
25	≈ 6	94	7.68 × 10 <sup>-4</sup>	72	9.78
37	≈ 4	94	1.15 × 10 <sup>-3</sup>	48	1.68
60 <sup>[a]</sup>	n.d.	n.d.	n.d.	n.d.	n.d.

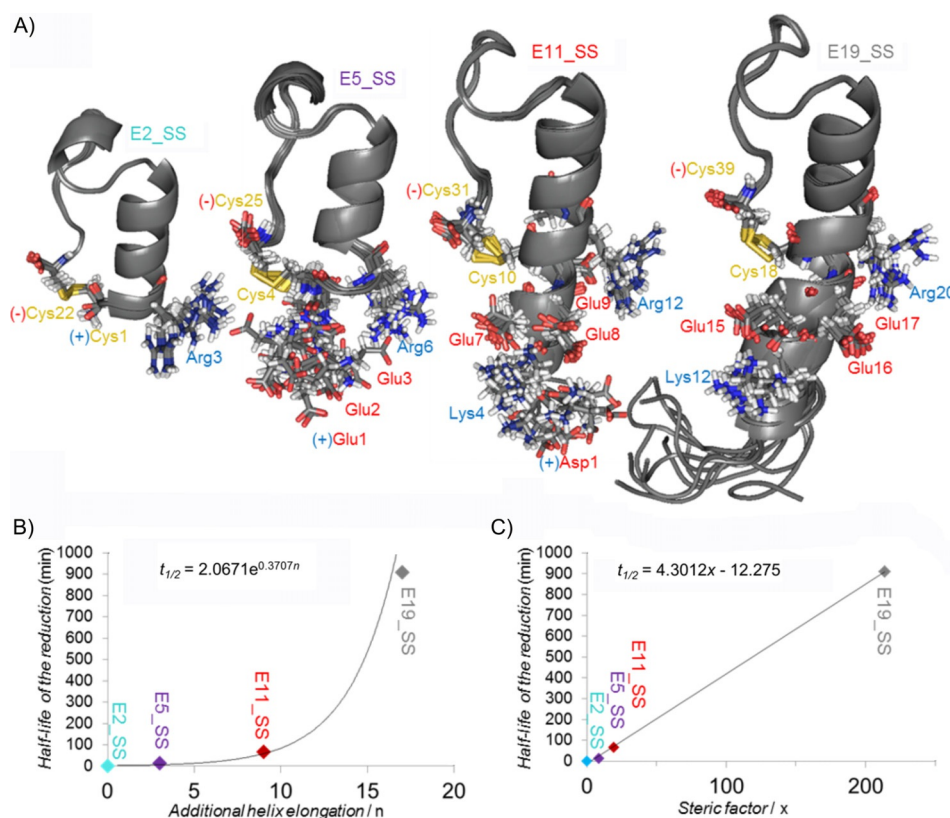
[a] n.d.: not determined.

tion energy of thiol–disulfide exchange between methylthiolate and oxidized DTT was calculated to be 62 kJ mol<sup>-1</sup>.<sup>[76]</sup> Both FUV-ECD and NMR spectroscopy derived structural information support the high conformational similarity between E19\_SS and E19\_2SH; therefore, *E*<sub>a</sub> is likely to be used for the redox reaction, rather than for the conformational switch between the two conformational states (Table 1). Based on the NMR spectroscopy derived signal integral analysis, the reduction was almost complete (≈ 94%) and no sign of precipitation was detected at any temperature. Additional experiments were performed to investigate the effect of the protein/RA ratio as a practical perspective (Figure S7). The above-described NMR

spectroscopy methodology provides high-resolution information about the reduction mechanism, relative to that of the more rapid NUV-ECD approach, and thus, details of the reduction of all four –SS– protein models were obtained through NMR spectroscopy.

### Kinetics of SS bond reduction influenced by steric factors

An appropriate reduction protocol was required to unambiguously determine the 3D structures of the above-introduced pure reduced states. Thus, in agreement with the above discussion, only mild conditions (15 °C and twofold molar excess of TCEP) were used for the reduction of the four different mini-proteins. Determining the structural properties and reduction rates under the same conditions allowed us to elucidate the basis of the observed differences in the reduction rates. We found that, at *T* = 15 °C, the *k*<sub>1</sub> values of these four model proteins, comprising of identical core structures, but different lengths, were indeed different: their *k*<sub>1</sub> and *t*<sub>1/2</sub> values strongly depended on their sizes and/or molecular weights. It appears as if “cutting back” on the α-helical segment strongly affects the SS bond reducibility, even though the SS bonds of all four models are near the surfaces (Figure 7A). To our great surprise, we recorded three orders of magnitude differences between the reduction rate constants (Table 3). Whereas the reduction



**Figure 7.** A) Ten superimposed structural ensembles of E19\_SS, E11\_SS, E5\_SS, and E2\_SS. Notably, all model proteins have their SS bonds at the surface, but their N termini are of different lengths and charges; thus, affecting the SS bond reducibility. Charged residues at pH 7, close to the reaction center are explicitly depicted: the negatively charged side chains are highlighted in red, whereas the positive ones are in blue. The C-terminal negative charge of COO<sup>-</sup> is marked by (–) and the amino group –NH<sub>3</sub><sup>+</sup> of the N terminus by (+). Correlation between the reduction half-life versus B) helix length and C) steric factors (twofold excess of TCEP, 1.7 mM protein, 15 °C) is reported. (Figure S10 shows the correlation between the reduction half-life and helix length in the case of DTT reduction.)



**Table 3.** Kinetic parameters of the SS bond reduction of the four model proteins. For each reduction, a protein concentration of 1.7 mM was used with twofold excess of TCEP and DTT at 15 °C. For detailed results of parameter estimation, see Figures S8–S9. NMR spectroscopy derived structural properties of the outer helix are also shown.

	TCEP				DTT				Properties of outer helix					
	Elapsed time to steady state [h]	Conversion rate [%]	$k_1$ [L mmol <sup>-1</sup> min <sup>-1</sup> ]	$t_{1/2}$ [min]	Relative standard deviation of $k_1$ [%]	Elapsed time to steady state [h]	Conversion rate [%]	$k_1$ [L mmol <sup>-1</sup> min <sup>-1</sup> ]	$t_{1/2}$ [min]	Relative standard deviation of $k_1$ [%]	Outer helix length (i)	RMSD of outer helix	$[\sum \text{CSD}_{(i)}^{\text{H}\alpha}]/i$	Steric factor <sup>[a]</sup>
E19_SS	≈76 h	87	$2.71 \times 10^{-4}$	909	7.30	n.d.	n.d.	n.d.	30 545 <sup>[b]</sup>	n.d.	17	1.41	0.11	213.74
E11_SS	≈5–6 h	94	$3.68 \times 10^{-3}$	67	2.61	138	≈84	$1.52 \times 10^{-4}$	1659	45.152	9	0.55	0.26	19.26
E5_SS	≈1 h	93	$1.85 \times 10^{-2}$	14	3.35	9–10	95	$2.18 \times 10^{-3}$	115	1.064	3	0.39	0.14	8.56
E2_SS	<5 min	100	$2.59 \times 10^{-1}$	≈1	15.37	5	100	$4.04 \times 10^{-3}$	62	5.315	0	0	0	0

[a] Steric factor comprises the following factors: the length of the outer helix, the RMSD, and the reciprocal value of  $[\sum \text{CSD}_{(i)}^{\text{H}\alpha}]/i$ . [b] Half-life of E19\_SS reduction by DTT was calculated according to the equation of the dependence of the half-life on outer helical length.

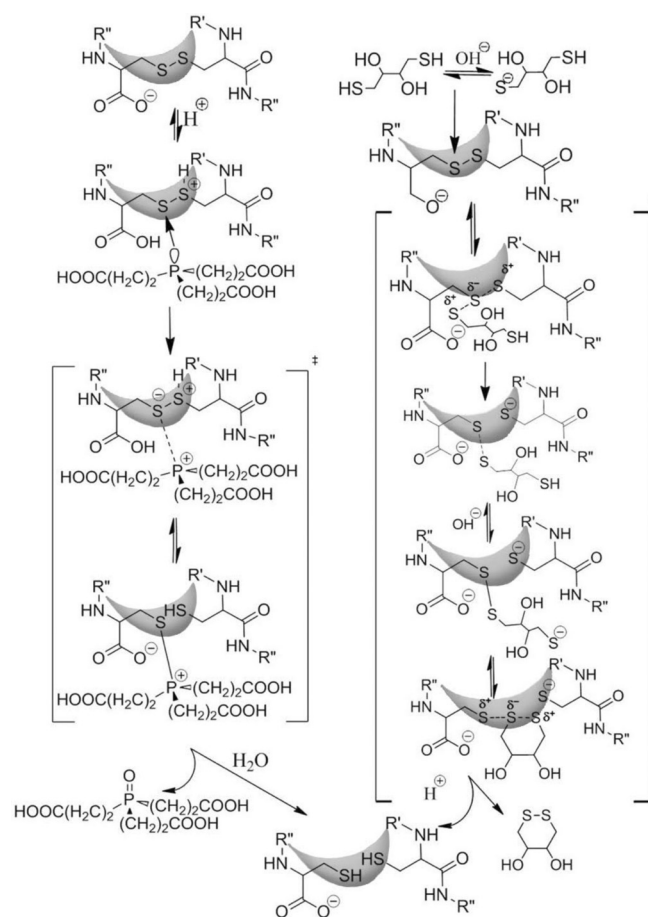
of E2\_SS is still extremely fast,  $t_{1/2}^{\text{E2-SS}} < \approx 1$  min, that of E5\_SS occurs on the timescale of minutes:  $t_{1/2}^{\text{E5-SS}} \approx 14$  min. E11\_SS, which is elongated by six residues (1.5 turns of  $\alpha$ -helix) with respect to that of E5\_SS, exhibits about a fourfold increase in  $t_{1/2}$  ( $t_{1/2}^{\text{E5-SS}} \approx 14$  min  $\rightarrow$   $t_{1/2}^{\text{E11-SS}} \approx 67$  min). Finally, the unstructured short octapeptide tail HAEGTFTS- further lengthens  $t_{1/2}$  by about 13-fold ( $t_{1/2}^{\text{E11-SS}} \approx 67$  min  $\rightarrow$   $t_{1/2}^{\text{E19-SS}} \approx 909$  min). The conversion rate was close to complete for the shorter peptide of E2\_SS, whereas the reduction of E19\_SS was only 88% complete. The kinetic parameters of all four model proteins were determined by using a twofold molar excess of DTT, at pH 7, and  $T = 15$  °C. The mechanism of SS bond reduction by DTT is also  $S_{\text{N}}2$ ,<sup>[77]</sup> but the determined  $t_{1/2}$  values are significantly longer than those obtained by using the same molar excess of TCEP; however, the observed overall tendency and conclusion appear to be the same (Table 3).

Because the well-folded Trp cage motifs are identical (based on their CSD cage values; Table 1) in all four model proteins, the observed  $k_1$  differences must be associated with the structural properties of their  $\alpha$ -helices and the eventually appearing unstructured tail. Although the dataset is limited ( $n = 3$  or 4), as the simplest approach, the length of the  $\alpha$ -helix ( $n$ ) and the half-lives ( $t_{1/2}$ ) of reduction could be correlated, leading to an exponential dependence for both TCEP ( $t_{1/2} = 2.06e^{0.371n}$ ,  $R^2 = 0.95$ ) and DTT ( $t_{1/2} = 50.47e^{0.377n}$ ,  $R^2 = 0.98$ ) as the RA (Figure 7B). To take into account the additional structural descriptors for a more complete characterization, we derived the steric factor ( $x$ ) for these protein models [Eq. (5)]:

$$x = \left( \frac{1}{[\sum \text{CSD}_{(i)}^{\text{H}\alpha}]/i} \right) \times \text{RMSD} \times n \quad (5)$$

in which the reciprocal of the helicity ( $[\sum \text{CSD}_{(i)}^{\text{H}\alpha}]/i$ ) and the bulkiness (RMSD) of the outer helical part were both calculated with respect to the length of the N terminus ( $n$ ; Table 3). We observed a linear dependence of the steric factors on the reduction half-lives as a function of the length of the N terminus

(Figure 7C). Some, but not all, of the above  $k_1$  ( $t_{1/2}$ ) differences can be explained by structural differences of the outer helix because both solvent exposure and local charges around the



**Scheme 1.** A generalized mechanism of TCEP- and DTT-assisted mechanisms of SS bond reduction in proteins. Functional group R- stands for the N terminus of the protein systematically elongated here: in E19\_SS the R group is equal to H<sup>+</sup>-HGEGTFTSDLSKQMEEE-, in E11: H<sup>+</sup>-DLSKQMEEE-, in E5\_SS: R = H<sup>+</sup>-EEE-, and in E2\_SS it is simply H<sup>+</sup>. A brief description of the detailed reaction mechanism is provided for both TCEP and DTT in the text.

SS bonds are also different. In Scheme 1, we provide a summary of the mechanistic explanation, including all of these factors and viewpoints.

### Rate-determining steric and electronic factors of SS bond reduction

Apart from the steric effect of the helical part emphasized above, the  $S_N2$  mechanism of TCEP-driven reduction has to be discussed in terms of electrostatic effects.<sup>[19]</sup> In general, attack is more favorable and effective on those structures in which the C terminus is neutral. According to the average  $pK_a$  of the cysteine carboxyl group ( $pK_a = 1.92$ ) at pH 7, the proportion of  $\text{COOH}/\text{COO}^-$  is low: 1/12000. The rate-determining step is cleavage of the SS bond.<sup>[78]</sup> During the  $S_N2$  reaction, the nucleophilic P atom of TCEP attacks one of the SS bonds, forming a thiophosphonium salt (an  $\text{S}^--\text{P}^+$  ion-pair complex; Scheme 1).

Nucleophilic attack ( $n \rightarrow \sigma^*$ ) is facilitated by the favorable arrow-shaped (tetrahedral:  $105^\circ$ ) steric arrangement of the nonbonding electron pair of the P atom of TCEP. The main portion of the activation Gibbs free energy of reduction is consumed by splitting of the SS bond and not by the steric rearrangement of the intermediate structure.<sup>[79]</sup> Better solvation of the thiol and zwitterion results in a lower activation Gibbs free energy of the reaction. Next, the positively charged  $-\text{S}-\text{P}^+-[(\text{CH}_2)-\text{COOH}]_3$  complex hydrolyzes rapidly and results in the phosphine oxide and free  $-\text{SH}$  groups of the protein.

Both charged and aromatic side chains can participate, and thus, intimately influence the efficacy of TCEP-mediated reduction (Figure 7A). The nucleophilic phosphine attacks the C-proximal cysteine because the intermediate cation can be stabilized by the proximal  $\text{COO}^-$  group of the C-terminal cysteine. A positive charge near the SS bond could enhance the reaction through electrostatic compensation of the N-proximal leaving thiolate group, whereas a negative charge might slow down the  $S_N2$  reaction.<sup>[80,81]</sup> Direct through-bond effects of any charged side chain can be ignored because they are separated by several  $\sigma$  bonds from the negative  $\text{COO}^-$  group. Although the inductive or direct  $\sigma$ -bond effects are negligible, both steric and spatial electrostatic effects in the vicinity of N-proximal cysteine play a major role in the reduction rate. At pH 7, the positively charged Arg near the SS bond in the inner helix may facilitate reduction; however, it is distant from the SS bond (Figure 7A), and thus, a direct charge-controlled interaction is less likely to occur. On the other hand, the positively charged N-terminal  $-\text{NH}_3^+$  can directly catalyze the instantaneous reduction<sup>[82]</sup> of E2\_SS ( $t_{1/2}^{\text{E2\_SS}} \approx 1$  min) because  $\text{H}-\text{N}-\text{C}^\alpha-\text{C}^\beta-\text{S}$  of the cysteine forms a five-membered pseudo-ring that facilitates intramolecular NS proton transfer.<sup>[83]</sup> Thus, upon TCEP attack, these ideal local electrostatic compensations may stabilize the intermediate thiophosphonium salt, shifting the reaction equilibrium towards splitting of the SS bond. Furthermore, because the leaving thiolate anion is only positioned at the N terminus of the well-folded  $\alpha$ -helix, the positive charge of the  $\alpha$ -helix macrodipole also promotes progress to the reduced state.<sup>[84,85]</sup> Moreover, due to the small protein size, the SS bond is most exposed to solvent and reagent in E2\_SS.

As the N terminus is elongated on the  $\alpha$ -helix from E2\_SS toward E19\_SS, the “catalyzing”  $-\text{NH}_3^+$  group of the N terminus moves further away from the SS bond, and the effect of the macrodipole gradually vanishes; thus, the reduction rate is reduced ( $t_{1/2}$  increases; Table 3). The role of this positive charge was directly probed by acetylating the N terminus, Ac-E2\_SS, and, as expected, the half-life of reduction increased significantly:  $t_{1/2}^{\text{E2\_SS}} \approx 1$  min  $\rightarrow$   $t_{1/2}^{\text{Ac-E2\_SS}} \approx 8$  min (in both cases, a protein concentration of 1.7 mM and twofold excess of TCEP were used).

The N-terminal elongation of E2\_SS by three Glu residues results in E5\_SS. As expected, the reduction rate is slower:  $t_{1/2}^{\text{E5\_SS}} \approx 14$  min. Although only a tripeptide is added to the dynamic N terminus, reaching the SS bond still becomes harder for both reagent and/or solvent molecules. In addition, the 3D structure (Figure 7A) shows that the three negatively charged Glu side chains (at pH 7) are flanked by the N-proximal cysteine and the positively charged N terminus, and thus, effectively neutralize the catalytic effect. The structure of the ensemble determined by means of NMR spectroscopy shows a distance fluctuation from 3.7 to 10.7 Å between 4Cys  $\text{C}\beta$  and 1Glu  $\text{NH}_3^+$ , whereas that of 4Cys  $\text{C}\beta$  and 1Glu  $\text{COO}^-$  fluctuates between 3.4 and 12.4 Å (Figure 7). Thus, SS bond protonation requires an active contribution from the medium; but proton transfer is perturbed by the proximity of the negatively charged glutamate side chains.

Further elongation of E5\_SS by the hexapeptide -DLSKQM- leads to E11\_SS. Under the same conditions, the reduction of this even larger model protein occurs more slowly ( $t_{1/2}^{\text{E11\_SS}} \approx 67$  min). The glutamate side chains are more oriented by the longer  $\alpha$ -helix of E11\_SS (Figure 7): whereas 8Glu<sup>-</sup> turns outward, both 7Glu<sup>-</sup> and 9Glu<sup>-</sup> flank the SS bond from two sides. Residues 7Glu<sup>-</sup> with 4Lys<sup>+</sup> and 9Glu<sup>-</sup> with 12Arg<sup>+</sup> are capable of forming salt bridges in close vicinity, and thus, could partly compensate for the slowing effect of the negatively charged side chains. E11 was found to be more helical than that of longer E19;<sup>[56]</sup> thus we find here that both E11\_SS and E11\_2SH have more compact  $\alpha$ -helices than those of E19\_SS and E19\_2SH, according to both  $[\Sigma \text{CSD}^{\text{H}\alpha(i)}]_i$  NMR spectroscopy measurements and FUV-ECD spectral properties. We believe that, in addition to partly compensated for negative electrostatic effect(s), mainly steric effects of the elongated and stiffer  $\alpha$ -helix cause the longer value of  $t_{1/2}^{\text{E11\_SS}}$  with respect to that of  $t_{1/2}^{\text{E5\_SS}}$ .

Finally, E11\_SS elongated by the -HGEGTFTS- octapeptide results in E19\_SS—the largest model protein used herein—for which the longest half-life ( $t_{1/2}^{\text{E19\_SS}} = 909$  min) is measured. E19\_SS has the same electrostatic pattern in the vicinity of the SS bond as that of E11\_SS, but its reduction rate is about 15 times slower than that of E11\_SS. Although the -HGEGTFTS- segment is far from the SS bond ( $d_{\text{Thr-18Cys}} = 11\text{--}14$  Å; Figure 7) and cannot influence reduction by electrostatic interactions, its higher internal dynamics (low  $S^2$  value),<sup>[56]</sup> as a steric effect, must slow the SS bond reduction rate further. In fact, the latter increase, in terms of  $t_{1/2}$ , is a good estimation of the magnitude of a purely steric effect of an unstructured polypeptide chain on reduction rate.

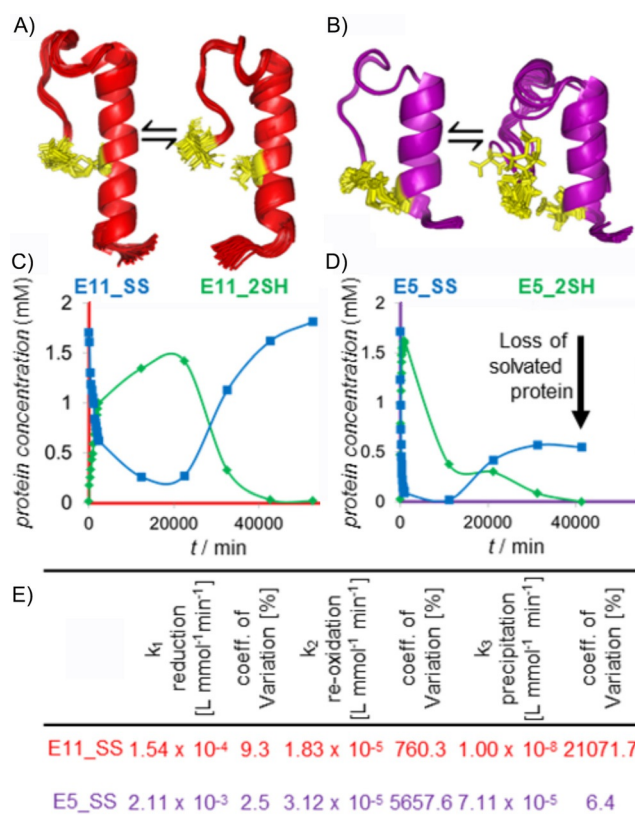
### Differences in reduction kinetics and mechanism with alternative reagents

There are a few distinct differences in terms of the general mechanism of SS reduction by TCEP and DTT (Scheme 1). 1) As an initializing step, deprotonation of the thiol group of DTT is required for successful nucleophilic attack, which depends on the pH of the medium. According to the Henderson–Hasselbach equation,<sup>[86]</sup> taking into account the acidic dissociation constant of DTT ( $pK_{a1}=9.2$  and  $pK_{a2}=10.1$ ) at pH 7, deprotonated thiolate concentration is about three to four times lower than that of the overall DTT concentration. After successful nucleophilic attack on the SS bond, a linear  $-S-S-$  transition complex has to be formed, in which the negative charge is located on the two leaving S atoms.<sup>[87]</sup> An intramolecular protonation, as for TCEP, also stabilizes the thiol anion leaving group if DTT is used, and thus, enhances the reaction rate. Therefore, a positive inductive/steric effect increases, whereas a negative effect decreases the reduction rate. 2) Contrary to TCEP, the active species of DTT has a negative charge. Therefore, charged amino acid side chains close to the SS bond will directly affect attack by the nucleophilic RA. In line with these observations, both the negative C terminus and the SS bond flanking glutamate side chains repel DTT; thus contributing to a significant and large-scale decrease in reaction rate (Table 3). 3) Moreover, complete reduction by DTT consists of two steps: after the first attack, the free SH group of the peptide–DTT complex has to cleave the previously formed SS bond, whereas DTT closes into a six-membered ring (Scheme 1). All of these factors jointly decrease the reduction rate if DTT is used instead of TCEP (Table 3). These considerations make it even more striking that, although several proteins with various numbers of SS bonds per molecule, such as  $\alpha$ -lactalbumin, lysozyme, and oxytocin, were reported to be completely reduced in 5 min by 10 mM DTT at pH 5.5 and 70 °C,<sup>[88]</sup> we found that the reduction of miniproteins (e.g., E11\_SS) with a single and exposed SS bond might take up to 138 h (Table 3).

### Spontaneous SH reoxidation accompanied by polymerization

Incomplete conversion, despite the presence of a large excess of the RA, provided evidence for the reoxidation of the reduced SS bond of the studied model systems. To study this process in detail, the in situ reoxidation of the DTT-reduced protein samples at room temperature in sealed NMR tubes (pH 7, 15 °C, twofold excess of DTT) was monitored for several weeks. Spontaneous reoxidation of E2\_2SH, E5\_2SH, and E11\_2SH by dissolved O<sub>2</sub> was clear after four weeks (Figure 8). The reoxidation rates ( $k_2$ ) have comparable orders of magnitude to that of the reduction rates (lower by one order of magnitude), but reoxidation has a pronounced role only after reaching steady state, at which the concentration of the already reduced peptides becomes significant.

Reoxidation can take place both intra- and intermolecularly. Whereas the former leads to a decrease of overall conversion rates, the latter results in the formation of random molecular



**Figure 8.** The 50-membered structure ensembles of A) E11\_SS⇌E11\_2SH and B) E5\_SS⇌E5\_2SH. The fold of E11\_2SH is more compact than that of E5\_2SH, which has more “open” conformers, in which the Cys residues are far from each other. This allows intermolecular, rather than intramolecular, re-oxidation. The dissolved oxidized and reduced protein concentrations of C) E11\_SS⇌E11\_2SH and D) E5\_SS⇌E5\_2SH (oxidized: blue; reduced: green) as a function of time. In the case of E5\_SS⇌E5\_2SH, the initial concentration decreased by 68%, whereas, at the end of a complete redox cycle, the concentration of E11\_SS⇌E11\_2SH remained the same. E) Estimated parameters of the complete redox cycles. (The  $k_1$  values are slightly different from those in Table 3, for which the estimation comprises data only for phase 1.) Notably, in these long-term experiments, the rate of O<sub>2</sub> diffusion characterized by the rate constant  $k_4$  was also involved. Figure S11 contains all data for parameter estimation of E11\_SS, E5\_SS, and E2\_SS.

clusters, which may lead to precipitation. According to our semiquantitative analysis based on the recorded <sup>1</sup>H NMR spectra, the integral changes of the Trp H $\epsilon$ 1 resonances both in the oxidized and reduced forms of the protein during reduction with DTT show a decrease in concentration over the observed period of redox time for both E2\_SS and E5\_SS. Precipitation can be more intense if the protein concentration is higher. According to our present observations, increasing the length of the  $\alpha$ -helix within the Trp cage proteins stabilizes the soluble protein fraction. This means that the elongated N terminus, namely, the outer helix in the case of E11\_2SH, effectively shields the free SH– groups of the reduced protein, and thus, prevents any intermolecular reoxidation, whereas shorter variants, such as E2\_2SH and E5\_2SH, yield a significant amount of polymer formation. Due to the diversity of open 3D folds of both E5\_2SH and E2\_2SH, spontaneous intramolecular ring closure is hindered and less likely to happen. The N-terminal Cys of E11\_2SH is placed and fixed at the highly ordered inner

helix, with a reduced internal mobility of Cys18, and thus, mostly intramolecular ring closures take place. In the case of E5\_2SH, intermolecular SS bond formation is allowed, but may be limited just by Brownian motion and concentration. A comparison of the polymerization rates ( $k_3^{E11-SS} < k_3^{E5-SS}$ ) with different N-terminal lengths also supports this concept (Figure 8E).

E2\_SS was N-acetylated to eliminate the reduction rate-enhancing effect of the positively charged N terminus,  $-NH_3^+$ , in the vicinity of the SS bond. Upon acetylation,  $t_{1/2}$  has indeed increased ( $t_{1/2}^{E2-SS} = \approx 1$  min  $t_{1/2}^{Ac-E2-SS} = \approx 8$  min), but, in addition, the reaction reaches its steady state at a low conversion rate (50%). During reduction, almost immediately, both of the appropriate signal integrals of Ac-E2\_SS and Ac-E2\_2SH start to decrease, with a foamy precipitate gradually forming in the NMR tube. The isolated and HPLC-purified precipitant was identified as a polymer of the parent miniprotein by means of MS (Figure S12). Oligomer formation and soluble protein concentration decrease were more advanced for Ac-E2\_SS than that of E2\_SS (Figure S13). Due to the absence of the shielding effect of the outer  $\alpha$ -helix, the free thiol moiety of the N terminus is accessible for additionally reduced peptides in which the two free SH groups can hook peptide chains together. The polymer can grow until another free N terminus and acetylated C-terminal thiol-containing peptide closes polymerization. In addition, for Ac-E2\_2SH, intramolecular N $\rightarrow$ S acyl transfer could take place,<sup>[89]</sup> blocking some of the SH groups from promoting oligo- and polymerization through intermolecular SS bond formation.

## Conclusion

The SS-bond cyclized exenatide derivative and its variants were synthesized. Both the oxidized (E19\_SS) and reduced (E19\_2SH) forms, along with the parent molecule, E19, and all three of their truncated variants (E11\_SS, E11\_2SH, E11, E5\_SS, E5\_2SH, E5, E2\_SS, E2\_2SH, and E2) comprised the same Trp cage/SS/SH bond motif as that of their core structures. The SS bond stabilized model proteins showed improved thermostability and 3D fold compactness, with respect to their reduced and parent forms. Key residues for receptor binding remained in position in all of these models; therefore, E19\_SS might be promising agonists for GLP-1R and as a lead compound for type 2 diabetes mellitus.

The reduction rate of E19\_SS was found to be unexpectedly slow compared with that reported in the literature. The reaction takes hours ( $t_{1/2} = 48$  min), even at 37 °C, although the protein is small, and its single SS bond is exposed at the surface, and thus, accessible for reducing reagents. All four Trp cage variants studied herein have an almost equally compact core structure, with  $\alpha$ -helical segments of different length and internal mobility. By performing a complete NMR spectroscopy based structure elucidation, we found that the progress of reduction could be monitored by means of  $^1H$  NMR by using selected resonance frequencies. We have established that these four model proteins of different  $\alpha$ -helical lengths have significantly different reduction rate constants. Although it is generally complicated to discriminate each factor that affects the SS

bond reduction rate, the present set of miniproteins enabled them to be deciphered separately. We have focused special attention on the importance of the intramolecular protonation of the SS bond; this step greatly enhances the reaction rate. From CPMG measurements, we found that, at steady state, selected residues in the vicinity of the SS bond presented a slow exchange on the micro- to millisecond timescale of motion. This redox cycle lasts as long as active RA can be found in solution. We found that structural, steric, and electrostatic factors influenced the reduction rate greatly, resulting in almost three orders of magnitude differences in reduction half-lives ( $t_{1/2}$ ) for otherwise structurally similar and globularly folded model proteins.

Notably, in addition to intramolecular reoxidation within the redox cycle, intermolecular oxidation could also occur. The rate of these two concerted reactions depended on 1) the internal dynamics of the backbone conformers in the proximity of the SS bond, and 2) the shielding effect of the  $\alpha$ -helix on the SS bond. Intramolecular N $\rightarrow$ S acyl transfer in Ac-E2\_SS inhibits intramolecular reoxidation, but increases intermolecular reoxidation, which leads to oligo- and polymerization.

We found that easy-to-collect NUV-ECD spectral properties were indeed useful for monitoring the SS $\rightarrow$ SH reaction, even quantitatively, without the time-consuming assignment of the high-resolution NMR spectroscopy data. If the SS bond were situated in the vicinity of an aromatic cluster, NUV-ECD spectral changes could be used to monitor the transformation, which was proportional to the extent of reduction and clearly signaled when steady state had been reached. Thus, we encourage the use of CD spectroscopy for monitoring protein reduction rate in the manufacture of recombinant proteins (e.g., insulin, human monoclonal IgG antibodies) on a large scale, to control and provide information on the state of SS-SH bonds.

## Experimental Section

**ECD:** FUV-ECD spectra were recorded on a Jasco J810 spectrophotometer by using a 1.0 mm path length cuvette with protein concentrations of 20–30  $\mu M$ . Data accumulation was performed over a range of 185–260 nm, with 0.2 nm step resolution at a scan rate of 50  $nm\ min^{-1}$  with a 1 nm bandwidth. The spectral accumulations were resolved between 5 and 85 °C in steps of 5 °C. The temperature was controlled by using a Peltier-type heating system. Each spectrum baseline was processed by subtracting the solvent spectrum from that of the protein and the raw ellipticity data were converted into mean residue molar ellipticity units,  $[\theta]_{MR}$ .

**Reduction monitoring by NUV-ECD:** The spectra were recorded on a Jasco J810 spectrophotometer by using a 10 mm path length cuvette with protein concentrations of 120–150  $\mu M$ . Data accumulation was performed over a range of 240–325 nm, with 0.2 nm step resolution at a scan rate 50  $nm\ min^{-1}$  with a 1 nm bandwidth. The sample was tempered by using a Peltier-type heating system. Each spectrum baseline was processed by subtracting the solvent spectrum from the peptide spectrum and the raw ellipticity data were normalized by the concentration  $[\theta]$ . Reduction was followed for 75 h. Each intensity  $[\theta]$  at 266, 281, 287, and 293 nm was converted into concentration by using Equation (6).

$$c(t) = \frac{A_{\infty} - A}{A_{\infty} - A_0} [SS]_0 \quad (6)$$

**NMR spectroscopy:** All  $^1\text{H}$  NMR spectroscopy experiments were performed on a Bruker Avance III 700 MHz spectrometer equipped with a z-gradient 5 mm probe head operating at 700.13 MHz for  $^1\text{H}$ , whereas  $^{31}\text{P}$  NMR spectroscopy experiments were performed on a Bruker Avance 250 spectrometer with a 5 mm SB quad probe head.

**Monitoring reduction kinetics:** Peptide samples were prepared between 0.8 and 1.8 mM in 50 mM  $\text{NaH}_2\text{PO}_4\text{--Na}_2\text{HPO}_4$  buffer (600  $\mu\text{L}$ , pH 6.95), with 10%  $\text{D}_2\text{O}$ . A 0.1 M solution of NaOH was used to set the pH to 7. Sodium trimethylsilylpropanesulfonate (DSS) was added as the internal proton reference standard, set to  $\delta = 0.0$  ppm under all conditions.  $^1\text{H}, ^1\text{H}$  2D homonuclear spectra were recorded for the oxidized peptide; thereafter, upon the addition of a different excess of 0.5 M TCEP or DTT, reduction was observed by recording a series of 1D  $^1\text{H}$  spectra ( $n_s = 64$  or 128 scans). Finally, at the end point,  $^1\text{H}, ^1\text{H}$  homonuclear 2D spectra were recorded on the reduced peptide. Data sets were processed by using TopSpin 3.2 software. The conversion rate was determined by using the relative integral of the Trp H $\epsilon$ 1 signal in the oxidized ( $\text{Int}_{\text{OX}}$ ) and reduced ( $\text{Int}_{\text{RED}}$ ) form. Each integral was normalized to the integral of DSS. The concentrations were determined by the ratio of the oxidized and reduced integrals and the initial protein concentration.

**Structure determination:**  $^1\text{H}$  NMR spectroscopy assignments were completed by using  $^1\text{H}, ^1\text{H}$  COSY and  $^1\text{H}, ^1\text{H}$  TOCSY spectra, and then the distance restraints were determined based on  $^1\text{H}, ^1\text{H}$  NOESY spectra. Spin locks for  $^1\text{H}, ^1\text{H}$  TOCSY were 80 ms, whereas the mixing time for  $^1\text{H}, ^1\text{H}$  NOESY was 150 ms. CCP NMR<sup>[90]</sup> spectroscopy was used for resonance assignment, crosspeak calibration, and structure refinement. CNS Solve 1.3,<sup>[91]</sup> Aria 2.0 standard iteration protocol, and water refinement were used for 10-membered structure ensemble calculations. All structural figures were illustrated by using PyMOL software.

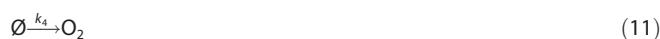
**CPMG effect:** Backbone  $^{15}\text{N}$ -longitudinal ( $R_1$ ) and transverse ( $R_2$ ) relaxation rates and the heteronuclear  $^1\text{H}, ^{15}\text{N}$  cross-relaxation rate constant (NOE) of E5, E5\_SS, and E5\_2SH were measured at 288 K. For each crosspeak ( $i$ ),  $R_{2,i}$  values were calculated by using Equation (7):

$$R_{2,i} = \frac{-\ln\left(\frac{I_i}{I_{\text{ref}}}\right)}{\frac{t_{\text{CPMG}}}{s}} \quad (7)$$

in which  $I_i$  is the intensity of the given crosspeak in the  $i$ th spectrum,  $I_{\text{ref}}$  is the intensity of the given crosspeak in the reference spectrum, and  $t_{\text{CPMG}}$  is the relaxation period of the CPMG measurement. The  $R_2$  values per residues were plotted against  $\nu_{\text{CPMG}}$  [Hz]. Quantitative analysis of the CPMG graph reveals those residues that show CPMG effects in the protein.

**Peptide synthesis and purification:** Proteins were prepared by means of standard solid-phase peptide synthesis or bacterial expression methods, as published previously.<sup>[92]</sup> Proteins were purified by means of reversed-phase HPLC on a  $\text{C}_{18}$  column by using a gradient of water/acetonitrile eluents. (Eluent A: 0.1% trifluoroacetic acid (TFA) in water; eluent B: 0.08% TFA and 80% acetonitrile in water.)

**Parameter estimation:** Kinetic parameter estimation was based on the integral of selected NMR signals considered to be proportional to the concentration of the relevant species. The mechanism taken into account is that given by Equations (8), (9), (10) and (11):



in which SS is the reduced model protein with an intramolecular S–S bond; 2SH is the same protein with the S–S bond reduced to two –SH groups; and the symbol  $\emptyset$  means a different phase from that of the reaction mixture, that is, the polymer aggregate as a sink in the first case and the gas phase as a source in the second case. Notably, in some cases, in which polymer precipitation ( $k_3$ ) and/or oxygen diffusion ( $k_4$ ) from the gas phase proved not to be present (indicated by largely nonsignificant estimated parameters concerning these processes), these steps have been omitted from the fitted mechanism.

For parameter estimation, the COPASI 4.16 (Build 104) Biochemical System Simulator software (<http://copasi.org/>) was used, with the parameter estimation option of the Levenberg–Marquardt method. The result of the estimation procedure did not depend on the choice of the initial parameters within a large interval; thus, there was one stable optimum for the fit of the model only. Confidence interval half-widths and relative standard deviations based on them were calculated from the estimated standard deviations, which suggested a Student distribution with  $n-p$  degrees of freedom, in which  $n$  is the number of data points in the concentration versus time measurements and  $p$  is the number of parameters estimated.

To determine the half-life and initial concentration of the SS species, a kinetic analysis of the temporal evolution of the reactions was performed. Both reduction and oxidation proved to be second-order reactions, which was not only supported by the good fit of the model, but also by the fact that, with this mechanism, the measured ( $[\text{SS}]_{0,\text{meas}}$ ) and calculated ( $[\text{SS}]_{0,\text{calcd}}$ ) initial concentrations of the model proteins were in very good agreement. From the kinetic analysis, the initial concentration of oxygen ( $[\text{O}_2]_0$ ) could also be estimated, except for one case in which the uncertainty of this parameter was very large, due to the lack of sufficient experimental data.

Because reduction follows second-order kinetics, the half-life ( $t_{1/2}$ ) of model proteins also depends on the actual concentration of the RA in the reaction mixture [Eq. (12)]:

$$t_{1/2} = \frac{1}{k_1(c_{\text{Red},0} - c_{\text{SS},0})} \ln\left(2 - \frac{c_{\text{SS},0}}{c_{\text{Red},0}}\right) \quad (12)$$

in which  $k_1$  is the rate constant of reduction,  $c_{\text{SS},0}$  is the initial concentration of the model protein, and  $c_{\text{Red},0}$  is the initial concentration of the RA. [Notably, Eq. (12) is valid only if  $c_{\text{Red},0}$  is greater than  $c_{\text{SS},0}$ —as in the current case. If  $c_{\text{SS},0}$  exceeds  $c_{\text{Red},0}$ , but it is not higher than that of twice the value of  $c_{\text{Red},0}$ , then the two initial concentrations should be flipped in both the difference and the fraction. If  $c_{\text{SS},0}$  exceeds  $c_{\text{Red},0}$  by more than a factor of two, then the SS protein concentration cannot become as low as half of the initial concentration, due to reduction.] For this reason, the half-life is less indicative of the rate of hydrolysis; the correct comparison

of the rates can be made based on the rate constant(s) of the second-order reaction(s).

## Acknowledgements

We thank Dóra K. Menyhárd and István Pintér for scientific discussions, Frank Löhr for CPMG measurements, and Zsanett Szegvári and András Koltai for their enthusiastic assistance. NMR spectrometer measurement time (700 MHz Bruker) was courtesy of a MedInProt Grant Facilitating Access to Instruments from the Hungarian Academy of Sciences. The Centre for Biomolecular Magnetic Resonance at the University Frankfurt (BMRZ) provided access for the CPMG measurements in project iNext, a Horizon 2020 programme of the European Union (H2020 grant no. 653706). This research project was supported by the State of Hungary and cofinanced by the European Regional Development Fund (VEKOP-2.3.3-15-2016-00009 and VEKOP-2.3.2-16-2017-00014) and the K116305 OTKA grant of the NKFIH of the Hungarian Academy of Sciences. This paper was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences (to V.F.)

## Conflict of Interest

The authors declare no conflict of interest.

**Keywords:** aggregation · kinetics · protein models · reduction · sulfur

- [1] C. Grek, D. M. Townsend, *Endoplasmic Reticulum Stress Dis.* **2014**, *1*, 4–17.
- [2] S. L. Xu, S. Sankar, N. Neamati, *Drug Discovery Today* **2014**, *19*, 222–240.
- [3] C. I. Andreu, U. Woehlbier, M. Torres, C. Hetz, *FEBS Lett.* **2012**, *586*, 2826–2834.
- [4] H. A. Khan, B. Mutus, *Front. Chem.* **2014**, *2*, 70.
- [5] G. Batta, T. Barna, Z. Gaspari, S. Sandor, K. E. Koeber, U. Binder, B. Sarg, L. Kaiserer, A. K. Chhillar, A. Eigentler, E. Leiter, N. Hegedues, I. Pocsi, H. Lindner, F. Marx, *FEBS J.* **2009**, *276*, 2875–2890.
- [6] J. W. H. Wong, S. Y. W. Ho, P. J. Hogg, *Mol. Biol. Evol.* **2011**, *28*, 327–334.
- [7] T. J. Bechtel, E. Weerapana, *Proteomics* **2017**, *17*, 1600391.
- [8] I. K. Jordan, F. A. Kondrashov, I. A. Adzhubei, Y. I. Wolf, E. V. Koonin, A. S. Kondrashov, S. Sunyaev, *Nature* **2005**, *433*, 633–638.
- [9] M. Cemazar, S. Zahariev, J. J. Lopez, O. Carugo, J. A. Jones, P. J. Hore, S. Pongor, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 5754–5759.
- [10] M. A. Wouters, S. W. Fan, N. L. Haworth, *Antioxid. Redox Signaling* **2010**, *12*, 53–91.
- [11] A. Holmgren, *Annu. Rev. Biochem.* **1985**, *54*, 237–271.
- [12] C. H. Lillig, C. Berndt, *Antioxid. Redox Signaling* **2013**, *18*, 1654–1665.
- [13] D. Butera, K. M. Cook, J. Chiu, J. W. H. Wong, P. J. Hogg, *Blood* **2014**, *123*, 2000–2007.
- [14] P. J. Hogg, *Nat. Rev. Cancer* **2013**, *13*, 425–431.
- [15] I. Azimi, J. W. H. Wong, P. J. Hogg, *Antioxid. Redox Signaling* **2011**, *14*, 113–126.
- [16] U. Siebenlist, G. Franzoso, K. Brown, *Annu. Rev. Cell Biology* **1994**, *10*, 405–455.
- [17] J. R. Auclair, J. L. Johnson, Q. Liu, J. P. Salisbury, M. S. Rotunno, G. A. Petsko, D. Ringe, R. H. Brown, D. A. Bosco, J. N. Agar, *Biochemistry* **2013**, *52*, 6137–6144.
- [18] F. Hatahet, L. W. Ruddock, *Antioxid. Redox Signaling* **2009**, *11*, 2807–2850.
- [19] K. S. Jensen, R. E. Hansen, J. R. Winther, *Antioxid. Redox Signaling* **2009**, *11*, 1047–1058.
- [20] “Thiol–Disulfide Interchange”, R. Singh, G. M. Whitesides in *The Chemistry of Sulfur-Containing Functional Groups* (Ed.: S. Patai), Wiley, London, **1993**, pp. 633–658.
- [21] R. Bhattacharyya, D. Pal, P. Chakrabarti, *Protein Eng. Des. Sel.* **2004**, *17*, 795–808.
- [22] T. E. Creighton, D. P. Goldenberg, *J. Mol. Biol.* **1984**, *179*, 497–526.
- [23] H. Nakamoto, J. C. A. Bardwell, *Biochim. Biophys. Acta Mol. Cell Res.* **2004**, *1694*, 111–119.
- [24] O. Carugo, M. Cemazar, S. Zahariev, I. Hudaky, Z. Gaspari, A. Perczel, S. Pongor, *Protein Eng.* **2003**, *16*, 637–639.
- [25] I. Hudáky, Z. Gaspari, O. Carugo, M. Cemazar, S. Pongor, A. Perezcel, *Proteins Struct. Funct. Bioinf.* **2004**, *55*, 152–168.
- [26] K. Ito, K. Inaba, *Curr. Opin. Struct. Biol.* **2008**, *18*, 450–458.
- [27] R. Daniels, P. Mellroth, A. Bernsel, F. Neiers, S. Normark, G. von Heijne, B. Henriques-Normark, *J. Biol. Chem.* **2010**, *285*, 3300–3309.
- [28] K. Inaba, *J. Biochem.* **2009**, *146*, 591–597.
- [29] N. A. Patil, J. Tailhades, R. A. Hughes, F. Separovic, J. D. Wade, M. A. Hosain, *Int. J. Mol. Sci.* **2015**, *16*, 1791–1805.
- [30] R. S. T. Charles, K. Padmanabhan, R. V. Arni, K. P. Padmanabhan, A. Tulinsky, *Protein Sci.* **2000**, *9*, 265–272.
- [31] Q. Liu, Q. Q. Huang, M. K. Teng, C. M. Weeks, C. Jelsch, R. G. Zhang, L. W. Niu, *J. Biol. Chem.* **2003**, *278*, 41400–41408.
- [32] J. Singh, J. M. Thornton, *FEBS Lett.* **1985**, *191*, 1–6.
- [33] T. R. Ioerger, C. G. Du, D. S. Linthicum, *Mol. Immunol.* **1999**, *36*, 373–386.
- [34] H. J. Leung, G. Xu, M. Narayan, H. A. Scheraga, *J. Pept. Res.* **2005**, *65*, 47–54.
- [35] J. L. Lau, M. K. Dunn, *Bioorg. Med. Chem.* **2018**, *26*, 2700–2707.
- [36] H. Liu, K. May, *mAbs* **2012**, *4*, 17–23.
- [37] J. A. Burns, J. C. Butler, J. Moran, G. M. Whitesides, *J. Org. Chem.* **1991**, *56*, 2648–2650.
- [38] D. J. Cline, S. E. Redding, S. G. Brohawn, J. N. Psathas, J. P. Schneider, C. Thorpe, *Biochemistry* **2004**, *43*, 15195–15203.
- [39] J. C. Han, G. Y. Han, *Anal. Biochem.* **1994**, *220*, 5–10.
- [40] F. Duchardt, I. R. Ruttekkolk, W. P. R. Verdurmen, H. Lortat-Jacob, J. Buerck, H. Hufnagel, R. Fischer, M. van den Heuvel, D. W. P. M. Loewik, G. W. Vuister, A. Ulrich, M. de Waard, R. Brock, *J. Biol. Chem.* **2009**, *284*, 36099–36108.
- [41] M. E. Levison, A. S. Josephson, D. M. Kirschenbaum, *Experientia* **1969**, *25*, 126–127.
- [42] W. R. Gray, *Protein Sci.* **1993**, *2*, 1732–1748.
- [43] J. Eng, W. A. Kleinman, L. Singh, G. Singh, J. P. Raufman, *J. Biol. Chem.* **1992**, *267*, 7402–7405.
- [44] J. L. Iltz, D. E. Baker, S. M. Setter, R. K. Campbell, *Clin. Therapeutics* **2006**, *28*, 652–665.
- [45] D. J. Drucker, M. A. Nauck, *Lancet* **2006**, *368*, 1696–1705.
- [46] G. Cantini, E. Mannucci, M. Luconi, *Trends Endocrinol. Metabol.* **2016**, *27*, 427–438.
- [47] J. J. Meier, *Nat. Rev. Endocrinol.* **2012**, *8*, 728–742.
- [48] D. Donnelly, *British J. Pharmacol.* **2012**, *166*, 27–41.
- [49] A. Hareter, E. Hoffmann, H. P. Bode, B. Goke, R. Goke, *Endocr. J.* **1997**, *44*, 701–705.
- [50] J. W. Neidigh, R. M. Fesinmeyer, N. H. Andersen, *Nat. Struct. Biol.* **2002**, *9*, 425–430.
- [51] B. Barua, J. C. Lin, V. D. Williams, P. Kummner, J. W. Neidigh, N. H. Andersen, *Protein Eng. Des. Sel.* **2008**, *21*, 171–185.
- [52] S. Runge, H. Thogersen, K. Madsen, J. Lau, R. Rudolph, *J. Biol. Chem.* **2008**, *283*, 11340–11347.
- [53] Y. Zhang, B. F. Sun, D. Feng, H. L. Hu, M. Chu, Q. H. Qu, J. T. Tarrasch, S. Li, T. S. Kobilka, B. K. Kobilka, S. Skiniotis, *Nature* **2017**, *546*, 248–253.
- [54] P. Rovó, V. Farkas, O. Hegyi, O. Szolomajer-Csikos, G. K. Toth, A. Perczel, *J. Pept. Sci.* **2011**, *17*, 610–619.
- [55] P. Hudáky, P. Straner, V. Farkas, G. Varadi, G. Toth, A. Perczel, *Biochemistry* **2008**, *47*, 1007–1016.
- [56] P. Rovó, V. Farkas, P. Straner, M. Szabo, A. Jermendy, O. Hegyi, G. K. Toth, A. Perczel, *Biochemistry* **2014**, *53*, 3540–3552.
- [57] G. R. S. Hartig, T. T. Tran, M. L. Smythe, *Protein Sci.* **2005**, *14*, 474–482.
- [58] A. Kumar, R. Schweitzer-Stenner, B. M. Wong, *Chem. Commun.* **2019**, *55*, 5701–5704.
- [59] R. Schweitzer-Stenner, S. E. Toal, *Biophys. J.* **2018**, *114*, 1046–1057.
- [60] V. Farkas, I. Jakli, G. K. Toth, A. Perczel, *Chem. Eur. J.* **2016**, *22*, 13871–13883.

- [61] P. Rovó, P. Straner, A. Lang, I. Bartha, K. Huszar, L. Nyitray, A. Perczel, *Chem. Eur. J.* **2013**, *19*, 2628–2640.
- [62] A. Perczel, K. Park, G. Fasman, *Anal. Biochem.* **1992**, *203*, 83–93.
- [63] A. Perczel, M. Hollosi, G. Tusnady, G. Fasman, *Protein Eng.* **1991**, *4*, 669–679.
- [64] R. W. Woody, *Biopolymers* **1978**, *17*, 1451.
- [65] “Circular Dichroism of Peptides”, R. W. Woody in *The Peptides, Vol. 7* (Eds.: V. J. Hruby), Academic Press, New York, **1985**, pp. 15–114.
- [66] *Circular Dichroism and the Conformational Analysis of Biomolecules* (Ed.: G. D. Fasman), Springer, Boston, **1996**.
- [67] A. Bundi, K. Wüthrich, *Biopolymers* **1979**, *18*, 285–297.
- [68] A. J. Doig, R. L. Baldwin, *Protein Sci.* **1995**, *4*, 1325–1336.
- [69] A. Mirzahosseini, B. Noszal, *Sci. Rep.* **2016**, *6*, 37596.
- [70] M. Tollinger, N. R. Skrynnikov, F. A. A. Mulder, J. D. Forman-Kay, L. E. Kay, *J. Am. Chem. Soc.* **2001**, *123*, 11341–11352.
- [71] B. Herbert, F. Hopwood, D. Oxley, J. McCarthy, M. Laver, J. Grinyer, A. Goodall, K. Williams, A. Castagna, P. G. Righetti, *Proteomics* **2003**, *3*, 826–831.
- [72] Z. Wang, T. Rejtar, Z. S. Zhou, B. L. Karger, *Rapid Commun. Mass Spectrom.* **2010**, *24*, 267–275.
- [73] P. Liu, B. W. O’Mara, B. M. Warrack, W. Wu, Y. Huang, Y. Zhang, R. Zhao, M. Lin, M. S. Ackerman, P. K. Hocknell, G. Chen, L. Tao, S. Rieble, J. Wang, D. B. Wang-Iverson, A. A. Tymiak, M. J. Grace, R. J. Russell, *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 837–844.
- [74] G. G. Smith, B. Silvadesol, *Science* **1980**, *207*, 765–767.
- [75] D. E. Schwass, J. W. Finley, *J. Agric. Food Chem.* **1984**, *32*, 1377–1382.
- [76] P. A. Fernandes, M. J. Ramos, *Chem. Eur. J.* **2004**, *10*, 257–266.
- [77] J. Houk, G. M. Whitesides, *J. Am. Chem. Soc.* **1987**, *109*, 6825–6836.
- [78] L. E. Overman, E. M. O’Connor, *J. Am. Chem. Soc.* **1976**, *98*, 771–775.
- [79] O. Dmitrenko, C. Thorpe, R. D. Bach, *J. Org. Chem.* **2007**, *72*, 8298–8307.
- [80] G. H. Snyder, M. J. Cennerazzo, A. J. Karalis, D. Field, *Biochemistry* **1981**, *20*, 6509–6519.
- [81] C. L. Wu, C. Belenda, J. C. Leroux, M. A. Gauthier, *Chem. Eur. J.* **2011**, *17*, 10064–10070.
- [82] O. Gawron, S. Mahboob, J. Fernando, *J. Am. Chem. Soc.* **1964**, *86*, 2283–2286.
- [83] C. G. Moore, B. R. Trego, *Tetrahedron* **1962**, *18*, 205–218.
- [84] T. Kortemme, T. E. Creighton, *J. Mol. Biol.* **1995**, *253*, 799–812.
- [85] L. W. Guddat, J. C. A. Bardwell, J. L. Martin, *Struct. Fold. Design* **1998**, *6*, 757–767.
- [86] H. N. Po, N. M. Senozan, *J. Chem. Educ.* **2001**, *78*, 1499–1503.
- [87] R. D. Bach, O. Dmitrenko, C. Thorpe, *J. Org. Chem.* **2008**, *73*, 12–21.
- [88] M. Scigelova, P. S. Green, A. E. Giannakopoulos, A. Rodger, D. H. G. Crout, P. J. Derrick, *Eur. J. Mass Spectrom.* **2001**, *7*, 29–34.
- [89] J. Tailhades, N. A. Patil, M. A. Hossain, J. D. Wade, *J. Pept. Sci.* **2015**, *21*, 139–147.
- [90] W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, P. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides, E. D. Laue, *Proteins Struct. Funct. Bioinf.* **2005**, *59*, 687–696.
- [91] A. T. Brunger, *Nat. Protoc.* **2007**, *2*, 2728–2733.
- [92] P. Stráner, N. Taricska, M. Szabo, G. K. Toth, A. Perczel, *Curr. Protein Pept. Sci.* **2016**, *17*, 147–155.

---

Manuscript received: July 30, 2019

Accepted manuscript online: September 2, 2019

Version of record online: November 18, 2019