Proteins' fold compactness alters disulfide-bond reducibility by 3 orders of magnitude: a comprehensive kinetic case study on the reduction of different size Trp-cage model proteins

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

Abstract: We derived a novel approach to monitor disulfide bond reduction in the vicinity of aromatic cluster(s) by using the near-UV range (266–293 nm) of ECD spectra. By using combined NMR- and ECD-spectroscopy we have determined the 3D-fold characteristics and the associated reduction rate constants (k) of E19_SS – a highly thermostable; disulfide bond reinforced 39-amino acid long Exenatide mimetic – and its N-terminally truncated derivatives at different experimental conditions. Single SS-bond reduction of the E19_SS model (using 18-fold excess of TCEP, pH = 7, 37 °C) takes hours, 20-30 times longer than expected, thus would not reach completion applying the commonly used reduction protocols. We found that structural, steric and electrostatic factors influence the reduction rate, resulting in magnitude differences in reduction half-times ($900 > t_{\frac{1}{2}} > 1$ min) even for structurally similar, well-folded derivatives of a small model-protein.

Introduction

Forming a multitude of temporary H-bonds, H_2O molecules interact with backbone amides and thus, loosen the polar intramolecular interactions (salt-bridges, H-bonds, cation/anion↔π-interactions, *etc.*) holding together proteins' secondary and tertiary structures. Structural disulfide bonds – stable in harsh oxidative extracellular environment – on the contrary, maintain proteins' native fold by fixing and protecting them from thermal fluctuation induced elevated internal dynamics. SS-bond formation is perhaps the most fundamental post-translational modification that stabilizes globular proteins 3D-fold. The absence of regulated disulfide formation leads to diseases including diabetes [1], cancer [2], neurodegenerative conditions [3] and cardiovascular diseases [4]. Non-native SS-bond pairing evokes backbone misfolding, jeopardizing both function and bioactivity, although some proteins may present alternative disulfide states and still get similarly well-folded forms. [5] In protein evolution the presence of disulfide bonds shows a significant correlation with complexity of the organism. [6] Approximately 50% of all the cysteines found in proteins form disulfide-bonds [7], thereby these cysteines become the most conserved among all amino-acid, despite being added late to the genetic code during protein evolution. [8] Due to the unique pairing pattern of cysteines, SS-bonds stabilize proteins' 3D-fold unambiguously [9].

Contrary to the structural disulfides, redox-active disulfides are highly dynamic and their formation is reversible. The redox potential of the surrounding environment controls these proteins regulation and cellular localization. [10] Intramolecular formation of these redox-active disulfides is common for oxoreductases (thioredoxin [11] or glutaredoxin [12] family) and allosteric disulfides [13][14][15], while the intermolecular disulfide linkage results in glutathionylated [16] or cysteinylated [17] small molecule-protein adducts. The redox potential and stability of the SS-bond is highly dependent on several factors, such as the pKa of the tiols (the standard pKa is 8.5, however pKa values are known from 3.5 to 12.8 depending on local environment) [18], the SS-bond introduced strain of the protein structure and the entropic cost of the SS-bond formation. [19][20] The Cys residues of an SS-bond are typically distant in the primary sequence, 49% of the SS-bond forming cysteines are more than 25 residues apart from each other. [21] Disulfide-bond formation is thermodynamically more favorable if the cysteines are already in spatial vicinity in the tertiary structure before oxidation by the native fold itself [22], otherwise – in the absence of chaperones assisting the folding [23] - the protein precipitates. Adjacent cysteines oxidized to SS-bond are rare, though examples can be found among enzymes, receptors and toxins. [24][25] The disulfide bond or SS-bond pattern in prokaryotic proteins is formulated by ribosomal mRNA translation followed by oxidation and posttranslational modifications catalyzed by various enzymes located in the periplasm (DsbC, DsbG, DsbD) [26] or cytoplasm (DsbA, DsbB). [27][28] 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). [29]

For hundreds of proteins of known 3D structures, SS-bonds form the nucleus of their core. Hydrophobic and/or aromatic residues (*e.g.* Trp and Tyr) may condense around the SS-bond and form the network of key interactions that determines the 3D-structure of a large number of different protein families. [21] In at least 50% of the protein families, this type of interaction is invariant. In dozen of proteins (*e.g.* Tick anticoagulant peptide [30], phospholipase A2 [31]) SS-unit(s) are reinforced by associated aromatic-aromatic interactions [32][33] and *vica-versa*. For instance, Tyr92 of RNase-A effectively shields the solvent-exposed nearby SS-bond (Cys40-Cys95) from reducing agents and thus, Tyr92 helps to maintain the native fold of the protein. [34]

When SS-bonds are reduced, the free thiol groups of cysteines often adopt an ensemble of local conformers loosening neighboring residues compactness as well. In the era of manufacturing recombinant proteins (*e.g.* Insulin), SS-bond cyclized peptides (*e.g.* vasopressin, oxytocin, desmopressin, octreotide) [35] and human monoclonal IgG antibodies [36] on a large scale by biopharmaceutical industry, it is vital to have reliable and fully tested methods for SS-bond reduction in an aqueous solution.

Besides β-mercaptoethanol or DTT, more recently TCEP is commonly used as it has an elevated chemical stability and reduces SS-bonds more effectively and quickly even at lower pHs. The non-volatile, SH-free and thus odorless TCEP is

[a] D. Horváth, N. Taricska, P. Stráner, V. Farkas, A. Perczel
Laboratory of Structural Chemistry & Biology and MTA-ELTE
Protein Modeling Research Group at the Institute of Chemistry
Eötvös Loránd University
H-1518, 112, PO Box 32, Budapest, Hungary
E-mail: perczel@ghem.elte.hu
[b] E. Keszei
Chemical Kinetics Laboratory, Institute of Chemistry
Eötvös Loránd University
H-1518, 112, PO Box 32, Budapest, Hungary
[c] G. K. Tóth
Department of Medical Chemistry, Faculty of General Medicine
University of Szeged
Szeged, Hungary

Supporting information for this article is given via a link at the end of the document.
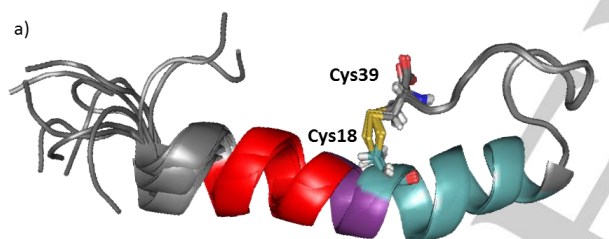
hydrophilic, soluble (>24 h) in aqueous solutions and buffers at a wide *pH* range (2.0 < *pH* < 9.0) [37], but it is less likely to permeate the hydrophobic protein core, which limits its applicability. Esterification to trimethyl TCEP (tmTCEP) could be a proper alternative of TCEP for SS-bond reduction in a hydrophobic medium and for non-accessible, core disulfide-bonds. [38] TCEP is claimed to selectively and completely reduce water-soluble alkyl disulfides over a wide *pH* range within a few minutes (< 5 min). [39] Some protocols recommend using 1–100 molar equivalents of TCEP over protein concentration. [40][41] The reduction time and the proper temperature greatly depends on the nature of the protein, but generally, elevated temperature and/or TCEP [42] concentration and longer times make the reduction more complete, but these conditions also initiate a multitude of side reactions, which are yet poorly described.

Exendin-4 [43] or Exenatide [44] (synthetic name) – used in clinical practice since 2005 - is an incretin mimetic [45] GLP-1 analogue, a ninety-three residue long peptide with complex physiological actions [46] in multiple organs, resulting in an effective treatment of type 2 *Diabetes Mellitus* [47]. Exenatide acts as an agonist of the GLP-1 receptor [48] (GLP-1R). Its amphipathic helix binds to the extracellular domain of the GLP-1R, the mainly unstructured *N*-terminal activates the receptor [49] while the structure stabilizing Trp-cage [50][51] fold is not directly involved in interactions to GLP-1R. [52][53] We have synthesized and studied the 3D-fold of several dozens of Trp-cage folds including Exenatide-4 analogs such as E19 [54][55], a 39 amino acid comprising protein of comparable bioactivity but of improved water solubility. As a 3D-fold compactness enhancing "natural tool", we introduced two solvent exposed Cys residues into E19 making E19_SS **(Figure 1)** looping residue 18


Figure 1. a) Structure ensemble of E19_SS and **b)** amino acid sequences of GLP-1, Exenatide, the parent E19 and its truncated derivatives E11, E5, E2 as well as their SS analogues: E19_SS, E11_SS, E5_SS, E2_SS and their reduced 2SH analogues: E19_2SH, E11_2SH, E5_2SH, E2_2SH. The position of SS-bridge is highlighted by stick representation and underlined as <u>C</u>. Sequences of E19 is divided into 6 major parts: 1) 1-8 unstructured N-terminal, 2) 9-14 outer helix, 3) 15-17 kink region, 4) 18-27 inner helix, 5) 28-34 3_{10} helix, 6) poly-proline region. This apportionment of the sequence coincides with the truncation of the peptides.

to 39 in E19_A18C_S39C (E19_2SH). E19_2SH oxidized to E19_SS spontaneously by atmospheric O_2 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

Tyr22, surrounded by explicit negative charges (Glu15, Glu16, Glu17). Although E19_SS is small in size, (MW: 4334.9 Da) its reduction took several hours to reach equilibrium, by using >10 molar excess of TCEP in water at room temperature. As the SS-bond reduction time turned out to be significantly longer than expected based on literature data and common laboratory practice, we launched a comparative study including 3 designed and truncated analogs of E19_SS, namely E11_SS, E5_SS and E2_SS as well. Note that the model systems thus created (**Figure 1**) oxidize spontaneously and rapidly adopt the Trp-cage 3D-fold. [56] Moreover, the "loop size" created by the SS-bond, in other words, the number of residues between the two reacting cysteines is 20 amino-acid, close to the average value (~17) observed in thousands of proteins. [57]

E11_SS was designed by removing the '**HGEGTFTS**-'tail, the unstructured GLP-1R activating *N*-terminal eight residues of E19_SS. Shortened by an additional six residues eliminates also removes the outer-helical part of E19_SS, namely the '**HGEGTFTS-DLSKQM-**'subunit [56] affording E5_SS. Although fourteen residues shorter than E19-SS, E5_SS still adopts a compact Trp-cage fold and comprises the entire interface for binding to GLP-1R. [58] Finally, in E2_SS, the entire *N*-terminus preceding Cys18 of E19_SS was omitted, namely '**HGEGTFTS-DLSKQ-EEE-**' was cleaved, to get a folded protein with the SS-bond at its surface fully exposed, considered as a construct ready for a rapid SS-bond reduction **(Figure 1)**.

Hereby we discuss the detailed atomic structure and properties both the oxidized and the reduced form in comparison with the parent miniproteins. We introduce possible spectroscopic approaches that make the monitoring of the reduction progress available, fast and easy. We define and compare the reduction kinetics of the four model proteins of different α-helical length. The effect of the compactness on the protein fold, the accessibility and the local explicit charges of the SS-bond, the reagent type on reduction rate and the mechanism are explanted in this paper.

Results and Discussion

3D-fold as characterized by far-UV ECD spectra

The FUV-ECD spectra of Trp-cage proteins (*e.g.* Exenatide, E19, E19_SS) are typically the weighted sums of the C-type (folded & highly helical) and U-type (unfolded) base curves **(Figure 2/a),** assigned and verified by NMR [59] [60]. As the temperature increases, the shape of the FUV-ECD spectra changes: those of the parent proteins – E2, E5, E11, E19 – acquire more and more U-type characteristics, as they unfold gradually. The temperature dependent FUV-ECD spectra for all 4 SS-bond enforced model peptides were recorded between 5 °C and 85 °C (by a step of 5 °C resulting in 17 spectra for each proteins) **(SFigure 1)**. Aside from E2_SS, the SS-bond containing mutants have similar FUV-ECD spectra as their parent proteins at low temperatures. On the other hand, as the SS-bond makes the 3D-folds of disulfide variants more rigid, they preserve more their C-type characteristic and delay the unfolding even at higher temperatures. Once the SS-bond is reduced (see in details later), the spectral properties of the SH-variants revert to those of parent proteins. Their 3D-scaffold compactness decreases as the temperature increases, less apparent in case of E2 & E2_2SH, as they both present – already at 5°C – an ensemble of dynamic backbone structures.
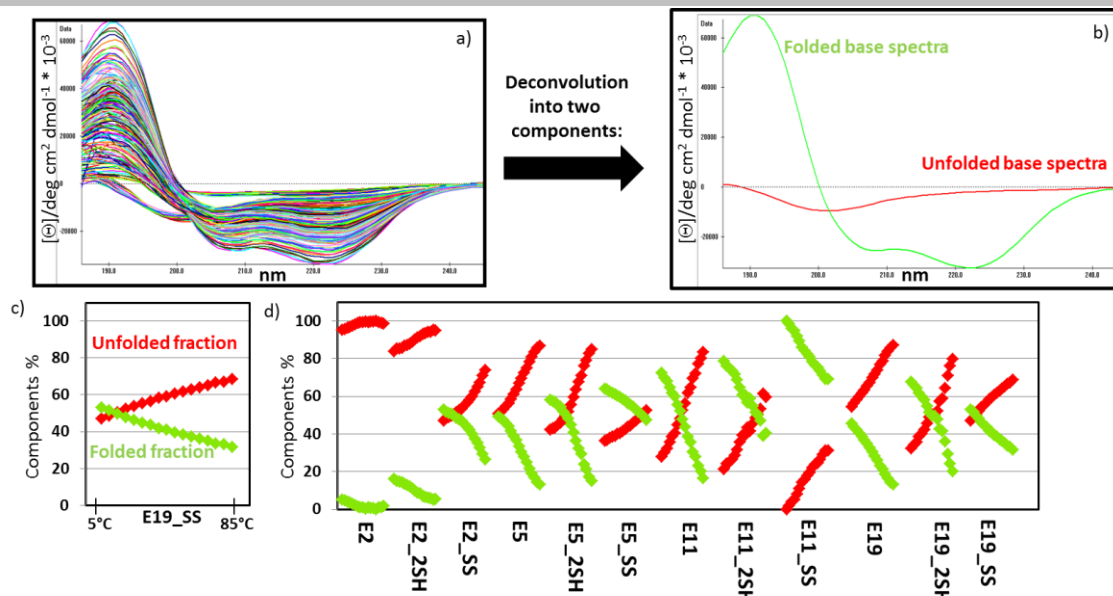
Figure 2. a) Temperature-dependent FUV-ECD spectra-series (204 in total) of the 4 primal peptides (E2, E5, E11, E19) and their 4 reduced (_2SH) & 4 oxidized (_SS) variants. **b)** The two pure ECD curves were derived from the ensemble analysis of the 204 ECD-spectra by using CCA+. Pure component 1 (red) stands for the unfolded/U-type, while 2 (green) for the folded/C-type backbone structure. **c)** The associated relative propensities (%) 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).

Ensemble deconvolution [61][62] of the 204 (12x17) 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, since as the pure ECD curves were successfully assigned.[59][63][64][65] The results indicate that **Figure 2/b**): *i*) SS-bond stabilizes the less folded protein scaffolds more effectively: *e.g.* while 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 case of E11 and E11_SS is only 28 % and last but not least for E19 and E19_SS is 7 %. (**Figure 2/d**) *ii*) The ratio of the folded, helical components increases going from E2 to E5 and E11, however the compact α-helix content of E19_SS, E19_2SH & E19 is lower than those of E11_SS, E11_2SH & E11 as the unfolded 8-residue long *N*-terminal part elevates the overall backbone dynamics, destabilizing the compact 3D-fold. *iii*) All 4 reduced proteins (E2_2SH, E11_2SH, E19_2SH) have higher helix content (~7–15%) than the parent proteins in a course of the whole temperature range. *iv*) 3D-folds stabilized by SS-bonds are less temperature-sensitive. (**Figure 2/c-d**)

Proteins 3D-fold determined and characterized by NMR

The ensemble of the temperature-dependent FUV-ECD spectra confirms that an SS-bond preserves the model proteins' fold and increases thermostability. NMR analysis carried out at 15°C allowed further characterization of the 3D structures of each variant. Fold, chemical shift and secondary chemical shift information [66] were derived from the appropriate 2D-homonuclear NMR experiments (^{1}H-^{1}H-COSY, ^{1}H-^{1}H-TOCSY and ^{1}H-^{1}H-NOESY) at T = 15 °C, the ensemble of the 10 lowest energy structures were analyzed. This comprehensive analysis conducted at 15°C provides the following useful structural descriptors, the RMSD of the 3D-structures, the average chemical shift deviance of backbone Hα protons per residue ($\frac{\Sigma CSD_{(i)}^{H\alpha}}{i}$) in the helical segment and the compactness of the Trp-cage core by secondary chemical shift sum of selected protons: CSD$_{cage.}$ (**Table 1**)

The comparison of the helices of different lengths is more straightforward if the helical segment is divided into four partitions: 1) the outer α-helix, 2) the kink region in the vicinity of the SS-bond and 3) the inner α-helix. (**Figure 3**) The outer helix

Table 1. Selected measures characterizing the degree of the model protein's folding (T=15 °C & pH=7)

	E2			E5			E11			E19		
	Parent	_SH	_SS	Parent	_SH	_SS	Parent	_SH	_SS	Parent	_SH	_SS
degree of the fold by FUV-ECD (%)[a],[b]	3.7	14.8	51.8	51.5	42.8	63.4	70.0	76.0	96.0	43.7	65.8	51.6
backbone RMSD (Å)[c]	1.5	0.3	0.1	1.6	0.1	0.2	1.5	0.3	0.3	1.2	0.7	0.7
CSD$_{cage}$[e]	3.8	9.6	11.3	10.3	10.4	11.4	10.9	11.0	11.5	10.9	11.1	11.7
$\frac{\Sigma CSD_{(i)}^{H\alpha}}{i}$[f]	0.3	0.4	0.5	0.4	0.5	0.5	0.5	0.6	0.5	0.5	0.6	0.5

[a] T= 15°C, c$_{(protein)}$= 20-30 µM @ pH~7 (typical conditions applied for CD measurements)
[b] calculated % from the joint deconvolution (CCA+) of 204 T-dependent FUV-ECD spectra (c.f. Figure 2)
[c] T= 15°C, c$_{(protein)}$= 0.8-1.8 mM @ pH~7 (typical conditions applied for ^{1}H-NMR measurements)
[d] RMSD of the 10 best structure's all backbone atoms
[e] Xf-cage values [51][55] were used to correlate characterize protein's fold. The following „H" atoms were involved in calculation: W25Hε1, L26Hα, G30Hα2, P31Hβ2, R35Hα, P37Hα, P37Hβ2, P38Hδ1, P38Hδ2
[f] The average chemical shift deviance of backbone Hα protons per residue
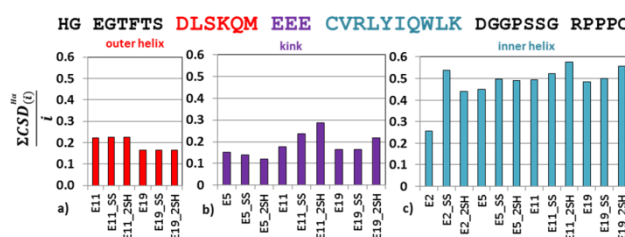
Figure 3. The average chemical shift deviance of backbone Hα protons per residue ($\frac{\Sigma CSD_{(i)}^{H\alpha}}{i}$) in the three different helical regions: **a)** outer helix (i=6), **b)** kink region (i=3), **c)** inner helix (i=10). Higher residual values imply a more structured α-helix.

compactness seems to be affected by the length of the α-helix. Interestingly, this part of E11 variants is slightly more structured, although this segment is the terminal part (usually flexible and unstructured) unlike in E19 variants where this helical segment is flanked. (**Figure 3/a**) The above tendency is true for the kink-region too, but here, the presence and state (_SS or _2SH) of the SS-bond also differentiates. (**Figure 3/b**) These distant

helical parts have generally lower $\frac{\Sigma CSD_{(i)}^{H\alpha}}{i}$ values compared to those of the inner helix. The compactness of the inner helices is similar (expect for E2). Interestingly, reduced longer polypeptides show slightly increased $\frac{\Sigma CSD_{(i)}^{H\alpha}}{i}$ values, which may be the indicative of ring tension in the SS-bond cyclized variants in these systems. (**Figure 3/c**)

^{1}H-NMR studies also confirm that all model proteins - except E2 - have a common, compact and folded, Trp-cage core structure at T=15°C (**Table 1** & **SFigure 2**) regardless to the differently structured tails attached to them (**Table 1**). E2 is dominantly unfolded even at low temperature (15°C), but as SS-bond fixes together the N- and the C-terminus of E2_SS and the hydrophobic core folds properly. Interestingly, even in the E2_2SH the A1C and A22C substitution make the Trp-cage fold more compact as compared to E2. In line with literature data, cysteines promote and stabilize α-helices, if located at their N-terminus. [67] The cage values of the longer reduced peptides are close to their oxidized counterparts (**Table 1**). The NMR data reveal that the longer the α-helix, the more structured the Trp-cage is, in all the studied cases.

In general, it seems that the reduced (SH-) proteins' core is almost as well-folded as those of the 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) → E2_SS ($CSD_{cage}^{E2\_SS}$ = 11.3).

Oxidized and reduced states defined by 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 the event of reduction NUV-ECD spectra (instead of Far-UV) had to be used. The interpretation of the observed chiroptical property changes of the Trp/Tyr/**SS** → Trp/Tyr/2**SH** (**Figure 4/a**) complex chromophore system is less straightforward, as the assignment of "pure" NUV-ECD spectra has not yet been completed. The conformation dependent fine structure of Tyr/Trp chromophores [59] (260 ≤ λ ≤ 320 nm) comprises the [1]Lb of Tyr (~276 nm with a shoulder at ~287 nm), [1]Lb of Trp (~281 and ~293 nm), and [1]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 nearby 260–270 nm. For the current proteins with SS-bonds a larger negative band was recorded (**SFigure 3**). The peaks of Trp and Tyr in the SS-bond containing proteins shifted to the negative ellipticity range, which was not seen in case of the parent proteins (E2, E5, E11, E19 [59], where the bands of these amino acids have been detected in the positive range (except Trp band at ~293nm). The reduction kinetics of E19_SS → E19_2SH was monitored over time as band intensities at ~266, ~281, ~287, and ~295 nm increases from larger negative to smaller negative and/or positive values becoming similar to those of the parent proteins. (**Figure 4/a, SFigure 3**) We were encouraged to use NUV-ECD spectral changes to monitor SS to SH reduction in proteins if embedded in a suitable molecular environment like that of the Tc-motif. Below we give the NUV-ECD and ^{1}H-NMR analyses of the reduction of E19_SS to determine its kinetic properties.

Due to the acidic nature of TCEP, to avoid any pH shift, a phosphate buffer (50 mM, pH= 7) was used and thus a near physiological pH was kept. Han [39] and Whitesides [37] described the chemical instability of TCEP above pH= 7 in 300–400 mM phosphate buffer. They found that the autoxidation of TCEP depends on how the reagent is stored (open air and/or capped vials), whether the solution is stirred and on the elapsed time (24 ≤ t ≤ 72 h) of the storage. However, here we have monitored TCEP stability by using ^{31}P-NMR in 50 mM phosphate buffer,

and found no significant spectral changes connected to TCEP oxidation or degradation at room temperature over 14 days.
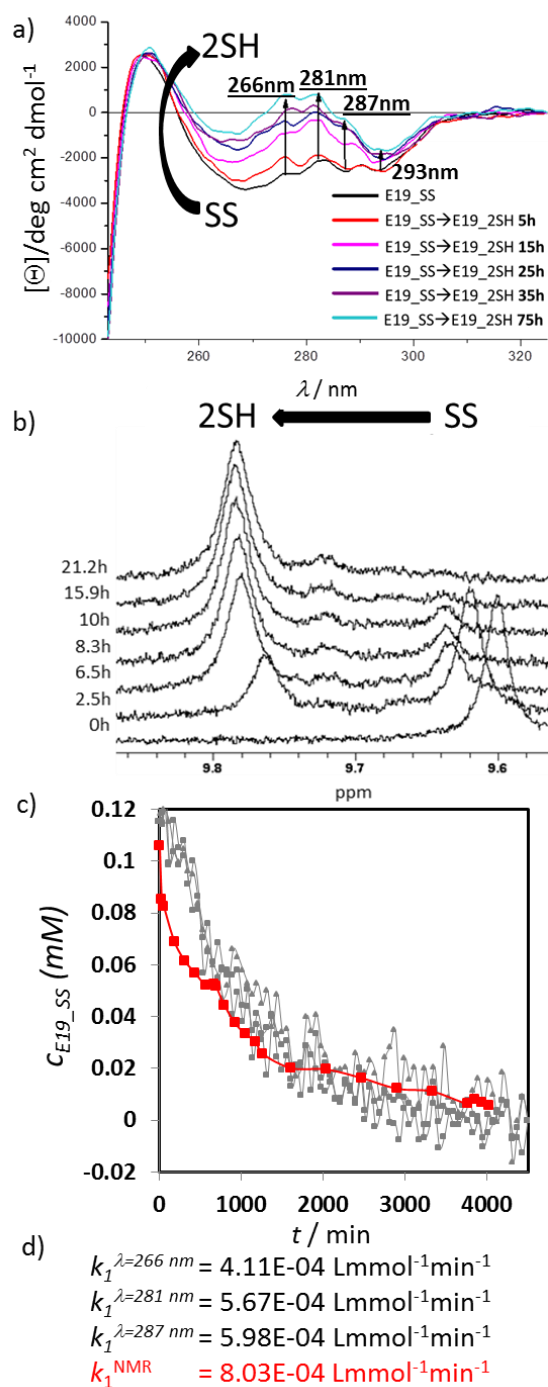
Figure 4. a) NUV-ECD spectral changes measured for the reduction of E19_SS (~ 0.113 mM E19_SS, pH = 7, 15°C, 18xTCEP) at four different wavelengths (266 nm, 281 nm; 287 nm; 293 nm). No spectral changes were observed after ~55 h (3300 min). **b)** ^{1}H-NMR spectra of the E19_SS → E19_2SH reduction (c ~ 0.115 mM, pH = 7, 15 °C, 18xTCEP) in water. The chemical shift of the indole Hε1 of Trp25 is used to monitor reduction: Hε1 upfield shifts from 9.60 ppm (SS) to 9.78 ppm (2SH during the reduction). Reaching steady state the integral ratio of E19_2SH and E19_SS was found as 92 to 8%. **c)** Concentration change of E19_SS (mM) measured during reduction by different approaches   plotted as a function of time. **d)** Complemented with the calculated rate constants. (c.f. modeling reduction kinetic)

Reduction of the E19_SS protein was followed by recording NUV-ECD spectra (~0.113 mM E19_SS, pH= 7, 15 °C, 18-fold excess of TCEP, cell length = 10 mm) at four different wavelengths (266 nm, 281 nm, 287 nm and 293 nm). Thus, by following band intensity changes of selected (one or more) of the [1]Lb of Tyr or Trp transitions we could monitor the redox state of the SS/SH-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 to monitor its reduction/oxidation state even when 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 concentrations using the following equation: $c(t) = \frac{A_\infty - A}{A_\infty - A_0}[SS]_0$. Steady state was reached conclusively after ~55 h. We determined the rate constant k_1 at each wavelength by parameter estimation as: $k_1^{\lambda=266nm}$ = 4.11E–4 Lmmol^{-1}min^{-1}, $k_1^{\lambda=281nm}$ = 5.67E–4 Lmmol^{-1}min^{-1}, $k_1^{\lambda=287\ nm}$ = 5.98E–4 Lmmol^{-1}min^{-1}. **(Figure 4/c, SFigure 4)** The deviations of the fitted model from the measured data at 293 nm were remarkably large; therefore, parameter estimation was not performed on this dataset. NUV-ECD monitoring enables to observe the clean and obvious changes in the spectra but it does not make possible to extract the absolute value of the concentration $[SS]_\infty$ at the endpoint of the reaction. Based only on the intensity of the molar ellipticity, it cannot be decided if the 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 5mm Ø NMR tube (~0.113 mM E19_SS, pH = 7, 15 °C, 18-fold TCEP) by recording ^{1}H-NMR resonances (**Figure 4/b**). By using both SS and SH state integrals of the signals at selected resonance frequencies (*e.g.* Hε^{Trp}), the ^{1}H-NMR driven quantitative analysis of the reduction was performed (**Figure 3/e**) and the rate constant was determined as k_1^{NMR} = 8.03E–4 Lmmol^{-1}min^{-1}. Although 18-fold excess of TCEP was used, ^{1}H-NMR data shows that at the steady state, $\partial[E19\_SS]/\partial t$ ~0 and $\partial[E19\_2SH]/\partial t$ ~0, reduction is incomplete and about 8% of E19_SS remains oxidized. Comparing the calculated reaction rates of the two methods (NMR & CD) shows that not only magnitudes are the same, but

values can be quite similar as well. The reduction rate of NUV-ECD measured at λ = 287nm is closest to k_1^{NMR} (**Figure 4/c**). 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 integrals of ^{1}H-NMR, it is possible to determine the rate of the conversion and see evidence for the reversibility of the redox system as well. Taking into account the incomplete conversion despite the presence of the 18-fold excess reducing agent, the role of the dissolved oxygen and the re-oxidation should also be included in the kinetic mechanism.

The concept of the reversible redox system

Physiologic solutions contain dissolved O_2 from the air and thus Cys –SHs of any protein might oxidize spontaneously to SS-bond(s). The apparent rate constant depends on several micro equilibrium constants, explicitly not elaborated here. [68] However, it certainly depends on how wide the conformational space of the reduced molecular fold is. Furthermore, the concentration of the dissolved O_2 (and thus, T and p), diffusion rate of TCEP, protein concentration are all rate influencing factors. As our model protein forms a coupled reaction cycle, once E19_2SH is oxidized, by the excess of TCEP E19_SS will be reduced, however, instantaneously reoxidized by the dissolved O_2. **(Figure 5)** Before exploring the mechanism of this redox cycle related electron transfer processes, it has to be mentioned 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 O_2 is consumed completely from the dissolved air. However, if the concentration of the reducing agent declines faster than that of O_2, then oxidation will dominate the process and spectral properties will change accordingly. **(Figure 5)** It is hard to *a priori* predict the endpoint of the latter process, as unlike the oxidized fold of a protein, the reduced one could have a multitude of backbone conformers in exchange at a various timescale of motion (*e.g.* µs to ms). Among these 3D-folds of the reduced state the "closed-
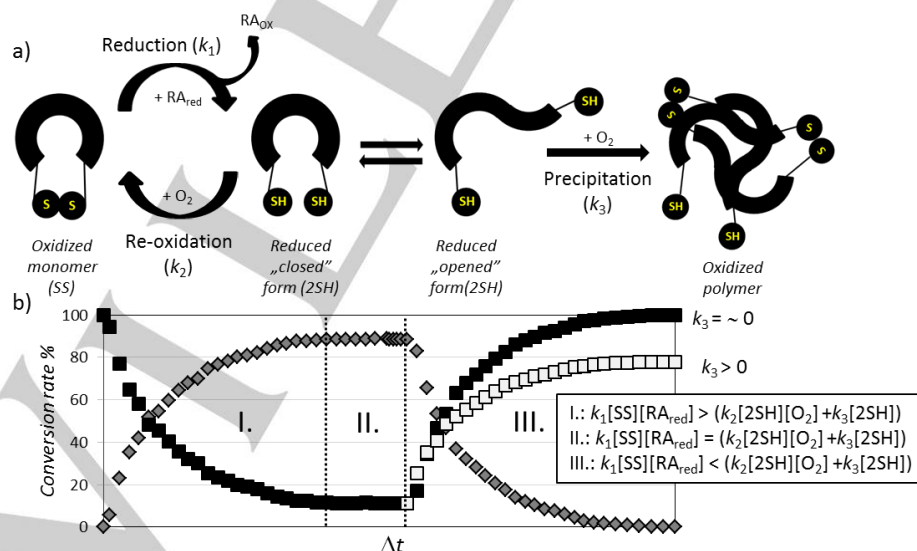
Figure 5 a) The redox cycle schematically: The oxidized state (SS) in the presence of reducing agent (e.g.: TCEP), becomes reduced (2SH), where "open & closed" conformers stand in equilibrium. In the case of the closed conformer, where the –SH groups are closely fixed to each other, intramolecular re-oxidation can occur in the presence of O_2, while the open conformer is more likely to aggregate due to intermolecular interaction. **b)** Three stages of the theoretical redox setups provide the state when reduction dominates the overall process (**I**), a steady state (**II**) and a state (**III**) where the excess of dissolved O_2 and the absence of reducing agent oxidize back the reduced state. Black square stands for the relative concentration of the oxidized form, gray diamond for the reduced form. If precipitation happens ($k_3 > 0$), then at the endpoint of the redox cycle the soluble protein concentration decreased compared to the initial one.

SH" forms (**SFigure 2/II**), where both the *C*- and *N*-terminus are close to each other, lead only to intramolecular re-oxidation. Furthermore, if "open-SH" backbone forms get highly populated (*e.g.* as is the case of E2_2SH **SFigure 2/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 µs to ms of motion was successfully attempted by using CPMG-NMR [69]. Here we give the example of the successfully characterize shorter derivative E5, and its variants: E5_SS and E5_2SH. We found that the backbone NHs of Glu3, Cys4, Val5, Arg6, Tyr8 & Cys25 residues (**Figure 1/b**) present only such slower motion and only in the case of the reduced form E5_2SH. Considering the fact that all these NHs are close to both Cys residues (**Figure 6**), the CPMG data suggested for the E5_SS↔E5_2SH molecular system that using an excess of the reducing agent, either E5_2SH presents alternative backbone structures interconverting at a slow exchange rate, or the remaining oxidized form in trace quantity (1-8 % see later the conversion rate) constantly exchanges states with the reduced form. The minor coexisting oxidized form 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 of the above described slow exchange; however, both scenarios of motion can occur in a concerted way.
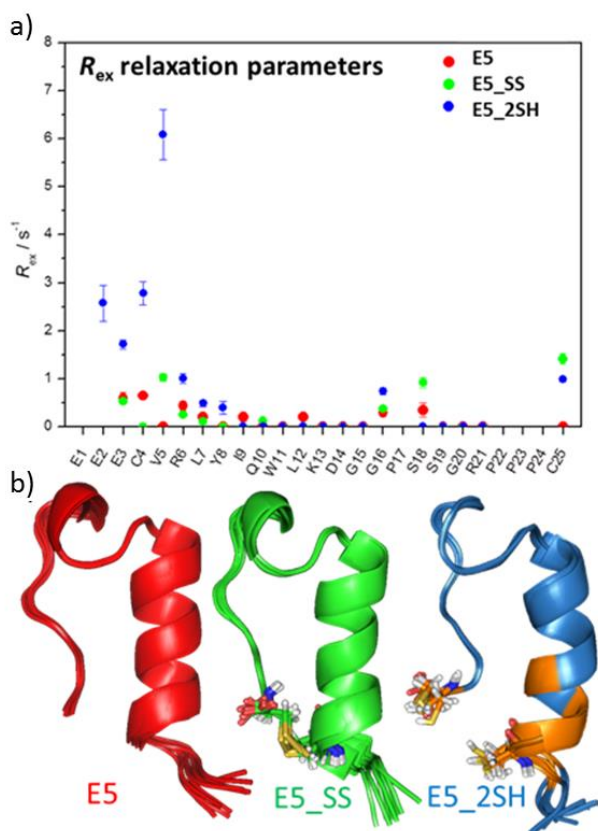

Figure 6. a) CPMG determined NH R_{ex} values of E5 (red), E5_SS (green), E5_2SH (blue), and **b)** their backbone structures, with the key Cys residues highlighted. Slow exchange was measured for backbone NHs of E3, C4, V5, R6, Y8 & C25 of E5_2SH only. Note that residues giving R_{ex} are in the proximity of the Cys residues colored orange.

Modelling of the SS-bond reduction kinetics

The SS-bond reduction by TCEP is a bimolecular nucleophilic substitution (S_N2) reaction [37]. Bimolecular by nature, both the concentration of the oxidized form of the protein [_SS] and that of the reducing agent (RA) determine the rate of the reduction. In an ideal case we should consider only the nucleophilic attack of the reducing agent (k_1), but as we explained previously, in practice we also have to take into account the back-oxidation (k_2) taking place simultaneously, and in some cases, depending on the size and shape of the protein, the precipitation (k_3) (**Figure 5**). The mechanism of the reduction therefore, can be described by the following equations:

$$SS + RA_{red} \xrightarrow{k_1} 2SH + RA_{ox}$$
$$2SH + O_2 \xrightarrow{k_2} SS$$
$$2SH \xrightarrow{k_3} SS_{precipitated}$$
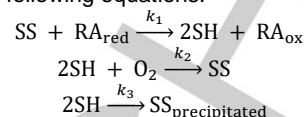
Fitting this model to the NMR determined concentration-time functions, k_1, k_2 and k_3 can be determined, and half-life times can be calculated. We focused on the determination of the reduction rate constants k_1, therefore, the sampling was more frequent in the reduction phase. (Stage I. **Figure 5**) As a result of the parameter estimation (see materials and methods) k_2 and k_3 are very often either negligibly small, or – due to the lack of enough data – cannot be confidently estimated. Comparing these key kinetic parameters allows us to describe and compare the reduction kinetics of the SS containing miniproteins under various experimental conditions. Some protocols of the literature apply extreme conditions to obtain short reduction times, such as high temperature, clearly unsuitable for keeping protein's integrity (*e.g.* 50-80°C), or > 20-fold molar excess of reagent. Carrying out reduction of E19_SS (0.8 mM) using 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 is distinctly different at 60 °C compared to that at 15 °C, presenting many more unfolded states. The folded fraction of E19_2SH is 64 % at 15 °C, while 41 % at 60 °C according to FUV-ECD. Instead of intramolecular re-oxidation, undesirable intermolecular re-oxidation might occur between particles. (Reducing E19_SS for 120 min, followed by centrifugation gave practically zero soluble protein concentration.) In general, reduction and re-oxidation at higher *T* (*e.g.* ≥ 60°C) is expected to be less effective, being accompanied by multiple side-reactions such as β-elimination [70] (occurring already at a lower *T*) [71][72] , racemization [73][74] aggregation, *etc.*. In principle, the reduction rate can be enhanced at lower *T* by increasing TCEP molar ratio (15–20-fold molar excess); however, this also triggers obscure unwanted processes (**SFigure 5**). Experiments were repeated at different temperatures (15 °C, 25 °C, 37 °C) using 0.8 mM protein and 18-fold excess of TCEP. (**Table 2, SFigure 6**) *Arrhenius*-equation allows deriving the activation energy (E_a) of the redox reaction, resulting in ~ 44.3 kJ/mol value. For comparison, the

Table 2. Kinetic parameters of the temperature depending E19_SS reduction using 0.8 mM protein and 18-fold excess of TCEP. For the detailed results of parameter estimation see **SFigure 6**.

T [°C]	elapsed time to steady state [h]	Conversion rate [%]	k_1 [L][mmol]$^{-1}$[min]$^{-1}$	$t_{1/2}$ [min]	relative standard deviation of k_1 [%]
15	~ 15	92	3.05E−04	181	3.27
25	~ 6	94	7.68E−04	72	9.78
37	~ 4	94	1.15E−03	48	1.68
60	n.d.	n.d.	n.d.	n.d.	n.d.

activation energy of the thiol-disulfide exchange between methylthiolate and oxidized DTT was calculated as 62 kJ/mol. [75] Both FUV-ECD and NMR derived structural information support the high conformational similarity between E19_SS and E19_2SH, therefore E_a 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 derived signal-integral analysis, the reduction was almost complete (~

94%) and no sign of precipitation was detected neither at any temperature.

The above described NMR methodology provides high resolution information about the reduction mechanism compared to that of the more rapid NUV-ECD approach and thus, the details of the reduction of all four -SS- protein models were completed by NMR.

Additional experiments were performed to investigate the effect of the protein/reducing agent ratio and found for the current protein models that i) for the *initial protein concentration* (**Table 3/a**) the reduction of the shortest E2 peptide in two

Table 3. SS-bond reduction as a bimolecular nucleophilic substitution (S_N2) reaction both the concentration of the initial oxidized protein and the reagent influences the reduction rate. To quantify this extent **a)** the initial peptide concentrations were altered for the shortest E2_SS and the longest E19_SS reductions. The experiments were carried out at 15 °C, for E2_SS reductions ~2 fold of TCEP excess were used while in case of E19_SS ~16 fold excess of TCEP. **b)** In case of E11_SS the excess of TCEP was altered, while concentration of E11_SS was set to 0.73 mM and the temperature to 15 °C. For the detailed results of parameter fitting see **SFigure 7**.

a)	[SS] [mM]	elapsed time to steady state [h]	Conversion rate [%]	k_1 [L][mmol]$^{-1}$[min]$^{-1}$	$t_{1/2}$ [min]	relative standard deviation of k_1 [%]
E2_SS	1.05	~ 15 min	100	1.37E–01	3	16.92
E2_SS	1.74	< 5 min	100	2.59E–01	~ 1	15.37
E19_SS	0.12	~ 40 h	92	8.03E–04	466	12.37
E19_SS	0.81	~ 15 h	92	3.05E–04	181	3.27

b)	[RA$_{red}$] [mM]	elapsed time to steady state [h]	Conversion rate [%]	k_1 [L][mmol]$^{-1}$[min]$^{-1}$	$t_{1/2}$ [min]	relative standard deviation of k_1 [%]
E11_SS	17.61	~ 28 h	91	1.64E–04	337	2.98
E11_SS	8.90	~ 35 h	89	2.61E–04	423	4.80

different concentration was rather fast using as little as only 2-fold excess of TCEP, while in the case of E19_SS, the difference between different initial protein concentration is found to be more relevant. **(SFigure 7)** Furthermore, ii) the *excess of TCEP* (**Table 3/b**) is such that when increasing the TCEP excess by 2 fold, it makes the reduction of the E11_SS 20% faster. **(SFigure 8)** Contrary to reduction of E19_SS using 17-fold TCEP excess **(SFigure 5)** no additional side product was observed. In conclusion, it might be sensible to apply moderate/milder conditions to avoid any undesirable reaction although to reach a complete reduction might take longer time.

SS-bond reduction kinetics influenced by steric factor

A proper reduction protocol was required to unambiguously determine the 3D structures of the above introduced pure reduced states. Thus in line with the above only mild conditions (15 °C and 2-fold molar excess of TCEP) was used for reduction of the four different miniproteins. The knowledge of the structural properties and the reduction rates under the same conditions allowed us to elucidate the basis of the observed magnitude differences in their reduction rates. We have found that at *T*= 15 °C, k_1 of these four different length model proteins comprising of otherwise identical core structures are indeed different: their k_1 and $t_{1/2}$ values strongly depend on their size and/or molecular weights. It seems as if 'cutting back' the α-helical segment size strongly affects the SS-bond reducibility, although the SS-bonds of all four models are on their surfaces. (**Figure 7/a**). To our
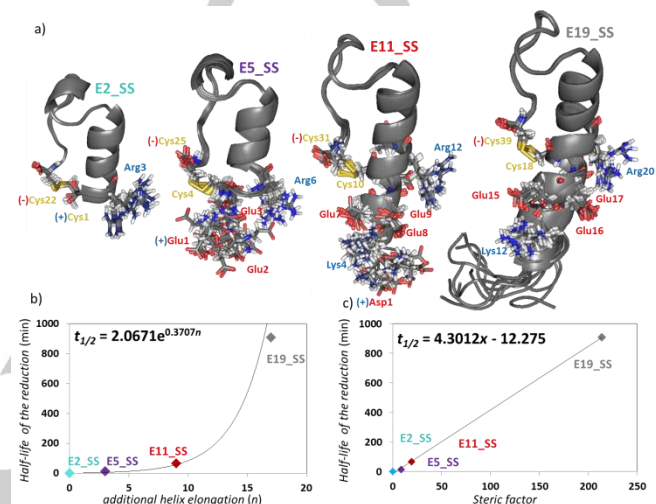
Figure 7. a) 10 superimposed structural ensembles of E19_SS, E11_SS, E5_SS & E2_SS. Note that all model proteins have their SS-bond at the surface but their N-terminus is of different length and charges thus, affecting the SS-bond reducibility. Charged residues at pH= 7, close to the reaction centre are explicitly depicted: the negatively charged side chains are highlighted by red, while the positive ones by blue. The C-terminus negative charge of the COO$^-$ is marked by (–) and the amino group -NH$_3^+$ of the N terminus by (+). Correlation between the reduction half-life **b)** vs. helix length and **c)** vs. steric factor (2-fold excess of TCEP, 1.7 mM protein, 15 °C) is reported. (**SFigure 10** shows the correlation between the reduction half-life vs. helix length in case of DTT reductions.)

great surprise, we have got three magnitude differences between their reduction rate constants (**Table 4**). While the reduction of E2_SS is still extremely fast, $t_{1/2}^{E2\_SS} < ~ 1$ min, that of E5_SS occurs on the time scale of minutes: $t_{1/2}^{E5\_SS} ~ 14$ min.

Table 4. Kinetic parameters of the SS-bond reduction of the 4 model proteins. For each reduction, 1.7 mM protein concentration was used with two fold excess of TCEP and DTT at 15 °C. For the detailed results of parameter estimation see SFigure 8-9. NMR 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]$^{-1}$[min]$^{-1}$	$t_{1/2}$ [min]	relative standard deviation of k_1 [%]	elapsed time to steady state [h]	Conversion rate [%]	k_1 [L][mmol]$^{-1}$[min]$^{-1}$	$t_{1/2}$ [min]	relative standard deviation of k_1 [%]	Outer helix length (*i*)	RMSD of outer helix	$\frac{\Sigma CSD_{(i)}^{H\alpha}}{i}$	Steric factor[a]
E19_SS	~ 76 h	87	2.71E–04	909	7.30	n.d.	n.d.	n.d.	30545[b]	n.d.	17	1.41	0.11	213.74
E11_SS	~ 5–6 h	94	3.68E–03	67	2.61	138 h	~ 84	1.52E–04	1659	45.152	9	0.55	0.26	19.26
E5_SS	~ 1 h	93	1.85E–02	14	3.35	9–10 h	95	2.18E–03	115	1.064	3	0.39	0.14	8.56
E2_SS	<5 m	100	2.59E–01	~ 1	15.37	5h	100	4.04E–03	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 $\frac{\Sigma CSD_{(i)}^{H\alpha}}{i}$

[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.

E11_SS – elongated by 6 residues (1.5 turns of α-helix) with respect to E5_SS – makes a ~ fourfold increase in $t_{1/2}$ ($t_{1/2}^{E5\_SS}$ ~ 14 min → $t_{1/2}^{E11\_SS}$ ~ 67 min). Finally, the unstructured short octapeptide tail HAEGTFTS- further lengthens $t_{1/2}$ by ~ 13-fold ($t_{1/2}^{E11\_SS}$ ~ 67 min → $t_{1/2}^{E19\_SS}$ ~ 909 min). Conversion rate was close to complete for the shorter peptide of E2_SS, while the reduction of E19_SS was completed only to 88%.The kinetic parameters of all four model proteins were determined by using

two-fold molar excess of DTT, at *pH*= 7 and *T*= 15 °C. Mechanism of the SS-bond reduction by DTT is also S_N2 [76] but the determined $t_{1/2}$ values are significantly longer than those obtained using the same molar excess of TCEP, however the observed overall tendency and conclusion looks the same (**Table 4**).

As the well folded Trp-cage motive are identical (based on their CSD Cage values (**Table 1**)) in all of the 4 model proteins,
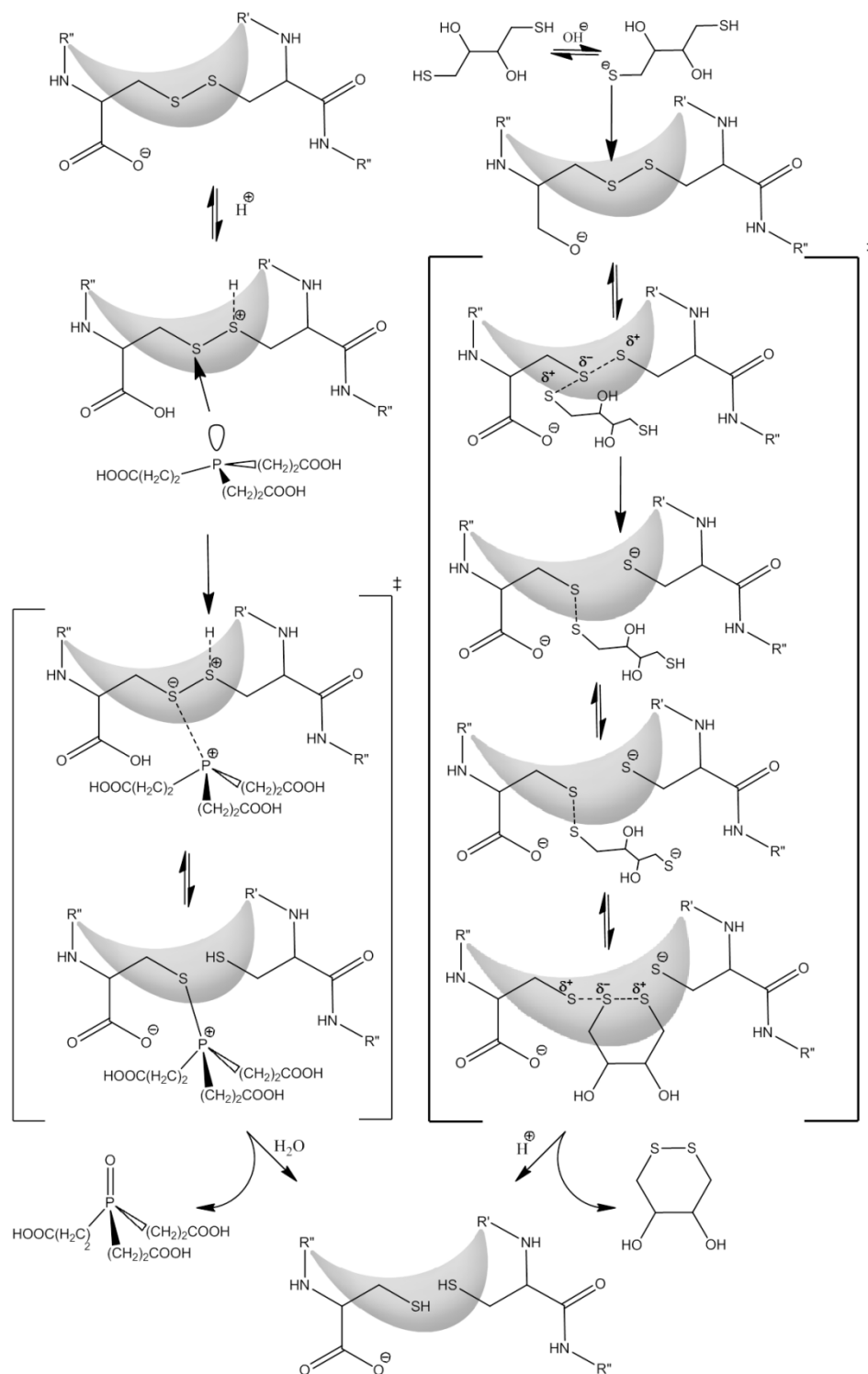
Figure 8. The generalized mechanism of TCEP and DTT assisted mechanism of SS-bond reduction in proteins. Functional group R- stands for the N-terminus of the protein systematically elongated here: in E19_SS the R-group is equal to **H⁺-HGEGTFTSDLSKQMEEE-**, in E11: **H⁺-DLSKQMEEE-**, in E5_SS: R = **H⁺-EEE-**, while in E2_SS simply with **H⁺**. A brief description of the detailed reaction mechanism is provided both for TCEP and DTT in the text.

the observed k_1 differences must be associated with the structural properties of their α-helices and the eventually appearing unstructured tail. Though the dataset is limited ($n = 3$ or 4), as the most simple approach, the length of the α-helix (n) and the half-lifes ($t_{1/2}$) of the reduction could be correlated leading to an exponential dependence in case of using 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 reducing agent. (**Figure 7/b**) To take into account the additional structural descriptors for a more complete characterization, we derived for these protein models the steric factor (x), beside the length of the N-terminals. The helicity ($\frac{\Sigma CSD_{(i)}^{H\alpha}}{i}$) and the bulkiness (RMSD) of the outer helical part were both taken into account. **(Table 4)** We observed a linear dependence of the steric factors on the reduction half-lifes as function of the length of the N-terminals. (**Figure 7/c**) Some, but not all of the above k_1 ($t_{1/2}$) differences can be explained by structural differences of the outer helix, as both their solvent exposure and local charges around the SS-bonds are different also. Below we summarize a mechanistic explanation, including all of these factors and viewpoints. (**Figure 8**)

SS-bond reduction rate determining steric and electronic factors

Beside the steric effect of the helical part emphasized above the S_N2 mechanism of the TCEP driven reduction has to be discussed in terms of electrostatic effects. [19] In general, the attack is more favorable and effective on those structures where 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 COOH/COO$^-$ is low: 1/12000. The rate determining step is the cleavage of the SS-bond. [77] During the S_N2 reaction, the nucleophilic P-atom of the TCEP attacks one of the SS-bonds, forming a thiophosphomium salt (an S$^-$-P$^+$ ion-pair complex) (**Figure 8**).

Nucleophilic attack (n → σ*) is facilitated by the favorable arrow shape (tetrahedral: 105°) steric arrangement of the non-bonding electron-pair of the P-atom of TCEP. Because the three -COOH groups of TCEP are H-bonded at $pH = 7$, though not taking part of the reaction, their orientation enhance the P-nucleophile [38] and also increase the proton activity. The main portion of the activation Gibbs-free energy of the reduction is consumed by the splitting of the SS-bond and not by the steric rearrangement of the intermediate structure. [78] The better (the thiol and) the zwitterion solvated, the lower the activation Gibbs-free energy of the reaction is. Next the positively charged -S-P$^+$-[(CH$_2$)-COOH]$_3$ complex hydrolyzes rapidly and results in the phosphine oxide and the free –SH groups of the protein.

However, all charged amino acids in the sequential and /or spatial vicinity of an SS-bond should be considered, as both charged and aromatic side chains can participate and thus, intimately influence the efficacy of TCEP mediated reduction (**Figure 7/a**). The nucleophile phosphine attacks the C-proximal cysteine because the intermediate cation can be stabilized by the proximal COO$^-$ group of the C-terminal cysteine. A positive charge nearby the SS-bond could enhance the reaction *via* electrostatic compensation of the *N*-proximal leaving thiolate group, while a negative charge might slow down the S_N2 reaction. [79][80] Direct through-bond effects of any charged side-chain can be ignored as they are separated by several σ-bonds from the negative COO$^-$. While the inductive or direct σ-bonds effects are negligible, both steric and spatial electrostatic effects in the vicinity of *N*-proximal cysteine play a major role on the reduction rate.

The positively charged Arg at $pH = 7$ nearby the SS-bond in the inner helix may facilitate the reduction; however, it is away from the SS-bond (**Figure 7/a**), thus a direct charge-controled interaction is less likely to occure. On the other hand, the positivly charged *N*-terminus -NH$_3^+$ can directly catalyze the instantaneous reduction [81] of E2_SS ($t_{1/2}^{E2\_SS} \sim 1$ min) as H-N-

C$^\alpha$–C$^\beta$–S of the cysteine forms a 5-member pseudo-ring facilitating intramolecular N→S proton transfer. [82] Thus, when TCEP attacks, presumably these ideal local electrostatic compensations stabilize the intermedier thiophosphonium salt by shifting the reaction equilibrium towards the splitting of the SS-bond. Not alone the optimally positioned charges nearby the SS-reaction center make the reduction conditions favorable for E2_SS. As the *N*-proximal leaving thiolate anion is positioned just at the *N*-terminal of the well folded α-helix, the positive charge of the α-helix macro-dipole stabilizes the intermediate structure. [83][84] Moreover, due to the small protein size the SS-bond is the most exposed to solvent and reagent in E2_SS.

As the *N*-terminus gets elongated lengthening the α-helix from E2_SS toward E19_SS, the "catalyzer" –NH$_3^+$ group of the *N*-terminus gets further away from the SS-bond and the effect of the macro-dipole gradually vanishes and thus, reduction rate is reduced ($t_{1/2}$ increases). (**Table 4**) The role of this positive charge was directly probed by acetylating the *N*-terminus, Ac-E2_SS, and as expected, half-time of the reduction increased significantly: $t_{1/2}^{E2\_SS} \sim 1$ min → $t_{1/2}^{Ac-E2\_SS} \sim 8$ min. (In both cases 1.7 mM protein concentration and 2-fold excess of TCEP was used.)

The *N*-terminal elongation of E2_SS by three Glu residues results E5_SS. As expected, reduction rate gets slower: $t_{1/2}^{E5\_SS} \sim 14$ min. Although only a tripeptide is added to the dynamic *N*-terminus, it makes harder for both reagent and/or solvent molecules to reach the SS-bond. In addition, the 3D-structure **(Figure 7/a)** 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 neutralizes the catalytic effect. NMR structure ensemble shows a 3.7 Å ↔ 10.7 Å distance fluctuation between 4Cys Cβ and 1Glu NH$_3^+$, while that of 4Cys Cβ and 1Glu COO$^-$ fluctuates between 3.4 Å and 12.4 Å. (**Figure 7**) Thus, SS-bond protonation requires the active contribution of the medium; but the proton transfer is perturbed by the proximity of the glutamates negative sidechains.

Further elongation of E5_SS by hexapeptide -DLSKQM- leads to E11_SS. Using the same conditions, the reduction of this even larger model protein occur slower ($t_{1/2}^{E11\_SS} \sim 67$ min). The glutamate sidechains of E5_SS are more oriented by the longer α-helix of E11_SS (**Figure 7**): while 8Glu$^-$ turns outward, both 7Glu$^-$ and 9Glu$^-$ flanks from both side the SS-bond. 7Glu$^-$ with 4Lys$^+$ and 9Glu$^-$ with 12Arg$^+$ are capable to form salt-bridges in close vicinity, thus could partly compensate the slowing effect of the flanking negative sidechains. E11_SS has the more compact α-helical segment among these model proteins, as perhaps the negatively charged side-chain of 1Asp$^-$ self-compensates with its own *N*-terminus –NH$_3^+$. E11 was found to be more helical than the longer E19 [56], thus we find here that both E11_SS and E11_2SH have more compact α-helices than E19_SS and E19_2SH, according to both the $\frac{\Sigma CSD_{(i)}^{H\alpha}}{i}$ NMR-measure and the FUV ECD-spectral properties. We think that beside the partly compensated negative electrostatic effect(s), mainly the steric effects of the elongated and stiffer α-helix causes the longer $t_{1/2}^{E11\_SS}$ with respect to $t_{1/2}^{E5\_SS}$.

Finally E11_SS elongated by the -*HGEGTFTS*-octapeptide results in E19_SS, the largest model protein used here for which the longest half-life ($t_{1/2}^{E19\_SS} = 909$ min) is measured. As E19_SS has the same electrostatic pattern in the vicinity of the SS-bond as that of E11_SS nevertheless, its reduction rate is about 15 times slower than that of E11_SS. As -*HGEGTFTS*- segment is far from the SS bond, it cannot influence the reduction by electrostatic interaction, **(**$d_{7Thr-18Cys}$: 11-14 Å, **Figure 7)** its higher internal dynamics (low S^2 value) [56] – as a steric effect – must slow 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 by using alternative reagents

There are a few distinct differences in terms of general mechanism of disulfide reduction by TCEP and DTT. **(Figure 8)** *i)* As an initializing step, the deprotonation of the thiol group of DTT is required for a successful nucleophile attack, which depends on the *pH* of the medium. According to the Henderson-Hasselbach equation [85] and the acidic dissociation constant of the DTT (pK_{a1} = 9.2 and pK_{a2} = 10.1), at *pH* = 7, the deprotonated thiolate concentration is about 3–4–fold lower than that of the overall DTT concentration. After the successful nucleophile attack on the SS-bond, a linear -S-S-S- transition complex has to be formed in which the negative charge is located on the two leaving S-atoms. [86] An intramolecular protonation, as for TCEP, also stabilizes the thiol anion leaving-group when DTT is used, and thus enhances the reaction rate. Therefore, a positive inductive/steric effect increases, while a negative decreases the reduction rate. *ii)* Contrary to TCEP, the active species of DTT bears a negative charge. Therefore, charged amino acid side chains close to the SS-bond will directly affect the attack of the nucleophile reducing agent. In line with these, both the negative *C*-terminus and the SS-bond flanking glutamate side chains repel DTT contributing to a significant and large-scale reaction rate decrease. **(Table 4)** *iii)* Moreover, the 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, while the DTT closes into a six-member ring. **(Figure 8)** All these factors jointly decrease reduction rate when DTT is used compared to TCEP (**Table 4**). However, these considerations stand contrary with general practice: several proteins of various numbers of disulfide bonds per molecule – like α-lactalbumin, lysozyme, and oxytocin – can completely be reduced in 5 minutes by 10 mM DTT in *pH* = 5.5 ammonium-acetate buffer at 70 °C. [87]

Spontaneous SH re-oxidation accompanied by polymerization

The incomplete conversion despite the presence of a large excess of the reducing agent showed evidence for the re-oxidation of the reduced disulfide bond of the studied model systems. To study this process in more detailed, the *in situ* re-oxidation of the DTT reduced protein samples at room temperature in a sealed NMR tubes (*pH* = 7, 15 °C, 2-fold excess of DTT) was monitored for several weeks. Spontaneous re-oxidation of the E2_2SH, E5_2SH, E11_2SH by the dissolved O₂ was obvious after 4 weeks. **(Figure 9)** The re-oxidation rates (k_2) have comparable order of magnitude to the reduction rates (lower by one order of magnitude), but the re-oxidation has a pronounced role only after reaching the steady-state, when the concentration of the already reduced peptides become significant.

Re-oxidation can take place both intra- and inter-molecularly. While the former leads to the decrease of overall conversion rates the latter results in formation of random molecular cluster which may lead to precipitation. According to our semi-quantitative analysis based on the recorded ¹H-NMR spectra, the integral changes of the Trp Hε1 resonances both in the oxidized and reduced form of the protein during reduction with DTT show a decrease in concentration over the observed period of redox time both for E2_SS and E5_SS. Precipitation can be more intense if protein concentration is higher. According to our present observations, increasing the length of the α-helix within the Trp-cage proteins stabilizes the soluble protein fraction. This means, that the elongated *N*-terminus, namely the outer helix in case of E11_2SH, effectively shields the free SH-
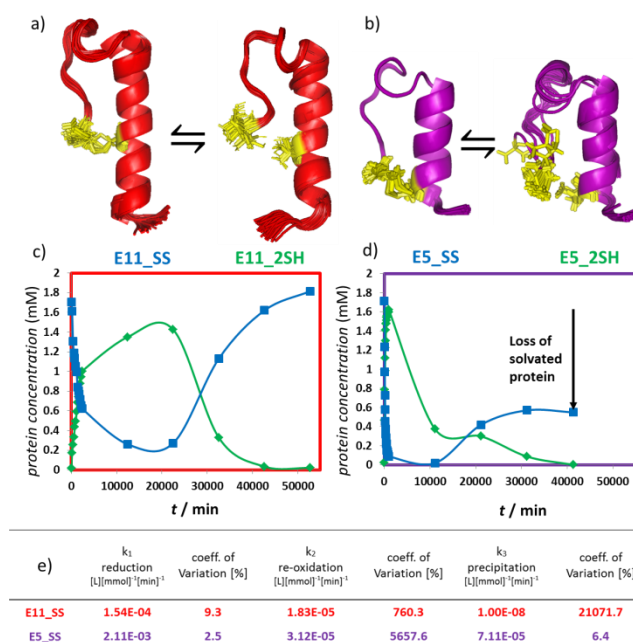
Figure 9. The 50 membered structure ensemble of **a)** E11_SS ⇌ E11_2SH and **b)** E5_SS ⇌ E5_2SH. The fold of E11_2SH is more compact than that of E5_2SH, which has more "opened' conformers, where the Cys residues get far from each other. This allows intermolecular re-oxidation instead of intramolecular one. 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 case of E5_SS ⇌ E5_2SH, the initial concentration decreased by 68%, while at the end of a complete redox-cycle, the concentration of E11_SS ⇌ E11_2SH remains the same. **e)** Estimated parameters of the complete redox cycles. (k_1 values are slightly different compared to those in **Table 4,** where the estimation comprises data only for phase 1.) Note that, in these long-term experiments, the rate of O₂ diffusion characterized by the rate constant k_4 was also involved. **SFigure 11** contains all data of parameter estimation for E11_SS, E5_SS, E2_SS.

groups of the reduced protein and thus, prevents any intermolecular re-oxidation, while shorter variants like E2_2SH and E5_2SH yield to a significant amount of polymer formation. Due to the diversity of opened 3D-folds of both E5_2SH, E2_2SH the 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 closure takes place. In case of E5_2SH intermolecular SS-bond formation is allowed, but may limited just by the Brownian motion and the concentration. Comparison of the polymerization rates ($k_3^{E11\_SS} < k_3^{E5\_SS}$) with the different *N*-terminal lengths supports also this concept. **(Figure 9/e)**

E2_SS was *N*-acetylated to eliminate the reduction rate enhancing effect of the positively charged *N*-terminus, -NH₃⁺, in the vicinity of the SS-bond. Upon acetylation, $t_{1/2}$ has indeed increased ($t_{1/2}^{E2\_SS}$= ~1min → $t_{1/2}^{E2\_SS}$= ~8min); but in addition, the reaction reached 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 started to decrease, with a foamy precipitate gradually sedimenting in the NMR tube. The isolated and HPLC purified precipitant was identified as a polymer of the parent miniprotein by MS. (**SFigure 12**) Oligomer formation and soluble protein concentration decrease was more advanced for Ac-E2_SS compared to E2_SS. (**SFigure 13**) Due to the absence of the shielding effect of the outer α-helix, the free thiol moiety of the *N*-terminus is accessible for additionally reduced peptides where the two free SH-groups after oxidation hook peptid chains together. The polymer can grow until another free *N*-terminus and acetylated *C*-terminus thiol-containing peptide closes

polymerization. In addition, for Ac-E2_2SH, the intramolecular N→S acyl transfer could take place [88] blocking some of the SH-groups promoting oligo- and polymerization *via* intermolecular SS-bond formation.

Conclusions

Disulfide-bond cyclized Exenatide derivate and its variants were synthesized. Both the oxidized (E19_SS) and the reduced form (E19_2SH) 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) comprise the very same Trp-cage/SS/SH-bond motif as their core structures. The SS-bond stabilized model proteins showed improved thermo-stability and 3D-fold compactness with respect to their reduced and parent forms. Key residues for the receptor binding remain in position in all of these models; therefore E19_SS might be promising agonists for GLP-1 receptor and thus, for developing as lead compound for type 2 *Diabetes Mellitus*.

The reduction rate of the E19_SS was found to be unexpectedly slow compared to literature data. The reaction took hours ($t_{1/2}$ = 48 min) even at 37 °C, although the protein is small and it's single SS-bond is exposed at the surface and is thus accessible for reducing reagents. All four Trp-cage variants studied in this work have an almost equally compact core-structure with α-helical segments of different length and internal mobility. By performing a complete NMR-based structure elucidation, we found that the progress of reduction can be monitored by ^{1}H-NMR using selected resonance frequencies. We have established that these four model proteins of different α-helical lengths have significantly different reduction rate constants. Although it is generally complicated to discriminate each factor affecting the SS-bond reduction rate, the present set of miniproteins of different length enables to decipher them separately. We have focused special attention to the importance of the intramolecular protonation of the SS-bond: a step greatly enhancing the reaction rate. Using CPMG-measurements we found that at the steady state, selected residues at the vicinity of the SS-bond present a slow exchange at the μs-ms timescale of motion. This redox cycle lasts as long as the active reducing agent can be found in the solution. We found that structural, steric and electrostatic factors influence the reduction rate greatly, resulting in almost 3 orders of magnitude differences in reduction half-times ($t_{1/2}$), for otherwise structurally similar and globularly folded model proteins.

Notably, in addition to the intramolecular re-oxidation within the redox cycle, intermolecular oxidation could also occur. The rate of these two concerted reactions depend on 1) the internal dynamics of the backbone conformers at the SS-bond proximity and 2) the shielding effect of the α-helix on the SS-bond. Intramolecular *N→S* acyl transfer in Ac-E2_SS inhibits the intramolecular, while increases the intermolecular re-oxidation, which leads to oligo- and polymerization.

We found that easy to collect NUV-ECD spectral properties are indeed useful for monitoring the SS→SH reaction even quantitatively, without the time-consuming assignment of the high resolution NMR data. If the SS-bond is situated in vicinity of an aromatic cluster, NUV-ECD spectral changes can be used to monitor the transformation being proportional to the extent of the reduction and clearly signalling when steady-state is reached. Thus, we encourage the use of CD-spectroscopy for monitoring protein reduction rate when manufacturing 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

Abbreviations

Parent model proteins: E2, E5, E11 and E19; SS-bonded / oxidized model proteins: E2_SS, E5_SS, E11_SS and E19_SS; Reduced model proteins: E2_2SH, E5_2SH, E11_2SH and E19_SS; TCEP: Tris-(2-carboxyethyl)-phosphine; DTT: 1,4-Dithio-D-threitol; ECD: Electronic circular dichroism spectroscopy; FUV: Far UV range; GLP-1: Glucagon-like peptide-1 IgG: Immunoglobulin G; NUV: Near UV range; RMSD: Root-mean-square deviation of atomic positions; S-S, SS: disulfide bond; SCS: Secondary chemical shift; CSD: Chemical shift deviation; CPMG-effect: *Carr-Purcell-Meiboom-Gill*-effect / experiment; NMR: Nuclear magnetic resonance spectroscopy; Trp-cage, TC: Tryptophan-cage

Electronic Circular Dichroism:

Far-UV ECD spectra were recorded on a Jasco J810 spectrophotometer using 1.0 mm path length cuvette with protein concentration 20-30 μM. Data accumulation were performed over a range of 185-260 nm with 0.2 nm step resolution where the scan rate 50 nm/min with a 1 nm bandwidth. The spectra accumulations were absolved between 5°C – 85°C per steps of 5°C. The temperature was controlled by a Peltier-type heating system. Each spectrum baselines were processed by subtracting the solvent spectra from that of the protein and the raw ellipticity data were converted into mean residue molar ellipticity units $[\Theta]_{MR}$.

Reduction monitoring by Near-UV ECD: spectra were recorded on a Jasco J810 spectrophotometer using 10 mm path length cuvette with protein concentration 120-150 μM. Data accumulation were performed over a range of 240-325 nm with 0.2 nm step resolution where the scan rate 50 nm/min with a 1 nm bandwidth. The sample was tempered by a Peltier-type heating system. Each spectrum baselines were processed by subtracting the solvent spectra from peptide spectra and the raw ellipticity data were normalized by the concentration $[\Theta]$. The reduction was followed for 75 hours. Each intensity $[\Theta]$ at 266nm, 281nm, 287nm & 293nm were converted into concentration using the following equation $(t) = \frac{A_\infty - A}{A_\infty - A_0}[SS]_0$.

Nuclear Magnetic Resonance:

All ^{1}H experiments were performed on a Bruker Avance III 700 MHz equipped with z-gradient 5-mm probe-head operating at 700.13 MHz for ^{1}H, while ^{31}P experiments were carried out on a Bruker Avance 250 spectrometer with 5mm SB quad probe-head.

Monitoring reduction kinetic. Peptide samples were prepared between 0.8-1.8 mM dissolving in 600μl of 50mM, *pH*=6.95 NaH$_2$PO$_4$-Na$_2$HPO$_4$ buffer, with 10% D$_2$O. 0.1 M NaOH was used to set the *pH* to 7. DSS was added as the internal proton reference standard set to 0.0 ppm under all conditions. [1]H-[1]H 2D-homonuclear spectra were recorded from the oxidized peptide, thereafter adding a different excess of 0.5M TCEP or DTT the reduction was observed by 1D ^{1}H spectrum series (ns=64 or 128 scans), finally at the endpoint [1]H-^{1}H homonuclear 2D spectra were recorded from the reduced peptide. Data sets were processed using TopSpin 3.2. The conversion rate was determined using the relative integral of the Trp Hε1 peak in the oxidized (Int_{OX}) and reduced (Int_{RED}) form. Each integral were normalized to the integral of the DSS. The concentrations were determined by the ratio of the oxidized and reduced integrals and the initial protein concentration.

Structure determination: Proton NMR assignations were completed using ^{1}H-^{1}H COSY and [1]H-^{1}H TOCSY spectra, then the distance restraints were determined based on [1]H-^{1}H NOESY spectra. Spin locks for [1]H-^{1}H TOCSY were 80ms while the mixing time for [1]H-^{1}H NOESY were 150ms. CCPNMR [89] was used for resonance assignment, cross-peak calibration and

structure refinement. CNS Solve 1.3 [90], Aria 2.0 standard iteration protocol and water refinement were used for 10-membered structure ensemble calculation. All structural figures are illustrated using Pymol software.

CPMG effect: Backbone ^{15}N-longitudinal (R_1) and transverse (R_2) relaxation rates and the heteronuclear $^{1}H-^{15}N$ cross-relaxation rate constant (NOE) of E5, E5_SS, E5_2SH were measured at 288K. For each cross-peaks (*i*) $R_{2,i}$ values were calculated using the equitation below:

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

where I_i is the intensity of the given cross-peak in the i-th spectrum, I_{ref} the intensity of the given cross-peak in the reference spectrum, t_{CPMG} is the relaxation period of the CPMG measurement. The R_2 values per residues were plotted against the ν_{CPMG} (Hz). Quantitative analysis of the CPMG graph results those residues that show CPMG-effect in the protein.

Peptide synthesis and purification:
Proteins were prepared by standard solid-phase peptide synthesis or bacterial expression methods as we published previously. [91] Proteins were purified by reverse-phase HPLC using C-18 column using a gradient of water/acetonitrile eluents. (Eluent A: 0.1% TFA in water and 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 the following:

$$SS + Red \xrightarrow{k_1} 2SH + Ox$$
$$2SH + O_2 \xrightarrow{k_2} SS$$
$$2SH \xrightarrow{k_3} \varnothing$$
$$\varnothing \xrightarrow{k_4} O_2 ,$$
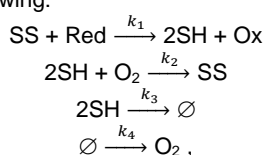
where SS is the reduced model protein with an intramolecular S–S bond, 2SH is the same protein with the S–S bond reduced to two –SH groups, and the symbol $\varnothing$ means a different phase than the reaction mixture; *i. e.* the polymer aggregate as a sink in the first and the gas phase as a source in the second case. Note that in some cases – where polymer precipitation (k_3) and/or oxygen diffusion (k_4) from the gas phase proved not to be present (indicated by largely non-significant estimated parameters concerning these processes) – these steps have been omitted from the fitted mechanism.

For the 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 suggesting Student distribution with $n - p$ degrees of freedom, where n is the number of data in the concentration *vs.* 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 – using this mechanism – the measured ($[SS]_{0,meas}$) and calculated ($[SS]_{0,calc}$) initial concentrations of the model proteins were in very good agreement. From the kinetic analysis, the initial

concentration of oxygen ($[O_2]_0$) could also be estimated, except for one case where the uncertainty of this parameter was very large, due to the lack of enough experimental data.

As the reduction follows second-order kinetics, the half-life ($t_{1/2}$) of model proteins depends also on the actual concentration of the reducing agent in the reaction mixture as follows:

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

where k_1 is the rate constant of the reduction, $c_{SS,0}$ is the initial concentration of the model protein, and $c_{Red,0}$ that of the reducing agent. (Note that this formula is valid only if $c_{Red,0}$ is greater than $c_{SS,0}$ – as in the current case. If $c_{Red,0}$ exceeds $c_{SS,0}$ but it is not higher than twice the value of $c_{Red,0}$, then the two initial concentrations should be flipped in both the difference and the fraction. If $c_{SS,0}$ exceeds $c_{Red,0}$ by more than a factor of 2, then the SS protein concentration cannot become as low as half of the initial concentration, due to the reduction.) For this reason, the half-life is less indicative concerning the rate of the hydrolysis; the correct comparison of the rates can be made based on the rate constant(s) of the second-order reaction(s).

[1]   C. Grek, D.M. Townsend, *Endoplasmic Reticulum Stress Disease* **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 Letters* **2012**, 586, 2826-2834.
[4]   H. A. Khan, B. Mutus, *Frontiers in Chemistry* **2014**, 2.
[5]   G. Batta, T. Barna, Z. Gaspari, S. Sandor, K. E. Koever, U. Binder, B. Sarg, L. Kaiserer, A. K. Chhillar, A. Eigentler, E. Leiter, N. Hegedues, I. Pocsi, H. Lindner, F. Marx, *Febs Journal* **2009**, 276, 2875-2890.
[6]   J. W. H. Wong, S. Y. W. Ho, P. J. Hogg, *Molecular Biology and Evolution* **2011**, 28, 327-334.
[7]   T. J. Bechtel, E. Weerapana, *Proteomics* **2017**, 17.
[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, *Proceedings of the National Academy of Sciences of the United States of America* **2003**, 100, 5754-5759.

[10] M. A. Wouters, S. W. Fan, N. L. Haworth, *Antioxidants & Redox Signaling* **2010**, 12, 53-91.

[11] A. Holmgren, *Annual Review of Biochemistry* **1985**, 54, 237-271.

[12] C. H. Lillig, C. Berndt, *Antioxidants & 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, *Nature Reviews Cancer* **2013**, 13, 425-431.

[15] I. Azimi, J. W. H. Wong, P. J. Hogg, *Antioxidants & Redox Signaling* **2011**, 14, 113-126.

[16] U. Siebenlist, G. Franzoso, K. Brown, *Annual Review of 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, *Antioxidants & Redox Signaling* **2009**, 11, 2807-2850.

[19] K. S. Jensen, R. E. Hansen, J. R. Winther, *Antioxidants & Redox Signaling* **2009**, 11, 1047-1058.

[20] R. Singh and G. M. Whitesides. "Thiol-Disulfide Interchange." in *The Chemistry of Sulfur-Containing Functional Groups*, (Eds.: S. Patai), J. Wiley and Sons, London, **1993**, pp. 633-658.

[21] R. Bhattacharyya, D. Pal, P. Chakrabarti, *Protein Engineering Design & Selection* **2004**, 17, 795-808.

[22] T. E. Creighton, D. P. Goldenberg, *Journal of Molecular Biology* **1984**, 179, 497-526.

[23] H. Nakamoto, J. C. A. Bardwell, *Biochimica Et Biophysica Acta-Molecular Cell Research* **2004**, 1694, 111-119.

[24] O. Carugo, M. Cemazar, S. Zahariev, I. Hudaky, Z. Gaspari, A. Perczel, S. Pongor, *Protein Engineering* **2003**, 16, 637-639.

[25] I. Hudaky, Z. Gaspari, O. Carugo, M. Cemazar, S. Pongor, A. Perezel, *Proteins-Structure Function and Bioinformatics* **2004**, 55, 152-168.

[26] K. Ito, K. Inaba, *Current Opinion in Structural Biology* **2008**, 18, 450-458.

[27] R. Daniels, P. Mellroth, A. Bernsel, F. Neiers, S. Normark, G. von Heijne, B. Henriques-Normark, *Journal of Biological Chemistry* **2010**, 285, 3300-3309.

[28] K. Inaba, *Journal of Biochemistry* **2009**, 146, 591-597.

[29] N. A. Patil, J. Tailhades, R. A. Hughes, F. Separovic, J. D. Wade, M. A. Hossain, *International Journal of Molecular Sciences* **2015**, 16, 1791-1805.

[30] R. St Charles, K. Padmanabhan, R. V. Arni, K. P. Padmanabhan, A. Tulinsky, *Protein Science* **2000**, 9, 265-272.

[31] Q. Liu, Q. Q. Huang, M. K. Teng, C. M. Weeks, C. Jelsch, R. G. Zhang, L. W. Niu, *Journal of Biological Chemistry* **2003**, 278, 41400-41408.

[32] J. Singh, J. M. Thornton, *Febs Letters* **1985**, 191, 1-6.

[33] T. R. Ioerger, C. G. Du, D. S. Linthicum, *Molecular Immunology* **1999**, 36, 373-386.

[34] H. J. Leung, G. Xu, M. Narayan, H. A. Scheraga, *Journal of Peptide Research* **2005**, 65, 47-54.

[35] J. L. Lau, M. K. Dunn, *Bioorganic & Medicinal Chemistry* **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, *Journal of Organic Chemistry* **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, *Analytical Biochemistry* **1994**, 220, 5-10.

[40] F. Duchardt, I. R. Ruttekolk, 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, *Journal of Biological Chemistry* **2009**, 284, 36099-36108.

[41] M. E. Levison, A. S. Josephson, D. M. Kirschenbaum, *Experiantia* **1969**, 25, 126-127

[42] W. R. Gray, *Protein Science* **1993**, 2, 1732-1748.

[43] J. Eng, W. A. Kleinman, L. Singh, G. Singh, J. P. Raufman, *Journal of Biological Chemistry* **1992**, *267*, 7402-7405.

[44] J. L. Iltz, D. E. Baker, S. M. Setter, R. K. Campbell, *Clinical 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 in Endocrinology and Metabolism* **2016**, *27*, 427-438.

[47] J. J. Meier, *Nature Reviews Endocrinology* **2012**, *8*, 728-742.

[48] D. Donnelly, *British Journal of Pharmacology* **2012**, *166*, 27-41.

[49] A. Hareter, E. Hoffmann, H. P. Bode, B. Goke, R. Goke, *Endocrine Journal* **1997**, 44, 701-705.

[50] J. W. Neidigh, R. M. Fesinmeyer, N. H. Andersen, *Nature Structural Biology* **2002**, 9, 425-430.

[51] B. Barua, J. C. Lin, V. D. Williams, P. Kummler, J. W. Neidigh, N. H. Andersen, *Protein Engineering Design & Selection* **2008**, 21, 171-185.

[52] S. Runge, H. Thogersen, K. Madsen, J. Lau, R. Rudolph, *Journal of Biological Chemistry* **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, G. Skiniotis, *Nature* **2017**, *546*, 248-+.

[54] P. Rovo, V. Farkas, O. Hegyi, O. Szolomajer-Csikos, G. K. Toth, A. Perczel, *Journal of Peptide Science* **2011**, 17, 610-619.

[55] P. Hudaky, P. Straner, V. Farkas, G. Varadi, G. Toth, A. Perczel, *Biochemistry* **2008**, 47, 1007-1016.

[56] P. Rovo, 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 Science* **2005**, 14, 474-482.

[58] S. Runge, H. Thogersen, K. Madsen, J. Lau, R. Rudolph, *Journal of Biological Chemistry* **2008**, 283, 11340-11347.

[59] V. Farkas, I. Jakli, G. K. Toth, A. Perczel, *Chemistry a European Journal* **2016**, 22, 13871-13883.

[60] P. Rovo, P. Straner, A. Lang, I. Bartha, K. Huszar, L. Nyitray, A. Perczel, *Chemistry a European Journal* **2013**, 19, 2628-2640.

[61] A. Perczel, K. Park, G. Fasman, *Analytical Biochemistry*, **1992**, 203, 83 – 93

[62] A. Perczel, M. Hollsi, G. Tusndy, G. Fasman, *Protein Engineering*, **1991**, 4, 669 – 679

[63] R. W. Woody, *Biopolymers* **1978**, 17, 1451.

[64] R. W. Woody, Circular dichroism of peptides, in *The Peptides*, vol. 7 (Eds.: V. J. Hruby), Academic, New York, **1985**, pp. 15–114.

[65] J. D. Fasman, in *Circular Dichroism and the Conformational Analysis of Biomolecules*, Springer, Boston, **1996**.

[66] A. Bundi, K. Wuthrich, *Biopolymers* **1979**, 18, 285-297.

[67] A. J. Doig, R. L. Baldwin, *Protein Science* **1995**, 4, 1325-1336.

[68] A. Mirzahosseini, B. Noszal, *Scientific Reports* **2016**, *6*.

[69] M. Tollinger, N. R. Skrynnikov, F. A. A. Mulder, J. D. Forman-Kay, L. E. Kay, *Journal of the American Chemical Society* **2001**, 123, 11341-11352.

[70] 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.

[71] Z. Wang, T. Rejtar, Z. S. Zhou, B. L. Karger, *Rapid Communications in Mass Spectrometry* **2010**, 24, 267-275.

[72] 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*, Journal of the American Society for Mass Spectrometry* **2010**, 21, 837-844.

[73] G. G. Smith, B. Silvadesol, *Science* **1980**, 207, 765-767.

[74] D. E. Schwass, J. W. Finley, *Journal of Agricultural and Food Chemistry* **1984**, 32, 1377-1382.

[75] P. A. Fernandes, M. J. Ramos, *Chemistry-a European Journal* **2004**, 10, 257-266.

[76] J. Houk, G. M. Whitesides, *Journal of the American Chemical Society* **1987**, 109, 6825-6836.

[77] L. E. Overman, E. M. Oconnor, *Journal of the American Chemical Society* **1976**, 98, 771-775.

[78] O. Dmitrenko, C. Thorpe, R. D. Bach, *Journal of Organic Chemistry* **2007**, 72, 8298-8307.

[79] G. H. Snyder, M. J. Cennerazzo, A. J. Karalis, D. Field, *Biochemistry* **1981**, 20, 6509-6519.

[80] C. L. Wu, C. Belenda, J. C. Leroux, M. A. Gauthier, *Chemistry a European Journal* **2011**, 17, 10064-10070.

[81] O. Gawron, S. Mahboob, J. Fernando, *Journal of the American Chemistry Society*, **1964**, 86, 2283–2286.

[82] C. G. Moore, B. R. Trego, *Tetrahedron* **1962**, 18, 205-218.

[83] T. Kortemme, T. E. Creighton, *Journal of Molecular Biology* **1995**, 253, 799-812.

[84] L. W. Guddat, J. C. A. Bardwell, J. L. Martin, *Structure with Folding & Design* **1998**, 6, 757-767.

[85] H. N. Po, N. M. Senozan, *Journal of Chemical Education* **2001**, 78, 1499-1503.

[86] R. D. Bach, O. Dmitrenko, C. Thorpe, *Journal of Organic Chemistry* **2008**, 73, 12-21.

[87] M. Scigelova, P. S. Green, A. E. Giannakopulos, A. Rodger, D. H. G. Crout, P. J. Derrick, *European Journal of Mass Spectrometry* **2001**, 7, 29-34.

[88] J. Tailhades, N. A. Patil, M. A. Hossain, J. D. Wade, *Journal of Peptide Science* **2015**, 21, 139-147.

[89] 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-Structure Function and Bioinformatics* **2005**, 59, 687-696.

[90] A. T. Brunger, *Nature Protocols* **2007**, 2, 2728-2733.

[91] P. Straner, N. Taricska, M. Szabo, G. K. Toth, A. Perczel, *Current Protein & Peptide Science* **2016**, 17, 147-155.

Graphical Abstract
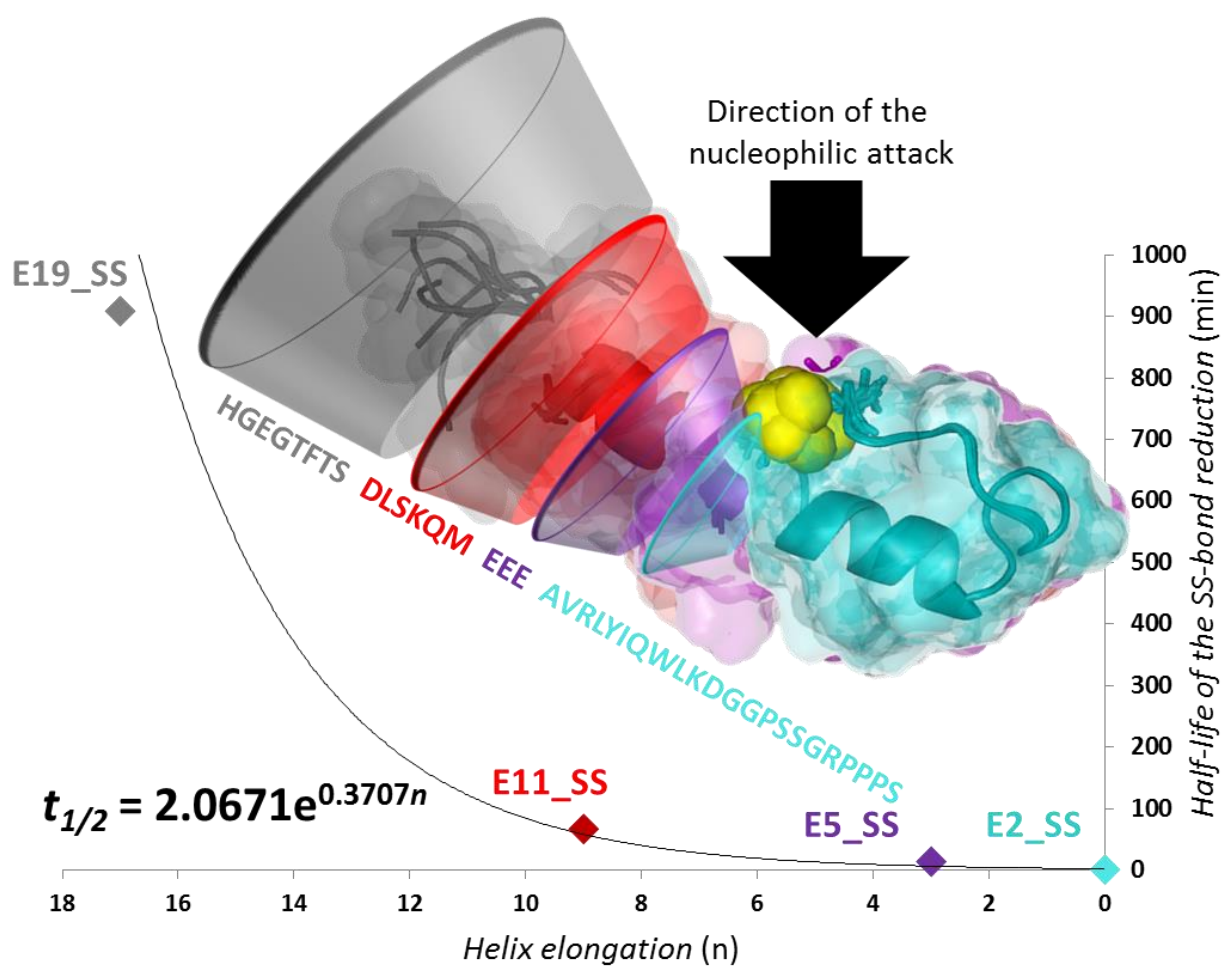
Structural, steric and electrostatic factors influence the SS-bond (yellow-spheres) reduction rate resulting in 3 orders of magnitude differences in reduction half-times ($t_{1/2}$) in case of N-proximal truncated SS-bond cyclized Exenatide derivates.