Modulation of Structure and Dynamics by Disulfide Bond Formation in Unfolded States

Robert Silvers,[†] Friederike Sziegat,[†] Hideki Tachibana,[‡] Shin-ichi Segawa,[§] Sara Whittaker,[⊥] Ulrich L. Günther,[⊥] Frank Gabel,[¶] Jie-rong Huang,[¶] Martin Blackledge,[¶] Julia Wirmer-Bartoschek,[†] and Harald Schwalbe*,[†]

[†]Institute for Organic Chemistry and Chemical Biology, Center of Biomolecular Magnetic Resonance, Goethe University Frankfurt, Max-von-Laue-Strasse 7, 60438 Frankfurt am Main, Germany

[‡]School of Biology-Oriented Science and Technology, Kinki University, Kinokawa, 649-6493, Japan

[§]School of Science and Technology, Kwansei Gakuin University, Sanda, 669-1337, Japan

[⊥]The Henry Wellcome Building for Biomolecular NMR Spectroscopy, School of Cancer Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, U. K.

[¶]Institut de Biologie Structurale Jean-Pierre Ebel, CEA-CNRS-UJF UMR 5075, 41 Rue Jules Horowitz, Grenoble 38027, France

Ⓢ *Supporting Information*

ABSTRACT: During oxidative folding, the formation of disulfide bonds has profound effects on guiding the protein folding pathway. Until now, comparatively little is known about the changes in the conformational dynamics in folding intermediates of proteins that contain only a subset of their native disulfide bonds. In this comprehensive study, we probe the conformational landscape of non-native states of lysozyme containing a single native disulfide bond utilizing nuclear magnetic resonance (NMR) spectroscopy, small-angle X-ray scattering (SAXS), circular dichroism (CD) data, and modeling approaches. The impact on conformational dynamics varies widely depending on the loop size of the single disulfide variants and deviates significantly from random coil predictions for both NMR and SAXS data. From these experiments, we conclude that the introduction of single disulfides spanning a large portion of the polypeptide chain shifts the structure and dynamics of hydrophobic core residues of the protein so that these regions exhibit levels of order comparable to the native state on the nanosecond time scale.
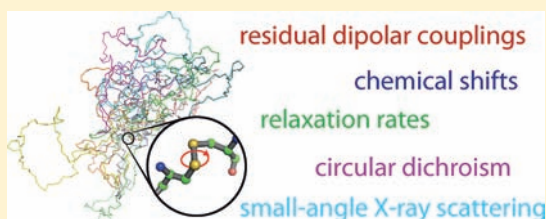
INTRODUCTION

Intramolecular disulfide bonds play crucial roles in protein folding and stability. Approximately 15% of the total human proteome and ∼65% of extracellular, secreted proteins and the extracellular portions of membrane proteins contain disulfide bonds.[1] In oxidative folding, the collapse of hydrophobic residues and formation of disulfide bonds are linked, and both restrain the conformational space accessible to unfolded states of proteins, lower their entropy, and likely enhance cooperativity in protein folding. During oxidative folding, not only native disulfide bonds can form, but also non-native species can appear as major intermediates, as shown in hallmark studies for the protein bovine pancreatic trypsin inhibitor (BPTI).[2,3] Proteins with more than a single disulfide bond can populate several stable folding intermediates. Most of the early folding intermediates with only a subset of the native disulfide bonds show unfolded state features characterized by a highly dynamic ensemble of rapidly interconverting conformations rather than a persistent global fold.[4−6] The investigation of these flexible protein states at a residue-specific level can be achieved using nuclear magnetic resonance (NMR) spectroscopy. In particular, local and global dynamics and residual structure of the unstructured state can be investigated using chemical shifts, *J* couplings, and relaxation data. Residual dipolar couplings (RDCs) complete these data, as they report on both dynamics and residual structure. The compactness of the unstructured state can be characterized using diffusion-ordered NMR spectroscopy (DOSY) but also by complementary techniques, notably small-angle X-ray scattering (SAXS). In order to predict the structure and dynamics of these ensemble states, theoretical models help to corroborate experimental data. Random flight models by Flory[7,8] treat the polypeptide chain as a homopolymer and derive global properties including the change in compactness from geometric and statistical modeling. Random coil models[9,10] describe the polypeptide chain as a heteropolymer and can be used to generate ensembles allowing the simulation of experimental data on a per-residue basis.

Hen egg white lysozyme (HEWL) has served as a model protein for oxidative folding in several biochemical and biophysical studies.[11−19] It consists of 129 residues arranged in two domains: the α-domain, which comprises residues 1−35 and 85−129, and the β-domain spanning residues 36 and 84.
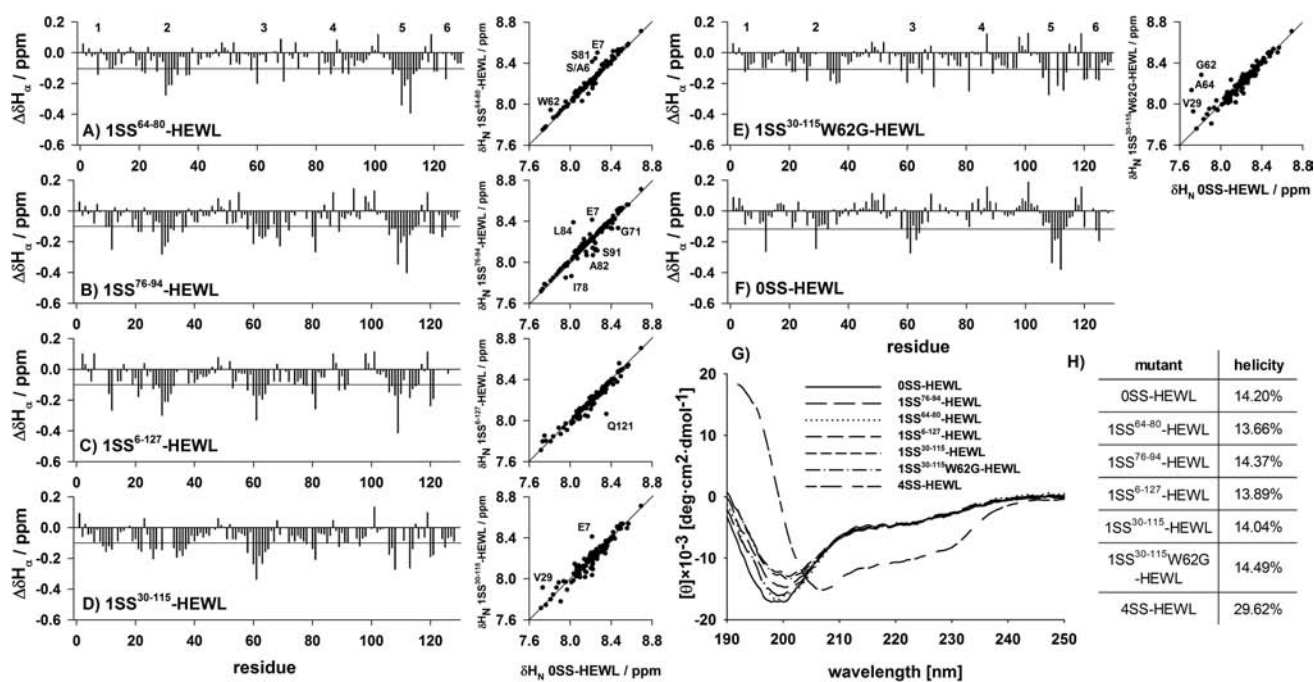
Figure 1. Residual secondary structure of lysozyme mutants. The left panel of (A−F) shows H_α chemical shift deviations from random-coil values ($\Delta\delta = \delta exp − \delta rc$) of $1SS^{64−80}$-HEWL, $1SS^{76−94}$-HEWL, $1SS^{6−127}$-HEWL, $1SS^{30−115}$-HEWL, $1SS^{30−115}$W62G-HEWL, and 0SS-HEWL, respectively. Samples were prepared with a concentration of 300 μM in water at pH 2.0. The horizontal black line at −0.1 ppm indicates thresholds for chemical shift indexing for α helices. The right panel of (A−E) shows δH_N chemical shift correlations between $1SS^{64−80}$-HEWL, $1SS^{76−94}$-HEWL, $1SS^{6−127}$-HEWL, $1SS^{30−115}$-HEWL, and $1SS^{30−115}$W62G-HEWL, respectively, and 0SS-HEWL. The correlation can be described by the function $y = x$, as indicated by a solid line. (G) shows circular dichroism spectra of $1SS^{64−80}$-HEWL, $1SS^{76−94}$-HEWL, $1SS^{6−127}$-HEWL, $1SS^{30−115}$-HEWL, $1SS^{30−115}$W62G-HEWL, 0SS-HEWL, and folded native 4SS-HEWL. Samples were prepared with a concentration of 20 μM in water at pH 2.0. (H) shows the content of α helices estimated using the approach of Rohl and Baldwin.[20]

The native protein has eight cysteine residues forming a total of four disulfide bonds, two in the α-domain (C6−C127 and C30−C115), one in the β-domain (C64−C80), and one interdomain disulfide bond (C76−C94). Jarrett et al. showed with the help of immunochemical pulsed-labeling methods that the α-domain forms first during oxidative refolding.[11] Intermediates containing only two of the four disulfide bonds (C30−C115 and C6−C127) located in the α-domain ($2SS^\alpha$-HEWL) can be detected by monoclonal antibodies that bind to the natively folded α-domain. A native β-domain can only be detected in three disulfide containing variants (C30−C115, C6−C127, C64−C80). The fourth disulfide bond (C76−C94), which is located between both domains, seems to contribute only minor changes to the conformation and stabilizes the final interdomain structure. These findings are corroborated by our previous NMR studies on $2SS^\alpha$-HEWL and $2SS^\beta$-HEWL.[4] Whereas the α-domain of $2SS^\alpha$-HEWL shows a native fold even at pH 2.0 in the absence of urea, $2SS^\beta$-HEWL is unfolded under these conditions. Furthermore, it has been reported that the first disulfide bond to form during lysozyme folding is C30−C115 since it is present in all isolated 3SS-HEWL variants during oxidative refolding.[21] The use of water at pH 2.0, or even 8 M urea at pH 2.0, is justified because at this pH the exchange of backbone protons is limited to a level that is optimal for NMR spectroscopy while limiting unspecific intermolecular interactions and aggregation via electrostatic repulsion. This strategy is well established, as shown in various studies of non-native states of proteins.[22−24]

Besides disulfide bond formation, the hydrophobic interactions drive productive protein folding. Recent studies on lysozyme mutants without disulfide bonds (0SS-HEWL) reflect the significance of hydrophobic patches and their interaction within the ensemble of non-native states of lysozyme.[5,6] Within these states, six hydrophobic patches have been identified, and four out of six contain tryptophan residues W28 (cluster 2), W62/W63 (cluster 3), W108/W111 (cluster 5), and W123 (cluster 6). The impact of these long-range interactions detected by NMR on native disulfide bond formation has been investigated by Mishima et al.[13] The disulfide exchange equilibrium of all four single disulfide mutants of lysozyme with only a single disulfide bond is affected by W111G or W123G mutations. Similar studies were used to investigate the effect of a W62G mutation on the disulfide exchange equilibrium of these four single disulfide mutants.[15] These results strongly suggest that long-range interactions between several hydrophobic clusters within lysozyme modulate disulfide bond formation, and hence control early protein folding events.

In this study, we present comprehensive data on permanently trapped early single disulfide folding intermediates populated during the oxidative refolding of HEWL. The following single disulfide mutants, $1SS^{6−127}$-HEWL, $1SS^{30−115}$-HEWL, $1SS^{64−80}$-HEWL, $1SS^{76−94}$-HEWL and $1SS^{30−115}$W62G-HEWL, were studied by high resolution NMR spectroscopy, SAXS, and circular dichroism (CD) studies. We map out the effect on dynamics and residual structure on a local and global level. Disulfide bond formation therefore stabilizes hydrophobic clusters during folding, suggesting a coevolution of hydrophobic clusters and disulfide bonds to guide folding and avoid misfolding.

Table 1. Compactness of Single Disulfide Mutants Compared to 0SS-HEWL

	% decrease in R_g from Zimm/Stockmeyer	R_g from flexible meccano	R_h from flexible meccano	R_g from SAXS	R_h from DOSY
0SS-HEWL	(0.0)	33.0 Å (0.0%)	28.8 Å (0.0%)	35.2 Å (0.0%)	32.5 ± 0.4 Å (0.0%)
1SS$^{64-80}$-HEWL	(−8.9)	30.5 Å (−7.5%)	27.8 Å (−3.5%)	35.0 Å (−0.6%)	32.5 ± 0.4 Å (0.0%)
1SS$^{76-94}$-HEWL	(−8.9)	30.5 Å (−7.6%)	27.8 Å (−3.5%)	32.2 Å (−8.5%)	30.6 ± 1.0 Å (−5.8%)
1SS$^{6-127}$-HEWL	(−31.0)	22.9 Å (−30.7%)	24.9 Å (−13.5%)	31.2 Å (−11.1%)	27.3 ± 0.8 Å (−16.0%)
1SS$^{30-115}$-HEWL	(−32.3)	23.6 Å (−28.6%)	25.1 Å (−12.8%)	30.9 Å (−12.2%)	24.4 ± 0.5 Å (−24.9%)
1SS$^{30-115}$W62G-HEWL	(−32.2)	23.6 Å (−28.6%)	25.1 Å (−12.8%)	27.6 Å (−21.6%)	30.7 ± 1.3 Å (−5.5%)

■ RESULTS

Residual Secondary Structure in Single Disulfide Mutants. Single disulfide mutants were investigated by circular dichroism (CD) spectroscopy (Figure 1G) to derive the degree of helicity (Figure 1H).[20] All five single disulfide mutants show great similarity to 0SS-HEWL in the shape of the CD curve and in its helical content (∼14%) and little resemblance to the native state (∼29%). Lack of persistent secondary structure is further supported by low chemical shift dispersion detected in heteronuclear H, N correlation NMR spectra (see Figures S1−S7 and Tables S1−S5, Supporting Information (SI)). Figure 1A−F (left panel) compares H_α chemical shift deviations of the lysozyme mutants from random coil chemical shifts derived from model peptides.[25−28] Although all single disulfide mutants as well as 0SS-HEWL are mostly unfolded, there are significant deviations from random coil behavior, principally mapping out areas with α-helical propensity. Most prominent regions are identified around tryptophan clusters W28 and W108/W111. The central hydrophobic cluster around residues W62/W63 shows a strong α-helical propensity for fully unbranched 0SS-HEWL without disulfide bonds as well as for 1SS$^{30-115}$-HEWL, 1SS$^{76-94}$-HEWL, and 1SS$^{6-127}$-HEWL. For 1SS$^{64-80}$-HEWL, however, this is not true, most likely because of the proximity of the C64−C80 disulfide bond. In the case of 1SS$^{30-115}$W62G-HEWL, this α-helical propensity is also perturbed by the tryptophan to glycine mutation at position 62. The close proximity of two tryptophan residues and its structural interplay seem to induce structure in this region, which is then lost by W62G mutation or by introduction of a neighboring disulfide bond. Figure 1 A−E (right panel) depicts H_N chemical shift correlations between each single disulfide mutant and 0SS-HEWL. All single disulfide mutants show a strong correlation to 0SS-HEWL with deviations restricted to regions close to disulfide bonds or point mutations.

Compactness of Single Disulfide Mutants. According to polymer theory, unfolded polypeptide chains can be described as a random coil polymer where each amino acid is one polymer unit linked by peptide bonds. For unbranched and noncyclic polymers meeting the Θ condition, the radius of gyration is defined by $R_g = c(N/6)^{1/2}$,[7,8] where N is the number of units in the polymer and c is a measure of chain stiffness. For branched and cyclic polymers, a tool set for the calculation of the radius of gyration was introduced by Zimm and Stockmeyer in 1949.[29] Single disulfide mutants represent such branched polymers and can be described as a polymer with cyclic and branched elements of different size. The R_g is then given by eq 1 as follows:

$$R_g = \frac{c}{\sqrt{N}} \sqrt{\left(\frac{N_1^2}{2} - \frac{N_1^3}{3N}\right) + \left(\frac{N_2^2}{2} - \frac{N_2^3}{3N}\right) + \left(\frac{N_r^2}{6} - \frac{N_r^3}{12N}\right)}$$

(1)

where N is the number of units in the polymer, N_1, N_2, and N_r are the number of units in the two linear polymers and one cyclic polymer chain, respectively, and c is a measure of chain stiffness. Since c is not known a priori, we utilize the ratio of both (linear and branched) radii to compare the single disulfide mutants with the unbranched polypeptide chain on a relative scale. According to these predictions, the radius of gyration R_g of single disulfide mutants is predicted to be 8.9, 8.9, 32.3, and 31.0% smaller than 0SS-HEWL for 1SS$^{64-80}$-HEWL, 1SS$^{76-94}$-HEWL, 1SS$^{30-115}$-HEWL, and 1SS$^{6-127}$-HEWL, respectively (Table 1). Calculations predict 1SS$^{64-80}$-HEWL and 1SS$^{76-94}$-HEWL to show only small deviation from 0SS-HEWL, whereas 1SS$^{30-115}$-HEWL and 1SS$^{6-127}$-HEWL are predicted to be significantly more compact than 0SS-HEWL.

Next, we performed ensemble calculations to validate and expand these findings derived from polymer theory. The flexible meccano algorithm[30] was used to generate 10 000 structures of each of the five single disulfide mutants as well as 0SS-HEWL, which were subsequently analyzed using HYDRO-PRO,[31,32] yielding the radius of gyration (R_g) and radius of hydration (R_h) for each structure. The size distributions of R_g and R_h values for each single disulfide mutant and 0SS-HEWL is shown in Figure S8A−F (SI). These simulations predict the R_g values of single disulfide mutants to be 7.5, 7.6, 28.6, and 30.7% smaller than 0SS-HEWL for 1SS$^{64-80}$-HEWL, 1SS$^{76-94}$-HEWL, 1SS$^{30-115}$-HEWL, and 1SS$^{6-127}$-HEWL, respectively. Hence, the results from polymer theory and from the flexible meccano approach are remarkably similar, although the ensemble simulations yield R_g values slightly less compact than predicted by polymer theory. Although by utilizing the ratio of $R_g^{branched}/R_g^{linear}$ the majority of excluded volume effects should be eliminated, small differences between the two prediction methods remain, most likely because of such effects imposed by the disulfide bonds. In order to experimentally determine the global parameters R_g and R_h, we utilized two complementary techniques: small-angle X-ray scattering (SAXS) and diffusion-ordered NMR spectroscopy (DOSY), respectively. SAXS data were analyzed by ensemble optimization method (EOM) as seen in Figure 2 (left and middle panel).[33] Distributions selected by EOM are larger than predicted by distributions generated by flexible meccano (pool distribution). The more extended character of the
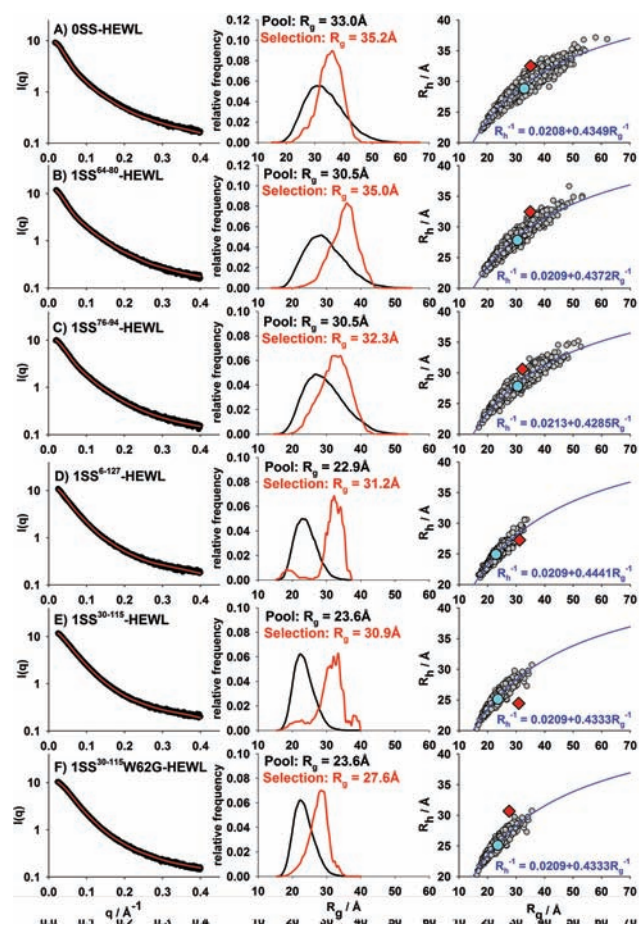
Figure 2. Compactness of lysozyme mutants by SAXS and DOSY. Panels (A−F) (left column) show the SAXS profiles (black dots) and EOM fit (red line) of (A) 0SS-HEWL, (B) 1SS$^{64-80}$-HEWL, (C) 1SS$^{76-94}$-HEWL, (D) 1SS$^{6-127}$-HEWL, (E) 1SS$^{30-115}$-HEWL, and (F) 1SS$^{30-115}$W62G-HEWL. Panels (A−F) (middle column) show the EOM results with R_g distributions for the pool ensemble (black) and selected ensemble (red). Panels (A−F) (right column) show the experimental R_g and R_h values derived from SAXS and DOSY measurements (red diamonds). The R_g/R_h distributions derived from ensemble calculations are shown in gray, the fit is shown, and the fit equation is displayed in blue. Mean R_g/R_h derived from distributions are plotted as cyan dots. Samples were prepared with a concentration of typically 60 μM in water at pH 2.0.

ensembles can be due to electrostatic effects not taken into account by the flexible meccano algorithm. For 1SS$^{30-115}$-HEWL and 1SS$^{6-127}$-HEWL, this effect is especially apparent, pointing to an even more extended ensemble, which is lost by W62G mutation. Figure 2 (right panel) shows the correlation between R_g and R_h for the ensemble calculation (gray dots) as well as the R_g/R_h correlation for all single disulfide mutants and 0SS-HEWL.

To a good approximation, the correlation of R_g and R_h can be described by $R_h^{-1} = 0.208 + 0.4349 R_g^{-1}$. Similar correlations can also be obtained from ensembles of single disulfide mutants revealing the same dependency. The overall shape of the Kratky plots in SAXS data and the absence of a Porod−Debye plateau for all five single disulfide mutants and 0SS-HEWL underlines the unfolded nature of these mutants (see Figure S9 (SI)). Interestingly, the Kratky plot profiles of the single disulfide mutants show a tendency toward 4SS-HEWL, defined by the number of stationary points. The first derivative of the Kratky

plot of 0SS-HEWL reveals no stationary point, whereas the Kratky plot of 4SS-HEWL show two. The single disulfide mutants 1SS$^{64-80}$-HEWL and 1SS$^{76-94}$-HEWL have only one broad stationary point (plateau), 1SS$^{6-127}$-HEWL, 1SS$^{30-115}$-HEWL, and 1SS$^{30-115}$W62G-HEWL have two. The experimental data obtained from SAXS and DOSY measurements depicted in Figure 2 (right column) are in good agreement with the R_g/R_h correlation of the flexible meccano simulations.

Disulfide Bonds Modulate Protein Dynamics. Residual dipolar couplings (RDCs) were measured in gel-aligned samples to obtain information on the orientation of all internuclear vectors in the conformational ensemble. RDCs are sensitive to both structural and dynamical aspects of the polypeptide chain on a local and global level. Figure 3A−F shows RDCs for all five single disulfide mutants and 0SS-HEWL. In 0SS-HEWL, several regions stand out; they are located in the hydrophobic clusters: A6, W28, W62/W63, W108/W111, and W123. This pattern can be found in the RDC data of all single disulfide mutants, however with several exceptions. In 1SS$^{64-80}$-HEWL, the central hydrophobic patch around W62/W63 shows a higher degree of random coil behavior, in good agreement with the loss of secondary structure detected by analysis of chemical shift deviations from random coil predictions. This is also true for the hydrophobic patch around W62/W63 in the 1SS$^{30-115}$W62G-HEWL mutant. ^{15}N relaxation rate distributions provide information about protein dynamics.[4−6,34] The position and shape of these distributions indicate the mean overall flexibility of the molecule, and the deviation is a measure for persistence of flexibility over the protein chain. Previously, we could also show that depressions in the R_2 profile upon mutation reflect long-range interactions in unfolded protein states. Therefore, relaxation rate distribution analysis allows the direct comparison of different single disulfide bond mutants (see Figure S10 (SI)). The relaxation rate distribution of 1SS$^{64-80}$-HEWL and 1SS$^{76-94}$-HEWL display a high similarity to 0SS-HEWL with only few residue values observed in the region of native 4SS-HEWL, whereas the relaxation rate distribution for 1SS$^{6-127}$-HEWL, 1SS$^{30-115}$-HEWL, and 1SS$^{30-115}$W62G-HEWL differ significantly from either 0SS-HEWL or 4SS-HEWL. Their distribution is much broader, spreading from 0SS-HEWL to 4SS-HEWL and even beyond, indicating that these mutants show less overall flexibility in contrast to 1SS$^{64-80}$-HEWL and 1SS$^{76-94}$-HEWL. To allow more detailed insight into protein dynamics of these four single disulfide bond mutants, relaxation rates are plotted as a function of sequence (Figure 4, left panel). As expected from distribution analysis, the mutants 1SS$^{64-80}$-HEWL and 1SS$^{76-94}$-HEWL (Figure 4A,B, left panel) show only little deviation from relaxation rates of 0SS-HEWL, whereas the mutants 1SS$^{6-127}$-HEWL, 1SS$^{30-115}$-HEWL and 1SS$^{30-115}$W62G-HEWL (Figure 4C−E, left panel) deviate significantly from relaxation rates of 0SS-HEWL.

Significant deviations of R_2 relaxation rates are most pronounced around disulfide bonds because of three different effects. The flexibility at these regions is restricted since each cysteine residue has three bonds to other residues, as compared to residues that have only two (or one) bonds to other residues.

Second, the introduction of a covalent constraint between two parts of the two polypeptide chains contributes to the bulkiness around disulfide bonds. Additionally, for some residues, R_2 relaxation rates are elevated because of chemical exchange effects caused by disulfide bond rotation. Interestingly, both small-sized single disulfide bond mutants ([64−80]
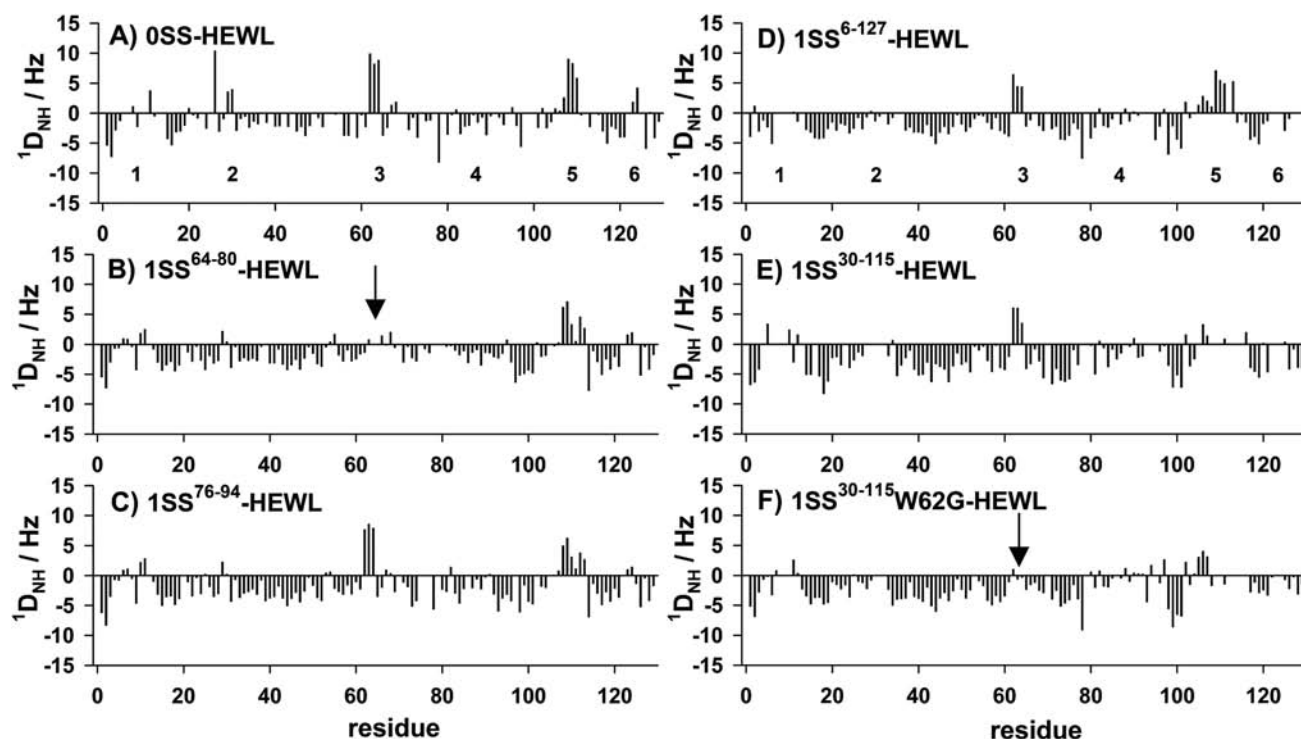
Figure 3. Residual dipolar couplings (RDCs) of lysozyme mutants. (A–F) show RDCs of 0SS-HEWL, 1SS$^{64-80}$-HEWL, 1SS$^{76-94}$-HEWL, 1SS$^{6-127}$-HEWL, 1SS$^{30-115}$-HEWL, and 1SS$^{30-115}$W62G-HEWL, respectively. RDCs were normalized to 10 Hz HDO splitting. Significant deviations from 0SS-HEWL are indicated by an arrow. Samples were prepared with a concentration of typically 300 μM in water at pH 2.0.

= 16 residues, [76−94] = 18 residues) with disulfide bonds located in the β domain show only local deviations from the relaxation rates of 0SS-HEWL, and the rates of the α domain remain mostly unchanged. Transverse relaxation rates above 5 Hz are mostly restricted to the β-domain containing the disulfide bond, rendering the polypeptide chain in general as flexible as 0SS-HEWL. On the other hand, high relaxation rates for the large-sized single disulfide bond mutants ([6−127] = 121 residues, [30−115] = 85 residues) with their disulfide bond located within the α domain are not only restricted to the α domain, but the β domain also shows significant deviations from 0SS-HEWL backbone flexibility. Introduction of the long looped disulfides therefore rigidifies the entire polypeptide chain. Apart from the regions around the disulfide bonds, high relaxation rates are observed for G26, Q57, W63, L84, T89, V99, and W108 for 1SS$^{30-115}$-HEWL and for H15, G22, G26, A30, Q57, W63, L84, T89, V99, and A110 for 1SS$^{6-127}$-HEWL. Interestingly, the central hydrophobic cluster around residues W62/W63 splits up into two distinct clusters that are much more apparent in the large SS-loop sized mutants. Residue I58 also shows a higher rigidity comparable to W62/W63, but with residues much more flexible in between. Furthermore, in region 70−105 not only bulky hydrophobic residues, but also aliphatic residues such as leucine and isoleucine form hydrophobic clusters. Within these clusters, residues I58, W62/W63, I78/L83/L84, I88, and I98/V99 seem to play an important role. Interestingly, the amino acid isoleucine appears to play an important part in hydrophobic networking in the 1SS$^{6-127}$-HEWL and 1SS$^{30-115}$-HEWL variants. These aliphatic residues, which are not significantly elevated in relaxation analyses of unfolded yet unbranched lysozyme mutants, seem to participate in hydrophobic networking within these mutants. Reduced spectral density analysis of all five single disulfide mutants in

respect to 0SS-HEWL and 4SS-HEWL reveal an interesting view on protein dynamics (Figure 4, right panel, and Figure S11 (SI)). While none of the residues of 1SS$^{76-94}$-HEWL reach a dynamic state similar to 4SS-HEWL, relaxation data in the mutants 1SS$^{6-127}$-HEWL, 1SS$^{30-115}$-HEWL, and 1SS$^{30-115}$W62G-HEWL tend toward the folded state and for some residues seem as rigid as the native state on the sub-nanosecond time scale. Some residues even exceed the dynamical properties of 4SS-HEWL indicating the presence of chemical exchange. Such chemical exchange broadening is likely induced by rotation around the disulfide bond that can be described as a two-site exchange process. It was found by ab initio self-consistent molecular orbital calculations that the S−S bond of cystins (χ_3) exist in two main conformations, i.e., χ_3 = ±90°.[35] This was further corroborated by extensive chemical exchange studies on disulfide bond dynamics in BPTI.[36−38] The correlation of R_2 and $R_{1\rho}$ rates sheds a light on residues with chemical exchange contribution to the relaxation rates (see Figure S12 (SI)). Residues with chemical exchange contributions are almost exclusively restricted to residues close to the disulfide bond. With the exception of 1SS$^{76-94}$-HEWL, all mutants show some level of chemical exchange effects in the vicinity of the disulfide bond. For 1SS$^{64-80}$-HEWL residues R61 and G67 show little chemical exchange, for 1SS$^{6-127}$-HEWL residue C6 shows significant chemical exchange effects, and for 1SS$^{30-115}$W62G-HEWL residues V29 and R114 show chemical exchange. Residues in the vicinity of disulfide bonds in 1SS$^{30-115}$-HEWL were invisible because of line broadening. Additionally, in 1SS$^{6-127}$-HEWL and 1SS$^{30-115}$-HEWL, residues W62 and W63 as well as H15, G22, G26, and D119 show chemical exchange, respectively, indicating a significant impact of these disulfide bonds on the dynamics of the polypeptide far from the disulfide bond. These results are also further fully
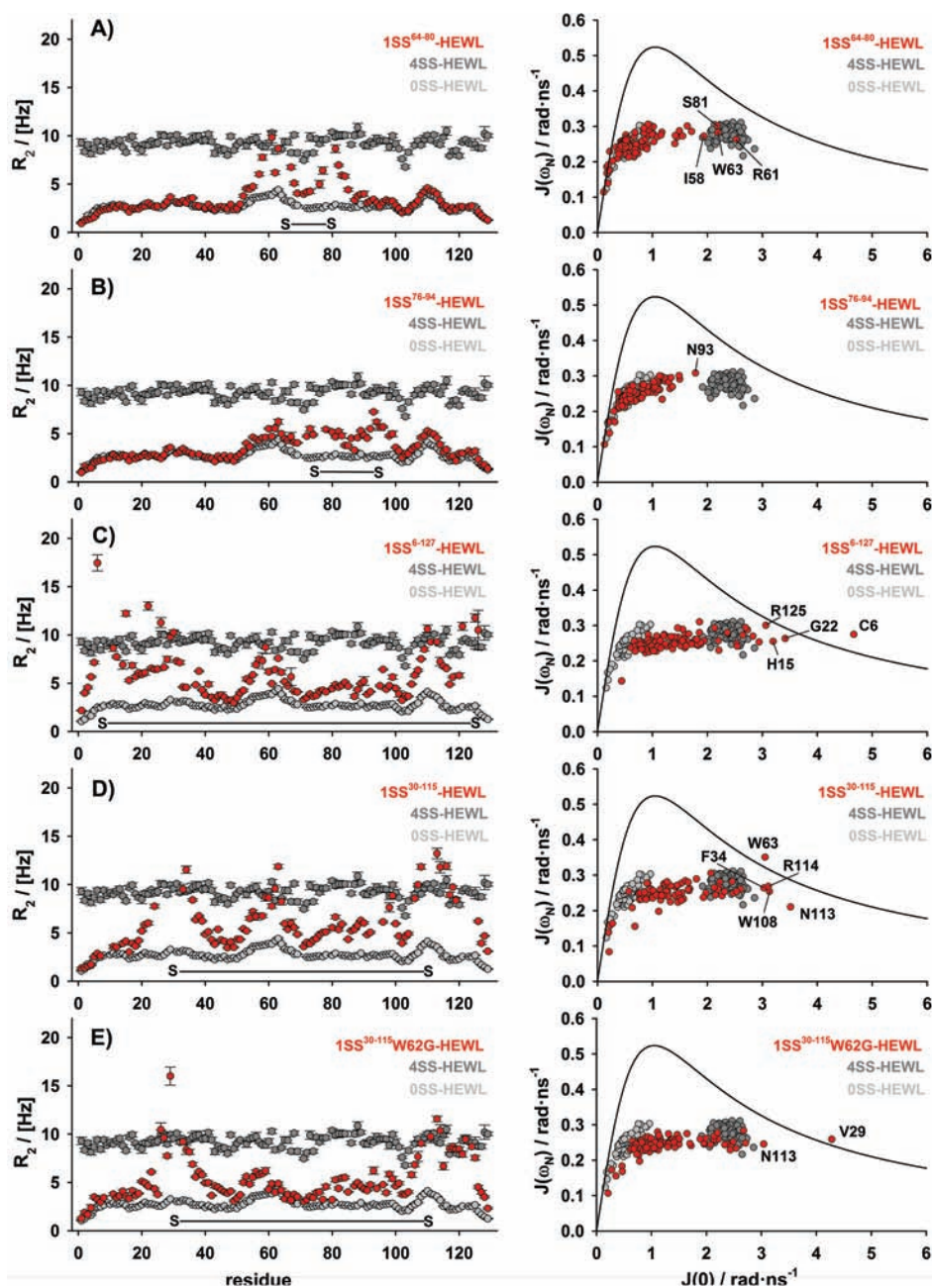
Figure 4. Backbone dynamics of lysozyme mutants. The left panel of (A−E) shows in red the R_2 relaxation rates of 1SS$^{64-80}$-HEWL, 1SS$^{76-94}$-HEWL, 1SS$^{6-127}$-HEWL, 1SS$^{30-115}$-HEWL, and 1SS$^{30-115}$W62G-HEWL, respectively. Additionally, each plot shows the R_2 relaxation rates of 0SS-HEWL (light gray) and folded native 4SS-HEWL (dark gray). The right panel of (A−E) shows in red the reduced spectral densities of 1SS$^{64-80}$-HEWL, 1SS$^{76-94}$-HEWL, 1SS$^{6-127}$-HEWL, 1SS$^{30-115}$-HEWL, and 1SS$^{30-115}$W62G-HEWL, respectively. Significant residues are indicated. Additionally, each plot shows the reduced spectral densities of 0SS-HEWL (light gray) and folded native 4SS-HEWL (dark gray). The solid line indicates the limit case where J is determined by a single Lorentzian and the dynamics are determined by a single motional mode. Samples were prepared with a concentration of typically 300 μM in water at pH 2.0.

supported by relaxation dispersion experiments (Figure S13 (SI)). These findings show that in single disulfide mutants there is still rotation around the disulfide bond, which becomes restricted during folding, presumably as a result of increased compaction and formation of a hydrophobic core.

■ CONCLUSIONS

Single disulfide bonds can have substantial effects on conformational rigidity and dynamics of a polypeptide chain. In fact, the impact of disulfide bonds on backbone dynamics have been shown to vary from minor, locally restricted changes

in dynamics to substantial changes that influence the backbone dynamics of the entire polypeptide chain (Figure 5, Videos S1−S7 (SI)). The present, comprehensive study provides key insight into local and global structure and dynamics of all native single disulfide mutants and allows comparison to the unbranched (0SS-HEWL) and the four disulfide restrained (4SS-HEWL) counterparts. Although the influence of the C64−C80 disulfide bond is only local, it is of significant importance for the region around W62/W63. Chemical shifts, RDCs, and relaxation data suggest that by introduction of the C64−C80 disulfide bond, order in this region is decreased. In
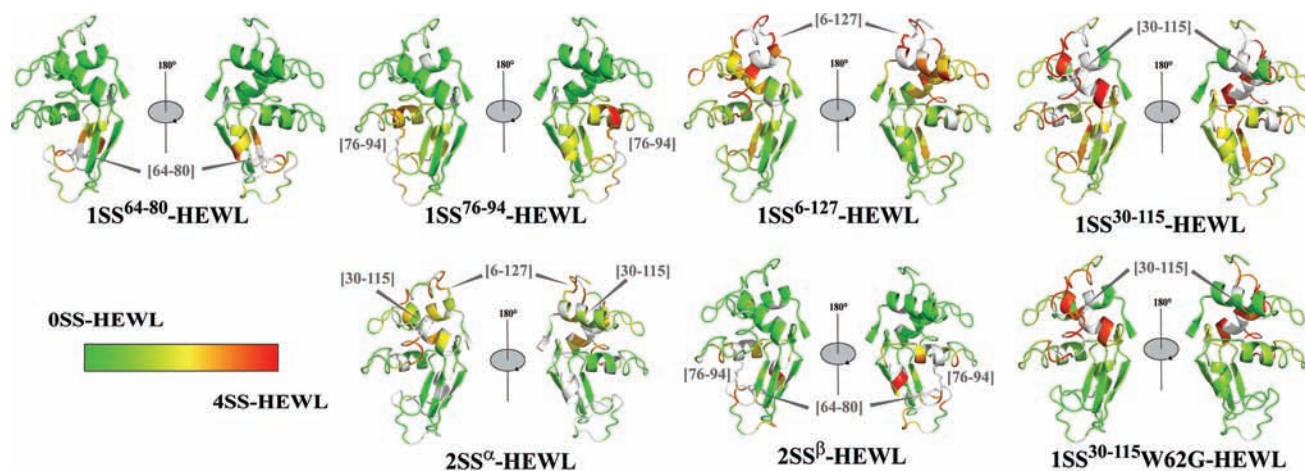
Figure 5. Schematic representation of the degree of folding imposed by disulfide bonds in unfolded 1SS and 2SS HEWL mutants plotted onto the folded structure. The range of $R_{1\rho}$ data is set to 0 to 1 for 0SS-HEWL and 4SS-HEWL, respectively. The $R_{1\rho}$ data of 1SS and 2SS HEWL mutants are calculated according to this range. Values below 0 are set to 0 and values above 1 are set to 1. The gradient from 0−0.5−1 is color coded by green−yellow−red. Structures were generated using the PyMOL software package (PDB ID: 6LYZ). Videos S1−S7 of rotating structures are available in the Supporting Information.

the final native-state structure, these residues are part of a loop. We postulate that the introduction of this disulfide bond disrupts the tendency of this region to form α-helices, thus promoting protein maturation during late oxidative folding.

A general property of a branched polymer is the number of residues within the closed cycle. For single disulfide mutants, this number varies between 16 and 121 residues for 1SS$^{64−80}$-HEWL and 1SS$^{6−127}$-HEWL, respectively. The longer the closed structure, the higher is the influence on the entire polymer chain. This effect can be described by the loop entropy $\Delta S = \alpha k_B \ln N$, where N is the number of residues in the loop, k_B is the Boltzmann's constant, and α is a polymer-specific constant. Hence, the loop entropy indicates that greater loop sizes are accompanied by a higher loss in entropy than smaller loops. This loss in entropy is correlated to the loss in conformational space within the polypeptide and is directly observed by changes in the both local and global experimental parameters.

Formation of the α domain disulfide bonds (C6−C127 and C30−C115) are of higher importance for HEWL folding than the β domain disulfide bonds (C64−C80 and C76−C94).[11] The disulfide bond C30−C115 is thought to be the first disulfide bond formed during folding.[21] The dynamical properties of many residues in 1SS$^{30−115}$-HEWL is comparable to 4SS-HEWL, which suggests that the formation of this particular disulfide bond generates a higher rigidity within the α and β domain, thus facilitating further folding processes. To discuss this on a more detailed level, the differences between 1SS$^{30−115}$-HEWL to 0SS-HEWL have to be taken into account. Within a region of residue 50 to 105, there are five major deviations from the basal flexibility of 0SS, namely, around G26, Q57, W63, L84, T89, and V99. All of the corresponding parts of the polypeptide chain are located in proximity to each other in the native structure at the surface between the α and β domains. All residues form a hydrophobic network within the native structure of HEWL, where one residue is close to at least one other residue (3.5−5 Å). The elevation of R_2 relaxation rates of these specific residues, and the loss of this elevation upon introduction of a W62G mutation, suggests that these essential hydrophobic contacts within the unfolded 1SS$^{30−115}$-HEWL are dynamically preformed in order to facilitate and

accelerate HEWL folding. Therefore, disulfide bonds forming prior to or during protein folding not only are capable of merging certain subdomains, but also seem to govern the overall dynamics of the polypeptide chain facilitating protein folding. This is consistent with the postulated role of these two disulfide bonds during lysozyme folding. Since these disulfide bonds are thought to only stabilize structure, they seem to be responsible for rearrangements during late folding processes.

In summary, the investigation details the interplay between hydrophobic collapse and disulfide bond formation in single disulfide variants of lysozyme in unprecedented detail. The theory and experimental approaches developed here will impact further biophysical studies of structure between the truly unfolded and the native folded state. The understanding of the impact of disulfide bond break and formation in the context of transport across membrane, in redox signaling and oxidative chemistry in cells, will be significantly improved on the basis of the experiments described here.

■ METHODS

Constructs. DNA sequences of 4SS-HEWL, 0SS-HEWL, and 0SSW62G-HEWL were available in pET11a plasmids.[39] For 1SS$^{64−80}$-HEWL, 1SS$^{76−94}$-HEWL, and 1SS$^{6−127}$-HEWL, the plasmid containing the 0SS-HEWL gene was used as a template to reintroduce cysteine residues by two consecutive site-directed mutageneses using the QuikChange site-directed mutagenesis kit (Stratagene, now Agilent Technologies, Santa Clara, CA, USA). For 1SS$^{30−115}$-W62G-HEWL, the plasmid containing 0SSW62G-HEWL was used in a similar fashion. 1SS$^{30−115}$-HEWL was then generated by reintroduction of tryptophan at position 62. For further details, see the Supporting Information.

Expression and Purification. Expression of recombinant lysozyme mutants was performed as described earlier.[39] Lysozyme mutants were overexpressed in *Escherichia coli* BL21(DE3) strain (Invitrogen, Carlsbad, CA, USA) using LB medium or, for isotope labeling, M9 minimal medium. Cultures were induced at OD$_{600}$ = 0.8 with 2 mM IPTG and allowed to grow for 3 h after induction. Inclusion bodies were prepared as described earlier and purified by ion-exchange chromatography. Fractions containing lysozyme were merged and stored at −80 °C. Urea-denatured 4SS-HEWL was introduced into a refolding buffer at a final concentration of 3 μM. After 12 h at 25 °C, the refolding mixture was concentrated by ultrafiltration, and the soluble fraction was dialyzed against pH 7.0

6852

dx.doi.org/10.1021/ja3009506 | *J. Am. Chem. Soc.* 2012, 134, 6846−6854

water and freeze-dried for storage at −20 °C. Urea-denatured 0SS-HEWL was dialyzed against pH 2.0 water and freeze-dried. The samples were further purified with reverse-phase HPLC on a Waters μBondasphere C4 column (19 × 150 mm). Urea-denatured 1SS-HEWL was introduced into an oxidizing buffer at 25 °C for 24 h at the protein concentration of 3 μM for air-oxidation as previously described.[16] The solution was concentrated by ultrafiltration, and the soluble fraction was dialyzed against pH 2.0 water and freeze-dried. The samples were further purified with reverse-phase HPLC on a Waters μBondasphere C4 column (19 × 150 mm). Monodispersity of samples was checked by polyacrylamide gel electrophoresis and MALDI-TOF mass spectrometry. The yield of the purified 0SS-HEWL, 1SS-HEWL, and 4SS-HEWL ranged from 8 to 12, 2 to 4, and 3 to 5 mg per liter of medium, respectively. For further details, see the Supporting Information.

Circular Dichroism Spectroscopy. Circular dichroism (CD) spectra were obtained using a Jasco J-810 spectropolarimeter. All measurements were performed at 20 °C using a quartz cell with 0.2 mm path length. Baseline corrected CD spectra were recorded at 0.5 nm increments between 180 and 250 nm. Protein solutions were prepared in water at pH 2.0 at a typical concentration of 20 μM.

Small-Angle X-ray Scattering (SAXS). The SAXS measurements were recorded on the ID14−3 BioSAXS beamline at the European Synchrotron Radiation Facility (ESRF Grenoble, France). Sample−detector distance was 2.6 m, and the X-ray wavelength used was 0.931 Å (13.32 keV). Fifty microliters of each protein solution (as well as corresponding buffers) were loaded in a flow-through quartz capillary cell at 20 °C. Total exposure time was 100 s per sample. The 2D diffraction patterns were normalized to an absolute scale and azimuthally averaged to obtain the intensity profiles $I(q)$, within BSxCuBE (ESRF beamline data collection software). Solvent contributions were averaged and subtracted from the associated protein sample using the program PRIMUS.[40] The radii of gyration were extracted using the ensemble optimization method (EOM).[33] An ensemble pool of 10 000 structures generated with the flexible meccano approach were used for EOM. For further details, see the Supporting Information.

Ensemble Calculations. The flexible meccano algorithm was utilized to generate 10 000 structures of 0SS-HEWL, 1SS$^{64-80}$-HEWL, 1SS$^{76-94}$-HEWL, 1SS$^{6-127}$-HEWL, 1SS$^{30-115}$-HEWL, and 1SS$^{30-115}$W62G-HEWL each. In the case of single disulfide mutants, the disulfide bond was incorporated as an additional constraint: the C_β/C_β atom distance of the two cysteine residues is to be under 6 Å. These ensembles were subsequently analyzed by HYDROPRO to yield radius of gyration (R_g) and hydrodynamic radius (R_h) distributions. For further details, see the Supporting Information.

Resonance Assignment and Chemical Shift Analysis. All nuclear magnetic resonance (NMR) spectra were recorded on a Bruker DRX600 MHz spectrometer, and all experiments involving heteronuclear experiments were acquired on uniformly ^{15}N-labeled samples at a concentration of typically 300 μM. Data was processed using TOPSPIN 2.1 software (Bruker BioSpin). For resonance assignment typically 3D-^{1}H,^{15}N-NOESY-HSQC and 3D-^{1}H,^{15}N-TOCSY-HSQC spectra were recorded, and resonance assignment was carried out using the CARA software package. The resonance assignment was facilitated by the existence of assignments for 0SS-HEWL and 4SS-HEWL. Proton chemical shifts were referenced to 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS). ^{15}N chemical shifts were calculated relative to DSS using the gyromagnetic ratios of ^{15}N and ^{1}H. Chemical shift deviations from random coil model were calculated according to Wishart et al.[27,28]

Backbone Dynamics and Relaxation. Relaxation analysis was carried out with data sets obtained from heteronuclear ^{1}H−^{15}N relaxation rates of backbone amides. All experiments involving heteronuclear experiments were acquired on uniformly ^{15}N-labeled samples at a concentration of typically 300 μM. For that, ^{15}N transverse (R_2) and longitudinal (R_1) relaxation rate constants, rotating frame longitudinal relaxation rates ($R_{1\rho}$), and the heteronuclear ^{1}H−^{15}N-NOE (hetNOE) were measured. Additionally, CPMG relaxation experiments were recorded. All experiments were processed and analyzed using Topspin 2.1 and the Sparky software package (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco, CA, USA). Heteronuclear relaxation was analyzed by reduced spectral density mapping translating R_2, R_1, and hetNOE data into spectral densities.[41] For further details, see the Supporting Information.

Residual Dipolar Couplings. Residual dipolar couplings (RDCs) were measured in stretched 7% polyacrylamide gels. For gel preparation, a gel preparation system from New Era Enterprises was used. Spectra of isotropically and anisotropically tumbling samples were measured at 293 K and at a sample concentration of typically 300 μM. A pseudo-three-dimensional spectrum was recorded in a J-modulated manner.[42] RDCs were obtained by the subtraction of couplings isotropically and anisotropically tumbling samples. RDCs were normalized to 10 Hz HDO (deuterium protium oxide) splitting.

Diffusion-Ordered Spectroscopy. Translational diffusion was measured using diffusion-ordered spectroscopy (DOSY) in 100% D$_2$O at pH 2.0 (corrected for D$_2$O) at 20 °C and at a typical sample concentration of 300, 150, and 75 μM. Dioxane was used as internal diffusion standard in 5-fold molar excess compared to the protein sample. The experiments were recorded according to established methods. For further details, see the Supporting Information.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Detailed experimental section, Figures S1−S12, Tables S1−S5, and Videos S1−S7. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

schwalbe@nmr.uni-frankfurt.de

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Mossuto, M. F.; Bolognesi, B.; Guixer, B.; Dhulesia, A.; Agostini, F.; Kumita, J. R.; Tartaglia, G. G.; Dumoulin, M.; Dobson, C. M.; Salvatella, X. *Angew. Chem., Int. Ed. Engl.* **2011**, *50*, 7048.

(2) Creighton, T. E. *Biochem. J.* **1990**, *270*, 1.

(3) Weissman, J. S.; Kim, P. S. *Science* **1991**, *253*, 1386.

(4) Collins, E. S.; Wirmer, J.; Hirai, K.; Tachibana, H.; Segawa, S.; Dobson, C. M.; Schwalbe, H. *ChemBioChem* **2005**, *6*, 1619.

(5) Klein-Seetharaman, J.; Oikawa, M.; Grimshaw, S. B.; Wirmer, J.; Duchardt, E.; Ueda, T.; Imoto, T.; Smith, L. J.; Dobson, C. M.; Schwalbe, H. *Science* **2002**, *295*, 1719.

(6) Wirmer, J.; Schlorb, C.; Klein-Seetharaman, J.; Hirano, R.; Ueda, T.; Imoto, T.; Schwalbe, H. *Angew. Chem., Int. Ed. Engl.* **2004**, *43*, 5780.

(7) Flory, P. J. *Principles of Polymer Chemistry*; Cornell University Press: Ithaca, NY, 1953.

(8) Flory, P. J. *Statistical Mechanics of Chain Molecules*; Hanser Publishers: Munich, 1989.

(9) Fiebig, K. M.; Schwalbe, H.; Buck, M.; Smith, L. J.; Dobson, C. M. *J. Phys. Chem.* **1996**, *100*, 2661.

(10) Smith, L. J.; Fiebig, K. M.; Schwalbe, H.; Dobson, C. M. *Folding Des.* **1996**, *1*, R95.

(11) Jarrett, N. M.; Djavadi-Ohaniance, L.; Willson, R. C.; Tachibana, H.; Goldberg, M. E. *Protein Sci.* **2002**, *11*, 2584.

(12) Matsuo, K.; Watanabe, H.; Tate, S.; Tachibana, H.; Gekko, K. *Proteins* **2009**, *77*, 191.

(13) Mishima, T.; Ohkuri, T.; Imoto, T.; Ueda, T. *Biosci., Biotechnol., Biochem.* **2007**, *71*, 2072.

(14) Noda, Y.; Yokota, A.; Horii, D.; Tominaga, T.; Tanisaka, Y.; Tachibana, H.; Segawa, S. *Biochemistry* **2002**, *41*, 2130.

(15) Ohkuri, T.; Shioi, S.; Imoto, T.; Ueda, T. *J. Mol. Biol.* **2005**, *347*, 159.

(16) Sakamoto, K.; Hirai, K.; Kitamura, Y.; Yamazaki, K.; Yusa, M.; Tokunaga, N.; Doi, G.; Noda, Y.; Tachibana, H.; Segawa, S. *Biopolymers* **2009**, *91*, 665.

(17) Tachibana, H. *FEBS Lett.* **2000**, *480*, 175.

(18) Yokota, A.; Hirai, K.; Miyauchi, H.; Iimura, S.; Noda, Y.; Inoue, K.; Akasaka, K.; Tachibana, H.; Segawa, S. *Biochemistry* **2004**, *43*, 6663.

(19) Yokota, A.; Izutani, K.; Takai, M.; Kubo, Y.; Noda, Y.; Koumoto, Y.; Tachibana, H.; Segawa, S. *J. Mol. Biol.* **2000**, *295*, 1275.

(20) Rohl, C. A.; Baldwin, R. L. *Biochemistry* **1997**, *36*, 8435.

(21) van den Berg, B.; Chung, E. W.; Robinson, C. V.; Dobson, C. M. *J. Mol. Biol.* **1999**, *290*, 781.

(22) Meier, S.; Grzesiek, S.; Blackledge, M. *J. Am. Chem. Soc.* **2007**, *129*, 9799.

(23) Wirmer, J.; Peti, W.; Schwalbe, H. *J. Biomol. NMR* **2006**, *35*, 175.

(24) Peti, W.; Smith, L.; Redfield, C.; Schwalbe, H. *J. Biomol. NMR* **2001**, *19*, 153.

(25) Schwarzinger, S.; Kroon, G. J.; Foss, T. R.; Chung, J.; Wright, P. E.; Dyson, H. J. *J. Am. Chem. Soc.* **2001**, *123*, 2970.

(26) Wishart, D. S.; Bigam, C. G.; Holm, A.; Hodges, R. S.; Sykes, B. D. *J. Biomol. NMR* **1995**, *5*, 67.

(27) Wishart, D. S.; Sykes, B. D.; Richards, F. M. *J. Mol. Biol.* **1991**, *222*, 311.

(28) Wishart, D. S.; Sykes, B. D.; Richards, F. M. *Biochemistry* **1992**, *31*, 1647.

(29) Zimm, B. H.; Stockmayer, W. H. *J. Chem. Phys.* **1949**, *17*, 1301.

(30) Bernado, P.; Blanchard, L.; Timmins, P.; Marion, D.; Ruigrok, R. W.; Blackledge, M. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 17002.

(31) de la Torre, J. G.; Huertas, M. L.; Carrasco, B. *Biophys. J.* **2000**, *78*, 719.

(32) Ortega, A.; Amorós, D.; García de la Torre, J. *Biophys. J.* **2011**, *101*, 892.

(33) Bernado, P.; Mylonas, E.; Petoukhov, M. V.; Blackledge, M.; Svergun, D. I. *J. Am. Chem. Soc.* **2007**, *129*, 5656.

(34) Schwalbe, H.; Fiebig, K. M.; Buck, M.; Jones, J. A.; Grimshaw, S. B.; Spencer, A.; Glaser, S. J.; Smith, L. J.; Dobson, C. M. *Biochemistry* **1997**, *36*, 8977.

(35) Ohsaku, M.; Allinger, N. L. *J. Phys. Chem.* **1988**, *92*, 4591.

(36) Otting, G.; Liepinsh, E.; Wuethrich, K. *Biochemistry* **1993**, *32*, 3571.

(37) Grey, M. J.; Wang, C.; Palmer, A. G. *J. Am. Chem. Soc.* **2003**, *125*, 14324.

(38) Szyperski, T.; Luginbühl, P.; Otting, G.; Güntert, P.; Wüthrich, K. *J. Biomol. NMR* **1993**, *3*, 151.

(39) Schlorb, C.; Ackermann, K.; Richter, C.; Wirmer, J.; Schwalbe, H. *J. Biomol. NMR* **2005**, *33*, 95.

(40) Konarev, P. V.; Volkov, V. V.; Sokolova, A. V.; Koch, M. H. J.; Svergun, D. I. *J. Appl. Crystallogr.* **2003**, *36*, 1277.

(41) Krizova, H.; Zidek, L.; Stone, M. J.; Novotny, M. V.; Sklenar, V. *J. Biomol. NMR* **2004**, *28*, 369.

(42) Tjandra, N.; Grzesiek, S.; Bax, A. *J. Am. Chem. Soc.* **1996**, *118*, 6264.

6854

dx.doi.org/10.1021/ja3009506 | *J. Am. Chem. Soc.* 2012, 134, 6846−6854