Characterisation of Disulfide-Bond Dynamics in Non-Native States of Lysozyme and Its Disulfide Deletion Mutants by NMR

Emily S. Collins,^[a] Julia Wirmer,^[a] Kenichi Hirai,^[b] Hideki Tachibana,^[c] Shin-ichi Segawa,^[b] Christopher M. Dobson,^[d] and Harald Schwalbe^{*[a]}

This report describes NMR-spectroscopic investigations of the conformational dynamics of disulfide bonds in hen-egg-white lysozyme substitution mutants. The following four systems have been investigated: $2SS^{\alpha}$, a lysozyme variant that contains C64A, C76A, C80A and C94A substitutions, was studied in water at pH 2 and 3.8 and in urea (8 m, pH 2); $2SS^{\beta}$ lysozyme, which has C6S, C30A, C115A and C127A substitutions, was studied in water (pH 2) and urea (8 m, pH 2). The NMR analysis of heteronuclear ¹⁵N-relaxation rates shows that the barrier to disulfide-bond isomerisation can vary substantially in different lysozyme mutants and depends on the residual structure present in these states. The investigations reveal cooperativity in the modulation of micro- to millisecond dynamics that is due to the presence of multiple disulfide bridges in lysozyme. Mutation of cysteines in one of the two structural domains substantially diminishes the barrier to rotational isomerisation in the other domain. However, the interactions between hydrophobic clusters within and across the domains remains intact.

Introduction

The conformation and dynamics of non-native proteins can have widely different characteristics. These can range from completely random-coil states, molten globule and partially folded conformations to those that, although mostly unstructured, still retain some residual native structure. Over the last decade, high-resolution NMR spectroscopic studies of isotopically labelled proteins have allowed a detailed delineation of the dynamics of non-native states and have also shown that both local and global residual structures can exist in these states. Such a description at atomic resolution is of importance for understanding the properties of the starting point of protein refolding. They could also be of importance for the understanding of the kinetics and probability of misfolding.

The structure of proteins can be stabilized by disulfide bonds. The dynamics around disulfide bonds can vary substantially in different protein states. Despite the fact that free rotation can occur around the S–S bond in the native state of a protein, in general it is slow and not detectable in many proteins by NMR spectroscopy. For example, in native wild-type lysozyme all four disulfide bridges are found to be locked into one predominant conformation on the NMR time scale. However, there are also reports on slow disulfide dynamics in native proteins. NMR studies on native bovine pancreatic trypsin inhibitor (BPTI) have revealed^[1,2] that some of the residues around the C14-C38 disulfide bridge exist in two identifiable conformations between 277–293 K, and have elevated T_2 values at 309 K. This suggests that there is rotation around the S-S bond between C14 and C38. Similar changes in relaxation rates have been reported for interleukin 4.[3] Slow chemical exchange involving different conformations of the disulfide bond has also been identified in cyclic peptides.^[4-6]

Little is known about this phenomenon for non-native states of proteins. Most NMR investigations thus far have focused either on proteins that lack disulfide bridges,^[7-9] on states in which cysteine bridges have been permanently reduced^[10,11] or in which some or all cysteine residues have been substituted, for example, by alanine.^[12-14] van Mierlo et al.^[13,15] studied an analogue of the predominate folding intermediate of BPTI in which only one of the three native disulfide bonds was present, between C30 and C51, and the other cysteine residues were replaced by serine. This mutant was partly folded, which was thought to be the reason why it predominates, and was in equilibrium with a completely unfolded form. Backbone-dynamics studies showed that the regions that contained the

[a] Dr. E. S. Collins, Dr. J. Wirmer, Prof. Dr. H. Schwalbe Institute for Organic Chemistry and Chemical Biology Center for Biomolecular Magnetic Resonance Johann Wolfgang Goethe University Marie-Curie-Straße 11, 60439 Frankfurt am Main (Germany) Fax: (+49) 69-798-29515 E-mail: schwalbe@nmr.uni-frankfurt.de
[b] Dr. K. Hirai, Prof. Dr. S.-i. Segawa School of Science and Technology, Department of Physics

Kwansei Gakuin University 2-1 Gakuen, Sanda, 669-1337 (Japan)

[c] Dr. H. Tachibana Department of Biology, Faculty of Science and Graduate School of Science and Technology, Kobe University Rokkodai, Nada-ku, Kobe, 657-8501 (Japan)

- [d] Prof. C. M. Dobson
- Department of Chemistry, University of Cambridge Lensfield Road, Cambridge, CB2 1EW (UK)
- Supporting information for this article is available on the WWW under http://www.chembiochem.org or from the author.

other disulfide bridges in the wild-type BPTI were flexible. The authors interpreted this flexibility as necessary for the rapid formation of the remaining disulfide bridging.

In this report, we show that the rotational interconversion of the R/S stereoisomers around disulfide bonds is one of the key determinants of microsecond to millisecond dynamics of non-native states of proteins. It influences the conformational heterogeneity of proteins when unfolded in high concentrations of denaturants. Furthermore, it affects the rate of disulfide-bond formation in oxidative refolding, which constitutes one rate-limiting step in the folding of proteins with multiple disulfide bonds. Oxidative folding has been studied in great detail in vitro, for example, for ribonuclease A,^[16-20] BPTI^[21-23] and henegg-white lysozyme.^[24-27] Initial in vivo investigations into the process have also been performed, for example, on the HIV-1 envelope glycoprotein.[28] However, many aspects of the role of disulfide bridges in protein folding remain poorly understood. Some proteins can attain their native fold prior to the forma-

tion of disulfide bridges. This indicates that in such instances the disulfide bridges provide a stabilising effect in the final structure. Other proteins fail to fold in the absence of some disulfide bridges; for example, the absence of two of the disulfide bridges in the β -domain of hen-egg-white lysozyme (between cysteines 64–80 and 76–94) destroys its ability to refold.^[29,30]

We have studied hen-egg-white lysozyme mutants in which the disulfide bonds have been removed. Lysozyme contains two structural domains: the α -domain, which contains residues 1–35 and 85–129, and the β -domain, which comprises residues 36–84.^[31–33] It has eight cysteine residues that form two disulfide bridges in the α -domain (C6–C127 and C30–C115) and two in the β -domain (C64–C80 and C76–C94), as shown in Figure 1.

Using ¹⁵N transverse heteronuclear-relaxation rates (R_2), we have investigated the backbone dynamics of two lysozyme mutants in which either the disulfide bonds in the α - or β -domain have been removed. Previous dynamical studies of ly-

sozyme^[10, 34, 35] that was either denatured in urea (8 M) or dissolved in water at pH 2, have indicated the presence of hydrophobic clusters that are mostly centred around the tryptophan residues, and of us-ms exchange processes that are centred around the disulfide bridges. This study builds on the previous report by Noda et al. who studied the doubly disulfide-bridged variant of lysozyme, $2SS^{\alpha}$, which has intact disulfide bridges in the α -domain only. This group was able to show



Figure 1. Ribbon diagram of the structure of wild-type hen-egg-white lysozyme derived from its crystal structure^[31] (6LYZ) from the Protein Data Bank (http://www.rcsb.org/pdb/ index.html). The eight cysteine residues that form the disulfide bridges are shown in yellow and their positions in the sequence are indicated. Separate diagrams of the four disulfide bridges are also shown, viewed along the S–S bond. Angles were measured in Insight (Biosym/MSI, CA, USA).

that the protein is partially folded in water at pH 3.8 whereas the 2SS^{β} variant, which has intact disulfide bridges in the β -domain only, was unfolded under the same conditions.^[29,30,36]

Results

The current investigations have been performed with two lysozyme mutants in two different solvent systems. Mutant 2SS^{α}, which contains the substitutions C64A, C76A, C80A and C94A and in which only the α -domain disulfides are intact, was studied in water (pH 3.8) and urea (8 M, pH 2). The second mutant, 2SS^{β}, with the substitutions C6S, C30A, C115A and C127A and in which only the β -domain disulfides are intact, was studied in water (pH 2) and urea (8 M, pH 2). Table 1 lists the lysozyme variants that have either been studied here or previously and summarises their dynamical characteristics. References are included for variants that are the subject of previous studies; those presented here are highlighted in bold.

| Table 1. Summary of the behaviour of different lysozyme species under various denaturing and nondenaturing conditions. | | | | |
|--|--|--|--|--|
| Lysozyme species | in water | in urea (8 м, pH 2) | | |
| 4SS-WT | folded ^[43] | Random-coil behaviour, nonrandom clusters and µs–ms dynamics observed ^[10] | | |
| 0SS-WT | random-coil behaviour, nonrandom clusters observed ^[34] | random-coil behaviour, nonrandom clusters observed ^[10, 34] | | |
| 0SS-W62G | random-coil behaviour, no nonrandom clus- ters ^[34] | not studied | | |
| 4SS-W62G | folded, HSQC spectrum is identical to WT spectrum | random-coil behaviour, μs–ms dynamics almost eliminated ^[34] | | |
| 2SS ^{α[a]} | at pH 3.8 partly folded | random-coil behaviour | | |
| $2SS^{\beta[a]}$ | at pH 2.0 random-coil behaviour | random-coil behaviour | | |
| [a] Hen-egg-v | white lysozyme mutants studied in this work. | | | |

Resonance assignments of lysozyme 2SS variants

Assignments of the ¹H and ¹⁵N resonances of 2SS^{α} and 2SS^{β} were carried out by using 3D-NOESY-HSQC and ¹H,¹⁵N-HSQC experiments and are given in the Supporting Information. In most parts, resonance assignments could be obtained by using the strategy outlined before for non-native protein states.^[10, 37]

Of the 129 residues, 107 could be assigned for $2SS^{\alpha}$ in water (pH 3.8); resonance assignments were mainly missing for residues in the β -domain. The assignments obtained mostly agreed with those of Noda et al.^[30] The ¹H,¹⁵N-HSQC spectrum of partly folded $2SS^{\alpha}$ in water (pH 3.8) contains more than the expected 127 backbone NH resonances. Some of the unassigned peaks were considerably weaker than the other resonances. This indicates the presence of a minor, less folded conformer of the lysozyme variant as was also seen for a partially folded intermediate of BPTI.^[13] Five of these weaker resonances with ¹⁵N chemical shifts of around 108–110 ppm—a region where only glycine resonances are observed in a random-coil protein-were observed in addition to the twelve assigned glycines. The weaker residues increase in relative intensity as the temperature increases. This indicates that the equilibrium moves more towards the unfolded conformation (Figure S1 in the Supporting Information). For all the denatured 2SS variants, a substantial proportion of the peaks could be assigned (107 for $2SS^{\alpha}$ in 8 M urea, 108 for $2SS^{\beta}$ in 8 M urea and 96 for $2SS^{\beta}$ in water; all solutions were at pH 2). In general, resonances of residues from the β -domain were more difficult to assign. This was either due to the presence of mutations in this region, or because the peaks are significantly broadened due to the presence of the disulfide bridges.

Since the chemical shifts of the denatured variants are mostly very similar to those of wild-type (WT) lysozyme,^[10] similar conclusions on the presence and position of secondary structures in the mutants can be drawn. The chemical shifts are generally very close to values measured for unstructured peptides^[38] and significant changes in H_a shifts are seen in particular regions of the sequence (between residues 19–32, 58–63, 106–112 and 120–125). A further discussion of residual secondary structure as evidenced from these chemical shifts has been published previously for other lysozyme mutants^[10] and is beyond the scope of the present publication.

Interpretation of backbone dynamics

Measurement of the backbone dynamics of denatured proteins can reveal important aspects of the structural characteristics and motions of the non-native states. Common measurements for the investigation of backbone dynamics in proteins are the ¹⁵N-heteronuclear relaxation rates R_1 and R_2 and the steadystate ¹H,¹⁵N-NOE. For unstructured proteins, R_2 rates provide the most information about residual structure and processes, such as isomerisation of disulfide bridges, as they are sensitive to motions on both the ns–ps and ms–µs timescales. Exchange processes thar occur on the ms–µs timescale, such as the rotational isomerisation of the disulfide bond, are revealed through increased line widths in spectra and elevated R_2 rates.

Millisecond-to-microsecond timescale exchange processes have been investigated in folded BPTI and an SH3 domain by using relaxation-dispersion techniques.^[39,40] In the present investigation, we attempted to determine the contribution from the isomerisation of the disulfide bridges to R_2 rates in the 2SS variants of lysozyme, using relaxation-dispersion techniques (data not shown). However, it was not possible to obtain an R_{ex} value from these measurements, because the exchange rates in the 2SS variants of lysozyme are too fast; they most reliably quantify processes in the order of hundreds of microseconds to several milliseconds.^[41] Also, some of the most interesting peaks with regard to millisecond exchange are broadened beyond detectability.

Backbone dynamics of W62G (0SS-W62G, 4SS-W62G) and WT lysozyme (0SS-WT, 4SS-WT)

The R_2 relaxation rates of disulfide reduced and methylated WT lysozyme (0SS-WT) and of the disulfide reduced and methylated W62G mutant (0SS-W62G) have been studied before^[10,34] and are reproduced in Figure 2. Mutant W62G, has been



Figure 2. R_2 values for a) 4SS-W62G in urea (8 m, pH 2) and 0SS-W62G in water, pH 2. The continuous line shows the modelled R_2 values for a branched-polypeptide chain by using Equation (2) and an R_2^{int} of 0.29 s⁻¹. b) 4SS-WT and 0SS-WT in urea (8 m, pH 2). The continuous line shows the fitting of the R_2 values of 4SS-WT to Equation (4) where $R_2^{\text{int}} = 0.29 \text{ s}^{-1}$, $R_2^{\text{exch}} = 10.9 \text{ s}^{-1} \lambda_1 \lambda_2 = 7$.

CHEMBIOCHEM

shown to approximate a "true" random-coil polypeptide chain in terms of its overall diffusive properties^[35] when in the S-methylated state (0SS-W62G). In this mutant, all nonlocal interactions between locally defined clusters are significantly reduced. The mutation reduces the relaxation rates in the areas of hydrophobic clustering seen in 0SS-WT lysozyme by inhibiting the long-range interactions between clusters. It defines the random-coil behaviour for the disulfide mutants discussed in the following.

The R_2 rates of the oxidised form of W62G (4SS-W62G) are also reproduced in Figure 2 a. The variation of the R_2 values is significantly lower than for denatured WT oxidised lysozyme (4SS-WT; Figure 2 b). However, the positions of the maxima still correspond to the location of the disulfide bridges. This indicates that the lack of long-range interactions reduces the motional constraints centred around the disulfide bonds in 4SS-WT and allows a faster rate of rotational isomerisation around the disulfide bridges.

Backbone dynamics of $2SS^{\alpha}$ in urea (8 M, pH 2)

The R_2 values for the backbone NH's of 2SS^{α} are shown in Figure 3. The R_2 rates for 95 residues reveal considerable variation between 3.2 and 11.4 s⁻¹. Whilst the R_2 values of 2SS^{α} in



Figure 3. R_2 values for 2SS^{*a*} and 0SS-WT in urea (8 M, pH 2). The black continuous line shows the fitting of the R_2 data to models for a branched polymer chain as described in the text, where $R_2^{\text{int}}=0.25 \text{ s}^{-1}$, $R_2^{\text{exch}}=4.36 \text{ s}^{-1}$ and $\lambda_1, \lambda_2=7$.

the α -domain have a very similar pattern to those of 4SS-WT, they vary considerably from 4SS-WT in the β -domain. In contrast, the R_2 values for the residues in the β -domain closely resemble those of 0SS-WT. This indicates that the hydrophobic clustering seen for wild-type lysozyme is undisturbed by the cysteine to alanine substitutions.

We therefore conclude that the R_2 rates can be dissected into contributions that arise from the hydrophobic clusters and from the disulfide bridges. In fact, the R_2 values of the α domain of 2SS^{α} lysozyme and the $R_{1\rho}$ values of urea-denatured 4SS-WT are correlated (Figure S2 in the Supporting Information). This observation arises presumably because $R_{1\rho}$ values are much less susceptible to exchange on the μ s-ms timescale than R_2 values.^[1] R_2 and $R_{1\rho}$ values of 4SS-WT (Figure 2b, and Supporting Information Figure S2) differ drastically.^[10] These differences are presumed to be due to the fact that the regions that surround the cysteines in wild-type lysozyme are subject to considerable μ s-ms timescale motions, which is most likely due to rotational isomerisation of the disulfide bond. We conclude that the disruption of two out of four disulfide bridges in 2SS^{α} reduces the μ s-ms timescale motions. The $R_{1\rho}$ rates of lysozyme 2SS^{α} were also measured to confirm the disappearance of μ s-ms motions (see Figure S2 in the Supporting Information) and were found to be very similar but slightly lower than the R_2 rates. Therefore, a small contribution from exchange effects is present in the measured R_2 rates.

Backbone dynamics of $2SS^{\beta}$ in urea (8 M, pH 2)

The R_2 relaxation rates of $2SS^{\beta}$ in urea (8 M, pH 2) are quite different to those of $2SS^{\alpha}$ (Figure 4).



Figure 4. R_2 values for 2SS^β and 0SS-WT in urea (8 м, pH 2). The black continuous line shows the fitting of the R_2 data to models for a branched-polymer chain as described in the text, where $R_2^{\text{int}} = 0.25 \text{ s}^{-1}$, $R_2^{\text{exch}} = 4.25 \text{ s}^{-1}$ and $\lambda_1, \lambda_2 = 7$.

 R_2 rates could be measured for 106 residues. For this variant, the R_2 rates of the α -domain generally fit well to the OSS-WT R_2 values, whilst those of the β -domain correlate well with the $R_{1\rho}$ values of 4SS-WT lysozyme. This indicates that a reduced amount of μ s-ms exchange is occurring in the β -domain since the R_2 values have a similar pattern to 4SS-WT in urea but are considerably lower. The $R_{1\rho}$ values of 2SS^{β} (Figure S3 in the Supporting Information) are reduced to 6 s⁻¹ compared to the R_2 values of the same variant. This indicates that some exchange on the μ s-ms timescale affects the R_2 rates.

Backbone dynamics of $2SS^\beta$ in water, pH 2

The deviations from random-coil predictions in the R_2 rates of OSS-WT are more pronounced when studied in water (pH 2) than urea (8 M, pH 2). Therefore, the dynamics of 2SS^{β} were also measured in water (pH 2), that is, conditions under which 2SS^{α} is not completely unfolded. The R_2 rates for 92 assigned peaks are shown in Figure 5. The chemical shift dispersion of



Figure 5. R_2 values for 2SS^{β} in water, pH 2 and 2SS^{β} in urea, pH 2.

 $2SS^{\beta}$ in water (pH 2) is reduced compared to the other mutants that were studied. In addition, several peaks are too broad to be seen in any of the spectra. This indicates that they must have even larger R_2 values than those measured. An example for this is W62, which has a very distinctive chemical shift in all mutants, but which cannot be detected above the noise in $2SS^{\beta}$. This suggests, that although they could not be assigned, there could nevertheless be several residues in the β -domain with R_2 values of 15 s⁻¹ or greater.

In the β -domain and for residues 1–36, the R_2 values of 2SS^{β} in water (pH 2) are very similar to those of the same variant in 8 μ urea. For residues 95–115, however, some of the R_2 values of 2SS^{β} in water are considerably higher than for the variant in 8 μ urea. The $R_{1\rho}$ rates of 2SS^{β} in water are slightly lower than the R_2 rates. This indicates that only a small amount of μ s-ms timescale exchange is occurring (Figure S4 in the Supporting Information).

Backbone dynamics of 2SS^a in water, pH 3.8

At pH 3.8, $2SS^{\alpha}$ becomes partly folded. Therefore, R_2 values of the partly folded state of $2SS^{\alpha}$ were measured in water, pH 3.8 at 293 K (Figure 6).



Figure 6. R_2 values for partly folded 2SS^{α} in water, pH 3.8 and denatured 2SS^{α} in urea (8 M, pH 2).

The profile of the R_2 values is reasonably flat in the folded region of the protein, as would be expected, and the average R_2 value is 10.1 s⁻¹ for all residues and 10.2 s⁻¹ for α -domain residues. Some greater variation is observed in the $\beta\text{-domain},$ but it was not possible to measure R_2 values for all of this domain due to a lack of assignments in the region. In total, 103 assigned R_2 values could be measured for the partly folded lysozyme. R_2 values were also measured for most of the unassigned peaks and the majority of these are around 5-6 s⁻¹. Thus, they are more similar to the values seen for denatured lysozyme; this indicates that they arise from the unfolded region of the variant. The unassigned resonances in the region typical for glycine NH's also have low R_2 values, thus supporting the proposition that they are from a completely unfolded conformer which is in equilibrium with the partly folded conformer. The published R_2 rates of native lysozyme were measured at 35 °C and 500 MHz by Buck et al.^[42] and Boyd and Redfield^[43] and have an average value of 7.6 s^{-1} . However, if the change in frequency and the effect of the 15° difference in temperature on the solution viscosity and hence the overall correlation time is taken into account, then a value of $\sim 11 \text{ s}^{-1}$ would be expected for WT lysozyme at 20 °C and 600 MHz (assuming an S² value of 0.9 for all residues). Therefore, the R_2 rates of the α -domain of 2SS^{α} in water (pH 3.8) are just a small amount lower than those of the native lysozyme. This suggests that the protein is just slightly more dynamic than the wild type, probably due to missing structural constraints in the β -domain. When comparing the R_2 values of the unfolded (8 μ urea, pH 2) to those of the partly folded 2SS^{α} (water, pH 3.8), it is worth noting that the values of the residues that immediately surround C30, C115 and, to a lesser extent, C127, are very similar in both measurements. Since, the R_2 values for residues that surround cysteines in the partly folded form of $2SS^{\alpha}$ are not significantly higher than the average R_2 value of 10.1 s⁻¹ (for assigned residues), there is little evidence that millisecond-timescale chemical exchange is occurring.

Discussion

Modelling of R_2 relaxation rates

 R_2 rates of non-native states of proteins, in particular hen-egg white lysozyme,^[10, 34, 35] have been modelled quite successfully in the past by using a simple model that assumes that an unfolded protein behaves like an unbranched polymer in solution. Primary factors that are responsible for an increase in the R_2 rates—above an intrinsic R_2 rate that is determined by the temperature and viscosity of the solution—are the length of the polymer chain (i.e.: the number of residues) and the persistence length of the chain. The addition of different terms into an initial equation [Eq. (1)] that models the behaviour of a "random coil" polypeptide chain, can characterise the effect of the introduction of i) branching in the chain (i.e., addition of disulfide bridges) [Eq. (2)], ii) hydrophobic clusters [Eq. (3)], or iii) exchange processes on the dynamics of the non-native

CHEMBIOCHEM



Figure 7. a) Example of model R_2 rates for an unbranched polypeptide chain, such as 0SS-WT. Values were obtained by using Equations (1) (....) and (3) (....). The six hydrophobic clusters in non-native states of 0SS-WT are indicated. b) Model R_2 rates for a branched polypeptide chain, such as 4SS-WT. Values were obtained by using Equations (2) (....) and (4) (.....). The positions of the cysteine residues in lysozyme are indicated above the graph.

states [Eq. (4)]. In Figure 7 a graphical illustration of the shape of the fitting produced with each equation is shown.

$$R_2(i) = R_2^{\text{int}} \sum_{j=1}^{N} \exp\left(-\frac{i-j}{\lambda_1}\right)$$
(1)

$$R_2(i) = R_2^{\text{int}} \sum_{j=1}^{N} \exp\left(-\frac{dm_{ij}}{\lambda_1}\right)$$
(2)

$$R_{2}(i) = R_{2}^{int} \sum_{j=1}^{N} \exp\left(-\frac{dm_{ij}}{\lambda_{1}}\right) + \sum_{cluster} R_{cluster} \exp\left(-0.5\left(\frac{|i-\chi_{cluster}|}{\lambda_{cluster}}\right)^{2}\right)$$
(3)

$$R_2(i) = R_2^{\text{int}} \sum_{j=1}^{N} \exp\left(-\frac{dm_{ij}}{\lambda_1}\right) + R_2^{\text{exch}} \sum_{k=1}^{N_{\text{cys}}} \exp\left(-\frac{|(i-\mathsf{Cys}_k)|}{\lambda_2}\right) \quad (4)$$

 R_2^{int} is a measure of the intrinsic relaxation rate; *i* and *j* represent residue numbers; *N* is the total chain length of the protein (i.e., number of residues); λ_1 is the persistence length of the polypeptide chain (in number of residues); λ_{cluster} and λ_2 are the persistence lengths due to the effects of either the hydrophobic clusters or the exchange contribution from the disulfide bonds, respectively. dm_{ij} is a topological distance matrix which is used to count the number of covalent bonds along the shortest path between residue *i* and *j*. R_{cluster} and χ_{cluster} describe the amplitude and position that characterise each cluster and R_2^{exch} is the amplitude of the exchange contribution from the disulfide bonds. Cys_k is the position of each cysteine residue. Since in the case of an unbranched chain dm_{ij} reduces to just (*i*–*j*), Equation (3) can be used to model both branched and unbranched chains.

The R_2 relaxation rates of OSS-W62G can, to a good approximation, be characterized by a model of an *unbranched* polypeptide chain given in Equation (1). The increased R_2 relaxation rates (Figure 2a) of 4SS-W62G have been modelled as a *branched* polypeptide chain, as given in Equation (2). The R_2 rates of the lysozyme variants were analysed by using models for the dynamical behaviour of either branched or unbranched polymer chains, as described previously.^[10,34] The parts of the protein where cysteine residues were substituted to alanine

were analysed by using the model for an unbranched polypeptide chain. However, with the inclusion of a second term to model hydrophobic clusters, the data were fitted to Equation (3). The parts containing native cysteines were analysed with the model for a branched chain by using Equation (4). The model values for a branched polypeptide chain just undergoing segmental motion were obtained with Equation (2).

The fitting of the R_2 rates of urea denatured 2SS^{α} to a combination of the models for *branched* and *unbranched* polymer chains [Eqs. (3) and (4)] is shown in Figure 3. Relaxation rates of residues in the α -domain were fit to the model for a *branched* chain and those for residues in the β -domain to the model for an *unbranched* chain. Fitting of the data gave a value for R_2^{exch} of 4.36 s⁻¹, a R_2^{int} of 0.25 s⁻¹ and a λ_1, λ_2 value of 7. The R_2^{exch} is thus reduced by more than a factor of two compared to the value of 10.9 s⁻¹ that was obtained for 4SS-WT.

The fitting of the R_2 data for $2SS^{\beta}$ in urea (8 m, pH 2) to the models for *unbranched* and *branched* polymer chains is pictured in Figure 4. The four clusters present in the α -domain can be distinctly identified. The clusters in the β -domain are difficult to identify due to the summed effect of hydrophobic clusters and elevated dynamics around the disulfide bridges. As was seen for urea denatured $2SS^{\alpha}$, the value for $R_2^{\text{exch}} = 4.25 \text{ s}^{-1}$ is lower than the value obtained for 4SS-WT ($R_2^{\text{exch}} = 10.9 \text{ s}^{-1}$), but λ_1, λ_2 values of 7 are still obtained along with a R_2^{int} of 0.25 s⁻¹.

Correlation of the location of hydrophobic clusters and disulfide dynamics

In the R_2 rates of native lysozyme there is little evidence that the residues around the cysteines are affected by slow chemical exchange processes. This suggests that the rate of rotational isomerisation of the disulfide bridging is most probably slowed or completely retarded by the native structure. Two out of the four disulfides (64–80 and 76–94) in native lysozyme adopt a χ_3 angle of around 90°, while the other two adopt an angle of around –90° (Figure 1). These conformations are typical for disulfide bridges in proteins^[44] and suggest that disulfide isomerisation can be interpreted by a two-site exchange between χ_3 angle of \pm 90°.

In a random-coil denatured-polymer chain containing a single disulfide bond there are no significant and additional barriers to the rotational isomerisation of this bond. However, the introduction of hydrophobic clusters provides an energetic barrier to disulfide bond rotational isomerisation. Disruption of hydrophobic clusters by a mutation such as W62G, results in much lower R_2 rates in the 4SS-W62G mutant than in 4SS-WT when denatured in 8 M urea. Here, we examine the effect of the removal of disulfide bridging in individual domains on the rate of rotational isomerisation of the remaining disulfides.

From the investigation of the relaxation properties of four different variants of lysozyme, the following conclusions can be drawn:

1) When the 2SS variants of lysozyme are denatured, either in 8 m urea and/or water at pH 2, the hydrophobic clusters

appear to be undisturbed by the presence or absence of particular disulfide bridges—the R_2 rates over the parts of the sequence that lack disulfide bonds are generally very similar to those in OSS-WT lysozyme. The disulfide bonds therefore do not induce, but might stabilise, hydrophobic clusters.

- 2) The R_2 rates of 2SS^{α} in 8 \bowtie urea are more similar in the α domain to the $R_{1\rho}$ than the R_2 rates of the 4SS-WT denatured lysozyme. This indicates that a large part of the barrier to disulfide isomerisation has disappeared. The amount of μ s-ms chemical exchange seen to be affecting the R_2 rates is dramatically reduced. We conclude that the removal of the two β -domain disulfide bridges of lysozyme also has an effect on the α -domain disulfide bridges; this suggests a degree of cooperation between these two domains. As the number of disulfide bridges increases so does the barrier to disulfide-bond isomerisation. In the β -domain, the protein has similar properties to 0SS-WT, with just the interactions between hydrophobic residues affecting the R_2 rates.
- 3) $2SS^{\beta}$ exhibits converse behaviour. It has R_2 rates that are similar to 0SS-WT in the α -domain, but which are close to the $R_{1\rho}$ rates of 4SS-WT in the β -domain. The R_2^{exch} rate for both variants is decreased by more than a factor of two (Table 2). This supports the notion of cooperativity between the disulfide bonds in the two domains.

| Table 2. Summary of calculated $R_2^{\rm int}$, λ_1 , λ_2 and $R_2^{\rm exch}$ values for different lysozyme variants. | | | | | |
|--|-----------------------------------|------------------------|---------------------------------------|--|--|
| Lysozyme species | $R_2^{\rm int}$ [s ⁻] | λ_1, λ_2 | R_2^{exch} [s ⁻] | | |
| 4SS-WT | 0.29 | 7 | 10.9 | | |
| 0SS-WT (8м urea) | 0.25 | 7 | n/a | | |
| 2SS ^α | 0.25 | 7 | 4.36 | | |
| 2SS ^β | 0.25 | 7 | 4.25 | | |

4) The folded regions of the partly folded 2SS^{α} in water (pH 3.8) behave much like the corresponding areas in native lysozyme. The unstructured parts, however, behave like a completely random-coil protein. For this state, an additional unfolded conformer can be identified. The simultaneous existence of this additional unfolded conformer with the partly folded conformation of $2SS^{\alpha}$ is interesting particularly, as the nonoxidative folding pathway of lysozyme is known to consist of at least two parallel routes.[45,46] The majority of molecules fold through a slow pathway with folding intermediates. However, approximately 20% of the molecules fold through a faster pathway that involves no intermediates.^[45,47,48] In the slower pathway, the first intermediate formed is one where the α -domain is protected from hydrogen exchange, but the β -domain is still exposed.

Disulfide bridges in protein folding

In the oxidative folding of a protein, the introduction of a disulfide bridge is thought to have a mainly entropic effect on

the unfolded state; by imposing distance and angle constraints between the C_{β} and C_{α} atoms of the two cysteine residues that are involved in forming the bridge, the conformational entropy of the system is greatly reduced.^[49-51] The farther apart the two cysteines are in the primary structure, the greater this decrease in conformational entropy will be. Disulfide bridges have also been suggested to have an enthalpic effect by stabilising local interactions such as those between hydrophobic residues which cause local clusters.^[16] The entropic effect might be one reason why 2SS^{α} is partly folded under nativelike conditions, whereas the $2SS^{\beta}$ is unstructured. The cysteine residues that form the disulfide bonds in the α -domain are much further apart in the primary sequence than those in the $\beta\text{-domain.}$ Consequently, when they form disulfide bonds the conformational entropy decreases much more than the formation of those in the β -domain. Whether the disulfide bridges in lysozyme have a stabilising effect on hydrophobic interactions is less clear; the presence or lack of disulfide bonds seem to have no effect on the hydrophobic interactions as judged by the dynamics of the unfolded states. Some hydrophobic collapse occurs before the formation of disulfide bridges in the folding of lysozyme, and almost all hydrophobic clusters identified by dynamics surround cysteine residues. These findings suggest that in fact the hydrophobic interactions help to bring the cysteine residues into a position where they can form a disulfide bridge.

In order to look at the effect disulfide bridges on the folding of hen-egg-white lysozyme, it is also useful to compare the structural details of variants of lysozyme that contain different numbers of disulfide bridges. In the absence of disulfide bridges-caused either by the S-methylation of the cysteine residues, or their substitution with alanines or serine-the protein is unfolded even under native-like conditions. $^{\scriptscriptstyle [52,53]}$ The 1SS variants of lysozyme, which have one native disulfide bond, are also unstructured.^[54] The different 3SS variants of lysozyme have been studied by NMR spectroscopy and X-ray crystallography. These variants were generated either by selective carboxymethylation of C6 and C127 (CM^{6,127} lysozyme),^[25] isolation of a kinetically trapped three-disulfide intermediate des-[76-94]^[26] or cysteine-to-alanine/serine substitution to produce three-disulfide variants C64A/C80A, C76A/C94A and C30A/ C115A.^[55] All of the 3SS variants have CD spectra that are similar to native lysozyme, but differences in their structural properties have been detected. CM^{6,127} lysozyme has a virtually identical structure to native lysozyme. In the 3SS[6-127,64-80,76–94] variant, one helix and a helix interface in the α domain are either unstructured or have large fluctuations. 3SS[6–127,30–115,76–94] has regions in the β -domain that remain unstructured and flexible, whereas 3SS[6-127,30-115,64–80] has a compact native-like structure in both the α and β domains.

By using Equation (2), R_2 rates for oxidised lysozyme with different numbers of disulfide bridges and with and without (here only the first part of the equation is used) hydrophobic clustering can be predicted (Figure 8). Examination of these calculated rates shows that the introduction of a disulfide bond into lysozyme changes the R_2 rates differently depending

CHEMBIOCHEM



Figure 8. Graphs showing predicted R_2 rates for different denatured variants of lysozyme. The values were calculated by using Equation (2). a) 4SS-WT, b) 2SS^{α}, c) 2SS^{β}, d) 1SS[6-127], e) 1SS[30-115], f) 1SS[64-80] and g) 1SS[76-94]. Gray dotted lines in a), b) and c) illustrate the predicted R_2 rates when the fitted hydrophobic clusters of 0SS-WT are added to those calculated from Equation (2).

on the position of the bond. For example, the R_2^{max} for 1SS[30– 115] is significantly higher than that of 1SS[6–127]. The calculated R_2 rates for the 2SS variants illustrate that the introduction of a second disulfide bond to the α -domain of an α domain 1SS variant has a much greater effect on the R_2 rates of the protein as a whole, than when the bond is introduced in the β -domain of a β -domain 1SS mutant. The calculated rates for the 4SS protein are almost a linear combination of the α - and β -domain rates from 2SS^{α} and 2SS^{β} with small additional increases that are due to the proximity of SS[30–115] and SS[76–94].

The calculated R_2 rates correlate well with the structural details known for lysozyme variants. The introduction of SS[30– 115] creates the largest increase in R_2 rates and therefore causes the most constraints to the motions of the protein. This could explain why the 3SS variant that lacks this disulfide bond is the least structured of the 3SS mutants and also why des[30–115] has not been observed in the oxidative folding pathway of lysozyme.^[26] This stabilising effect of the SS[30– 115] bond could also partly explain—in addition to the entropic effects mentioned earlier—why $2SS^{\alpha}$ is partly folded and $2SS^{\beta}$ completely unfolded.

In conclusion, the work presented here supports the framework for the interpretation of relaxation rates in partly oxidised non-native states of proteins. In the absence of hydrophobic clusters, the relaxation rates can be interpreted by a topological description of the polymer. In turn, deviations from these predictions clearly correlate with additional barriers on torsion angles. Together with mutational studies, NMR spectroscopy allows the dissection of the contribution of hydrophobic clusters and disulfide bridges to the dynamic heterogeneity of non-native protein states. Due to their slow timescales, these effects inevitably result in significant barriers in the course of protein folding.

Experimental Section

Preparation of hen-egg-white lysozyme mutants: The two disulfide-bridge variants of lysozyme, 2SS^{α} (C64A, C76A, C80A and C94A) and 2SS^{β} (C6S, C30A, C115A and C127A) were prepared as described by Tachibana et al. and Noda et al.^[29,30] Protein was overexpressed in *E. coli* grown in M9 media, purified from inclusion bodies and oxidatively refolded as detailed by Tachibana et al. and others.^[29,36,56] Preparative reverse-phase HPLC was used to separate the different 2SS isomers.^[29,36]

NMR measurements: Lyophilised samples of the 2SS lysozyme variants were dissolved in either H₂O (90%)/D₂O (10%), pH 2 or 3.8, or urea (8 m) with D₂O (10%) at pH 2. Sample concentrations were $\sim 1 \text{ mm}$ for 2SS^{α} and 300 μ m for 2SS^{β} in water or urea.

All NMR experiments were recorded on either a Bruker DRX 600 spectrometer with a z axis gradient ¹H{¹³C,¹⁵N} triple resonance probe or a Bruker AV 900 MHz spectrometer equipped with a 5 mm ${}^{1}H{}^{13}C{}^{15}N$ XYZ-Grad TXI probe. For the assignment of the peaks in the ¹H,¹⁵N-HSQC spectra of the unfolded 2SS variants, 3D ¹H-¹H-¹⁵N NOESY-HSQC spectra were recorded at 293 K with mixing times of 200-250 ms. A natural abundance HNCA was recorded at 900 MHz by using a 5 mm ${}^{1}H{}^{13}C{}^{15}N{}$ cryoprobe, on $2SS^{\beta}$ in water at 293 K. All relaxation experiments were measured at 600 MHz at 293 K. Backbone NH ¹⁵N-relaxation times were measured with a standard CPMG pulse sequence^[57,58] either as a set of 9-10 2D experiments or in an interleaved fashion. For each sample, seven different R₂ delays were used between 18 and 252 ms; two or three delay values were repeated. Experiments were acquired with either 8 or 16 scans with a delay of 3 s between scans.

In order to assess the temperature dependence of the partly folded 2SS^{α} in water a series of ¹H $^{-15}$ N HSQC spectra were measured at 5 K intervals between 293 K and 308 K.

 $R_{1\rho}$ measurements were performed with a pulse sequence derived from that of Korzhnev et al.,^[59] with adiabatic pulses before the spin-lock pulse.^[60] Experiments were measured interleaved with a spin-lock field strength of 1.7 kHz and a 2 kHz ¹⁵N offset. For each sample, 5–7 delay values between 20 and 160 ms were recorded. These were interleaved and three delay values were repeated for error analysis. A delay time of 3.5 s between scans was used. Data processing was carried out with either XWinNMR (Bruker) or NMRpipe.^[61] Spectra were analysed by using XEASY^[62] and Felix2000 (Biosym/MSI, CA, USA). R_2 and $R_{1\rho}$ relaxation times were determined from peak heights as described by Stone et al.^[63,64] by using programs provided by A. G. Palmer.^[65]

Acknowledgements

We thank Dr. Wolfgang Bermel, Bruker, Germany and Dr. Christian Richter for assistance with the $R_{1\rho}$ pulse sequence, and the European Union for the Marie Curie postdoctoral fellowship to E.S.C. and general support in the project, UPMAN.

Keywords: disulfide bonds · lysozyme · mutagenesis · NMR spectroscopy · protein folding

- [1] T. Szyperski, P. Luginbuhl, G. Otting, P. Guntert, K. Wüthrich, J. Biomol. NMR 1993, 3, 151.
- [2] G. Otting, E. Liepinsh, K. Wüthrich, Biochemistry 1993, 32, 3571.
- [3] C. Redfield, J. Boyd, L. J. Smith, R. A. Smith, C. M. Dobson, *Biochemistry* 1992, 31, 10431.
- [4] B. Donzel, B. Kamber, K. Wüthrich, R. Schwyzer, *Helv. Chim. Acta* 1972, 55, 947.
- [5] G. Jung, M. Ottnad, Angew. Chem. 1974, 86, 856; Angew. Chem. Int. Ed. 1974, 13, 818.
- [6] K. D. Kopple, Y.-S. Wang, A. G. Cheng, K. K. Bhandary, J. Am. Chem. Soc. 1988, 110, 4168.
- [7] B. J. Stockman, A. Euvrard, T. A. Scahill, J. Biomol. NMR 1993, 3, 285.
- [8] M. Jourdan, M. S. Searle, Biochemistry 2001, 40, 10317.
- [9] S. A. Gorski, C. S. Le Duff, A. P. Capaldi, A. P. Kalverda, G. S. Beddard, G. R. Moore, S. E. Radford, J. Mol. Biol. 2004, 337, 183.
- [10] H. Schwalbe, K. M. Fiebig, M. Buck, J. A. Jones, S. B. Grimshaw, A. Spencer, S. J. Glaser, L. J. Smith, C. M. Dobson, *Biochemistry* **1997**, *36*, 8977.
- [11] M. A. Lietzow, M. Jamin, H. J. Jane Dyson, P. E. Wright, J. Mol. Biol. 2002, 322, 655.
- [12] C. Redfield, B. A. Schulman, M. A. Milhollen, P. S. Kim, C. M. Dobson, *Nat. Struct. Biol.* **1999**, *6*, 948.
- [13] C. P. van Mierlo, N. J. Darby, J. Keeler, D. Neuhaus, T. E. Creighton, J. Mol. Biol. 1993, 229, 1125.
- [14] T. N. Niraula, T. Konno, H. Li, H. Yamada, K. Akasaka, H. Tachibana, Proc. Natl. Acad. Sci. USA 2004, 101, 4089.
- [15] C. P. van Mierlo, N. J. Darby, T. E. Creighton, Proc. Natl. Acad. Sci. USA 1992, 89, 6775.
- [16] W. J. Wedemeyer, E. Welker, M. Narayan, H. A. Scheraga, Biochemistry 2000, 39, 4207.
- [17] Y. Konishi, T. Ooi, H. A. Scheraga, Biochemistry 1982, 21, 4734.
- [18] D. M. Rothwarf, H. A. Scheraga, *Biochemistry* **1993**, *32*, 2680.
- [19] D. M. Rothwarf, Y. J. Li, H. A. Scheraga, Biochemistry 1998, 37, 3767.
- [20] D. M. Rothwarf, Y. J. Li, H. A. Scheraga, Biochemistry 1998, 37, 3760.
- [21] T. E. Creighton, Science 1992, 256, 111.
- [22] M. Dadlez, P. S. Kim, *Biochemistry* **1996**, *35*, 16153.
- [23] J. P. Staley, P. S. Kim, Protein Sci. 1994, 3, 1822.
- [24] W. L. Anderson, D. B. Wetlaufer, J. Biol. Chem. 1976, 251, 3147.
- [25] S. E. Radford, D. N. Woolfson, S. R. Martin, G. Lowe, C. M. Dobson, *Bio-chem. J.* **1991**, *273*, 211.
- [26] B. van den Berg, E. W. Chung, C. V. Robinson, C. M. Dobson, J. Mol. Biol. 1999, 290, 781.
- [27] V. Guez, P. Roux, A. Navon, M. E. Goldberg, Protein Sci. 2002, 11, 1136.
- [28] A. Land, D. Zonneveld, I. Braakman, FASEB J. 2003, 17, 1058.
- [29] H. Tachibana, T. Oka, K. Akasaka, J. Mol. Biol. 2001, 314, 311.
- [30] Y. Noda, A. Yokota, D. Horii, T. Tominaga, Y. Tanisaka, H. Tachibana, S. Segawa, *Biochemistry* 2002, 41, 2130.
- [31] C. C. Blake, D. F. Koenig, G. A. Mair, A. C. North, D. C. Phillips, V. R. Sarma, *Nature* **1965**, *206*, 757.

- [32] L. J. Smith, M. J. Sutcliffe, C. Redfield, C. M. Dobson, J. Mol. Biol. 1993, 229, 930.
- [33] H. Schwalbe, S. B. Grimshaw, A. Spencer, M. Buck, J. Boyd, C. M. Dobson, C. Redfield, L. J. Smith, *Protein Sci.* 2001, 10, 677.
- [34] J. Klein-Seetharaman, M. Oikawa, S. B. Grimshaw, J. Wirmer, E. Duchardt, T. Ueda, T. Imoto, L. J. Smith, C. M. Dobson, H. Schwalbe, *Science* 2002, 295, 1719.
- [35] J. Wirmer, C. Schlorb, J. Klein-Seetharaman, R. Hirano, T. Ueda, T. Imoto, H. Schwalbe, *Angew. Chem.* **2004**, *116*, 5904; *Angew. Chem. Int. Ed.* **2004**, *43*, 5780.
- [36] H. Tachibana, K. Ohta, H. Sawano, Y. Koumoto, S. Segawa, *Biochemistry* 1994, 33, 15008.
- [37] T. M. Logan, E. T. Olejniczak, R. X. Xu, S. W. Fesik, J. Biomol. NMR 1993, 3, 225.
- [38] G. Merutka, H. J. Dyson, P. E. Wright, J. Biomol. NMR 1995, 5, 14.
- [39] M. Tollinger, N. R. Skrynnikov, F. A. Mulder, J. D. Forman-Kay, L. E. Kay, J. Am. Chem. Soc. 2001, 123, 11341.
- [40] M. J. Grey, C. Wang, A. G. Palmer, 3rd, J. Am. Chem. Soc. 2003, 125, 14324.
- [41] D. M. Korzhnev, V. Y. Orekhov, F. W. Dahlquist, L. E. Kay, J. Biomol. NMR 2003, 26, 39.
- [42] M. Buck, J. Boyd, C. Redfield, D. A. MacKenzie, D. J. Jeenes, D. B. Archer, C. M. Dobson, *Biochemistry* 1995, 34, 4041.
- [43] J. Boyd, C. Redfield, J. Am. Chem. Soc. 1998, 120, 9692.
- [44] J. M. Thornton, J. Mol. Biol. 1981, 151, 261.
- [45] S. E. Radford, C. M. Dobson, P. A. Evans, Nature 1992, 358, 302.
- [46] T. Kiefhaber, Proc. Natl. Acad. Sci. USA 1995, 92, 9029.
- [47] C. M. Dobson, P. A. Evans, S. E. Radford, Trends Biochem. Sci. 1994, 19, 31.
- [48] S. K. Kulkarni, A. E. Ashcroft, M. Carey, D. Masselos, C. V. Robinson, S. E. Radford, Protein Sci. 1999, 8, 35.
- [49] D. C. Poland, H. A. Scheraga, J. Chem. Phys. 1965, 43, 2071.
- [50] S. H. Lin, Y. Konishi, M. E. Denton, H. A. Scheraga, *Biochemistry* 1984, 23, 5504.
- [51] C. N. Pace, G. R. Grimsley, J. A. Thomson, B. J. Barnett, J. Biol. Chem. 1988, 263, 11 820.
- [52] K. Yutani, A. Yutani, A. Imanishi, T. Isemura, J. Biochem. 1968, 64, 449.
- [53] V. P. Saxena, D. B. Wetlaufer, Biochemistry 1970, 9, 5015.
- [54] H. Tachibana, FEBS Lett. 2000, 480, 175.
- [55] A. Yokota, K. Hirai, H. Miyauchi, S. limura, Y. Noda, K. Inoue, K. Akasaka, H. Tachibana, S. I. Segawa, *Biochemistry* 2004, 43, 6663.
- [56] H. Sawano, Y. Koumoto, K. Ohta, Y. Sasaki, S. Segawa, H. Tachibana, *FEBS Lett.* **1992**, *303*, 11.
- [57] L. E. Kay, L. K. Nicholson, F. Delaglio, A. Bax, D. A. Torchia, J. Magn. Reson. 1992, 97, 359.
- [58] N. A. Farrow, R. Muhandiram, A. U. Singer, S. M. Pascal, C. M. Kay, G. Gish, S. E. Shoelson, T. Pawson, J. D. Forman-Kay, L. E. Kay, *Biochemistry* 1994, 33, 5984.
- [59] D. M. Korzhnev, N. R. Skrynnikov, O. Millet, D. A. Torchia, L. E. Kay, J. Am. Chem. Soc. 2002, 124, 10743.
- [60] F. A. A. Mulder, R. A. de Graaf, R. Kaptein, R. Boelens, J. Magn. Reson. 1998, 131, 351.
- [61] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, A. Bax, J. Biomol. NMR 1995, 6, 277.
- [62] C. Bartels, T. H. Xia, M. Billeter, P. Guntert, K. Wüthrich, J. Biomol. NMR 1995, 6, 1.
- [63] M. J. Stone, W. J. Fairbrother, A. G. Palmer, 3rd, J. Reizer, M. H. Saier, Jr., P. E. Wright, *Biochemistry* **1992**, *31*, 4394.
- [64] M. J. Stone, K. Chandrasekhar, A. Holmgren, P. E. Wright, H. J. Dyson, *Biochemistry* 1993, 32, 426.
- [65] http://cpmcnet.columbia.edu/dept/gsas/biochem/labs/palmer/.

Received: May 10, 2005