Journal of Biomolecular NMR (2005) 33: 95–104
DOI 10.1007/s10858-005-2063-y

Article

Heterologous expression of hen egg white lysozyme and resonance assignment of tryptophan side chains in its non-native states

Christian Schlörb[a], Katrin Ackermann[a,b], Christian Richter[a], Julia Wirmer[a] & Harald Schwalbe[a,*]

[a]*Institute for Organic Chemistry and Chemical Biology, Center for Biomolecular Magnetic Resonance, Johann Wolfgang Goethe-University Frankfurt, Marie-Curie-Strasse 11, D-60439, Frankfurt, Germany;* [b]*Present Address: Dr. Senckenbergische Anatomie, Anatomie III, Johann Wolfgang Goethe-University Hospital, Theodor-Stern-Kai 7, D-60590, Frankfurt, Germany*

Received 4 May 2005; Accepted 5 August 2005

Key words: inclusion bodies, isotope enrichment, lysozyme, non-native proteins, resonance assignment, tryptophan side chains

Abstract

A new protocol is described for the isotope (^{15}N and ^{13}C,^{15}N) enrichment of hen egg white lysozyme. Hen egg white lysozyme and an all-Ala-mutant of this protein have been expressed in *E. coli*. They formed inclusion bodies from which mg quantities of the proteins were purified and prepared for NMR spectroscopic investigations. ^{1}H,^{13}C and ^{15}N main chain resonances of disulfide reduced and S-methylated lysozyme were assigned and its residual structure in water pH 2 was characterized by chemical shift perturbation analysis. A new NMR experiment has been developed to assign tryptophan side chain indole resonances by correlation of side chain and backbone NH resonances with the C^{γ} resonances of these residues. Assignment of tryptophan side chains enables further residue specific investigations on structural and dynamical properties, which are of significant interest for the understanding of non-natives states of lysozyme stabilized by hydrophobic interactions between clusters of tryptophan residues.

Introduction

The native state and different non-native states of hen egg white lysozyme have been investigated extensively by nuclear magnetic resonance spectroscopy. Redfield et al. provided the first resonance assignments for a sizeable protein using homonuclear correlation spectroscopy (Redfield and Dobson, 1988). Detailed structural investigations including the first solution structure of the protein were reported by Redfield and Dobson (1988) and Smith et al. (1991, 1993). Buck et al. analyzed the dynamics of native hen egg white lysozyme by measuring T_1 and T_2 relaxation times of ^{15}N main chain and side chain nuclei (Buck et al., 1995). The structural and dynamical properties of denatured hen lysozyme were investigated using ^{15}N-resolved 3D NMR experiments. Coupling constants, ^{15}N relaxation rates, NOE data and chemical shift deviations from random coil chemical shifts of both oxidized and reduced lysozyme in 8 M urea together led to a better understanding of residual structure (Schwalbe et al., 1997) and side chain conformations (Hennig et al., 1999) in non-native states of this protein. A partially folded state of lysozyme in a trifluorethanol/water mixture was characterized by investigation of NOE patterns, H^{α} chemical shift perturbations, $^{3}J(H^{N},H^{\alpha})$-coupling constants and protection from hydrogen exchange of a ^{15}N labeled sample of this protein (Buck et al.,

*To whom correspondence should be addressed. E-mail: schwalbe@nmr.uni-frankfurt.de

1996). Recently, it could also been shown that long-range interactions and extensive hydrophobic clusters are present in lysozyme even under denaturing conditions (Klein-Seetharaman et al., 2002; Wirmer et al., 2004). These conclusions have been derived from the analysis of chemical shift perturbations, transverse relaxation rates, and diffusion constants of ^{15}N labeled wild type lysozyme and respective single point mutants (Wirmer et al., 2004).

However, the previous relaxation studies on non-native lysozyme solely relied on information based on the backbone assignments. Assignment of tryptophan side chain resonances in addition to the backbone resonances allows gaining valuable site-specific information on structural and dynamical properties of these residues. This information is particularly interesting when looking at non-native and native-like interactions in non-native states of proteins, where tryptophans play an important role in mediating such contacts and the formation of hydrophobic clusters (Klein-Seetharaman et al., 2002). The indole proton and nitrogen resonances in the tryptophan side chain are especially valuable for structural and dynamical analysis as they are reasonably resolved and eligible for heteronuclear relaxation and photo-CIDNP methods (Mok and Hore, 2004).

Heterologous expression of recombinant proteins is normally a prerequisite for isotope labeling of proteins for nuclear magnetic resonance spectroscopy. For the isotope labeling of recombinant hen egg white lysozyme (HEWL), several protocols have been published and were used in the previous NMR studies of this protein. Expression of ^{15}N and ^{13}C,^{15}N labeled HEWL in the filamentous fungus *Aspergillus niger* (MacKenzie et al., 1996; Spencer et al., 1999) was used for the refinement of the solution structure of hen egg white lysozyme (Schwalbe et al., 2001) and for conformational NMR studies on non-native lysozyme (Buck et al., 1995, 1996; Schwalbe et al., 1997; Hennig et al., 1999). Heterologous expression of ^{15}N isotope labeled HEWL has also been reported for the methylotropic yeast *Pichia pastoris* (Mine et al., 1999). However, the usage of eukaryotic expression systems such as *P. pastoris* and *A. niger*, is time consuming and expensive. In contrast, heterologous expression of proteins in *E. coli* usually is less time consuming, easier to implement, and isotope enrichment is less cost intensive due to a lower need for labeled media. However, for most of the recombinant proteins expressed in *E. coli*, the N-terminal methionine is not eliminated post-translationally unlike in the eukaryotic expression systems mentioned above.

Investigations on non-native states of HEWL can either be performed under denaturing conditions as mentioned above or by a permanent destruction of structure-stabilizing disulfide-bridges in the molecule. This can be achieved by reduction and permanent blocking of the cysteine residues or by means of site-specific mutation. Both approaches have been applied to HEWL in this study. The eight cysteines forming the four disulfide-bridges in native HEWL have been reduced and methylated gaining permanent non-native states, designated HEWL-S^{Me}. The mutational approach to obtain non-native HEWL without the necessity of any chemical modification or presence of chaotropic agents involves the replacement of all cysteine residues with alanines. The result of this is a very similar state as HEWL-S^{Me} and is referred to as all-Ala-HEWL.

In this paper, we report on the optimization of an expression protocol to produce HEWL from *E. coli*. ^{13}C,^{15}N labeled protein was used to develop a set of two new NMR experiments to assign the side chain resonances of tryptophan residues. These usually cannot be assigned from ^{15}N-resolved 3D data of only ^{15}N labeled proteins when investigated in their non-native states.

A previously published experiment for the sequence-specific assignment of tryptophan ring ^{1}H,^{13}C and ^{15}N resonances in proteins (Löhr et al., 2002) could not be applied in our case. The described HN(CDCG)CB experiment correlates the indole protons in tryptophan side chains with $^{13}C^{\beta}$ resonances in a series of magnetization transfers via one-bond scalar couplings. In the case of non-native proteins like the all-Ala mutant of hen egg white lysozyme, the $^{13}C^{\beta}$ resonances are not well enough resolved to allow for the unambiguous assignment of the tryptophan side chain resonances.

The use of combinations of conventional tryptophan side chain specific TOCSY and NOESY spectra (Slupsky et al., 1998) fail in unfolded proteins due to the massive overlap of resonances in the aromatic side chain. In such a strategy, the $^{13}C^{\beta}$ resonances in tryptophan side chains are correlated with the aromatic $^{1}H^{\delta 1}$

resonances via a selective (HB)CB(CGCD)HD experiment (Yamazaki and Kay, 1993). The $^{1}H^{\delta 1}$ resonances are then correlated to the $^{15}N^{\varepsilon}$ and $^{1}H^{N\varepsilon}$ resonances by exploiting their scalar and NOE couplings to the $^{1}H^{\delta 1}$ nucleus.

Since the existing experiments for the assignment of tryptophan side chain resonances in native proteins are not suitable in the case of some non-native proteins due to the resonance overlap, a set of two new NMR experiments has been developed. In these experiments, the tryptophan side chain indole resonances are assigned by the correlation of side chain and backbone NH resonances with the C^{γ} resonances of these residues. The C^{γ} resonances exhibit a reasonable resolution even in non-native HEWL, which is a prerequisite for the unambiguous assignment.

Materials and methods

Cloning, expression and purification

The synthetic genes encoding hen egg white lysozyme (HEWL) with an N-terminal methionine and its respective all-Ala mutant with all cysteine residues replaced by alanines, were obtained from Entelechon, Regensburg, Germany in vector pCR4-TOPO, and cloned into the pET11a expression vector. The uniformly ^{15}N,^{13}C-labeled proteins were overproduced in *E. coli* strain BL21(DE3) using M9 minimal medium containing $^{15}NH_4Cl$ (Martek) and $^{13}C_6$-glucose (Silantes). Cultures were induced with 2 mM IPTG at an OD_{600} of ~0.8, and grown for an additional 3 h before harvesting. Cells were lysed by sonication in buffer containing 50 mM Tris, 25% sucrose, and 1 mM EDTA, pH 7.5, the insoluble fraction was washed with the same buffer and with buffer containing 20 mM Tris, 1% Triton X-100, and 1 mM EDTA, pH 7.5. The resulting pellets (inclusion bodies) were solubilized in 20 mM Tris, 50 mM NaCl, 8 M urea, 5 mM EDTA, and 0.1 M DTT, pH 7.5, and loaded onto a fast-flow CM-Sepharose (Sigma) column equilibrated with 50 mM Tris, 50 mM NaCl, 4 M urea, 1 mM EDTA, and 5 mM 2-mercaptoethanol. The column was eluted with a linear gradient of NaCl (final concentration 300 mM) in 4 M urea. Fractions containing HEWL or all-Ala-HEWL were pooled, concentrated to a final concentration of about 1 mg/ml using centrifugal filter devices (Vivaspin 15, 5000 MWCO, Vivascience). In the case of HEWL, the cysteine groups were methylated as described before (Heinrikson, 1971) in buffer containing 8 M urea and 25% (v/v) acetonitrile. Reduced S-methylated HEWL was dialysed against deionized water pH 2, purified by reversed-phase HPLC using a linear water-to-acetonitrile gradient, and freeze-dried. All-Ala-HEWL was dialyzed against water pH 2, purified by reversed phase HPLC, and freeze dried. After solvation of the dried protein in deionized water pH 2 containing 10% (v/v) D_2O, the NMR samples had protein concentrations of approximately 0.5 mM.

The synthetic gene encoding for hen egg white lysozyme was also cloned into the expression vector pJC40 to introduce an N-terminal deca-his tag. Expression and purification of the his-tagged protein followed the same protocol as described above with an additional immobilized-metal affinity chromatography (IMAC) using a Ni-NTA matrix before the ion-exchange column. Protein was eluted from the Ni-NTA column using a step gradient of 10–500 mM imidazol in 50 mM Tris, 50 mM NaCl, 4 M urea, and 20 mM 2-mercaptoethanol, pH 7.5. Vectors can be obtained upon request from the authors.

Refolding

After ion exchange chromatography purification, urea was added to a final concentration of 8 M and the protein was concentrated 3 mg/ml using centrifugal filter devices. Subsequently, HEWL was refolded by rapid 20-fold dilution into buffer containing 50 mM Tris, 50 mM KCl, 1 mM ED-TA, 5 mM reduced glutathion, and 0.5 mM oxidized glutathione. After dilution the solution containing the refolded HEWL was dialyzed twice for 12 h against deionized water pH 6.8 and freeze-dried. Freeze-dried protein was dissolved in water pH 3.8 containing 10% D_2O to a concentration of about 15 μM. The yield of pure and refolded lysozyme was approximately 50%.

Assignment of the backbone

NMR experiments were carried out at 293 K on Bruker AVANCE spectrometers, operating at proton frequencies of 700 or 800 MHz and

equipped with 5 mm triple-resonance pulsed-field z-gradient probes. The NMR-data were processed using XWIN-NMR 3.5 (Bruker Biospin) and analyzed with XEASY (Bartels et al., 1995) and CARA (Keller, 2004). The sequential backbone resonance assignment of the reduced S-methylated hen egg white lysozyme in deionized water pH 2 followed the standard triple resonance strategy for $^{13}C,^{15}N$-labeled proteins (Sattler et al., 1999). The assignments were obtained with HNCACB (Wittekind and Mueller, 1993), CBCA(CO)NH (Grzesiek and Bax, 1993), HNCO (Grzesiek and Bax, 1992), HN(CA)CO (Clubb et al., 1992), HBHA(CO)NH (Grzesiek and Bax, 1993) and (H)CC(CO)NH (Montelione et al., 1992) experiments. ^{1}H chemical shifts were referenced to TMSP at 0.00 ppm and ^{13}C and ^{15}N chemical shifts were calculated from the ^{1}H frequency (Wishart et al., 1995). A $^{1}H,^{15}N$-HSQC-spectrum of refolded lysozyme was recorded at 293 K on a Bruker AVANCE-800 spectrometer equipped with a $^{1}H,^{13}C,^{15}N$ triple-resonance cryoprobe. The amide ^{1}H and ^{15}N resonances of the tryptophan residues of the all-Ala-HEWL mutant, which are necessary for the side chain assignment, have been assigned by comparison of the $^{1}H,^{15}N$ HSQC spectra of HEWL-S^{Me} and all-Ala-HEWL.

Assignment of the tryptophan side chains

Tryptophan side chain indole $^{1}H^{N\varepsilon}/^{15}N^{\varepsilon}$ resonances have been assigned by the newly developed HN(CACB)CG and HN(CD)CG experiments with the uniformly $^{13}C,^{15}N$ labeled all-Ala-HEWL sample on a Bruker AVANCE-700 spectrometer equipped with a $^{1}H,^{13}C,^{15}N$ triple-resonance cryoprobe. Both experiments have been run as two-dimensional versions by omitting the ^{15}N chemical shift evolution. The HN(CACB)CG experiment was performed with spectral widths of 10 ppm ($^{13}C^{\gamma}$) and 14 ppm ($^{1}H^{N}$), 64 scans per increment and 128 ($^{13}C^{\gamma}$) times 512 ($^{1}H^{N}$) complex points, resulting in a total experiment time of 8.5 h. The HN(CD)CG experiment was recorded in 2 h with 16 scans and 128 ($^{13}C^{\gamma}$) times 512 ($^{1}H^{N}$) complex points. Spectral widths for the $^{13}C^{\gamma}$ and $^{1}H^{N\varepsilon}$ were set to 3 and 14 ppm, respectively. The three-dimensional version of the HN(CD)CG experiment was recorded in 60 h with 16 scans and 128 ($^{1}H^{N\varepsilon}$) times 128 ($^{13}C^{\gamma}$) times 512 ($^{1}H^{N}$) complex points.

The spectra were processed using the Bruker TOPSPIN 1.3 software. Linear prediction as implemented in the Bruker software package was applied in the case of the indirect domain of the HN(CACB)CG spectrum.

Results and discussion

Expression, purification and refolding

Expression of hen egg white lysozyme and its respective all-Ala-mutant in *E. coli* yielded up to 25 mg $^{13}C,^{15}N$ doubly labeled lysozyme per liter M9 minimal medium. Most of the protein was produced in the first hour after induction. The yield for the ^{15}N labeled proteins was about the same as for the doubly labeled. The bulk amount of the overproduced lysozyme went into inclusion bodies as shown in the SDS-PAGE analysis of the different steps of the solubilization process. After purification with CM sepharose ion exchange chromatography, the protein was already more than 95% pure as judged from SDS-PAGE analysis. HEWL and all-Ala-HEWL eluted from the ion exchange column at a salt concentration of around 150 mM. After S-methylation of disulfide-reduced HEWL and HPLC purification of all-Ala-HEWL and the methylated HEWL, both proteins were pure as judged from MALDI-TOF mass spectrometry.

The yield of native lysozyme after refolding by rapid dilution into refolding buffer was about 50%, the remainder precipitated due to misfolding or was lost during HPLC purification. The recorded $^{1}H,^{15}N$-HSQC spectrum of the refolded lysozyme at pH 3.8 and 293 K (Figure 1) is identically to a previously published spectrum of native lysozyme expressed in *Aspergillus niger* at 308 K and the same pH (Buck et al., 1995).

Backbone resonance assignment

The resonance assignment of H^{N}, N, H^{α}, H^{β}, C^{α}, C^{β} and C' nuclei for reduced, S-methylated HEWL using the standard heteronuclear two- and three-dimensional NMR experiments was straightforward. Some of the H^{N} and N^{H} resonances, especially the glycines, serines and threonines, were easily assigned by comparison with existing $^{1}H,^{15}N$-HSQC spectra for S-methylated

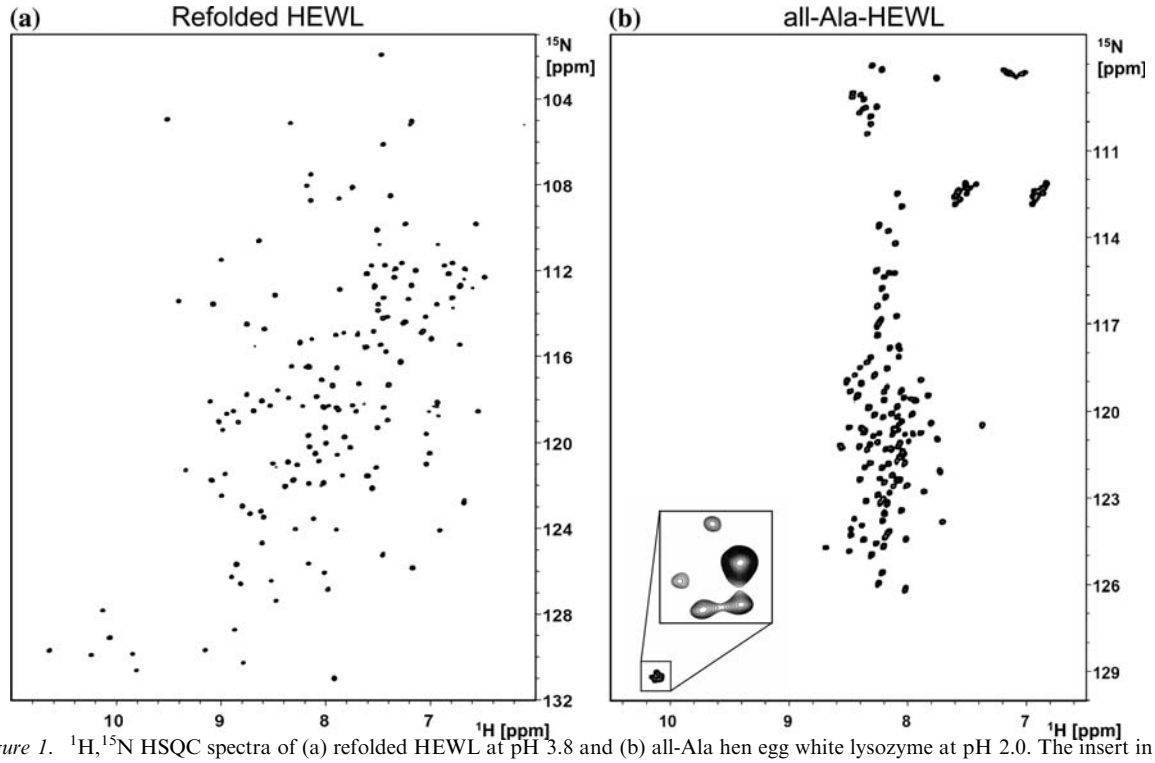


Figure 1. $^{1}H,^{15}N$ HSQC spectra of (a) refolded HEWL at pH 3.8 and (b) all-Ala hen egg white lysozyme at pH 2.0. The insert in (b) shows the zoomed tryptophan side chain indole region. Both spectra have been acquired at a temperature of 293 K. Note that the scales for both spectra are different.

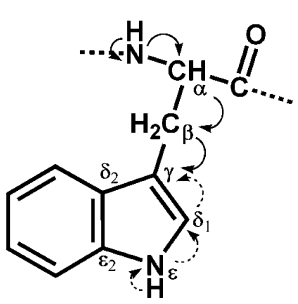
HEWL in water pH 2 (Grimshaw, 1999) and HEWL in 8 M urea (Schwalbe et al., 1997; Hennig et al., 1999). These initial assignments formed the basis for the sequential resonance assignments using the CBCA(CO)NH and HNCACB spectra. (H)CC (CO)NH spectra were used to identify the amino acid types of the residues. For the assignment of the carbonyl carbon nuclei HN(CA)CO and HNCO spectra were used.

For 112 out of the 130 residues of HEWL-S^{Me} the H^{N}, N^{H}, H^{α}, H^{β}, C^{α}, C^{β} and C′ could be assigned completely, for 10 additional residues partial assignments have been found. The chemical shifts were stored in the BMRB database under accession code 6622.

Tryptophan side chain resonance assignment

The experiments for the assignment of indole resonances in tryptophan side chains of non-native proteins depend on the selectivity of pulses and transfers within the tryptophan side chains (Figure 2).

The HN(CACB)CG experiment (Figure 3) correlates the backbone amide $^{1}H^{N}/^{15}N$ resonances with the $^{13}C^{\gamma}$ resonances in the side chain and the HN(CD)CG experiment (Figure 4) correlates the side chain indole $^{1}H^{N\varepsilon}/^{15}N^{\varepsilon}$ resonances with the $^{13}C^{\gamma}$ resonances (Figure 5). This



H^{N}_{bb}	7.2 – 8.0 ppm
C_{α}	~ 55 ppm
C_{β}	~ 27 ppm
C_{γ}	~ 110 ppm
$C_{\delta 1}$	~ 127 ppm
$C_{\delta 2}$	~ 129 ppm
$C_{\varepsilon 2}$	~ 138 ppm
H^{N}_{sc}	9.9 - 10.0 ppm

Figure 2. Magnetization transfer pathway for the assignment of tryptophan side chain indole protons in the HN(CACB)CG experiment (full arrows) and the HN(CD)CG experiment (dashed arrows). Typical chemical shifts in tryptophan residues of non-native lysozyme as found during the development of the two assignment experiments in this work are shown on the right.

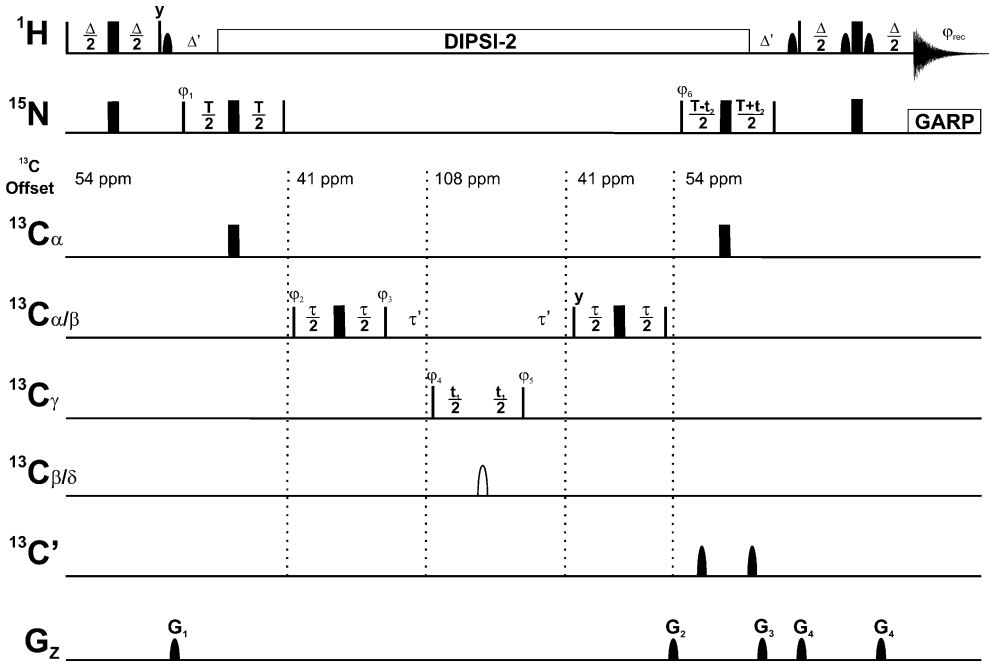


Figure 3. Pulse sequence of the 3D-HN(CACB)CG experiment for the correlation of the backbone amide proton with the aromatic $^{13}C^{\gamma}$ resonances in tryptophan side chains. Narrow and wide filled bars correspond to rectangular 90° and 180° pulses applied with RF field strengths of 23.6 kHz (^{1}H) and 6.8 kHz (^{15}N), respectively. RF field strengths of 12.9 kHz and 28.7 kHz on ^{13}C resonances are used for the standard Q_3 and Q_5 Gaussian cascades (Emsley and Bodenhausen, 1992). Selective pulses and gradients are indicated by semi-ellipses. The default pulse phase is x. The pulse sequences were optimized on a Bruker spectrometer with the Bruker typical phase nomenclature. Fixed delays are adjusted as follows: $\Delta = 4.6$ ms ($1/(2*^{1}J_{NH})$), $\Delta' = 4.9$ ms, $T = 24.8$ ms ($1/(4*^{1}J_{CN})$), $\tau = 7.2$ ms ($1/(4*^{1}J_{CC})$), $\tau' = 8.5$ ms ($1/(2*^{1}J_{CC})$). Proton and nitrogen carrier frequencies are centered at the water (4.7 ppm) and the amide ^{15}N region (118 ppm), respectively. The carbon carrier frequency changes during the course of the experiment as indicated by vertical dashed lines and the value of the ^{13}C offset. The shaped 180° decoupling pulses on carbonyl carbon resonances are Q_3 Gaussian cascades with durations of 2 ms and have offsets of 172 ppm. The shaped 180° decoupling pulse on the C^{β} and C^{δ} carbon resonances during the carbon t_1 time is based on a Q_3 Gaussian cascade and has duration of 1.7 ms. About 1 ms water flipback square pulses are applied after the first INEPT step and during the backtransfer from ^{15}N to ^{1}H. Asynchronous GARP decoupling (Shaka et al., 1985) is used to suppress ^{15}N–^{1}H heteronuclear scalar coupling during acquisition. Proton decoupling using the DIPSI-2 (Shaka et al., 1988) composite pulse decoupling is applied during most of the pulse sequence. The pulsed field gradients of 1 ms length are sine-bell shaped, applied along the z-axis and have the following strengths: $G_1 = 27.5$ G cm^{-1}, $G_2 = 22$ G cm^{-1}, $G_3 = 33$ G cm^{-1}, $G_4 = 16.5$ G cm^{-1}. Phase cycling is: $\phi_1 = 4(x), 4(-x)$; $\phi_2 = 16(x), 16(-x)$; $\phi_3 = y, -y$; $\phi_4 = 8(x), 8(-x)$; $\phi_5 = 32(x), 32(-x)$; $\phi_6 = 2(x), 2(-x)$; $\phi_{rec} = R, 2(-R), R, -R, 2R, -R$, where $R = 2(x), 4(-x), 2(x)$. In addition, ϕ_4 is incremented in a States-TPPI (Marion et al., 1989) manner to achieve quadrature detection in the ω_1 direction.

assignment strategy exploits the fact that the $^{13}C^{\gamma}$ resonances are comparably well resolved in non-native states of proteins and that they are well separated from other aromatic carbon resonances.

The $^{13}C^{\gamma}$ resonances have been assigned in the HN(CACB)CG experiment; from this correlation the $^{1}H^{N\varepsilon}/^{15}N^{\varepsilon}$ resonances have been assigned using the HN(CD)CG experiment. Both experiments are of the 'out-and-back'-type and have been derived from the HNCACB experiment (Wittekind and Mueller, 1993). All delay durations have been optimized to selectively observe correlations within tryptophan side chains. Solvent is suppressed by soft-WATERGATE pulse schemes (Piotto et al., 1992). The HN(CACB)CG pulse sequence comprises four successive steps to transfer magnetization via ^{1}J one-bond couplings from the $^{1}H^{N}$ to the $^{13}C^{\gamma}$. Amplitude modulated shaped pulses are used for the inter-carbon transfers and the decoupling of the $^{1}J_{C\gamma C\delta 1}$ and $^{1}J_{C\gamma C\delta 2}$ couplings during t_1 and to refocus C^{β} chemical shifts during $2\tau' + t_1$. The delay τ' for the magnetization transfer from the $^{13}C^{\beta}$ to the $^{13}C^{\gamma}$ is adjusted to approximately $1/(2^{1}J_{C\beta C\gamma})$, conducting a HMQC scheme. To gain maximal transfer efficiency the $^{13}C^{\alpha}$–$^{13}C^{\beta}$ transfer delay is set to $1/(2^{1}J_{C\alpha C\beta})$, unlike in a standard HNCACB pulse sequence, where this delay duration is adjusted to $1/(4^{1}J_{C\alpha C\beta})$.

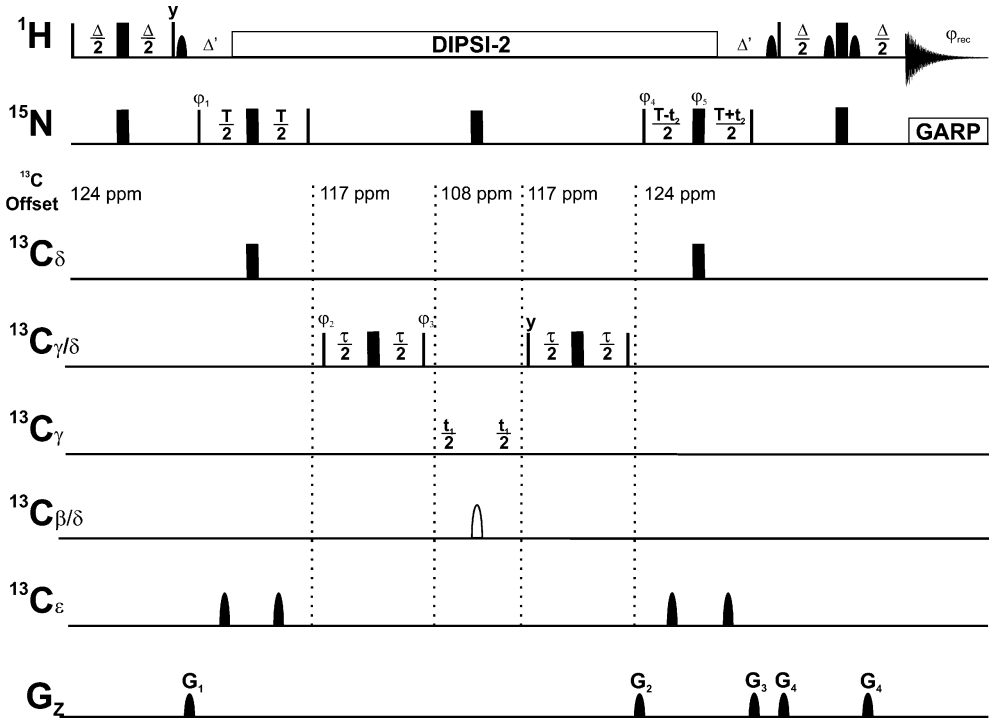


Figure 4. Pulse sequence of the 3D-HN(CD)CG experiment for the correlation of the side chain indole proton resonances with the aromatic $^{13}C^{\gamma}$ resonances in tryptophan side chains. Narrow and wide filled bars correspond to rectangular 90° and 180° pulses applied with RF field strengths of 23.6 kHz (^{1}H) and 6.8 kHz (^{15}N), respectively. RF field strengths of 12.9 kHz and 28.7 kHz on ^{13}C resonances are used for the standard Q_3 and Q_5 Gaussian cascades (Emsley and Bodenhausen, 1992). Selective pulses and gradients are indicated by semi-ellipses. The default pulse phase is x. The pulse sequences were optimized on a Bruker spectrometer with the Bruker typical phase nomenclature. Fixed delays are adjusted as follows: $\Delta = 4.6$ ms ($1/(2*^{1}J_{NH})$), $\Delta' = 4.9$ ms, $T = 27.4$ ms ($1/(4*^{1}J_{CN})$), $\tau = 6.4$ ms ($1/(2*^{1}J_{CC})$). Proton and nitrogen carrier frequencies are centered at the water (4.7 ppm) and the indole ^{15}N region (129.2 ppm), respectively. The carbon carrier frequency changes during the course of the experiment as indicated by vertical dashed lines and the value of the ^{13}C offset. The shaped 180° decoupling pulses on aromatic C^{ε} resonances are Q_3 Gaussian cascades with durations of 2 ms and have offsets of 141 ppm. The shaped 180° decoupling pulse on the C^{β} and C^{δ} carbon resonances during the carbon t_1 time is based on a Q_3 Gaussian cascade and has duration of 1.15 ms. 1 ms water flipback square pulses are applied after the first INEPT step and during the backtransfer from ^{15}N to ^{1}H. Asynchronous GARP decoupling (Shaka et al., 1985) is used to suppress ^{15}N–^{1}H heteronuclear scalar coupling during acquisition. Proton decoupling using the DIPSI-2 (Shaka et al., 1988) composite pulse decoupling is applied during most of the pulse sequence. The pulsed field gradients of 1 ms length are sine-bell shaped, applied along the z-axis and have the following strengths: $G_1 = 27.5$ G cm^{-1}, $G_2 = 22$ G cm^{-1}, $G_3 = 33$ G cm^{-1}, $G_4 = 16.5$ G cm^{-1}. Phase cycling is: $\phi_1 = 8(x), 8(-x)$; $\phi_2 = x$; $\phi_3 = y, -y$; $\phi_4 = 2(x), 2(-x)$; $\phi_5 = 4(x), 4(-x)$; $\phi_{rec} = 2R, 2(-R)$, where $R = 2(x), 2(-x)$. In addition, ϕ_2 and ϕ_3 are incremented in a States-TPPI (Marion et al., 1989) manner to achieve quadrature detection in the ω_1 direction.

The HN(CD)CG experiment transfers magnetization from the $^{1}H^{N\varepsilon}$ to the $^{13}C^{\gamma}$ in three successive steps: The initial INEPT transfer step from the indole proton to the nitrogen is followed by the transfer to the $^{13}C^{\delta 1}$ nucleus and the successive transfer to the $^{13}C^{\gamma}$. The pulses for the magnetization transfer between $^{13}C^{\delta 1}$ and the $^{13}C^{\gamma}$ are selective as well as the decoupling pulse on β and δ_1/δ_2 carbons during t_1. The offset for this transfer is on the C^{γ} and all pulses are off resonance on the $C^{\gamma}/C^{\delta 1}$ chemical shift region (117 ppm). The introduction of selective pulses on $^{13}C^{\delta}$ and $^{13}C^{\beta}$ during t_1 allows to obtain a high resolution in the carbon dimension, which is essential for the unambiguous assignment in non-native proteins. This an important advantage compared to the previously existing TROSY-HN(CDCG)CB experiment (Löhr et al., 2002).

Figure 5 demonstrates the result of assignment approach for the tryptophan side chain indole $^{1}H^{N\varepsilon}/^{15}N^{\varepsilon}$ resonances using HN(CACB)CG and HN(CD)CG experiments.

The resonances in the backbone amide region of a ^{1}H,^{15}N HSQC correlation spectrum are cor-

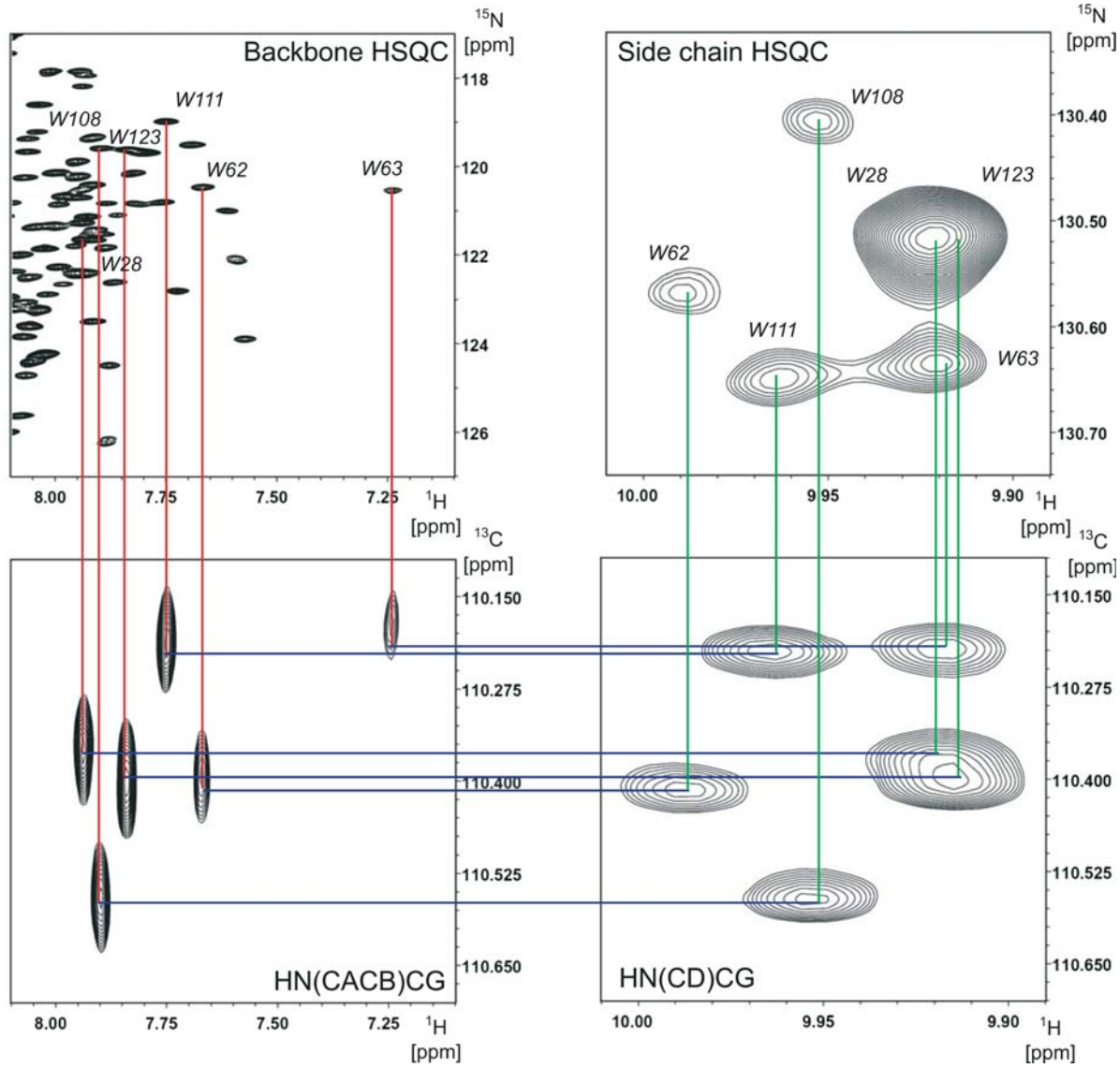


Figure 5. Results of the assignment strategy for the tryptophan side chain indole $^{1}H^{N}/^{15}N^{\varepsilon}$ resonances by correlation with the aromatic $^{13}C^{\gamma}$ resonances. The tryptophan peaks in the $^{1}H,^{15}N$ HSQC spectrum of the backbone amide region (upper left) is correlated with the $^{1}H,^{13}C$ plane of the HN(CACB)CG spectrum (lower left) to assign $^{13}C^{\gamma}$ resonances of the tryptophan side chains. $^{13}C^{\gamma}$ resonances in the $^{1}H,^{13}C$ planes of the HN(CACB)CG and the HN(CD)CG spectrum are correlated via their $^{13}C^{\gamma}$ chemical shifts. The resonances in the HN(CD)CG spectrum are correlated to the $^{1}H,^{15}N$ HSQC spectrum via their ^{1}H chemical shifts.

related to the peaks in the $^{1}H^{N}/^{13}C^{\gamma}$ plane of the HN(CACB)CG spectrum by the distinct $^{1}H^{N}$ resonance frequencies. The HN(CACB)CG and HN(CD)CG spectra are connected via the $^{13}C^{\gamma}$ chemical shifts. From the distinct $^{1}H^{N\varepsilon}$ chemical shifts in the $^{1}H^{N\varepsilon}/^{13}C^{\gamma}$ plane of the HN(CD)CG spectrum, the $^{1}H^{N\varepsilon}/^{15}N^{\varepsilon}$ resonances in a $^{1}H,^{15}N$ HSQC correlation spectrum of the tryptophan side chain indole region have been assigned. In the HN(CD)CG spectrum and the indole region of the $^{1}H,^{15}N$ HSQC spectrum the resonances for the tryptophans 28 and 123 of hen egg white lysozyme overlap significantly, but the other four tryptophans are clearly separated and unambiguously assignable. The peaks of the tryptophans 62 and 108 are considerably broadened compared to the other peaks. This coincides with their crucial role in hydrophobic clustering in non-native

lysozyme (Klein-Seetharaman et al., 2002; Wirmer et al., 2004). The assignment in the HN(CD)CG spectrum has been double-checked by letting the experiment run in the three-dimensional version with the ^{15}N chemical shifts evolved.

Conclusion

A novel expression system for uniformly isotope labeled hen egg white lysozyme and mutants thereof has been established. Production of the isotope enriched protein is possible with a relatively high yield of more than 20 mg per liter culture medium. Expression and purification are straight forward and fast. Refolding of the purified protein is possible and pure fractions of native protein can be obtained. However, the yield of refolding is in the same range as for other published methods (Fischer, 1996; Hevehan and De Bernadez Clark, 1997).

This protocol enables NMR studies on the structural and dynamical properties of both native and non-native states of hen lysozyme. In addition, studies of the folding pathway of this protein are also possible.

The introduction of an all-Ala mutant of hen egg white lysozyme, in which all of the cysteine residues are replaced by alanines, provides a fast and simple way of obtaining lysozyme in a permanent non-native state. The absence of cysteine residues prohibits the formation of a branched polypeptide chain via disulfide bridges like in native hen egg white lysozyme. All-Ala-HEWL in water pH 2 resembles the non-native states of reduced and S-methylated wild-type HEWL under the same conditions, but is easier to prepare and is not subject of the disadvantages of the methylation process, such as incomplete reduction or methylation or unwanted methylation of residues other than cysteine.

Properties of the all-Ala mutant and the S-methylated wild-type are very similar as judged from chemical shifts and relaxation analysis.

The new experiments for the assignment of tryptophan side chain indole protons allow new insights in structural and dynamical properties of non-native proteins, such as all-Ala-HEWL, since many important interactions in native and non-native proteins occur on the level of the side chains. Residue specific information on the solvent accessibility of tryptophan side chains as obtained by photo-CIDNP experiments could yield important insights into residual structure and hydrophobic clustering. Relaxation studies of single tryptophan indole resonances would allow for information on dynamical properties of these side chains and the comparison with the backbone data.

Acknowledgements

We thank Jens Wöhnert for fruitful discussions. The support for CS by the Fonds der Chemischen Industrie is gratefully acknowledged. This work was supported in part by the EU project UPMAN, the BMRZ is supported by the state of Hesse.

References

Bartels, C., Xia, T.H., Billeter, M., Güntert, P. and Wüthrich, K. (1995) *J. Biomol. NMR*, **5**, 1–10.
Buck, M., Boyd, J., Redfield, C., MacKenzie, D.A., Jeenes, D.J., Archer, D.B. and Dobson, C.M. (1995) *Biochemistry*, **34**, 4041–4055.
Buck, M., Schwalbe, H. and Dobson, C.M. (1996) *J. Mol. Biol.*, **257**, 669–683.
Clubb, R.T., Thanabal, V. and Wagner, G. (1992) *J. Magn. Reson.*, **97**, 213–217.
Emsley, L. and Bodenhausen, G. (1992) *J. Magn. Reson.*, **97**, 135–148.
Fischer, B. (1996) *Experientia Suppl.*, **75**, 143–161.
Grimshaw, S. B. (1999) PhD Thesis, University of Oxford, United Kingdom.
Grzesiek, S. and Bax, A. (1992) *J. Am. Chem. Soc.*, **114**, 6291–6293.
Grzesiek, S. and Bax, A. (1993) *J. Biomol. NMR*, **3**, 185–204.
Heinrikson, R.L. (1971) *J. Biol. Chem.*, **246**, 4090–4096.
Hennig, M., Bermel, W., Dobson, C.M., Smith, L.J. and Schwalbe, H. (1999) *J. Mol. Biol.*, **288**, 705–723.
Hevehan, D.L. and De Bernadez Clark, E. (1997) *Biotech. Bioeng.*, **54**, 221–230.
Keller, R. (2004) *The Computer Aided Resonance Assignment Tutorial* CANTINA Verlag, Goldau.
Klein-Seetharaman, J., Oikawa, M., Grimshaw, S.B., Wirmer, J., Duchardt, E., Ueda, T., Imoto, T., Smith, L.J., Dobson, C.M., and Schwalbe, H. (2002) *Science*, **295**, 1719–1722.
Löhr, F., Katsemi, V., Betz, M., Hartleib, J. and Rüterjans, H. (2002) *J. Biomol. NMR*, **22**, 153–164.
MacKenzie, D.A., Spencer, J.A., Le Gal-Coeffet, M.F. and Archer, D.B. (1996) *J. Biotechnol.*, **46**, 85–93.
Marion, D., Ikura, M., Tschudin, R. and Bax, A.J. (1989) *J. Magn. Reson.*, **89**, 496–514.
Mine, S., Ueda, T., Hashimoto, Y., Tanaka, Y. and Imoto, T. (1999) *FEBS Lett.*, **448**, 33–37.
Mok, K.H. and Hore, P.J. (2004) *Methods*, **34**, 75–87.
Montelione, G.T., Lyons, B.A., Emerson, S.D. and Tashiro, M. (1992) *J. Am. Chem. Soc.*, **114**, 10974–10975.

Piotto, M., Saudek, V. and Sklenar, V. (1992) *J. Biomol. NMR*, **2**, 661–5.
Redfield, C. and Dobson, C.M. (1988) *Biochemistry*, **27**, 122–136.
Sattler, M., Schleucher, J. and Griesinger, C. (1999) *Progr. Nucl. Magn. Res. Spec.*, **34**, 93–158.
Schwalbe, H., Fiebig, K.M., Buck, M., Jones, J.A., Grimshaw, S.B., Spencer, A., Glaser, S.J., Smith, L.J. and Dobson, C.M. (1997) *Biochemistry*, **36**, 8977–8991.
Schwalbe, H., Grimshaw, S.B., Spencer, A., Buck, M., Boyd, J., Dobson, C.M., Redfield, C. and Smith, L.J. (2001) *Protein Sci.*, **10**, 677–688.
Shaka, A.J., Barker, P.B. and Freeman, R.J. (1985) *J. Magn. Reson.*, **64**, 547–552.
Shaka, A.J., Lee, C.J. and Pines, A. (1988) *J. Magn. Reson.*, **77**, 274–293.
Slupsky, C.M., Gentile, L.N. and McIntosh, L.P. (1998) *Biochem. Cell Biol.*, **76**, 379–390.
Smith, L.J., Sutcliffe, M.J., Redfield, C. and Dobson, C.M. (1991) *Biochemistry*, **30**, 986–996.
Smith, L.J., Sutcliffe, M.J., Redfield, C. and Dobson, C.M. (1993) *J. Mol. Biol.*, **229**, 930–944.
Spencer, A., Morozov-Roche, L.A., Noppe, W., MacKenzie, D.A., Jeenes, D.J., Joniau, M., Dobson, C.M. and Archer, D.B. (1999) *Prot. Expr. Purif.*, **16**, 171–180.
Wirmer, J., Schlörb, C., Klein-Seetharaman, J., Hirano, R., Ueda, T., Imoto, T. and Schwalbe, H. (2004) *Angew Chem. Int. Ed. Engl.*, **43**, 5780–5785.
Wishart, D.S., Bigam, C.G., Yao, J., Abildgaard, F., Dyson, H.J., Oldfield, E., Markley, J.L. and Sykes, B.D. (1995) *J. Biomol. NMR*, **6**, 135–140.
Wittekind, M. and Mueller, L. (1993) *J. Magn. Reson.*, **101**, 201–205.
Yamazaki, T. and Kay, L.E. (1993) *J. Am. Chem. Soc.*, **115**, 11054–11055.