Comparison of the accuracy of protein solution structures derived from conventional and network-edited NOESY data

CHARLES G. HOOGSTRA滕,¹ SOHEUI CHOE, WILLIAM M. WESTLER,
AND JOHN L. MARKLEY

Department of Biochemistry, University of Wisconsin—Madison, Madison, Wisconsin 53706

(RECEIVED May 24, 1995; ACCEPTED August 15, 1995)

Abstract

Network-editing experiments are variants of the basic NOESY experiment that allow more accurate direct measurement of interproton distances in macromolecules by defeating specific spin-diffusion pathways. Two network-editing approaches, block-decoupled NOESY and complementary-block-decoupled-NOESY, were applied as three-dimensional, heteronuclear-edited experiments to distance measurement in a small protein, turkey ovomucoid third domain (OMTKY3). Two-hundred and twelve of the original 655 distance constraints observed in this molecule (Krezel AM et al., 1994, *J Mol Biol* 242:203–214) were improved by their replacement by distances derived from network-edited spectra, and distance geometry/simulated annealing solution structure calculations were performed from both the unimproved and improved distance sets. The resulting two families of structures were found to differ significantly, the most important differences being the hinge angle of a β -turn and an expansion of the sampled conformation space in the region of the reactive-site loop. The structures calculated from network-editing data are interpreted as a more accurate model of the solution conformation of OMTKY3.

Keywords: BD-NOESY-HSQC; CBD-NOESY-HSQC; cross relaxation; network editing; ¹⁵N NMR assignments; NOESY; nuclear magnetic resonance; protein structure; spin diffusion

The accuracy and reliability of structural models derived by NMR is a key issue in the methodology of structure determination. In the context of a drug-development study, for example, a relatively minor structural distortion can lead to a tremendous amount of wasted effort on the part of synthetic chemists. NMR data directly determine only short-range constraints, and those with relatively low precision, but yield those constraints in great abundance such that converged and apparently highly precise structures are obtained. Therefore, the extent to which the accuracy and precision of distance measurements will affect the quality of derived structural models has been a source of controversy for some time.

Two different approaches can be taken in improving the accuracy of distance measurements derived from NMR data. One approach makes use of a model structure in the analysis of the full relaxation matrix (Boelens et al., 1988; Yip & Case, 1989;

Borgias & James, 1990). The other approach, which we have been pursuing, makes use of specialized NOE experiments designed to defeat magnetization-exchange pathways that degrade distance measurements. The general approach, which is known as magnetization exchange network editing (MENE) (Macura et al., 1992, 1994), involves the cancellation of selected cross-relaxation or chemical-exchange pathways through the manipulation of magnetization during the mixing time. Examples of such experiments are the block-decoupled NOESY (BD-NOESY) experiment, which measures distances within a chemical shift-based spectral block free of spin-diffusion contributions mediated by protons outside of the block (Hoogstraten et al., 1993), and the complementary-block-decoupled NOESY (CBD-NOESY) experiment, which measures distances between two protons in different spectral blocks free of spin-diffusion pathways involving at least one transfer within a single block (Hoogstraten et al., 1995a). If the proton NMR spectrum of a protein is divided into one block containing the amide and aromatic-ring protons and a second block containing the aliphatic protons, the application of these two experiments in combination allows the measurement of a large fraction of the observable distances in the system with a very significant reduction in errors due to spin diffusion.

Reprint requests to: John L. Markley, Department of Biochemistry, University of Wisconsin—Madison, 420 Henry Mall, Madison, Wisconsin 53706; e-mail: markley@nmrfam.wisc.edu.

¹ Present address: Department of Chemistry and Biochemistry, Campus Box 215, University of Colorado at Boulder, Boulder, Colorado 80309-0215.

As reported here, we have applied the network-editing approach to the refinement of the solution structure of turkey ovomucoid third domain (OMTKY3). Ovomucoid, a proteinase inhibitor consisting of three tandem homologous domains, is the second most abundant protein in avian egg white. Comparative studies of ovomucoids and ovomucoid third domains from more than 100 species of birds have yielded significant insight into structure–function relationships in this class of molecule (Laskowski et al., 1987). Turkey ovomucoid third domain is a 56-residue polypeptide, which contains several residues of disordered linker at the N-terminus and three covalent disulfide bonds. High-resolution crystal structures of OMTKY3 in complex with α -chymotrypsin (Fujinaga et al., 1987) and *Streptomyces griseus* proteinase B (Read et al., 1983) are available, as is a 1.5-Å structure of the closely related (one amino acid change) silver pheasant ovomucoid third domain (OMSVP3) in uncomplexed form (Bode et al., 1985). Complete ^1H NMR resonance assignments of OMTKY3 are available (Robertson et al., 1988), and solution structures have been derived from conventional NOE and coupling constant data (Krezel et al., 1994). In previous studies, comparison was made of structures calculated from NOESY data at a single mixing time and from an improved data set that incorporated distances derived from simultaneous fitting of NOESY and ROESY buildup curves; the improved structures were found to be closer by positional RMS deviation (RMSD) to a set of X-ray structures of ovomucoid third domains than the structures calculated from unimproved data (Krezel et al., 1994). The NMR assignments and solution structures have also been published for modified OMTKY3 (OMTKY3*), in which the reactive-site peptide bond (between residues 18 and 19) is hydrolyzed (Rhyu & Markley, 1988; Walkenhorst et al., 1994).

The purpose of the work described here is to evaluate the effect of the improvement of a subset of interproton distances in OMTKY3 using MENE on a derived solution structure. To allow the analysis of large numbers of cross peaks using normalization on the diagonal, we have implemented BD-NOESY and CBD-NOESY in three-dimensional heteronuclear-edited form on uniformly ^{15}N -labeled OMTKY3. In cross-relaxation spectra, normalization on the corresponding diagonal or direct peak removes the effect of autorelaxation, lengthens the linear buildup region, and corrects for spectrum-to-spectrum variations (Macura et al., 1986); in heteronuclear-edited spectra, this procedure also corrects for the effects of variations in coherence transfer efficiency and heteronuclear relaxation rates. In the interest of a clear demonstration of the effects of constraint accuracy and precision, we have only analyzed those NOE constraints that were used in the original structure determination (Krezel et al., 1994); additional NOE cross peaks resolved by virtue of the improved peak dispersion in multidimensional, multinuclear spectroscopy (Clare & Gronenborn, 1991) were not used because they would have biased the comparison.

To facilitate these studies, we have assigned the ^{15}N NMR resonances of OMTKY3, a procedure that entails little difficulty in a protein of this size for which the proton assignments are known. In addition, we have performed ^{15}N T_1 and $T_{1\rho}$ relaxation time measurements on uniformly labeled OMTKY3, because such measurements provide an estimate of the correlation time for isotropic rotation (τ_0), which is useful in the interpretation of buildup data. Relaxation studies also allow the identification of residues for which the backbone ^{15}N atom is rigidly

attached to the protein framework (in other words, does not undergo significant internal motion). Because internal motion may quench NOESY cross-peak intensity (Neuhaus & Williamson, 1989), the application of explicit lower bounds to distance constraints is only appropriate when its presence can be ruled out.

One problem in studies of this type is that there is no rigorous method to determine the absolute accuracy of a set of structures determined by NMR, because the “true” structure is not known. The approach used in the present investigation was to calculate, using identical methods, families of structures from two different input data sets, one of which is known, on methodological grounds, to contain fewer errors due to spin diffusion than the other (Hoogstraten et al., 1993, 1995a). If the two sets of structures are significantly different from each other as compared to the spread of structures within each set, it may be argued a priori that the structures based on the more accurate input are to be preferred, and that real refinement has been achieved. An independent measure of accuracy may also be inferred by comparison of the derived results with experimental quantities that were not used in the calculations; therefore, we also have assessed the ability of each set of structures to reproduce the observed chemical shifts for H^α atoms in the molecule (Williamson & Asakura, 1993; Williamson et al., 1995). Finally, we compare the derived solution sets to the three high-resolution crystal structures of ovomucoid third domains mentioned above. The X-ray structures, however, provide only a problematic measure of absolute structural accuracy because of the limited accuracy of the crystallographic structures themselves and the possibility for real differences between the solution and crystal states. In addition, the crystal structures considered here are either of OMTKY3 in complex with another protein or of a slightly different uncomplexed ovomucoid third domain (OMSVP3); the crystal structure of uncomplexed OMTKY3 has not been reported.

Results and discussion

^{15}N shift assignments for OMTKY3

The ^1H shift assignments for OMTKY3 have been reported (Robertson et al., 1988), but no ^{15}N assignments for this protein have been published. With the high digital resolution of the three-dimensional experiments reported here (see below), essentially all of the backbone amide resonances were resolved in the proton–nitrogen plane, and the ^{15}N signals could be identified by assigning cross peaks in three-dimensional HSQC-TOCSY and -NOESY spectra. The ^{15}N shift assignments derived in this fashion are reported in the Electronic Appendix, along with the frequencies at which the amide protons were observed. Proton chemical shifts are reported relative to internal DSS. ^{15}N chemical shifts are reported relative to liquid NH_3 , as calculated using the proton DSS line and the ratio of absolute spectrometer frequencies for the two nuclei (Edison et al., 1994). For residues 2, 4, and 5, two sets of peaks corresponding to the NH group were observed, possibly indicating some structural or conformational heterogeneity in the N-terminal linker. This result is not surprising because this region has been observed to be disordered in previous structural studies of OMTKY3 (Fujinaga et al., 1987; Krezel et al., 1994). The amide proton resonance frequencies are very close to those reported previously (Robertson et al., 1988), with the exception of the reactive-site loop res-

idues Thr-17 and Glu-19; the differences may be due to slight differences in solution conditions between previous and current protein preparations. All peaks analyzed in the remainder of this work were consistent with the chemical shifts reported in the assignment table.

Heteronuclear relaxation studies

For all T_1 and $T_{1\rho}$ magnetization decay curves analyzed, single-exponential decay was observed to within the limits of the data. Relaxation rates could be obtained for all non-proline residues except for Tyr-11 and Ala-15, which overlapped with each other, and Thr-47, which overlapped with a side-chain amide group of an asparagine residue. These rates are summarized in Figure 1. A full analysis of the internal backbone dynamics of OMTKY3 would be best performed with additional data, such as heteronuclear NOEs (Torchia et al., 1993) or spin-lock power varia-

tion for $T_{1\rho}$ (Szyperski et al., 1993), and is deferred to another place; we merely observe here that significant deviations from the most common values cluster in the flexible N-terminus, reactive-site loop, and reverse turns, as expected.

For 15 residues identified on the basis of secondary structure and hydrogen exchange rates as part of the rigid framework of the molecule, the mean values of T_1 and $T_{1\rho}$ were $(0.365 \pm 0.018 \text{ s})$ and $(0.261 \pm 0.020 \text{ s})$, respectively. The $T_1/T_{1\rho}$ ratio for these residues is thus (1.40 ± 0.13) . Equations that take account of relaxation due to both dipolar and chemical-shift anisotropy contributions (Kay et al., 1992) can be used to convert these values into estimates of the correlation time for isotropic tumbling of OMTKY3 under the experimental conditions. The results of this analysis are 1.8 ns (1.6–2.0 ns) or 5.1 ns (4.5–5.6 ns) from T_1 (double-valued), 1.9 ns (1.7–2.1 ns) from $T_{1\rho}$, and 2.0 ns (1.4–2.4 ns) from the ratio. All three values are seen to be highly consistent. The calculation from the ratio is a valuable check, because it is less dependent on certain experimental errors than either direct calculation, but it is expected to give a less precise determination (Szyperski et al., 1993); therefore, we consider a range consistent with the two direct calculations and report the correlation time as $1.8 \pm 0.3 \text{ ns}$. We earlier reported a correlation time for this protein of 6.0 ns in D_2O solution (Fejzo et al., 1990); this value was determined at much greater protein concentration than the current work and was apparently affected by aggregation and viscosity effects. The value reported here is in good agreement with recent results for the global dynamics of BPTI, a protein of similar size, in dilute solution (Szyperski et al., 1993).

Three-dimensional spectra

The three-dimensional network-editing experiments yielded high-quality data with excellent water suppression. An example of the data obtained is shown in Figure 2, which displays slices through CBD-NOESY-HSQC and NOESY-HSQC spectra of OMTKY3 at 750 MHz at the ω_2 and ω_3 frequencies of the amide group of Lys-34. NOEs to the methylene protons of Asn-33 are indicated. The $K^{34} H^N/N^{33} H^{\beta 2}$ NOE is drastically reduced in intensity in CBD-NOESY-HSQC data compared to NOESY-HSQC data, indicating that the NOESY-HSQC result is badly contaminated with spin diffusion, presumably mediated by the $N^{33} H^{\beta 3}$ proton. In contrast, the $K^{34} H^N/N^{33} H^{\beta 3}$ NOE is at essentially the same relative intensity in both spectra and is interpreted as arising largely from direct cross relaxation. The three-dimensional experiments are thus seen to have the same ability to discriminate between direct and spin-diffusion contributions to cross peaks as the two-dimensional sequences reported previously (Hoogstraten et al., 1995a), with the additional advantage of the greater peak dispersion provided by multidimensional spectroscopy.

Distance measurements

The reliability of distance measurements from network-editing data can be verified in a number of ways. In the case of BD-NOESY, a number of backbone amide-backbone amide NOEs could be analyzed using two different cross peaks (at the direct-transfer frequencies of either residue involved); these cases should in principle provide two independent determinations of the same cross-relaxation rate. Of the 20 cases where this was

Fig. 1. Heteronuclear relaxation rates in OMTKY3. Backbone ^{15}N T_1 (top) and $T_{1\rho}$ (bottom) relaxation rates in OMTKY3. Rates for Ala-2 and Ala-3 are off scale and not shown; blanks at other positions correspond to prolines or residues with overlapped signals.

Fig. 2. Three-dimensional CBD-NOESY-HSQC data. Slices are shown through three-dimensional CBD-NOESY-HSQC (A) and NOESY-HSQC (B) at 160 ms and 750 MHz at the ω_2/ω_3 frequencies of the backbone amide of Lys-34. NOEs to the β -methylene protons of Asn-33 are indicated.

possible, for 12, the derived cross-relaxation rates were equivalent to within the error in the fit; in only 1 case was the difference greater than 50%. (In these cases, the smaller rate was used in the distance calculation, as it would yield the more conservative constraint.) Experimental distances may also be compared with distances known to characterize regions of stable secondary structure (Wüthrich, 1986). For BD-NOESY, the observable $H^N-H^N(i, i+1)$ distances within the α -helix in OMTKY3 gave a median cross-relaxation rate of 0.389 s^{-1} , corresponding to an interproton distance of 2.53 Å ; in the crystal structure in complex with α -chymotrypsin (Fujinaga et al., 1987), the median of the corresponding distances is 2.69 Å . For CBD-NOESY, the $H^N-H^N(i, i+3)$ distance was observable in five cases within the α -helix; the median experimental distance for these was 3.23 Å , compared with a median of 3.57 Å from the X-ray structure (Fujinaga et al., 1987) and an accepted value for this distance of 3.4 Å in a perfect helical conformation (Wüthrich, 1986). Distances derived from network-editing data are thus seen to reproduce expected distances in regions of regular secondary structure, albeit with a slight residual underestimation; this underestimation is well within the error bars assumed in the translation of distances into upper distance bounds, and so does not lead to inaccurate constraints in the cases analyzed here.

Comparison of unrefined and refined structures

Fifty converged structures were calculated using either the "crude" data set, consisting of distances previously calibrated from NOESY data at a single mixing time (Krezel et al., 1994), or the "improved" data set, in which a fraction of the distance constraints were replaced with data derived from more accurate network-editing experiments. Dihedral-angle and hydrogen-bond constraints were translated directly from the original work

(Krezel et al., 1994) and used in identical fashion for both data sets. RMSD and dihedral angle deviation (DHAD) analyses (Havel & Wüthrich, 1985) of these structures are summarized in Table 1. "Pairwise" RMSD and DHAD values are the mean of comparisons within a structure set. Comparisons to the three high-resolution X-ray structures mentioned above are also shown, along with RMSD and DHAD comparisons between the structural sets. For superpositions, RMSDs, and χ^1 DHAD involving the OMSVP3 structure, the amino acid differing between the two molecules (residue 18) was omitted; all other RMSD calculations were performed for residues 7–56, ϕ and ψ DHAD calculations for residues 7–55, and χ^1 DHAD calculations for residues 7–56, excluding alanines and glycines.

Both at the global (as represented by positional RMSD) and local (as represented by DHAD) levels, the differences between the two sets of structures are greater than the spread observed within either single set. This observation indicates that the two sets of structures differ from each other to a significant extent; given the a priori arguments that improved input is in all cases to be preferred, this indicates that the "improved" structure set represents a true refinement of the structures derived from "crude" data. Differences between the sets of structures may be analyzed in more detail by reference to the plots of positional RMSD versus protein sequence shown in Figure 3. For both backbone and heavy-atom RMSD, the most significant differences cluster in the β -turn containing residues 25–28 and in the reactive-site loop, centered on residues 18 and 19. For the β -turn, the two sets diverge from each other to a much greater extent than the spread within either set, indicating the presence of a significant difference in local structure beyond that represented by the inherent imprecision of the structure determination. This difference is interpreted, as discussed above, as an inaccuracy in the structures calculated from unimproved data that has been removed by incorporation of network-editing data. In the reactive-site loop, in contrast, the pairwise RMSD within the

Table 1. Analysis of derived sets of structures

Parameter ^a	Crude	Improved
Pairwise RMSD (bb, ha)	0.78, 1.35	0.86, 1.41
RMSD between sets (bb, ha)	0.91, 1.47	
RMSD from 1CHO (bb, ha)	0.91, 1.49	0.98, 1.55
RMSD from 2OVO (bb, ha)	1.25, 1.80	1.25, 1.79
RMSD from 3SGB (bb, ha)	1.23, 1.83	1.26, 1.87
Pairwise DHAD (ϕ, ψ, χ^1)	25.6, 24.2, 59.4	31.2, 26.9, 57.4
DHAD between sets (ϕ, ψ, χ^1)	32.2, 29.3, 59.6	
DHAD from 1CHO (ϕ, ψ, χ^1)	28.4, 29.0, 55.5	32.4, 29.2, 55.8
DHAD from 2OVO (ϕ, ψ, χ^1)	35.9, 36.1, 105.8	38.0, 36.3, 105.2
DHAD from 3SGB (ϕ, ψ, χ^1)	36.4, 31.0, 108.4	37.2, 31.7, 105.0
SD for H^N (ppm)	0.292	0.284
RMS violations		
Crude constraints	0.0355	0.1009
Improved constraints	0.0547	0.0370

^a bb, backbone atoms; ha, all non-hydrogen atoms; 1CHO, crystal structure of OMTKY3 in complex with α -chymotrypsin (Fujinaga et al., 1987); 2OVO, crystal structure of uncomplexed OMSVP3 (Bode et al., 1985); 3SGB, crystal structure of OMTKY3 in complex with *S. griseus* protease B (Read et al., 1983); SD, standard deviation of calculated from observed chemical shifts. All values are means across the sets of structures; minimized average structures were not derived.

Fig. 3. Analysis of positional RMSD versus protein sequence. Plots of (top) backbone and (bottom) heavy-atom RMSD values for each residue for residues 5–56 in OMTKY3 for structures calculated in this work. Indicated parameters are: solid line, pairwise RMSD for structures calculated from crude data; long dashes, pairwise RMSD for structures calculated from improved data; short dashes, RMSD between the two structural sets.

structures calculated from improved data is significantly larger than that within the structures calculated from crude data, and the divergence between the two sets reflects the larger value. This observation represents an expansion of the conformation space explored by “improved” as compared to “crude” structures and may be interpreted as the constriction of the structure set by the inaccurate data to an unjustified level of apparent precision. The reactive-site loop is the major contributor to the increased global pairwise RMSD figure for improved as compared to crude structures (Table 1).

The connection between the structural changes and the distance constraints may be seen in the analysis of violations of each data set by each structural set (Table 1). The improved structures show a considerably greater global RMS violation of the crude constraints than do the structures calculated from those crude constraints. Likewise, but to a lesser extent, the crude structures show a greater RMS violation of the improved constraints than do the structures calculated from those improved constraints. These results indicate that the structural differences observed are in fact due to changes in the distance constraints. A plot of summed NOE violations versus protein sequence (Fig. 4) maps these changes to the observed structural shifts. Interestingly, the improved structures show drastically increased violations of the crude constraints in the region of the β -turn discussed above (residue 28), and the crude structures also show a (somewhat smaller) increase in violations of the improved constraints in the same region, supporting the interpretation that the two data sets are consistent with two noncongruent regions of conformation space in this region of the protein. In contrast, the improved structures show significantly increased violations of the crude constraints in the region of the reactive-site loop, whereas the crude structures show no significantly in-

Fig. 4. Violations of constraint sets. Top: Violations of 655 crude NOE constraints by structures calculated from crude (solid line) and improved (dashes) data sets. Bottom: Violations of 655 improved constraints by structures calculated from crude (solid line) and improved (dashed line) data sets. For each residue, absolute values of violations of all inter- and intraresidue NOE constraints involving that residue were summed and divided by the total number of NOE constraints in the data set involving that residue and by the number of structures in the set.

creased violations of the improved restraints. This observation indicates that the expansion of the improved structures in this region of the protein is directly related to the requirements of the experimental constraints, in that the improved constraints are consistent with a larger region of conformation space than the crude constraints, presumably due to the removal of inaccuracies due to spin diffusion present in the unedited NOESY data.

The nature of the structural change in the β -turn is illustrated in Figure 5, which shows a superposition of the polypeptide backbone for the unminimized average structures for the two structure sets with the width of the ribbon drawn proportional to the backbone RMSD. The difference is seen to arise from a hinging of the entire loop region at the α -carbons of residues 25 and 29. This hinging gives rise to a difference between the two average structures at the C^α atom of Asp-27, the extreme of the loop, of 2.1 Å, and correspondingly smaller deviations at other portions of the loop. In this presentation, the nonoverlapping character of the ribbons in the region of Asp-27 indicates that the difference between the average structures shown is greater than that expected due to the inherent imprecision of the determinations, as seen in Figure 3. Little change is observed throughout the remainder of the structure. The changes in the reactive-site loop involve an expansion rather than a displacement of the set of structures (Fig. 3); this change is not clearly visible in the perspective shown in Figure 5.

Comparison with X-ray crystallographic studies

We have also compared the crude and improved structure sets to three published high-resolution crystal structures of ovom-

Fig. 5. Average structures of OMTKY3. Comparison of polypeptide backbone for residues 7–56 of unminimized average structures calculated from crude (blue) and improved (red) data. For calculation of average structures, structures within a set were superimposed by reference to the backbone atoms for residues 7–56; the two average structures were superimposed in the same way for comparison. Widths of the ribbons are proportional to the backbone pairwise RMSD for the appropriate set of structures. Figure prepared using INSIGHT-II 2.3.5 software package (Biosym Technologies, Inc., San Diego, California).

coid third domains, at the global (RMSD) and local (DHAD) level (Table 1). At the global level, the improved structures show increased RMSD from the 1CHO and 3SGB structures of OMTKY3 in complex with protease as compared with the crude structures; the two structure sets show equivalent RMSDs to the 2OVO structure of uncomplexed OMSVP3. Plots of these statistics versus protein sequence, shown in Figure 6, indicate a strong localization of the differences between the improved structures and the complexed crystal structures to the turn containing residue 27; the improved structures, by contrast, are slightly closer to the 2OVO structure in this region than are the crude structures. In short, the effect of network-editing data has been to shift the structures from a conformation more characteristic of OMTKY3 in complex with protease to one more characteristic of ovomucoid third domains in uncomplexed form. Because the solution studies were performed on free OMTKY3, this observation supports the conclusion that the improved structures are a more accurate solution model than the crude structures.

At the local level, the improved structures diverge slightly more than the unimproved structures from the ϕ and ψ angles observed in the crystal structures, but better approach crystal χ^1 values. These observations may reflect expansion of the conformational space covered by the improved as opposed to the crude structures; note that χ^1 DHAD, the only statistic of the

Fig. 6. Backbone RMSD comparison with crystallographic structures. Backbone RMSD of crude (solid line) and improved (dashed line) structures from X-ray structures of ovomucoid third domains. Top: 1CHO crystal structure of OMTKY3 in complex with α -chymotrypsin (Fujinaga et al., 1987), 1.8 Å resolution. Center: 2OVO crystal structure of uncomplexed OMSVP3 (Bode et al., 1985), 1.5 Å resolution, residue 18 not analyzed. Bottom: 3SGB crystal structure of OMTKY3 in complex with *S. griseus* protease B (Read et al., 1983), 1.8 Å resolution.

three by which the improved structures cover a smaller region of conformational space than the crude structures, is also the only statistic by which the improved structures are closer on average to the three crystal structures than are the crude structures.

Chemical-shift analysis

Finally, we have adduced the ability to reproduce observed chemical shifts as a measure of structural accuracy. Although the accuracy of chemical-shift calculations themselves is imperfect, such an analysis is supported by recent work showing a correlation between the resolution of crystal structures and the consistency of chemical shifts calculated from these structures with observed shifts (Williamson et al., 1995). We have used the chemical-shift calculation algorithm of Williamson and coworkers (Williamson & Asakura, 1993; Williamson et al., 1995) to calculate the standard deviation (SD) between calculated and observed H^α chemical shifts averaged across the 50-member structural sets. Calculations were performed for residues 7–56, excluding glycines. For the crude structures, this calculation gave a mean SD of 0.293 ppm; the corresponding statistic for the im-

proved structures was 0.284 ppm. Both sets of structures thus fall among the better NMR solution models of a large number evaluated with this algorithm and correspond to crystallographic models of roughly 2.0 Å resolution (Williamson et al., 1995). The small difference between the two values is within the error in both the calculations and the chemical-shift referencing, suggesting that the relative accuracy of these two sets of structures cannot be distinguished by current chemical-shift calculation algorithms. A plot of deviations versus protein sequence showed no localized regions in which one set of structures reproduced α -proton chemical shifts better than the other (data not shown), supporting this interpretation.

Conclusions

The results presented in this work demonstrate that it is possible, in an experimental situation, to use network-editing techniques to derive a substantial fraction of the interproton distances in an NMR structural study with improved accuracy. Only a subset of the total number of such constraints obtainable from three-dimensional network-editing spectra were used in this study: so as to preserve the comparisons based only on differences in constraint accuracy and precision but not in the total number of constraints, only those interproton distances observed in an earlier, two-dimensional NOESY study of this molecule were analyzed. Three-dimensional spectroscopy, however, renders additional distances observable by virtue of improved peak dispersion (Clare & Gronenborn, 1991); analysis of those cross peaks as well would have increased the fraction of "improved" interproton distances observed. Increases in the number of constraints made possible by the use of longer mixing times for network-editing spectra (Hoogstraten et al., 1993) were also not considered here. Nevertheless, the divergence of the two structural sets from each other suggests that the use of network-editing data to remove spin-diffusion contributions has resulted in an increase in structural accuracy. The expansion of the "improved" structures in the region of the active site indicates that the structures calculated from a simple NOESY analysis reflect an unjustified level of precision in this important region of the structure and re-emphasize that a very small pairwise RMSD is an inadequate measure of structural quality (James, 1994; Zhao & Jardetzky, 1994).

The NOE data set for OMTKY3 consists of approximately 12 constraints per residue. This represents a highly refined distance set, as significant redundancy has been shown to set in above approximately 10 constraints per residue, although below this threshold the quality of structures depends critically on the number of constraints (Clare et al., 1993). Distance accuracy improvements using network-editing data are only likely to be of interest for data sets at or above this level of completeness because the structural quality improvements in this work are not as dramatic as those observed due to, for example, an increase from 6 to 12 constraints per residue. In addition, because CBD-NOESY has the most dramatic effect on NOEs involving methylene groups, stereospecific assignments are a necessity for the full power of this method to be realized. MENE analysis is thus most appropriate for accuracy improvement and additional refinement of already high-quality structures. Techniques both for massive NOE cross-peak analysis and stereospecific assignments, however, are widely available (Clare & Gronenborn, 1993). In addition, the increased accuracy of MENE data for diastereo-

topic groups should assist in the determination of stereospecific assignments. Network-editing techniques should thus be applicable to a wide variety of experimental systems.

We previously reported a different technique for the refinement of distance accuracy in protein structures in which NOESY and ROESY buildup curves are fit simultaneously to a single improved value of the cross-relaxation rate (Fejzo et al., 1989). As with the present analysis of the MENE approach, we used structure calculations of OMTKY3 as a means of evaluating the effects of the refinement (Krezel et al., 1994); this puts us in the position of being able to compare these two approaches. MENE analysis removes certain classes of spin-diffusion effects from the cross-relaxation data, whereas NOESY/ROESY comparisons permit one to detect spin diffusion so that its effects can be accounted for. Thus, it is possible to use comparisons or simultaneous analysis of network-edited NOESY and ROESY spectra (for example, BD-NOESY and BD-ROESY) as a means of identifying the presence of unedited spin diffusion (Hoogstraten et al., 1995b), although this approach was not utilized extensively here. The NOESY/ROESY method can be applied to situations in which the spectral distribution does not allow efficient MENE analysis (aliphatic-aliphatic cross peaks in proteins, for example), whereas MENE experiments can measure distances accurately in geometries for which the standard experiments are badly perturbed by spin diffusion even at very short mixing times. It is anticipated that the MENE approach will be more suitable than NOESY/ROESY comparisons to refinement of larger structures, because of deleterious $T_{1\rho}$ effects on ROESY data acquired with long mixing times. MENE data will be less sensitive (although not completely insensitive) to such effects, and compensatory sensitivity improvements are possible due to the use of mixing times closer to the cross-peak sensitivity optimum (Hoogstraten et al., 1993, 1995a).

Network-editing experiments provide a flexible framework for the careful analysis of cross relaxation in biological macromolecules (Macura et al., 1992, 1994). The particular scheme used in this work, the application of three-dimensional versions of BD-NOESY and CBD-NOESY in combination, can be implemented with only a moderate increase in the time required for data acquisition and analysis. A key point is that both BD-NOESY and CBD-NOESY are nonselective experiments and allow the analysis of large numbers of NOEs in a single NOESY-type experiment. The benefits in structural accuracy seem to warrant the additional effort. It should be pointed out that the improved structures resulting from this approach could be refined further by means of direct or iterative relaxation-matrix refinement using a combination of conventional and network-edited NOESY spectra.

Materials and methods

Protein preparation

Uniformly labeled, recombinant OMTKY3 was overexpressed in *Escherichia coli* as a fusion protein with staphylococcal nuclease as described for the related protein OMCH13 (Hinck et al., 1993). The overproduction plasmid was modified for the production of OMTKY3 by mutation of the three amino acids different from OMCH13 and the mutation of two of the internal methionine residues in the nuclease moiety to alanines (A.P. Hinck, W.F. Walkenhorst, and S. Choe, unpubl. work). The sole

nitrogen source used in the growth medium was $^{15}\text{NH}_4\text{Cl}$ (98+ atom % enriched, Cambridge Isotope Laboratories, Woburn, Massachusetts). After purification of the fusion protein, OMTKY3 was separated from staphylococcal nuclease by cyanogen bromide cleavage (50-fold molar excess CNBr at 10 mg/mL protein, stirred at room temperature for 24 h) and was purified by cation-exchange chromatography on a Mono S column on an FPLC system (Pharmacia Biotech, Uppsala). The column buffer was 50 mM MES, pH 6.2. Protein was eluted with a shallow (0–0.1 M in 120 mL) NaCl gradient begun immediately following sample loading. The purified protein had a specific activity of 2.1×10^5 U/mg for inhibition of the chymotrypsin-catalyzed hydrolysis of succ-Ala-Ala-Pro-Phe-para-nitroanilide (DelMar et al., 1979), as compared to a value of 2.0×10^5 U/mg measured for an authentic sample of OMTKY3 prepared from turkey egg whites. Two-dimensional ^{15}N - ^1H NMR correlation spectroscopy also provided an index of purity. The OMTKY3 concentration in the NMR sample was 2.6 mM, and the pH was adjusted to 4.1. The same sample was used for all experiments.

^{15}N relaxation time measurements

All NMR experiments were performed at 298 K. To analyze the global and local dynamics of OMTKY3, we determined T_1 and $T_{1\rho}$ values for the backbone ^{15}N resonances of this molecule at 50.6 MHz (11.7 Tesla; proton frequency of 500 MHz). T_1 measurements were performed with the pulse sequence of Kay et al. (1992), including 180° pulses during the relaxation delay to suppress cross-correlation effects. Spectra were acquired at relaxation delays of 40, 80, 120, 180, 240, 300, 360, 420, 480, 560, 640, 720, 800, and 900 ms. $T_{1\rho}$ spectra were acquired with a slight modification of the pulse sequence of Peng et al. (1991) at relaxation delays of 40, 60, 80, 120, 140, 160, 200, 240, 280, 320, 360, 400, 440, and 480 ms. The spin-lock field strength was 2,910 Hz. For all spectra, 256 t_1 (^{15}N) slices of 4K points and 16 scans each were acquired, sweep widths of 6,009.6 Hz (^1H) and 1,602.6 Hz (^{15}N) were used, solvent was suppressed with low-power presaturation, and spectra were processed using FELIX 2.3 (Biosym Technologies Inc., San Diego, California) with 2 Hz exponential line-broadening in t_2 , a squared-cosine window function in t_1 , and baseplane flattening with routines internal to FELIX. Peak intensities were measured by volume integration over a rectangular footprint and relaxation times were determined by the fitting of each time course to a single-exponential decay with two parameters ($A \cdot \exp[-t/T_x]$, $T_x = T_1$ or $T_{1\rho}$).

For analysis of global dynamics, a subset of residues known from prior evidence to be associated with the rigid framework of the molecule was chosen. Residues were chosen if they belonged to regions of regular secondary structure and had slowly exchanging amide protons by ^1H NMR (Robertson et al., 1988); 15 residues met these criteria. Mean values of ^{15}N T_1 and $T_{1\rho}$ for these residues were used to derive estimates of the correlation time for molecular tumbling. Relaxation due both to dipole-dipole interactions with the directly bound proton and to chemical-shift anisotropy was considered (Kay et al., 1992). The effects of cross correlation and (for $T_{1\rho}$) of off-resonance effects were neglected. For the purposes of structural analysis, residues were classified as belonging to the rigid framework of the molecule if both their T_1 and $T_{1\rho}$ values were within 1.5

SDs of the mean for the 15 residues selected above; 27 of the 56 residues in OMTKY3 met this criterion.

Cross-relaxation experiments

For this work, cross-relaxation measurements were performed in three-dimensional heteronuclear (^{15}N)-edited form as NOESY-HSQC, BD-NOESY-HSQC, and CBD-NOESY-HSQC. The pulse sequences used for the network editing experiments are shown in Figure 7; the two sequences shown, as well as the NOESY experiment, differ only in processes during the mixing time and in slight readjustments of gradient strengths to optimize solvent suppression. The mixing sequences for BD-NOESY (A) and CBD-NOESY (B) are identical to those published previously (Hoogstraten et al., 1993, 1995a), except that the selective pulses in BD-NOESY are alternated with short delays to reduce the total number of pulses applied (Macura et al., 1994). The gradients are arranged in a pattern of coherence destruction rather than coherence selection, as described by Bax and Pochapsky (1992); because no coherences are dephased during

Fig. 7. Three-dimensional network-editing pulse sequences. NMR pulse sequences for the BD-NOESY-HSQC (A) and CBD-NOESY-HSQC (B) experiments. Upper portion shows the ^1H RF pulses, ^{15}N RF pulses, and z -gradient pulses for the framework NOESY-HSQC sequence; network-editing mixing sequences are indicated separately (A, BD-NOESY; B, CBD-NOESY). Shaded triangles for RF pulses indicate low-power selective pulses. Diagonal shading indicates spin lock. τ_N , longitudinal relaxation period; SL_x , transverse relaxation period (spin-locked); τ_p , precession delay; Δ , $1/4J$ delay for evolution of heteronuclear coupling. Phase cycle: ϕ_1 : 0, π ; ϕ_2 : 0, 0, π , π ; receiver: 0, π , π , 0.

incremented delays, no additional refocusing steps or separate accumulation of P- and N-type transients are required. Gradient G_1 destroys unwanted coherences at the beginning of the mixing time that might otherwise be acted on in unpredictable fashion by the network-editing sequences, whereas gradients G_3 and G_4 select for z-magnetization during the coherence transfer steps of the INEPT and reverse INEPT sequences. The gradient pairs G_2 and G_5 compensate for imperfections in the refocusing 180° pulses (Bax & Pochapsky, 1992). Evolution on the nitrogen resonances occurs during t_2 . The use of gradients allowed the reduction of the phase cycle to four steps, as shown in Figure 7; in approximately 3 days of experimental time, 256 real points in t_1 (^1H) and 128 real points in t_2 (^{15}N) could be acquired. Water suppression was accomplished with the gradients and trim pulses alone; no presaturation was used. An acquisition-dimension data size of 4,096 points was used in all spectra.

BD-NOESY-HSQC spectra were acquired at 500 MHz at mixing times of 80, 120, 160, and 240 ms. The mixing sequence consisted of a train of 5-ms hyperbolic secant (hsec) pulses with phase factor 12 truncated at 12% and phase cycled as $(0 \pi \pi 0)$ alternating with 5-ms delays. The shaped pulse power level was adjusted to 1 dB above the minimum level that gave good inversion. A NOESY-HSQC at 500 MHz and 160 ms was acquired for comparison. A CBD-NOESY-HSQC spectrum with cross relaxation adjusted to be comparable to a 160-ms NOESY was acquired at 750 MHz. For this spectrum, the longitudinal relaxation delay was set to 17.50 ms, the spin-lock period to 11.12 ms, the precession delay to 33.0 μs , and the mixing sequence was repeated $N = 4$ times to the full mixing period. A spin-lock power of 3,788 Hz was used. The selective inversion pulse was a 1.1-ms IBURP2 pulse (Geen & Freeman, 1991) applied 1,800 Hz upfield of solvent to invert the aliphatic spectral region. A NOESY-HSQC spectrum was acquired at 750 MHz and 160 ms for comparison. Finally, a 40-ms TOCSY-HSQC spectrum was acquired at 500 MHz using the same basic pulse sequence and DIPSI-2 mixing (Shaka et al., 1988) to facilitate ^{15}N resonance assignments. All spectra were processed using 3-Hz exponential line-broadening in the acquired dimension and squared-cosine bell window functions in the nonacquired dimensions. Residual solvent was suppressed with convolution difference techniques and the spectra were extended by 25% in the ^{15}N dimension and 33% in the nonacquired ^1H dimension by means of complex linear prediction.

Distance measurement from network-editing data

Cross peaks in the network-editing spectra were assigned by comparison to the NOESY-HSQC spectra and interproton distance assignments from a previous structural study of this protein (Krezel et al., 1994). Intensities of cross and diagonal peaks were measured from volume integrals over a rectangular footprint in three dimensions containing the peak. All calculations were based on the ratio of cross-peak intensities to the intensity of the direct peak at the same ω_2/ω_3 frequency. For BD-NOESY-HSQC data, buildup series of cross- to diagonal-peak ratios were constructed and fitted to a quadratic polynomial with zero constant term. The linear term was interpreted as the cross-relaxation rate (Macura et al., 1986), and the conversion to a calculated distance made use of the correlation time calculated from ^{15}N relaxation data. Error bounds on the calculated dis-

tance (d) were assigned according to propagated error from uncertainties in the linear term of the fit and the correlation time estimate, but in no case were distance bounds set tighter than $\pm 0.5 \text{ \AA}$ ($d < 3.0 \text{ \AA}$) or $\pm 1.0 \text{ \AA}$ ($d > 3.0 \text{ \AA}$). In addition, explicit lower bounds were assigned according to this protocol only for distances between two backbone amide protons both of which were attached to the rigid framework of the molecule according to the relaxation-time criteria described above; in all other cases, the lower bound was set to the van der Waals contact distance of 1.9 $\text{\AA}$.

For CBD-NOESY, the cross-peak ratios were corrected for the effects of resonance offset on the cross-relaxation rate during the ROESY periods (Griesinger & Ernst, 1987). Intensities involving methyl or unresolved methylene groups were divided by the number of protons contributing to the resonance. Distances were then calibrated using the two-spin approximation and assuming an r^{-6} dependence of relative cross-peak intensity on interproton distance (Wüthrich, 1986). The mean of five observed $d_{\alpha\text{N}}$ cross peaks in β -sheet regions of the protein was taken as the calibration intensity, and the corresponding distance was set to 2.2 $\text{\AA}$ (Wüthrich, 1986). The lower distance bound was set to 1.9 $\text{\AA}$ in all cases, and the upper distance bound was set to $d + 0.7 \text{ \AA}$ ($d < 3.0 \text{ \AA}$), $d + 1.0 \text{ \AA}$ ($3.0 \text{ \AA} < d < 4.0 \text{ \AA}$), or $d + 1.5 \text{ \AA}$ ($d > 4.0 \text{ \AA}$). Data collected at 500 MHz (BD-NOESY) and 750 MHz (CBD-NOESY) were not combined or compared until the final "improved" distance constraint list was assembled.

Construction of data sets and structure calculations

The original NMR structure determination of OMTKY3 (Krezel et al., 1994) was performed with the now-obsolete computer algorithm DSPACE. Therefore, the 655 distance constraints used in this calculation were translated directly into the format of the X-PLOR/dg program (Brünger, 1992) to form the "crude" data set. Distance constraints derived by analysis of NOESY and ROESY buildup curves were not used. Upper distance bounds for the crude data set were exactly as described in the original work, and lower bounds were set to the van der Waals contact distance of 1.9 $\text{\AA}$, as in the original work. The "improved" data set was constructed by substituting the 40 distance constraints derived from BD-NOESY-HSQC and the 157 distance constraints derived from CBD-NOESY-HSQC for the corresponding constraints in the "crude" data set. In addition, 5 of the cross peaks expected from conventional NOESY data in BD-NOESY-HSQC and 10 in CBD-NOESY-HSQC were of insufficient intensity to be observed, although comparison with corresponding NOESY-HSQC spectra indicated that the network editing spectra had sufficient signal-to-noise for the observation of these peaks had they appeared at the same intensity as in NOESY. These cross peaks were interpreted as containing significant contributions due to spin diffusion in NOESY data and therefore to be potentially underestimated in the original data set; the corresponding upper bounds were loosened by 2 $\text{\AA}$ in the "improved" data set. Distances to methyl groups were referred to the carbon atom, and the upper bound was increased by the carbon-hydrogen bond length; for all other unresolved groups, distances were referred to the geometric center of the group, and the upper bound was increased by the distance from the individual atoms to the geometric center. Constraints involving resolved methylene groups that lacked stereospecific assignments

were incorporated using a high-dimensional potential (Habazettl et al., 1990; Brünger, 1992) to allow the structure to "float" between the two possible stereospecifically assigned states. All calculations included the 29 dihedral angle constraints and the 17 hydrogen bond constraints used previously for this molecule (Krezel et al., 1994), except that the hydrogen bond donated by the backbone H^N of Ser-9 was reassigned from the side chain of Glu-10 to the side chain of Asp-7 (A.D. Robertson, pers. comm.). Hydrogen bonds were imposed by constraining the H–O distance to ≤ 2.3 Å and the N–O distance to 2.5–3.3 Å.

Sets of structures for each of the two data sets were calculated using the distance geometry-simulated annealing protocol as implemented in the software X-PLOR/dg (Nilges et al., 1988; Brünger, 1992). Initial structures were generated with all-atom metric matrix embedding including random metrization of four atoms and simulated annealing regularization (Brünger, 1992). Residue-by-residue replacement of coordinates with those derived from a template coordinate set was not performed. Structures were subjected to four cycles of high-temperature simulated annealing refinement (Nilges et al., 1988; Brünger, 1992). For the first two cycles, the initial temperature was 1,500 K, and 3,000 cooling steps were used, whereas for the last two, the initial temperature was 1,000 K, and 2,000 cooling steps were used. The three disulfide bonds in OMTKY3 were imposed as pseudo-NOE constraints for embedding and as covalent bonds during refinement. For each of the two data sets, 50 structures were calculated that had no NOE violations greater than 0.4 Å, no dihedral angle violations greater than 4°, and good covalent geometry. The resulting two sets of structures are referred to as the "crude" and "improved" structure sets.

Supplementary material in Electronic Appendix

The Electronic Appendix contains the following: one table giving NOE constraints for the "crude" data set, one table giving NOE constraints for the "improved" data set, one table giving dihedral angle constraints, and one table giving hydrogen bond constraints, all formatted for input into the X-PLOR/dg software; one table giving amide nitrogen and amide proton chemical-shift assignments for OMTKY3 as observed in this work; two files giving the Bruker DMX pulse sequence programs used in the BD-NOESY-HSQC and CBD-NOESY-HSQC experiments. All files in standard ASCII (MS-DOS platform).

Acknowledgments

This work was carried out at the National Magnetic Resonance Facility at Madison with support from NIH grants RR02301 and GM35976. Equipment in the NMR Facility was purchased with funds from the NIH National Center for Research Resources, Biomedical Research Technology Program (grant RR02301), the University of Wisconsin, the NSF Biological Instrumentation Program (grant DMB-8415048), the NIH Shared Instrumentation Program (grants RR02781 and RR08438), the NSF Academic Infrastructure Program (grant BIR-0214394), and the U.S. Department of Agriculture. C.G.H. was supported by a Howard Hughes Medical Institute Predoctoral Fellowship. The authors are grateful to Prof. Slobodan Macura of the Mayo Foundation for many helpful discussions and to Prof. Michael P. Williamson of MRC/Cambridge for providing software for chemical-shift calculations and a preprint of his paper referenced here. NMR data and structures reported in this work will be submitted to the appropriate macromolecular structure databases: Brookhaven Protein Data Bank (<http://www.pdb.bnl.gov/>, accession codes 1OMT and 1OMU for crude and improved structures, respectively) and BioMag ResBank (<http://www.bmr.b.wisc.edu/>).

References

- Bax A, Pochapsky SS. 1992. Optimized recording of heteronuclear multi-dimensional NMR spectra using pulsed field gradients. *J Magn Reson* 99:638–643.
- Bode W, Epp O, Huber R, Laskowski M Jr, Ardelt W. 1985. The crystal and molecular structure of the third domain of silver pheasant ovomucoid (OMSVP3). *Eur J Biochem* 147:387–395.
- Boelens R, Koning TMG, Kaptein R. 1988. Determination of biomolecular structures from proton–proton NOE's using a relaxation matrix approach. *J Mol Struct* 173:299–311.
- Borgias BA, James TL. 1990. MARDIGRAS—A procedure for matrix analysis of relaxation for discerning geometry of an aqueous structure. *J Magn Reson* 87:475–487.
- Brünger AT. 1992. *X-PLOR version 3.1: A system for X-ray crystallography and NMR*. New Haven, Connecticut: Yale University Press.
- Clore GM, Gronenborn AM. 1991. Structures of larger proteins in solution: Three- and four-dimensional heteronuclear NMR spectroscopy. *Science* 252:1390–1399.
- Clore GM, Gronenborn AM. 1993. Determination of structures of larger proteins in solution by three- and four-dimensional heteronuclear magnetic resonance spectroscopy. In: Clore GM, Gronenborn AM, eds. *NMR of proteins*. Boca Raton, Florida: CRC Press. pp 1–32.
- Clore GM, Robien MA, Gronenborn AM. 1993. Exploring the limits of precision and accuracy of protein structures determined by nuclear magnetic resonance spectroscopy. *J Mol Biol* 231:82–102.
- DelMar EG, Largman C, Brodrick JW, Geokas MC. 1979. A sensitive new substrate for chymotrypsin. *Anal Biochem* 99:316–320.
- Edison AS, Abildgaard F, Westler WM, Mooberry ES, Markley JL. 1994. Practical introduction to the theory and implementation of multinuclear, multidimensional NMR experiments. *Methods Enzymol* 239:3–79.
- Fejzo J, Zolnai Zs, Macura S, Markley JL. 1989. Analysis of laboratory-frame and rotating-frame cross-relaxation buildup rates from macromolecules. *J Magn Reson* 82:518–528.
- Fejzo J, Zolnai Zs, Macura S, Markley JL. 1990. Quantitative evaluation of two-dimensional cross-relaxation NMR spectra of proteins. Interpretation distances in turkey ovomucoid third domain. *J Magn Reson* 88:93–110.
- Fujinaga M, Sielecki AR, Read RJ, Ardelt W, Laskowski M Jr, James MNG. 1987. Crystal and molecular structures of the complex of α -chymotrypsin with its inhibitor turkey ovomucoid third domain at 1.8 Å resolution. *J Mol Biol* 195:397–418.
- Geen H, Freeman R. 1991. Band-selective radiofrequency pulses. *J Magn Reson* 93:93–141.
- Griesinger C, Ernst RR. 1987. Frequency offset effects and their elimination in NMR rotating-frame cross-relaxation spectroscopy. *J Magn Reson* 75:261–271.
- Habazettl J, Cieslar C, Oschkinat H, Holak TA. 1990. ¹H NMR assignments of sidechain conformations in proteins using a high-dimensional potential in the simulated annealing calculations. *FEBS Lett* 268:141–145.
- Havel TF, Wüthrich K. 1985. An evaluation of the combined use of nuclear magnetic resonance and distance geometry for the determination of protein conformation in solution. *J Mol Biol* 182:281–294.
- Hinck AP, Walkenhorst WF, Westler WM, Choe S, Markley JL. 1993. Overexpression and purification of avian ovomucoid third domains in *Escherichia coli*. *Protein Eng* 6:221–227.
- Hoogstraten CG, Westler WM, Macura S, Markley JL. 1993. Improved measurement of longer proton–proton distances in proteins by relaxation network editing. *J Magn Reson B* 102:232–235.
- Hoogstraten CG, Westler WM, Macura S, Markley JL. 1995a. NOE measurements in the absence of spin diffusion: Application to methylene groups in proteins and effects on local structural parameters. *J Am Chem Soc* 117:5610–5611.
- Hoogstraten CG, Westler WM, Mooberry ES, Macura S, Markley JL. 1995b. Analysis of transverse cross relaxation within a simplified relaxation submatrix: BD-ROESY. *J Magn Reson B* 109:76–79.
- James TL. 1994. Assessment of quality of derived macromolecular structures. *Methods Enzymol* 239:416–439.
- Kay LE, Nicholson LK, Delaglio F, Bax A, Torchia DA. 1992. Pulse sequences for removal of the effects of cross-correlation between dipolar and chemical-shift anisotropy relaxation mechanism on the measurement of heteronuclear T1 and T2 values in proteins. *J Magn Reson* 97:359–375.
- Krezel AM, Darba P, Robertson AD, Fejzo J, Macura S, Markley JL. 1994. Solution structure of turkey ovomucoid third domain as determined from nuclear magnetic resonance data. *J Mol Biol* 242:203–214.
- Laskowski M Jr, Kato I, Ardelt W, Cook J, Denton A, Empie MW, Kohr WJ, Park SJ, Parks K, Schatzley BL, Schoenberger OL, Tashiro M, Vichot G, Whatley HE, Wiczorek A, Wiczorek M. 1987. Ovomucoid third domains from 100 avian species: Isolation, sequences, and hyper-

- variability of enzyme-inhibitor contact residues. *Biochemistry* 26:202-221.
- Macura S, Farmer BT II, Brown LR. 1986. An improved method for the determination of cross-relaxation rates from NOE data. *J Magn Reson* 70:493-499.
- Macura S, Fejzo J, Hoogstraten CG, Westler WM, Markley JL. 1992. Topological editing of cross-relaxation networks. *Israel J Chem* 32:245-256.
- Macura S, Westler WM, Markley JL. 1994. Two-dimensional exchange spectroscopy of proteins. *Methods Enzymol* 239:106-144.
- Neuhaus D, Williamson M. 1989. *The nuclear Overhauser effect in structural and conformational analysis*. New York: VCH.
- Nilges M, Clore GM, Gronenborn AM. 1988. Determination of three-dimensional structures of proteins from interproton distance data by hybrid distance geometry-dynamical simulated annealing calculations. *FEBS Lett* 229:317-324.
- Peng JW, Thanabal V, Wagner G. 1991. 2D heteronuclear NMR measurements of spin-lattice relaxation times in the rotating frame of X nuclei in heteronuclear HX spin systems. *J Magn Reson* 94:82-100.
- Read RJ, Fujinaga M, Sielecki AR, James MNG. 1983. Structure of the complex of *Streptomyces griseus* protease B and the third domain of the turkey ovomucoid inhibitor at 1.8 Å resolution. *Biochemistry* 22:4420-4433.
- Rhyu GI, Markley JL. 1988. Two-dimensional NMR studies of Kazal proteinase inhibitors. 2. Sequence specific assignments and secondary structure of reactive site modified turkey ovomucoid third domain. *Biochemistry* 27:2529-2539.
- Robertson AD, Westler WM, Markley JL. 1988. Two-dimensional NMR studies of Kazal proteinase inhibitors. 1. Sequence-specific assignments and secondary structure of turkey ovomucoid third domain. *Biochemistry* 27:2519-2529.
- Shaka AJ, Lee CJ, Pines A. 1988. Iterative schemes for bilinear operators; application to spin decoupling. *J Magn Reson* 77:274-293.
- Szyperski T, Luginbühl P, Otting G, Güntert P, Wüthrich K. 1993. Protein dynamics studied by rotating frame ^{15}N spin relaxation times. *J Biomol NMR* 3:151-164.
- Torchia DA, Nicholson LK, Cole HBR, Kay LE. 1993. Heteronuclear NMR studies of the molecular dynamics of staphylococcal nuclease. In: Clore GM, Gronenborn AM, eds. *NMR of proteins*. Boca Raton, Florida: CRC Press, Inc. pp 190-219.
- Walkenhorst WF, Krezel AM, Rhyu GI, Markley JL. 1994. Solution structure of reactive-site hydrolyzed turkey ovomucoid third domain by nuclear magnetic resonance and distance geometry methods. *J Mol Biol* 242:215-230.
- Williamson MP, Asakura T. 1993. Empirical comparisons of models for chemical-shift calculations in proteins. *J Magn Reson B* 101:63-71.
- Williamson MP, Kikuchi J, Asakura T. 1995. Application of ^1H NMR chemical shifts to measure the quality of protein structures. *J Mol Biol* 247:541-546.
- Wüthrich K. 1986. *NMR of proteins and nucleic acids*. New York: Wiley.
- Yip P, Case DA. 1989. A new method for refinement of macromolecular structures based on nuclear Overhauser effect spectra. *J Magn Reson* 83:643-648.
- Zhao D, Jardetzky O. 1994. An assessment of the precision and accuracy of protein structures determined by NMR: Dependence on distance errors. *J Mol Biol* 239:601-607.