Localized solution structure refinement of an F45W variant of ubiquitin using stochastic boundary molecular dynamics and NMR distance restraints

PAUL B. LAUB,^{1,2} SEPIDEH KHORASANIZADEH,^{1,3} AND HEINRICH RODER^{1,2}

¹ Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111

² Department of Biochemistry & Biophysics, University of Pennsylvania, Philadelphia, Pennsylvania 19104

³ Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104

(RECEIVED December 27, 1994; ACCEPTED February 28, 1995)

Abstract

The local structure within an 8-Å radius around residue 45 of a recombinant F45W variant of human ubiquitin has been determined using 67 interproton distance restraints measured by two-dimensional proton NMR. Proton chemical shift evidence indicates that structural perturbations due to the F45W mutation are minimal and limited to the immediate vicinity of the site of mutation. Simulated annealing implemented with stochastic boundary molecular dynamics was applied to refine the structure of Trp 45 and 10 neighboring residues. The stochastic boundary method allowed the entire protein to be reassembled from the refined coordinates and the outlying unrefined coordinates with little distortion at the boundary. Refinement began with four low-energy indole ring orientations of F45W-substituted wild-type (WT) ubiquitin crystal coordinates. Distance restraints were derived from mostly long-range NOE cross peaks with 51 restraints involving the Trp 45 indole ring. Tandem refinements of 64 structures were done using either (1) upper and lower bounds derived from qualitative inspection of NOE cross-peak intensities or (2) quantitative analysis of cross-peak heights using the program MARDIGRAS. Though similar to those based on qualitative restraint, structures obtained using quantitative NOE analysis were superior in terms of precision and accuracy as measured by back-calculated sixth-root *R* factors. The six-membered portion of the indole ring is nearly coincident with the phenyl ring of the WT and the indole NH is exposed to solvent. Accommodation of the larger ring is accompanied by small perturbations in the backbone and a 120° rotation of the χ^2 dihedral angle of Leu 50.

Keywords: molecular dynamics; nuclear Overhauser effect; protein structure; relaxation matrix; simulated annealing; site-directed mutagenesis

Ubiquitin has become an important model protein for studies of protein-folding mechanisms using thermodynamic, kinetic, and structural approaches (Briggs & Roder, 1992; Cox et al., 1993; Khorasanizadeh et al., 1993; Wintrobe et al., 1994). With only 76 residues and a monomeric globular structure, ubiquitin is small enough to permit a detailed structural characterization of its folding process. A 1.8-Å resolution crystal structure of ubiquitin has been reported (Vijay-Kumar et al., 1987) and proton NMR assignments have been independently determined by two groups (Di Stefano & Wand, 1987; Weber et al., 1987). The packing of three strands of a curved β -sheet against an α -helix creates an extensive hydrophobic core (Fig. 1; Kinsinger et al., 1994). The presence of a simple, well-developed hydropho-

bic core contributes to the remarkable stability ($T_m \approx 90^\circ\text{C}$ at pH 4.0) (Wintrobe et al., 1994) of ubiquitin and offers the opportunity to investigate the effects of mutational alteration of packing on folding and stability.

To facilitate equilibrium and kinetic folding studies, Khorasanizadeh et al. (1993) introduced a fluorescence probe, using cassette mutagenesis, to replace the mostly buried Phe 45 residue with a tryptophan. The F45W mutation produced no observable change in biological activity and reduced the stability of the folded structure by less than 0.4 kcal mol⁻¹. Comparison of two-dimensional ¹H NMR spectra of the wild-type (WT) and F45W variant showed only small chemical shift changes in the NH-C α H cross peak region for most backbone protons. Those resonances showing chemical shift differences greater than 0.1 ppm were limited to six residues, all located within 6 Å of C γ of residue 45. In addition, the NOE connectivity patterns of the α -helix and β -sheets agree with those of the WT (Di Stefano

Reprint requests to: Heinrich Roder, Institute for Cancer Research, Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, Pennsylvania 19111; e-mail: h_roder@fccc.edu.

Fig. 1. Ribbon diagram of ubiquitin showing side chains of residues in the vicinity of the site of mutation (Phe 45), generated with the program Molscript (Kraulis, 1991).

& Wand, 1987; Weber et al., 1987), indicating that the secondary structure is conserved in the mutant.

Given the importance of site-directed mutagenesis in current research on protein structure, folding, and function, there is an obvious need for efficient and accurate methods to determine the structural consequences of mutations. In previous NMR studies, the solution structure of each protein variant was determined *de novo* without reference to the WT structure. For example, Hurle et al. (1992) investigated the effects of mutations in bovine pancreatic trypsin inhibitor, using the laborious distance geometry and restrained molecular dynamics approach. Structural perturbations found were small (in some cases within the uncertainty of their structures) and limited to the vicinity of the sites of mutation. Hence, much of the data collection and refinement was ultimately applied to regions of the protein unaffected by mutation, regions whose structure was already known through the crystal or NMR structure of the WT protein.

In contrast, the refinement method presented here makes use of the known structure of the WT protein as a starting point and is applied locally about the site of mutation. Distance restraints are obtained from intensities of only those NOESY cross peaks originating from the changed residue or showing large differences in chemical shift with respect to WT spectra. Stochastic boundaries are used to diminish distortions at the interface of the refined region with the rest of the protein. Consequently, refinement becomes considerably faster and more efficient. The method is demonstrated here using F45W ubiquitin as an example. We expect this approach to be more generally applicable for determining the solution structure of protein variants with conservative amino acid substitutions.

To assess how the conversion of cross-peak intensities into distance restraints might affect the structure, two sets of 64 F45W structures were refined in tandem using restraints obtained (1) by qualitative interpretation of NOE intensities (Clore & Gronenborn, 1989) and (2) by quantitative analysis of intensities using the program MARDIGRAS (Borgias & James, 1990).

In the latter approach, spin diffusion is confronted directly by complete relaxation matrix analysis.

Results

WT ubiquitin is illustrated schematically in Figure 1 and Kinsinger 1. Phe 45 lies in a hydrophobic pocket composed of Ala 46, Ile 61, and Leu 67, yet its side chain retains 13% surface accessibility. This site of mutation is located in a turn-rich region of ubiquitin; in particular, residues 45–48, 51–54, and 62–65 all form reverse turns, whereas residues 56–59 form a single turn of 3_{10} -helix. According to the temperature factors reported by Vijay-Kumar et al. (1987), the turns at residues 51–54 and at residues 62–65 are among the most flexible regions of the protein. One edge of the pocket is lined by residues 64–68, which are part of the C-terminal strand of the β -sheet.

Conformational perturbations introduced by the mutation are expected to be reflected in changes in the chemical shifts of the backbone atoms in the vicinity of the mutation (Osapay & Case, 1991; Hurle et al., 1992). In the case of F45W, we found chemical shift changes larger than ± 0.1 ppm for backbone amide or α proton resonances (Fig. 2) of only six residues, Ala 46, Lys 48, Leu 50, Ile 61, Gln 62, and Ser 65, all of which are in the immediate vicinity of Phe 45 (Khorasanizadeh et al., 1993). The greatest differences (F45W – WT) in chemical shift were observed for the amide proton of Ala 46 (-0.38 ppm) and one of the γ methylene protons of Ile 61 (0.53 ppm). These protons are located directly above and below the plane of the phenyl ring in the WT structure. In contrast to the ring of Phe 45 in the WT structure, which undergoes rapid 180° rotations on the chemical shift time scale, the Trp ring in F45W exhibits only one set of six single-proton resonances, which indicates that the ring is in a unique orientation. These observations suggested to us the feasibility of restricting structural refinement to the region about the site of mutation.

Using restrained molecular dynamics (r-MD), our refinement focused on the residues with at least one atom within $6 \text{ \AA}$ of residue 45. The effective radius, however, is approximately $8 \text{ \AA}$,

Fig. 2. Difference in chemical shifts between the F45W mutant and WT. Filled symbols represent amide protons; unfilled symbols represent α protons. Residues within $6 \text{ \AA}$ of the C_γ atom of Phe 45 are shown as diamonds; all other residues are shown as circles.

because all atoms of a residue were included if only one atom is within 6 Å. Residues with all atoms more than 9 Å distant were held fixed and removed from the refinement. Mediating the transition between the two regions was a stochastic boundary layer serving as a heat bath. Motion of residues in this layer were governed by the Langevin equation (Chandler, 1987). The stochastic boundary molecular dynamics used here is adapted from that developed for studying active-site dynamics in enzymes (Brooks et al., 1985; Brooks & Karplus, 1986, 1989).

Generation of initial structures for molecular dynamics

Initial structures for use with r-MD refinement were obtained by beginning with the WT coordinates (1UBQ) deposited by Vijay-Kumar et al. (1987) in the Brookhaven Protein Data Bank (Bernstein et al., 1977). Crystallographic water molecules, all of which are located on the surface, were discarded. The F45W mutation was introduced with molecular graphics. Subsequently, XPLOR 3.0 (Brünger, 1992) was used to build in hydrogens and perform a 16×16 -point grid search about the χ^1 and χ^2 dihedral angles of Trp 45 for low-energy conformations. As expected, for most conformations of the indole side chain, potential energy was very high, being dominated by Van der Waals energy resulting from steric clashes. Contours in Figure 3 describe low-energy regions found.

Four low-energy structures with distinct indole conformations (labeled A, B, C, and D in Fig. 3) were chosen for further refinement. The low-energy region of the χ^1 dihedral does not deviate far from that of the WT value, suggesting that the only favorable conformations for the indole ring are in the pocket previously occupied by the phenyl ring in the WT structure. By

Fig. 3. Potential energy contours calculated from incremental rotation of the χ^1 and χ^2 dihedral angles of Trp 45 in F45W ubiquitin. The six contours range from -235 to -150 kcal/mol. Filled circles labeled A, B, C, and D identify structures selected for further refinement. For reference, the coordinates of wild-type ubiquitin (labeled WT) and the fully-refined F45W mutant (labeled F45W) are also shown.

their χ^2 dihedrals, all four conformations are roughly coplanar with the WT Phe 45 ring; however, in structure A, the indole ring is flipped by about 180° relative to the other structures.

Each of the structures A–D obtained by grid search was subjected to stochastic boundary r-MD at 600 K. Sixteen coordinate sets were saved at 2-ps intervals for subsequent annealing using restraints derived by qualitative analysis of cross-peak intensity or by MARDIGRAS. The eight families of structures thus obtained are designated QA, QB, QC, QD, MA, MB, MC, and MD, where M (MARDIGRAS) and Q (qualitative) describe the restraint type and A–D describe the initial structure.

Derivation of distance restraints from NOESY cross-peak intensities

NOESY cross-peak intensities from a single mixing time of 150 ms were converted into lower and upper distance bounds for use as restraints in r-MD refinement by two different methods. Qualitative restraints were obtained by sorting intensities into strong, medium, and weak categories and assigning upper bounds of 2.7 Å, 3.3 Å, and 6.0 Å, respectively (Clare & Gronenborn, 1989; Clare et al., 1993). Given the long mixing time and the abundance of low-intensity cross peaks, we found it necessary to extend the upper bound for weak cross peaks from 5.0 Å to 6.0 Å. Lower bounds were set to 1.8 Å in all cases. For methylene pseudoatoms and for methyl groups, the upper bound was extended by 1.0 Å and by 1.5 Å, respectively (Liu et al., 1992).

Cross peaks were sorted according to normalized intensity with strong intensities defined to be between 0.2 and 1.0, medium between 0.1 and 0.2, and weak less than 0.1 (Table 1). Cutoffs between intensity classes were initially somewhat arbitrary. However, in the iterative refinement used, restraints whose upper bounds were systematically violated by 0.1 Å or more in all annealed structures were reclassified to the next weaker intensity class by adjustment of the strong–medium or medium–weak intensity cutoff. Through repeated annealing runs, intermediate structures were used to assess violations from weak-intensity cross peaks. Upper bounds of six restraints from very weak NOEs all involving the indole ring H δ 1 proton were consistently violated. Because the assignment of H δ 1 is not in question, spin diffusion across the plane of the indole ring is apparently the cause. These restraints were therefore removed from the restraint set used in later annealing runs.

The iterative approach used to correctly classify cross-peak intensities was also employed to assign stereospecifically the geminal methyls of Leu 50 and Leu 67. Theoretical intensities of intermediate structures in the refinement were back calculated using CORMA 3.0 (Keepers & James, 1984; Borgias & James, 1989) and compared to the measured intensities. It was not possible, however, to assign methylene protons in this way; consequently, all methylene proton pairs were represented by pseudoatoms placed at their center of mass.

A corresponding set of lower and upper bounds for the same 67 restraints was generated using the program MARDIGRAS 1.0 (Borgias & James, 1990). Using the experimental intensities and a database of the fixed interproton distances in proteins (e.g., the distance between the proton pair of a methylene group), MARDIGRAS converges to the maximally self-consistent relaxation rate matrix from which ideal distances are derived. Because the set of experimental intensities is never a complete

Table 1. Summary of NOE cross peaks and derived restraints

Method	Qualitative	MARDIGRAS
Number of NOE cross peaks used as restraints ^a :		
Strong ^b (0.2–1.0)	12	12
Medium (0.1–0.2)	17	17
Weak (<0.1)	52	52
Total	81	81
Distribution of restraints ^c :		
Upper bound <3 Å	2	2
3–4 Å	3	11
4–5 Å	23	11
5–6 Å	19	12
6–7 Å	20	31
Intraresidue	9	9
$i \rightarrow i + 1$	9	9
$i \rightarrow i + 2$	3	3
$i \rightarrow i + 3$	0	0
$i \rightarrow i + 4$	1	1
Long range	45	45
Total	67	67
Mean lower bound (Å)	1.80	2.99
Mean breadth (Å)	3.70	2.61
Mean upper bound (Å)	5.50	5.60

^a The number of restraints (67) is less than the number of NOE cross peaks used (81) because intensities from pairs of resolved, but not stereospecifically assigned methylenes were treated as a single cross peak having the mean intensity. Of the 67 restraints, 51 involved the indole ring and 35 involved methyl groups.

^b Intensities varied from 0.0103 to 1.000, where 1.0 is the strongest NOE cross-peak intensity used.

^c A listing of upper and lower restraint bounds is included in the Supplementary Material in the Electronic Appendix.

set, theoretical intensities from a trial structure are used by MARDIGRAS to fill in the intensity matrix. Distance bounds are obtained from the ideal distance by using a user-supplied estimate of the noise in the spectra. In general, the bounds are not symmetric with respect to the target distance; the upper bound is typically further from the target distance than the lower bound. Notably, MARDIGRAS uses the theory of Tropp (1980) to account for the effect of methyl rotation on relaxation rates, a feature particularly important here, given the abundance of NOEs involving methyl groups. MARDIGRAS has been employed in the refinement of DNA (Gochin & James, 1990; Stolarski et al., 1992) and peptide structures (Davis et al., 1993). In the latter case, a very similar conformation of the same ω -conotoxin GVIA peptide was independently obtained by Skalicky et al. (1993) using distance restraints derived from NOE buildup rates.

Because MARDIGRAS is largely, but not completely, independent of the trial structure used (Thomas et al., 1991), distance bounds were generated for each of the four initial structures, A–D. The degree of independence of target distances from trial structure is illustrated in Figure 4, where the two trial structures compared, A and C, have indole ring orientations inverted with respect to each other (cf. Fig. 3). In addition, MARDIGRAS was run using the correlation time of 4.1 ns rad^{-1} , measured for WT ubiquitin at 30 °C (Schneider et al.,

Fig. 4. Restraint target distances (A) before ($R^2 = 0.16$) and (B) after ($R^2 = 0.91$) use of MARDIGRAS. For each NOESY cross peak used to derive a restraint, the corresponding interproton distances measured in grid search structures A and C are plotted in A. The clustering of points near the diagonal in B demonstrates that target distances computed by MARDIGRAS are largely independent of the initial structure. Outlying points, shown as open circles in B, are associated with distances >10 Å in structure A. The corresponding intensity, back calculated by MARDIGRAS, is too small to be handled accurately.

1992), as well as 3.6 ns rad^{-1} and 4.6 ns rad^{-1} . From these 12 runs, the maximum upper bound and the minimum lower bound for each restraint were selected.

Local refinement with stochastic boundaries

Molecular dynamics involved the partitioning of the protein into three concentric regions centered on the C_γ carbon of Trp 45 in each of structures A–D such that all atoms of each residue are contained in the same single region.

Residues having at least one atom within 6.0 Å of the center define the inner region, which undergoes normal r-MD. The effective radius of the inner region is approximately 8 Å. These 11 residues (44–48, 50, 59, 61, and 65–67) are responsible for all but two of the cross peaks used in generating distance restraints. These residues also account for all residues in Van der Waals contact with Trp 45 and account for all residues showing a chemical shift difference between WT and F45W ubiquitin of >0.1 ppm. Surrounding the inner region is the stochastic boundary layer composed of 13 residues (3, 4, 23, 43, 49, 51, 56, 58, 60, 62–64, and 68) having at least one atom within 9.0 Å of the center and not part of the inner region. Atoms in the boundary layer are governed by the Langevin equation (cf. section 12.1.1 in Brünger [1992]); their equation of motion includes additional terms describing random fluctuations, velocity-

dependent frictional damping, and harmonic restraints on heavy atom positions, as described in detail by Brooks and Karplus (1989).

The remaining residues compose the outer, or reservoir, region. Located in regions of the protein unaffected by mutation and without distance restraints, the coordinates of these residues are not used in dynamics and annealing. The outer region atoms are implicitly represented as a potential of mean force, roughly approximated by the harmonic restraints applied to heavy atoms in the boundary layer.

Structures composed of the 24 residues of the inner region and boundary layer were annealed by reducing the temperature from 600 to 400 K in 5.0 ps (5,000 steps), from 400 to 200 K in 1.5 ps, and from 200 to nearly 0 K in 1.0 ps. Subsequently, 500 steps of conjugate gradient (CG) minimization were applied.

Complete structures containing all 76 residues were obtained by merging the annealed coordinates with the unmodified coordinates of residues in the outer region. Distortions at the interface, primarily for backbone bond angles in the buffer region, were removed by a brief 25-step CG minimization applied to each reassembled structure. For this purpose, the minimization did not go to convergence nor was it intended to. However, in 2 of the 48 structures annealed with qualitative restraints and in 3 structures annealed with MARDIGRAS restraints, inversion of chirality at a buffer region C α carbon could not be corrected; such structures were discarded. Overall, minimization resulted in minimal changes in heavy atom coordinates (~ 0.03 Å).

Assessment of annealing

Success of annealing is measured by reduction of the RMS restraint violation and of the sixth-root *R* (*R*6th) quality factor

(Thomas et al., 1991) when comparing values for each of the four grid search structures, A–D, to the mean values of annealed structures (Table 2). The *R*6th factor is defined as

$$\sum_i \left| (a_o^i)^{1/6} - (a_c^i)^{1/6} \right| / \sum_i (a_o^i)^{1/6},$$

where a_o^i is the observed intensity and a_c^i is the corresponding calculated intensity. In the summations, *i* indexes those intensities from which restraints were derived.

Both the RMS violations and the *R*6th quality factor assess the fit of structures to the NOESY data. However, in comparing back-calculated NOEs from the annealed structure with their measured intensities, the *R*6th quality factor is a more direct measure of accuracy and, unlike the RMS violations, is uninfluenced by the breadth of distance restraints. Rather than constructing an *R*-factor in direct analogy to the crystallographic *R*-factor, we have chosen to use the sixth-root *R*-factor in order to provide greater consideration of weak intensities representing longer distances (Thomas et al., 1991).

Both in terms of RMS violations and *R*6th quality factor, the MA family of annealed structures is distinctly inferior to all other structures (Table 2). The substantially elevated mean potential energy (excluding the NOE penalty energy) and deviations in their mean RMS bond length from equilibrium indicate that the MA structures are badly strained.

The indole ring conformation of structure A is inverted with respect to structures B–D. In Figure 3, structure A ($\chi^2 = 67^\circ$) lies above the horizontal high-energy ridge, whereas structures B–D ($\chi^2 = -93^\circ$ to -155°) lie below. Inspection of the structures by molecular graphics suggests that in the case of the MA structures, annealing became frustrated in attempting to flip the

Table 2. Structural statistics by distance bounds method

	Initial structure ^a	Refined structures	
		Qualitative	MARDIGRAS
Potential energy ^b (kcal/mol)	A	-225 ± 8	-92 ± 31
	B–D	-225 ± 2	-228 ± 3
Mean bond length	A	0.0114	0.0151
RMSD (Å)	B–D	0.0115	0.0116
Mean RMS restraint violation (Å) ^c	A	1.363	0.196
	B	1.043	0.104
	C	0.650	0.093
	D	0.747	0.094
	Average (B–D)	0.083	0.097
<i>R</i> 6th quality factor (Å)	A	0.159 ± 0.003	0.194 ± 0.006
	B	0.167 ± 0.002	0.125 ± 0.002
	C	0.159 ± 0.007	0.129 ± 0.004
	D	0.160 ± 0.003	0.130 ± 0.002
	Average (B–D)	0.162 ± 0.006	0.128 ± 0.003
Mean pairwise RMS (Å) ^d	B, C, D	0.88 ± 0.24	0.68 ± 0.20

^a A, B, C, and D refer to the indole ring conformations of structures selected from the grid search.

^b Calculated from the four lowest-energy structures annealed from each of A–D. The energy excludes contributions from the distance restraints.

^c Values of RMS restraint violation for structures A, B, C, and D are with respect to the qualitative restraints. Corresponding values with respect to MARDIGRAS restraints are similar.

^d Excludes atoms from residues not refined.

ring, leaving χ^1 ($-134 \pm 2^\circ$) and χ^2 ($-85 \pm 4^\circ$) dihedrals for Trp 45 distant from both those of the initial structure A ($\chi^1 = 165^\circ$, $\chi^2 = 67^\circ$) and those of families MB, MC, and MD ($\chi^1 = -167^\circ$, $\chi^2 = -115^\circ$, labeled F45W in Fig. 3). Surprisingly, with the looser qualitative restraints, all 16 structures of the QA family annealed to conformations similar to families QB, QC, and QD. In each of two 600 K r-MD runs, the ring inversion occurred early in the high temperature r-MD run before any of the 16 sets of QA coordinates for annealing were saved. Apparently, during the 600 K dynamics, the qualitative restraints were sufficient to induce a ring flip but loose enough to allow the flip without a prohibitive energetic barrier.

For the statistical analysis of the annealed structures (Tables 2, 3), we thus consider the MA (and QA) families separately. All structures from families MB, MC, and MD were pooled and are compared with the corresponding structures (QB, QC, and QD) obtained using the qualitative restraints. To a first approximation, the pooled structures obtained using qualitative and MARDIGRAS restraints are very similar, particularly in their mean potential energies (-225 versus -228 kcal mol $^{-1}$, respectively), mean RMS restraint violation (0.083 versus 0.097 Å), and mean Trp 45 χ^1 (-172 versus -167°) and χ^2 (-127 versus -115°) dihedral angles. Inspection by molecular graphics shows that the primary difference is in the side-chain conformations for Leu 50 and Leu 67, which are well determined by many distance restraints.

Closer analysis shows structures from families MB, MC, and MD to be superior to their qualitative restraint counterparts both in terms of accuracy described by the *R*6th value (0.128 versus 0.162 , respectively) and in terms of precision described by the mean pairwise RMS deviation (RMSD) in coordinates (0.68 versus 0.88 Å). Use of the null hypothesis shows that in each case the difference in means is statistically significant ($p < 0.0001$). This correlation of accuracy and precision is also seen when comparing among the individual families MB, MC, and MD. Variability among them is small, as indicated by the size of the standard deviations for pooled MARDIGRAS structures. However, of the 12 MARDIGRAS structures lowest in potential energy, 10 are from MB. Likewise, when computed for MB, MC, and MD separately, the *R*6th factor is lowest for MB (0.125 Å, in comparison to 0.129 Å for MC and 0.130 Å for MD), as is

the pairwise RMSD in coordinates (0.47 Å, in comparison to 0.70 Å for MC and 0.55 Å for MD).

Structures refined with the qualitative restraints have lower mean RMS restraint violation (0.083 versus 0.097 Å), but this may reflect the wider mean breadth of the qualitative restraint set (Table 1). In addition, among the structures refined with MARDIGRAS restraints, two restraints are violated by 0.35 – 0.40 Å in all 48 structures of families MB, MC, and MD. Excluding these two restraints in annealing produced structures essentially indistinguishable from structures with the two restraints included. With the removal of the two restraints, the mean RMS restraint violation (0.058 Å) became less than that of structures obtained with qualitative restraints (0.083 Å). Oddly, removing the two restraints elevated the *R*6th quality factor (0.138) over that obtained (0.128) by keeping them in. For this reason, and for lack of any reason to discard them, these restraints were retained in the MARDIGRAS restraint set.

Conformation of the Trp 45 indole ring

In Figure 5 and Kinemage 2, the 10 lowest-energy structures from families MB, MC, and MD are aligned by least squares on the heavy atoms of the backbone. Although the indole ring is defined with high precision, variation among structures is primarily present in the loop composed of residues 45–48. According to Table 3, the side chain of Leu 67 is defined with lower precision. Disorder in the side chains of Lys 48 and Tyr 59 likely results in part from the absence of cross peaks (and thus restraints) involving these side chains. No cross peaks were observed for the side chain of Lys 48, whereas side-chain cross peaks of Tyr 59 were limited to the β protons. Interpretation of NOEs involving Tyr 59 is difficult because the ring rotates rapidly on the chemical shift time scale. Residues 62–64, located at the lower left of Figure 5 and Kinemage 2, show higher precision because they are located in the boundary layer and therefore experience damped stochastic motion. They are not considered to be part of the refinement.

Fig. 5. Bundle of 10 lowest-energy structures taken from families MB–MD. Residue 62–64, located at the lower left, show higher precision as they are located in buffer layer and therefore experience damped stochastic motion.

Table 3. Side-chain orientations (degrees) of refined F45W ubiquitin structures compared to WT crystal structure

		WT	Qualitative ^a	MARDIGRAS ^a
Phe/Trp 45	χ^1	179	-172 ± 1	-167 ± 3
	χ^2	-109	-127 ± 3	-115 ± 3
Leu 50	χ^1	-51	-43 ± 31	-79 ± 1
	χ^2	-173	123 ± 83	-60 ± 3
Ile 61	χ^1	-70	-78 ± 4	-71 ± 1
	χ^2	176	181 ± 3	163 ± 3
Leu 67	χ^1	-44	-64 ± 2	-61 ± 11
	χ^2	-171	175 ± 2	146 ± 25
RMSD with WT (Å)			0.87 ± 0.13	0.83 ± 0.13

^a Statistics exclude structures from families QA and MA.

Fig. 6. Superposition of the average F45W structure (filled) from Figure 5 with the X-ray structure for WT ubiquitin (open).

The larger aromatic side chain of Trp 45 compared to Phe 45 in the WT structure results in small perturbations of the backbone (Fig. 6; Kinemage 3). Among the side chains in close proximity to the indole ring (Ala 46, Leu 50 and 67, Ile 61), only the χ^2 dihedral of Leu 50 shows a rotamer change between the WT (-172.7°) and the mean F45W ($-60.0 \pm 3.0^\circ$) conformation (Table 3). The reason for the preference of this alternative rotamer in F45W is not clear, as the geminal methyls of Leu 50 can fit without obvious steric clashes into the F45W structures at both the WT value of the χ^2 dihedral and at the values found here.

The apparent changes in average conformation for the side chains of Lys 48 and Tyr 59, like their disorder, probably reflect the lack of restraints to these side chains to a greater extent than internal motion or real structural difference. In the WT crystal structure, hydrogen bonds exist between the ϵ -amino of Lys 48 and the carbonyl of Ala 46 and between the hydroxyl oxygen of Tyr 59 and nitrogen of Glu 51 (Vijay-Kumar et al., 1987). The stochastic boundary protocol works best where all atoms of a residue are partitioned into the same region. For Lys 48 and Tyr 59, the proximity of backbone atoms of these residues to the site of mutation results in the placement of the entire residue, including the unrestrained side chains, in the refinement region.

Comparing the 45 structures of families MB, MC, and MD with the WT structure, the greatest mean RMS change in backbone heavy atom coordinates is seen in the loop composed of residues 45–48 (Fig. 7). Perturbation of the backbone is not accompanied by the gain or loss of any backbone hydrogen bonds. As a result of the mutation, the surface accessibility of the side chain of residue 45 increases from 13% for Phe to $18 \pm 2\%$ for Trp.

Discussion

In a molecular dynamics simulation of WT ubiquitin at 300 K (Braatz et al., 1992), the backbone heavy atoms of residues 45–48 varied with a 0.92-Å RMSD over the 3,000-ps duration, indicating that this loop is one of the most flexible regions of the protein. These results are consistent with our finding that

among the refined residues, the greatest variation in backbone heavy atom coordinates appears in residues 45–48 (Fig. 7). The flexibility of this region may explain our observation that the Phe to Trp mutation is accommodated with little change in the structure and activity of the protein. Mutations having little effect on biological function and/or stability tend to be found in flexible, solvent-exposed parts of proteins (Matthews, 1993), whereas mutations involving buried side chains often result in impaired biological activity (e.g., Rennell et al., 1991).

With the exception of the side chains of Lys 48 and Tyr 59 for which there are no restraints, the residues in the inner refinement region are well defined. Of their side chains, only the χ^2 dihedral of Leu 50 shows a change in rotamer as a result of mutation. In their study of the structural effects of core mutations in T4 lysozyme, Baldwin et al. (1993) also found that mutations are accompanied primarily by small perturbations of the back-

Fig. 7. Mean RMS difference in backbone heavy atom coordinates between 45 F45W structures from families MB, MC, and MD and the WT structure. Residues composing the 6-Å inner region are denoted by filled bars; remaining residues are from the boundary layer.

bone rather than by rotameric changes of side chains. The indole ring of Trp 45, defined by 51 NOE restraints, is coplanar with the phenyl ring in the WT structure and fits into the same hydrophobic pocket with the indole NH exposed to solvent. Residues showing the greatest change in chemical shift due to mutation, Ala 46 and Ile 61 (Fig. 2), lie directly above and below the indole ring.

Progress toward a solution structure of WT ubiquitin has been reported by several groups (Weber et al., 1988; Schneider et al., 1992); however, actual coordinates are not available. For this reason, the crystal coordinates of Vijay-Kumar et al. (1987) were used as a starting point for the refinement of F45W ubiquitin presented here. There are strong indications that the structure of ubiquitin in solution is very similar to that obtained by crystallography. For example, in their analysis of sequential and cross-strand NOE intensities, Weber et al. (1987) found complete agreement of the secondary structure with that in the 1.8-Å crystal structure. Moreover, qualitative interpretation of NOEs to the His, Tyr, and two Phe rings indicates that their extent of burial is similar to that in the crystal structure. In a preliminary report, Weber et al. (1988) extended their NMR refinement to include the side chains and confirmed the close agreement to the crystal structure.

In their study of the backbone dynamics and structure of ubiquitin, Schneider et al. (1992) found a general correspondence with the crystal structure, but also noted several subtle differences in backbone hydrogen bonding patterns. Significant among these are the hydrogen bonds found between O45 and NH47 and between O53 and NH23. For the former, differences in hydrogen bonding patterns may reflect the flexibility of the loop composed of residues 45–48. For the latter, in both WT ubiquitin and in the F45W structures, the carbonyl group of Gly 53 is oriented toward the solvent and unable to hydrogen bond to the helix containing Ile 23.

More generally, in a survey of 12 proteins whose structures were independently determined by both NMR and single crystal diffraction, Billeter (1992) found a high degree of similarity between structures produced by each method. Where they were noted, discrepancies were attributed to genuine disorder (as for side chains on the surface), improper use of H-bond restraints in the NMR refinement, perturbations from small co-crystallizing ions, and from crystal lattice contacts. In an analogous survey, Wagner et al. (1992) conclude similarly that the primary differences between corresponding NMR and crystal structures of small proteins are primarily found on the surface. As we have noted, disorder or a lack of restraints may explain the different conformations of the Lys 48 and Tyr 59 side chains in the WT and F45W structures; we cannot at present distinguish the two.

We have found that the MARDIGRAS method with its narrower mean widths of restraints (Table 1) results in annealed structures with improved precision and accuracy (Table 2) compared to the conventional qualitative analysis of NOE intensities. Using the exceptionally precise NMR structure of the 56-residue immunoglobulin G protein (Gronenborn et al., 1991) as a model system for studying the limits of accuracy and precision in protein structures determined by NMR, Clore et al. (1993) found that the ratio of precision to accuracy is independent of the number and breadth of interproton distance restraints. Though based on a different measure of accuracy, our results do support correlation of accuracy and precision.

In contrast to the report by Clore et al. (1993), we find that use of quantitatively determined distance restraints (e.g., as determined by MARDIGRAS) do result in a statistically significant improvement in the accuracy and precision of the resulting structures. This may be due to the fact that our restraint set is heavily biased by long-range NOEs connecting protons on the indole ring to neighboring residues (Table 1) and is hardly representative of a restraint set used in refining an entire protein. Consequently, the mean breadth of the qualitative restraints is very large, extending from 1.8 Å to a mean upper bound of 5.50 Å. With restraints this loose, it is not surprising that the narrower MARDIGRAS restraints produced a statistically different ensemble of annealed structures. MARDIGRAS-derived restraints differ little from the qualitative restraints in the mean upper bound (5.60 versus 5.50 Å, respectively), but their average lower bound is significantly higher, resulting in a narrower distribution.

Overall, our experience with MARDIGRAS resembles that of Thomas et al. (1991). Using simulated NOE intensity data, they found that structures generated by restrained distance geometry improved in accuracy as the mean restraint breadth became narrower, going from a qualitative determination of restraints (widest) to restraints resulting from MARDIGRAS (narrowest). These results emerge whether accuracy is computed as atomic RMSDs with the "true" structure (i.e., the structure from which the simulated intensities were back calculated) or by direct comparison of intensities using any of four different *R* factors.

The utility of local structural refinement has been demonstrated here for a mutant protein with a conservative amino acid change. Use of the WT structure as an initial structure permits one to limit the assignment and integration of NOESY cross peaks to those within about 6 Å of the site of mutation. For the F45W variant, this allowed us to focus our attention on the unambiguous and well-resolved cross peaks involving indole ring protons. That similar sets of structures were obtained using distance bounds derived qualitatively and by using MARDIGRAS demonstrates the robustness of refinement. However, it is clear that use of MARDIGRAS produces structures with better accuracy and precision. Distortion and strain at the interface between the region of the protein subjected to refinement and that held fixed has been avoided through use of stochastic boundaries. We anticipate that local refinement will be especially useful to those studying the effects of mutations on stability, folding, and function of proteins by NMR.

Materials and methods

Sample preparation and NMR spectroscopy

The F45W variant of ubiquitin was constructed by cassette mutagenesis (Ecker et al., 1987) and purified according to the procedures of Khorasanizadeh et al. (1993). Proton NMR spectra were recorded at 600.13 MHz on a Bruker AM 600 spectrometer. Phase-sensitive NOESY spectra were collected at 30 °C, pH 5.8, 4 mM protein concentration in 90% H₂O/10% D₂O with a mixing time of 150 ms. Felix (Biosym, San Diego, California) was used to process the spectra, which consisted of 750 *t*/2 increments of 2,048 complex point data sets, each the average of 64 acquisitions (Khorasanizadeh et al., 1993). A 60° shifted sine bell was used in processing with the final data matrix having 2,048 by 2,048 real points for a resolution of 3.6 Hz

per point. Proton resonance assignments for F45W ubiquitin are supplied as Supplementary Material in the Electronic Appendix.

Trp 45 indole placement by grid search

XPLOR 3.0 (Brünger, 1992) was used to rotate each of the χ^1 and χ^2 dihedrals in 22.5° increments, generating a total of 256 structures. XPLOR lacks built-in capability for dihedral rotation. Consequently, rotation was accomplished by fixing all atoms but those to be rotated, applying a dihedral angle restraint placed about the target angle measure, and performing rigid body minimization to convergence. Nonbonded terms were excluded from the potential energy function so that atoms could sample all conformations, including those otherwise sterically inaccessible. Following rotation, 50 steps of CG minimization were applied to all atoms using a potential energy function with the Van der Waals and electrostatic terms now included and with the dihedral restraint excluded. Slight atomic shifts (typically <0.2 Å) during the second minimization made the low-energy region of χ^1 - χ^2 dihedral space (Fig. 3) substantially larger and, we believe, more realistic. In neither minimization were distance restraints used.

MARDIGRAS and CORMA

MARDIGRAS used the intensity of the HD1 HE1 proton pair on the indole ring for normalization and an absolute normalized noise level of 0.010. Normalization on the basis of all intensities produced similar bounds. CORMA was run with a single isotropic correlation time of 4.1 ns rad⁻¹ (Schneider et al., 1992). Both programs used the methyl jump option 3, the most precise and realistic representation of methyl motion. On a 20-MHz Silicon Graphics 4D25TG workstation, a single run of CORMA consumed ca. 32 CPU min for the 630 protons of F45W, whereas a single MARDIGRAS run, using five iterations to converge, required ca. 270 CPU min.

Stochastic boundary layer

Langevin dynamics in the stochastic boundary layer used a friction coefficient of 50 ps⁻¹ for all heavy atoms and 0 ps⁻¹ for hydrogen (Brooks et al., 1985). Force constants for the harmonic restraints were calculated on the basis of averaged mean square atomic fluctuations for backbone heavy atoms (0.410 Å²) and side-chain heavy atoms (0.732 Å²) (Brooks & Karplus, 1989). Single isolated residues as well as contiguous stretches of residues in the inner region should be surrounded or terminated by residues in the boundary layer; otherwise, unrealistically large displacements of the backbone atoms of untethered inner region residues may occur. Consequently, several residues (51, 58, 63, and 64) that were part of the unrefined outer region were re-partitioned into the boundary layer. The XPLOR input file implementing annealing with stochastic boundaries is included as Supplementary Material in the Electronic Appendix.

Distance restraints

Distance restraints were implemented using a potential of zero between the lower and upper bounds and quadratically increasing potential outside. A force constant of 20 kcal mol⁻¹ Å⁻²

was applied throughout structure generation dynamics, annealing, and CG minimization.

Restraints involving methylene and methyl spin groups used $\langle r^{-6} \rangle$ distance averaging. In addition, methyl group restraints computed with MARDIGRAS had their upper bounds reduced by 0.08 to 0.22 Å and their lower bounds reduced by 0.17–0.28 Å. These modifications were necessary because MARDIGRAS accounts for methyl rotation using the theory of Tropp (1980), in contrast to XPLOR. To a first approximation, methyl rotation affects cross relaxation so that distance averaging is effectively $\langle r^{-3} \rangle$. The small subtractive modifications of restraint bounds, then, are meant to make the MARDIGRAS bounds commensurate with the shorter average distances computed by the $\langle r^{-6} \rangle$ averaging of XPLOR 3.0. XPLOR 3.1 now includes an $\langle r^{-3} \rangle$ averaging option. The corrections, derived from a computer model of Tropp's equations and consistent with related data presented in Table 1 of Liu et al. (1992), are a function of both distance and orientation with respect to the methyl group. The orientation dependence is conservatively accounted for by using the minimum orientation-dependent correction for the upper bound and the maximum correction for the lower bound.

Nonbonded interactions

In order to better sample conformational space during dynamics, the electrostatics component of the potential energy function was removed, the standard Lennard-Jones component was replaced by a quartic, purely repulsive component, and the atomic van der Waals radii were uniformly multiplied by 0.8 (Nilges et al., 1988; Clore et al., 1993). In the CG minimizations following annealing, the standard representation of nonbonded interactions, using a dielectric constant of 4.0, was employed. The all-hydrogen topology and parameter sets topallh6x.pro and parmallh3x.pro distributed with XPLOR 3.0 were used throughout.

Miscellaneous

The contours of Figure 3 were created with the in-house program cntour.f based on the algorithm of Watson (1992). Watson's algorithm is notable because it directly contours unevenly spaced data points. In the present case, the regular spacing of points in the grid search of χ^1 - χ^2 dihedral space was lost with the 50-step CG minimization applied to each structure after rotation. MOLSCRIPT (Kraulis, 1991) was used to create Figures 1, 5, and 6. Percent accessible surface area is the accessible surface area of the side chain in the presence of the protein divided by the accessible surface area of the side chain alone. Calculation was done on heavy atoms using XPLOR 3.0 and a probe radius of 1.6 Å. Two methods of Student's *t*-test were used in comparing structures obtained using restraints generated qualitatively and by MARDIGRAS. One method assumed (as is usual) equal variances, and the other (Press et al., 1992) removed this assumption. In all cases, both methods were in agreement about rejection of the null hypothesis at the significance level used (0.0001).

Acknowledgments

We thank Drs. Iain Peters (Hawaii Biotechnology Group, Inc.) for the preparation of the F45W recombinant clone, Tom James (UCSF) for

providing the source code for CORMA 3.0 and MARDIGRAS 1.0, He Liu (UCSF) for expert advice in the use of MARDIGRAS, and Robert Beckman (Biophysics Research Division, University of Michigan) for discussions on the representation of NOE intensities as distance restraints. We benefited from the resources of the DNA Synthesis Facility and the Research Computing Services of Fox Chase Cancer Center, supported by an NIH core grant CA06927. This work was supported by NIH grants GM44881 and GM35926, and an appropriation from the Commonwealth of Pennsylvania to the Institute for Cancer Research.

References

- Baldwin EP, Hajiseyediavadi O, Baase WA, Matthews BW. 1993. The role of backbone flexibility in the accommodation of variants that repack the core of T4 lysozyme. *Science* 262:1715–1718.
- Bernstein FC, Koetzle TF, Williams GJB, Meyer EF Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. 1977. The Protein Data Bank: A computer-based archival file for macromolecular structures. *J Mol Biol* 112:535–542.
- Billeter M. 1992. Comparison of protein structures determined by NMR in solution and by X-ray diffraction in single crystals. *Q Rev Biophys* 25:325–377.
- Borgias BA, James TL. 1989. Two-dimensional nuclear Overhauser effect: Complete relaxation matrix analysis. *Methods Enzymol* 176:169–183.
- 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.
- Braatz JA, Paulsen MD, Ornstein RL. 1992. 3 Nsec molecular dynamics simulation of the protein ubiquitin and comparison with X-ray crystal and solution NMR structures. *J Biomol Struct & Dyn* 9:935–949.
- Briggs MS, Roder H. 1992. Early hydrogen-bonding events in the folding reaction of ubiquitin. *Proc Natl Acad Sci USA* 89:2017–2021.
- Brooks CL III, Brünger AT, Karplus M. 1985. Active site dynamics in protein molecules: A stochastic boundary molecular-dynamics approach. *Biopolymers* 24:843–865.
- Brooks CL III, Karplus M. 1986. Theoretical approaches to the solvation of biopolymers. *Methods Enzymol* 127:369–400.
- Brooks CL III, Karplus M. 1989. Solvent effects on protein motion and protein effects on solvent motion: Dynamics of the active site region of lysozyme. *J Mol Biol* 208:159–181.
- Brünger AT. 1992. *X-PLOR (version 3.0) manual*. New Haven, Connecticut: Yale University.
- Chandler D. 1987. *Introduction to modern statistical mechanics*. New York: Oxford University Press.
- Clore GM, Gronenborn AM. 1989. Determination of three-dimensional structures of proteins and nucleic acids in solution by nuclear magnetic resonance spectroscopy. *CRC Crit Rev Biochem* 24:479–564.
- 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.
- Cox JPL, Evans PA, Packman LC, Williams DH, Woolfson DN. 1993. Dissecting the structure of a partially folded protein; circular dichroism and NMR studies of peptides from ubiquitin. *J Mol Biol* 234:483–492.
- Davis JH, Bradley EK, Miljanich GP, Nadasdi L, Ramachandran J, Basus VJ. 1993. Solution structure of ω -conotoxin GVIA using 2-D NMR spectroscopy and relaxation matrix analysis. *Biochemistry* 32:7396–7405.
- Di Stefano DL, Wand AJ. 1987. Two-dimensional ^1H NMR study of human ubiquitin: A main chain directed assignment and structure analysis. *Biochemistry* 26:7272–7281.
- Ecker DJ, Butt TR, Marsh J, Sternberg EJ, Morgolis N, Monia BP, Jonnalagadda S, Khan MI, Weber PL, Mueller L, Crooke ST. 1987. Gene synthesis, expression, structures, and functional activities of site-specific mutants of ubiquitin. *J Biol Chem* 262:14213–14221.
- Gochin M, James TL. 1990. Solution structure studies of $d(\text{AC})_4 \cdot d(\text{GT})_4$ via restrained molecular dynamics simulations with NMR constraints derived from two-dimensional NOE and double-quantum-filtered COSY experiments. *Biochemistry* 29:11172–11180.
- Gronenborn AM, Filpula DR, Essig NZ, Achari A, Whitlow M, Wingfield PT, Clore GM. 1991. A novel, highly stable fold of immunoglobulin binding domain of streptococcal protein G. *Science* 253:657–661.
- Hurle MR, Eads CD, Pearlman DA, Seibel GL, Thomason J, Kosen PA, Kollman P, Anderson S, Kuntz ID. 1992. Comparison of solution structures of mutant bovine pancreatic trypsin inhibitor proteins using two-dimensional nuclear magnetic resonance. *Protein Sci* 1:91–106.
- Keepers JW, James TL. 1984. A theoretical study of distance determinations from NMR two-dimensional nuclear Overhauser effect spectra. *J Magn Reson* 57:404–426.
- Khorasanizadeh S, Peters ID, Butt TR, Roder H. 1993. Folding and stability of a tryptophan-containing mutant of ubiquitin. *Biochemistry* 32:7054–7063.
- Kraulis PJ. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr* 24:946–950.
- Liu H, Thomas PD, James TL. 1992. Averaging of cross-relaxation rates and distances for methyl, methylene, and aromatic ring protons due to motion or overlap. Extraction of accurate distances iteratively via relaxation matrix analysis of 2D NOE spectra. *J Magn Reson* 98:163–175.
- Matthews BW. 1993. Structural and genetic analysis of protein stability. *Annu Rev Biochem* 62:139–160.
- Nilges M, Clore GM, Gronenborn AM. 1988. Determination of three-dimensional structure of proteins from interproton distance data by hybrid distance geometry-dynamical simulated annealing calculations. *FEBS Lett* 229:317–324.
- Osapay K, Case DA. 1991. A new analysis of proton chemical shifts in proteins. *J Am Chem Soc* 113:9436–9444.
- Press WH, Teukolsky SA, Vetterling WT, Flannery BP. 1992. *Numerical recipes in C, 2nd ed.* New York: Cambridge University Press.
- Rennell D, Bouvier SE, Hardy LW, Poteete AR. 1991. Systematic mutation of bacteriophage T4 lysozyme. *J Mol Biol* 222:67–88.
- Schneider DM, Dellwo MJ, Wand AJ. 1992. Fast internal main-chain dynamics of human ubiquitin. *Biochemistry* 31:3645–3652.
- Skalicky JJ, Metzler WJ, Ciesla DJ, Galdes A, Pardi A. 1993. Solution structure of the calcium channel antagonist ω -conotoxin GVIA. *Protein Sci* 2:1591–1603.
- Stolarski R, Egan W, James TL. 1992. Solution structure of EcoRI DNA octamer containing 5-fluorouracil via restrained molecular dynamics using distance and torsion angle constraints extracted from NMR spectral simulations. *Biochemistry* 31:7027–7042.
- Thomas PD, Basus VJ, James TL. 1991. Protein solution structure determination using distances from two-dimensional nuclear Overhauser effect experiments: Effects of approximations on the accuracy of derived structures. *Proc Natl Acad Sci USA* 88:1237–1241.
- Tropp J. 1980. Dipolar relaxation and nuclear Overhauser effects in non-rigid molecules: The effect of fluctuating internuclear distances. *J Chem Phys* 72:6035–6043.
- Vijay-Kumar S, Bugg CE, Cook WJ. 1987. Structure of ubiquitin refined at 1.8 Å resolution. *J Mol Biol* 194:531–544.
- Wagner G, Hyberts SG, Havel TF. 1992. NMR structure determination in solution: A critique and comparison with crystallography. *Annu Rev Biophys Biomol Struct* 21:167–198.
- Watson DF. 1992. *Contouring: A guide to the analysis and display of spatial data*. New York: Pergamon Press.
- Weber PL, Brown SC, Mueller L. 1987. Sequential ^1H NMR assignments and secondary structure identification of human ubiquitin. *Biochemistry* 26:7282–7290.
- Weber PL, Ecker DJ, Marsh J, Crooke ST, Mueller L. 1988. NMR-derived structures of ubiquitin and ubiquitin mutants. *Trans ACA* 24:91–105.
- Wintrodde PL, Makhatazde GI, Privalov PL. 1994. Thermodynamics of ubiquitin unfolding. *Proteins Struct Funct Genet* 18:246–253.