PROTEIN STRUCTURE REPORT

Solution structure of lysine-free (K0) ubiquitin

Tao Huang, Jess Li, and R. Andrew Byrd*

Structural Biophysics Laboratory, Center for Cancer Research, National Cancer Institute, Frederick, Maryland 21702

Received 30 October 2013; Accepted 28 February 2014

DOI: 10.1002/pro.2450

Published online 3 March 2014 proteinscience.org

Abstract: Lysine-free ubiquitin (K0-Ub) is commonly used to study the ubiquitin-signaling pathway, where it is assumed to have the same structure and function as wild-type ubiquitin (wt-Ub). However, the K0-Ub ^{15}N heteronuclear single quantum correlation NMR spectrum differs significantly from wt-Ub and the melting temperature is depressed by 19°C , raising the question of the structural integrity and equivalence to wt-Ub. The three-dimensional structure of K0-Ub was determined by solution NMR, using chemical shift and residual dipolar coupling data. K0-Ub adopts the same backbone structure as wt-Ub, and all significant chemical shifts can be related to interactions impacted by the K to R mutations.

Keywords: ubiquitin; K0-Ub; NMR; CS-Rosetta

Introduction

Protein ubiquitination is a posttranslational modification that regulates almost all aspects of cell life. In this process, ubiquitin (a 76 amino acid protein) is covalently attached to a lysine on the targeted proteins through an isopeptide bond formed between the carboxylate of the C-terminal glycine residue of ubiquitin and a free N^ϵ of a lysine of the substrate. Subsequently, a lysine residue on the first ubiquitin can covalently link through another isopeptide bond to a second ubiquitin and so on to create a polyubiquitin chain on target proteins. There are seven lysine residues in ubiquitin at positions 6, 11, 27, 29, 33, 48, and 63, which can be used in isopeptide link-

ages and create many types of polyubiquitinated proteins. These polyubiquitin chains can have different topologies, can vary in length, and have different functional consequences.¹ Functional investigations of ubiquitin pathways often utilize lysine-substituted variants, which have either a single lysine or all lysines substituted with arginine.² These variants, particularly lysine-free ubiquitin (K0-Ub), are used to define specific ubiquitin modifications and for detecting ubiquitination of specific substrates.^{3–7} In spite of its wide applications in biology and biochemistry, no structural study has been reported for K0-Ub, although structural equivalence between wild-type ubiquitin (wt-Ub) and single K-to-R variants^{8,9} have been reported. As lysine represents almost 10% of the residues of ubiquitin, it is of interest whether the substitution alters the three-dimensional (3D) fold or stability of ubiquitin. Furthermore, as NMR is widely used in ubiquitination studies, including the mutants of ubiquitin^{8–15} and conformations of polyubiquitin chains,^{8,16–18} it is important to characterize the NMR spectra of K0-Ub and validate the solution structure. Hence, we have used solution

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: Intramural Research Program of the National Institutes of Health; Grant sponsor: National Cancer Institute; Grant sponsor: Center for Cancer Research.

*Correspondence to: R. Andrew Byrd; Structural Biophysics Laboratory, Center for Cancer Research, National Cancer Institute, Frederick, MD 21702. E-mail: byrdr@mail.nih.gov

NMR to characterize the 3D structure of K0-Ub and interpret the unexpected and dramatic differences in NMR spectra.

Results and Discussion

NMR spectroscopy of K0-Ub

The ^{15}N - ^1H heteronuclear single quantum correlation (HSQC) spectrum was collected for K0-Ub and compared with wild-type ubiquitin [Fig. 1(A)]. K0-Ub showed a well-dispersed HSQC spectrum with uniform peak intensity and the expected number of backbone amide peaks, indicating that K0-Ub adopts an ordered 3D structure. However, the HSQC spectrum of K0-Ub is substantially different from wt-Ub, with more than half of the peaks shifting to an extent as to be unreliably traceable from wt-Ub to K0-Ub. As the ^{15}N - ^1H HSQC spectrum represents the chemical environment of each amino acid residue and is considered a “fingerprint” of the protein structure, the large difference in HSQC spectrum might come from significant differences in the chemical environment and raise the question as to whether K0-Ub adopts a distinct 3D structure.

The NMR chemical shifts for the backbone atoms (H^{N} , N , C , C^{α} , C^{β} , H^{α} , and H^{β}) of K0-Ub were assigned by standard triple resonance experiments. For 70 of the 72 observed N - H^{N} peaks in the ^{15}N - ^1H HSQC spectrum, complete resonance assignments were possible, shown in Supporting Information Figure S1. Residues E24 and G53 were not observable, due to exchange with water, as found in wild-type ubiquitin.¹⁹ The backbone chemical shift perturbation (CSP) was calculated between wt-Ub and K0-Ub, shown in Figure 1(B). The CSP values range from 0 to 0.6 ppm, and the CSP for 33 out of 66 conserved residues are larger than 0.1 ppm. A few residues which have large CSPs are adjacent to lysines, such as T7 (next to K6) and A28 (between K27 and K29), indicating that the lysine to arginine substitution imparts a nearest neighbor effect. In the case of T7, there is an additional effect from a backbone hydrogen bond between T7 and K11. If the strength of this hydrogen bond is impacted by the K11R mutation, then T7 could be expected to have an additional shift. In fact, the amide group of T7 shifts 0.36 ppm. Additionally, some of the lysine side chain $\text{N}^{\epsilon}\text{H}_3^+$ groups (e.g., K11, K27, K29, and K33) form extensive hydrogen bonds with backbone carbonyls and salt bridges with side chain carboxylates of Glu and Asp, which were observed from the X-ray crystal structure²⁰ and solution NMR studies²¹ (Supporting Information Table S1). The hydrogen bonds and salt bridges from lysine $\text{N}^{\epsilon}\text{H}_3^+$ are summarized in Figure 1(B). Many of the residues with large CSPs are hydrogen bond acceptors from lysine NH_3^+ , or have a secondary interaction with one of the affected residues or one residue adjacent. The lysine to argi-

Figure 1. Spectra comparison between ubiquitin and K0-Ub. (A) ^{15}N - ^1H HSQC spectra overlay of ubiquitin (blue) and K0-Ub (red). (B) CSP between ubiquitin and K0-Ub. Average CSP of ^1H and ^{15}N chemical shifts is calculated by the formula: $\text{CSP} = [(\Delta\delta_{\text{NH}}^2 + \Delta\delta_{\text{N}}^2/25)/2]^{1/2}$ (ppm), where $\Delta\delta_{\text{NH}}$ and $\Delta\delta_{\text{N}}$ are the amide proton and nitrogen shift differences for conserved residues between ubiquitin and K0-Ub, respectively. “KR” represents lysine to arginine mutations, “P” represents proline and residues 24 and 53 are missing due to solvent exchange. Hydrogen bonds are connected with gray lines and salt bridges are connected with black lines.

nine substitutions will alter the distances between donor ($\text{N}^{\epsilon}\text{H}_3^+$ becoming $\text{N}^{\eta 1}\text{H}_2$ or $\text{N}^{\eta 2}\text{H}_2$) and acceptor oxygen atoms, hence perturbing or abolishing the hydrogen bonds and salt bridges. It is also possible to have different rotamers of the arginine sidechain to achieve the hydrogen bond with one of the amino groups. For example, K27 is involved in a salt bridge with D52, which has a large CSP. Also, K27 participates in a backbone hydrogen bond with the I23 carbonyl oxygen. If the K27R mutation alters this hydrogen bond strength or direction slightly, then the CSP of I23 (0.22 ppm) could result. Similarly, K27 N^{ϵ} participates in a hydrogen bond to the backbone carbonyl of Q41. The Q41 NH correlation peak is not significantly shifted, but the adjacent R42 exhibits a large CSP (0.39 ppm). Careful examination of the observed CSPs in K0-Ub indicates that the majority of the shifted residues correlates with interactions impacted by the conversion of lysine to arginine or are adjacent to one of these interactions. Hence, the possibility arises that there could be substantive conformational affects due to

the simultaneous mutation of all seven lysine residues.

Solution structure of K0-Ub

The secondary structure of K0-Ub was determined from chemical shift assignments using PECAN²² and TALOS+²³ for K0-Ub and compared with ubiquitin (PDB ID: 1UBQ or 1D3Z). Both PECAN and TALOS+ reveal that K0-Ub adopts the same secondary structure as ubiquitin, with average deviations of phi and psi angles 3.4° and 5.9° from wt-Ub, respectively (Supporting Information Fig. S2). The 3D structure of K0-Ub was calculated using the chemical shift-based Rosetta (CS-Rosetta) method,²⁴ with the addition of H^N-N residual dipolar coupling (RDC) data. All assigned H^N, N, C, C^α, C^β, and H^α chemical shifts plus 52 N-H^N RDCs, representing a random selection of 80% of complete RDC data, were used in the calculations. The remaining 13 RDCs (20% of the data) were used for cross-validation. The RDC data significantly improved the convergence of the ensemble of structures, shown in Figure 2(A). The ensemble of the 20 lowest-energy structures [Fig. 2(B)] exhibited excellent Ramachandran statistics and had an overall backbone root mean squared deviation (RMSD) of 0.30 Å in the structured region (Table I). The best model, with the lowest Rosetta energy, is in good agreement with the RDC values withheld from the Rosetta calculation, and the RDC Q-factor is 0.20, determined using PALES.²⁵ The lowest energy model of K0-Ub has a backbone RMSD of 0.62 Å compared to the crystal structure of wt-Ub (PDB: 1UBQ), and the superposition of K0-Ub and wt-Ub is shown in Figure 2(C). Interestingly, 65 RDCs measured for K0-Ub fit very well to the solution structure of wt-Ub (PDB: 1D3Z), exhibiting a Q-factor of 0.15. These results confirm that K0-Ub adopts the same 3D structure as wild-type ubiquitin.

The sensitivity of the ¹⁵N-¹H HSQC spectrum to the hydrogen bonding and salt-bridge network within a protein has been observed previously,^{26,27} where the structure remains unchanged but the ¹⁵N-¹H chemical shifts are significantly affected. Despite the large number of chemical shift differences (50% of ubiquitin residues have CSP values larger than 0.1 ppm), K0-Ub retains the same structure as wild type. This situation might seem surprising in light of the widely held view that chemical shift affects imply either ligand binding or structural changes; rather, the data suggest that interpretation of amide H^N-N CSP data include consideration of hydrogen bonding perturbations. We have also observed a small population, generally <10–15%, of K0-Ub molecules with altered ¹⁵N-¹H chemical shifts, but the same C, C^α, C^β, H^α, and H^β chemical shifts. This occurs for many peaks in the HSQC spectrum. Assignment of these signals and analysis of the secondary structure from TALOS+

Figure 2. Solution structure of K0-Ub compared with wt-Ub. (A) Plot of rescored Rosetta all-atom energy versus C^α RMSD relative to the lowest-energy model (bold dot on vertical axis). (B) 20 lowest-energy CS-Rosetta models of K0-Ub superimposed. (C) Superposition of best K0-Ub model (green) with ubiquitin (orange; PDB: 1UBQ).

shows that the minor population adopts the same fold as the major population and these shifts represent alternate conformations in the hydrogen bonding and salt-bridge network, which occur in mostly random members of the molecular ensemble and not as a single minor species. See Supporting Information for further discussion.

Thermal stability of K0-Ub

The thermal stability of K0-Ub, assessed by the melting temperature (T_m), was examined using differential scanning calorimetry (DSC; Supporting Information Fig. S3). K0-Ub is a fairly thermostable protein, exhibiting a T_m of 71.9°C; however, this value is 19°C lower than that of wt-Ub (T_m = 90.6°C). The decreased T_m of K0-Ub is consistent with the loss or perturbation of hydrogen bonds and salt bridges that occur in the substitution of lysine to arginine. Nevertheless, K0-Ub remains a stable, folded analog of wt-Ub and should not represent any difficulties or substantively altered behavior in biochemical experiments.

Summary

We have examined the structure of K0-Ub by solution NMR methods and demonstrated that K0-Ub

Table I. Structural Statistics for the Ensemble of 20 Calculated Structures of K0-Ub

RMSD values	All	Ordered ^a	Selected ^b
All backbone atoms	1.4 Å	0.3 Å	0.3 Å
All heavy atoms	1.6 Å	0.8 Å	0.8 Å
Structure quality factors—overall statistics			
	Mean score	SD	Z-score ^c
Procheck G-factor ^a (phi/psi only)	−0.09	N/A	−0.04
Procheck G-factor ^a (all dihedral angles)	0.25	N/A	1.48
Verify 3D	0.44	0.0336	−0.32
ProsaII (−ve)	0.91	0.0385	1.08
MolProbity clashscore	4.06	1.7587	0.83
Ramachandran plot summary from procheck ^b			
Most favored region	95.5%		
Allowed region	4.5%		
Generously allowed region	0.0%		
Disallowed region	0.0%		

^a Residues with sum of phi and psi order parameters >1.8. Ordered residue ranges: 2–51, 54–71.

^b Residues selected based on: Dihedral angle order parameter, with S(phi) + S(psi) ≥ 1.8. Selected residue ranges: 2–51, 54–71.

^c With respect to mean and standard deviation for a set of 252 X-ray structures <500 residues, of resolution ≤ 1.80 Å, R-factor ≤ 0.25 and R-free ≤ 0.28; a positive value indicates a “better” score.

retains the same secondary and tertiary structure as wild-type ubiquitin. This is similar to findings for single K-to-R mutants^{8,9} and for cases that examined hydrophobic core residues,^{8,9,15} showing that the hydrophobic core of ubiquitin could be redesigned without altering the 3D structure. The structural equivalence is consistent with our observations that K0-Ub is readily (and equivalently to wt-Ub) loaded by the E1 ubiquitin activating enzyme onto E2 conjugating enzymes.²⁸ The chemical shift assignments and solution structure obtained in these studies enables confident, NMR-based monitoring of ubiquitination and conformational preferences of K0-Ub in general mechanistic studies of the ubiquitination pathway.

Materials and Methods

Protein expression and purification

The coding sequence of K0-Ub in a pET vector was the kind gift of Dr. Allan M. Weissman. Uniformly ¹⁵N- or ¹⁵N/¹³C-labeled K0-Ub were expressed in *Escherichia coli* (BL21DE) cells grown on M9 minimal medium enriched with 0.1% ¹⁵NHCl₄ alone or together with 0.3% uniformly ¹³C-labeled D-glucose. Protein expression was induced by adding 1 mM isopropyl β-D-1-thiogalactopyranoside at OD₆₀₀ ~ 0.8 for 3 h at 37°C. *E. coli* cells were harvested and lysed in 50 mM Tris, pH 7.4 using a microfluidizer. Soluble cell lysate were passed over a Q-sepharose column (GE Healthcare), the flow-through containing K0-Ub was collected and concentrated. K0-Ub was further purified using a Superdex S75 column in 50 mM Tris, pH 7.2 (GE Healthcare).

NMR spectroscopy

NMR samples were prepared in 50 mM Tris, pH 7.2 and experiments were performed at 25°C on a

Bruker Avance-III 700 MHz spectrometer, equipped with a triple-resonance cryogenic probe. The protein samples were 0.7 mM for resonance assignment experiments. The assignment of the backbone resonances of K0-Ub were accomplished by recording and analyzing HNCACB, CBCA(CO)NH, HNCO, and HBHACONH triple resonance datasets.²⁹ Data processing and analysis were performed using the nmrPipe³⁰ and Sparky³¹ software packages, respectively. Backbone H^N-N RDCs were measured using an in-phase anti-phase (IPAP) experiment³² with ¹⁵N K0-Ub aligned in 4% (w/v) C12E5 polyethylene glycol/*n*-hexanol media.³³

Structure calculation

Structures of K0-Ub were calculated using chemical shift-based Rosetta (Rosetta 3.4) using the standard protocols^{24,34} and incorporating backbone N, H^N, C, C^α, H^α, and C^β chemical shifts together with 80% N-H RDC data (52 RDCs randomly selected from total 65 RDCs in structured region from 1 to 71). The calculations generated 10,000 all-atom models, and the 3000 low-energy models were extracted and further rescored against chemical shifts. The final 20 lowest energy structures were chosen as representative of the calculation. The remaining 20% of the RDCs (13 RDCs, not used in calculation) were used to cross-validate the final models by PALES, using the alignment tensor fitted from the 52 RDCs that were used in calculation.²⁵ Validation of the structure was performed using protein structure validation suite (PSVS) (http://psvs-1_5-dev.nesg.org).

Data deposition

The coordinates of the solution structure have been deposited in the PDB as entry 2MI8. The chemical

shift assignments and RDC data have been deposited in the BioMagResBank as entry 19670.

Differential scanning calorimetry

The thermal stabilities of wt-Ub and K0-Ub were measured with a VP-DSC MicroCalorimeter (MicroCal, GE Healthcare). The concentrations of the two proteins were 1.2 mg/mL in 50 mM Tris (pH 7.2). The heating rate used was 1°C/min, and the scanning was performed from 25 to 100°C.

Acknowledgment

The authors thank Marzena A. Dyba and Sergey G. Tarasov for assistance with DSC measurements and Yang Shen, Nikolaos Sgourakis, Oliver Lange for help with CS-Rosetta calculation. None of the authors have any conflict of interest regarding the material or data presented in this publication.

References

- Komander D, Rape M (2012) The ubiquitin code. *Annu Rev Biochem* 81:203–229.
- Hospenthal MK, Freund SM, Komander D (2013) Assembly, analysis and architecture of atypical ubiquitin chains. *Nat Struct Mol Biol* 20:555–565.
- Ikeda H, Kerppola TK (2008) Lysosomal localization of ubiquitinated Jun requires multiple determinants in a lysine-27-linked polyubiquitin conjugate. *Mol Biol Cell* 19:4588–4601.
- Oshikawa K, Matsumoto M, Nakayama KI (2012) Comprehensive study of protein ubiquitylation sites by conjugation of engineered lysine-less ubiquitin. *Seikagaku* 84:479–487.
- Oshikawa K, Matsumoto M, Oyama K, Nakayama KI (2012) Proteome-wide identification of ubiquitylation sites by conjugation of engineered lysine-less ubiquitin. *J Proteome Res* 11:796–807.
- Sun T, Guo J, Shallow H, Yang T, Xu J, Li W, Hanson C, Wu JG, Li X, Massaeli H, Zhang S (2011) The role of monoubiquitination in endocytic degradation of human ether-a-go-go-related gene (hERG) channels under low K⁺ conditions. *J Biol Chem* 286:6751–6759.
- Ziv I, Matiuhi Y, Kirkpatrick DS, Erpapazoglou Z, Leon S, Pantazopoulou M, Kim W, Gygi SP, Haguenauer-Tsapis R, Reis N, Glickman MH, Kleinfeld O (2011) A perturbed ubiquitin landscape distinguishes between ubiquitin in trafficking and in proteolysis. *Mol Cell Proteomics* 10:M111 009753.
- Eddins MJ, Waradan R, Fushman D, Pickart CM, Wolberger C (2007) Crystal structure and solution NMR studies of Lys48-linked tetraubiquitin at neutral pH. *J Mol Biol* 367:204–211.
- Zhang N, Wang Q, Ehlinger A, Randles L, Lary JW, Kang Y, Haririna A, Storaska AJ, Cole JL, Fushman D, Walters KJ (2009) Structure of the S5a:k48-linked diubiquitin complex and its interactions with rpn13. *Mol Cell* 35:280–290.
- Johnson EC, Handel TM (1999) Effect of hydrophobic core packing on sidechain dynamics. *J Biomol NMR* 15:135–143.
- Johnson EC, Lazar GA, Desjarlais JR, Handel TM (1999) Solution structure and dynamics of a designed hydrophobic core variant of ubiquitin. *Structure* 7:967–976.
- Lazar GA, Desjarlais JR, Handel TM (1997) De novo design of the hydrophobic core of ubiquitin. *Protein Sci* 6:1167–1178.
- Lazar GA, Johnson EC, Desjarlais JR, Handel TM (1999) Rotamer strain as a determinant of protein structural specificity. *Protein Sci* 8:2598–2610.
- Haririna A, Verma R, Purohit N, Twarog MZ, Deshaies RJ, Bolon D, Fushman D (2008) Mutations in the hydrophobic core of ubiquitin differentially affect its recognition by receptor proteins. *J Mol Biol* 375: 979–996.
- Hirano T, Serve O, Yagi-Utsumi M, Takemoto E, Hiromoto T, Satoh T, Mizushima T, Kato K (2011) Conformational dynamics of wild-type Lys-48-linked diubiquitin in solution. *J Biol Chem* 286:37496–37502.
- Lai MY, Zhang D, Laronde-Leblanc N, Fushman D (2012) Structural and biochemical studies of the open state of Lys48-linked diubiquitin. *Biochim Biophys Acta* 1823:2046–2056.
- Castaneda CA, Kashyap TR, Nakasone MA, Krueger S, Fushman D (2013) Unique structural, dynamical, and functional properties of k11-linked polyubiquitin chains. *Structure* 21:1168–1181.
- Cornilescu G, Marquardt JL, Ottiger M, Bax A (1998) Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. *J Am Chem Soc* 120:6836–6837.
- Vijay-Kumar S, Bugg CE, Cook WJ (1987) Structure of ubiquitin refined at 1.8 Å resolution. *J Mol Biol* 194: 531–544.
- Esadze A, Li DW, Wang T, Bruschweiler R, Iwahara J (2011) Dynamics of lysine side-chain amino groups in a protein studied by heteronuclear 1H-15N NMR spectroscopy. *J Am Chem Soc* 133:909–919.
- Eghbalnia HR, Wang L, Bahrami A, Assadi A, Markley JL (2005) Protein energetic conformational analysis from NMR chemical shifts (PECAN) and its use in determining secondary structural elements. *J Biomol NMR* 32:71–81.
- Shen Y, Delaglio F, Cornilescu G, Bax A (2009) TALOS plus : a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* 44:213–223.
- Shen Y, Lange O, Delaglio F, Rossi P, Aramini JM, Liu G, Eletsky A, Wu Y, Singarapu KK, Lemak A, Ignatchenko A, Arrowsmith CH, Szyperski T, Montelione GT, Baker D, Bax A (2008) Consistent blind protein structure generation from NMR chemical shift data. *Proc Natl Acad Sci USA* 105:4685–4690.
- Zweckstetter M, Bax A (2000) Prediction of sterically induced alignment in a dilute liquid crystalline phase: aid to protein structure determination by NMR. *J Am Chem Soc* 122:3791–3792.
- Das R, Mariano J, Tsai YC, Kalathur RC, Kostova Z, Li J, Tarasov SG, McFeeters RL, Altieri AS, Ji X, Byrd RA, Weissman AM (2009) Allosteric activation of E2-RING finger-mediated ubiquitylation by a structurally defined specific E2-binding region of gp78. *Mol Cell* 34:674–685.
- Stagno JR, Altieri AS, Bubunenkov M, Tarasov SG, Li J, Court DL, Byrd RA, Ji X (2011) Structural basis for RNA recognition by NusB and NusE in the initiation of transcription antitermination. *Nucleic Acids Res* 39: 7803–7815.
- Das R, Liang YH, Mariano J, Li J, Huang T, King A, Tarasov SG, Weissman AM, Ji X, Byrd RA (2013) Allosteric regulation of E2-E3 interactions promote a processive ubiquitination machine. *EMBO J* 32:2504–2516.
- Sattler M, Schleucher J, Griesinger C (1999) Heteronuclear multidimensional NMR experiments for the

- structure determination of proteins in solution employing pulsed field gradients. *Prog Nucl Mag Res Sp* 34: 93–158.
29. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) Nmrpipe—a multidimensional spectral processing system based on Unix pipes. *J Biomol NMR* 6:277–293.
 30. Goddard TD, Kneller DG SPARKY 3, San Francisco: University of California.
 31. Cordier F, Dingley AJ, Grzesiek S (1999) A doublet-separated sensitivity-enhanced HSQC for the determination of scalar and dipolar one-bond J-couplings. *J Biomol NMR* 13:175–180.
 32. Ruckert M, Otting G (2000) Alignment of biological macromolecules in novel nonionic liquid crystalline media for NMR experiments. *J Am Chem Soc* 122: 7793–7797.
 33. Sgourakis NG, Lange OF, DiMaio F, Andre I, Fitzkee NC, Rossi P, Montelione GT, Bax A, Baker D (2011) Determination of the structures of symmetric protein oligomers from NMR chemical shifts and residual dipolar couplings. *J Am Chem Soc* 133:6288–6298.