

Published in final edited form as:

J Mol Biol. 2009 September 11; 392(1): 208–217. doi:10.1016/j.jmb.2009.07.018.

Prokaryotic ubiquitin-like protein Pup is intrinsically disordered

Xiang Chen¹, William C. Solomon¹, Yang Kang¹, Francisca Cerda-Maira², K. Heran Darwin², and Kylie J. Walters^{1,*}

¹Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, Minneapolis, MN 55455, USA

²Department of Microbiology, New York University School of Medicine, New York, NY 10016, USA

Abstract

The prokaryotic ubiquitin-like protein Pup targets substrates for degradation by the *Mycobacterium tuberculosis* proteasome through its interaction with Mpa, an ATPase that is thought to abut the 20S catalytic subunit. Ubiquitin, which is assembled into a polymer to similarly signal for proteasomal degradation in eukaryotes, adopts a stable and compact structural fold that is adapted into other proteins for diverse biological functions. We used NMR spectroscopy to demonstrate that unlike ubiquitin, the 64 amino acid protein Pup is intrinsically disordered with small helical propensity in the C-terminal region. We found that the Pup:Mpa interaction involves an extensive contact surface that spans S21–K61 and that the binding is in the “slow” exchange regime on the NMR time scale, thus demonstrating higher affinity than most ubiquitin:ubiquitin receptor pairs. Interestingly, during the titration experiment, intermediate Pup species were observable, suggesting the formation of one or more transient state(s) upon binding. Moreover, Mpa selected one configuration for a region undergoing chemical exchange in the free protein. These findings provide mechanistic insights into Pup’s functional role as a degradation signal.

Keywords

prokaryotic ubiquitin-like protein; Pup; tuberculosis; intrinsically disordered

Introduction

Proteasomes are ATP-dependent, multi-subunit proteases found in all domains of life. Like their eukaryotic counterparts, prokaryotic proteasomes are self-compartmentalized proteases¹. To date, only bacteria found in the class Actinomycetes are known to have proteasomes^{2; 3; 4; 5}. Despite the presence of bacterial proteases structurally and biochemically similar to eukaryotic proteasomes, it was not understood how proteins were targeted for degradation, as ubiquitin, the post-translational modifier that tags proteins for degradation, is found only in eukaryotes⁶. Proteins with a ubiquitin-like fold are present in bacteria; however, none has demonstrated covalent attachment to other proteins. Recently, a small protein

© 2009 Elsevier Ltd. All rights reserved.

*To whom correspondence should be addressed. E-mail: walte048@umn.edu; Telephone (612) 625-2688; Facsimile (612) 625-2163 (KJW).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

modifier, prokaryotic ubiquitin-like protein (Pup), was found to target proteins for proteolysis by the *Mycobacterium tuberculosis* proteasome^{7, 8}.

Pup has a C-terminal glutamine, which is deamidated to glutamate by Dop⁹. PafA⁹ ligates Pup to substrates to form an isopeptide bond between Pup's C-terminus and the ε-amino group of substrate lysines^{7, 8}. It also binds non-covalently to the *Mycobacterium* proteasomal ATPase Mpa⁷, which forms a hexameric ring that presumably unfolds and translocates substrates into the bacterial 20S core particle for degradation. The Mpa:Pup interaction likely recruits pupylated substrates for degradation, and Mpa itself is also covalently modified by Pup to become a degradation substrate⁷.

Despite their functional similarity, secondary structure prediction programs suggest that Pup does not have a canonical ubiquitin fold (Figure 1a). We characterized Pup's structural and dynamic characteristics by NMR and CD spectroscopy to find that it is an intrinsically disordered protein. We have also found that it binds to Mpa through interactions that span S21–K61 with its strongest contacts towards the C-terminal end. We propose that Pup's stronger-binding C-terminal region serves as a targeting signal to dock degradation substrates to the prokaryotic proteasome complex while its unstructured N-terminal sequence contains the properties characteristic of a degradation initiation sequence.

Results

Pup's migratory behavior suggests that it is intrinsically disordered

We found that Pup runs at 14 kDa on an SDS gel as demonstrated previously⁷, rather than its calculated 6.9 kDa molecular weight. The delayed migration is most likely due to low SDS binding¹⁰, as Pup's primary sequence has a low hydrophobic amino acid content and is 30% glutamic or aspartic acid (Figure 1a). Pup also elutes earlier than expected during size exclusion chromatography however. It directly follows the 16.7 kDa ubiquitin receptor Rpn13 and elutes significantly earlier than 8.6 kDa ubiquitin (Figure 1b), which forms a compact structure (Figure 1c). Pup has abundant regions of low sequence complexity and is predicted to lack β-strands and to have limited α-helical content. PONDR¹¹ and IUPred¹² each predicted it to be an intrinsically disordered protein (Figure 1a). The elution volume of four well-behaved proteins of known molecular weight and Stokes radius (R_s) (myoglobin, ovalbumin, albumin, and ferritin) was used to estimate an R_s for Pup of 22 Å based on its elution volume (data not shown), as was done in previous work^{13; 14; 15}. Molecular weight was recently correlated with R_s for five known protein conformations (folded, molten globule, pre-molten globule, natively unfolded “pre-molten globule-like,” and natively unfolded “coil-like”)¹⁶. We calculated Pup's predicted R_s for each of these classes (Table 1) to find the best match with the natively unfolded coil-like class (Figure 1d).

Pup has limited secondary structure, with helical propensity in the region spanning A51–F54

We recorded a ¹H, ¹⁵N HSQC experiment on ¹⁵N labeled Pup, which demonstrated limited amide proton dispersion (left panel of Figure 2a); that of ubiquitin is included for comparison (right panel of Figure 2a). Complete amide and Cα chemical shift assignments were made by using a ¹H, ¹⁵N, ¹³C HNCA experiment (see Supplementary Figure 1 for an example). V55–V59 and R29 exhibited two resonances of equal intensity in ¹H, ¹⁵N HSQC spectra recorded with and without 10% glycerol (Supplementary Figure 2a) and these were not exchange broadened, suggesting slow chemical exchange between two distinct states. By using their

chemical shift difference and the equation $\tau = \frac{1}{k_{ex}} \gg \frac{2\pi}{\Delta\omega}$ where τ , k_{ex} and ω are the lifetime, exchange rate and chemical shift difference in Hertz between the two states, respectively¹⁷, we estimated a lifetime of greater than 300 ms. SDS-PAGE revealed that the additional set of

peaks were not derived from proteolysis (data not shown and Supplementary Figure 3). A population of Pup in low abundance elutes early and as a broadened peak (Figure 1b), suggesting the presence of reversible aggregation, which could potentially involve the hydrophobic residues among or proximal to those undergoing chemical exchange, including F54, V55, Y58, and V59 (Figure 1a). 2D exchange spectroscopy (EXSY) can be used to identify spins that exchange magnetization by chemical exchange¹⁸ and ‘N_Z-exchange’ HSQC-type spectra have been used to test for this directly^{19; 20}. Lack of chemical shift dispersion prohibited this analysis for most of the residues with two amide resonances (Figure 2a); however, EXSY crosspeaks do appear for R56, A57 and Y58 (Supplementary Figure 4), providing further evidence for reversible exchange between two distinct NMR states.

We plotted the difference between the chemical shift values of Pup’s C α and H α atoms relative to those of randomly coiled values applying sequence-dependent corrections²¹ to find very little deviation from the random coil values (Figure 2b). Only A51–A57 demonstrated a trend towards helicity with slight, but consistent C α downfield shifting and H α upfield shifting. A51–F54 also demonstrated helical propensity in an ¹⁵N-NOESY spectrum, in which non-sequential, inter-residue NOE interactions were observed only for E35–Q60 (Figure 2c and Table 2). Only one amide resonance is displayed for V55–V59 in Figure 2c; however, the other one exhibits identical NOEs, thus demonstrating similarity between the two structural states. Altogether, our results suggest that transient helicity likely occurs in the region spanning A51–A57; however, all secondary structure is too labile for characterization by NOESY interactions.

Circular dichroism (CD) experiments reveal that Pup does not undergo cooperative unfolding

CD spectroscopy is sensitive to secondary structure. A spectrum recorded on Pup reflects that of a disordered polypeptide with negative ellipticity near 200 nm and low ellipticity at 190 and 222 nm (Figure 3a). Moreover, induced thermal melting did not produce the spectral transition characteristic of cooperative protein unfolding, thus providing further evidence for Pup being an intrinsically disordered protein (Figure 3b).

NMR relaxation experiments indicate that Pup is a highly flexible protein

Amide longitudinal (R_N(N_Z); Figure 4a) and transverse (R_N(N_X); Figure 4b) relaxation rates as well as ¹⁵N heteronuclear NOE enhancements (hetNOE; Figure 4c) were used to probe the internal dynamics of Pup. Both sets of resonances are displayed for R29 and V55–V59 with one set arbitrarily in red. Compared to ubiquitin²², Pup exhibits significantly smaller hetNOE values, with an average of 0.16 (at 800 MHz; Figure 4c) compared to ubiquitin’s average of 0.75 (at 750 MHz)²², thus revealing an increase of high frequency motions. Pup’s last three C-terminal residues, its N-terminal end, and R29–K31 demonstrate enhanced flexibility (Figure 4c). R29–K31 and V55–V59 exhibit faster than average R_N(N_X) values (Figure 4b), providing further evidence that these two regions undergo conformational exchange. V55–V59 also demonstrated larger than average hetNOE values (Figure 4c), as expected by the presence of inter-residue NOE interactions in this region (Figure 2c).

The C-terminal region of Pup binds to Mpa

We used ¹H, ¹⁵N HSQC experiments to test whether our Pup sample was functional for Mpa binding. S21–K61 attenuated upon Mpa addition and new Mpa-bound Pup resonances appeared (Figure 5a and 5b and Supplementary Figure 5a). Mpa forms a hexameric ring and saturation was achieved at 1:1 Pup:Mpa hexamer (Supplementary Figure 5b). Since free and bound states are observed simultaneously for sub-stoichiometric molar ratios of Pup:Mpa hexamer, their interaction is in the “slow exchange” regime on the NMR time scale, which is indicative of strong binding. Whereas many bound-state resonances appear at sub-stoichiometric molar ratios and persist to a molar ratio of 1:1 Pup:Mpa hexamer (highlighted in Figure 5a and Supplementary Figure 5 with asterisks), others appear and then disappear with

Mpa addition (highlighted in Figure 5a and Supplementary Figure 5 with arrows). These data suggest the presence of intermediate binding states.

N50 and A51 are 100 and 94% obliterated, respectively, when Pup:Mpa hexamer are at 1:0.7 molar ratio; hence this segment exhibits the highest affinity for Pup. Binding is propagated throughout the C-terminal half of Pup however. At 1:1 Pup:Mpa hexamer, the unbound resonances of L40, D44, V46, N50 and A51 are obliterated, and one set of resonances for R29, V55, R56, A57, Y58, and V59 is almost obliterated; the second set of resonances was affected relatively little (Figure 5b, right panel). This latter finding indicates that Mpa selectively binds to one of the slow-exchange conformations. As discussed above, it is possible that the second set of resonances is derived from aggregation, which could make critical hydrophobic residues inaccessible to Mpa.

S21–E30's interaction with Mpa is weaker than that of the more C-terminal region with milder effects observed (Figure 5b and Supplementary Figure 5). The N-terminal 20 amino acids of Pup are not significantly affected by Mpa, suggesting that this region remains unbound and disordered. Future experiments are needed to fully characterize the Pup:Mpa complex and these efforts may be hampered by the difficulty of observing the Mpa-bound state of Pup. At 1:1 molar ratio, only ten bound-state resonances are observable despite the significant attenuation of 41 peaks (Figure 5a). Our finding that binding is initiated by the C-terminal half of Pup is consistent with a 2-hybrid experiment performed in *E. coli*, which demonstrated the C-terminal 26 amino acids of Pup to be sufficient for Mpa interaction⁷.

Discussion

We have found that unlike ubiquitin, Pup is intrinsically disordered. Degradation by eukaryotic proteasome typically requires substrates to be covalently attached to ubiquitin and to either harbor or be complexed with a protein containing an unstructured region^{23; 24; 25}. It is possible that Pup fulfills these two requirements in prokaryotes by serving as an adaptor that tethers substrates to Mpa and by harboring intrinsically disordered segments (Figure 5c). Disordered segments are significantly enriched in eukaryotic proteins and it is perhaps for this reason that the eukaryotic modifier ubiquitin does not require them. In one study, disordered regions of greater than 30 amino acids were predicted to exist in 2.0% of archaean, 4.2% of eubacterial and 33.0% of eukaryotic proteins²⁶. In mammals, ~75% of signaling proteins and half of all proteins contain disordered regions of greater than 30 amino acids²⁷. Moreover, ubiquitin serves as a regulator for proteasome-independent events, such as in DNA repair or endocytosis, whereas no such functionality has been ascribed yet to Pup. Only the GG motif is truly conserved between Pup and ubiquitin, a similarity that could have converged from a common need to access a sterically restricted active site during ligation.

It is not clear why intermediate state/s appear at sub-stoichiometric molar ratios of Pup:Mpa hexamer (Figure 5a). Disorder-to-order transitions have been observed as intrinsically unfolded proteins interact with their binding partner^{28; 29}. It is possible that an Mpa-driven conformational change occurs in Pup. We do not expect, however, that Mpa promotes intramolecular interactions in Pup to form a folded structure, but rather, that Pup's configuration changes to enhance its intermolecular interactions with Mpa. This hypothesis is based on Mpa's presumed role as an ATPase that participates in unfolding substrates for degradation.

Mpa-driven conformational changes in Pup are also suggested by the propagation of the binding surface throughout S21–K61. Interactions across this large segment will certainly restrict Pup's conformational freedom. Since Pup's C-terminal residue is ligated to degradation substrates^{7; 8}, the interaction of V59 and K61 with Mpa indicates that pupylated substrates will

be spatially close to the Pup binding site. This docking mechanism may be important for their robust capture. Were they ligated to the N-terminal end of Pup for example, the 20 following non-interacting randomly coiled amino acids might enable them to assume distances too far for reliable and expedient degradation and thereby restrict flux through the proteasome.

Materials and Methods

Sample preparation

Mpa and Pup were expressed and purified from *Escherichia coli* as fusion proteins with a histidine tag (for Mpa) and a chitin-binding domain and intein that undergoes self-cleavage in the presence of thiols (for Pup; New England Biolabs IMPACT™ Kit). Following cell lysis by sonication, the proteins were purified by affinity chromatography with a chitin column (New England Biolabs) for Pup and Ni-NTA agarose resin (Qiagen) for Mpa. On-column cleavage of Pup from the intein-tag was achieved by incubating the chitin beads with elution buffer (20 mM HEPES, 300 mM NaCl, 10% glycerol, pH 7.6) containing 50 mM DTT for 24 hours and at 4 °C. Pup was recovered from the column by elution with 5 volumes of elution buffer. Mpa was eluted from Ni-NTA resin by 250 mM imidazole. Further purification was achieved by size exclusion chromatography on an FPLC system. ¹⁵N and/or ¹³C labeled samples were produced by growth in M9 minimal media with ¹⁵N labeled ammonium chloride and/or ¹³C labeled glucose as the nitrogen and carbon sources, respectively. NMR and CD experiments were recorded with samples dissolved in 20 mM HEPES pH 6.5, 50 mM NaCl, 10% glycerol and at 25°C unless otherwise noted. NMR and CD spectra were also taken without glycerol and the spectra were almost identical in both cases (Supplementary Figure 2b). In all of our samples, Pup's C-terminal residue was glutamic acid rather than glutamine.

NMR experiments

Data processing was performed with NMRPipe³⁰ and the spectra visualized and analyzed with CARA (Diss. ETH Nr. 15947), and XEASY³¹. A 3D HNCA experiment was acquired at 700 MHz with a cryogenically cooled probe on 0.7 mM ¹⁵N, ¹³C labeled Pup. The resulting spectrum contained amide to Cα_i as well as Cα_(i-1) resonances for all spin systems. ¹⁵N dispersed NOESY (120 and 200 ms mixing time) and TOCSY experiments were recorded on ¹⁵N labeled Pup at 0.7 mM concentration and at 900 MHz (for the NOESY experiments) or 600 MHz (for the TOCSY experiment). Rates for ¹⁵N longitudinal R_N(N_Z) and transverse R_N(N_X) relaxation and magnitudes of the heteronuclear NOE enhancements were recorded at 800 MHz and with a cryogenically cooled probe. The EXSY experiment was recorded by monitoring the transfer of heteronuclear N_Z magnetization during mixing times ranging from 10 to 500 ms on a spectrometer operating at 800 MHz.

R_N(N_X) and R_N(N_Z) was derived by fitting data acquired with different relaxation delays to a single-exponential decay function, and error values were determined by repeating one data point. Two spectra were recorded for steady-state NOE intensities, one with 4 seconds of proton saturation to achieve the steady-state intensity and the other as a control spectrum with no saturation to obtain the Zeeman intensity. The control spectrum was repeated to determine error values. Heteronuclear NOE enhancements (hetNOE) were then calculated from the ratio described in Equation 1, as described in ³².

$$\text{XNOE} = \left| \left(\frac{\gamma_{\text{H}}}{\gamma_{\text{N}}} \right) \left[\frac{R_{\text{N}}(\text{H}^{\text{N}} \leftrightarrow \text{N}_{\text{Z}})}{R_{\text{N}}(\text{N}_{\text{Z}})} \right] \right| \quad (1)$$

Circular dichroism (CD) experiments

CD spectra were recorded on a Jasco J-815 spectropolarimeter on samples dissolved in 20 mM HEPES pH 6.5, 50 mM NaCl, 10% glycerol by using quartz cells with a path-length of 1 mm. For the thermostability measurements, a water circulation temperature control was used and spectra in the far-UV region recorded after five minutes of incubation at temperatures ranging from 20 to 80 °C.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgment

We are grateful to Dr. Hiroshi Matsuo for useful discussions and his critical reading of this manuscript. NMR data were acquired in the NMR facilities of the University of Minnesota and the University of Georgia. We are grateful to Dr. John Glushka of the University of GA for setting up experiments for us. The EXSY pulse sequence was generously provided to us by Dr. Marco Tonelli of the National Magnetic Resonance Facility at Madison. Data processing and visualization occurred in the Minnesota Supercomputing Institute Basic Sciences Computing Lab. The CD experiments were performed in the Biophysical Spectroscopy Facility of the Department of Biochemistry, Molecular Biology, and Biophysics. This research was funded by the National Institutes of Health (HL092774 to KHD and CA097004 to KJW).

References

1. Zwickl P, Baumeister W, Steven A. Dis-assembly lines: the proteasome and related ATPase-assisted proteases. *Curr Opin Struct Biol* 2000;10:242–250. [PubMed: 10753810]
2. Nagy I, Tamura T, Vanderleyden J, Baumeister W, De Mot R. The 20S proteasome of *Streptomyces coelicolor*. *J Bacteriol* 1998;180:5448–5453. [PubMed: 9765579]
3. Pouch MN, Cournoyer B, Baumeister W. Characterization of the 20S proteasome from the actinomycete *Frankia*. *Mol Microbiol* 2000;35:368–377. [PubMed: 10652097]
4. Darwin KH, Ehrst S, Gutierrez-Ramos JC, Weich N, Nathan CF. The proteasome of *Mycobacterium tuberculosis* is required for resistance to nitric oxide. *Science* 2003;302:1963–1966. [PubMed: 14671303]
5. Knipfer N, Shrader TE. Inactivation of the 20S proteasome in *Mycobacterium smegmatis*. *Mol Microbiol* 1997;25:375–383. [PubMed: 9282749]
6. Kerscher O, Felberbaum R, Hochstrasser M. Modification of proteins by ubiquitin and ubiquitin-like proteins. *Annu Rev Cell Dev Biol* 2006;22:159–180. [PubMed: 16753028]
7. Pearce MJ, Mintseris J, Ferreyra J, Gygi SP, Darwin KH. Ubiquitin-like protein involved in the proteasome pathway of *Mycobacterium tuberculosis*. *Science* 2008;322:1104–1107. [PubMed: 18832610]
8. Burns KE, Liu WT, Boshoff HI, Dorrestein PC, Barry CE 3rd. Proteasomal protein degradation in *Mycobacteria* is dependent upon a prokaryotic ubiquitin-like protein. *J Biol Chem* 2009;284:3069–3075. [PubMed: 19028679]
9. Striebel F, Imkamp F, Sutter M, Steiner M, Mamedov A, Weber-Ban E. Bacterial ubiquitin-like modifier Pup is deamidated and conjugated to substrates by distinct but homologous enzymes. *Nat Struct Mol Biol*. 2009
10. Tompa P. Intrinsically unstructured proteins. *Trends Biochem Sci* 2002;27:527–533. [PubMed: 12368089]
11. Peng K, Radivojac P, Vucetic S, Dunker AK, Obradovic Z. Length-dependent prediction of protein intrinsic disorder. *BMC Bioinformatics* 2006;7:208. [PubMed: 16618368]
12. Dosztanyi Z, Csizmok V, Tompa P, Simon I. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* 2005;21:3433–3434. [PubMed: 15955779]

13. Carvalho AF, Costa-Rodrigues J, Correia I, Costa Pessoa J, Faria TQ, Martins CL, Fransen M, Sa-Miranda C, Azevedo JE. The N-terminal half of the peroxisomal cycling receptor Pex5p is a natively unfolded domain. *J Mol Biol* 2006;356:864–875. [PubMed: 16403517]
14. Tcherkasskaya O, Uversky VN. Denatured collapsed states in protein folding: example of apomyoglobin. *Proteins* 2001;44:244–254. [PubMed: 11455597]
15. Denning DP, Uversky V, Patel SS, Fink AL, Rexach M. The *Saccharomyces cerevisiae* nucleoporin Nup2p is a natively unfolded protein. *J Biol Chem* 2002;277:33447–33455. [PubMed: 12065587]
16. Uversky VN. What does it mean to be natively unfolded? *Eur J Biochem* 2002;269:2–12. [PubMed: 11784292]
17. Rule, GS.; Hitchens, TK. *Fundamentals of protein NMR spectroscopy*. Springer, Dordrecht: Focus on structural biology; 2006.
18. Jeener J, Meier BH, Bachmann P, Ernst RR. Investigation of exchange processes by two-dimensional NMR spectroscopy. *J. Chem. Phys* 1979;71:4546–4553.
19. Farrow NA, Zhang O, Forman-Kay JD, Kay LE. A heteronuclear correlation experiment for simultaneous determination of ¹⁵N longitudinal decay and chemical exchange rates of systems in slow equilibrium. *J Biomol NMR* 1994;4:727–734. [PubMed: 7919956]
20. John M, Headlam MJ, Dixon NE, Otting G. Assignment of paramagnetic (¹⁵N)-HSQC spectra by heteronuclear exchange spectroscopy. *J Biomol NMR* 2007;37:43–51. [PubMed: 17096205]
21. Schwarzinger S, Kroon GJ, Foss TR, Chung J, Wright PE, Dyson HJ. Sequence-dependent correction of random coil NMR chemical shifts. *J Am Chem Soc* 2001;123:2970–2978. [PubMed: 11457007]
22. Lee AL, Wand AJ. Assessing potential bias in the determination of rotational correlation times of proteins by NMR relaxation. *J Biomol NMR* 1999;13:101–112. [PubMed: 10070752]
23. Prakash S, Tian L, Ratliff KS, Lehotzky RE, Matouschek A. An unstructured initiation site is required for efficient proteasome-mediated degradation. *Nat Struct Mol Biol* 2004;11:830–837. [PubMed: 15311270]
24. Prakash S, Inobe T, Hatch AJ, Matouschek A. Substrate selection by the proteasome during degradation of protein complexes. *Nat Chem Biol* 2009;5:29–36. [PubMed: 19029916]
25. Takeuchi J, Chen H, Coffino P. Proteasome substrate degradation requires association plus extended peptide. *EMBO J* 2007;26:123–131. [PubMed: 17170706]
26. Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J Mol Biol* 2004;337:635–645. [PubMed: 15019783]
27. Dunker AK, Silman I, Uversky VN, Sussman JL. Function and structure of inherently disordered proteins. *Curr Opin Struct Biol* 2008;18:756–764. [PubMed: 18952168]
28. Hershey PE, McWhirter SM, Gross JD, Wagner G, Alber T, Sachs AB. The Cap-binding protein eIF4E promotes folding of a functional domain of yeast translation initiation factor eIF4G1. *J Biol Chem* 1999;274:21297–21304. [PubMed: 10409688]
29. Sugase K, Dyson HJ, Wright PE. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* 2007;447:1021–1025. [PubMed: 17522630]
30. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 1995;6:277–293. [PubMed: 8520220]
31. Bartels C, Xia T-H, Billeter M, Güntert P, Wüthrich K. The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J. Biomol. NMR* 1995;6:1–10.
32. Peng, J.; Wagner, G. *Nuclear Magnetic Resonance Probes of Molecular Dynamics*. 1994. *Protein Mobility from Multiple ¹⁵N Relaxation Parameters*; p. 374–454.
33. Cole C, Barber JD, Barton GJ. The Jpred 3 secondary structure prediction server. *Nucleic Acids Res* 2008;36:W197–W201. [PubMed: 18463136]
34. Wootton JC, Federhen S. Analysis of compositionally biased regions in sequence databases. *Methods Enzymol* 1996;266:554–571. [PubMed: 8743706]

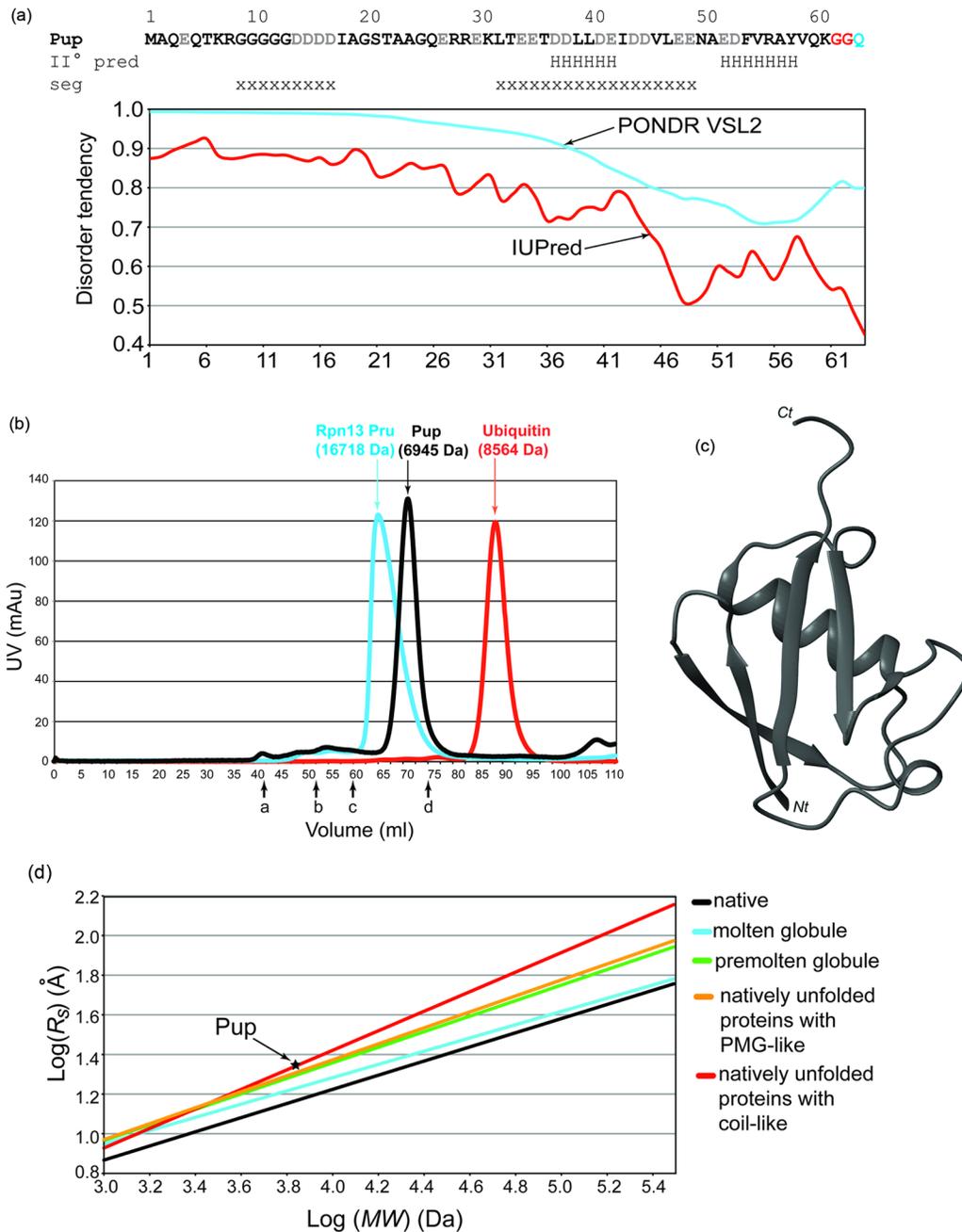


Figure 1. Prokaryotic ubiquitin-like protein Pup is rich in low complexity sequence and has a larger than expected hydrodynamic radius under native conditions

(a) The amino acid sequence of Pup is displayed with its aspartic and glutamic acids in grey and its diglycine motif and C-terminal glutamine in red and blue, respectively. Its predicted secondary structure (with the program Jpred³³) and sequence complexity (with the program seg³⁴) is displayed below the sequence. Helical regions and sequences of low complexity are displayed with 'H' and an 'X,' respectively. The results of the disorder prediction programs POND VSL2¹¹ and IUPred¹² are plotted as labeled; scores over 0.5 indicate predicted disorder. (b) Size exclusion chromatography by FPLC reveals that 6.9 kDa Pup has a larger hydrodynamic radius than 8.6 kDa ubiquitin. At the bottom of the chromatogram, we indicate

the reported elution volume for ferritin, albumin, ovalbumin and myoglobin with arrows labeled a, b, c, and d, respectively. (c) A ribbon diagram is displayed of ubiquitin to highlight its structural complexity. Ubiquitin targets proteins for degradation in eukaryotes, but unlike Pup forms a well-folded, compact structure. (d) Pup's apparent R_s based on its elution time in (b) matches that of a natively unfolded protein. Stokes radius, R_s ; MW, molecular weight.

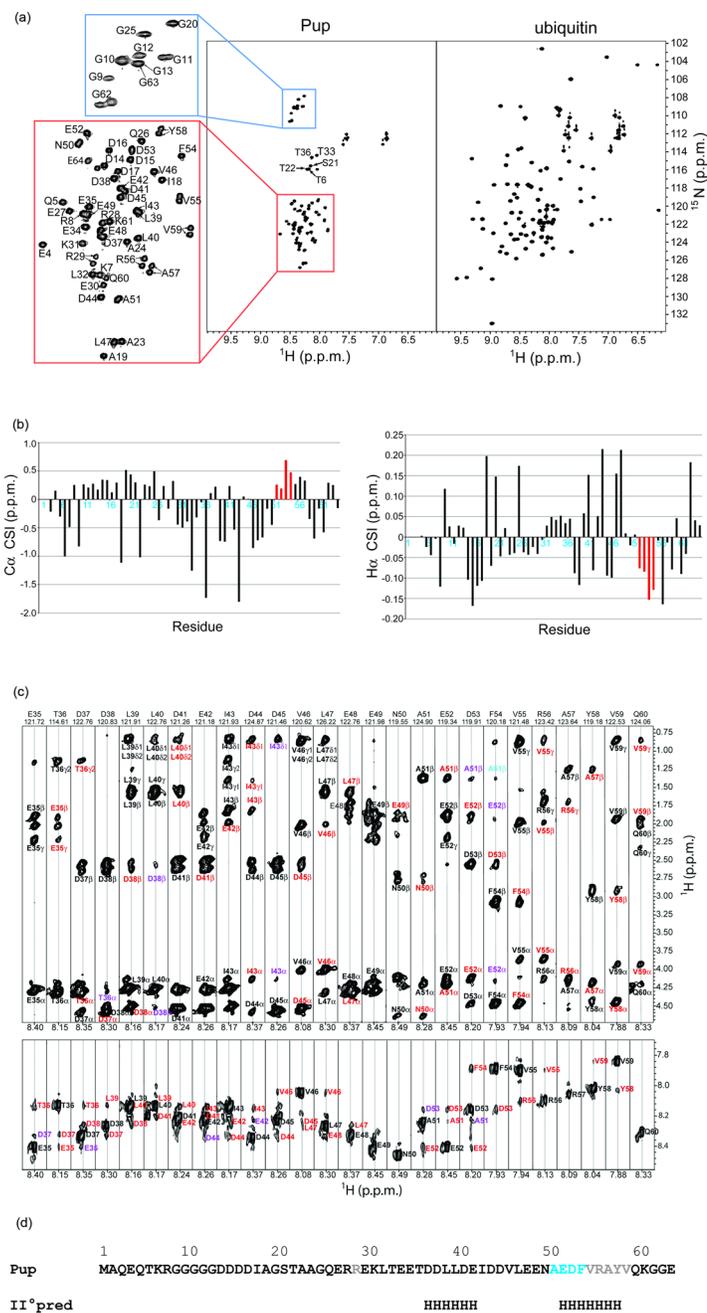


Figure 2. NMR data indicate that Pup contains a stable helix spanning A51–F54
 (a) ^1H , ^{15}N HSQC experiments reveal that Pup (shown on the left) lacks the amide chemical shift dispersion characteristic of a folded protein (as displayed for ubiquitin on the right). R29, V55–V59 exhibited two sets of amide crosspeaks, as labelled. (b) Chemical shift index (CSI) profile for Pup $\text{C}\alpha$ (left) and $\text{H}\alpha$ (right) atoms are displayed. Residues A51–F54 are highlighted in red, as they demonstrate the shifting characteristic of helical structures and were also demonstrated to be helical in an ^{15}N dispersed NOESY spectrum (Figure c). CSI values were determined by subtracting random coil chemical shift value from the assigned one. (c) Regions with long-range NOE interactions are displayed for an ^{15}N dispersed NOESY spectrum acquired on ^{15}N labeled Pup in 20 mM HEPES (pH 6.5) and 50 mM NaCl; no glycerol was

used for this experiment. Intra-residue, sequential, $i \rightarrow i+2$, and $i \rightarrow i+3$ interactions are displayed in black, red, purple, and blue, respectively. (d) Summary of NOESY and CSI data demonstrating that residues A51–F54 (highlighted in blue) exhibit a propensity towards helicity while the rest of the sequence is largely devoid of canonical secondary structure elements. Residues undergoing slow conformational exchange are indicated in grey.

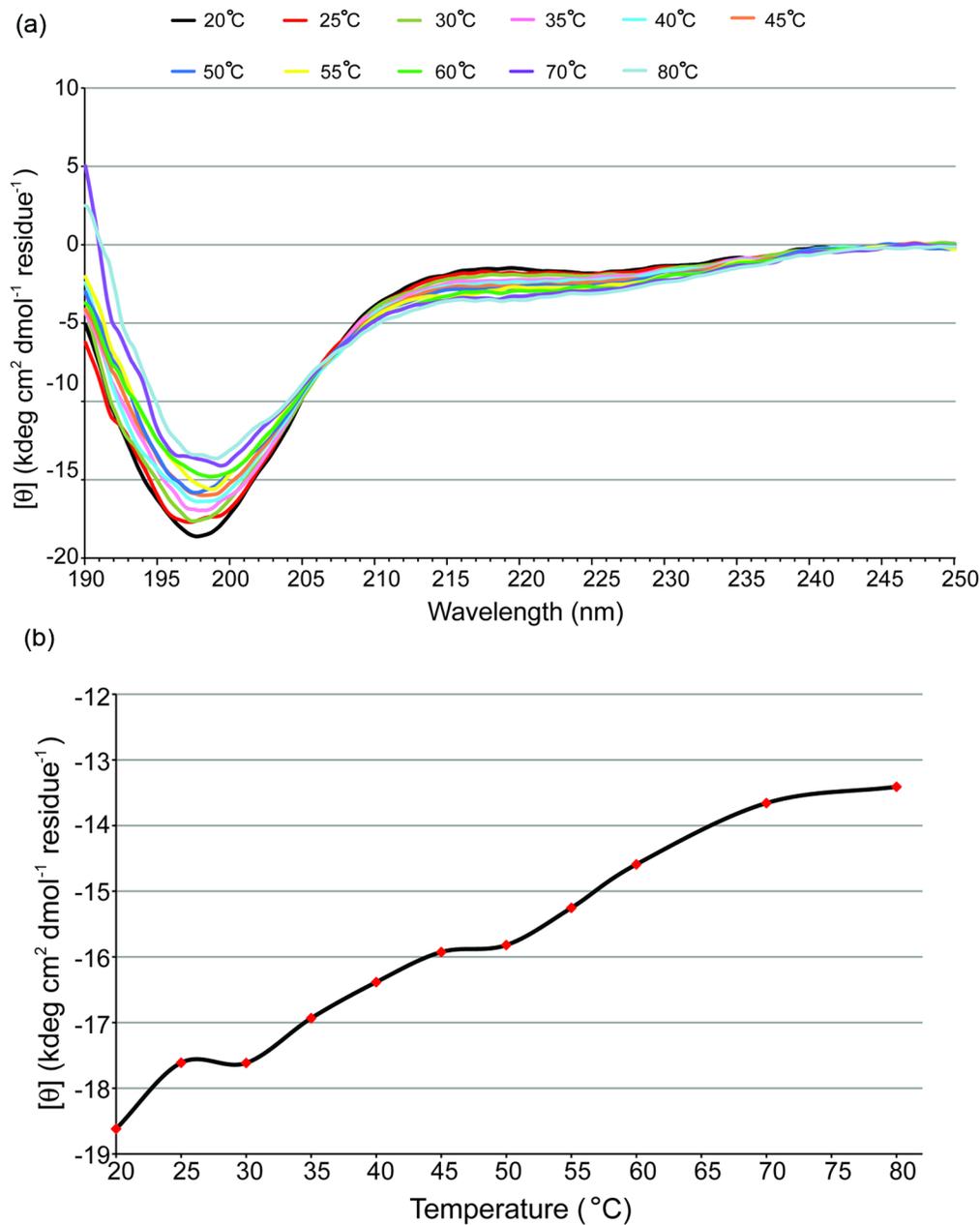


Figure 3. Circular dichroism (CD) data demonstrate that Pup does not unfold cooperatively
 (a) CD spectra in the far-UV region are monitored at several temperatures spanning from 20 to 80 $^{\circ}\text{C}$ on Pup by using a spectropolarimeter equipped with a water circulation temperature control. (b) Thermal unfolding transition curves are provided by plotting the change in ellipticity at 197.7 nm from (a) across temperature. No melting transition is observed for Pup.

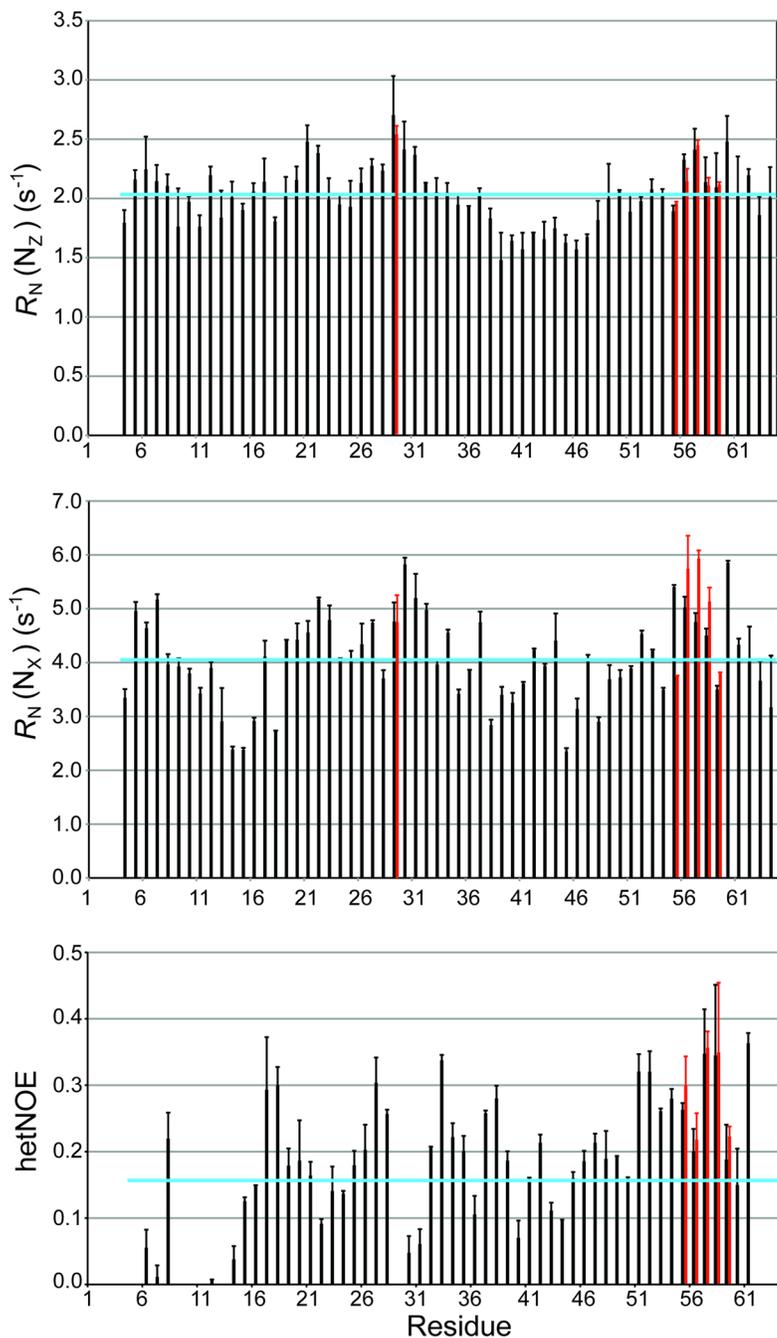


Figure 4. NMR relaxation experiments demonstrate that Pup is a dynamic protein with regions of conformational exchange

(a) Amide longitudinal ($R_N(N_Z)$) and (b) transverse ($R_N(N_X)$) relaxation rates and (c) ¹⁵N heteronuclear NOE enhancements (hetNOE) are displayed for Pup. Both sets of resonances are displayed for R29 and V55–V59 with the arbitrarily assigned “second” resonance displayed in red. The average value is displayed with a blue line to highlight sequence variations.

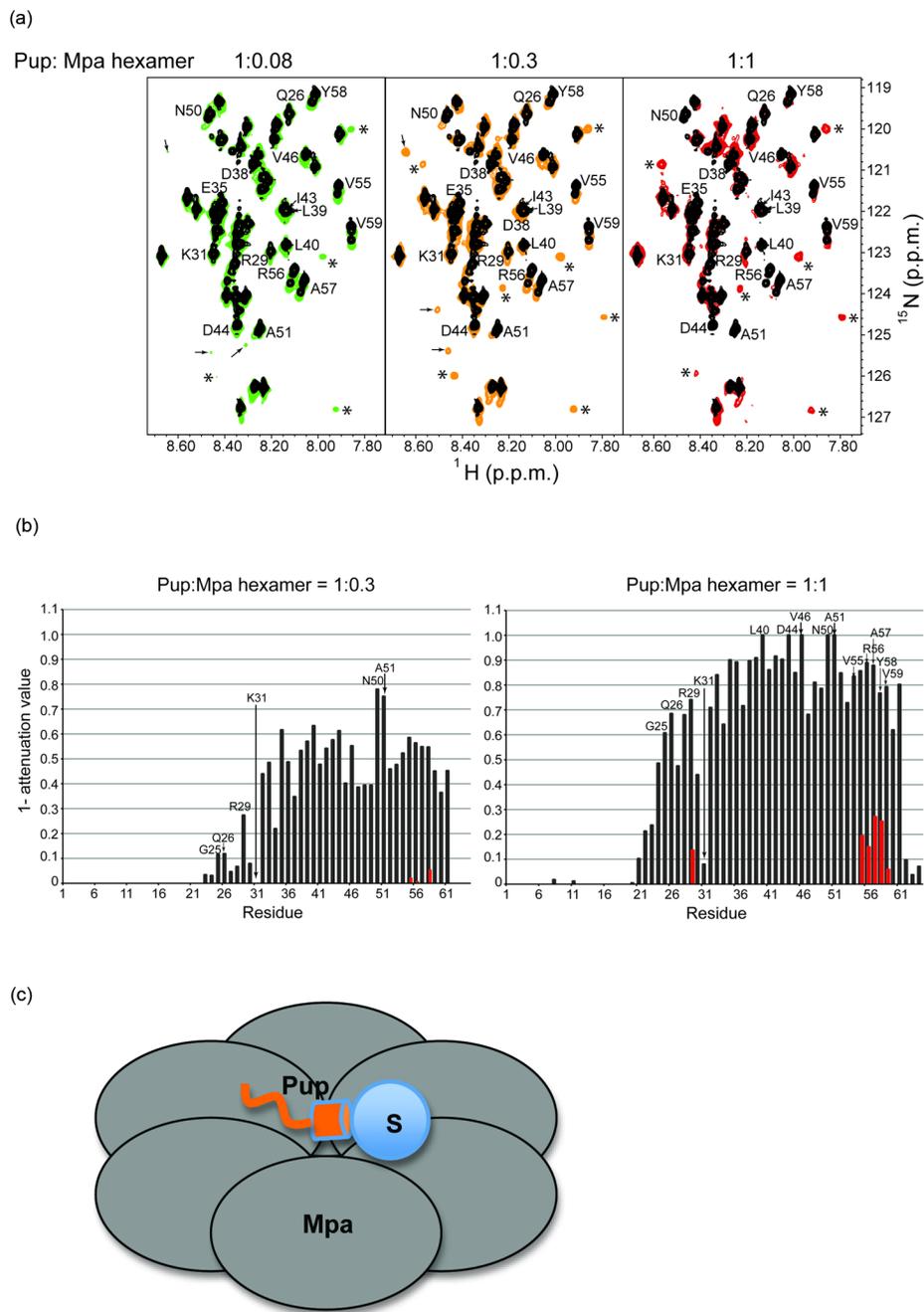


Figure 5. The C-terminal region of Pup binds to the *Mycobacterium* proteasomal ATPase Mpa
 (a) ^1H , ^{15}N HSQC titration experiment in which unlabeled Mpa is added to ^{15}N labeled Pup. Free Pup (black) is displayed superimposed onto Pup:Mpa hexamer at molar ratios of 1:0.08, 1:0.3, and 1:1 as indicated. Resonances from the Mpa-bound state that persist throughout the titration are highlighted with an asterisk, whereas those that appear and then disappear are marked with an arrow. The binding is in slow exchange and therefore the Mpa-bound state is not assigned. (b) Quantitative analysis of the Pup residues involved in Mpa hexamer binding plotting $1 - \text{attenuation} = 1 - \left(\frac{I_{1:0.3}}{I_{\text{free}}} \right)$ (left panel) or $1 - \left(\frac{I_{1:1}}{I_{\text{free}}} \right)$ (right panel) for each residue in

the Pup sequence. In this equation, $I_{1:0.3}$ and $I_{1:1}$ is the intensity of the unbound Pup resonance with 0.3 and equimolar ratio of Mpa hexamer; I_{free} is the intensity of the Pup resonance with no Mpa hexamer present. (c) A model depicting a pupylated substrate binding to Mpa, which forms a hexamer. Pup, its substrate, and Mpa are displayed in orange, blue and grey, respectively.

Table 1Theoretical Stokes radius (R_s) for Pup in different conformation states.

Conformation state	R_s (Å)
Natively folded	14.7 ± 1
Molten globule	17.0 ± 1
Pre-molten globule (PMG)	19.8 ± 1
Natively unfolded (PMG-like)	20.4 ± 1
Natively unfolded (coil-like)	22.0 ± 1

Table 2

Non-sequential, inter-residue NOE interactions observed from Pup E35–Q60.

$\text{HN}_i - \text{HN}_{i+2}$	E35 – D37
	E42 – D44
	A51 – D53
$\text{H}\alpha_i - \text{HN}_{i+2}$	T36 – D38
	D38 – L40
	I43 – D45
	E52 – F54
$\text{H}\beta_i - \text{HN}_{i+2}$	D38 – L40
	A51 – D53
	E52 – F54
$\text{H}\delta_i - \text{HN}_{i+2}$	I43 – D45
$\text{H}\beta_i - \text{HN}_{i+3}$	A51 – F54