CellPress

Structure
Previews

When Less Is More: Combining Site-Specific Isotope Labeling and NMR Unravels Structural Details of Huntingtin Repeats

Florent Delhommel[1,2] and Michael Sattler[1,2,*]
[1]Institute of Structural Biology, Helmholtz Zentrum München, Neuherberg, Germany
[2]Bavarian NMR Center and Center for Integrated Protein Science Munich at Department Chemie, Technical University of Munich, Garching, Germany
*Correspondence: sattler@helmholtz-muenchen.de
https://doi.org/10.1016/j.str.2020.06.005

In this issue of *Structure*, Urbanek et al. (2020a) combine site-specific isotope labeling and NMR spectroscopy to investigate opposing effects of flanking regions onto the conformation of the poly-Q region in Huntingtin. Poly-Q interactions with preceding residues promote an α-helical conformation while a following proline-rich region favors extended conformations.

Huntington's disease is an incurable genetic disorder that causes progressive neurodegeneration with a prevalence of about 8 cases per 100,000. The first symptoms generally occur between the thirties and forties of most patients, with an inexorable progression for 10 to 15 years that eventually leads to death. The disease is caused by the extension of CAG trinucleotide repeats in the first exon of the *huntingtin* gene that produces an extended poly-Q tract in the Huntingtin protein (Htt) beyond a pathological threshold of 35 repeats. A direct correlation between length of the repeat region and age of disease onset as well as its severity has been observed (Harding and Tong, 2018). The pathology manifests by the presence of various forms of aggregates including amyloid fibrils, large protein clusters, and smaller oligomers of Htt. While the aggregation process depends on the length of the poly-Q tract, its pathological role is still unclear.

Htt is a large scaffolding protein that is primarily expressed in neurons and interacts with many binding partners. The functions of its physiological and pathological forms are still poorly understood. However, it is known that the transcript of the first exon of *huntingtin* (*httex1*) is essential and sufficient to recreate the disease symptoms (Harding and Tong, 2018) and is produced by digestion of Htt by caspases. The first exon of *huntingtin* encodes a 17-amino-acid N-terminal region (N17) and a poly-Q tract of variable length, followed by an 11-proline repeat part of a larger proline-rich region (PRR). The two flanking regions of the poly-Q tract have been shown to influence the aggregation propensity of Htt harboring long poly-Q tracts; the presence of the N17 region increases aggregation, while the PRR region reduces it (Shen et al., 2016). The underlying molecular mechanisms by which the flanking regions modulate the aggregation propensity of the poly-Q tract have attracted particular attention but remained poorly understood. The analysis of the structural features of the Httex1 protein is particularly challenging as the polypeptide is highly dynamic and samples disordered and partially folded conformations, in addition to a propensity to aggregate.

NMR is very well suited to study dynamic and disordered regions in proteins. However, low-complexity regions with sequence redundancy such as the ones present in Httex1 remain challenging to investigate due to severe signal overlap. Multi-dimensional NMR spectroscopy has successfully been used to assign the resonances from the Htt poly-Q tract (Baias et al., 2017). These experiments showed that the end of the N17 region has an increased helical propensity that extends into the poly-Q tract but reduces progressively with the distance to the N17 region. However, due to the low sensitivity of these experiments, rather drastic conditions deviating significantly from the native cellular environment had to be used to obtain sufficient signal-to-noise, while still avoiding aggregation. Further biophysical studies revealed an increase in secondary structure content and a compaction of the protein with the increase in poly-Q tract length. These results suggest that the helix conformation could propagate from the N17 region into the start of the poly-Q tract and could be favored with longer tracts (Bravo-Arredondo et al., 2018). However, the lack of residue-level resolution in NMR experiments under more physiological conditions prevented acquisition of high-resolution information on the conformation of Httex1.

Urbanek et al. (2020a) now present a structural analysis of a non-pathologic Httex1 protein containing a short poly-Q tract. Their study was enabled by using site-specific isotopic labeling of individual glutamines within an otherwise unlabeled protein based on a modified protocol for unnatural amino acid labeling (Noren et al., 1989). First, the gene of interest is altered to introduce an unused amber codon at the desired position. Then, the isotopically labeled glutamine is loaded *in vitro* onto an artificial amber suppressor tRNA that recognizes the amber codon. The pre-loaded suppressor tRNA is then used for cell-free expression of the protein of interest, thereby avoiding scrambling with native unlabeled glutamines. The resulting cell-free expressed protein is fully unlabeled except for the targeted position. The authors used this versatile approach in the current work (Urbanek et al., 2020a) and to investigate the *cis*/*trans*-conformation equilibrium of prolines in the PRR of Httex1 in another recent publication (Urbanek et al., 2020b). This approach paves the way for NMR analysis of many proteins containing low-complexity amino acid sequences in which unambiguous chemical shift

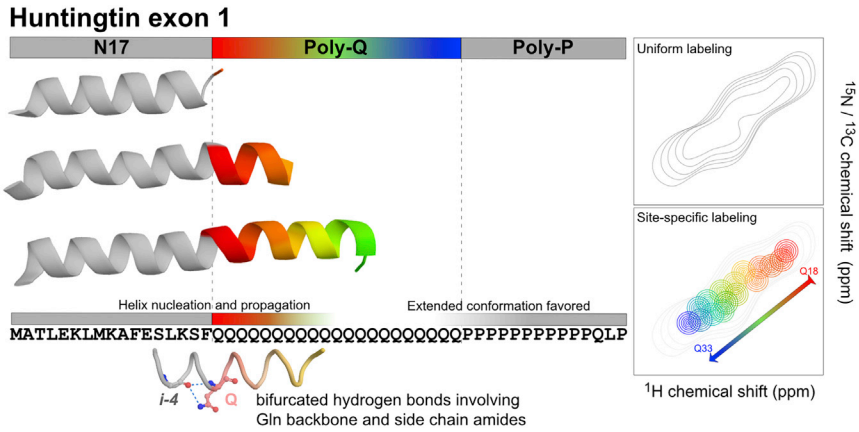

Figure 1. Propagation of Helical Conformation from the N17 Region to the Poly-Q Tract of Httex1 Revealed by NMR Spectroscopy and Site-Specific Isotope Labeling

Secondary NMR chemical shifts obtained from site-specific labeling show an increased helical propensity toward the end of the N17 region that extends into the poly-Q tract. Analysis of conformational ensembles reweighted with the NMR data shows that the helical conformation nucleates in the N17 region and propagates into the poly-Q tract to extend the helix to variable lengths as seen in crystallographic studies (PDB: 3IO6). On the contrary, the PRR favors an extended conformation for preceding residues in the poly-Q tract. The interactions between the first four glutamines of the poly-Q tract with the last four residues of the N17 appear to reinforce the helical propensity and thus mediate the conformational coupling between N17 and poly-Q tract. An example of bifurcated hydrogen bonding between glutamine backbone and side-chain amides and a (i-4) backbone carbonyl is shown (PDB: 3DAQ).

assignments are challenging, if not impossible.

Using their site-specific isotope-labeling approach, the authors (Urbanek et al., 2020a) study 22 unique samples, systematically labeling every glutamine present in Httex1. This extensive work allows unambiguous NMR analysis of the glutamine region, without signal overlap in the NMR spectra. Combining this approach and uniformly labeled samples, the authors achieve complete chemical shift assignment of Httex1, relying only on sensitive two-dimensional $^{1}H,^{15}N$ and $^{1}H,^{13}C$ correlation experiments. The high sensitivity and simplicity of these NMR spectra allow, for the first time, the studying of Httex1 with no additional fusion tags at high resolution in more physiologically relevant concentrations and buffer conditions while still avoiding aggregation.

The authors (Urbanek et al., 2020a) first investigate the secondary structure propensity of Httex1. NMR secondary chemical shifts are used to confirm that the end of the N17 region has a high propensity to form a transient α-helix. Interestingly, this helical propensity extends into the first residues of the poly-Q tract but reduces gradually with the distance from the N17 region and the proximity to the PRR. Based on an ensemble model obtained from the reweighting of population of conformations using NMR data, the authors propose that the helix formation nucleates within the N17 region and then propagates to the poly-Q tract, rationalizing the increased helical propensity closer to the N17 region. These results confirm previous observations on Httex1 conformation made in less-physiological conditions by NMR (Baias et al., 2017), by low-resolution biophysical methods (Bravo-Arredondo et al., 2018), and by crystallographic studies of MBP-fused Httex1 (Kim et al., 2009). Moreover, it extends the previous results by showing that the transient α-helical conformation nucleates within the N17 region and then propagates into the poly-Q tract, thus revealing an important effect of the N17 region on the poly-Q tract (Figure 1).

The site-specific amino acid labeling combined with computational modeling provides detailed structural insight into the folding regulation of this N-terminal helix. By analyzing the side-chain chemical shifts, the authors show that most glutamines in the poly-Q tract are solvent exposed, with flexible side chains. In contrast, the four glutamines closest to the N17 region show distinct chemical shifts indicative of ordered side chains and hydrogen bonding. The authors propose that the first glutamines of the poly-Q tract create bifurcated hydrogen bonds of their main-chain amide and their side-chain NH_2 group with the main-chain carbonyl group of residue i-4, located in the last residues of the N17 region. Similar observations were previously made in the pathological poly-Q tract of the androgen receptor (AR) (Escobedo et al., 2019). In the AR, this interaction and the helical propensity are particularly promoted when the interacting residues of the flanking region are hydrophobic. The authors now identify a similar mechanism for Httex1 by analyzing NMR chemical shifts of the first poly-Q residues in various mutants of the N17 region that modulate its hydrophobic content. This analysis demonstrates that the helical propensity correlates with the rigidity of the poly-Q residue side chains and depends on the hydrophobicity of the final residues of the N-terminal flanking region, similarly to what is observed in the AR.

The authors also investigate the influence of the flanking PRR onto the poly-Q tract conformation. The effect on secondary chemical shifts of a mutant decoupling poly-Q tract and PRR shows a stronger helix propensity at the end of the poly-Q tract. These results suggest that the PRR structural influence extends well into the preceding poly-Q tract to promote an extended conformation and thereby counteracts the helical propensity induced by the N17 region.

Intrigued by the similar behavior of poly-Q regions of the AR and Httex1, the authors analyzed the amino acid composition bias of regions flanking poly-Q tracts in the human proteome. Strikingly, hallmarks of the conformational features found for Httex1, i.e., increased α-helical propensity preceding the poly-Q region, and strong enrichment of prolines following the poly-Q region, are also seen for many other long poly-Q tracts. These findings reveal common molecular and sequence features shared by poly-Q regions and suggest that common conformational behavior of poly-Q tracts and their flanking regions could be related to their pathological role.

The results presented in this article confirm and highlight the importance of side-chain to main-chain interactions in the stabilization of α-helical secondary structures and demonstrate how proline-rich regions promote an extended conformation on neighboring residues by distal effects. The rigorous analysis presented

here offers new perspectives on distal conformational effects of flanking regions on transient secondary structures that are particularly relevant in intrinsically disordered regions of proteins and are often linked to pathological poly-peptide sequences (Milles et al., 2018). It also paves the way to study disease-related mechanisms of poly-Q sequences in cellular and in *in vivo* assays.

ACKNOWLEDGMENTS

We thank Hyun-Seo Kang and Santiago Lumbreras-Martinez for discussions. F.D. is supported by an EMBO Long term Fellowship (ALTF 243-2018). We acknowledge support by the Helmholtz Association Initiative and Networking Fund under project number ZT-I-0003 and the Deutsche Forschungsgemeinschaft (grant SFB1035) (to M.S.).

REFERENCES

Baias, M., Smith, P.E.S., Shen, K., Joachimiak, L.A., Żerko, S., Koźmiński, W., Frydman, J., and Frydman, L. (2017). Structure and dynamics of the Huntingtin Exon-1 N-Terminus: A solution NMR perspective. J. Am. Chem. Soc. *139*, 1168–1176.

Bravo-Arredondo, J.M., Kegulian, N.C., Schmidt, T., Pandey, N.K., Situ, A.J., Ulmer, T.S., and Langen, R. (2018). The folding equilibrium of huntingtin exon 1 monomer depends on its polyglutamine tract. J. Biol. Chem. *293*, 19613–19623.

Escobedo, A., Topal, B., Kunze, M.B.A., Aranda, J., Chiesa, G., Mungianu, D., Bernardo-Seisdedos, G., Eftekharzadeh, B., Gairí, M., Pierattelli, R., et al. (2019). Side chain to main chain hydrogen bonds stabilize a polyglutamine helix in a transcription factor. Nat. Commun. *10*, 2034.

Harding, R.J., and Tong, Y.F. (2018). Proteostasis in Huntington's disease: disease mechanisms and therapeutic opportunities. Acta Pharmacol. Sin. *39*, 754–769.

Kim, M.W., Chelliah, Y., Kim, S.W., Otwinowski, Z., and Bezprozvanny, I. (2009). Secondary structure of Huntingtin amino-terminal region. Structure *17*, 1205–1212.

Milles, S., Salvi, N., Blackledge, M., and Jensen, M.R. (2018). Characterization of intrinsically disordered proteins and their dynamic complexes: From in vitro to cell-like environments. Prog. Nucl. Magn. Reson. Spectrosc. *109*, 79–100.

Noren, C.J., Anthony-Cahill, S.J., Griffith, M.C., and Schultz, P.G. (1989). A general method for site-specific incorporation of unnatural amino acids into proteins. Science *244*, 182–188.

Shen, K., Calamini, B., Fauerbach, J.A., Ma, B., Shahmoradian, S.H., Serrano Lachapel, I.L., Chiu, W., Lo, D.C., and Frydman, J. (2016). Control of the structural landscape and neuronal proteotoxicity of mutant Huntingtin by domains flanking the polyQ tract. eLife *5*, 1–29.

Urbanek, A., Popovic, M., Morató, A., Estaña, A., Elena-Real, C.A., Mier, P., Fournet, A., Allemand, F., Delbecq, S., Andrade-Navarro, M.A., et al. (2020a). Flanking Regions Determine the Structure of the Poly-Glutamine in Huntingtin through Mechanisms Common among Glutamine-Rich Human Proteins. Structure *28*, 1–14.

Urbanek, A., Popovic, M., Elena-Real, C.A., Morató, A., Estaña, A., Fournet, A., Allemand, F., Gil, A.M., Cativiela, C., Cortés, J., et al. (2020b). Evidence of the reduced abundance of proline cis conformation in protein poly proline tracts. J. Am. Chem. Soc. *142*, 7976–7986.