Deciphering the Structural and Functional Roles of Poly-Alanine Repeats: Insights from the Combined Application of NMR and Molecular Simulations

Santiago Rodriguez*[1], Carlos Alberto Elena-Real[2], Ons Dakhlaoui[1], Amin Sagar[3], and Pau Bernado[1]

[1]Centre de Biologie Structurale, Université de Montpellier, INSERM, CNRS, 29 rue de Navacelles, 34090, Montpellier, France – CNRS, Université de Montpellier, Institut National de la Santé et de la Recherche Médicale - INSERM – France

[2]Centre de Biologie Structurale, Université de Montpellier, INSERM, CNRS, 29 rue de Navacelles, 34090, Montpellier, France – CNRS, Université de Montpellier, Institut National de la Santé et de la Recherche Médicale - INSERM – France

[3]BicycleTx Limited, Cambridge, UK – Royaume-Uni

Résumé

Abstract

Compositionally biased sequences confer unique structural and functional properties to proteins. Among them, homorepeats, stretches of a single amino acid, are particularly intriguing. Poly-alanine (poly-A) repeats stand out due to their abundance, specially in transcription factors, although their precise functional role remains to be fully elucidated. Moreover, aberrant poly-A expansions have been implicated in nine rare developmental diseases, suggesting previously unrecognized biological functions for these repeats. The inherent disorder and compositional bias of poly-A pose significant challenges for high-resolution investigations of their structure-function relationships. In this work, we present our efforts to tackle this challenge by combining theoretical and experimental approaches.

We have focused our structural study on Phox2B, a transcription factor essential for autonomic nervous system development. Phox2B contains two poly-A tracts of 9 and 20 consecutive alanines, which is the longest poly-A tract in the human proteome. Interestingly, expansions of +7 or +13 additional alanines on the longest tract lead to Congenital Central Hypoventilation Syndrome (CCHS), a rare autosomal dominant disorder that impairs autonomous respiration. To experimentally investigate the poly-A tracts in Phox2B, we employed the site-specific isotopic labeling (SSIL) approach developed in our group**1**, enabling the measurement of the chemical shifts of individual alanines by Nuclear Magnetic Resonance (NMR).

From a theoretical perspective, we performed Replica Exchange with Solute Tempering 3 molecular dynamics (REST3 MD) simulations on the 9- and 20-residue poly-A tracts using three force fields adapted to the study of intrinsically disordered proteins. Among them, amber03ws exhibited the best agreement with the NMR chemical shifts for both poly-A tracts.

*Intervenant

We further refined the MD trajectory by reweighting**2** it using the experimental data**3**. The resulting Phox2B structural model indicated the presence of fluctuating α-helices in the poly-A tracts, and helical populations that strongly depended on the length, the flanking sequences and the temperature. We then extended our REST3 MD simulations to the +7 and +13 poly-A expansions using the amber03ws force field. The simulations revealed an increase in helicity that correlated with the number of alanines present in the poly-A. Taken together, our results pave the way towards a structural understanding of poly-A function and the mechanisms underlying poly-A expansion disorders. Our theoretical strategy is particularly relevant for these aberrant expansions, for which the experimental manipulation remains extremely challenging.

Acknowledgements

The authors thank the CALMIP supercomputing center (Toulouse, France) for the HPC resources provided in the frame of the project P16032.

References

1Carlos A. Elena-Real, Annika Urbanek, Lionel Imbert, Anna Morató, Aurélie Fournet, Frédéric Allemand, Nathalie Sibille, Jérôme Boisbouvier, and Pau Bernadó. Site-Specific Introduction of Alanines for the Nuclear Magnetic Resonance Investigation of Low-Complexity Regions and Large Biomolecular Assemblies. ACS Chemical Biology 2023 *18* (9), 2039-2049. doi: 10.1021/acschembio.3c00288

2Bottaro S, Bengtsen T, Lindorff-Larsen K. Integrating Molecular Simulation and Experimental Data: A Bayesian/Maximum Entropy Reweighting Approach. *Methods Mol Biol.* 2020;2112:219-240. doi: 10.1007/978-1-0716-0270-6_15

3Antón, R., Treviño, M.Á., Pantoja-Uceda, D. *et al.* Alternative low-populated conformations prompt phase transitions in polyalanine repeat expansions. *Nat Commun* **15**, 1925 (2024). doi: 10.1038/s41467-024-46236-5