HOW A BIASED GENETIC CODE CREATES ORDER FROM PROTEINS' DISORDER

Rupert Faraway[1,2,#], Neve Costello Heaven[1,2,*], Holly Digby[1,2,*], Oscar G. Wilkins[1,2,3], Anob M. Chakrabarti[1,2,4], Ira A. Iosub[1,2], Lea Knez[1,7], Stefan L. Ameres[5], Clemens Plaschka[6], Jernej Ule[1,2,3,8,#]

[1]*The Francis Crick Institute, London, UK*

[2]*UK Dementia Research Institute at King's College London, London, UK*

[3]*UCL, UCL Queen Square Institute of Neurology, Department of Neuromuscular Diseases, London, UK*

[4]*University College London, UCL Respiratory, Division of Medicine, London, UK*

[5]*University of Vienna, Max Perutz Labs, Vienna BioCenter, Vienna, Austria*

[6]*Research Institute of Molecular Pathology, Vienna BioCenter, Vienna, Austria*

[7]*Ludwig-Maximilians-Universität München, Munich, Germany*

[8]*National Institute of Chemistry, Ljubljana, Slovenia*

E-mail: rupert.faraway@gmail.com, jernej.ule@kcl.ac.uk; #corresponding authors; * These authors contributed equally to this work

Protein dosage is regulated to maintain cellular homeostasis and health. The dosage of proteins containing disordered low complexity domains (LCDs) must be particularly well-controlled, yet no mechanism to maintain their mutual homeostasis has been identified. Here we report a mutual homeostatic mechanism that controls the concentration of such proteins, termed 'interstasis', in which proteins with similar LCDs co-regulate their combined dosage through mRNA-mediated negative feedback. We focused on the mechanism that exploits the fundamental multivalency of GA-rich RNA regions that encode charged LCDs, including those with arginine-enriched mixed charge domains (R-MCDs). Modest variations in the abundance of an R-MCD protein change the properties of nuclear speckles, a protein-RNA condensate, selectively trapping multivalent GA-rich mRNAs to promote their nuclear retention. This interstasis depends on the inherent characteristics of the genetic code, and on codon biases that are most pronounced in amniotes, which enhance the multivalency of GA-rich regions encoding the charged LCDs. The threshold of interstasis is modulated by CLK kinases, which affect the nuclear speckle localisation of proteins such as TRA2B, key binder of GA-rich RNAs. Notably, many classes of LCDs are encoded by RNA regions containing multivalency-enhancing codon biases, each preferentially bound by specific proteins, suggesting that interstasis might co-regulate many classes of functionally related LCD-containing proteins through dose-sensitivity of various types of protein-RNA condensates.