Researchers develop a library of elastin-like proteins to help in creating synthetic designs

September 23 2015, by Bob Yirka

(Phys.org)—A pair of researchers at Duke University has built a library of protein data that outlines the specific amino acid sequences that control changes of many elastin proteins. In their paper published in the journal *Nature Materials*, Felipe García Quiroz and Ashutosh Chilkoti describe their research, the making of their library, and their belief that what they have created will help in the development of new synthetic designs for possible use in medical applications.

Proteins are organic compounds essential to all living organisms, they are especially prevalent in components that have structure, such as muscle, skin, hair, etc. They provide structure by self-forming into different shapes under different conditions, two of which are solubility and temperature. Proteins are made of sequences of amino acids—the order and type of which drive the shape of the protein when certain conditions are met. Scientists still do not quite understand how proteins self assemble into the specific 3D shapes they take, nor which amino acids lead to which shapes, or indeed, how the order in which they exist contributes to those shapes.

To help provide a better understanding of how it all works, Quiroz and Chilkoti set about building a library of all the known elastin-like proteins, along with the shapes they take under different conditions. They based it on the sequences of five key amino acids found in the fibrous protein typically found in connective tissue, such as muscles. They then set about testing each entry in the library by growing samples of *E. coli* engineered to produce proteins that folded into different

shapes under different conditions. Most specifically noted was under which conditions the proteins shifted from being soluble to non-soluble and vice-versa. That work led them to developing a set of rules that loosely defined which amino acid sequences would result in which shapes under which conditions.

The research duo acknowledge that their rules are more like guidelines, but suggest the basis of what they have built can not only be made stronger with more research by them and others, but can be used to assist in creating synthetic proteins for use in developing targeted drugs. One example would be protein capsules that remain insoluble inside the body until a certain condition is met, at which point, a medication would be released.

More information: Sequence heuristics to encode phase behaviour in intrinsically disordered protein polymers, *Nature Materials* (2015) DOI: 10.1038/nmat4418

Abstract
Proteins and synthetic polymers that undergo aqueous phase transitions mediate self-assembly in nature and in man-made material systems. Yet little is known about how the phase behaviour of a protein is encoded in its amino acid sequence. Here, by synthesizing intrinsically disordered, repeat proteins to test motifs that we hypothesized would encode phase behaviour, we show that the proteins can be designed to exhibit tunable lower or upper critical solution temperature (LCST and UCST, respectively) transitions in physiological solutions. We also show that mutation of key residues at the repeat level abolishes phase behaviour or encodes an orthogonal transition. Furthermore, we provide heuristics to identify, at the proteome level, proteins that might exhibit phase behaviour and to design novel protein polymers consisting of biologically active peptide repeats that exhibit LCST or UCST transitions. These findings set the foundation for the prediction and encoding of phase

behaviour at the sequence level.