On the Role of Hydration in the Folding of Ordered Structures in Peptides

Thomas Dixon, New College (2016)

Protein folding is an essential life process. In all living cells, proteins are synthesised as linear strands, but most must exist as folded globular structures in order to carry out their biological function. The fold a protein adopts is consistent for a given sequence of amino acids (as encoded by DNA). Since proteins often have very specific biological functions, determined by their folded structure *in vivo*, cells have the ability to influence which functions are carried out on both the intracellular and extracellular level by tightly controlling DNA transcription and translation of RNA to protein. However, the mechanics behind this reproducibility of folding are not yet completely understood. In particular the specific role of water remains largely unknown, and a cause of scientific contention. (Ben-Naim, 2013)

DNA is the blueprint from which proteins are synthesized, but mutations arise in DNA. If these mutations arise on portions of DNA which encode a protein, the sequence of amino acids that make up the protein may be altered. This in turn can cause the protein to fold differently, adopting a different 3-dimensional structure, which may then have different functionality as compared to a non-mutated variant. Whilst the accumulation of mutations and subsequent environmental selection is the driving force for evolution, mutations are also the cause of a vast range of diseases. Mutated DNA in the germ line may lead to inherited genetic conditions where mutant proteins may undergo; a gain-of-function resulting in an autosomal dominant disease, e.g. hyperthyroidism (Weinstein, 2006); or a loss-of-function resulting in an autosomal recessive condition, e.g. cystic fibrosis (Kelley *et al.*, 1997). Alternatively, mutations acquired in somatic cells may also cause diseases, primarily cancers. Many mutations in proteins however have little effect on protein function, only effecting minimal changes to the folded structure, but there is currently no way to predict which mutations will cause significant protein misfolding detrimental to function.

The '*hydrophobic effect*' is usually cited as the principal entropic path for protein folding; where it is thought that hydrophobic residues cluster together to eliminate unfavourable water-hydrophobic species interactions. However, recent studies have suggested that water has a more active role in this process, where specifically water is thought to stabilise folded proteins (Huggins, 2016) and aid in nucleating the folding sites in order to guide the protein into its fully folded functional structure. (Blanco *et al.*, 1994; Busch *et al.*, 2013)

My research is focused on the β-turn; a common 4-residue structural motif found in loop regions in folded proteins (Rose *et al.*, 1985). Misfolding of this structure within proteins has implications in forming plaques, the mechanism of diseases such as Alzheimer's. (Hsiao *et al.*, 1996) The proline-glycine (PG) motif (Hsu *et al.*, 2006) is known to promote β-turn formation in loop regions. Short polypeptides containing this motif are ideal for investigating how water contributes to protein folding in solution. In solution, short PG containing peptides exist in an equilibrium between closed and open conformations in solution, allowing the effect of water on the folding process to be probed. Previous study (Busch *et al.*, 2013) on the glycyl-prolyl-glycinamide (GPG-NH_2) tripeptide in water suggested that water has an active role in mediating the turn, where a single water molecule engages in multiple hydrogen bonds with the peptide. This may therefore also be true for longer polypeptides and proteins, including *in vivo*. I am using both experimental and theoretical techniques to study this, with the aim of marrying together molecular dynamic modelling of systems *in silico* with models which are consistent with experimental data.

I am using a range of experimental and theoretical techniques. Nuclear Magnetic Resonance experimentation is used to gain information on the atomic chemical environments, which are influenced by the folded state of the peptide in solution. Molecular Dynamics is used to model the peptide solutions *in silico* to study associations of water

molecules to the peptide in a time-dependent manner, including mapping the changes in hydration as a result of peptide folding and unfolding. By combining information from neutron scattering on the peptide in solution with Empirical Potential Structure Refinement (EPSR) modelling, (Soper, 2012) the atomic scale interactions of the peptide with both solvent water molecules and itself can be modelled. Importantly, using the analysis program ANGULA, (Busch *et al.*, 2014) the peptides can be categorised from the EPSR model into different folding states, which allows the detailed assessment of how water interacts and the hydration changes around different atomic sites of peptides which adopt different conformations in solution. (Steinke *et al.*, 2017) For these systems, the information gained by using neutrons is crucial, as it provides insight to interactions on the atomic scale, particularly for hydrogen atoms which are invisible to many other techniques.

The hope is that all these methods will lead to a unified and consistent model, which could have the longer term impact of predicting results for one method from another, thus understanding better the structure and hydration of protein structures, and their implications in causing disease.

Bibliography

Arieh Ben-Naim. Inversion of the hydrophobic/hydrophilic paradigm demystifies the protein folding and self-assembly of problems. *International Journal of Physics*, 1(3): 66–71, 2013.

Francisco J Blanco, German Rivas, and Luis Serrano. A short linear peptide that folds into a native stable β-hairpin in aqueous solution. *Nature Structural & Molecular Biology*, 1(9):584–590, 1994.

Sebastian Busch, Chrystal D Bruce, Christina Redfield, Christian D Lorenz, and Sylvia E McLain. Water mediation is essential to nucleation of β-turn formation in peptide folding motifs. *Angewandte Chemie International Edition*, 52(49):13091–13095, 2013.

Sebastian Busch, Christian D Lorenz, Jonathan Taylor, Luis Carlos Pardo, and Sylvia E McLain. Short-range interactions of concentrated proline in aqueous solution. *The Journal of Physical Chemistry B*, 118(49):14267–14277, 2014.

Karen Hsiao, Paul Chapman, Steven Nilsen, Chris Eckman, *et al.* Correlative memory deficits, a beta elevation, and amyloid plaques in transgenic mice. *Science*, 274(5284): 98, 1996.

Hung-Ju Hsu, Hong-Ju Chang, Hung-Pin Peng, Shan-Sheng Huang, Ming-Yen Lin, and An-Suei Yang. Assessing computational amino acid β-turn propensities with a phage-displayed combinatorial library and directed evolution. *Structure*, 14(10): 1499–1510, 2006.

David J Huggins. Studying the role of cooperative hydration in stabilizing folded protein states. *Journal of Structural Biology*, 196(3):394–406, 2016.

Thomas J Kelley, Kirk Thomas, Laura JH Milgram, and Mitchell L Drumm. In vivo activation of the cystic fibrosis transmembrane conductance regulator mutant δf508 in murine nasal epithelium. *Proceedings of the National Academy of Sciences*, 94(6):2604–2608, 1997.

George D Rose, Lila M Glerasch, and John A Smith. Turns in peptides and proteins. *Advances in protein chemistry*, 37:1–109, 1985.

Alan K Soper. Computer simulation as a tool for the interpretation of total scattering data from glasses and liquids. *Molecular Simulation*, 38(14-15):1171–1185, 2012.

Nicola Steinke, Anna Genina, Christian D Lorenz, and Sylvia E McLain. Salt interactions in solution prevent direct association of urea with a peptide backbone. *The Journal of Physical Chemistry B*, 121(8):1866–1876, 2017.

Lee S Weinstein. Gain-of-function mutations in human genetic disorders. *eLS*, 2006.