Soft Matter at the nanoscale constitutes an information-transporting medium

F. Sumbul, S.A.E. Acuner-Ozbabacan and T. Haliloglu, 'Allosteric Dynamic Control of Binding', *Biophys. J.*, **109**, 1190-1201 (2015)

T. Lenaerts, J. Ferkinghoff-Borg, F. Stricher, L. Serrano, J. W. H. Schymkowitz and F. Rousseau, 'Quantifying information transfer by protein domains: Analysis of the Fyn SH2 domain structure', *BMC Structural Biology*, **8**:43 (2008)

Recommended with a commentary by Tom McLeish, Durham University

Information theory, with its origin in signal processing and the important concerns of signal-degradation by noise, has also found its way into neurology, from there to the physics of neural nets, and even, in speculative mode, to the reframing of the fundamental properties of space-time. However, soft matter physics is not commonly thought-of in terms of information-transfer. The most natural (non-cortical) area in which to think in this way is the role of soft matter in biology – the binding of signalling molecules ('effectors') to the larger, globular proteins, that in turn affects their own propensity to aggregate, to bind to DNA, or to catalyse reactions, 'looks' like a sort of information-transfer. Yet, once conceived in this way, it becomes apparent that reversibly-associating soft matter can be conceived in terms of information – proteins can provide lessons to colloids and polymers.

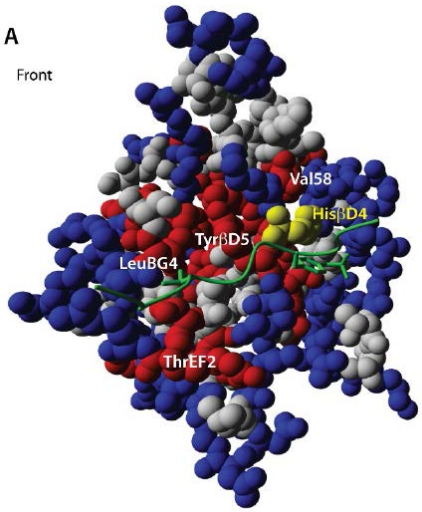
A full issue of the *Biophysical Journal* last September was devoted to the protein-version of information-transfer, which carries the special term 'allostery'. Simply put, a binding event at site A on a protein affects the equilibrium constant of another binding event at a distant site B. The strength of the allosteric effect is quantified by the differences in the binding free energies at A, ΔG, in the two cases of B bound and unbound, so is notated $\Delta\Delta G$. The information that A is occupied is somehow conveyed to site B. This can, of course, occur by structural change: a series of local rearrangements begins at the binding site A, propagates through the protein, and opens up (for example) the site at B. Thermal noise would degrade the effectiveness of such a signal in the usual way. But another mechanism works *with*, rather than *against*, the noise of thermal fluctuations. In this case the binding event at A modulates the effective mechanical properties locally, and so shifts the amplitude and structure of the thermally-activated normal modes of the protein. Such modes may also have amplitude at site B, and when this is the case the *entropy* of binding that arises from the further suppression of fluctuations at B is changed. No mean structural alteration is called-for; rather, the thermally activated mechanical fluctuations of the system act as a sort of 'carrier wave' for the information that certain binding sites are occupied or not. This 'allostery by thermal fluctuation' has been the object of experimental, theoretical and simulation for several years, and called for the special meeting that produced the journal issue.

I could have chosen almost any of the articles from the issue – and there are some excellent ones, including reviews, that are well-worth reading. The paper by Sumbul *et al.* drew my attention, however, as it analyses (by large-scale atomistic MD) a particular protein system in terms of its coarse-grained topology and structural integrity, pointing out the importance of 'hinge regions' of global slow modes in conveying the allosteric signal. For a given normal mode, these regions are defined (well, this is my reader's deduction – a rigorous definition seems to be wanting) as those which sustain a local maximum of bending deformation for that normal mode. Exploring a large

data-base of allosteric proteins and their mutants, the authors found that *a high proportion of single-point mutations that gave rise to changes in the allosteric free energy occurred in hinge regions*.

The paper then goes on to look at one example in detail, the human growth factor protein hGH, which binds to a larger protein (hGHR) to activate it. This binding was very strongly and allosterically modified by mutations at a site that corresponds closely to hinge regions for both of the first two global deformation modes of hGH. Mutations at the hinges were able to remove or reduce dynamic correlations across the protein substantially.

There is much less discussion of allostery in terms of information theory that one might expect. An exception is the paper by Lenaerts *et al*. These authors recast allostery within the frame of *mutual information*. They choose, again, to do this by way of a specific example – the SH2 domain of Fyn tyrosine kinase, a protein complex implicated in several signalling pathways. Mutual information *I(x,y)* is defined as a 2-point generalisation of Shannon entropy in terms of the joint probability distribution of two variables *x* and *y*:



$$I(X;Y) = I(Y;X) = H(X) + H(Y) - H(X,Y)$$

$$H(X) = -\sum_{x \in X} p(x) \log p(x) \quad \text{and} \quad H(X,Y) = -\sum_{x,y \in X,Y} p(x,y) \log p(x,y)$$

In the figure, *x* and *y* are taken to be the local conformational variables of particular groups on the protein. The mutual information function is colour-coded for all other residues *y*, in terms of their correlations with a fixed residue (labelled in yellow). Red indicates larger mutual information, blue less. Notable is the appearance of a network of connected pathways between residues with high mutual information. This network is both activated and modified in the functional process by which the attachment of an effector molecule induces the release of the entire protein domain from a complex (with the SH3 domain).

While it is compelling to explore the idea of mutual information in allosteric signalling, it's not clear that calculations of this measure in the unbound protein alone really give sufficient insight into the information carried by allosteric binding events. It is, after all, the binding of new structures to the larger molecules that constitute the information to be transferred. Nor is it clear that it gives a

clearer view of signalling pathways and connected networks than plots of simple correlation functions of local dynamics. There is no doubt that information is carried by modifying just these correlated fluctuation ensembles. For example, in perturbation theory there is an exact proportionality between the correlation function for the square amplitude of local pre-binding elastic fluctuations *u(x)* at two allosteric sites, and the allosteric free energy on changing the elastic constant at one of them by $\tilde{\kappa}$ (McLeish *et al. [1]*):

$$\Delta\Delta G = -\frac{1}{4kT}\tilde{\kappa}^2\left[\left\langle \left(u(x_1)\right)^2\left(u(x_2)\right)^2\right\rangle_0 - \left\langle \left(u(x_1)\right)^2\right\rangle_0\left\langle \left(u(x_2)\right)^2\right\rangle_0\right]$$

There is also, a very close relationship between the expression above for mutual information, and the theoretical expression for the allosteric free energy, if the latter were applied to the calculation of the entropic part of the free energy. The full allosteric free energy encompasses enthalpic terms as well as entropic. Furthermore, these can be intimately coupled to the entropic processes of fluctuation distributions (tighter fluctuations typically cluster around lower energy regions of conformation). For all these reasons, there is more to do before an information-theoretic approach to allosteric communication in proteins is really illuminated by it. However, there is a great deal of potential in thinking (and experimenting) this way in more general soft matter.

References

1. T.C. B. McLeish, T. L. Rodgers and M. R. Wilson,' Allostery without conformation change: modelling protein dynamics at multiple scales', *Phys. Biol.* **10** 056004 (2013)