Dynamic Complexity during Folding in molecular systems

Dmitry Nerukh, George Karvounis *, and Robert Glen

Unilever Cambridge Centre for Molecular Informatics
Department of Chemistry
University of Cambridge
Cambridge
CB2 1EW
UK

The idea

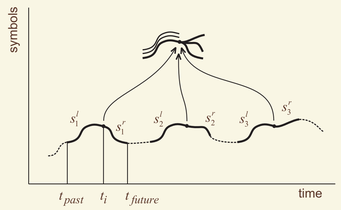
- The dynamics of molecular systems is **chaotic** having different degrees of randomness in different basins of attraction on the potential surface
- The aim is to capture the **emergence** of structure that appears in dynamical systems
- This intrinsic emergence leads to an increase of **computational capability** which we want to measure and analyse
- We address issues such as **innovation**, discovery and measurement of patterns, quantification of emergence and **self-organisation**
- In computational terms, we want to measure the **dynamic complexity** of a process
- Related with self-organisation and emergence COMPLEXITY estimates how sophisticated are the **dynamical laws** governing the evolution of a process

The method

- The approach by Crutchfield *et.al.* termed "Computational Mechanics" is adopted, which incorporates the ideas from Shannon entropy and algorithmic complexity theories, measuring the size of the informational description of a process

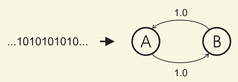
- Symbolic dynamics is analysed, that is real signal first converted into a sequence of symbols from the alphabet of a finite size

- The algorithm searches for regularities by analysing the pairs of "pasts" and "futures" in pieces of the process. Grouping the "pasts" that lead to the same "futures" it forms predictive states, "causal states":



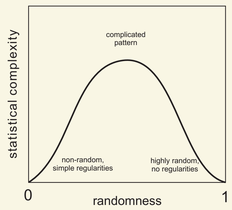

```
Causal State:
Pr(future|past1) = Pr(future|past2)
```

- The set of causal states and transitions between them constitutes an algorithmic machine (" -machine") that allows to build the statistically minimal "optimally predictive model" that is capable of reproducing the system's dynamics



- Finite Statistical Complexity is calculated from the distribution `Pr(S)` of the causal states as measured by the Shannon entropy:

$$C = -\Pr(S)\log_2\Pr(S)$$

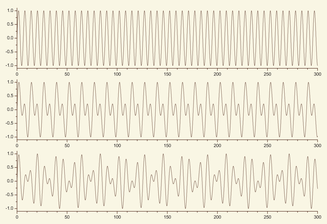


- The complexity units are bits, so it measures the informational contents of the process.

- Its values lie between fully predictable and completely stochastic mechanisms

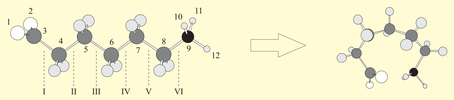
- Example:

```
Sin(x)                          low complexity
sin(ax)+sin(bx)  (a/b=2)
sin(ax)+sin(bx)  (a/b=10/7)     high complexity
```



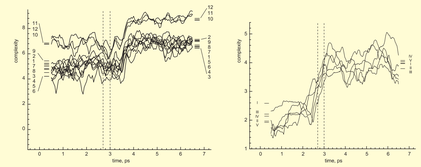
Test molecular systems

Zwitterion in vacuum



- Possesses two distinct dynamic states

- Complexity of atomic trajectories and dihedral angles show different values before and after the transition between the states

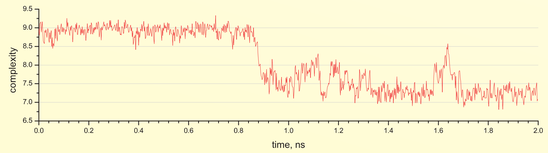


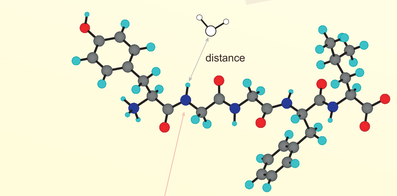
Leu-enkephaline in water

- Protein folding is considered as an intrinsic, self-driven phenomenon that emerges from chaotic dynamics of molecular system

- Trajectories of the protein as well as water atoms are used to calculate the complexity

- The dynamics of the backbone atoms demonstrate changes in the degree of complexity during the folding process



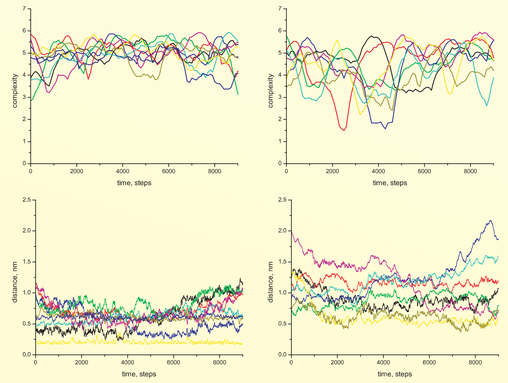


orientational dynamics complexity of surrounding water molecules in the vicinity of the turn formation moment (~4000)

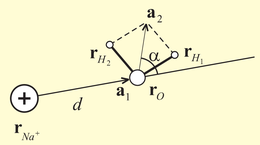
- Complexity of orientational dynamics (cos time dependence) of the water molecules located at various distances from Glycine is calculated

- The moment of turn formation is analyzed

- Around this moment water molecules from the second solvation shell exhibit a minimum in complexity



Na^+ ion in water



- Statistical complexity is sensitive to the details of water molecule dynamics in the solvation shell of the ion and in the bulk.

- The reorientational and translational motion of the bulk molecules show distinctively more complex character then those in the field of the ion.

Acknowledgments

This work is supported by
The Newton Trust and Unilever