

# Dynamic Complexity of Chaotic Transitions in High-Dimensional Classical Dynamics: Leu-Enkephalin Folding

Dmitry Nerukh, George Karvounis, and Robert C. Glen

Unilever Centre for Molecular Informatics, Department of Chemistry,  
Cambridge University, Cambridge CB2 1EW, UK  
Tel.: +44 (0)1223 763078; Fax: +44 (0)1223 763076.  
dn232@cam.ac.uk

**Abstract.** Leu-Enkephalin in explicit water is simulated using classical molecular dynamics. A  $\beta$ -turn transition is investigated by calculating the topological complexity (in the “computational mechanics” framework [J. P. Crutchfield and K. Young, *Phys. Rev. Lett.*, **63**, 105 (1989)]) of the dynamics of both the peptide and the neighbouring water molecules. The complexity of the atomic trajectories of the (relatively short) simulations used in this study reflect the degree of phase space mixing in the system. It is demonstrated that the dynamic complexity of the hydrogen atoms of the peptide and almost all of the hydrogens of the neighbouring waters exhibit a minimum precisely at the moment of the  $\beta$ -turn transition. This indicates the appearance of simplified periodic patterns in the atomic motion, which could correspond to high-dimensional tori in the phase space. It is hypothesized that this behaviour is the manifestation of the effect described in the approach to molecular transitions by Komatsuzaki and Berry [T. Komatsuzaki and R.S. Berry, *Adv. Chem. Phys.*, **123**, 79 (2002)], where a “quasi-regular” dynamics at the transition is suggested. Therefore, for the first time, the less chaotic character of the folding transition in a realistic molecular system is demonstrated.

**Keywords:** folding, enkephalin, dynamic complexity, computational mechanics.

## 1 Introduction

Self-organisation in molecular systems is ubiquitous (micelles, membranes, crystal structures etc.). This process is vital for the proper functioning of biomolecules, the most prominent example being protein folding. Despite the dynamic character of the process, the majority of the literature is mostly devoted to the structural aspects of protein folding; a significantly lower number of publications are devoted to folding dynamics and very few (if any) works exist that consider the problem from the self-organisation point of view. There are difficulties in simulating protein folding which derive mostly from the size of the molecular systems and ultimately, the complexity of both their structure and dynamics. However, research in the field has evolved and it

is now possible to investigate computational models of sufficient quality and size over a realistic time scale covering, in some cases, a complete folding event.

Protein folding is currently described using the notion of a “thermodynamic folding funnel”<sup>1</sup> which generally concentrates on the structural aspects of folding while explaining some aspects of the dynamics (particularly the selection of favoured transition structures). However, there is evidence that the funnel approach may fail to explain the folding process, its mechanisms and speed<sup>2</sup>. Therefore, we believe that direct analysis of the folding dynamics using the apparatus of non-linear dynamical systems and dynamic chaos is one of the most promising ways of moving towards understanding the protein folding phenomenon.

Within this framework, the study of dynamic complexity gives new insights into the details of the dynamics of molecular systems<sup>3</sup>. In particular, it is shown that the methodology we have developed is capable of distinguishing between different regimes of molecular motion which have different chaoticity<sup>4</sup>. This suggests that the computation of complexity may help to understand the features that characterise self-organisation of molecules in time frames much shorter than those predicted by e.g. the thermodynamic folding funnel approach.

The methodology we have developed is based on the computational mechanics theory developed by Crutchfield *et al*<sup>5,6,7,8</sup>. By “discovering” dynamic patterns rather than by postulating them (like, for example, in Fourier analysis), this theory is designed to elucidate the property of emergence in dynamic systems. That is, to uncover qualitatively new behaviour when it appears. This is exactly the essence of self-organisation and, to our view, a perception currently missing in the analysis of dynamic molecular systems.

Another aspect long recognized as fundamental to protein folding studies is of course the role and importance of water. There are a growing number of publications proving both the decisive role of water<sup>9-14</sup> and also the very different dynamics of water molecules at different locations with respect to the biomolecule<sup>15</sup>. This, on the one hand complicates the situation (since a significant number of water molecules now needs to be included into simulations of the essential “core” of the system) and on the other hand, indicates new directions of research to explain phenomenologically the observed folding events.

In a recent publication of a computer simulation of the folding of Leu-Enkephalin, we reported some key observations of the dynamics of the water shell around the peptide during folding<sup>16</sup>. Specifically, during the short time period of the folding event, at the moment of folding when  $i$  and  $i+3$   $C_\alpha$  atoms come together (this nomenclature is described in reference 15), the water network around the peptide undergoes significant collective changes. By making and breaking water-water and peptide-water hydrogen-bonds the whole network becomes more mobile at the moment of folding, in contrast to having a more “frozen” dynamics before and after the folding event.

Even though the formation of  $\beta$ -turns in proteins is well defined (through the distance between  $i$  and  $i+3$   $C_\alpha$  atoms) there is no explanation of the driving force for the molecule to undergo the transformation. There is no apparent reason why a protein fluctuating in the conformation close to a  $\beta$ -turn for a relatively long time suddenly and quickly enters into the transition regime and acquires the specific folded configuration. In the present study we link this behaviour to a general property of

dynamic systems – the phase space mixing or, loosely speaking, their chaoticity. By calculating the dynamic complexity of the system's trajectories we demonstrate a substantial decrease of chaoticity specifically at the moment of transition. We hypothesise that this phenomenon can be attributed to a general theory of transitions in molecular systems that predicts the non-chaotic character of the system's dynamics at these transitions<sup>17</sup>.

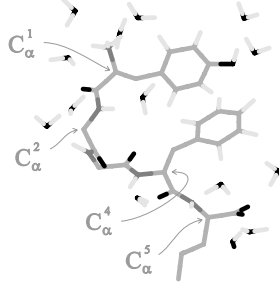
## 2 The System

All-atom classical molecular dynamics simulations of a Leu-Enkephalin molecule in explicit water have been carried out. Analysis of the dynamics of the atoms of the peptide, as well as the first solvation shell and some of the second solvation shell waters has been performed. The same analysis for a small group of water molecules in the bulk was performed for testing purposes.

The Leu-Enkephalin molecule ( $\text{NH}^{3+}$ -TYR-GLY-GLY-PHE-LEU-COO<sup>-</sup>), Fig. 1, was built using Sybyl6.9<sup>18</sup> in a linear conformation in a zwitterionic form (with N-terminal  $\text{NH}^{3+}$  and C-terminal  $\text{COO}^-$ ). It was solvated in a rectangular box of 5826 SPC<sup>19</sup> water molecules and energy minimized using the steepest descent method. The GROMOS96<sup>20</sup> force field was used in all simulations. Four simulations of 3ns (of 0.002 ps time steps) were performed with weak coupling to a water bath of constant temperature at 300 K (peptide and solvent individually), with a coupling time of 0.1ps<sup>21</sup>. Pressure coupling was also applied to a pressure bath with reference pressure of 1 bar and a coupling time of 0.1ps. All bond-lengths of the peptide were constrained using the SHAKE<sup>22</sup> algorithm. The SETTLE<sup>23</sup> algorithm was used to constrain the bond lengths and angles of the SPC waters. The simulations were performed using the GROMACS<sup>19</sup> package. A 1 nm cut-off distance for both van der Waals and Coulomb potentials was used. The system was allowed to equilibrate for the first 100ps before collecting data for analysis.

A  $\beta$ -turn was defined for 4 consecutive residues if the distance between the  $C_\alpha$  atom of residue  $i$  and the  $C_\alpha$  atom of residue  $i+3$  was less than  $7\text{\AA}$  and if the two central residues are not helical<sup>24</sup>. There are two possible  $\beta$ -turns in Leu-Enkephalin: 1-4 and 2-5. As has previously been reported in the literature<sup>25</sup> and observed in this study, the  $\beta$ -turn in this molecule is not stable, however, it survives for relatively long periods of time and is created very quickly. This kind of event in this small peptide is probably consistent with similar events found during protein folding.

We have chosen four moments of  $\beta$ -turn formation for preliminary inspection. We would like to stress that this formal  $\beta$ -turn definition is not exhaustive. In terms of both structure and dynamics, in all four cases we observed somewhat different results. These differences are reflected in the configuration of the side chains, dynamics of the  $C_\alpha$  distances, etc. Of course, each simulation starts from a different starting point and would be expected to be different. Two out of the four events exhibited the effect of formation of the  $\beta$ -turn and we have selected one transition for detailed analysis. This transition was chosen as it shows the most pronounced effect in dynamic complexity.



**Fig. 1.** A Leu-Enkephalin molecule with hydrogen-bonded waters at the moment of transition. The colours represent (from grey to black): H, C, N, O atoms.

### 3 The Method: Statistical Complexity of Deterministic Chaotic Systems

#### 3.1 General Idea

The complexity measure used in this work is described in detail in ref. 3. Here we give an outline of the theory and specific characteristics used to quantify dynamic complexity.

In the computational mechanics framework<sup>5,6,7</sup> symbolic dynamics is considered i.e. the signal consists of discrete symbols assigned to discrete time steps. Let a set of symbols corresponding to each time step  $t_i$  form a sequence  $S$ . To calculate the statistical complexity,  $S$  is decomposed into a set of left  $s_i^l$  (past) of length  $l$  and right  $s_i^r$  (future) of length  $r$  halves joined together at time points  $t_i$ . Consider all equivalent left subsequences  $s_k^l$ . Collect a set of all right subsequences following this unique left subsequence. Each right subsequence has its probability conditioned on the particular left one:  $\Pr(s^r | s_k^l)$ . The equivalence relation between any two left subsequences is defined as follows: two unique left subsequences  $s_i^l$  and  $s_j^l$  are equivalent if their right distributions are the same up to some tolerance value  $\delta$ :

$$\Pr(s^r | s_i^l) = \Pr(s^r | s_j^l) + \delta. \quad (1)$$

A set of all equivalent left subsequences forms an “equivalence class”. The equivalence classes have their own probabilities ( $\Pr(A_i)$ ) calculated from the probabilities of the constituent left subsequences.

The importance of the notion of equivalence classes is that they represent the states of the system that define the dynamics at future moments – the “causal states” (here equal to “equivalence classes” with corresponding probabilities). The time evolution of the system can be viewed as traversing from one causal state to the other with a transition probability equal to  $\Pr(s^r | s_i^l)$ . The set of the causal states together with the transition probabilities constitute a so called “ $\epsilon$ -machine”.  $\epsilon$ -machines represent the minimal computation necessary to reproduce the dynamics of the system.

A number of useful mathematical properties of the  $\varepsilon$ -machine are proved<sup>8</sup>. Among others is the fact that the  $\varepsilon$ -machines are monoids (semi-groups with an identity element). From this it follows that causal states are (i) sufficient statistics; (ii) unique; (iii) minimal. These and other properties serve as a fundamental basis for defining a complexity measure that stands out when compared to the majority of current approaches to quantify complexity of dynamic systems.

The statistical complexity is defined as the informational size of the  $\varepsilon$ -machine. The measure of this is the Shannon entropy of the causal states:

$$C_\mu \equiv -\sum_{A_i} \Pr(A_i) \log_2 \Pr(A_i), \quad (2)$$

where  $A_i$  are causal states. In contrast to Kolmogorov complexity, this measure provides a zero complexity for *both* extremes – a constant signal and a purely random process. The maximum value of complexity lies somewhere in between these two limits.

For the purposes of the present study “topological complexity” is important. It is defined as the size of the  $\varepsilon$ -machine:

$$C \equiv \log_2 N, \quad (3)$$

where  $N$  is the number of causal states  $A_i$ . For the reasons explained further, the highest of the transition probabilities (1) of causal states also carry important information about the dynamic patterns of the signal. To estimate it for the entire  $\varepsilon$ -machine, the highest transition probabilities averaged over all states were calculated:

$$D \equiv \frac{1}{N} \sum_i \max_j (\Pr_i^j), \quad (4)$$

where  $\Pr_i^j$  is a probability of a transition  $j$  of a state  $A_i$ .

### 3.2 Application to Molecular Dynamics

We are currently developing a general theory of dynamic complexity of molecular Hamiltonian systems. In this framework we define the complexity of the system as a statistical (or topological) complexity of a signal comprising of the realisations of the system’s phase-space trajectories symbolised using a particular partitioning method. This general definition can be exploited to investigate the properties of molecular systems using the results of chaotic non-linear dynamics from which some nontrivial implications follow.

However, such an analysis requires the information on the full-dimensional trajectories, possibly of infinite lengths. This is, of course, impossible to obtain in a simulation like the one used in the present work. Therefore, a simpler approach is employed, namely the one-atom projections of the full-dimensional trajectories are used and the length of the signal was chosen such that it is long enough to accumulate a useful statistic and short enough not to average out interesting effects of the system’s dynamics. Pieces of trajectory of 4ps long were used to calculate the values

of complexity. These values were then plotted at the times corresponding to the middle of the interval used. The procedure was repeated to obtain the time evolution of the complexity value of the system (Fig. 4,5).

Initially we used the three-dimensional Cartesian coordinates of the atoms as the signals to calculate the complexities. This approach, however, has the following disadvantages. The domain of the coordinates is very large compared to the area a trajectory covers during the time periods studied. In other words, the atoms only drift slowly in small regions (different for all atoms) of the allowed coordinate area. This is a major obstacle for the complexity method described above since the method is statistical in nature. Therefore, it requires a representative sampling of the whole domain of the signal.

Instead we used the velocities of the individual atoms as the signals. These quickly fluctuate around the origin and each undergoes a significant number of fluctuations during the time period of the investigation. Another advantage is that they are of comparable value for different atoms and that makes their comparison possible on the same grounds.

### 3.3 Symbolisation

The velocities were symbolised according to the following scheme. In spherical coordinates, the  $\varphi$  and the  $\theta$  angles were partitioned into  $n$  and  $n/2$  regions respectively and radius  $r$  was partitioned into  $m$  regions. The resulting spherical sectors were assigned a number (in an arbitrary order). A symbol was assigned to each value of the velocity depending upon which sector it lies.

We have tested the partitions for the values of  $n$  from 8 to 16, and  $m$  equal to 1 and 2. Overall, particular combinations of indices  $n$  and  $m$  were not essential; the total number of symbols ( $n \times \frac{n}{2} \times m$ ) was found to be the key parameter. The latter did influence the value of complexity significantly and the appearance of the complexity minimum during the  $\beta$ -turn formation was only observed for a narrow interval of the number of symbols (discussed in the “Results and Discussion” section). Indirectly, the number of symbols defines the time scale of the analysis: for a fixed time interval; small partitioning extracts more information on a fine-scale, while coarse partitioning is only sensible for very long run times.

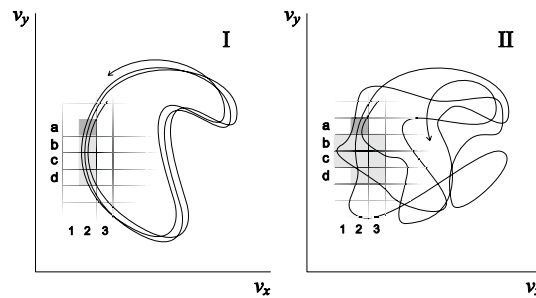
### 3.4 Dynamic Scenarios

There are particular reasons that we observe relations between the values of the parameters used for calculating complexities and the dynamics of the system.

If a trajectory consists of slowly changing values this leads to a situation where the symbolic sequence is comprised of a long series of repeating symbols. In this case, the majority of the transitions occur with the same symbol (for a history length much smaller than the repetition length) and the  $\varepsilon$ -machine consists of the states corresponding to each symbol with the prevailing probability of transition from each state to the same state. In this situation, the complexity reflects essentially the

statistics on the symbols and contains very little information on the dynamical patterns. Analysis of the highest transition probability (4) allows us to recognize this situation: the value of  $D$  in this case is high, reflecting the fact that most transitions are to the same causal state. The topological complexity in this case is close to the logarithm of the number of symbols used  $N_s$ :  $C = \log_2 N_s$ . We call this situation “slow” dynamics.

We should emphasize that the (relatively) short length of the signals introduces a specific feature to computational mechanics, different from its conventional meaning. Namely, when the signal is highly chaotic, that is when almost all subsequences entering probability (1) are different, the  $\epsilon$ -machine reconstruction algorithm does not recognise them as belonging to the same causal state, as should happen for infinitely long signals. Therefore, the large size of the  $\epsilon$ -machine (i.e. the topological complexity (3)) in this case signifies the high chaoticity (high degree of the phase space mixing) of the signal. The statistical complexity (2), however, can be of low value, depending on a particular distribution of subsequencies. This may contradict the normal interpretation of low complexity of a highly random (less structured) signal and should be taken into account in the analysis.



**Fig. 2.** Schematic representation of the trajectories of different chaoticity and symbolic sequences they produce. In case I the trajectory generates three identical symbolic sequences:  $\{a2,b2,c2,d2\}$ . In case II the sequences started with the same symbol  $a2$ , but because of high chaoticity generate three different sequences:  $\{a2,b2,b3,c3\}$ ,  $\{a2,b2,c2,d2\}$ ,  $\{a2,b1,c1,d2\}$ .

For chaotic signals generated by the atomic trajectories (evidence of the chaotic character of molecular trajectories is given in e.g. ref. 26, 27 through the calculation of positive Lyapunov exponents), particular importance can be assigned to the lengths of the left and right subsequence entered in the formula for probabilities (1), in other words, the memory of the system. The degree of chaoticity can be revealed by analysing the change in topological complexity  $C$  with increasing memory length. For signals having persistent dynamic structures the value of  $C$  will not change significantly, whereas for a chaotic signal, because of high sensitivity to the initial conditions,  $C$  will quickly rise with the length of the memory since quickly diverging trajectories will generate more and more new unique subsequences, i.e. causal states (Fig. 2).

## 4 Results and Discussion

The complexities obtained for the atoms investigated can be classified into two categories. All hydrogen atoms and some water oxygen atoms belong to the first category, the rest of the atoms – to the second category. The latter is characterised by the “slow” dynamics described in the previous chapter. Besides the apparent prevalence of a long series of the same symbol, this is also supported by the relatively high value of  $D$  (4). For both classes its value does not significantly change with time and lies in the range from 0.8 to 0.9 for “slow” dynamics and from 0.6 to 0.65 for “fast” dynamics.

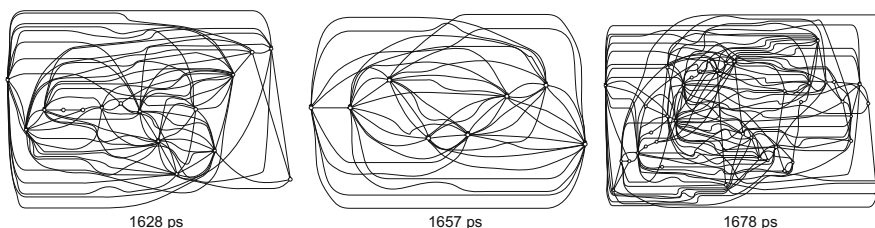
For the “fast” atoms (mostly the hydrogens of the system), the main result of this work is observed: we observe low topological complexity values exactly at the moment of  $\beta$ -turn formation, being high at all other times (Fig. 4,5). It should be stressed that this effect is only visible in a narrow range of the complexity algorithm parameters: the number of symbols from 8 to 36 and the memory length – 3-4 ps. For too many symbols or too long a memory length the dynamics appear “random” (chaotic) for all time periods. For too few symbols, when the symbolisation is too coarse, or for too short a memory length, the dynamics shows a value typically lower than the logarithm of the number of symbols (that does not change with time).

Even though the “gap” in topological complexity is in most cases self-evident on the graphs for sets of atoms, to quantify the minimum value of complexity the averaged complexities  $\bar{C}$  at each time  $t_i$  for  $u$  number of atoms was calculated:

$$\bar{C}_i = \frac{1}{u} \sum_{l=1}^u \frac{1}{b} \sum_{k=i-b/2}^{i+b/2} C_{lk}, \quad (5)$$

where  $C_{lk}$  is a topological complexity for atom  $l$  at time  $t_k$  and  $b$  (typically equal to 6) is the averaging interval. The averaging over time for  $b$  data points is needed because the “gap” in complexity for different atoms was sometimes misaligned for 1-2 data points, especially for the peptide’s atoms (Fig. 4). The averaging thus emphasises an effect that would otherwise be difficult to recognise.

Typical  $\varepsilon$ -machines obtained for the signals investigated are shown in Fig 3. The apparently less sophisticated structure can be seen for the transition moment of 1657



**Fig. 3.**  $\varepsilon$ -machines of a representative atomic trajectory at the moments before, at (1657ps), and after the  $\beta$ -turn formation. Dots represent the causal states, arcs – the transitions between the states. A state can lead to several other states with corresponding probabilities.



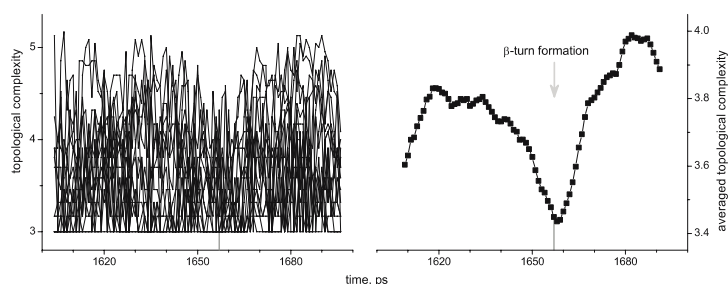
ps. It should be stressed that the  $\epsilon$ -machine structure reflects the dynamic patterns present in the signal, not the statistical features of the symbols comprising the signal. This is important because it proves that the dynamics does not become “slow” at the moment of transition, which is also confirmed by the value of  $D$  as discussed above.

The original and averaged topological complexity of the peptide’s atoms and its evolution in time are shown in Fig. 4. Not all atoms have the complexity minimum at exactly the transition time, having a small difference of a few ps. This makes the minimum in the complexity less pronounced in the summary graph of the original complexities (lower plot), but is clearly emphasised by the averaging procedure (upper plot). This is in contrast to the water atoms where the minimum falls exactly at the moment of transition (Fig. 5). The reasons for this may be that the peptide’s atoms interact with each other much more strongly and the transition path has more complicated character, whereas the weaker interactions of water, still being involved in the dynamics of transition, serve as a “bath”. Another possibility may be that in the full-dimensional space the low complexity value is observed for the wider period covering all the atoms, while for the projections studied here, it falls at slightly different times.

The complexities of the water atoms are shown in Fig. 5. Here the minimum in the complexity is most clearly visible for both the original values as well as for the averaged values of  $C$ .

To check if this behaviour is not simply a result of appropriately selected water molecules, we have performed the following test. At different moments (1616, 1640, 1657, 1672, and 1692 ps) a group of about 50 water molecules belonging to the first solvation shell of the peptide was chosen. The complexities of the atoms of these molecules were then calculated and plotted in a similar fashion to Fig. 5. If the effect (the decrease in complexity) was only the result of proximity of the water molecules to the peptide, it would be observed in all 5 test calculations at the chosen times. In reality it was only observed at 1657 ps, the moment of  $\beta$ -turn formation, therefore strongly implying that the dynamics of the system is the reason for this behaviour.

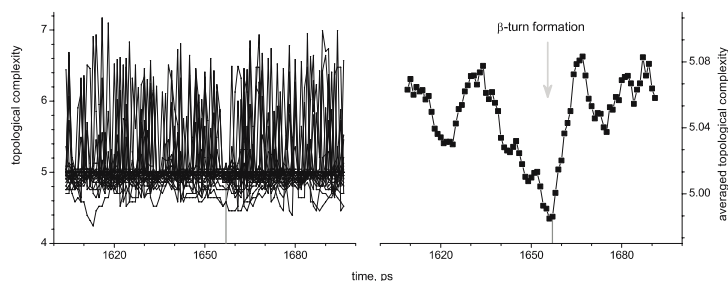
For the plots presented in Fig. 5 the complexities of 111 water atoms were calculated. Only one atom showed a high value at 1657 ps (excluded from the plot for clarity). This one atom is specific in that it is part of a peptide – water – peptide hydrogen-bonding bridge at the moment of transition. This, most probably, makes its dynamics special and different from the majority of water molecules neighbouring the peptide.



**Fig. 4.** Topological complexities of the peptide’s atoms. A symbolisation alphabet of size 8 and a history length of 3 ps were used. The  $\beta$ -turn transition is at 1657 ps.

Another subtle observation worth mentioning here is a much lower spread of values of complexities at the moment of transition, not only from above, but also from below (Fig. 5). In other words, the lines become on average slightly closer to each other. This is in line with our further explanation of the phenomenon.

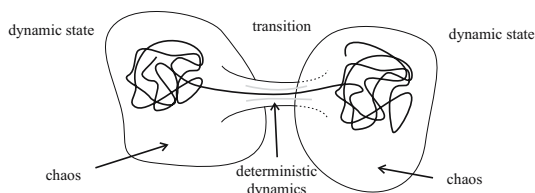
We would like to point out the connections between the phenomenon found here and our previous investigations. In ref. 16 we reported the change in the dynamics of the water network around Leu-Enkephalin during the  $\beta$ -turn formation. We observed that the water network exhibits a “freezing” behaviour before and after the transition with high mobility and the disruption of hydrogen-bonding at the moment of transition. We now see evidence of a completely different characteristic that directly investigates the dynamics of the atoms and that also demonstrates a fundamentally different character of the dynamics of the water before, at, and after the transition.



**Fig. 5.** Topological complexities of the water atoms. A symbolisation alphabet of size 32 and a history length of 4 ps were used. The  $\beta$ -turn transition is at 1657 ps.

Also, these results correlate well with our calculations of the complexity of a simpler (model) molecular system: a zwitterion that showed similar changes in statistical complexity during the transition between different conformational states<sup>4</sup>.

The authors of ref. 17 suggest that the dynamics of the states between the transitions is significantly chaotic while at the moment of the transition the dynamics becomes semi-chaotic or quasi-regular i.e. the system can maintain approximate constants of motion and possess fully deterministic dynamics. We hypothesise that the effect reported in this work is the manifestation of this phenomenon. Indeed, as discussed above, the low complexity value in our case corresponds to less chaotic motion. This indicates the appearance of simplified periodic patterns in the atomic



**Fig. 6.** Schematic illustration of the transitions between different phase-space regions of the system

motion, which could correspond to high-dimensional tori in the phase space. As a schematic illustration of the dynamics, the transition can be visualised as passing through a narrow “tunnel” connecting two states. In this situation the phase-space flow should “straighten” (or rather confined to the tori) in order to be transferred from one basin to the other, Fig. 6.

## 5 Conclusion

Summarising, the complexity of the dynamics of the Leu-Enkephalin – water system has been investigated paying particular attention to the moment of  $\beta$ -turn formation. The dynamic complexity of the atomic velocities has been calculated and its evolution during the fold transition has been analysed. It is found that exactly at the moment of transition, the complexity shows a significant drop that in this framework indicates a low degree of phase space mixing (less chaotic character) of the dynamics of the system.

Work is planned to extend this complexity analysis to other molecules and elementary folding events such as the formation of  $\alpha$  helices. We also plan to link this behaviour to what we think may be a general characteristic of biomolecular systems during their self-organisation, that is changes in the complexity of the dynamics of the system.

**Acknowledgement.** The work is supported by the Isaac Newton Trust and Unilever.

## References

1. N.D. Socci, J.N. Onuchic, P.G. Wolynes, Protein folding mechanisms and the multidimensional folding funnel, *Proteins*, **32** (2), 136 (1998).
2. S.V. Krivov and M. Karplus, Hidden complexity of free energy surfaces for peptide (protein) folding, *Proc. Nat. Acad. Sci.*, **101** (41), 14766 (2004).
3. D. Nerukh, G. Karvounis, and R.C. Glen, Complexity of classical dynamics of molecular systems. Part I: methodology, *J. Chem. Phys.*, **117**(21), 9611 (2002); Complexity of classical dynamics of molecular systems. Part II: finite statistical complexity of a water-Na<sup>+</sup> system, *ibid*, 9618.
4. D. Nerukh, G. Karvounis, and R.C. Glen, Quantifying the complexity of chaos in multi-basin multidimensional dynamics of molecular systems, *Complexity*, **10** (2), (2004).
5. J. P. Crutchfield and K. Young, Inferring statistical complexity Phys. Rev. Lett., **63**, 105 (1989).
6. J. P. Crutchfield and K. Young, Computation at the Onset of Chaos, in *Entropy, Complexity, and Physics of Information, SFI Studies in the Sciences of Complexity, VIII*, edited by W. Zurek (Addison-Wesley, Reading, Massachusetts (1990)).
7. J. P. Crutchfield, The Calculi of Emergence: Computation, Dynamics, and Induction, *Physica D* **75**, 11 (1994).
8. Cosma Rohilla Shalizi, *Causal Architecture, Complexity and Self-Organization in Time Series and Cellular Automata*, PhD thesis, University of Wisconsin at Madison, 2001.
9. D. Beck, D. Alonso, V. Daggett, A microscopic view of peptide and protein salvation, *Biophys. Chem.* **100**, 221 (2003).

10. J. M. Sorenson, G. Hura, A. K. Soper, A. Petsemidis and T. Head-Gordon; Determining the Role of Hydration Forces in Protein Folding, *J. Phys. Chem. B* **103** (26), 5413 (1999).
11. A.E.Garcia and G. Hummer, Water penetration and escape in proteins, *PROTEINS: Struct, Funct. and Genet.* **38**, 261-272 (2000).
12. S. Dennis, C.J. Camacho and S. Vajda, Continuum electrostatic analysis of preferred solvation sites around proteins in solution, *PROTEINS: Struct, Funct. and Genet.* **38**, 176 (2000).
13. R. L. Baldwin, Relation between peptide backbone solvation and the energetics of peptide hydrogen bonds, *Biophys. Chem.* **101-102**, 203 (2002).
14. S. M. Bhattacharyya, Z. Wang and A. H. Zewail, Dynamics of Water near a Protein Surface, *J. Phys. Chem. B* **107** (107), 13218 (2003).
15. P.W. Fenimore, H. Frauenfelder, B.H. McMahon, and R.D. Young, Bulk-solvent and hydration-shell fluctuations, similar to - and -fluctuations in glasses, control protein motions and functions, *Proc. Nat. Acad. Sci.*, **101** (40), 14408 (2004).
16. G. Karvounis, D. Nerukh, and R.C. Glen, Water network dynamics at the critical moment of a peptide's beta-turn formation: an MD study, *J. Chem. Phys.*, **121** (10), 4925 (2004).
17. T. Komatsuzaki and R.S. Berry, Chemical Reaction Dynamics: Many-Body Chaos and Regularity, *Adv. Chem. Phys.* **123**, 79 (2002).
18. Sybyl [molecular modeling package], version 6.8. St Louis (MO): Tripos Associates; (2000).
19. H. J. C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, J. Hermans, Interaction Model for Water in Relation to Protein Hydration, in *Intermolecular Forces*, edited by B. Pullman, (D. Reidel Publishing Company, Dordrecht, 1981), pp. 331-342.
20. W.R.P. Scott, P.H. Hunenberger, I.G. Tironi, A.E. Mark, S.R. Billeter, J. Fennen, A.E. Torda, T. Huber, P. Kruger, and W.F. van Gunsteren, The GROMOS Biomolecular Simulation Program Package, *J. Phys. Chem.*, **A 103**, 3596 (1999).
21. W.F. van Gunsteren and H.J.C. Berendsen, Computer simulation of molecular dynamics, *Angew. Chem. Int. Ed. Engl.*, **29**, 992 (1990).
22. J. P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, *J. Comput. Phys.*, **23**, 327 (1977).
23. S. Miyamoto and P. A. Kollman, An analytical version of the SHAKE and RATTLE algorithm for rigid water models, *J. Comp. Chem.*, **13**, 952 (1992).
24. E.G. Hutchinson, J.M. Thornton, PROMOTIF--A program to identify and analyze structural motifs in proteins, *Prot. Sci.*, **5**, 212 (1996).
25. D. van der Spoel, H. J. C. Berendsen, Molecular dynamics simulations of Leu-enkephalin in water and DMSO, *Biophys J.*, **72**, 2032 (1997).
26. M. Braxenthaler, R. Unger, D. Auerbach, J.A. Given, and J. Moult, Chaos in protein dynamics, *Proteins: structure, function, and genetics*, **29**, 417 (1997).
27. Huai-bei Zhou, Chaos in Biomolecular Dynamics, *J. Phys. Chem.*, **100**, 8101 (1996).