cecam Centre Européen de Calcul Atomique et Moléculaire	**Workshop Scientific Report**
Title	**Protein Folding Dynamics: Bridging the Gap between Theory and Experiment**
Organizers	**Fabio Pietrucci** (CECAM - Swiss Federal Institute of Technology Lausanne, Switzerland) **Ruhong Zhou** (IBM Watson Research Center, and Department of Chemistry, Columbia University, USA)

Scope of the workshop

The scope of the workshop was to bring the most prominent scientists in the field of protein folding dynamics together with younger researchers, with the aim of discussing the state-of-the-art in the field and the most promising current and future directions. In recent years experimental techniques showed considerable progress, allowing to probe folding events at high spatial resolution (down to single molecules) and also temporal resolution (to the limit of 100 ns and also for individual folding events). Computational techniques on the other hand greatly increased their reach thanks to improved hardware and to several advanced sampling methods, allowing for an ensemble thermodynamic and kinetic view of the folding process for small proteins. The workshop had a particular emphasis on the kinetics of protein folding, (un)folding intermediates, as well as discussions on effective strategies to interpret the former experimental data with the latter theoretical tools.

It should be noted that the top-level and number of invited speakers characterized the workshop as one of the most important which ever took place in Europe on the subject of protein folding dynamics.

Main outcomes of key presentations

The first day opened with the remarkable presentations of E. Shakhnovich from Harvard for theoretical work ("Protein folding in silico and in cells") and J. King from MIT for experimental work ("Sequential in vitro refolding of the Greek key domains of human eye lens gammaD-crystallin"), which set the stage in a very effective way for the rest of the workshop. In particular, the first presentation introduced the power and effectiveness of simplified mathematical models in capturing some key thermodynamic and kinetic aspects of experiments on complex systems. Likewise, the second presentation introduced the biological perspective of protein folding starting from the example of the human eye crystallin, where folding and misfolding processes show their complexity (both structural and in terms of time scales) and functional relevance (in this case being connected to the transparency of the crystallin, necessary for vision).

The multifaceted experimental approach to protein folding has been efficaciously reconstructed by the presentations of, among others, M. Vendruscolo ("Advances in the characterization of free energy landscapes of proteins by NMR spectroscopy"), W. A. Eaton ("Single molecule photon trajectories and transition paths in protein folding"), B. Schuler ("Surprises in the behavior of unfolded proteins from single molecule fluorescence spectroscopy"), and S. Jackson ("A tangled problem: the structure, function and folding of knotted proteins"). What is remarkable is the tight interconnection between advanced experimental techniques (NMR, fluorescence, mutagenesis, etc.) and sophisticated computational approaches, which often are mandatory to extract useful data from the experiments. We spot out the presentation of W. Peti ("Folding upon binding – an ensemble view"), which presented some encouraging results about the characterization of the structure, and interaction among intrinsically disordered proteins, which represent a large fraction of all proteins.

The state-of-the-art on simplified mathematical models of protein folding thermodynamics and kinetics have been reviewed in particular by (besides the already cited E. Shakhnovich) G. Hummer ("Diffusion models of protein folding"), E. Paci ("Free-Energy Surface of Proteins from Reversible Folding Simulations"), and D. Thirumalai ("Chaperonin-mediated protein folding: Realities and Models"). This last presentation is remarkable for the attempt to approach the complexity of folding in the real cellular environment, including the assistance of biological nanomachines like chaperonins, which so far eludes the possibility of high-detail simulations. In the same spirit is also the presentation of J.-E. Shea ("Simulations of Protein Aggregation in the Cellular Milieu") which tackle by a coarse-grained approach the intricated problem of protein-protein interaction in the crowded cellular environment.
Finally, several presentations have been devoted to the very detailed (but also computationally intensive) technique of atomistic molecular dynamics, which greatly enlarged our understanding of the folding dynamics. Progress in this field appears to depend on improving the accuracy of force-fields, as repeatedly pointed out during the workshop, as well as on enlarging the time scales. The presentation of S. Piana ("Equilibrium MD simulations of protein folding") showed how in some cases the impressive power of novel dedicated hardware can provide equilibrium folding trajectories of small proteins.

However to access folding times beyond the millisecond barrier and to reconstruct the free energy landscape enhanced sampling approaches look crucial, as reviewed by P. Bolhuis ("Understanding the kinetics of protein folding and aggregation through multiple simulation techniques"). An emerging paradigm in the analysis of simulation data is that of Markov state models, which has been effectively used in several studies presented in the workshop as a useful framework to reconstruct the macroscopic thermodynamics and kinetics and to compare with both experiments and simplified models. The presentation of V. Pande ("Insight from atomistic simulations of folding on the millisecond timescale and beyond") convincingly summarized the state of this field.

Report on selected discussions
eg. Were there interesting hints for new research? for new developments? for collaborations?

One of the key achievements of this workshop is indeed in the very open and constructive discussions. Many junior researchers were encouraged to participate in discussions. One area of extensive discussions is about the comparison and "converging/diverging" of the three common theories/views on protein folding: the diffusion-collision model, the folding funnel, and the network hub theory (based on Markov State Models). Interesting questions were raised on how Makov State Models (MSMs) can or cannot be reconciled with folding funnels, and/or two-state models. V. Pande, G. Hummer, and J. Onuchic and many others have contributed their views to this hot subject. Many new suggestions, advices, and even criticisms were raised in a very friendly manner for each of these three theories (which is not always true for this protein folding community and discussion can often get so intense that junior researchers are afraid to raise questions). For example, for the MSM model, one of the main criticisms was in the construction of the microstates and how different ways of construction can affect the final network and interpretation. New experiments were also suggested to test and validate these models, such as measuring the properties of folding intermediates, as well as unfolded ensembles (A. Szabo, W. Eaton, S. Jackson).
Another interesting area of broad discussions is in the importance of the crowded cellular environment. Both J. Shea and P. Bolhuis had shown the effect of crowded environment (using hydrophobic walls) on protein folding. How this will emerge into a more detailed description of the cellular environment is of great interest. Multiscale modeling, coarse-grained modeling, and other techniques were suggested and discussed. In addition, as a related topic, the cosovlent effects, such as with urea and GdnHCl, were also discussed heavily as another type of folding environment (R. Zhou, J. King, and S. Jackson). The related protein denaturization mechanism was also discussed in length, as this is one of the remaining problems in protein (un)folding, despite five decades of heavy debates.
The third area of extensive discussions was in the force filed assessment and development. With the massive increase of computer powers (IBM Blue Gene and DE Shaw Anton Supercomputers), both the simulation breadth and depth have increased dramatically in recent years, and thus, how well these commonly used force fields (CHARMM, OPLSAA, AMBER, GROMOS) behave is becoming a renewed hot subject. W. van Gunsteren, S. Piana, V. Pande, R. Zhou and many others had all tried to address this important question from different angles. One improvement to pursue for the standard force fields is to predict more accurately the balance between secondary structures, such as alpha helices vs. beta-sheets. Experimentalists, such as V. Munoz and W. Eaton, had particularly raised the issue with current force fields not being able to capture the temperature dependence. Discussions and suggestions were also made for the polarizable force fields, as well as polarizable water models. The AOMEBA polarizable force field was particularly mentioned and suggested for some folding simulations.
Other interesting questions and discussions include the current PDB database might reflect the bias of available crystallographic techniques but not represent faithfully the variety and complexity of folded proteins (J. King); unfolded ensembles might be overlooked which can carry a lot of useful information about protein folding and unfolding; and the number of knotted proteins are more than people used to think (S. Jackson).

To what extent were the **objectives** of the workshop achieved (strong points, weak points)?

The workshop excellently fulfilled the expectations of very high-level oral presentations and very open and constructive discussions. A second strong point has been the continuous comparison, throughout the whole workshop, of classical paradigms of protein folding (e.g., two-state folding behavior) with the rich dynamical picture (multiple heterogeneous folding pathways) coming from the latest highly detailed simulations and experiments. Another strong point has been the excellent average level of the 16 posters presented, which stimulated a long and well-participated poster session. Overall, as organizers, we very much appreciated the fair and constructive attitude of discussions among the participants, despite the belonging to different “schools”.

A weak point has been the absence of a larger number of biologists, which could have contributed a wider view over the scientific framework surrounding the folding problem. We also notice the lack of representatives from the structural prediction (statistical approaches) community: this has been due to the focus of the workshop on the dynamical processes, however discussions would have probably benefited from a contribution of experts of the latter field.

Do you have suggestions for new workshops/tutorials/conferences on the topic?

The workshop showed how important is to mix prominent experts with younger researchers and also beginning students. We suggest for future events on this topic to even increase the young-to-expert ratio (and the number of poster presentations) in order to include the maximum number of emerging researchers presenting fresh new ideas. The 4-days length of the workshop proved effective and appropriate, with most participants attending for the whole duration. Also the date (beginning of October) suited basically all the participants. We would be happy to organize again a similar event in 2012, as we have been strongly suggested by the participants.