



Laboratory of Molecular Biophysics

The Laboratory of Molecular Biophysics addresses how the structure and motions of proteins relate to disease, with a strong emphasis on diseases that involve the formation of protein aggregates. Members of the laboratory develop new methods to experimentally characterise the flexibility of proteins using Nuclear Magnetic Resonance (NMR) and molecular simulations and then apply these tools to study the molecular bases of specific diseases. The methodological aspect of the work performed by the lab is essential because there is a lack of tools to describe in detail the conformational properties of intrinsically disordered proteins - devoid of persistent secondary and tertiary structure - that are often involved in the formation of aggregates in protein aggregation diseases. The availability of these newly developed tools allows the lab to test the effect of mutation and environmental changes on the properties of these proteins, a procedure that is crucial for unravelling the molecular bases of this class of diseases.

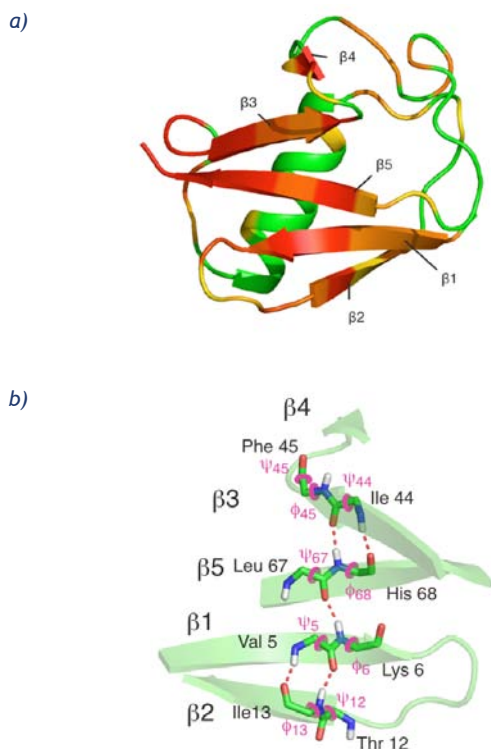
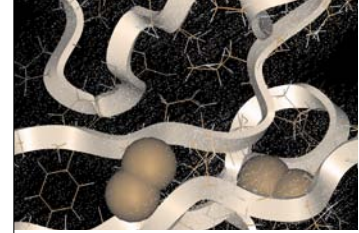


Figure 1. (a) Structure of ubiquitin coloured on the basis of the extent to which the backbone motions of the residues of the protein are correlated to those of residues distant in sequence, where red indicates strong correlation and green absence of correlation. (b) Network of hydrogen bonding interactions that dynamically link the backbone torsion angles of residues Ile 12 (strand $\beta 2$) to Phe45 (strand $\beta 3$), which are separated by a distance of ca. 15 Å.

During 2009 the lab focused its efforts on developing methods to determine the structure and dynamics of proteins from Residual Dipolar Couplings (RDCs) measured using NMR as well as on establishing structure-activity relationships in the aggregates formed by the protein lysozyme in non-neuropathic systemic amyloidoses. These efforts have led to the following: i) the identification of concerted conformational changes that transfer information across the structure of proteins through correlated backbone motions; ii) the development of a method to accurately characterise the structure and dynamics of disordered proteins; and iii) the identification of the structural and dynamical properties of the fibrils formed by lysozyme that render them toxic to cells.

Identification of long-range correlations in a surface patch of folded ubiquitin

Motions in folded proteins play key roles in biological functions ranging from enzyme catalysis to molecular recognition. Correlated motions, which occur when distinct sites in the structure fluctuate in concert, are candidate mechanisms for processes that require information transfer across large distances, such as allostery and signal transduction. In addition, they are thought to underlie folding co-operativity, which prevents proteins from undergoing local conformational changes that expose hydrophobic side chains and can trigger aggregation (for an example see last section). Correlated motions are challenging to observe because they involve concerted local conformational changes of low amplitude that are difficult to detect in the laboratory. RDCs are sensitive to the orientation of bond vectors and have recently revealed correlations in the motions of residues that are close in space.



To study the presence of long-range correlated motions in the backbone of ubiquitin, the laboratory has determined, in collaboration with Christian Griesinger's group at the Max Planck Institute for Biophysical Chemistry in Göttingen (Germany), an ensemble of conformations, named ERNST (ensemble refinement for native proteins using a single alignment tensor), which represents the structure and the sub-ms dynamics of the protein in solution. ERNST was obtained by using ensemble simulations restrained by 36 sets of RDCs and its quality is illustrated by its agreement with independently measured NMR parameters that are sensitive to backbone motions. The average structure of ERNST is also consistent with the static X-ray and the average NMR structures of ubiquitin. Moreover, a structural comparison of the ensemble members with a variety of bound conformations of ubiquitin confirmed the recent finding that conformational selection is the most likely mechanism by which this protein recognises its binding partners. Most importantly, ERNST cross-validates well against cross-correlated relaxation rates and trans-hydrogen bond scalar couplings. These NMR parameters are optimal for the analysis of correlated motions because they average on the same timescale as RDCs; however, unlike RDCs, they are sensitive to the relative motions of two distinct sites of the structure. The back-calculated values of cross-correlated relaxation and trans-hydrogen bond scalar couplings from ERNST are in better agreement with experimental data than those computed from the static X-ray structure, the average NMR structure and from other dynamic ensembles.

Using ERNST, the lab has identified and characterised, for the first time, correlations in the motions of residues separated by up to 15 Å (Figure 1) and has shown that they result from concerted conformational changes mediated by the hydrogen bonding network of the protein ubiquitin as well as by steric clashes. This novel collective motion spans four β -strands of ubiquitin and dynamically links residues that form a surface patch recognised by ubiquitin binding domains (UBDs) in ubiquitination, a post-translational modification that controls endocytosis, DNA repair and protein degradation. These results provide a detailed description of the mechanism by which conformational selection operates in molecular recognition and illustrate that the analysis of NMR parameters sensitive to correlated motions is a powerful approach to characterise information transfer in biology. In addition, our findings reveal the mechanism re-

Research Group Members

Group Leader:

Xavier Salvatella

Postdoctoral Fellows:

Santiago Esteban, Robert

Fenwick, Maria Francesca

Mossuto

PhD Student:

Eva De Mol



sponsible for folding co-operativity and are therefore relevant for our understanding of the origin of amyloidosis, as this disease is often triggered by mutations or environmental changes that lead to local unfolding and to the formation of potentially cytotoxic protein aggregates (for an example see last section).

Refinement of ensembles describing unstructured proteins

As shown in the characterisation of the dynamics of folded ubiquitin presented in previous section, RDCs are unique probes of the structural and dynamical properties of biomolecules in the sub-ms time scale and can be used as restraints in ensemble molecular dynamics simulations to study the relationship between macromolecular motion and biological function. Until now, however, this powerful strategy was applicable only to molecules that do not undergo shape changes in the timescale sampled

by RDCs, such as folded ubiquitin, thus preventing the study of key biological macromolecules such as intrinsically disordered proteins involved in protein aggregation diseases. To circumvent this limitation, researchers in the laboratory have developed the ERIDU algorithm - ensemble refinement of intrinsically disordered and unstructured molecules - that explicitly computes the individual alignment tensors of the ensemble members from their coordinates at each step in the simulation. As a first application, the laboratory has determined an ensemble of conformations that accurately describes the structure and dynamics of chemically denatured ubiquitin, a model system often used to develop algorithms to characterise the structural properties of disordered proteins.

In analogy to dynamic refinement of folded globular proteins, where simulations are initiated from average structures determined using X-ray crystallography or NMR spectroscopy, the researchers used statistical coil models (SCMs) as starting configuration because these represent the best structural descriptions available for unstructured proteins. They found that refinement causes significant structural corrections (Figure 2a) and produced an ensemble that was in complete agreement with experiments and presented transient mid-range inter-residue interactions between strands $\beta 1$ and $\beta 2$ of the native protein (Figure 2b), also observed in other studies based on trans-hydrogen bond scalar couplings and paramagnetic relaxation enhancements. The availability of ERIDU increases the range of systems that can be studied using ensemble simulations restrained by RDCs and will be particularly useful to characterise the conformational properties of intrinsically disordered proteins involved in protein aggregation diseases.

Structure-activity relationships in amyloid fibrils

Amyloid fibrils are non-covalent assemblies of proteins that form in the tissues of patients suffering from protein aggregation diseases, including sporadic and transmissible neurodegenerative disorders as well as various non-neuropathic amyloidoses. The observation that non-fibrillar species transiently populated in the early phase of *in vitro* aggregation are more cytotoxic than the corresponding mature amyloid fibrils suggests that they have generic and currently poorly understood structural properties and add to a significant body of evidence linking the onset of Alzheimer's and Parkinson's diseases with the formation of similar species in the brains of patients.

However, other neurodegenerative disorders, such as prion diseases, are caused by the propagation of infectious particles that carry all the information required to exhibit distinct phenotypic traits in identical hosts and, contrary to what is the case for non-fibrillar oligomers, are clearly fibrillar. These observations raise the possibility that the cytotoxicity of protein aggregates in the biological milieu is not directly related to their oligomeric nature but rather to structural properties common to non-fibrillar and certain fibrillar aggregates. Given that, in contrast to highly evolved native structures, the structures of protein aggregates are influenced by pH, buffer components, protein concentration and temperature, these prospects have led to intense research efforts aimed at establishing structure-activity relationships in protein aggregates.

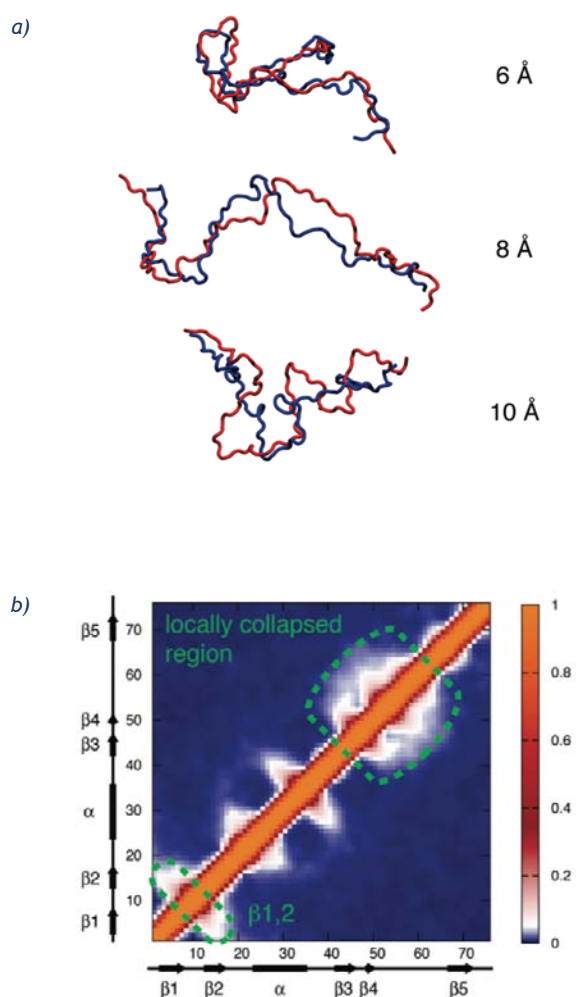


Figure 2. (a) Illustration of how the ERIDU algorithm corrects the conformation of three representative members of the SCM. (b) Contact map of the refined ensemble, where pairs of residues are coloured on the basis of the fraction of ensemble members where the Ca atoms are closer than 10 Å and where the mid-range inter-residue interactions detected using alternative techniques have been highlighted in green.

A particularly useful system to study these structure-activity relationships is lysozyme, a protein that is well characterized and forms amyloid deposits in patients suffering from familial lysozyme systemic amyloidosis, a disease that occurs when amyloidogenic mutations in the protein lead to the formation of partially unfolded amyloidogenic intermediates. By incubating lysozyme under various destabilising conditions, researchers in the lab produced fibrils differing in morphology, molecular structure and stability, thereby generating a range of cytotoxic effects (Figure 3). These results, obtained in collaboration with Christopher Dobson's lab at the University of Cambridge (UK), illustrate that the energy landscape of aggregation is significantly more rugged than the normal folding landscape and that the pathogenic properties of certain protein aggregates are related to the size of their cross- β core.

During 2009, the lab has made relevant contributions to the study of the molecular basis of protein aggregation diseases by developing methods to examine the conformational changes that occur in the earliest stages of disease as well as by identifying the properties of protein aggregates that lead to cytotoxicity. Our goals for 2010 include the application of these newly developed tools and principles to key protein aggregation diseases, such as Alzheimer's and Parkinson's disease, as well as to the rare neurodegenerative disease Spinal Bulbar Muscular Atrophy (SBMA), in which nuclear inclusions of Androgen Receptor (AR) molecules with an expanded polyglutamine tract cause motor neuron death.

Scientific output

Publications

De Simone A, Richter B, Salvatella X and Vendruscolo M. Toward an accurate determination of free energy landscapes in solution states of proteins. *J Am Chem Soc*, 131(11), 3810-11 (2009)

Robustelli P, Cavalli A, Dobson CM, Vendruscolo M and Salvatella X. Folding of small proteins by Monte Carlo simulations with chemical shift restraints without the use of molecular fragment replacement or structural homology. *J Phys Chem B*, 113(22), 7890-96 (2009)

Research networks and grants

Structural characterization of key conformational transitions in protein deposition diseases

Spanish Ministry of Science and Innovation, CTQ2009-08850-BQU (2009-2012)

Principal investigator: Xavier Salvatella

Support a grups de recerca

Agency for Administration of University and Research Grants, 2009-SGR-1514 (2009-2013)

Principal investigator: Xavier Salvatella

Collaborations

Analysis of protein dynamics

Christian Griesinger, Max Planck Institute for Biophysical Chemistry (Goettingen, Germany)

Analysis of RNA dynamics

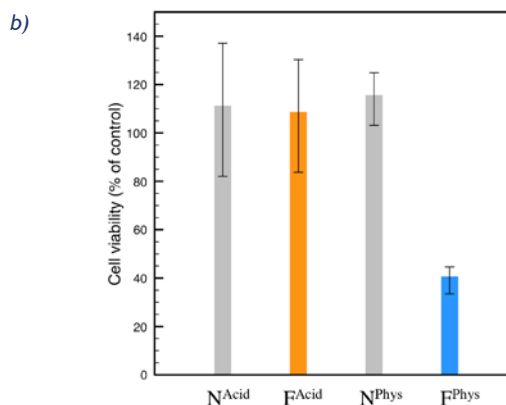
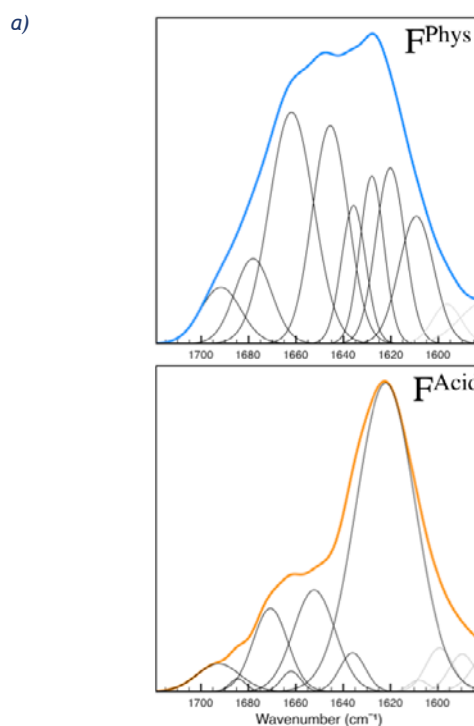


Figure 3. (a) Deconvolution of the FTIR spectrum of fibrils formed under physiological conditions (in blue) and those formed under acidic conditions (in orange) that present, respectively, a low and high degree of cross- β structure that gives rise to the band at 1620 cm^{-1} (right). (b) Viability of SH-SY5Y neuroblastoma exposed to native lysozyme under acidic pH (N^{Acid}), to fibrils formed under acidic conditions (F^{Acid}), to native lysozyme under physiological conditions (N^{Phys}) and to fibrils formed under physiological conditions (F^{Phys}).

Modesto Orozco, IRB Barcelona (Barcelona, Spain)

Structure-activity relationships in protein aggregates
Christopher Dobson, University of Cambridge (Cambridge, UK)

Use of protein native ensembles in drug discovery
Crystax Pharmaceuticals (Barcelona, Spain)

Use of protein native ensembles in protein-protein docking
Juan Fernández-Recio, Barcelona Supercomputing Centre (Barcelona, Spain)