SHARCNET Dedicated Resources Application

PART I: Applicant Information

Dr. Régis Pomès
Molecular Structure and Function
Hospital for Sick Children
555 University Avenue
Toronto, ON
M5G 1X8
pomes@sickkids.ca
1-416-813-5686

PART II: Program Applied for:

Dedicated Resources

PART III: Allocation

System on which CPU is requested: whale
CPU time requested: 140 processors x 6 months (180 days) = 604800 CPU hours
Storage required: 200 GB for 6 months

PART IV: Title of Research Project

Uncovering the Molecular Basis of Elasticity in Biological Tissues

PART V: Description of Research Project

Elastomeric proteins provide the elastic recoil necessary for biological machinery as diverse as the mammalian arterial wall, the capture spiral of spider webs, and the hinge of scallop shells. Of particular interest is elastin, which self-aggregates upon heating to form a fibrillar structure. Its durability and elasticity make elastin ideal for biomaterials development. A pathogenic type of protein aggregation is the formation of amyloid fibrils, which, like elastin, can be induced by increasing temperature. Elastin is essential in extensible tissues, including lungs, arteries and skin, whereas amyloid fibrils are associated with tissue-degenerative diseases, such as Alzheimer's. Although both elastin-like and amyloid-like materials result from protein self-organization, the molecular basis of their differing physical properties is poorly understood. There is currently little information concerning elastin's molecular structure. Its insolubility and flexibility have precluded the use of conventional structural determination methods, including crystallography and NMR. Using molecular dynamics (MD) simulations, which are not hindered by conformational disorder, we have demonstrated that elastin-like and amyloid-like peptides are separable on the basis of backbone hydration and peptide-peptide hydrogen bonding. In addition, we have discovered a remarkable correspondence between the structural tendencies of monomers and aggregated states. An analysis of diverse elastomeric and amyloidogenic sequences revealed a threshold in proline and glycine composition above which amyloid formation is impeded and elastomeric properties become apparent (Rauscher et al, 2006).

Our proposed study will elucidate the molecular determinants of elasticity in biological tissue. To this end, we will simulate one elastin-like peptide, $(GVPGV)_7$, and one amyloidogenic peptide, $(GA)_{18}$, using distributed replica sampling (DR), a method developed in our lab. By allowing exchanges between replicas of differing end-to-end distance, we will effectively "stretch" the peptides at equilibrium, and compute the potential of mean force of stretching. Importantly, we will be able to directly compare our computed elastic modulus to experimental measurements obtained using optical tweezers to stretch single molecules (in collaboration with Dr. Nancy Forde at Simon Fraser University), as well as measurements of materials composed of elastin-like peptides (in collaboration with Dr. Fred Keeley at the Hospital for Sick Children). We will also perform simulations of both $(GVPGV)_7$ and $(GA)_{18}$ at multiple temperatures using the DR technique. The purpose of this is two-fold. First, we will investigate the effect of temperature on the structure of both of these peptides, since increasing temperature is known experimentally to initiate both elastin-like and amyloid-

like peptide aggregation. Second, performing simulations with exchanges occurring between temperatures is a way to cross over barriers in the energy landscape of the peptides and sample all relevant conformations.

In summary, the goal of the proposed simulations is to obtain a detailed structural and thermodynamic description of the temperature-induced self-assembly and elasticity of elastin. Through experimental collaborations, the results of the simulations can be validated, and furthermore, provide insight on a molecular level for the observations of elastin's macroscopic behaviour. Ultimately, this understanding will advance the rational design of self-assembling biomaterials, such as artificial skin for burn victims and vascular grafts for heart patients.

PART Vb: Research Methodology

Both the stretching and temperature DR simulations we propose are unprecedented, and will be impossible without smart computational methods in combination with access to high performance computing. The simulations will be performed by imposing constraints on the ends of the chain with exchanges between constraints occurring to enhance sampling. We will probe a 2-to-3-fold extension of the chain, which is comparable to extensions measured for several natural elastomers, including elastin and spider silk. Repeating the calculation for the elastin-like sequence ($GVPGV_7$) and the non-elastic, amyloid-like sequence (GA_{18}) will provide insight into the molecular origin of elastomeric properties. Specifically, we will be calculating a potential-of-mean-force, which is directly related to force-extension curves obtained under conditions of reversibility. We have hypothesized that self-assembled elastomeric chains are partially hydrated and that this property is essential for extension and elastic recoil. In this model, water acts as molecular 'lubricant' favoring the interconversion between extended and collapsed conformations. It is generally accepted that rubber-like elastic recoil is driven by the conformational entropy of the polypeptide chains. However, there is at present very little insight into how this global effect relates to the fine balance of microscopic properties resulting in the distribution of conformations of the polypeptide chain, hydration, and extension. Such detailed understanding is required to explain why the elastic modulus of different elastomeric proteins, such as spider silks, can vary by orders of magnitude.

Statistical error in MD simulations commonly arises from insufficient sampling, and our proposed studies directly address this issue. First, we will take advantage of distributed replica sampling, a highly efficient simulation technique recently developed in our laboratory to extend the sampling capability of MD simulation techniques (Rodinger et al., 2006). Specifically, we will allow exchanges between temperatures in one of our simulations, and exchanges between chain extension in the other case. In both of these simulations, the sampling of the conformational landscape of the peptides will be enhanced by overcoming energetic barriers. This is analogous to the replica exchange method, but much more efficient, because replicas do not need to wait for another replica to finish before an exchange can occur. Second, the fact that the structure of elastin-like and amyloid-like peptides is disordered is a tremendous computational advantage because it indicates that the underlying energy landscape is defined by conformations that are very similar in energy, and these representative conformations exchange rapidly with one another. As a consequence, it will be possible to achieve statistical convergence, and our results will provide meaningful insight into the physical and structural basis of extension/recoil equilibria of elastomeric peptides. Importantly, our simulations will push the boundaries of modern MD simulations, with more than 12 microseconds of sampling for each peptide (from initial tests, we are able to simulate approximately one nanosecond per CPU day). We will be able to achieve complete statistical sampling of peptides on the nanoscale using our DR method, which will establish a framework for future studies of aggregates of these peptides, as well as other complex systems.

PART VI: Allocation Justification

The main reason for our request of dedicated resources is that our methodology requires many serial CPUs to be operating simultaneously. The distributed replica program monitors the temperature (or end-to-end distance) of each replica and uses this information to dictate which temperatures (or end-to-end distances) should be sampled next. By waiting in queue, it is infrequent that all of the required processors are simultaneously available. Without dedicated resources, it would require >2.5 years for these simulations, and there would not be appropriate coupling between replicas. We have been developing the methodology described using simpler test systems (octameric peptides) with SHARCNET for the past several months, and have demonstrated its utility (the manuscript is currently in preparation). However, in order to make it feasible to simulate the systems proposed here, we will require a minimum of 70 CPUs to be running concurrently. In order to simulate both systems at the same time, we are requesting the use of 140 CPUs for 6 months (a total of 604800 CPU hours). This is the projected amount of time it will take to achieve statistical

convergence based on preliminary simulations. Because we are running serial jobs, we attain 100% CPU efficiency, in addition to the enhanced sampling efficiency provided by DR compared to conventional MD. Analysis of the MD trajectories will be conducted using programs developed within our research group. In order to perform this analysis directly on SHARCNET as the simulations progress, we request the use of 200GB of storage for 6 months.

PART VII: Key Papers

1. Rauscher, S., Baud, S., Miao, M., Keeley, F. W., and Pomès, R. (2006) "Proline and Glycine Control Protein Self-Organization into Elastomeric or Amyloid Fibrils" *Structure*. **14**, 1667-1676.

2. Rodinger, T., Howell, P. L., and Pomès, R. (2006) "Distributed Replica Sampling" *Journal of Chemical Theory and Computation*. **2**, 725-731.

3. Rodinger, T., and Pomès, R. (2005) "Enhancing the accuracy, the efficiency and the scope of free energy simulations" *Current Opinion in Structural Biology*. **15**, 164-170.

PART VIII: Outcomes from Previous SHARCNET Awards

Not applicable

PART IX: Suggested Reviewers

1. Dr. Peter Tieleman, University of Calgary, tieleman@ucalgary.ca, (403) 220-2966
2. Dr. Guillaume Lamoureux, Concordia University, glamoure@alcor.concordia.ca, (514) 848-2424 x 5314
3. Dr. Patrick Lagüe, Université Laval, Patrick.Lague@rsvs.ulaval.ca, (418) 656-2311