The study of membraneless organelles, using string of beads integrative modeling

Ilia Bezgin, Rina Karnauch, Ofek Kaveh, Roy Maman

Abstract:

Many cellular compartments are not bound by membranes. These membraneless organelles, often called biomolecular condensates, form intracellular mechanisms of great physiological importance. The study of this field attracts many biologists and other researchers, but the complexity of biomolecular condensates makes them a difficult target to study. In this paper, we present a powerful computational tool - strings of beads dynamics, and demonstrate the possibility of using this tool to study the formation and dynamics of biomolecular condensates.

Introduction:

Biomolecular condensates are specialized subunits of molecules that form together and function as independent organelles. Unlike classical organelles, biomolecular condensates are membraneless organelles, meaning that they are not bounded together by a membrane. membraneless organelles are formed together through many different mechanisms. One of them, especially dominant in proteins and RNA condensate, is phase separation.

Phase separation is a process by which a well-mixed solution of macromolecules such as proteins or nucleic acids spontaneously separates into two phases: a dense and a dilute phase. The dense phase has liquid-like properties and enriches certain macromolecules while others are depleted, allowing the dense phase to function as a compartment (Alberti et al., 2019). In phase separation, thousands of identical molecules cluster and interact together, implying that small changes in molecular properties of components, by mutation for example, can propagate and dramatically impact macroscopic phenotypes of assemblies (Townsend et al., 2017).

Recent research has proved that condensates play fundamental roles in cellular organization, protein folding and physiology. As the research goes on, our

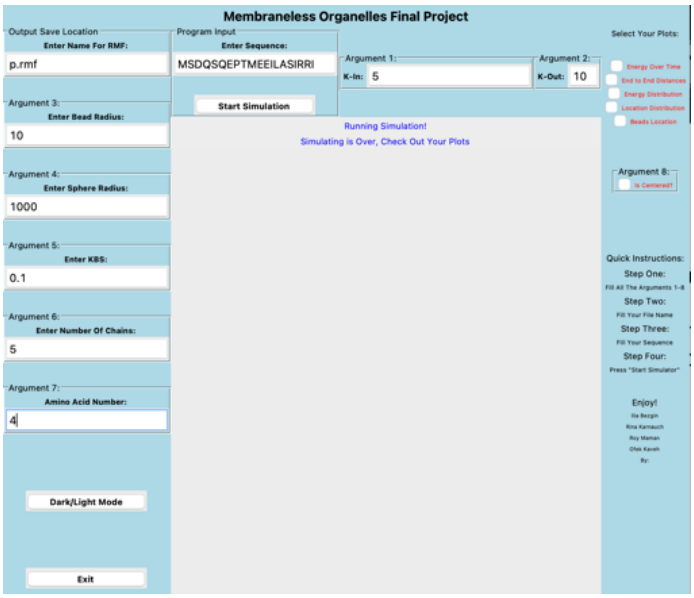
understanding of the mechanisms and regulation of condensate assembly improves, as well as their role in ageing and disease (Alberti et al., 2021).

The dynamics and self-organization of these systems involve elaborate spatio-temporal coordination of their elements, and thus are very hard to model.

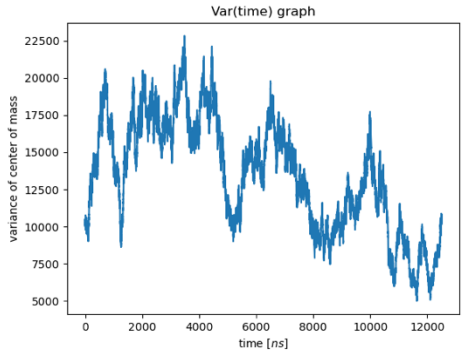
A convenient method for modeling complex dynamics is the strings-of-beads model. This model allows reducing the system complexity, by replacing short sequences of adjacent amino acids, with a representative sphere - a “bead”. A protein sequence is represented by a string of beads, of a chosen radius, with harmonic bonds, with a chosen effective spring constant ‘k’, between them. Different strings are interacting through specific predefined beads. The whole model then evolves with time, according to brownian motion formalism, and the dynamics of complex condensates can be studied. Our project implements a string of beads model, with integrated GUI, and offers a straightforward, easy to use, tool for modeling real membraneless organelles systems.

Results:

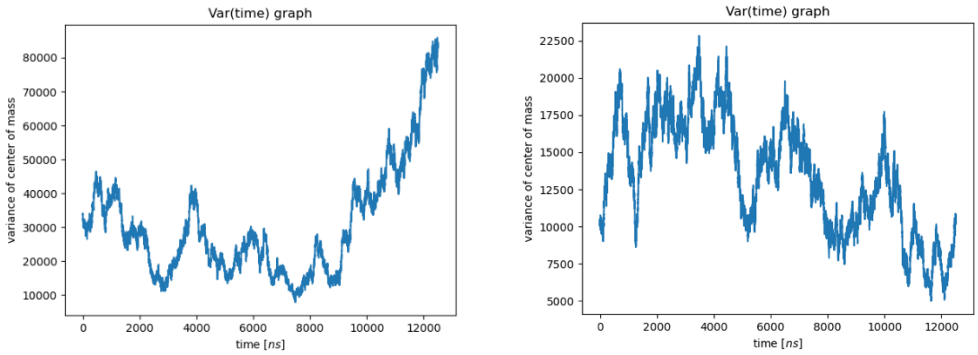
Using our tool, we ran about 50,000 frames (~10μs) with the following parameters:



For every run, we exported the following graph:



The graph shows the variance of strings locations (the location of each string was the center of its beads' mass) as function of time. We used this value as a measure for the proximity of the different strings in the simulation.
In the first step, we examined the behavior of this graph, for different amount of chains in the simulation:

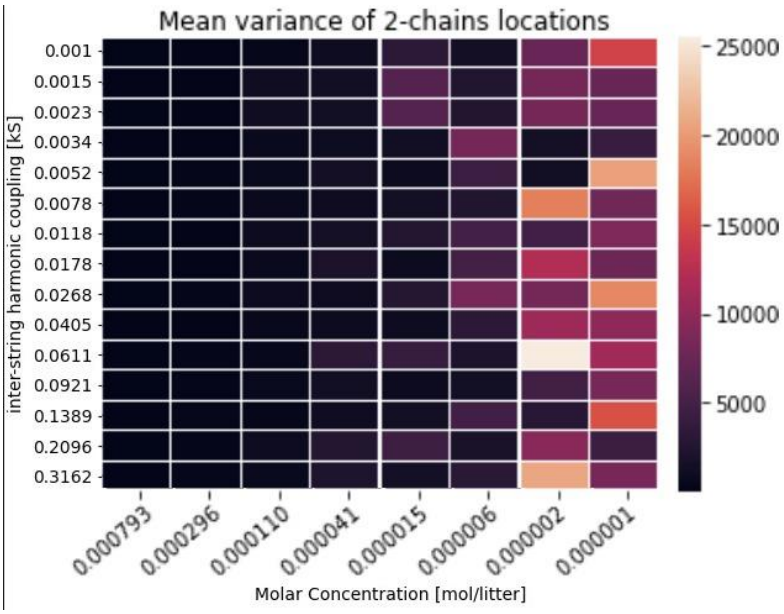


Variance of chains' center of mass locations, as function of time.
left – simulation of 2 chains. Right – simulation of 5 chains.

In a simulation of 2 chains, when the chains had been initialized relatively far from each other, we see that at some point, the 2 chains started to move on different directions. In contrast, when we simulate 5 chains (with the same conditions), we notice the opposite. The different chains got closer to each other. This result is not surprising. In our model, the chains evolve stochastically on a Brownian motion. When 2 chains are close enough to each other, the direction of their motion is biased towards an intercourse. As we increase the number of chains, we increase the chances of 2 chains to be close enough to each other, and from this point, the inter-chain coupling forces the distance between them to reduce.

To further our research, we tested the system (once with 5 chains, and once with 2 chains) with different values of K-Out (inter-string coupling coefficient) and molar concentration (in practice, we changed the bounding sphere radius). The results are 2 heatmaps (one for 2 chains, one for 5 chains), when the x and y axes are K-out and molar concentration, and the heat axis is the mean variance of chains' center of mass over time (starting from a temporal threshold). From these heatmaps, we want to learn something about the impact of molar concentration, as well as interactions strength, on

the formation of biomolecular condensates. Due to short simulation times (only ~10μs each), the information derived from the heatmaps would not be straightforward, but instead, would be correlated to the formation of condensates (we get the mean variance of distances, instead of absolute distances along some time interval).
Let's start from the easier case, a simulation of 2 chains:

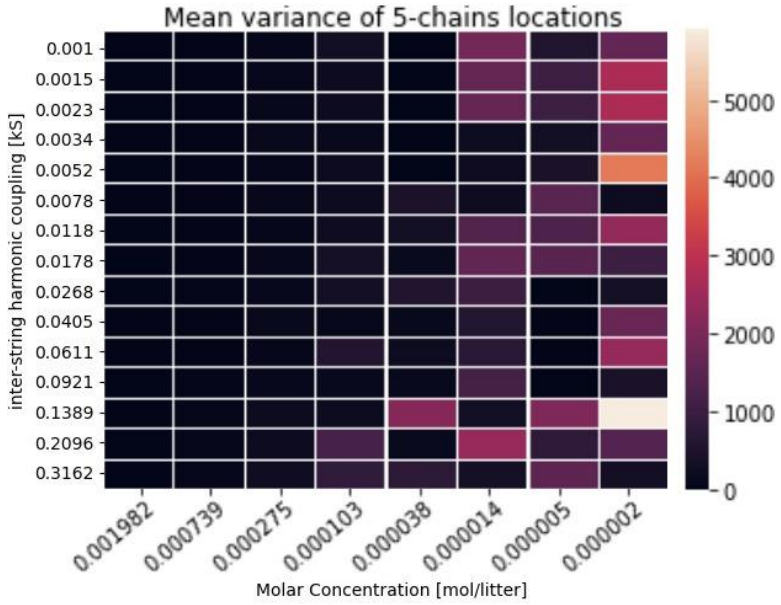


Mean variance of chains' center of mass locations, as function of the molar concentration and the inter-string harmonic coupling. Number of chains in the simulation: 2.

It seems that for our chosen working parameters, the molecular concentration is far more meaningful in determining the distance between the chains. The left columns, represents higher molecular concentrations, the distances are small. for low molecular concentrations, the distances are high, but there is no distinguishable trend along different rows, meaning that the distances along different K-outs seems to distribute randomly. Perhaps, much larger K-outs values are required to prevail the stochastic nature of the dynamics.

after we got some basic understanding, let's move to a more complex system. The same experiment is now done, this time with 5 chains in the simulation. The mean variance of distances for a system contains more than 2 chains, maybe misleading in some way (in a system of 2 chains only. The variance of distances contains information

of all the pairs in the simulation, while when we have more than one pair, we only get partial information).



Mean variance of chains' center of mass locations, as function of the molar concentration and the inter-string harmonic coupling. Number of chains in the simulation: 5.

The above graph is quite similar to the pervious one, but it contains less hot squares in the region of low concentration values, especially for higher K-values, meaning that in general, systems with more molecules are more likely to form condensates even without a significant change in concentrations. This phenomenon is easy to explain analytically - As we discussed earlier, the nature of chains' motion is in general stochastic, but it is biased towards lower energy areas. The inter-chain potential increases the probability of chains to get closer to other chains. Once 2 chains are closer than the truncation threshold (the value from which the harmonic coupling appears), the chains are more likely to go towards each other. When there are 2 chains, there is low chance for the chains to reach the threshold, but as we increase the number of chains, there are more terms in the probability space, so naturally the chance for any pair to reach this threshold gets higher.

Discussion:

in this work, we created a tool for integrative modeling based upon the strings-of-beads approach. In order to demonstrate its capabilities in real-life problems, we used our tool in order to study about the correlation between the molar density and chains proximity, for different inter-string interaction strengths, and for different amount of chains (2 chains and 5 chains). The interesting time constants for biomolecular condensates are ~μs and in order to learn new things, the simulations should be long as possible, and contain enough strings. Obviously, running the simulation for longer times and with more strings, consumes significant computational resources, therefore, our tool has its limitations and any researcher who wants to make a use of it, needs to take long running times into his considerations.

Methods:

Code. Based on the IMP library for Integrative modeling, our code takes a sequence of amino acids as an input, as well as configuration parameters, and outputs a 3D time-evolving simulation of strings of beads (in RMF format). The user is free to choose various parameters: amino acids - beads conversion ratio (how many amino acids are represented by each bead), the harmonic coupling between beads on the same string, the harmonic coupling between adjacent beads of different strings, number of string, the radius of the bounding sphere (the sphere that bounds the spatial space) and the harmonic coupling between the sphere and every bead. Our model only allows inter-string interactions between a single, constant bead on each string.