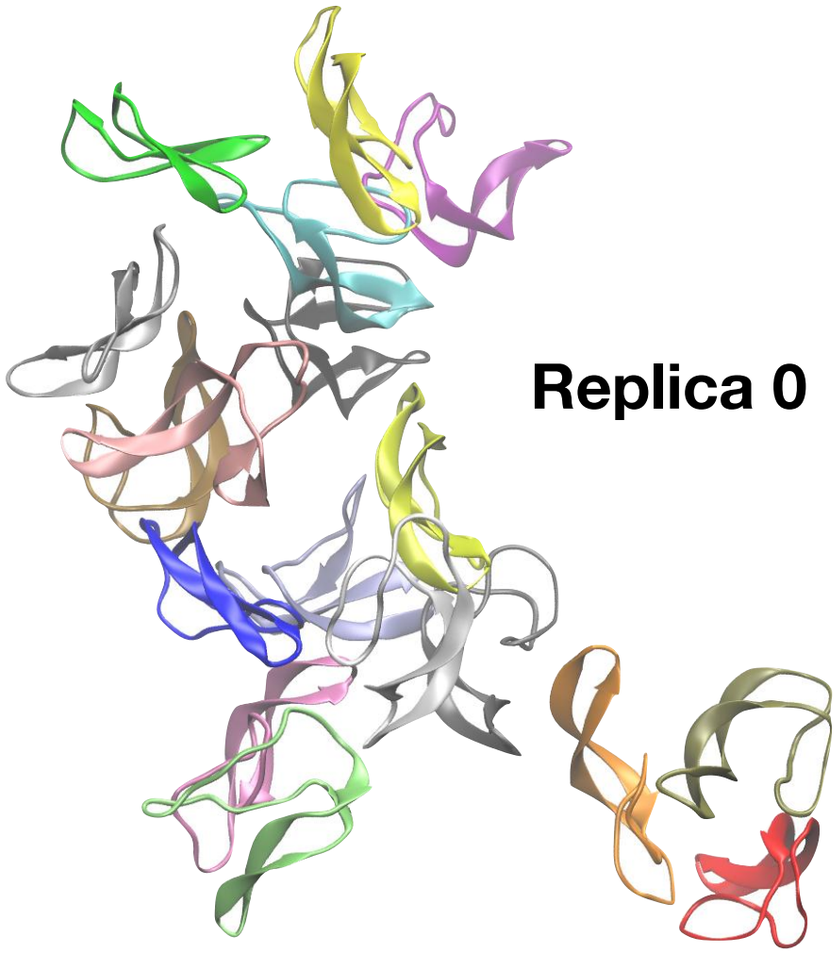
Prion Protein aggregation

Francesca Collu
Enrico Spiga
Irene Marzuoli

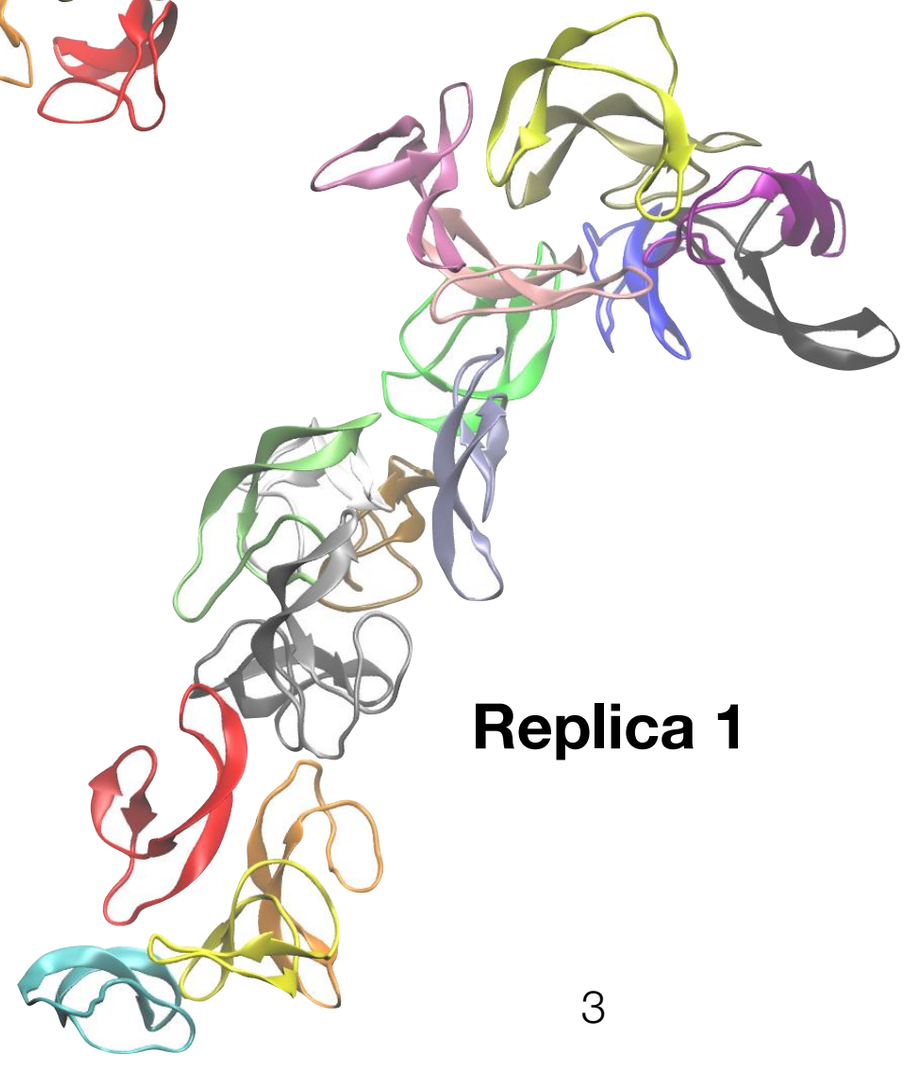
Structures of the final aggregates

Due to the nature of the box and extent of the final aggregate, in some of the replicas the final structure interacts with its periodic boundary copy. Such simulations (resulting in elongated geometries) must be discarded from the analysis as the spurious interaction is not physical.

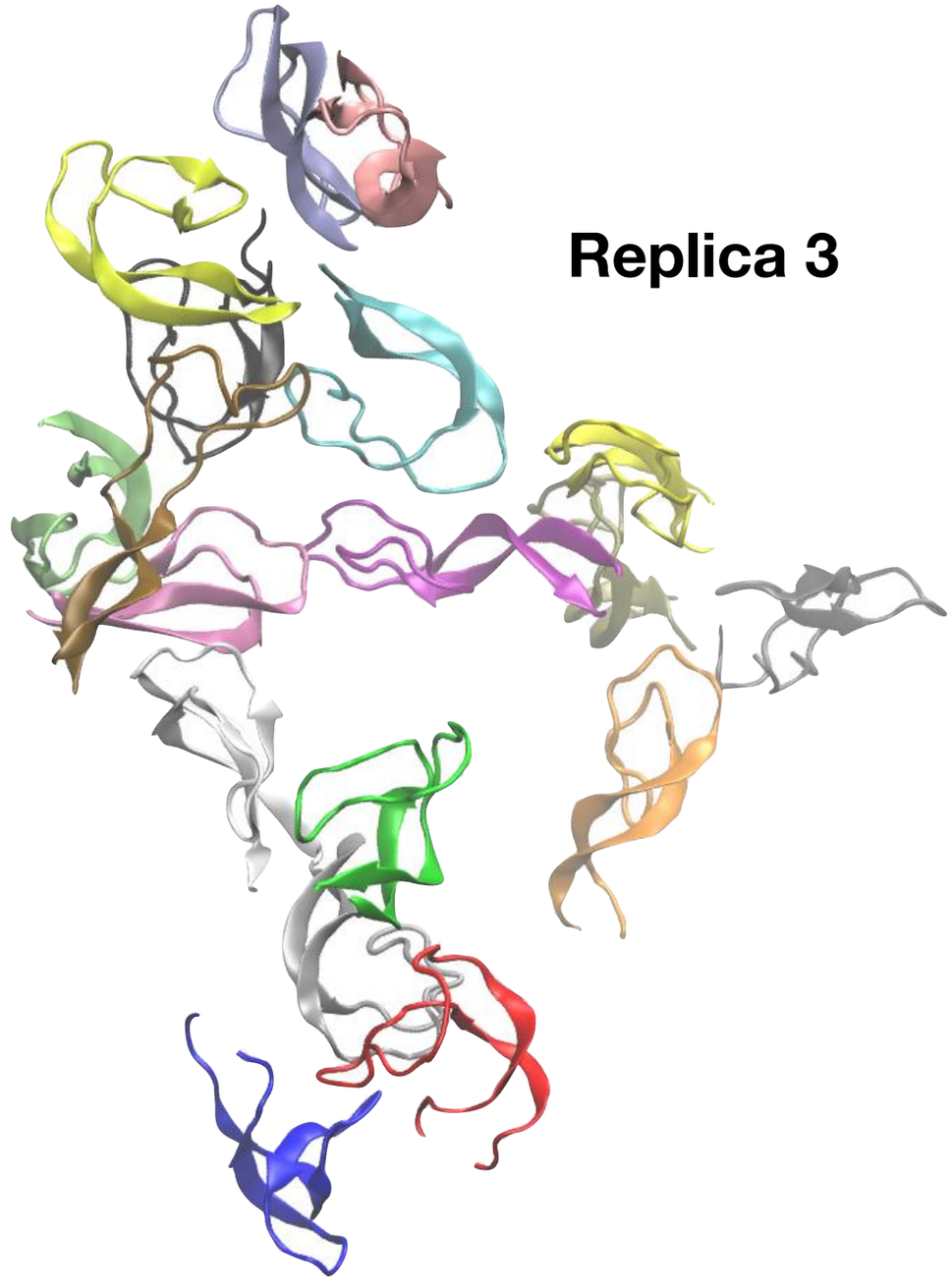
Correctly folded replicas



Replica 0

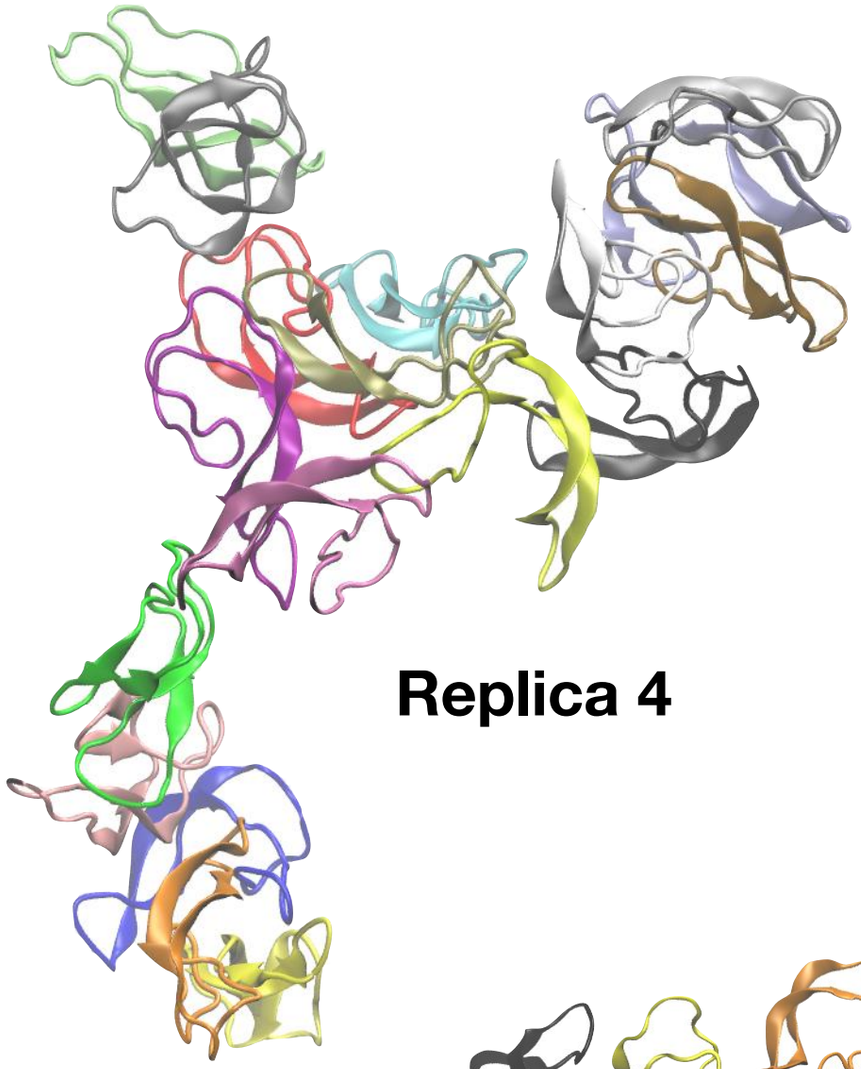


Replica 1

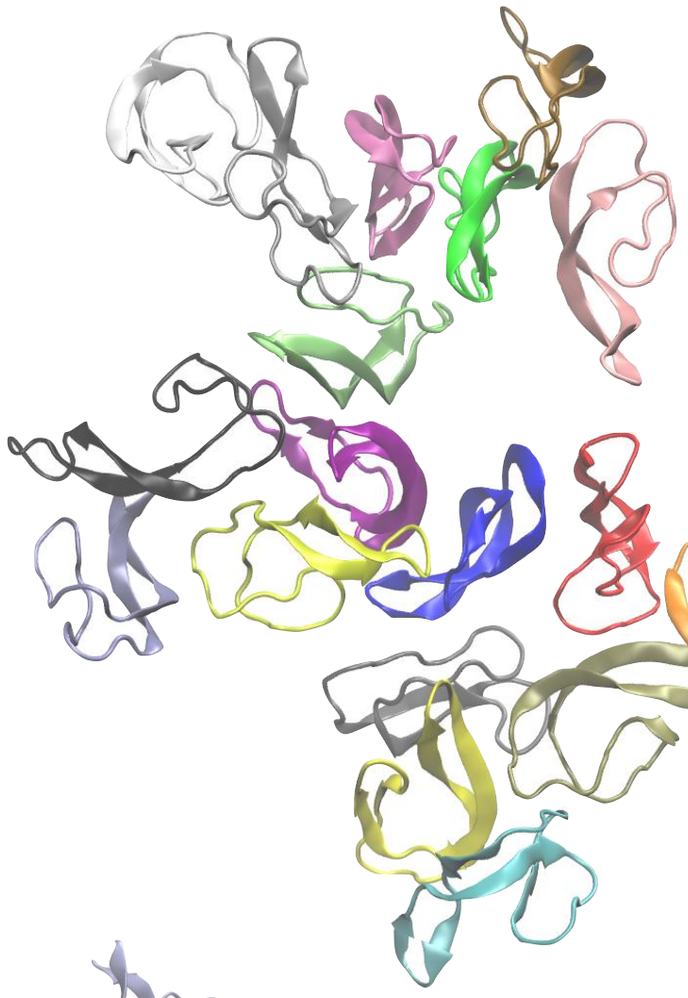


Replica 3

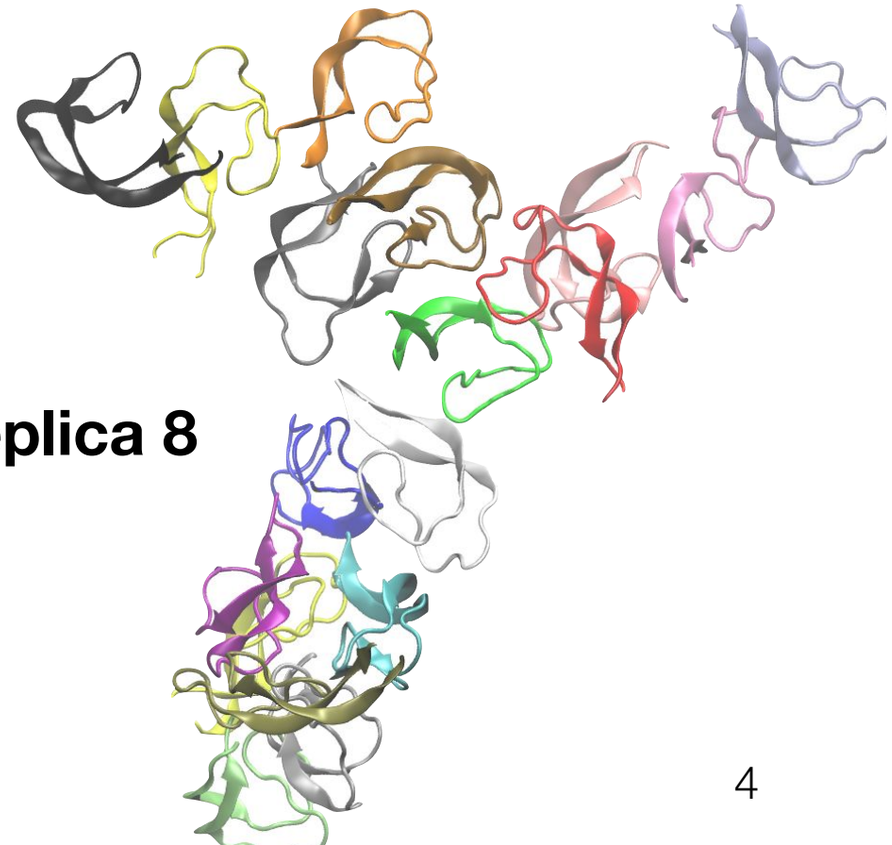
Correctly folded replicas



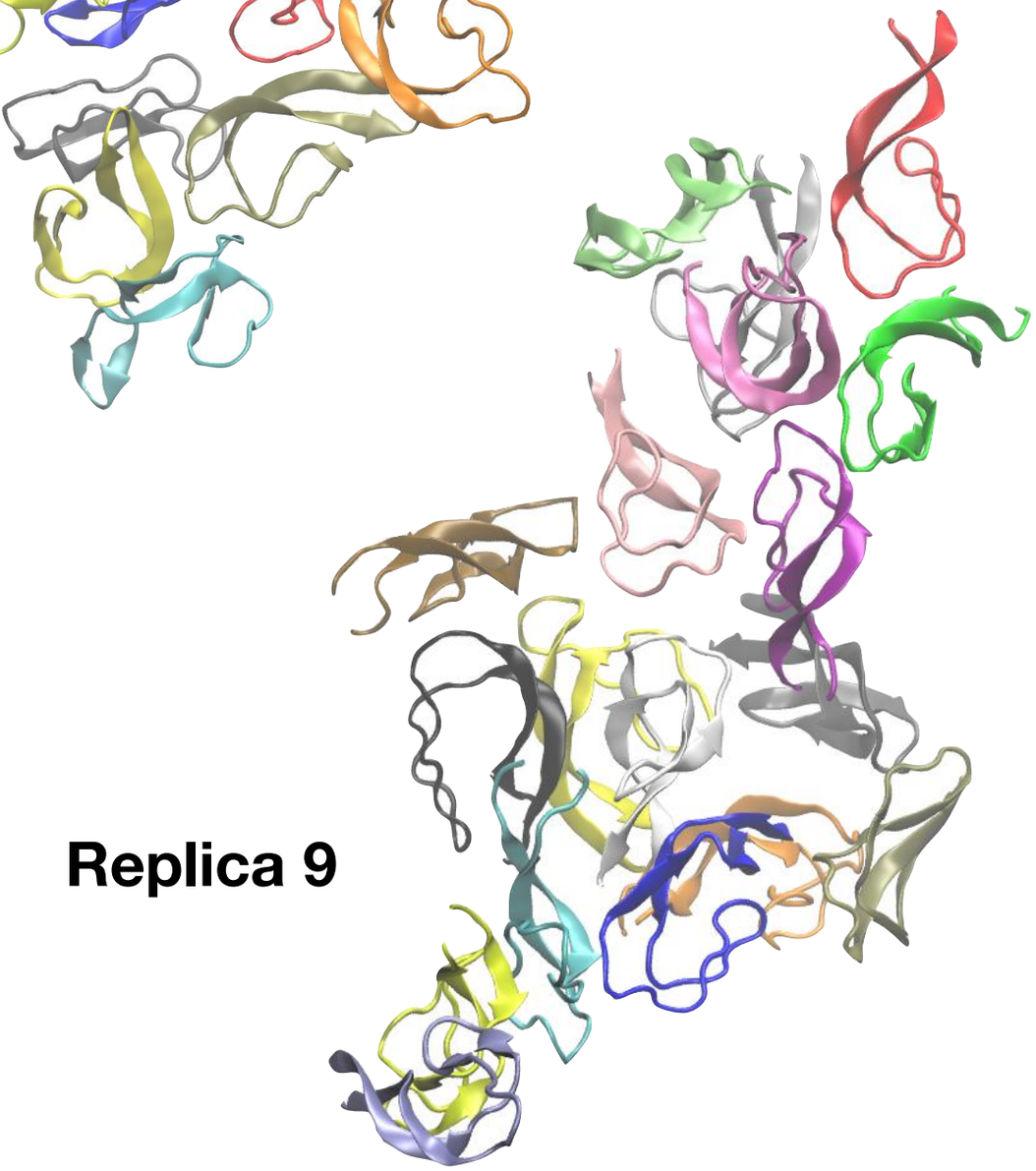
Replica 4



Replica 5

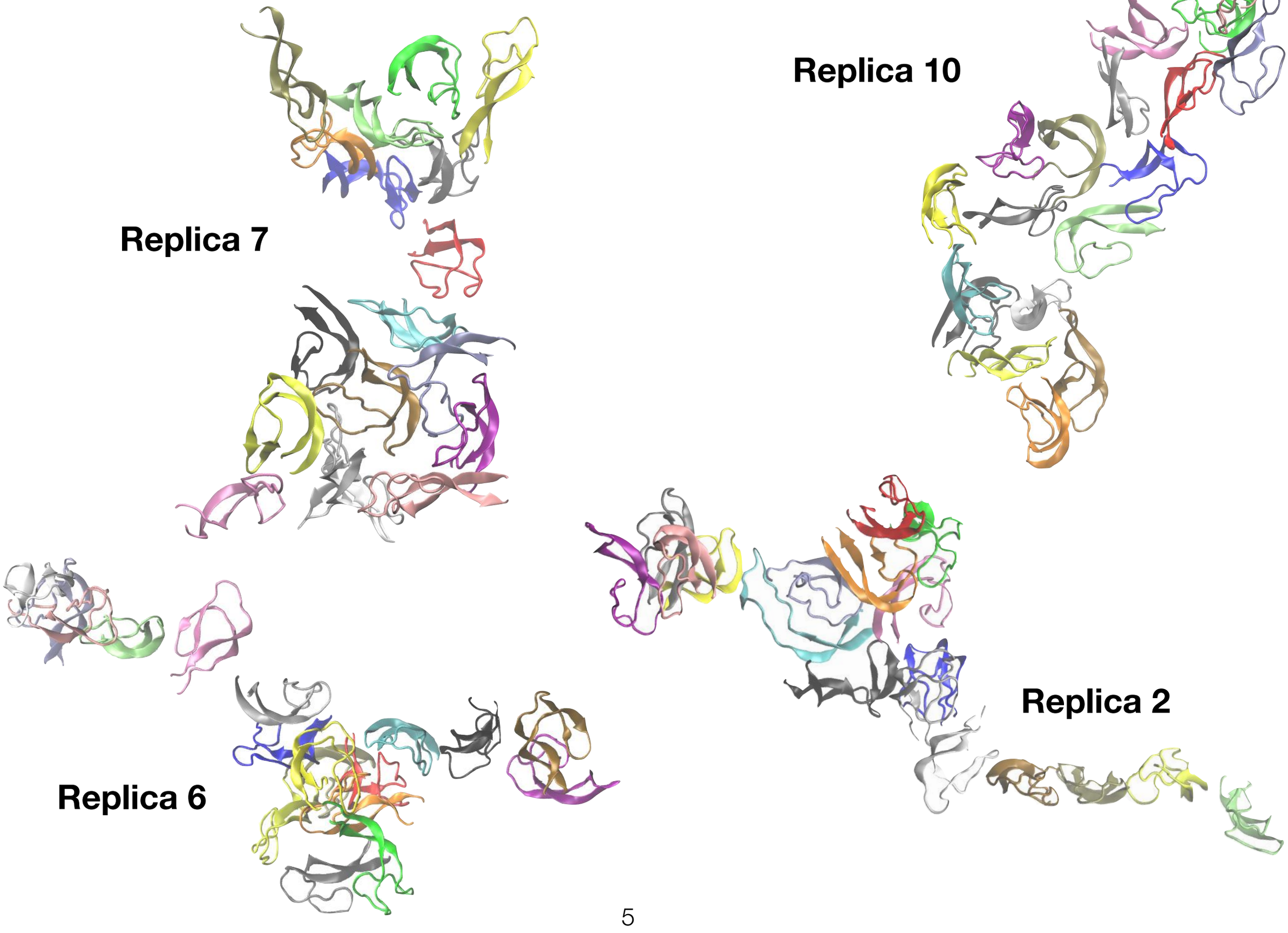


Replica 8



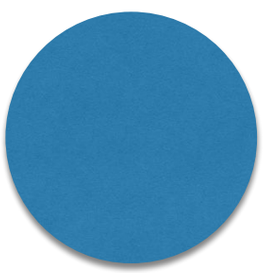
Replica 9

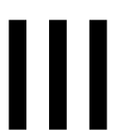
Elongated replicas (periodic interaction)

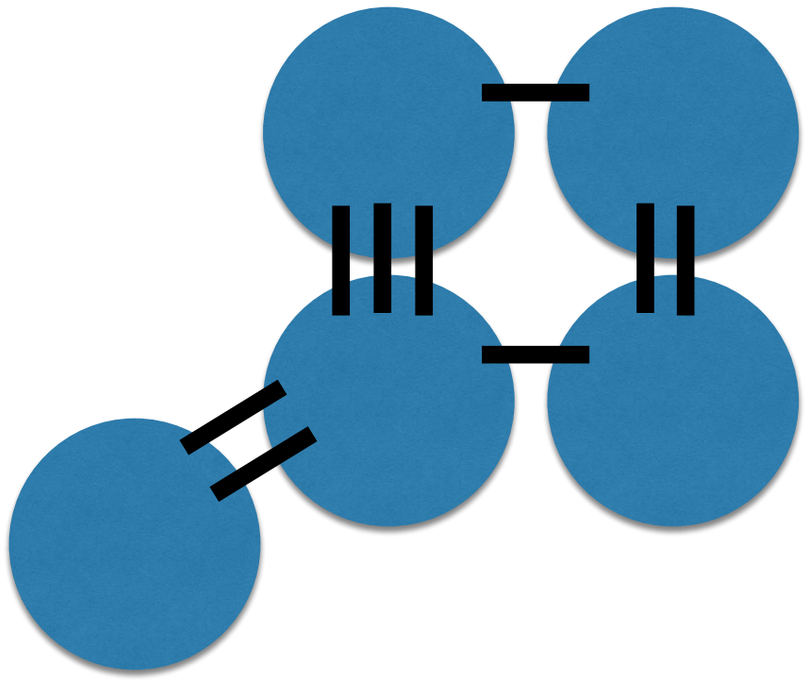


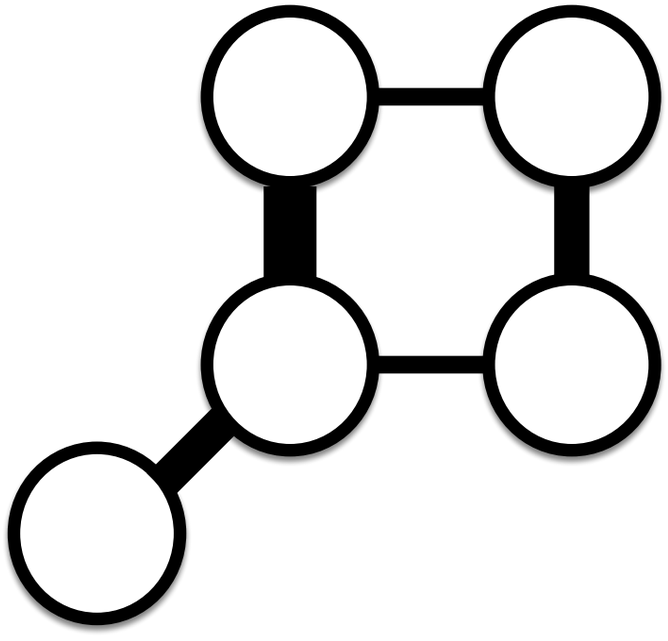
Processing the aggregation state as an undirected graph

Aggregation state ⟷ G(V,E)

Prion protein  ⟷ V=vertices

Interfaces (POPS)  ⟷ E=edges

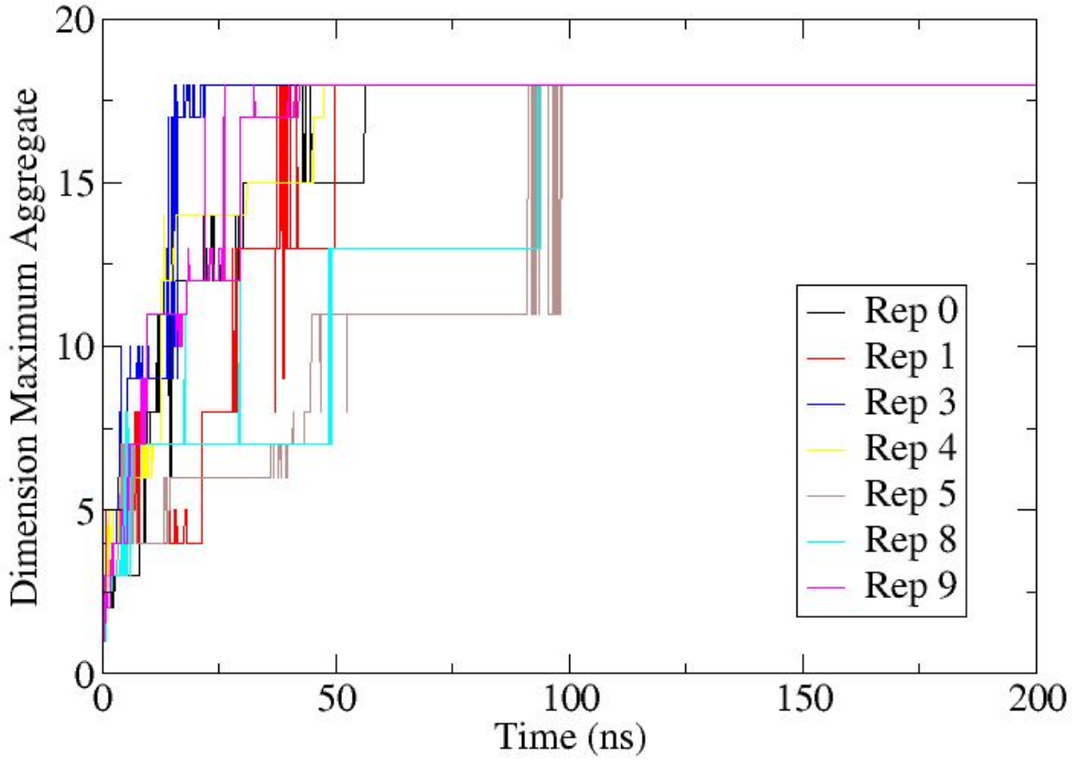


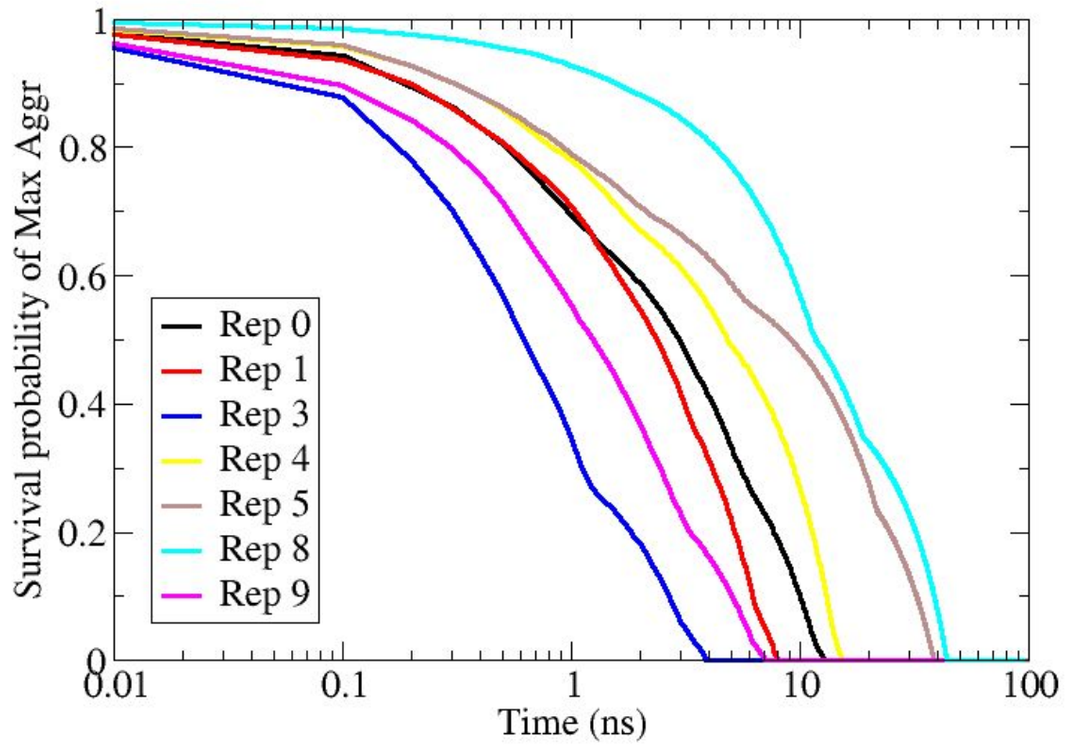


Survival Analysis of the Microstates

recomputed using more frames/ns and investigating the survival of single oligomers

Evolution in time and survival of maximum aggregate





Dimension of maximum aggregate in time (left) and its survival probability (right).

$$S_{TOT}(t) = \sum_{s=1}^{T-t} \frac{1}{T-t} \sum_{j} P_j(t,s)$$

$P_j(t,s)$ is equal to 1 if state j (in which the maximum aggregate has a size of j) survives for a time t from instant s. T (total simulation time), t and s are multiple of a discrete time interval Dt; j goes up to 17.

Recomputed on more frames/ns.

Survival of each aggregate as maximum aggregate

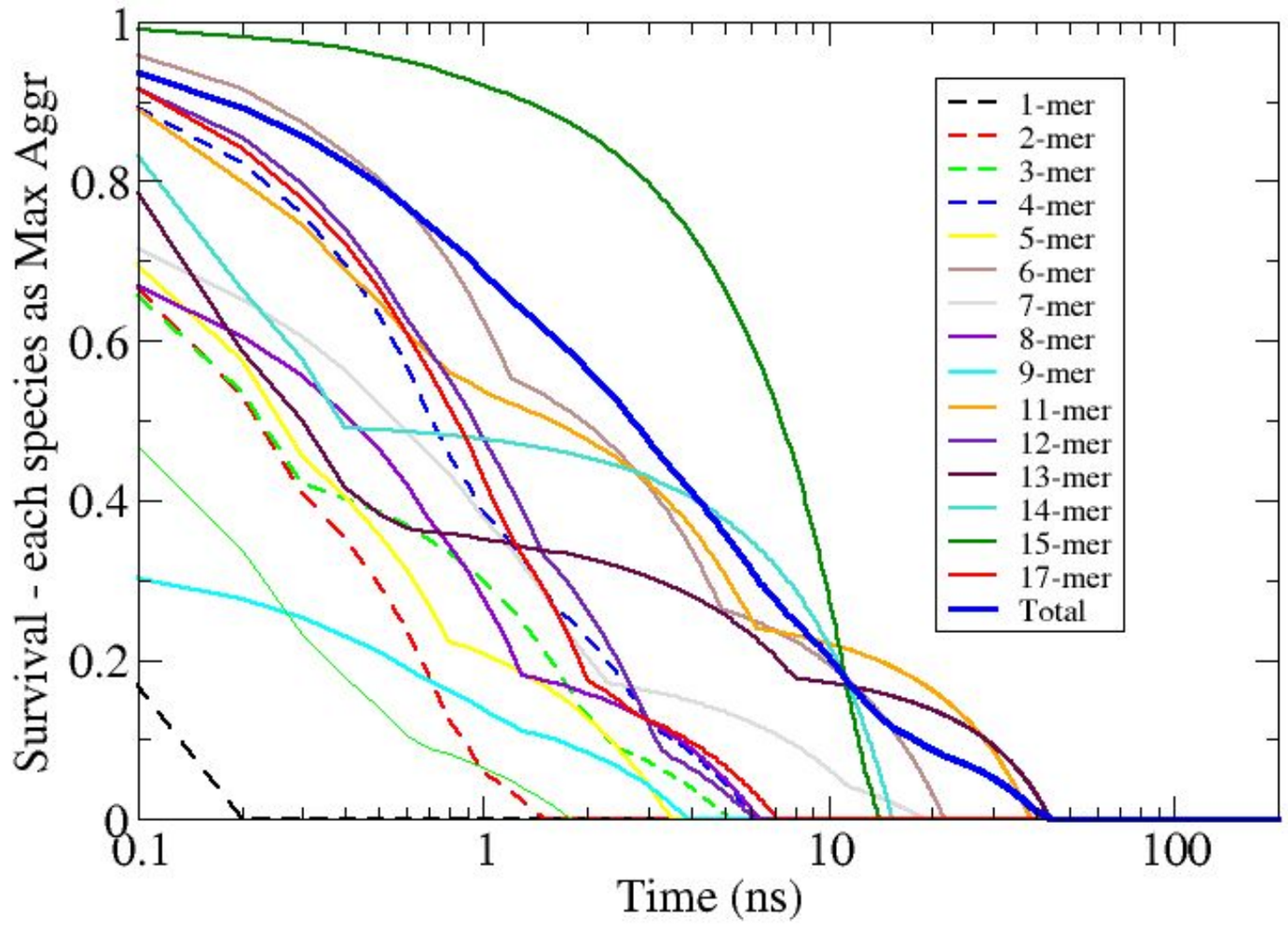
How long does each size maintain the role of maximum aggregate?

Survival probability for every size, averaged over replicas:

$$\langle S_j(t) \rangle = \sum_{replicas} \sum_{s=1}^{T-t} \frac{1}{T-t} P_j^{rep}(t,s)$$

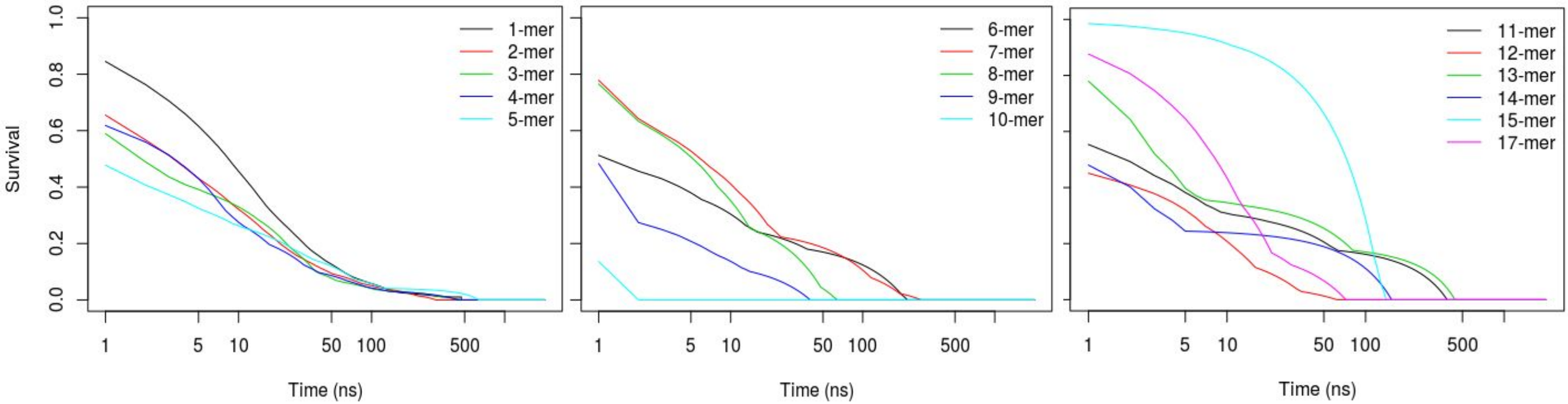
Plot of $<S_j(t)>$ for j in [1, 17]
Total corresponds to $<S_{TOT}>$.

Monomer and dimer have a small survival time: aggregates larger than size 2 are quickly formed.



Evolution in time of each cluster

Survival probability for each aggregate (identified by the ID of the prions in it), averaged over aggregate of the same size.



	Decay	**Explanation**
Small aggregates (< 6)	Exponential	Averaged on multiple copies
Large aggregates (> 10)	Linear tail	Single occurrence (not enough copies to have more than one at once)
Medium aggregates (6 - 10)	Linear	Multiple copies can occur but the decay is similar to the single copy one

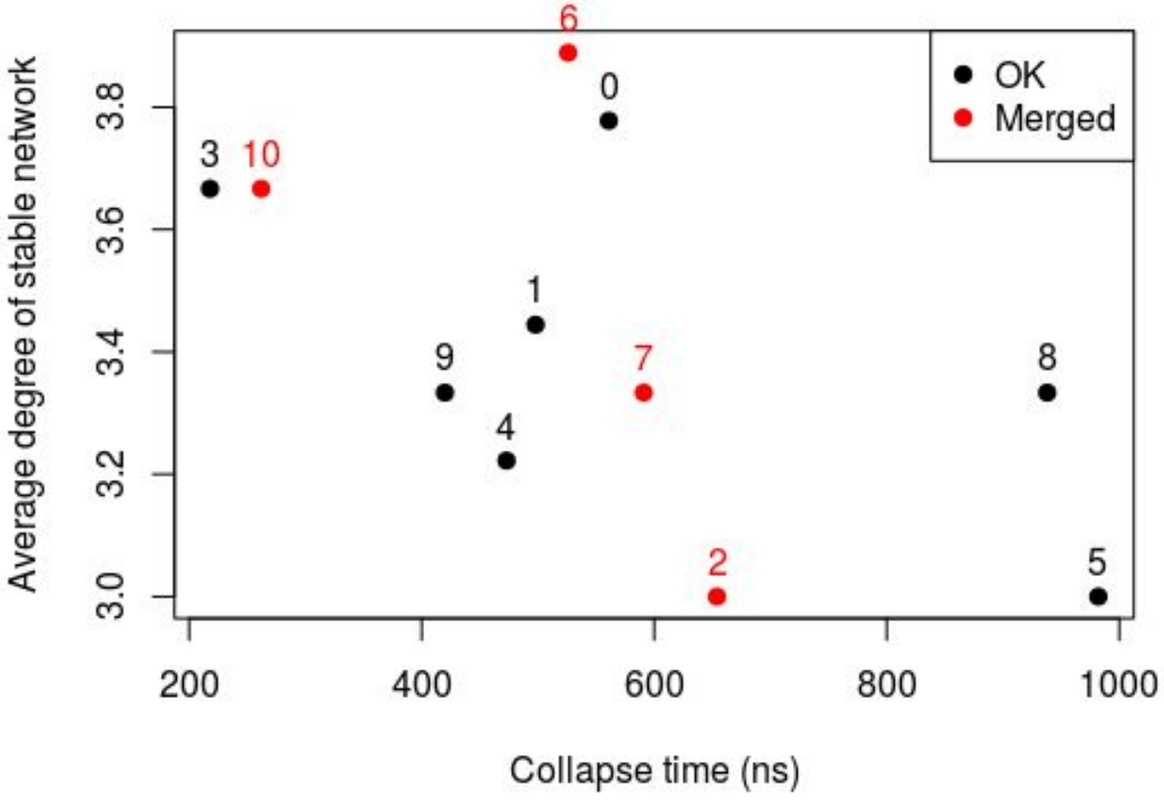
Network properties of the final aggregate

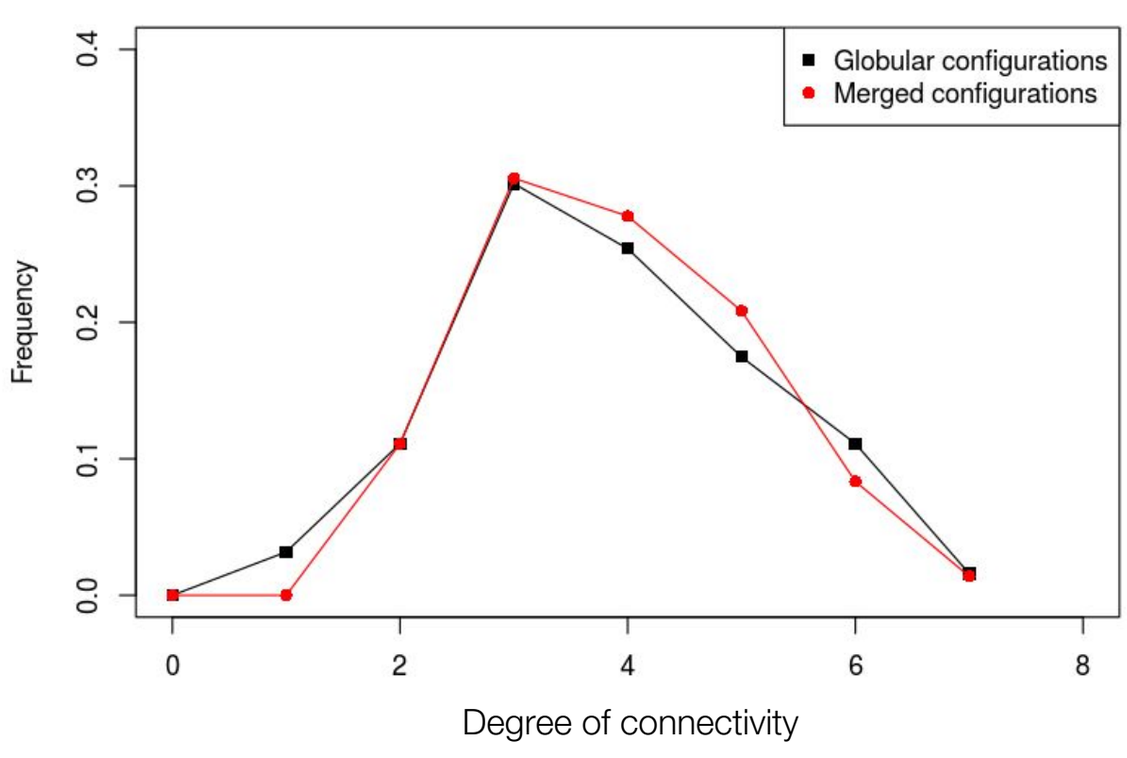
Investigating the connectivity and formation of the contacts between prions looking at the network properties of the structure

Connectivity in the structure of the final aggregate

The average connectivity in the final structure is independent from the aggregation time (left) and stays between 3 and 4 (even for the elongated replicas, in red).

The distribution of the connectivity across the nodes (i.e. each prion), averaged over the replicas, shows values up to 7 connections/prion. This is likely due to prions located at the center of the structure.





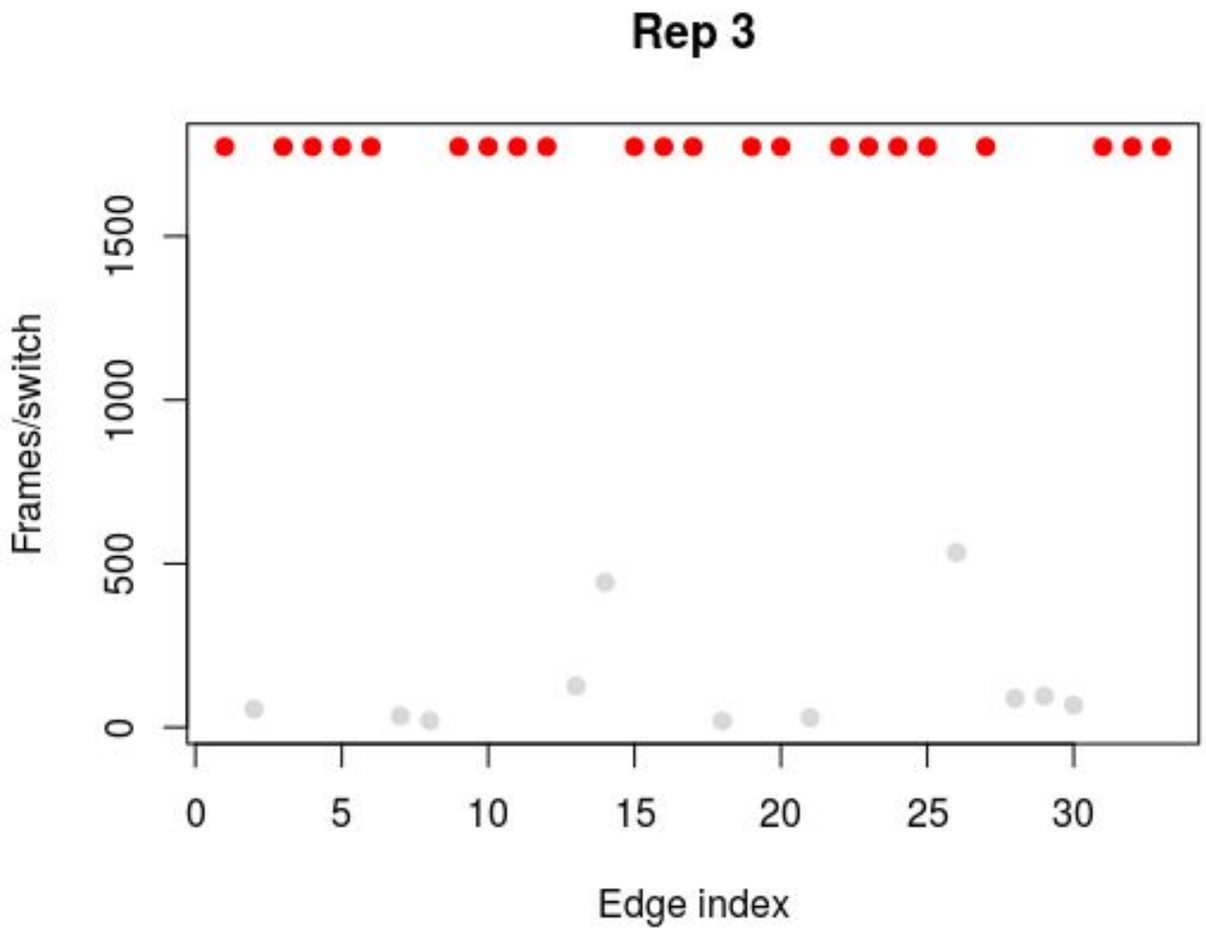
Analysis of the persistence of the bonds

After the aggregation time, we can identify two types of contacts between neighbour prions:

1. one class of contacts continuously present for each time (strong)
2. a second class of contacts forming and breaking (weak)

(Left) Persistence score for each bond in the final aggregate of Replica 3.

Persistence score:
Nr. of frames it is present, normalised over the number of times it forms and break.
Computed on time frames after aggregation time.



Are “weak” bonds forming and breaking or present only a few frames? Compare different time windows.

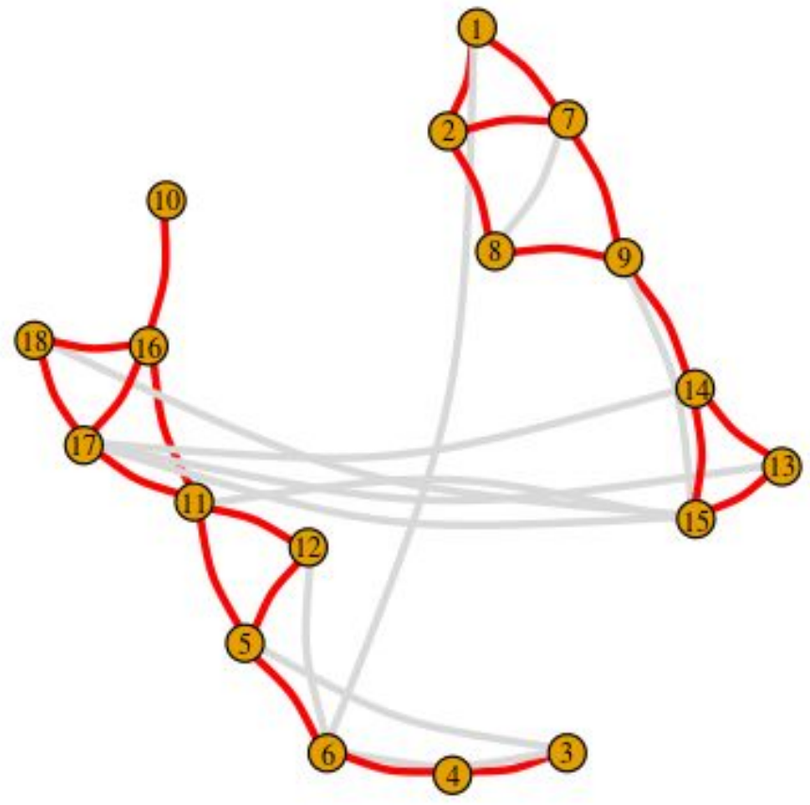
Analysis of the persistence of the bonds

Are “weak” bonds forming and breaking or present only a few frames?
Time window: [aggregation time; 200 ns]

Nr frames a bond is present/nr. of time it forms and break, analysing time frames from aggregation time till the end.

Color code:
red > 1500;
grey < 800;
Max score = 1780 (a bond always present from aggregation time till end)

Replica 3 - fast

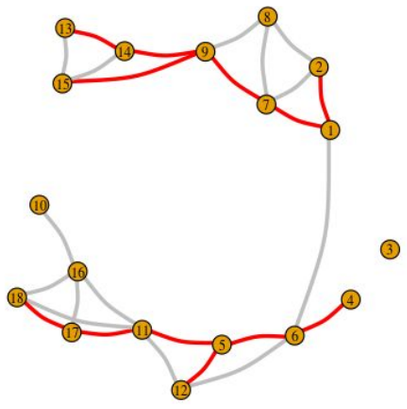


Analysis of the persistence of the bonds

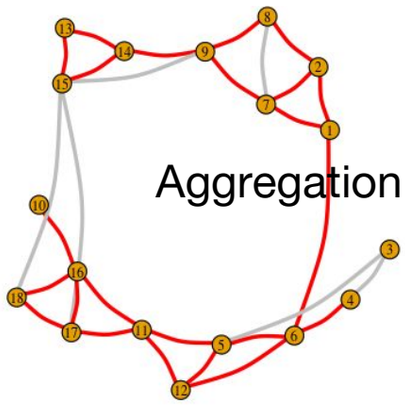
Are “weak” bonds forming and breaking or present only a few frames?
Time windows: 20 ns

Nr frames a bond is present/nr. of time it forms and break, analysing 20 ns time windows.
Color code: red > 150; grey < 100; Max score = 200

0-20 ns

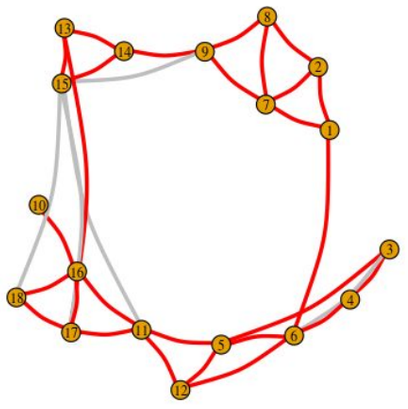


20-40 ns



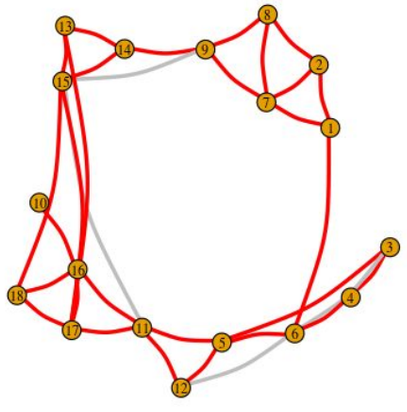
Aggregation (22 ns)

40-60 ns

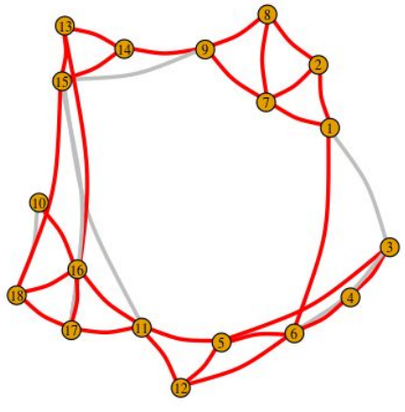


Replica 3 - fast

60-80 ns



80-100 ns



The bonds classified as non persistence in the analysis after the collapse time appear also over short, subsequent time windows so they are indeed always present but in an unstable way.

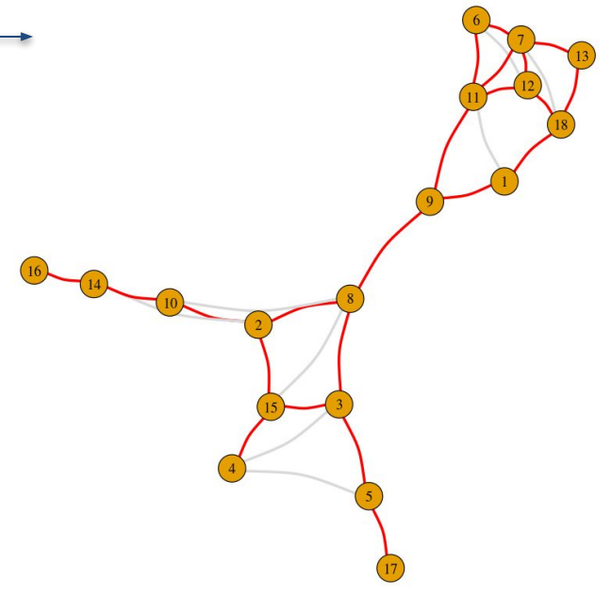
Analysis of the persistence of the bonds

Similar classification occurs for both fast aggregation time (Replica 3, previous slide) and slow ones (Replica 8, present slide)

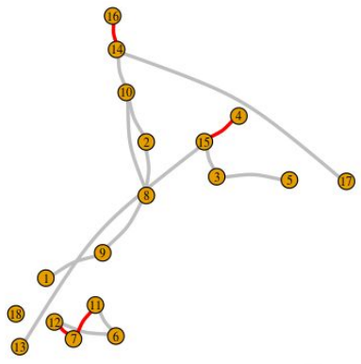
Analysis on timewindow [aggregation time; 200 ns]

Analysis on 20 ns time windows

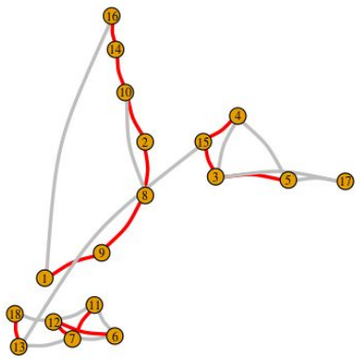
Replica 8 - slow



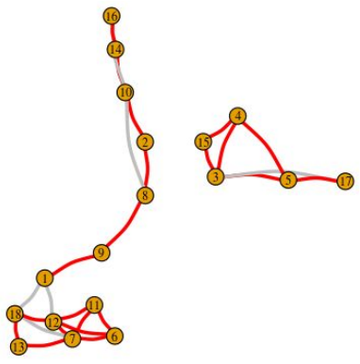
0-20 ns



20-40 ns

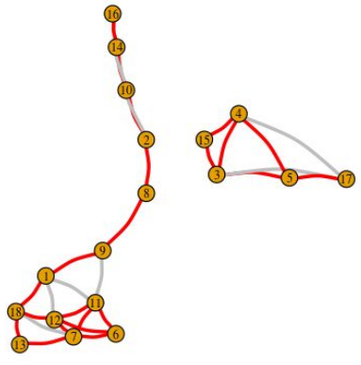


40-60 ns

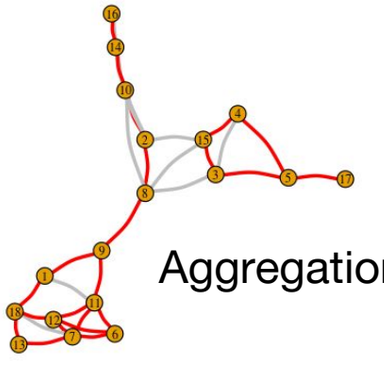


Replica 8 - slow

60-80 ns

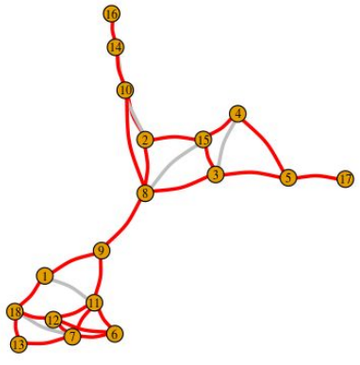


80-100 ns



Aggregation (94 ns)

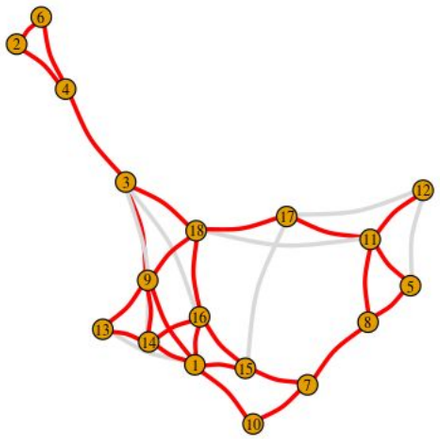
100-120 ns



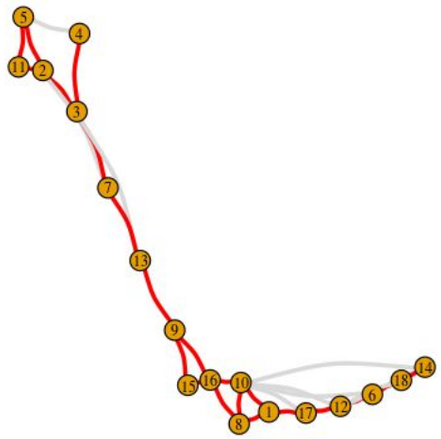
Analysis of the persistence of the bonds

Persistence of the bonds for each replica. Analysis over [aggregation time; 200 ns]

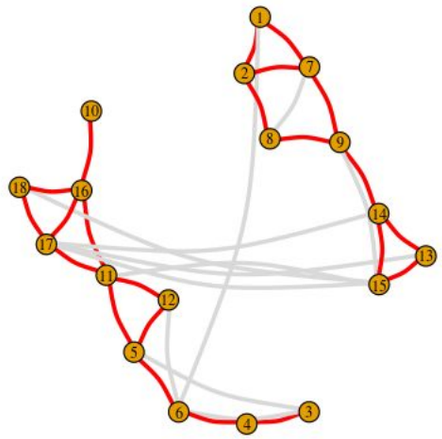
Replica 0 - medium



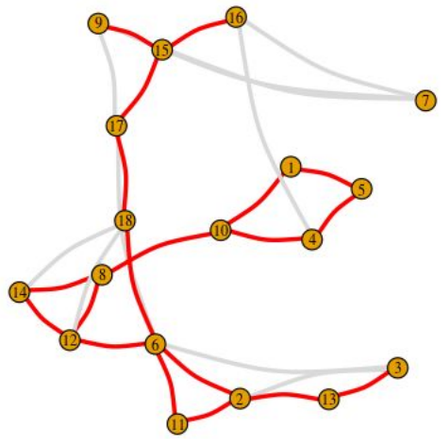
Replica 1 - medium



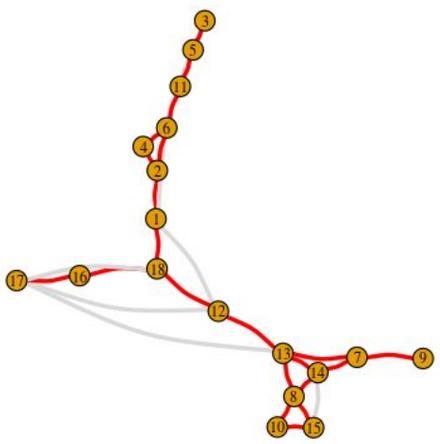
Replica 3 - fast



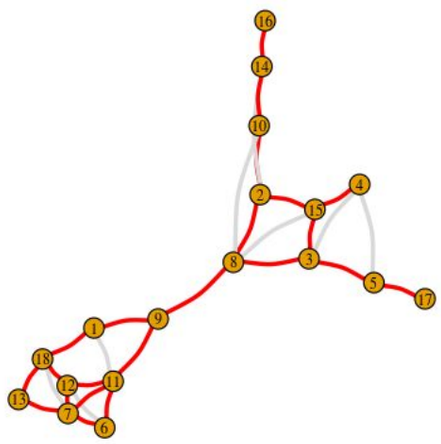
Replica 4 - medium



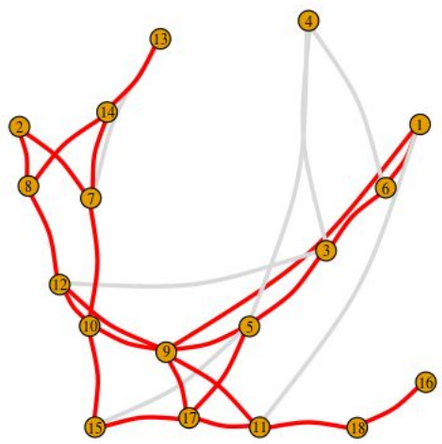
Replica 5 - slow



Replica 8 - slow



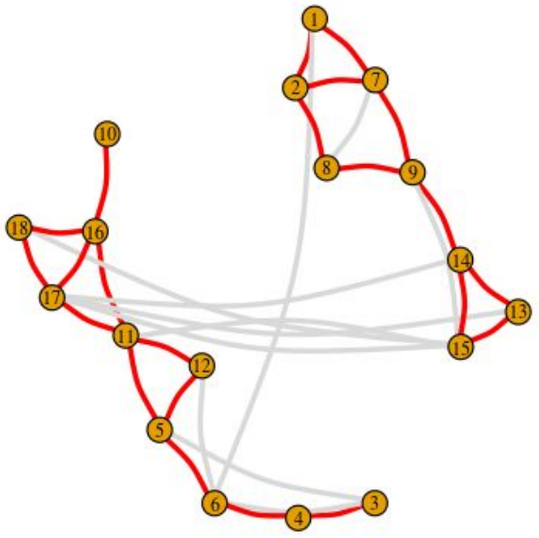
Replica 9 - medium



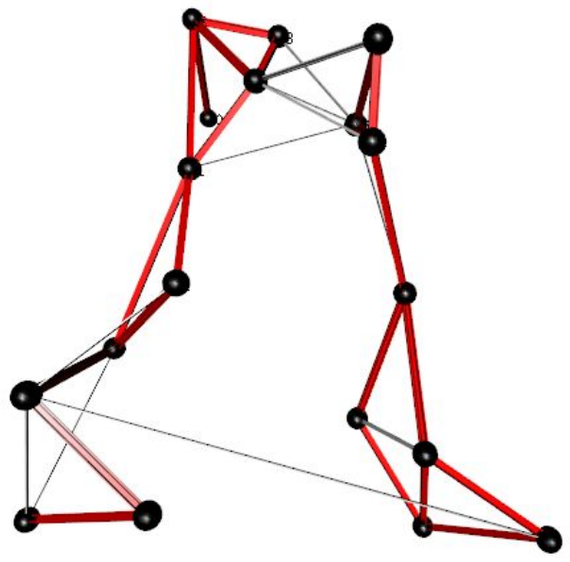
Spatial properties of the network

Spatial information? The graph itself does not have 3D information, but strong (red) bonds are occurring only if two molecules are spatially closed

Replica 3 - fast

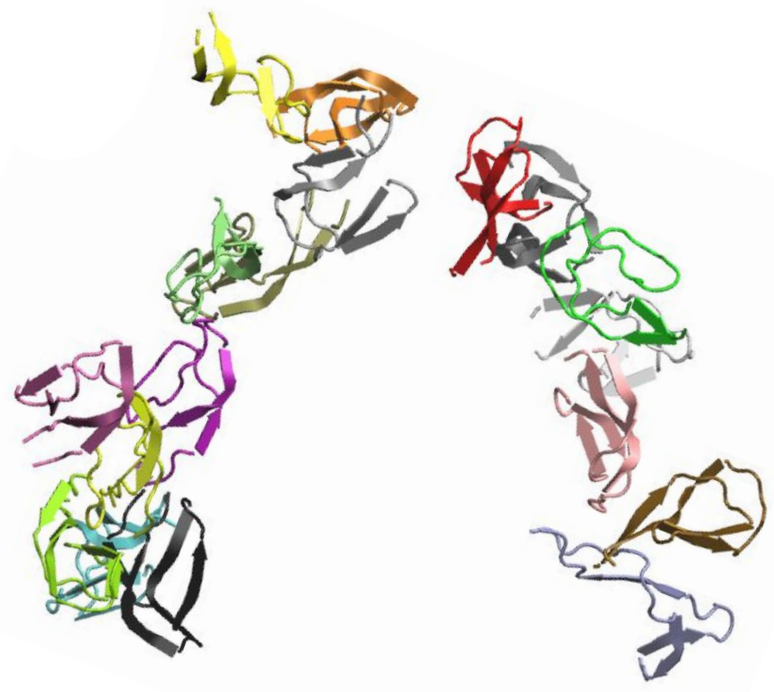


Graph



3D representation of the Graph, with nodes in the positions of the COM of the different prions

Final structure



Markov State Model of aggregation

=

Kinetic model of the aggregation process

How did we calculated the Markov State Model?

1. For all replicas

1.1 Choice of the lag time (the time elapsed between two consecutive structures)

2. For each replicas

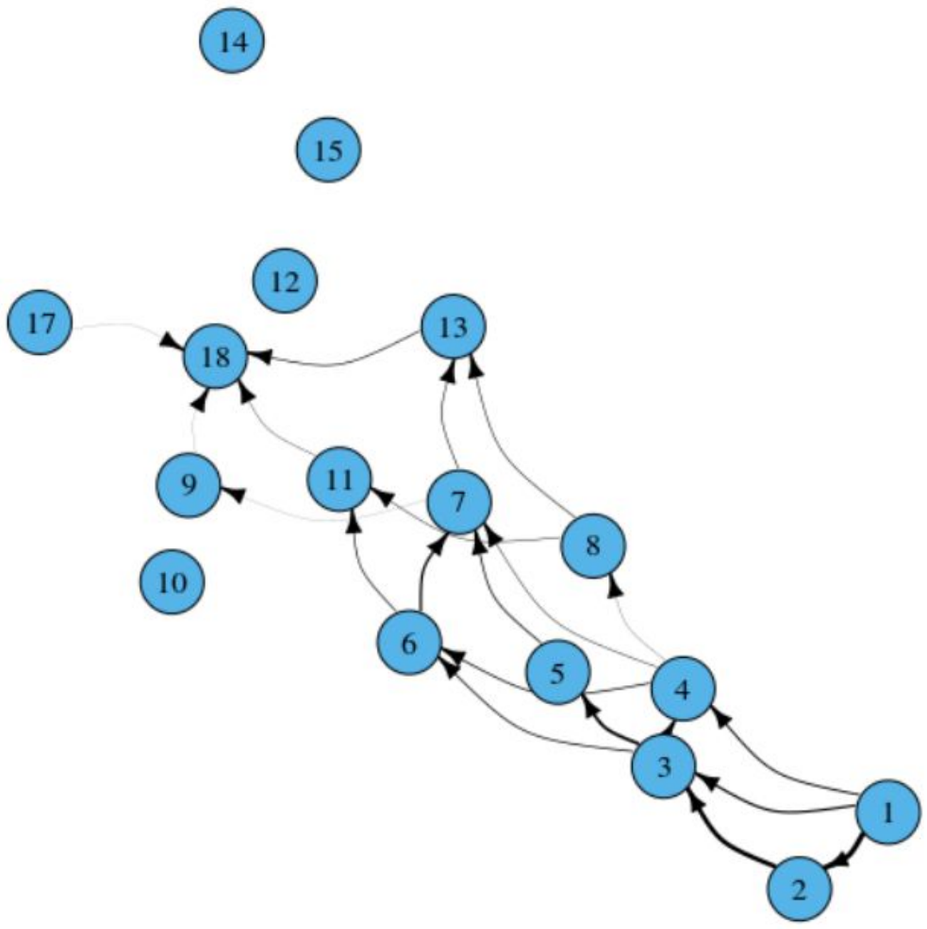
2.1. Creation of the structure according to the lag time
2.2 For each structure calculate the micro state (the maximum aggregate size)
2.3 Create the non redundant list of the observed micro states
2.4 Calculation of the Transition Count Matrix (the number of transitions between each pair of microstates)

3. Merging informations from all replicas

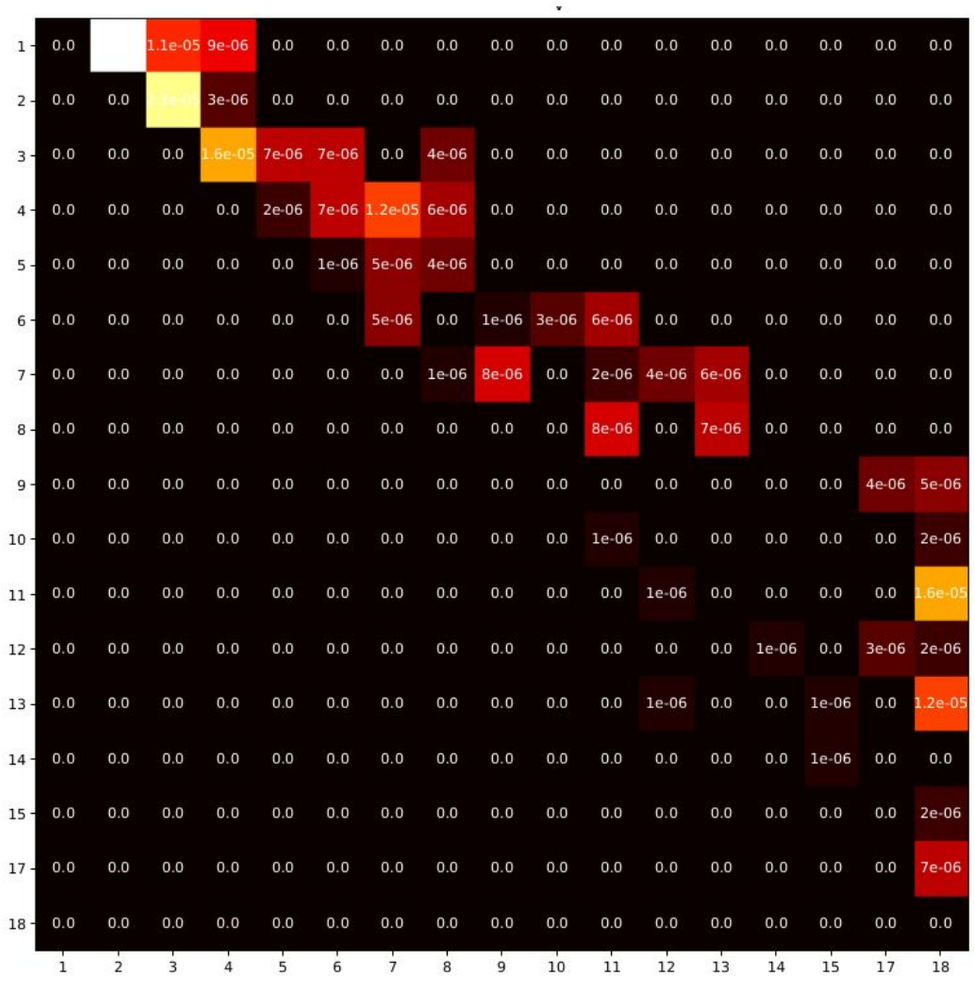
3.1. Make the non redundant list of the observed microstates
3.2 Calculation of the Transition Count Matrix (the number of transitions between each pair of microstates)
3.3 Symmetrization of the Transition Count Matrix
3.4 Calculation of the Transition Probability Matrix: row-normalization of the Transition Count Matrix

Markov Chain (Net flux) Merging all Replica

Schematic of the fluxes of the Markov state Model, filtering out fluxes < 5E-06
(Flux range [0, 26E-06])



Matrix of Net fluxes:
row = initial state, column = final state
Black 0, white 26E-06

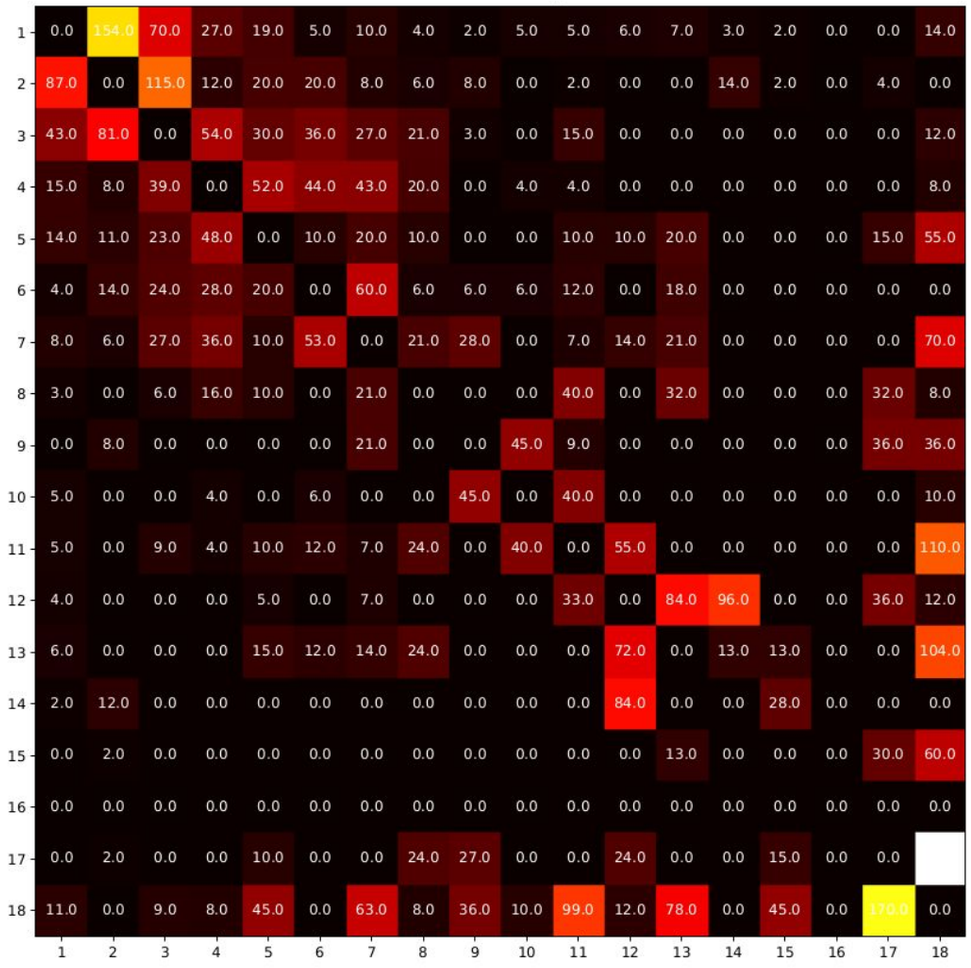


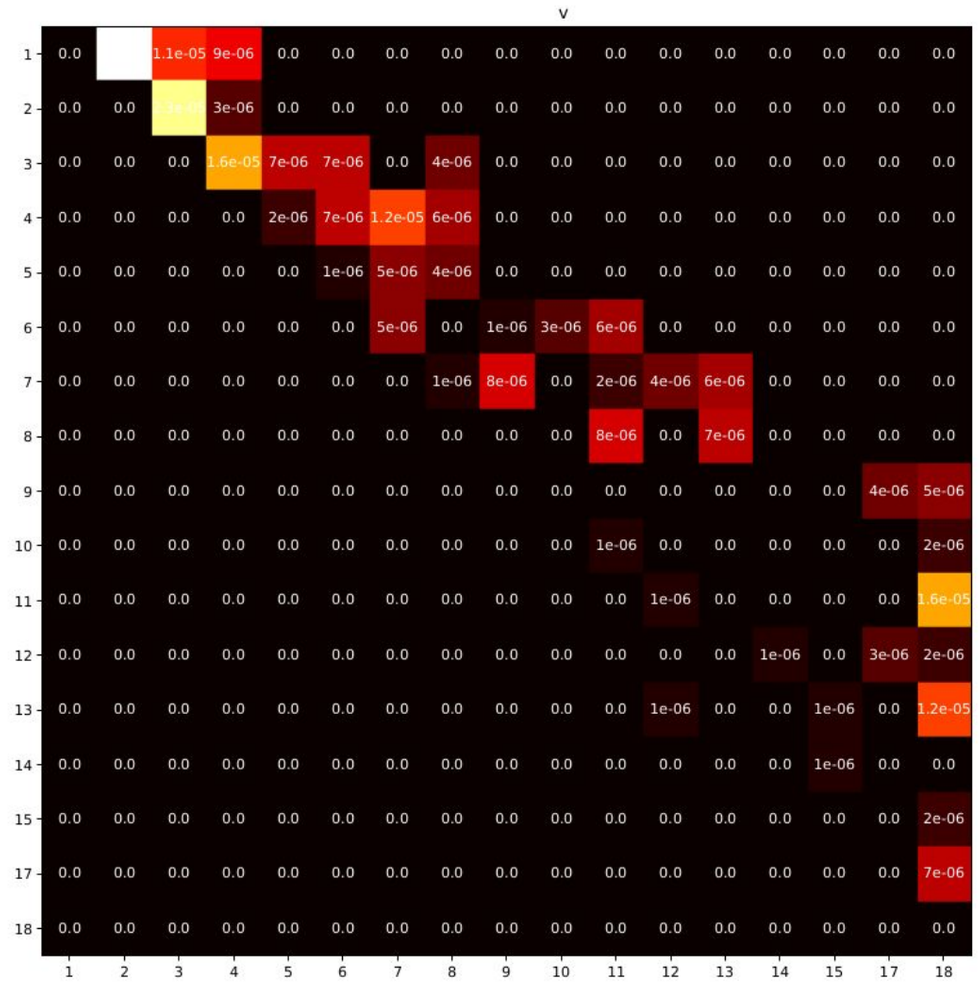
Confirming what found from the survival analysis, some large oligomers (e.g. 10-mer, 12-mer, …) are not contributing significantly to the pathways, being present (as Max Aggr) only occasionally. Aggregates up to 6-mers are involved in a more dynamical exchange.

Markov Chain (Net flux) Merging all Replica

Comparison between heat map of all transitions observed (left, color code: black = 0, white = 220 transitions, summing over all the replicas) and net fluxes of the MSM (right, color code: black = 0, white = 26E-06)

Row = initial state; column = final state.





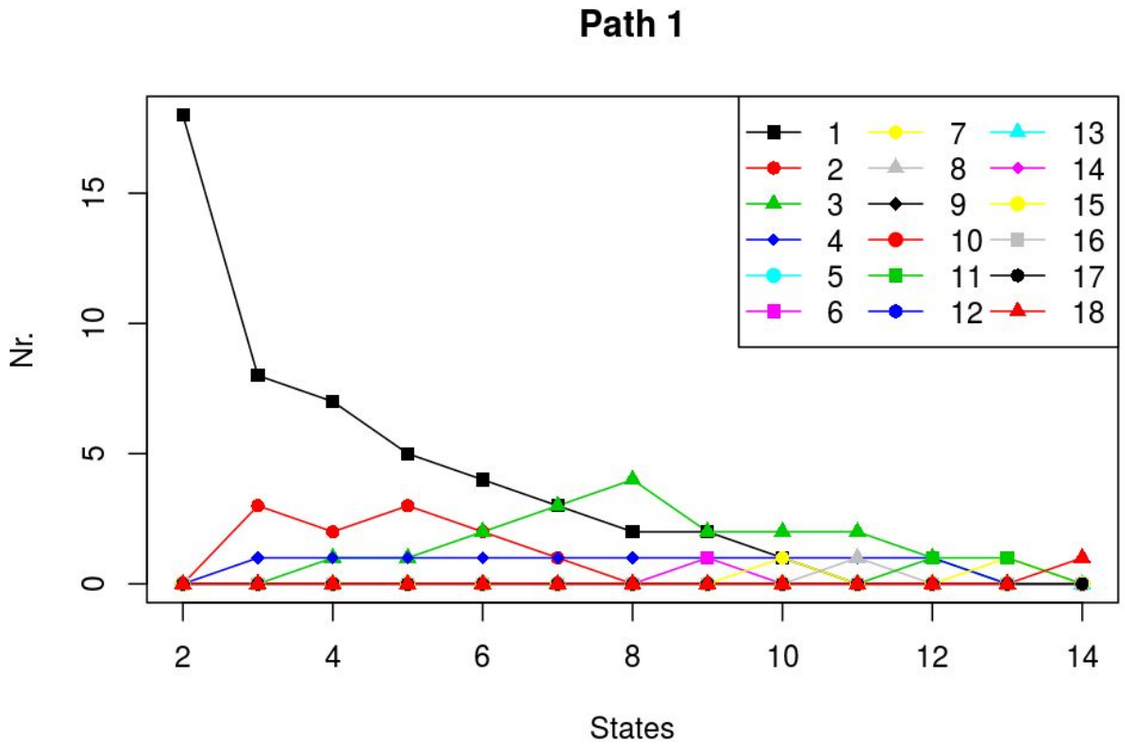
The net process is going toward the direction of larger aggregates, but on the short timescale the process is dynamic, including formation and rupture of aggregates.

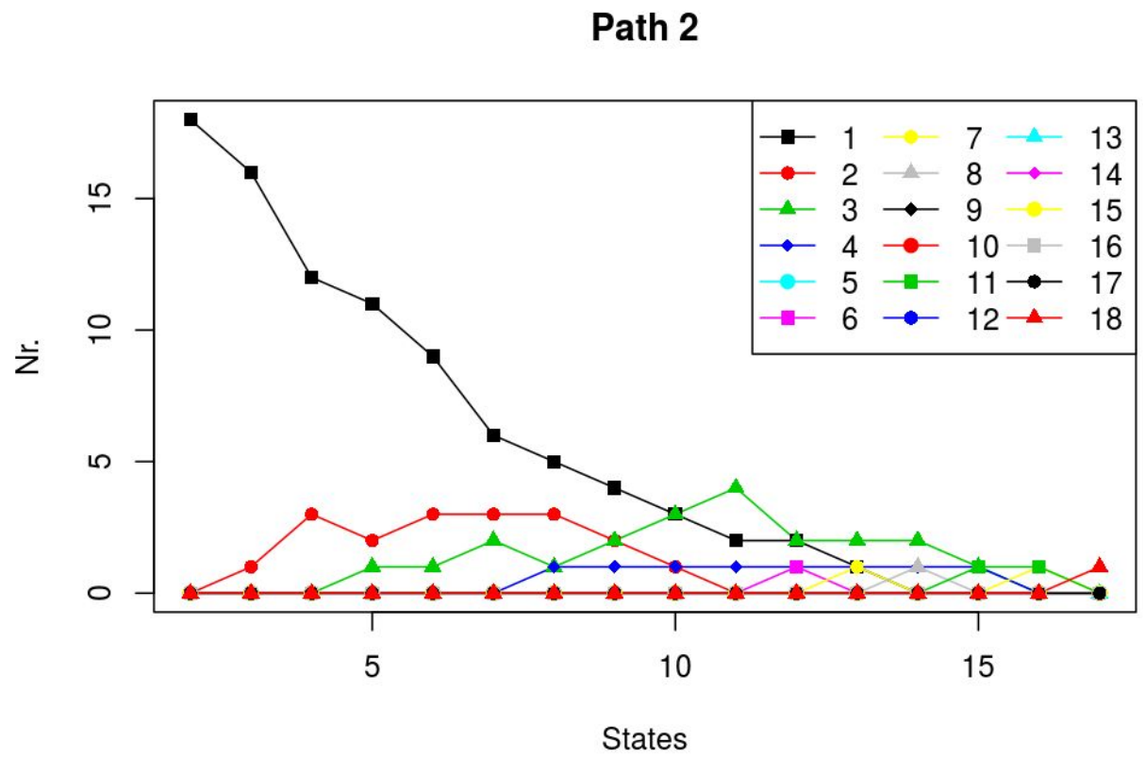
Markov Model with full size distribution at each time

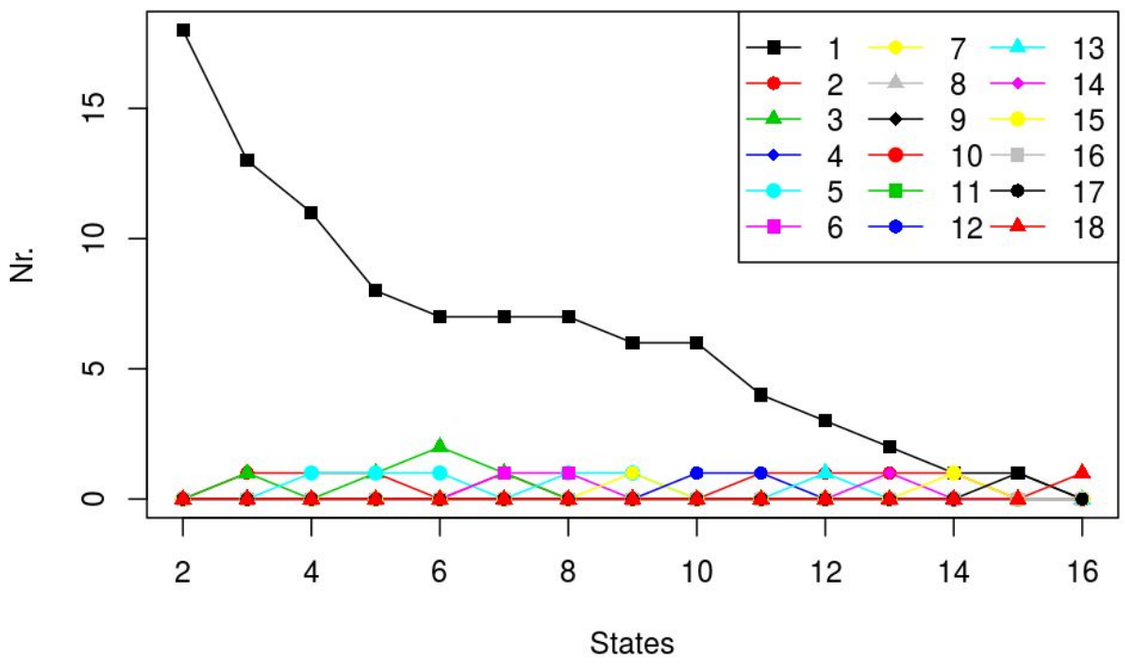
- MSM which consider as microstate the distribution of sizes at each time (i.e. how many monomers, dimers, …), average on all replicas, gives many equivalent pathways (i.e. with the same probability). Most of them includes many dimers, i.e. prions dimerize and then join in bigger aggregates.

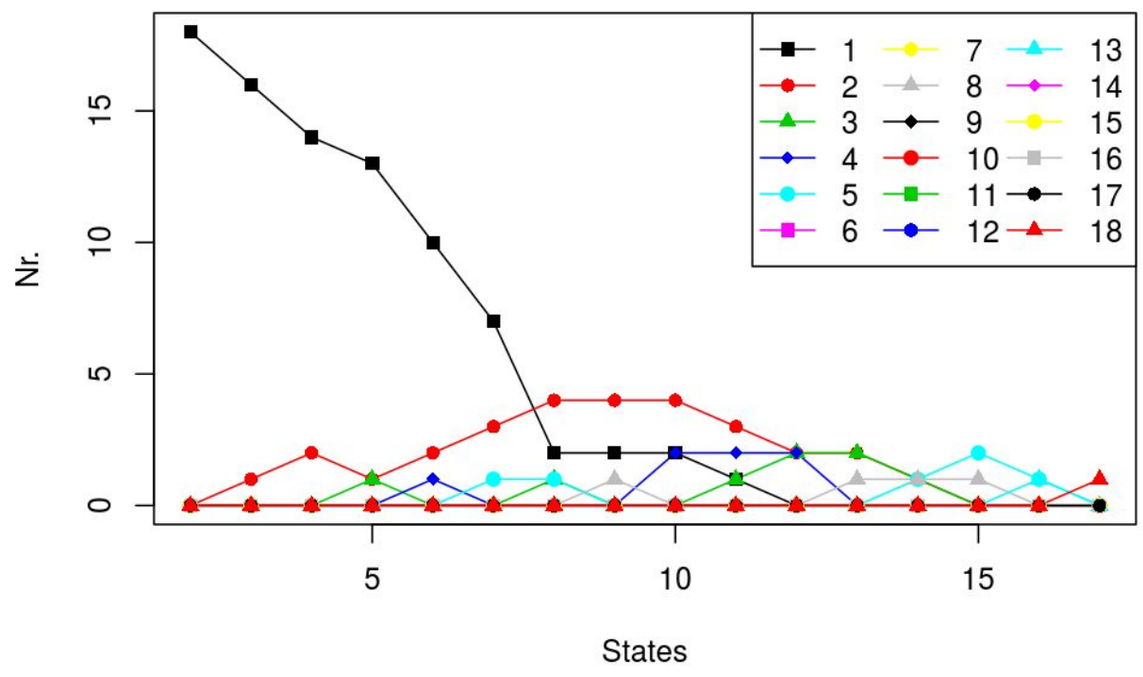
- First 10 pathways (up to 70% of flux)

Markov Model with full size distribution at each time

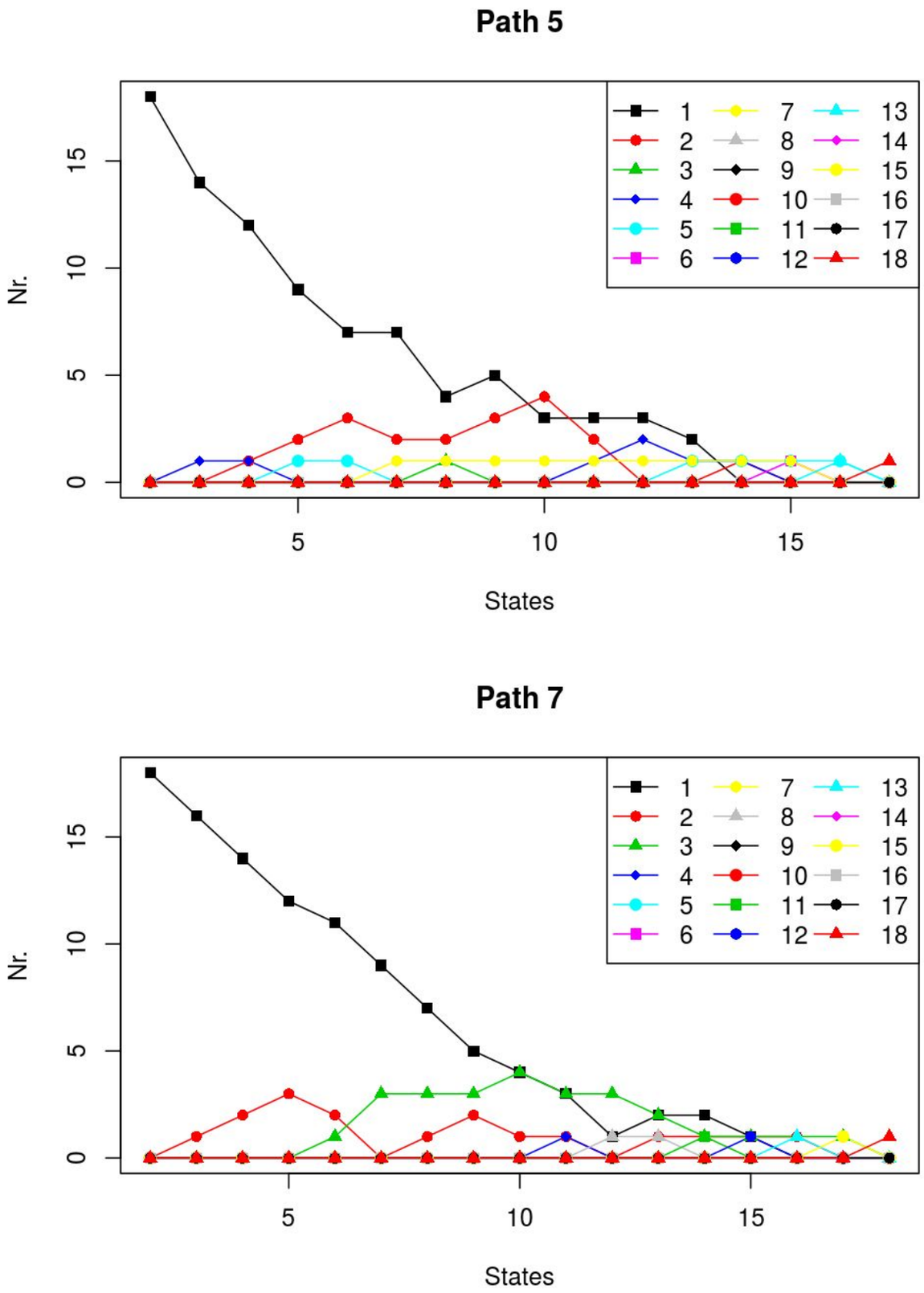


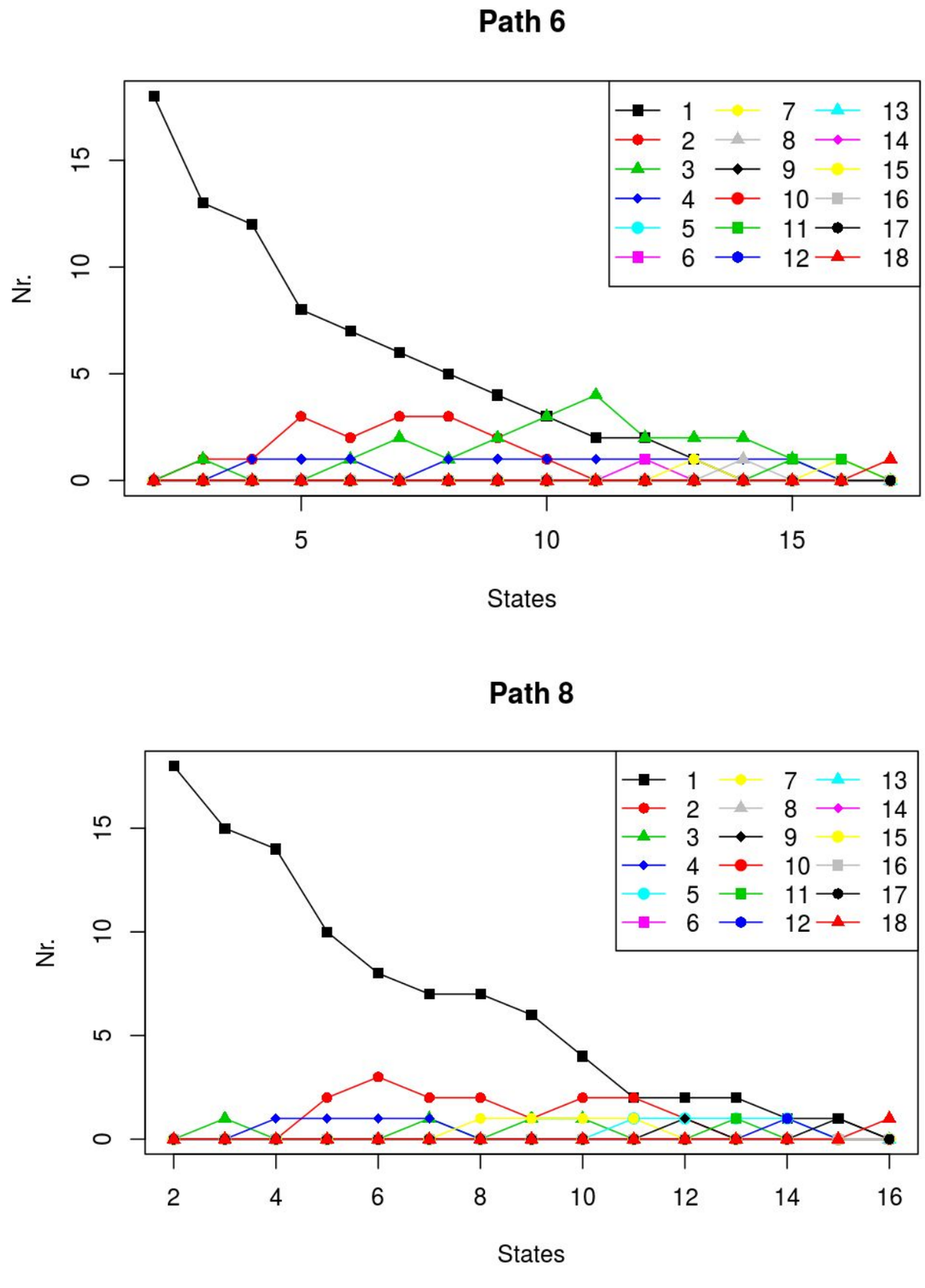




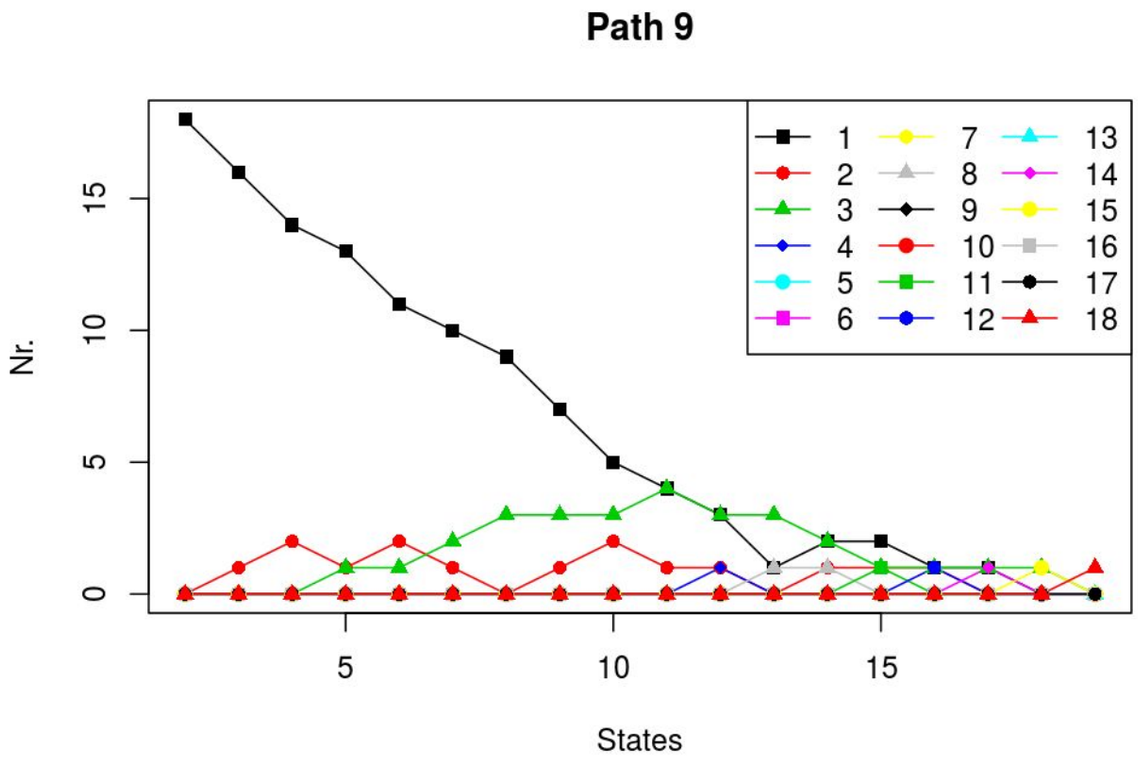


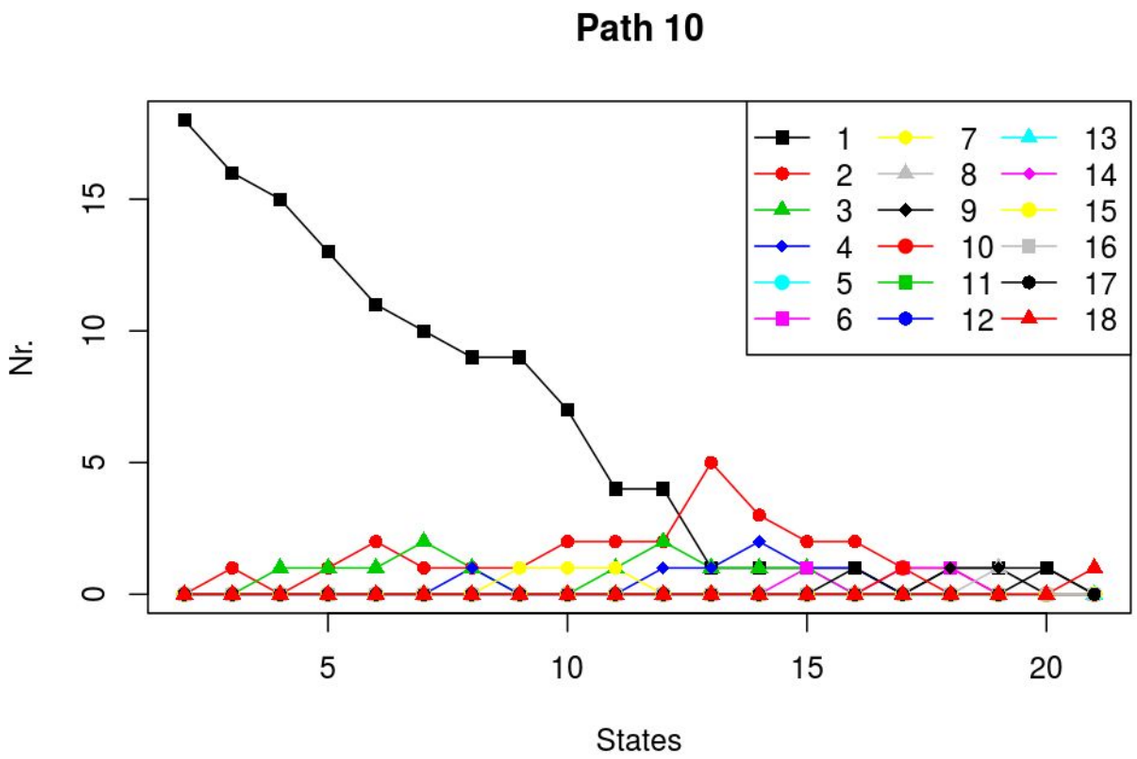
Markov Model with full size distribution at each time





Markov Model with full size distribution at each time





Caveats on Markov Chains Analysis

- the statistic is done over all the replicas (excluding the merged ones). A single replica analysis provides little statistics. No clear difference in the pathways is identified between fast and slow replicas.

- NO TIME INFORMATION