Progress report :
Repair probabilities for the RCL polymer with two random breaks

Ignacio Madrid

April 10, 2018

1   Introduction

The model simulated is described in figure 1. We study the probability of repair after two breaks, i.e., the probability of having encounters (see figure)

$$A_1 - A_2 \quad \text{or} \quad B_1 - B_2$$

before

$$A_1 - B_1 \quad \text{or} \quad A_1 - B_2 \quad \text{or} \quad A_2 - B_1 \quad \text{or} \quad A_2 - B_2$$

We define the first encounter time as

$$T_{m,n}^\epsilon := \inf\{t \geq 0 : ||(R_m)_t - (R_n)_t|| < \epsilon\}$$

where R_m and R_n are the positions of monomers m and n, and ϵ is called the encounter distance. So the repair probabilty is calculated in the simulations as

$$\mathbb{P}(Repair) = \mathbb{P}(T_{A_1,A_2}^\epsilon \wedge T_{A_1,A_2}^\epsilon \leq T_{A_1,B_1}^\epsilon \wedge T_{A_1,B_2}^\epsilon \wedge T_{A_2,B_1}^\epsilon \wedge T_{A_2,B_2}^\epsilon)$$

NB: By pure chance, the expected repair probability is $2/6 = 1/3$.
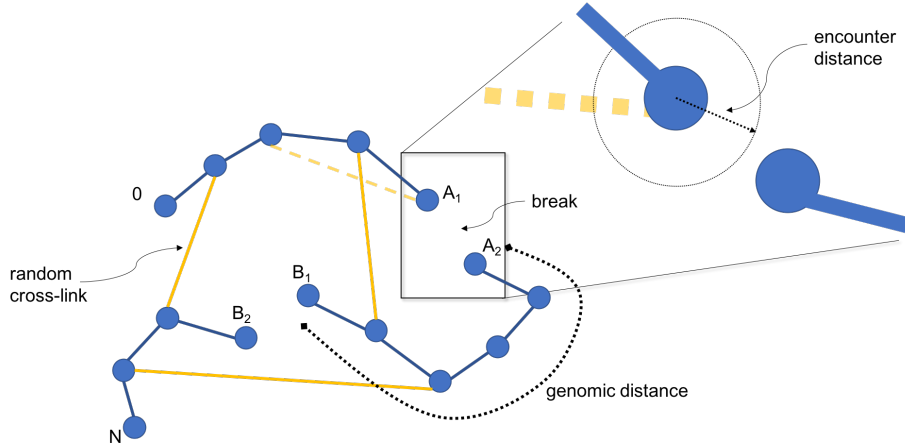
1

Figure 1: Model of the RCL polymer used for the simulations. Two random breaks were induced with a deterministic genomic distance (in number of monomers) between them. Random cross-links are built uniformly over all combinations of non-neighbor monomers. In the moment of cut, the cross-links with cleaved monomers (for instance, the cross-link with the monomer A_1 in the figure) may also be removed along with the cut bonds. Finally, we say there is a repair when two separated neighbor monomers encounter at a distance inferior to an encounter distance ϵ. On the other hand, if another non-neighbor combination of A_1, A_2, B_1, B_2 encounters at such distance we say it is a failure.

2   First Encounter Time (FET) distribution

Since an important peak was observed at instant 0 in the first simulations, we decided to wait some time after the induction of double strand breaks (DSBs), before measuring if any pair of monomers have encountered. Concerning the encounters events, if two or more pairs of monomers encounter at the same simulated instant, we toss a coin to choose uniformly over those pairs to be the first encounter.

The results of figure 2 indicate that the removal of the cross-links seems not affect the mean first encounter time, nor the ratio of fail/repair events. The estimated rate parameter (λ) is 0.08 for both cases. It should be noted that since the maximum time of simulation is set at 50 seconds (i.e. if nothing happens in 50 seconds the simulation is discarded) the distribution is biased and corresponds to $\mathcal{L}(FET|FET < 50)$. For instance, see figure 3 where the simulation maximum time has been increased to 200 seconds and has an estimated rate parameter of 0.04.
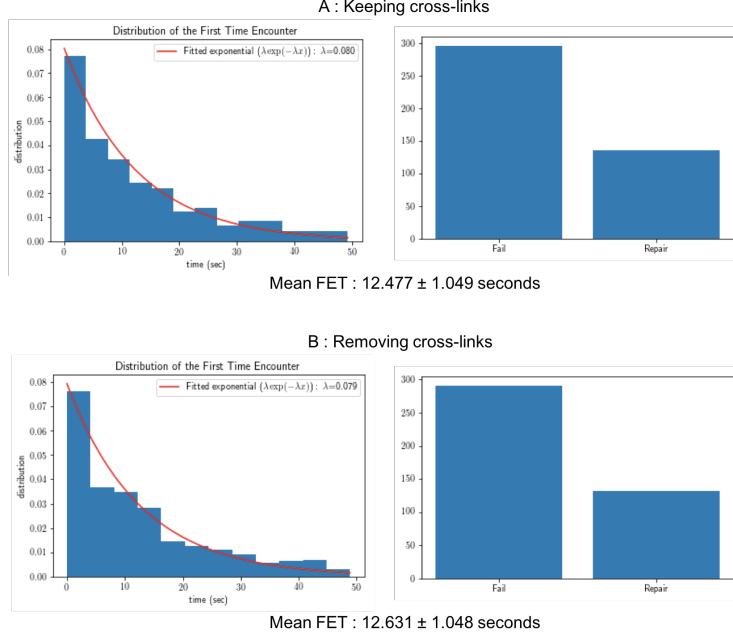
Figure 2: Distribution of the first encounter time keeping and removing the cross-links in the cleaved monomers. An exponential distribution has been fitted to each histogram. Right: Distribution of events (fail or repair). The results corresponds to a polymer of 100 monomers, 5 random cross-links and a waiting time of 25 seconds.

3   Repair probabilty

3.1   Effect of the number of cross-links

As we can see in figure 4, removing the cross-links in the separated monomers does not seem to induce a better repair probability when increasing the number of cross-links. However the genomic distance could have a more important role. Increasing the number of cross-links approaches the simulated probabilities to $1/3$, which is the result expected by pure chance. Indeed, increasing the number of CLs and therefore approaching all monomers not only increases the chances of a good matching, but also the bad match case. It is interesting however that systems which had a repair probability $\leq 1/3$ because had DSBs at a small genomic distance (so the chances of a bad matching where higher) improve their repair probabilities when we increase the number of cross-links. In other words, the compactness may help to repair polymers cut in a small neighborhood.

The most interesting remark though is the gap between the probabilities at a genomic distance of 1 (i.e. the two DSBs are connected, $A2$ and $B1$ being neighbors) and the other genomic distances. So far, no problems with the code
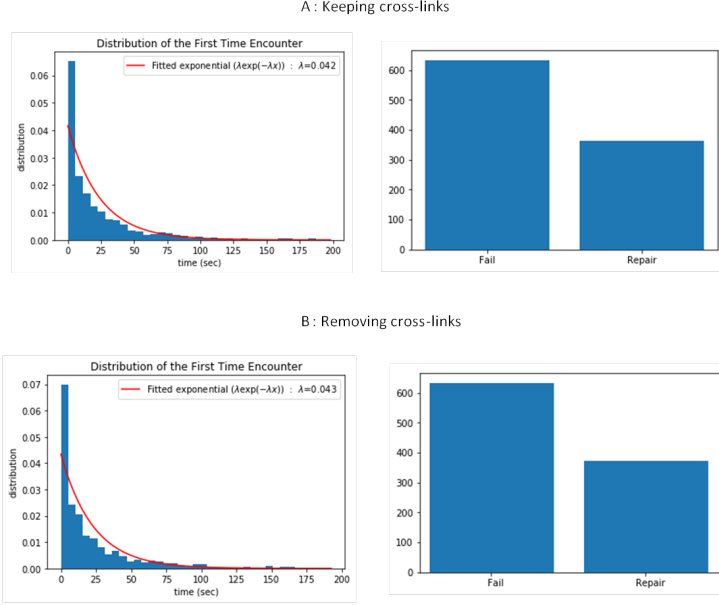
3

Figure 3: Distribution of the first encounter time keeping and removing the cross-links in the cleaved monomers with a maximum simulation time of 200 seconds.
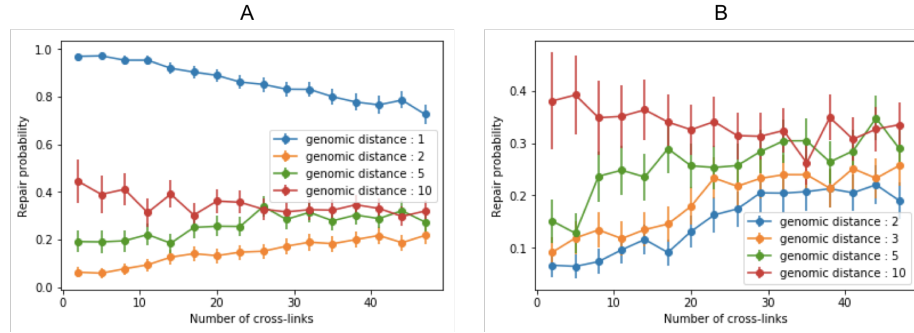

Figure 4: Polymer of 100 monomers. Repair probability against the number of random cross-links: A) keeping the cross-links with cleaved monomers ; B) removing the cross-links with cleaved monomers. 500 iterations for each. NB : In the case B) the lowest genomic distance is 2.

that could have leaded to a wrong calculation of the repair probability have been detected. Some hypothesis: the wrong combination $A_2 - B_1$ is no possible anymore, so in fact, by pure chance, we have a 2/5 probability of repair (instead of 2/6), and the combinations $A_2 - B_2$, $A_1 - B_1$ could be very rare. Besides,

as only valid cuts are allowed (i.e. the polymer rests fully connected after the breaks) the segment $A_2 - B_1$ is indeed connected to the upstream and/or the downstream fragment. The probability of being connected to only one of them is higher than the probability of being connected to both, so the segment could be pushed towards one of the free monomers rapidly and then induce a good ratio of repair.

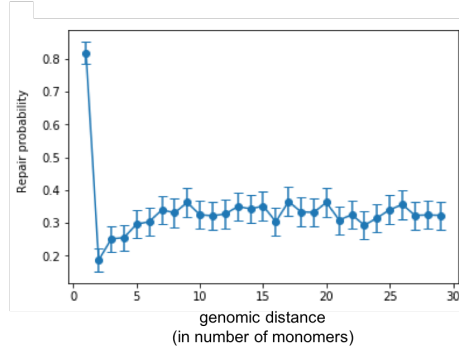## 3.2 Effect of the genomic distance between two DSBs

Figure 5: Repair probability against the genomic distance between the two DSBs. Polymer of 100 monomers with 25 random cross-links. Encounter distance of 0.1 μm (and b = 0.2 μm).

When the cross-links in the breaks are kept (fig. 5), we see that there is an important gap between the probability of repair for the genomic distance of 1 and the rest, as we have seen in the previous simulations.

The number of cross-links is relatively high (25 over 100 monomers) and we see that regardless how large the genomic distance could be, the probability of repair tends to the expected result by pure chance, 1/3. So, **it would be more interesting to analyze the effect of the genomic distance with a weak number of cross-links.**

In terms of the dependence on the encounter distance ε, the behavior seems to be the same regardless its value (figures 6 and 7), when we keep or remove the CLs in the cleaved monomers. In the curves, small values of ε have more variance since the number of realizations is fixed and the event of an encounter becomes rarer (NB: the maximum time of simulation is set at 10 seconds, if nothing happens the simulation is discarded):

3.3 Effect of the encounter distance

We see in figure 8 that encounter distances greater than the nanometric scale induces repair probability that tends to the expect pure chance value of 1/3 (a finer analyses could be also seen in figure 9 which confirms this trend). Again,
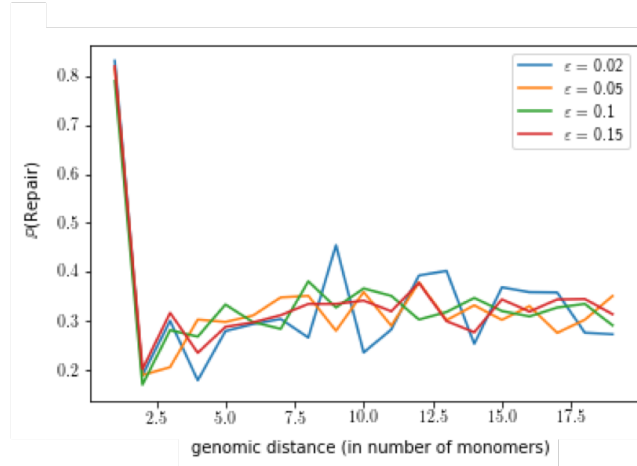
Figure 6: Repair probability against the genomic distance between the two DSBs and different encounter distances. Polymer of 100 monomers with 25 random cross-links. CL in the breaks are kept.

the variance in the first mean observations are greater since the events are rare. Studying nanometric encounter distances would require a greater number of simulations or rare events sampling techniques (?).

4 TODO

- Measure some statistical properties, such as the mean squared radius of gyration ton confirm the compactness induced by the random cross-links and study its shape.

- Measure the mean first encounter time in function of the parameters presented here.

- Perform the same tests with fixed cut loci.

- Add the effect of volume exclusion and other external statistical and dynamical forces.

- Implement the β-polymer and define how the DSBs are going to be performed. Study the same properties presented here.
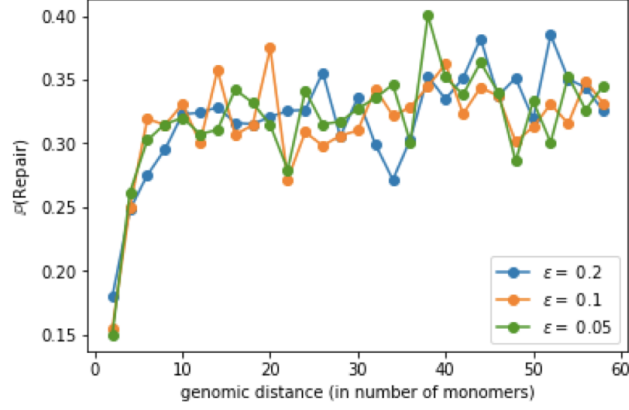
Figure 7: Repair probability against the genomic distance between the two DSBs and different encounter distances. Polymer of 100 monomers with 10 random cross-links. CL in the breaks are removed.
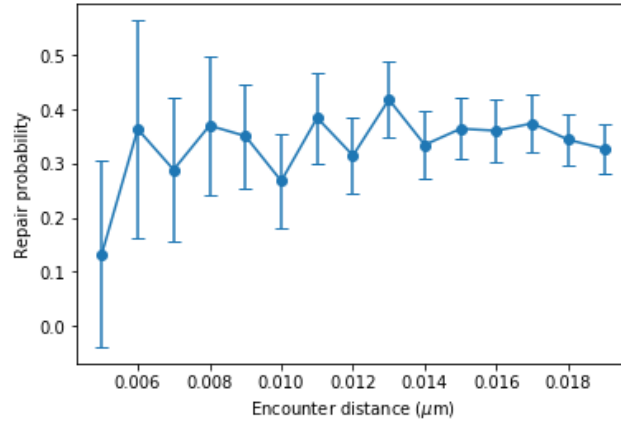

Figure 8: Probability and 95% confidence interval of repair, in function of the encounter distance that defines the repair. For the simulations, a random polymer of 100 monomers with a standard deviation of $b = 0.2\mu$m for the distance between neighbor monomers, and 4 random cross-links is used. CLs in the breaks are removed. DSBs are at a genomic distance of 10 monomers. 1500 realizations.
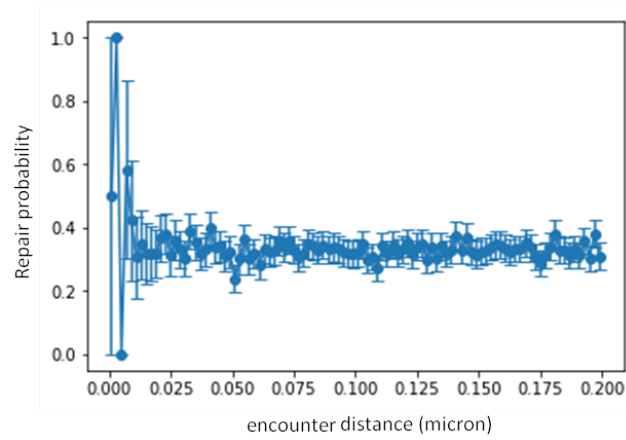
Figure 9: Probability and 95% confidence interval of repair, in function of the encounter distance that defines the repair in a finner grid of encounter distances.