

# Параметры

## 1. Параметры симуляций:

1. **STOP\_TIME** — длительность симуляции в секундах (биологическое время). Можно указать 0, и тогда критерием останова будет событие, при котором все целевые сайты окажутся занятыми хотя бы один раз.
2. **ENSAMBLE\_SIZE** — число независимых симуляций.
3. **RANDOM\_SEED** — seed для генератора случайных чисел. Для различного поведения генератора для множественных симуляциях нужно указать значение 0.
4. **COMPUTED\_AFFINITY\_PRECISION** — точность вычисления аффинности, количество знаков после запятой.
5. **DNA\_SECTOR\_SIZE** — размер сектора ДНК. ДНК разбивается на секторы и таким образом увеличивается скорость достижения пустых районов ДНК. Значение 0 для автоматического выбора.
6. **EVENT\_LIST\_SUBGROUP\_SIZE** — размер подсписка в списке событий. Используется только для прямого метода Гиллеспи. Значение 0 для автоматического выбора.
7. **EVENT\_LIST\_USES\_FR** — значение true, если используется метод первой реакции. Значение false, если используется прямой метод (Gillespie, 1976, 1977, 2007).

## 2. Параметры вывода:

1. **OUTPUT\_FOLDER** — название директории, где будут храниться результаты.
2. **OUTPUT\_FILENAME** — строковый идентификатор для всех файлов вывода, будет добавлен в начало названия всех файлов. Если оставить пустое поле, будет сгенерирован автоматически.
3. **PRINT\_INTERMEDIARY\_RESULTS\_AFTER** — время в секундах, через которое будут сгенерированы промежуточные результаты. Если указано значение 0, промежуточных результатов не будет.

4. **PRINT\_FINAL\_OCCUPANCY** — значение true, если нужно выводить финальную занятость в конце симуляций.
5. **DEBUG\_MODE** — значение true, если симулятор запущен в режиме debug (все действия выводятся в status файл). Для длительных симуляций параметр должен иметь значение false.
6. **OUTPUT\_TF** — названия тех ТФ, поведение которых необходимо отслеживать: симулятор выведет информацию о числе свободных молекул, числе свободных сайтов и пропорции свободных сайтов. Если молекул данного типа меньше 10, позиции каждой молекулы будут записаны в файл.
7. **OUTPUT\_TF\_POINTS** — число промежуточных временных точек, в которых записывается динамическое поведение ТФ-ов, которые отслеживаются.
8. **FOLLOW\_TS** — значение true, если симулятор будет выводить динамическое поведение целевых сайтов.
9. **OUTPUT\_AFFINITY\_LANDSCAPE** — значение true, если симулятор будет выводить профайл среднего времени ожидания в конце симуляции.
10. **OUTPUT\_BINDING\_ENERGY** — значение true, если симулятор будет выводить энергию присоединения вместо среднего времени ожидания.
11. **OUTPUT\_DNA\_OCCUPANCY** — значение true, если симулятор будет выводить занятость ДНК в конце симуляции.
12. **DNA\_OCCUPANCY\_FULL\_MOLECULE\_SIZE** — значение true, если связанная молекула будет оказывать влияние на занятость ДНК во всей области, которую занимает на ДНК. Значение false, если при подсчете занятости ДНК за область связывания молекулы принимается только первая позиция.
13. **OUTPUT\_SLIDING\_LENGTHS** — значение true, если симулятор будет выводить все длины слайдинга.
14. **WIG\_STEP** — значение шага в wig-файле фиксированного шага для вывода занятости.
15. **WIG\_THRESHOLD** — граница (как процент наивысшего пика) для фильтрации пиков в wig-файле. Значение -1 для автоматического выбора, значение 0 для того, чтобы не использовать границу.

### 3. Параметры видов ТФ:

#### 1. Загрузка ТФ:

1. **TF\_FILE** — имя csv-файла, содержащего данные ТФ-ов. Имя должно содержать относительный или абсолютный путь до файла. Указание файла обязательно, если ничего не указывать, симулятор сгенерирует ТФ-ы самостоятельно.
2. **TF\_COOPERATIVITY\_FILE** — имя csv-файла, содержащего данные кооперативности ТФ. Имя должно содержать относительный или абсолютный путь до файла. Указание файла необязательно.
3. **TS\_FILE** — имя файла, содержащего данные целевых сайтов ТФ. Имя должно содержать относительный или абсолютный путь до файла. Указание файла необязательно.

#### 2. Общие параметры ТФ-ов:

1. **TF\_READ\_IN\_BOTH DIRECTIONS** — значение true, если ТФ-ы садятся на обе цепи ДНК.
2. **SLIDING\_AND\_HOPPING\_AFFECTS\_TF\_ASSOC\_RATE** — значение true, если слайдинг и хоппинг влияют на скорость присоединения между молекулами ТФ и ДНК.
3. Остальные параметры указываются в файле с ТФ-ми. Если файл не указан, ТФ-ы генерируются случайно с параметрами, указанными ниже:
  1. **TF\_SPECIES\_COUNT** — число различных видов ТФ, которые необходимо сгенерировать.
  2. **TF\_COPY\_NUMBER\_MIN** — минимальное число молекул для типа ТФ.
  3. **TF\_COPY\_NUMBER\_MAX** — максимальное число молекул для типа ТФ.
  4. **TF\_DBD\_LENGTH\_MIN** — минимальная длина связывающего домена (п.о.).
  5. **TF\_DBD\_LENGTH\_MAX** — максимальная длина связывающего домена (п.о.).

6. **TF\_ES** — энергетический штраф за несовпадение (Gerland et al., 2002). Это фактически  $\epsilon$  из (Gerland et al., 2002), измеряется в К<sub>BT</sub>.
7. **TF\_SIZE\_LEFT** — длина левой окрестности связывающего домена ТФ, занимаемая на ДНК при связывании.
8. **TF\_SIZE\_RIGHT** — длина правой окрестности связывающего домена ТФ, занимаемая на ДНК при связывании.
9. **TF\_ASSOC\_RATE** — «скорость» присоединения ТФ-ов. Измеряется в с<sup>-1</sup>.
10. **TF\_PREBOUND\_PROPORTION** — процент связанных молекул на момент начала симуляций.
11. **TF\_PREBOUND\_TO\_HIGHEST\_AFFINITY** — значение true, если ТФ-ы на момент начала симуляций ТФ-ы сидят в сайтах наивысшей аффинности.

#### 4. ДНК — параметры, связанные с последовательностью ДНК:

##### 1. Загрузка последовательности ДНК:

1. **DNA\_SEQUENCE\_FILE** — fasta-файл, содержащий последовательность ДНК.
2. Если файл с ДНК не указан или пуст, ДНК генерируется случайно, согласно параметрам:
  1. **DNA\_LENGTH** — длина ДНК в п.о.
  2. **DNA\_PROPORTION\_OF\_A** — доля аденина в ДНК.
  3. **DNA\_PROPORTION\_OF\_T** — доля тимина в ДНК.
  4. **DNA\_PROPORTION\_OF\_C** — доля цитозина в ДНК.
  5. **DNA\_PROPORTION\_OF\_G** — доля гуанина в ДНК.
  6. **DNA\_BOUNDARY\_CONDITION** — граничные условия на ДНК: absorbing/reflexive/periodic.

#### 5. Параметры случайного блуждания ТФ:

##### 1. Предположения:

1. **CHECK\_OCCUPANCY\_ON\_BINDING** — значение true, если симулятор будет проверять занятость ДНК перед связыванием.
  2. **CHECK\_OCCUPANCY\_ON\_SLIDING** — значение true, если симулятор будет проверять занятость ДНК перед слайдингом.
  3. **CHECK\_OCCUPANCY\_ON\_REBINDING** — значение true, если симулятор будет проверять занятость ДНК перед связыванием сразу после отсоединения(re-binding).
2. Параметры ТФ по умолчанию — используются для ТФ, которые генерируются случайно, или когда соответствующая информация о ТФ отсутствует:
1. **TF\_IS\_IMMOBILE** — значение true, если ТФ после связывания становится неподвижным и остается в той же позиции до конца симуляций.
  2. **TF\_UNBINDING\_PROBABILITY** — вероятность отсоединения ТФ в течение следующего события.
  3. **TF\_SLIDE\_LEFT\_PROBABILITY** — вероятность слайдинга влево в течение следующего события.
  4. **TF\_SLIDE\_RIGHT\_PROBABILITY** — вероятность слайдинга вправо в течение следующего события.
  5. **TF\_JUMPING\_PROBABILITY** — вероятность совершить прыжок при отсоединении вместо того, чтобы вернуться на ДНК.
  6. **TF\_HOP\_STD\_DISPLACEMENT** — стандартное перемещение ТФ по время скачка (hop). Имеет нормальное распределение и измеряется в п.о.
  7. **TF\_NONSPECIFIC\_WAITING\_TIME** — время ожидания конкретного ТФ в секундах. Фактически это  $\tau_0 \cdot \exp(E_{ns})$  из (Gerland et al., 2002).
  8. **TF\_STEP\_LEFT\_SIZE** — длина шага при слайдинге влево, п.о.
  9. **TF\_STEP\_RIGHT\_SIZE** — длина шага при слайдинге вправо, п.о.
  10. **TF\_UNCORRELATED\_DISPLACEMENT\_SIZE** — длина независимого перемещения ТФ в п.о.: если длина скачка (hop)

больше этого значения, то скачок превращается в прыжок, и молекула должна сначала попасть к нуклеоплазме.

11. **TF\_STALLS\_IF\_BLOCKED** — значение true, если молекула ТФ остается в текущей позиции, когда не может переместиться посредством скачка, и значение false, если молекула отсоединяется в результате неудачного события скачка (hopping).
12. **TF\_COLLISION\_UNBIND\_PROBABILITY** — вероятность отсоединения молекулы ТФ при столкновении с другой молекулой.
13. **TF\_AFFINITY\_LANDSCAPE\_ROUGHNESS** — неоднородность аффинности для второстепенных ТФ (non-cognate), используется для случайной генерации «ландшафта» аффинности, изменяется в  $K_B T$ .