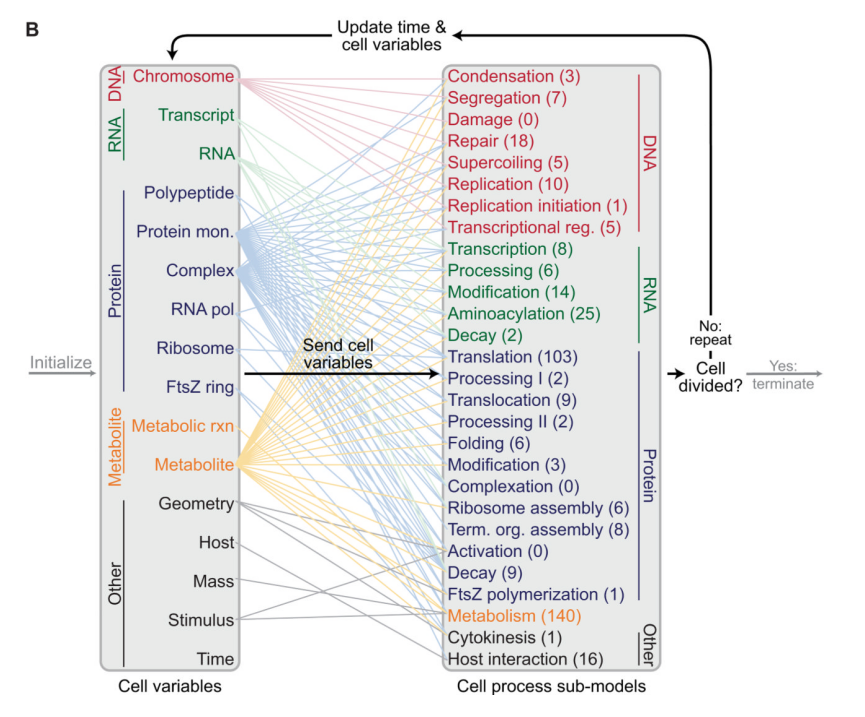
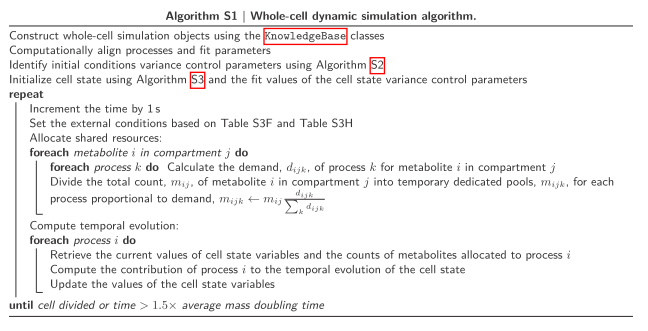
*Mycoplasma genitalium*: human urogenital parasite, circular DNA with 525 genes and 580 070 base pairs



410 genes in the simulation over the 525 genes of *Mycoplasma Genitalium*



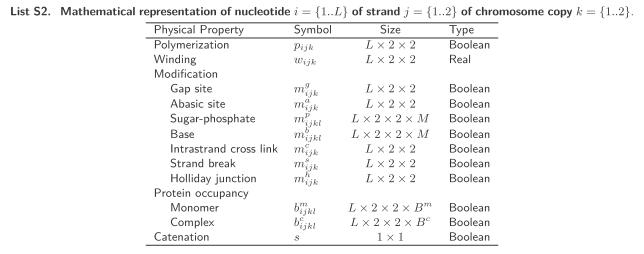
Independence assumption of the 28 sub-model processes? Done via resource allocation. How allocation of resource is performed. Highly unclear!

How initialization is performed?

Validation performed against:

1. Essentiality of each *Mycoplasma Genitalium* gene reported by Glass et al.;
2. Measured growth rates of 12 non-essential *Mycoplasma Genitalium* single-gene disruption strains;
3. Cytosolic concentrations of 39 *E. Coli* metabolites reported by Bennett et al. 392 and curated by Sundararaj et al.
4. Validation on cytosolic concentration was performed against *E. Coli* for *M. Genitalium*

|  |  |  |  |
| --- | --- | --- | --- |
| State name | Computation | Comments | Our state |
| Time | 1 s increment | Not possible to change the step length  All processes are run with the same time step  Maximum of 50 000 seconds (150% of M. Genitalium mean mass doubling cycle) | Time  Step lengths can be parameterized by the user  Different step lengths for different categories  Maximal step length specififed by the user |
| Geometry | Rod like shape: 2 models depending on the presence of the septum | Assumption: cell density and width are constant during the cell cycle  Heavily depends on the presence of the septum => how is formed the septum? Consistency between septum formation and cell cycle? | Volume  Same assumption is OK for our simulator  Is it possible to implement several shape: rod, spherical? |
| Mass | Sum of all elements in the cell: metabolites, chromosome, RNA, protein monomers and complexes | - | Mass |
| Chromosome | See List S2 (below) for representation | 525 (?) genes, 335 transcription units, 17 transcriptional regulators, 2 283 DnaA binding sites | DNA  Over 4 000 genes |
| FtsZ Ring | Polygon of FtsZ filaments: round down the number of fixed length filaments that can be present at the septum | - | !!! Nothing yet !!! |
| Metabolic reactions | Fluxes of metabolism  Performed via MOMA FBA with wild-type fluxes as target | 1 100 chemical reactions | - |
| Metabolite | Number of each metabolite | 722 metabolites | - |
| Polypeptide | Ribosome-bound mRNA  Length of nascent proteins  Sequence of aborted polypeptides  Proteolysis tag | Translation of 16 AA / second max | - |
| Protein complex | - | - | - |
| Protein Monomer | - | - | - |
| Ribosome | - | - | - |
| RNA | mRNA, rRNA, sRNA, tRNA, tmRNA | - | - |
| RNA polymerase | - | Gene regulation  Accesibility of DNA | !!! |
| Stimulus | Temperature  6 radiation types (Boolean)  3 stress (Boolean) | Temperature = 37° !!!  Stimulus seems to have no much effect on simulation (verify that DNA damage has any effect at all on transcription) | Environment  Allow constant and up/down shifts of nutriments  Should take into account also wholeCell bullshit |
| Transcript | - | 50 nucleotides / second | - |
| Host | Bullshit | Host interaction: only produces some proteins and localize it somewhere | !!! |



|  |  |  |  |
| --- | --- | --- | --- |
| Process name | Computation | Comments | Our process |
| Chromosome Condensation | SMC (Structural Maintenance of Chromosome) contribution to chromosome condensation | DNA supercoiling and segregation in specific classes  No macromolecular crowding and charge neutralization  SMC proteins bind to chromosome but do nothing! | No modeling planed yet!!! |
| Chromosome Segregation | Enzymatic separation and decatenation  Boolean for segregation (not well known) under 4 conditions: replicated; properly supercoiled; at least one free and active protein for 5 of the segregation proteins; enough GTP | Entropically driven chromosome separation not well known and not modeled | Planed |
| Cytokinesis | Pinching of membrane on septum region: circle, inscribe polygon, new smaller circle of circumference of the inscribed polygon | Starts only after segregation ends  Deterministic process even if each step is randomized | Planed |
| DNA Damage | No chemically induced damage | Bullshit  DNA damage does not seem to interact with transcription (TBC) | No modeling planed yet!!! |
| DNA Repair | - | Bullshit | No modeling planed yet!!! |
| DNA Supercoiling | Gyrase and Topoisomerase IV: 2 negative supercoil at each action with rate 1.2 and 2.5  Topoisomerase I: 1 positive supercoil at each action with rate 1 | No negative supercoil modeled by Topoisomerase IV because action is rare  Proccessitivity of gyrase and Topoisomerase IV  Supercoiling only affects the accessibility (and expression) of 3 transcription units: (gyrB, gyr A), (parC, pare) and (topA, MG119, MG120, MG121) | No modeling planed yet!!! |
| FtsZ Polymerization | ODE of Surovtsev et al. | Filaments of 9 subunits = 40 nm  Bullshit | No modeling planed yet!!! |
| Host Interaction | Bullshit: boolean of 6 properties  Not well known | No interaction with *Mycoplasma* behaviour | No modeling planed yet!!! |
| Macromolecular Complexation | Not well known  One step process (even if the real complexation is formed by several steps)  Alone in the world (does not interact with anything)  All specy rates are the same => rate computation => multinomial | All macromolecular complexation except formation of 30S and 50S (Ribosome Assembly), 70S (Translation), FtsZ ring (FtsZ Polymerization) and oriC DnaA complex (Replication Initiation)  201 macromolecules from 269 proteins and 5 rRNA | User specified |
| Metabolism | FBA with assumption: metabolism dynamics fast compared to 1 second step  FBA maximize metabolite production: maximize biomass flux with MOMA FBA with wild type fluxes target  Take into account thermodynamics and enzyme concentration | Intermediate energy carriers  Simplified metabolism due to the bacteria modeled: no pentose phosphate pathway, no amino acid, nucleotide, lipid and cofactor biosynthesis | Planed but need to take into account thermodynamics |
| Protein Activation | Boolean function of metabolite concentration, stimulus signal and temperature | No pH  Temperature = 37° | User specified |
| Protein Decay | Excess of decay machinery: nascent and mature proteins are degraded with Poisson distribution (half-life of 20 h)  Limitation of decay machinery: ??? | - | User specified |
| Protein Folding | Not well understood  One step  Boolean representation  Rate = min of (unfold protein, metabolites, chaperones)  Number = multinomial(rate) | - | User specified |
| Protein Modification | 3 modifications: protein phosphorylation, lipoyl transfer and α–glutamate ligation  Not well understood | Model is weird: rate is random with Poisson distribution. Stochasticity seems to be on the rate and not on the actual realization of proteins | User specified |
| Protein Processing I | N-terminal formylmethionine deformylation  N-terminal methionine cleavage  Not well understood | Both in a single step  Weird stochastic | User specified |
| Protein Processing II | Lipoprotein diacylglyceryl adduction  Lipoprotein and secreted protein signal peptide cleavage  Not well understood | Single step  Weird stochastic | User specified |
| Protein Translocation | Only membrane and extra-cellular translocation  Organelle translocation performed in Terminal Organelle Assembly process | Starts only once other modifications were performed | No modeling planed yet!!! |
| Replication | Leading and lagging strands  Polymerization of Okazaki fragments modeled but primer is uncharacterized: fragments length random with Poisson distribution  Rate limited by mean observed rate, availability of nucleotides, energy and proteins  Collision between replication and transcription machineries: replication always wins | One replication per cell cycle  Prevented by too long unwinding of DNA prevents protection of lagging strand or by leading polymerase too far from Okazaki fragments  End of replication triggers decatenation (???) of Chromosome by the Segregation process | Planed |
| Replication Initiation | Binding of DnaA-ATP to binding sites (R1 to R5 near oriC): complex of 7 DnaA-ATP on R1 to R4. 29th DnaA-ATP triggers initiation  Cooperation near oriC | Exact mechanism unknown  Binding sites with 3 affinity levels all over the chromosome but seems only R1 to R4 are taken into account to starts initiation: no effect of other binding sites except to delay it  No second initiation modeled at all | Planed  Replication Initiation has a strong invariant |
| Ribosome Assembly | Rate of ribosomal particle is the min of all participants  Not well known  3 states: rRNA transcript, protein monomer and fully formed ribosomal particle | - | Planed |
| RNA Decay | Poisson distributed for each RNA specy  One step degradation | Includes aborted mRNA but seems that it does not contain rRNA | Planed |
| RNA Modification | Non-coding RNA only  Kinetics not well known | (S28) same as (S20)  What is done with non-coding RNA? | Planed |
| RNA Processing | One step cleavage  Kinetics not well known | (S29) same as (S20) | Planed |
| Terminal Organelle Assembly | If organelle genes expressed, put it in the organelle | Bullshit, does nothing in the simulation except delaying cell cycle | No modeling planed yet!!! |
| Transcription | 4 states: free, non-specifically bounded, bounded to a promoter (only if sigma factor is present) and actively transcribing  Actively transcribing remains in this state: only if displaced by another protein (DNA polymerase; other RNA polymerase?) or stalled | - | Planed |
| Transcriptional regulation | Binding on promoter  Rate multiplier (?) change on polymerase  Not well known  Assumption: kinetically fast and energetically favorable | No regulator unbinding  5 regulators for 54 genes with 29 regulatory interactions | Planed with much more regulations |
| Translation | Initiation  Elongation  Termination | Are IF1 and IF2 really used???  Random somewhere? All possible reactions are performed, stochastic comes into play only in the random choice of the actual reaction that is performed between the different possible elongation steps  No regulation on translation? | Growth management with alarmone |
| tRNA Aminoacylation | Also accounts for tmRNA for stalled ribosome | One step aminoacylation | User specified |
| Stringent Response | - | Growth rate management possible | User specified |
|  |  |  | Need consistency between density, mass, volume and replication  Proof reading? |

Hard coded in several aspects. For instance:

1. it is not possible to perform a second round of replication in one cell cycle. It is thus not possible. Model bacterium cannot be simulated;
2. fundamental processes are hard coded (such as transcription or translation): it is not possible to change the process. Nothing can be done if alarmones were to be simulated.

The rates of reactions stochastic but the number of events is not as the reactions are performed until resources are exhausted.

Translation is not well described.

In our model, all the reactions are assumed to be exponentially distributed at the moment. Need for another modeling?

Need also different level of modeling if we want to aggregate some parts which we don’t care for a particular simulation.