BUILDING A WHOLE CELL IN 3D

Martina Maritan[1], Ludovic Autin[1], Jonathan Karr[2], Markus Covert[3], Arthur Olson[1], David Goodsell[1]
[1]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, San Diego, CA, USA; [2]Department of Genetics & Genomic Sciences, New York, NY, USA; [3]Department of Bioengineering, Stanford University, Stanford, CA, USA

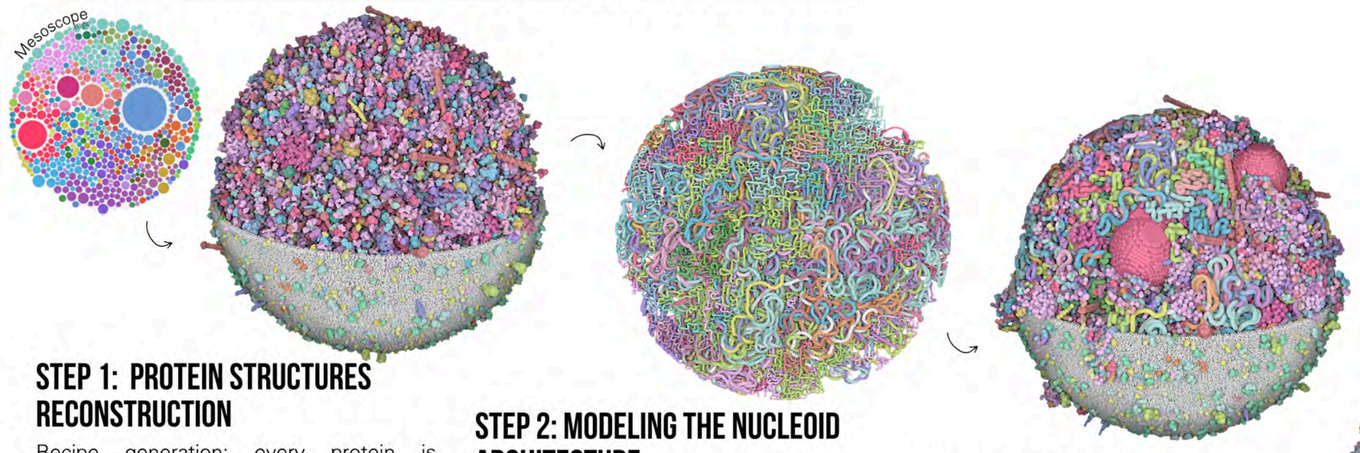
Mycoplasma genitalium (MG) is a bacterium whose life cycle has been computationally modeled and simulated (Karr, 2012). We used whole-cell (WC) simulation data to construct a physical whole-cell model.

↳ Data about protein structures, their oligomerization and membrane orientation come from primary sources of literature, UniProt, HHpred, OPM, PDB and a variety of homology modeling tools

↳ Data about cell size, protein concentrations, protein positions on the DNA, mRNA count, ribosome's location on mRNA were extracted from WC simulations

↳ Data integration of different data sources was performed through Mesoscope

CONSTRUCTING A WHOLE-CELL 3D MODEL: THE PIPELINE



STEP 1: PROTEIN STRUCTURES RECONSTRUCTION

Recipe generation: every protein is assigned a structural representation.

We used a semi-automated protocol to gather and validate structural data for each protein. Structures were retrieved from structural databases or generated via homology modeling. The final recipe includes 482 protein monomers and 201 protein complexes.

The recipe is assembled and curated on online interactive platform Mesoscope (Autin, 2020). Here color palettes and ingredients' concentrations are assigned.

STEP 2: MODELING THE NUCLEOID ARCHITECTURE

Fibrous structures such as nucleic acids are modeled using a lattice based method (Goodsell, 2018).

The model includes 580076nt of circular DNA, 57 DNA associated proteins, transcribing RNA polymerases, growing mRNA, free mRNA and mRNA complexed with ribosomes.

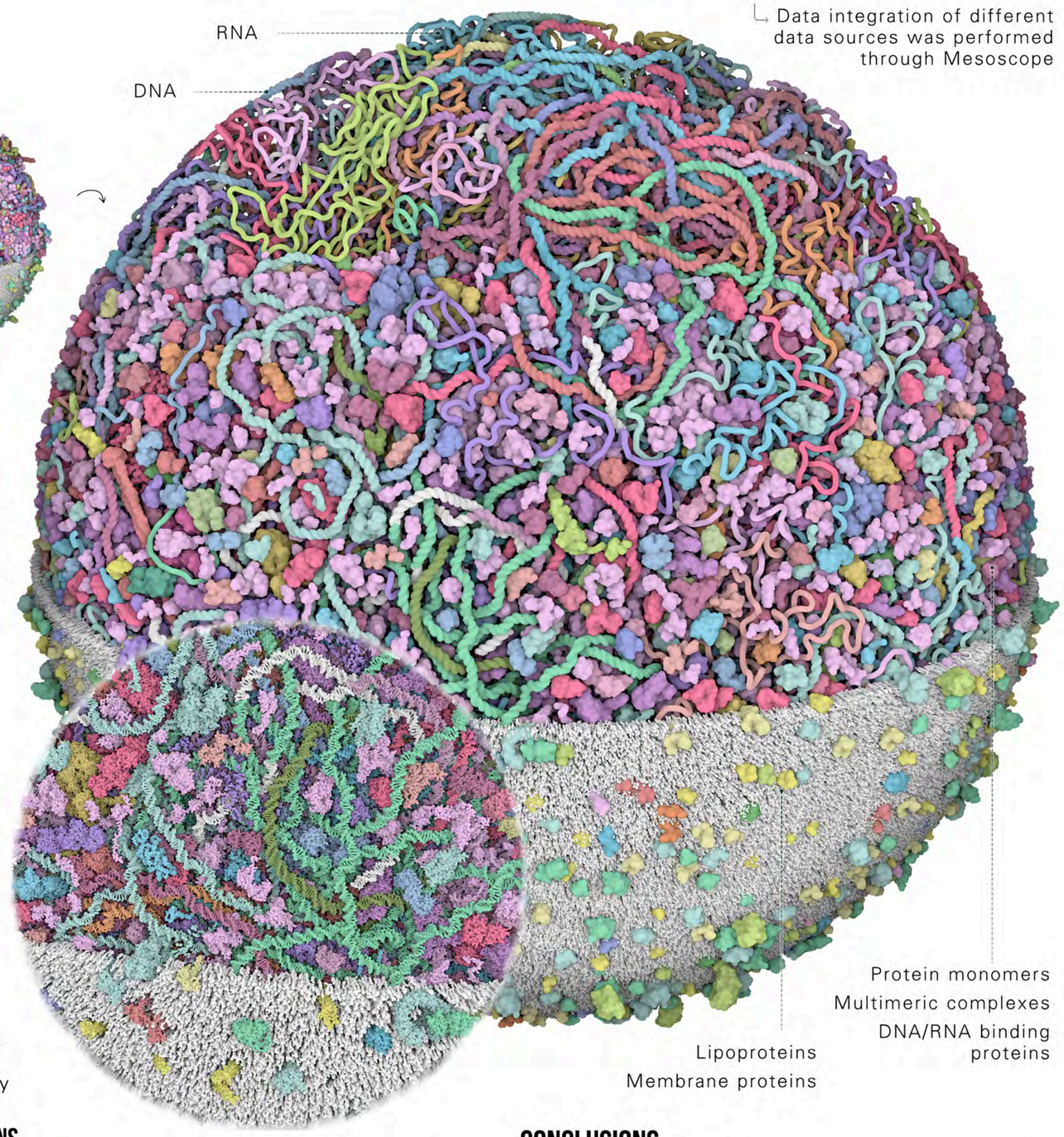
DNA/RNA lengths and the relative positions of binding proteins are based on data from whole-cell simulations.

STEP 3: SIMULATE THE 3D ORGANIZATION OF THE CELL

The recipe and nucleoid model are assembled and visualized in CellPACK.

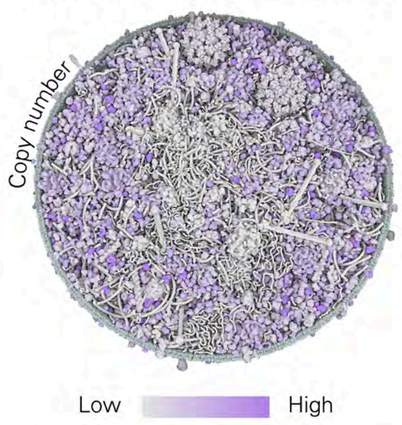
Recipe components are distributed around the nucleoid and relaxed within a defined cell volume using NVIDIA Flex.

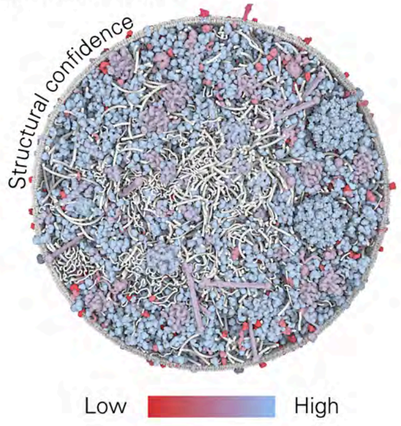
Ingredients are approximated to beads and subjected to relaxation in order to resolve overlaps and clashes.

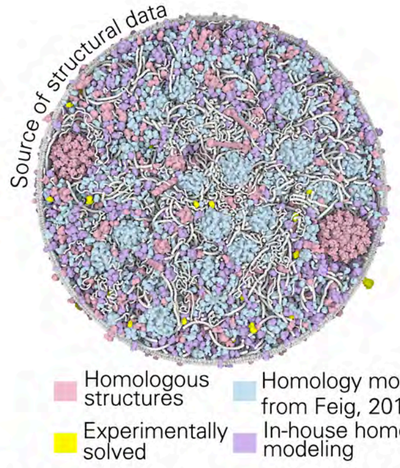


EXPLORING THE CELLULAR MESOSCALE WITH CELLPACK

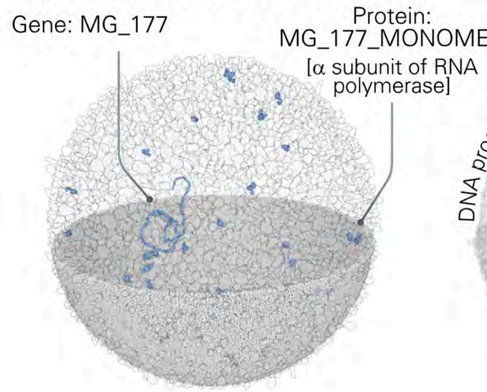
COLORING INGREDIENTS BASED ON USER DEFINED FEATURES



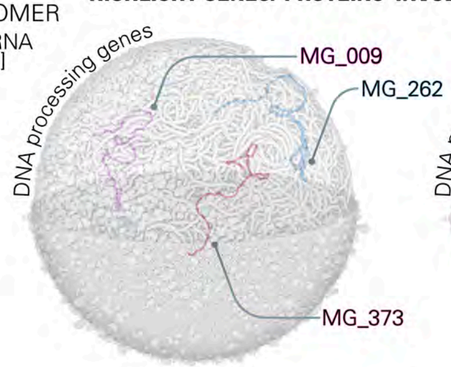


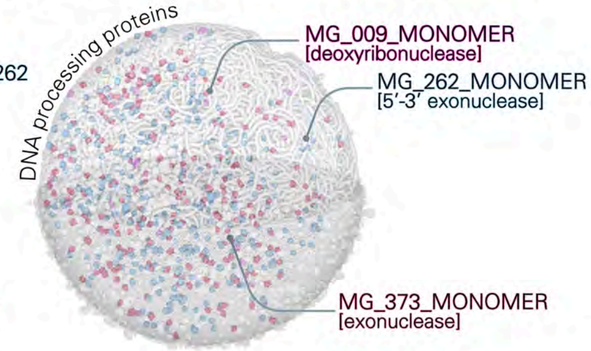


ISOLATE SINGLE GENES/PROTEINS



HIGHLIGHT GENES/PROTEINS INVOLVED IN SPECIFIC BIOLOGICAL FUNCTIONS





CONCLUSIONS

We built the first whole-cell 3D model of a Mycoplasma genitalium (MG) cell at atomic resolution by combining data from whole-cell system biology simulation and structural biology.

Our work integrates whole-cell simulations, structural data, homology modeling, bacterial nucleoid modeling tools, data curation methods, and molecular graphics.

We set up a solid workflow that allows us to generate and navigate 3D molecular-scale snapshots of an entire bacterial cell, providing a structural model for all Mycoplasma proteins and nucleic acids.

References:
Autin, L. Mesoscope: A Web-based Tool for Mesoscale Data Integration and Curation. doi:10.2312/molva.20201098 (2020).
Goodsell, D. S. et al. Lattice Models of Bacterial Nucleoids. The journal of physical chemistry. B 122, 5441-5447, doi:10.1021/acs.jpcb.7b11770 (2018).
Feig, M. et al. Complete atomistic model of a bacterial cytoplasm for integrating physics, biochemistry, and systems biology. Journal of molecular graphics & modelling 58, 1-9, doi:10.1016/j.jmgm.2015.02.004 (2015).
Karr, J. R. et al. A whole-cell computational model predicts phenotype from genotype. Cell 150, 389-401, doi:10.1016/j.cell.2012.05.044 (2012).

Scripps Research