

A Method for Modeling Oxygen Diffusion in an Agent-based Model with Application to Host-Pathogen Infection

Cheryl L. Sershen, Steven J. Plimpton and *Elebeoba E. May- *EMBS Member*

Abstract—This paper describes a method for incorporating a diffusion field modeling oxygen usage and dispersion in a multi-scale model of *Mycobacterium tuberculosis* (Mtb) infection mediated granuloma formation. We implemented this method over a floating-point field to model oxygen dynamics in host tissue during chronic phase response and Mtb persistence. The method avoids the requirement of satisfying the Courant–Friedrichs–Lewy (CFL) condition, which is necessary in implementing the explicit version of the finite-difference method, but imposes an impractical bound on the time step. Instead, diffusion is modeled by a matrix-based, steady state approximate solution to the diffusion equation. Presented in figure 1 is the evolution of the diffusion profiles of a containment granuloma over time.

I. INTRODUCTION

Multi-scale modeling of host-pathogen interactions has recently become a necessity since disease outcomes depend upon events occurring at many different biological scales. For example, in the extracellular environment within the body, immune cells and pathogens move freely, with their interactions determining disease outcomes. In the intracellular environment individual cells (both immune cells and invaders) enact their responses to extracellular events within their relative microenvironment. To a large extent the intracellular response is determined by metabolic pathways and protein interactions initiated at the genomic scale through gene regulatory networks. To truly portray the dynamics of disease *in silico*, so as to gain insight into disease dynamics not easily identified by either *in vitro* or *in vivo* studies, it is necessary to investigate and model dynamics

Research supported by DTRA/HJF subaward 2381, NIH/NHLBI award K25HL075105, Sandia National Laboratories' LDRD program. Sandia is a multi-program laboratory operated by Sandia Corporation, a Lockheed Martin Company for the US DoE's NNSA under contract DE-AC04-94AL85000.

Cheryl L. Sershen is a postdoctoral research scientist in the Department of Biomedical Engineering at the University of Houston, Houston TX 77204-5060 USA (e-mail: clershe@central.uh.edu).

Steven J. Plimpton is a Distinguished Member of Technical Staff at Sandia National Laboratories, Albuquerque, NM 87185 USA (e-mail: sjplimp@sandia.gov).

*Elebeoba E. May is an Associate Professor in the Department of Biomedical Engineering at the University of Houston, Houston TX 77204-5060 USA. Corresponding author (phone: 713-743-3851; fax: 713-743-0226; e-mail: eemay@central.uh.edu).

occurring at many scales of the biological system. Accordingly, we developed a multi-scale model of *Mycobacterium tuberculosis* infection granuloma formation with an agent-based model (ABM) to portray the extracellular environment coupled with an intracellular model for bacteria using a systems biological approach. This paper describes a method for simulating an oxygen field to allow measurement of the degree of hypoxia present in a given granuloma, using a steady-state solution that avoids the bound set on the time step by a fast diffusion coefficient.

II. AGENT-BASED MODEL (ABM)

To model the extracellular environment and the interactions occurring therein as an element in the multi-scale model (intracellular dynamics are modeled separately from the ABM), we developed proprietary software for implementing an agent-based model. Our software platform uses a Python based framework to encode rules flexibly, and C++ kernels and Octave libraries for solving partial differential equations quickly. We developed our core ABM in the manner of [1] and [2], then extended this framework to accommodate oxygen supply and consumption in the ABM. To define our ABM, we specify a regular grid of agents: bacteria (represented as a floating-point field), macrophages and T cells (represented as particle-objects), and set the size of each grid cell large enough to accommodate the largest cell type (the macrophage at 20 μm in diameter [1]). For a 2mm by 2mm tissue sample this gives a grid size of 100 X 100 cells. The simulations discussed here are 2-dimensional; the simulator also runs 3-dimensional models that are computationally more expensive.

Each agent is characterized by its state (e.g. resting or infected). For each cell under consideration, a neighborhood is defined relative to the specified cell (e.g. Moore neighborhood). We specify the initial system by assigning a state for each agent resident on the grid. We advance time by one time step according to the specified rules of diffusive movement (a qualitative change or mathematical function) and determine the new state of each agent in terms of the current state of the agent, composition of the agent's grid cell, and the states of the agents and composition of grid cells in its neighborhood. Rules may be

probabilistic or deterministic (e.g. a macrophage becomes activated with probability p when a gamma delta T cell is in its neighborhood).

Variables that represent continuous concentrations (such as chemotactic species) are modeled as a floating-point field on the grid. Chemotactic equations model the increase, degradation and spread of chemotactic species. Our ABMs profile the behavior of cells in the lung during irritation and infection, taking into account chemokine and cytokine levels and the amount of tumor necrosis factor alpha (TNF- α) released. These are all essential participants in capturing the emergent behavior of lung tissue cells during lung disruption or irritation.

III. MATERIALS AND METHODS

To capture oxygen dynamics we added an extra floating-point field that accounts for extracellular oxygen concentration to the ABM model. The biophysical parameters of oxygen uptake and absorption in the human lung were modeled by calculating the oxygen consumption (in molecules) in the various cell types over the $2mm \times 2mm$ grid. The number of oxygen molecules available in each grid cell was calculated by performing mass and atomic balances and using the ideal gas law:

$$PV = nRT \quad (1)$$

In the model this is done by considering:

- oxygen in and through the lung parenchyma from three sources: residual volume in the alveoli, inhaled air, and pulmonary blood volume. Model parameters correspond to a normal, non-exercising male, and respiratory ailments can be simulated by changing the core oxygen parameters, such as a slower oxygen diffusion coefficient and/or by adjusting the number of breaths per minute.
- Consumption of oxygen is by macrophages and bacteria. Oxygen consumption by the bacteria is based on data from *E. coli* [3]. Infected macrophages consume more oxygen than resting macrophages [4].

Diffusion of chemokine, cytokine, TNF- α , and oxygen, is formulated and solved as a two-dimensional, second order, parabolic diffusion (heat) equation:

$$u_t = D (u_{xx} + u_{yy}) + S \quad (2)$$

where u_t is the first order time partial derivative, u_{xx} and u_{yy} are the second order spatial partial derivatives with respect to x and y respectively (for the two-dimensional problem), S is the sink term on the grid (to accommodate consumers of oxygen), D is the

diffusion coefficient for the substance diffusing, with Dirichlet boundary conditions and initial condition $u(x, y, 0) = b$. Oxygen is delivered from the boundary cells throughout the tissue. For the oxygen field, $u(x, y, t)$ is the concentration of oxygen at the grid point (x, y) at time t .

Vascular source cells allow new macrophages and T cells to be introduced to the grid. In this version of our ABM, the oxygen consumption of the T cells is assumed to be negligible, since they are about half the size of the larger macrophage. The resulting ABM has many input parameters, for example the speed of various cells, probabilities, and qualitative and quantitative cell characterizations.

Oxygen diffuses into the current grid cell from its neighboring cells or out of the current grid cell to the neighboring cells. It can be shown via integration [5] that equation (2) is equivalent to solving the discrete system:

$$u(t+1) = u(1 - \lambda) + \lambda (u_N + u_S + u_E + u_W) / 4 + S \quad (3)$$

where u is $u(t)$, $u_N + u_S + u_E + u_W$ are the contributions from the adjacent grid cells at time t and S is the source/sink term. $\lambda = 4 * D * \Delta t / (\Delta x)^2$ (Courant–Friedrichs–Lewy (CFL) condition with $\Delta x = \Delta y$ [6, 7]) in the finite difference discretization and measures the proportion of oxygen diffusing into the cell at time t [2]. For the solution to be stable, it must be the case that $0 \leq \lambda < 1$. Thus the tradeoff between step size and grid size is linked to the diffusion constant in guaranteeing the stability of the explicit finite difference discretization method [5]. With the grid sizes employed in the ABM model, the required time step for oxygen diffusion ($D \approx 3.08e-5 \text{ cm}^2/\text{sec}$ [8]) would be at most 0.032 seconds. This renders the model computationally impractical since granuloma formation and dissemination requires simulations for 200-500 days of real time. By comparison, for models without oxygen, the chemokine diffusion rate is the limiting factor ($1.67e-08 \text{ cm}^2/\text{sec} \leq D \leq 1.167e-07 \text{ cm}^2/\text{sec}$) that gives a time step of 6 seconds ($0.1 < \lambda < 0.7$) so that explicit, time-dependent finite difference methods can be used to update the ABM [1].

For oxygen diffusion a steady state, matrix-based approach was adapted to solve equation (2). One assumption is that the lung has time to relax to equilibrium between field calculations and this is an appropriate assumption since substantial changes in the formation of the granuloma occur over relatively longer period of time. ABM solves the steady-state oxygen diffusion equation at intervals of 18 hours, as opposed to the finite-difference equations for chemokine, cytokine and TNF- α , which are updated every time step.

At steady state, $u_t = 0$, so it remains to discretize:

$$D(u_{xx} + u_{yy}) + S = 0. \quad (4)$$

Chemokine-biased movement motivates the dynamics of granuloma formation and once a macrophage joins the granuloma, it generally stays in that region, until bursting or death, due to the biased movement in response to chemokine secretion. Oxygen diffusion is formulated and solved as the two-dimensional, matrix equation [9]:

$$D^*Au = f \quad (5)$$

where A is the finite-difference, two-dimensional (5-point), block-tri-diagonal sparse Laplacian matrix, u is the lexicographically by y -lines ordered vector of grid cells for which we wish to compute the oxygen solution and f is the lexicographically ordered by y -lines vector of sources/sinks (oxygen input at the boundary and oxygen consumption by agents) at time t . Consumers of oxygen (sinks) are macrophages and bacteria.

For the ABM 100 X 100 problem, the matrix A is sparse, square, block tridiagonal, symmetric positive definite and has full rank with non-zero (extremely large) determinant, which requires $\sim O(n^3)$ operations to invert using Gauss-Jordan elimination and is subject to computer round-off error in the computing the inverse of A . Instead, equation (5) is solved much faster, without inverting A , in Octave using backslash, i.e. $v = [D^*A] \setminus f$ [10] via iterative conjugate gradient methods [11]. This solution approach was appropriate in considering the diffusion of oxygen in the ABM. It permits computation of the approximate solution to the diffusion equation, since the error in A is order $(\Delta x)^2$ based on the Taylor series derivation of the finite difference approximation of A , see for example [9], with $(\Delta x)^2 \sim e-08 \text{ cm}^2$ for the ABM problem. This implementation did not contribute significantly to the total time required in the ABM simulation, since the field equations are solved only every 18 hours.

IV. RESULTS

Figure 1 shows the results of our matrix computation on a containment granuloma at 25, 43 and 106 days on a grid size of $2\text{mm} \times 2\text{mm}$ for the granuloma and the corresponding oxygen field. As the granuloma forms in the tightly packed region of cells at the center of the grid, the oxygen concentration in the granuloma region drops steadily in response to the consumption by agents. The change in oxygen concentration impacts the effectiveness of the immune response of the macrophage and negatively affects the metabolic fitness of Mtb. Oxygen dynamics thus play a role in regulating the behavior of the macrophages and

bacteria, regulation that is modeled via intracellular models, which take into account the effect of oxygen on gene expression.

V. CONCLUSION

The ABM was developed to simulate Mtb-mediated granuloma formation and the oxygen field was added to allow measurement of the degree of hypoxia present in a given granuloma as well as the role of oxygen in regulating cellular response to varying levels of oxygen. We have described a method for avoiding instabilities in explicit finite difference schemes used in ABMs, induced by fast diffusion coefficients. We have addressed this problem using a steady state, matrix-based approach to derive an approximate solution for the oxygen field. The method can be executed quickly for the 100 X 100 grid of cells used in these models, and our tests show it is viable for larger grid sizes as well as three dimensions, and increased number of diffusing species. Our ABM is a central component in our multi-scale model of Mtb-induced granuloma formation and is linked through the oxygen field to other model components to characterize intracellular dynamics of host-Mtb interaction.

VI. FUTURE WORK

In addition to the two methods implemented here in the ABM, time-dependent finite-difference and steady-state matrix-based solution, we will give the user a choice of methods for solving a PDE over a field, that avoids the problems mentioned in Materials and Methods. We may wish to make the diffusion coefficient spatially dependent for certain diseases (ex.: infected or caseous tissues). We are also extending matrix-based methods to 3D models. These require much larger Laplacian matrices, e.g. for a small scale 3D problem such as that presented here, the Laplacian matrix size is in excess of $10^5 \times 10^5$ (memory resources of $\sim 80 \text{ GB}$). As part of that effort we plan to replace the Octave component of our software stack and implement solvers such as those available in Scipy or LAPACK via algorithms for sparse, block-tridiagonal matrices for increased speed of computation, which take advantage of the fact that A is sparse and require $\sim O(n)$ operations [12]. We are pursuing options for high performance cluster and supercomputing resources to overcome the large memory requirement for solving matrix-based problems in 3D.

We will use the freely available software DAKOTA written at Sandia National Laboratory to perform both uncertainty and sensitivity analysis to identify the significant oxygen parameter drivers of the model outcomes. DAKOTA enables Latin-hypercube sampling by providing a wrapper for the multi-scale model, which it sees as a black box: a sink for inputs (parameters) and a source for outputs (responses, i.e. model outcomes such as clearance, containment and dissemination). Among other model evaluation

measures, DAKOTA returns matrices of partial rank correlation coefficients that can be used to identify the significant drivers of the model under consideration.

Finally, we will identify via optimization techniques, optimal oxygen parameters supporting, for example, a clearance or containment outcome (e.g. optimal number of breaths per minute, optimal diffusion coefficient, etc.) by using the significant drivers identified in the sensitivity phase to build an objective function that summarizes the multi-scale model outcomes. Identifying these parameters may be of substantial value in determining prognosis and treatment of infection with Mtb.

ACKNOWLEDGMENT

The authors would like to thank Dr. Alfred O. Ong'iro for providing review and commentary for this paper.

REFERENCES

[1] J. Ray, J. Flynn, D. Kirschner, "Synergy between individual TNF-dependent function determines granuloma performance for controlling *Mycobacterium tuberculosis* infection," *Journal of Immunology*, 182, 2009.

[2] J. Segovia-Juarez, S. Ganguli, D. Kirschner, "Identifying control mechanisms of granuloma formation during m. tuberculosis infection using an agent-based model," *Journal of Theoretical Biology*, 231, 2004.

[3] M. Grieg, J. Hoogerheide, "The correlation of bacterial growth with oxygen consumption," *Journal of Bacteriology*, 41, 1941.

[4] L. Loose, "Characterization of macrophage dysfunction in rodent malaria," *Journal of Leukocyte Biology*, 36, 1984.

[5] S. Patankar, *Numerical Heat Transfer and Fluid Flow*. Hemisphere Publishing Corporation, 1980.

[6] G.B. Ermentrout, L. Edelstein-Keshet, "Cellular automata approaches to biological modeling," *J. Theoretical. Biology*, 160 (1), 97-133, 1993.

[7] C. Moler, *Numerical Computing with Matlab*. SIAM 2004.

[8] C. Hou, S. Gheorghiu, V. Huxley, P. Pfeifer, "Reverse engineering of oxygen transport in the lung: Adaptation to changing demands and resources through space-filling networks," *PLOS Computational Biology*, 6, 2010.

[9] W. Briggs, V. Henson, S. McCormick, *A multigrid tutorial*. SIAM 2000.

[10] Octave reference. <https://www.gnu.org/software/octave/doc/interpreter/Simple-Examples.html>, accessed 3-21-2014.

[11] L. Calvet, "Numerical Methods Library for Octave", www.ipb.pt/~balsa/teaching/MA08_09/UsersGuide.pdf, accessed 4-2-2014.

[12] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes: The Art of Scientific Computing* (3rd ed.). Cambridge University Press, 2007.

Figure 1: Containment granuloma (without TNF) at 25, 43 and 106 days and the corresponding steady state oxygen diffusion field

