Journal of Theoretical Biology 290 (2011) 60–72

Contents lists available at SciVerse ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Data-driven modeling of Alzheimer Disease pathogenesis

Thomas J. Anastasio *

Department of Molecular and Integrative Physiology, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Beckman Institute, 405 North Mathews Avenue, Urbana, IL 61801, USA

ARTICLE INFO

Article history:
Received 22 March 2011
Received in revised form
27 August 2011
Accepted 29 August 2011
Available online 5 September 2011

Keywords:
Neurological disease
Biological regulation
Term rewriting
Model checking
Complex systems

ABSTRACT

Alzheimer Disease (AD) is the most prevalent form of dementia and the sixth leading cause of death in developed world. A substantial amount of data concerning the pathogenesis of this neurological disorder is available, but the complexity of the interactions they reveal makes it difficult to reason about them. This paper describes a computational model that represents known facts concerning AD pathophysiology and demonstrates the implications of those facts in the aggregate. The computational model is written in a mathematical language known as Maude. Because a Maude specification is an executable mathematical theory, it can be used not only to simulate but also to logically analyze the system it models. This model is based on the amyloid hypothesis, which posits that AD results from the build-up of the peptide beta-amyloid. The AD model represents beta-amyloid regulation, and shows through model analysis how that regulation can be disrupted through the interaction of pathological processes such as cerebrovascular insufficiency, inflammation, and oxidative stress. The model demonstrates many other effects that depend in complex ways on interactions between elements. It also shows how treatments directed at multiple targets could be more effective at reducing beta-amyloid than single-target therapies, and it makes several experimentally testable predictions. The work demonstrates that modeling AD as an executable mathematical theory using a specification language such as Maude is a viable adjunct to experiment, which allows insights and predictions to be derived that take more of the relevant biology into account than would be possible without the aid of the computational model.

1. Background

Back in the heyday of complex-systems studies, a great deal of attention was devoted to simple systems that exhibit complex behavior, because it was obvious that complex systems exhibit complex behavior. Nowadays more effort is devoted to understanding complex systems, which can be loosely defined as systems composed of many different parts that interact in many different ways. Biological systems are complex systems of this type, and the goal of systems-biological modeling is to develop methods to represent, simulate, and analyze them. Systems-biological modeling approaches are critical for understanding highly complex disease processes such as neurological disorders. The purpose of this paper is to describe a data-driven approach to modeling the pathophysiology of Alzheimer Disease (AD). As a complex disorder that is also the sixth leading cause of death in the developed world, AD is an apt subject for complex systems modeling.

A concept-driven (as opposed to a data-driven) model implements an assumed mechanism (e.g. amplifier, filter, servomechanism, multi-stable attractor, etc.) and demonstrates how the properties of that mechanism provide insight into a biological or other natural phenomenon. In contrast, a data-driven model represents a set of interrelated experimental findings and demonstrates the implications of those findings in the aggregate. In a data-driven model, the data are represented in terms of declarations that specify how interactions between variables should change the values of the variables. The declarations can be as quantitative as available data allow but, in the absence of precise measurements, the values of variables can change between, for example, multiple, arbitrarily assigned integer-valued levels. Assuming that the representation of the data is valid, a data-driven model can be used to analyze the properties of a naturally occurring system and to generate experimentally testable predictions, the results of which can be used to correct or extend the model.

Some of the first data-driven models of biological systems took the form of Boolean-logic networks (Thomas and D'Ari, 1990), and that approach is undergoing continued development (Klamt et al., 2006, 2007). Other approaches to data-driven modeling of biological systems include Petri nets, various process calculi, and hybrid models; those of the last type are composed of differential (or difference) equations, to simulate continuous changes in variables, and rules, to simulate switch-like changes in parameters (for reviews see Hlavacek et al., 2006; Fisher and Henzinger, 2007).

* Tel.: +1 217 244 2895; fax: +1 217 244 5180.
E-mail address: tja@illinois.edu

0022-5193/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2011.08.038

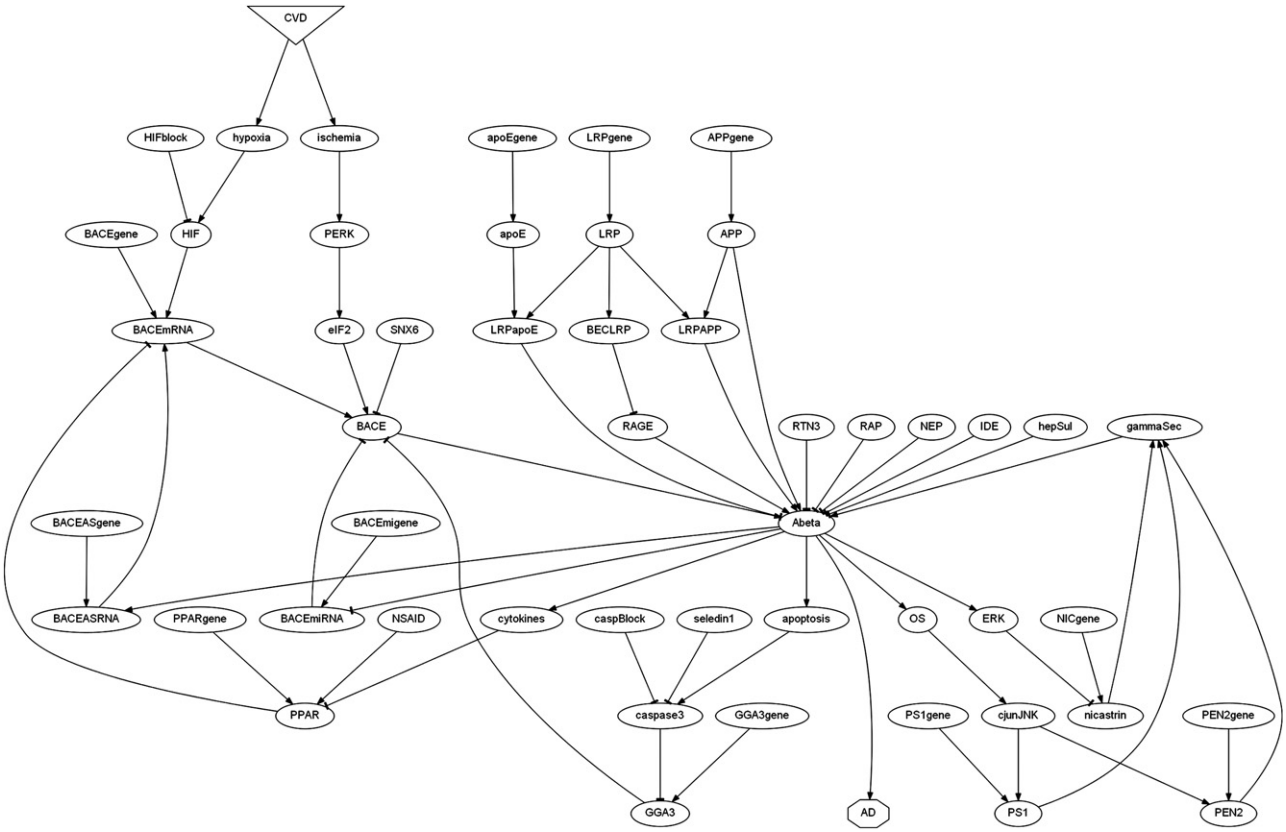

Fig. 1. Schematic of the data-driven model of Alzheimer Disease. The diagram depicts the influences that each element has on the other elements in the model. Arrowheads and tees represent positive and negative influences, respectively. The diagram was drawn using Graphviz software. The list of abbreviations is provided in Table 1.

The approach taken in this paper is most closely related to the Pathway Logic approach pioneered by Carolyn Talcott (Eker et al., 2002; Talcott, 2008). In Pathway Logic, molecular interactions are specified using a custom-designed graphical user interface to the Maude environment. The resulting Maude specifications are converted to Petri net models, which are then executed and analyzed using specialized Petri net tools.

The approach taken here is to represent interactions between molecules and conditions directly as declarations in Maude programs, and to use those programs as such in the Maude environment for simulation and analysis purposes. Maude is a mathematical modeling language in which systems can be modeled as sets of interrelated equations and rules. A Maude model constitutes an executable mathematical theory of the system it models (Clavel et al., 2007). Thus, a Maude model can be run as a computer program, to simulate the system it models, but also analyzed as a mathematical theory, to explore the properties of the system on a more general level. Because of its dual use potential, as a computer program and as a mathematical theory, a Maude model is known as a specification. Here Maude is used to specify a data-driven model of some of the interactions between molecules and conditions that are thought to underlie the complex etiology of Alzheimer Disease (AD). This Maude specification is used both for simulation purposes and as an object of analysis through state-space search and logical model checking.

Most of the data on AD biology is qualitative. To manage it, the molecules and conditions represented in the model are assigned arbitrary integer levels, and the equations and rules specify how changes in the levels of some model elements should change the levels of other elements. The model presented in this paper represents a small subset of the data on AD pathophysiology, but it is large enough that the nature of the interactions it models could not be discerned by inspection of the model diagram alone (Fig. 1). A longstanding goal of complex-systems studies has been to find general concepts that apply to whole classes of phenomena (Mainzer, 2007). In this spirit, we will show through simulation and analysis how the AD model reveals certain general properties that are likely to occur in real neurodegenerative disorders. However, the main focus here is on the details of AD etiology. Because all the variables are experimentally observable, simulation and analysis of the model automatically yields specific and testable hypotheses concerning the pathophysiology of AD. The results of the tests would have therapeutic value.

2. Methods

2.1. Biological background for the model

The model of AD pathophysiology is diagrammed in Fig. 1. The model is based on findings as reported in the primary literature. It is also, necessarily, based on a number of assumptions. The main assumption of the model is that AD results from build-up in the brain of the peptide beta-amyloid (or amyloid-beta, abbreviated Abeta). Although there are several other hypotheses concerning the cause of AD (see Section 4.3), the "amyloid hypothesis" is the leading hypothesis (Hardy and Selkoe, 2002). Efforts at AD therapeutics are directed both toward prevention and cure (Hardy and Selkoe, 2002). This model hopes to contribute to the former effort by focusing on the transition from the normal, healthy condition to the pathological state, rather than the state of advanced AD pathology in which brain levels of Abeta have been chronically elevated.

The Abeta peptide occurs in various lengths, but the 40 and 42 amino acid lengths are the most common (Selkoe, 2001). Although