MATHEMATICAL MODELLING OF BIOSYSTEMS

Applied Optimization

VOLUME 102

Series Editors:

Panos M. Pardalos
University of Florida, USA

Donald W. Hearn
University of Florida, USA

MATHEMATICAL MODELLING OF BIOSYSTEMS

Edited by

RUBEM P. MONDAINI
Federal University of Rio de Janeiro, Brazil

PANOS M. PARDALOS
University of Florida, Florida, USA

Springer

Prof. Rubem P. Mondaini
Federal University of Rio de Janeiro
UFRJ – COPPE – Centre of Technology
21.941-972 – P.O. Box 68.511
Rio de Janeiro – RJ
Brazil
rpmondaini@gmail.com

Prof. Panos M. Pardalos
University of Florida
Center for Applied Optimization
303 Weil Hall
P.O. Box 116595
Gainesville, FL, 32611-6595
USA
pardalos@ufl.edu

ISBN 978-3-540-76783-1 e-ISBN 978-3-540-76784-8

DOI 10.1007/978-3-540-76784-8

Applied Optimization ISSN 1384-6485

Library of Congress Control Number: 2008921240

AMS codes: 34-XX, 42-XX, 49-XX, 92-XX, 53-XX, 35-XX, 52-XX, 68-XX

Production: le-tex Jelonek, Schmidt & Vöckler GbR
Cover design: WMXDesign GmbH, Heidelberg

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Preface

The first idea for organizing this book was to collect some state of art contributions to the literature on mathematical modelling of biosystems written by representatives of research groups in the Americas. We have also invited a contribution from the Russian Federation written by A. Finkelstein. The importance of an interdisciplinary approach to this field of knowledge or of a biologically-inspired treatment of mathematical structures inherent to Nature, derives from the best insights professed in the last century by researchers like D'arcy Thompson, Rashevsky, Schroedinger, Ulam and Feynman. This already traditional avenue of Science in Latin America is now being followed by many multidisciplinary conferences all over the world and we think that in order to enhance the participation of young scientists on them, a book written by invited experts with many years of engagement on these interdisciplinary scientific activities was strictly necessary. We have chosen nine main themes to be addressed by these scientists as chapters of the present book. Each of them is aimed to correspond to a fresh start of the study of the selected theme at the level of first-year graduate students.

The first chapter by A. Goriely and collaborators of his research group from University of Arizona at Tucson, USA, emphasizes the mechanism of biological growth. The developments are based on insights aimed to generalize the classical theory of exact elasticity from observations of changes of geometry due to the dynamics of mechanical quantities during growth. Of particular importance is the unified treatment of geometry and mechanics and the deep scientific relevance of the theme with the recent discovery of biological systems such as the transition helix-spiral ribbon which can provide analytical solutions to the Föppl-von Karman equations.

The second chapter was written by F. W. Cummings from University of California Riverside, USA, as part of a successful program for studying the coupling of pattern to form in Biosystems. This is a research topic which has its motivation in the early work of D'arcy Thompson, one of the fathers of Mathematical Biology. The chapter contains some interesting reports on genetic regulatory domains closely coupled to the elements of the patterning

model. There is also the description of a model of patterning which couples to morphogenesis with an emphasis on Fibonacci patterns. It follows an application of the Gauss-Bonnet theorem for studying the geometry of epithelial tissues. A very interesting relationship between geometry and pattern is discussed by stressing the reciprocity of the effects of geometry and pattern on each other.

The contribution of A. Perelson and members of his research group at the Theoretical Division of Los Alamos National Laboratory, New Mexico, USA, is the content of the third chapter. This chapter deals with human immunodeficiency virus dynamics. The authors discuss some models in which parameters such as viral clearance rate, the death rate of active infected cells and viral load are related to emergence of drug resistance, the type of antiretroviral agent combination regimen, the level of patient adherence of treatment and drug efficacy. Due to polymorphism of these retroviruses, associated with a high rate of occurrence of mutant genomes in each offspring, dual strain modelling is studied. The viral and host factors relations are discussed regarding to therapeutic success.

The fourth chapter was written by C. Floudas from Princeton University, USA, and his collaborators. This is a nice review on the De Novo protein design problem or "inverse folding problem". The De Novo protein design framework has two stages: the amino acid sequence selection stage and the Fold Specifity or Fold validation stage. The authors have also reported on some improvements of the method related to the speeding up of the sequence selection stage as well as the fold validation stage. Another improvement is the possibility of true protein backbone flexibility through the introduction of models which map amino acid sequences onto a flexible and multi-structured template.

The fifth chapter introduces the development of a Heuristic for consistent Biclustering problems as developed by P. Pardalos and his research group from University of Florida, Gainesville, USA. After an objective characterization of the problem, an emphasis is also put on the possibility of constructing a consistent biclustering by deleting some features and/or samples from the given data set. The method is useful in the organization of microarray data for diagnostic of diseases in patients. These techniques allow for the introduction of well-posed optimization problems. Actually the approach used here is the iterative solution of continuous linear problems. The algorithm developed in this contribution was tested on Human Gene Expression Index.

The sixth chapter is the contribution of R. Mondaini and his research group from Federal University of Rio de Janeiro, Brazil to this interdisciplinary book. The chapter is a report on the fundamental problems of construction of an analytical formulation for an unconstrained optimization problem aimed at deriving the lowest upper value of the Steiner Ratio in $\mathbb{R}^3$ with Euclidean distance. These results are related with the modelling of biomolecular structures and are specially adequate to model the tertiary structure of proteins. A proof of the existence of a lower bound of the Steiner Ratio by Weierstrass theorem

for a Global minimum as well as a numerical implementation of an algorithm without derivatives will close the chapter with some perspective to continue this research field in which the author has worked for the last ten years.

The seventh chapter was written by J. Velasco-Hernández and B. Tapia-Santos from Mexico. This contributed chapter has some theorems dealing with mutated and phenotypic bacteria strains, in which resistance mechanisms (such as mutation and/or phenotypic switching) are elicited when exposed to certain biocides, considering their responses to the inhibitory challenge. Competition modelling in a chemostat is used to discuss these issues, considering nutrient intake and antibiotic sensitivity. The phenotypic switching resistance seems to be related to inhibitor concentration, not seen in mutation induced resistance.

The eighth chapter was written by J. Harte, from University of California, Berkeley, USA. A very nice and comprehensive review on the distribution and abundance of species. It has been written with the aim of stressing the problems of pattern formation and the description of process. A deep motivation is given to young research students for working the fundamentals of a future unified theory of ecology. A useful warning is also given by providing a comparison with the theoretical efforts accomplished until now to create a unified theory of particles and fields including gravity. Harte is specially acquainted with these developments due to his past expertise as a theoretical physicist working in particle physics.

The final chapter is the contribution of A. Finkelstein and his collaborators at the Laboratory of Protein Physics and Institute of Protein Research of the Russian Academy of Sciences, Russian Federation. His large expertise on protein physics has provided an excellent review of this subject which can be used now for an introduction to the theme by any dedicated research student. The general ideas are also an adequate introduction for undergraduates to decide to follow this research field in their future graduate studies. Special emphasis is put on to explain the "all-or-none" first order phase transitions in proteins from initial (native) states to the final (denatured) states. It follows also a discussion of the interesting Levinthal paradox with its solution and of protein folding in general.

We are indebted to Débora Mondaini, the elder daughter of the first named editor of this book. She has made a digression from her Ph.D. work to collaborate with the necessary editorial work to build up the book. We are also glad to Dr. Eduardo P. Marques, former Ph.D. student of the Optimization/Biomathematics group of COPPE/CT/UFRJ for reading the manuscript.

We sincerely hope that this volume will be helpful to all those researchers with a real interest on interdisciplinary themes of biological interest. The contributions collected here as chapters are an example of that kind of work aimed at developing the interdisciplinary research field of mathematical and computational modelling of Biosystems.

Rio de Janeiro,
December 2007

Rubem P. Mondaini
Panos Pardalos

Contents

List of Contributors

Stanislav Busygin
University of Florida/Center for Applied Optimization
Gainesville, FL 32611, USA
busygin@ufl.edu

Frederick W. Cummings
University of California Riverside/Department of Physics and Astronomy
San Anselmo, CA 94960, USA
fredcmgs@berkeley.edu

Zhilan Feng
Purdue University/Department of Mathematics
West Lafayette, IN 47907, USA
zfeng@math.purdue.edu

Alexei V. Finkelstein
Russian Academy of Sciences/Institute of Protein Research
4 Institutskaya str., Pushchino, Moscow Region, 142290, Russian Federation
alexey@finkelstein.ru

Christodoulos A. Floudas
Princeton University/Department of Chemical Engineering
Princeton, NJ 08544-5263, USA
floudas@titan.princeton.edu

Ho Ki Fung
Princeton University/Department of Chemical Engineering
Princeton, NJ 08544-5263, USA

Oxana V. Galzitskaya
Russian Academy of Sciences/Institute of Protein Research
4 Institutskaya str., Pushchino, Moscow Region, 142290, Russian Federation

Sergiy O. Garbuzynskiy
Russian Academy of Sciences/Institute of Protein Research
4 Institutskaya str., Pushchino, Moscow Region, 142290, Russian Federation

Alain Goriely
University of Arizona/Program in Applied Mathematics, RUMMBA
Tucson AZ 85721, USA
goriely@math.arizona.edu

John Harte
University of California/Energy and Resources Group
Berkeley, CA 94720 USA
jharte@berkeley.edu

Dmitry N. Ivankov
Russian Academy of Sciences/Institute of Protein Research
4 Institutskaya str., Pushchino, Moscow Region, 142290, Russian Federation

Rubem P. Mondaini
Federal University of Rio de Janeiro/Centre of Technology, COPPE, 21941-972, Rio de Janeiro, RJ, P.O. Box 68511, Brazil
rpmondaini@gmail.com

Dimitrios Morikis
University of California/Department of Bioengineering
Riverside, CA 92521, USA

Artyom Nahapetyan
University of Florida/Center for Applied Optimization
Gainesville, FL 32611, USA
artyom@ufl.edu

Panos Pardalos
University of Florida/Center for Applied Optimization
Gainesville, FL 32611, USA
pardalos@ufl.edu

Alan S. Perelson
Los Alamos National Laboratory/Theoretical Biology and Biophysics, Theoretical Division
Los Alamos, NM 87545, USA
asp@lanl.gov

Mark Robertson-Tessi
University of Arizona/Program in Applied Mathematics, RUMMBA
Tucson AZ 85721, USA

Libin Rong
Purdue University/Department of Mathematics
West Lafayette, IN 47907, USA
rong@math.purdue.edu

Michael Tabor
University of Arizona/Program in Applied Mathematics, RUMMBA
Tucson AZ 85721, USA

Brenda Tapia-Santos
Universidad Veracruzana/Facultad de Matemáticas
Xalapa, Ver., A.P. 270, C.P. 91090 Mexico
bretasa@gmail.com

Martin S. Taylor
Johns Hopkins University/School of Medicine
Baltimore, MD 21205, USA

Rebecca Vandiver
University of Arizona/Program in Applied Mathematics, RUMMBA
Tucson AZ 85721, USA

Jorge X. Velasco-Hernández
Instituto Mexicano del Petróleo/Programa de Matemáticas Aplicadas y Computación
México, D.F. 07730, Mexico
velascoj@imp.mx

Li Zhang
University of California/Department of Chemistry
Riverside, CA 92521, USA

Elastic Growth Models

Alain Goriely, Mark Robertson-Tessi, Michael Tabor, Rebecca Vandiver

Program in Applied Mathematics, RUMMBA (Research Unit in Mathematics, Mechanics, Biology, and Applications), University of Arizona, Tucson AZ85721
goriely@math.arizona.edu

Summary. Growth is involved in many fundamental biological processes such as morphogenesis, physiological regulation, or pathological disorders. It is, in general, a process of enormous complexity involving genetic, biochemical, and physical components at many different scales and with complex interactions. The purpose of this paper is to provide a simple introduction to the modeling of elastic growth. We first consider systems in one-dimensions (suitable to model filamentary structures)to introduce the key concepts. Second, we review the general three-dimensional theory and show how to apply it to the growth of cylindrical structures. Different possible growth mechanisms are considered.

Key words: Biological growth, growing rods, morphoelasticity, Mooney-Rivlin material.

1 Introduction

Biological growth is a fascinating process of tremendous complexity that has attracted the attention of generations of biologists and remains today a fundamental scientific problem. Surprisingly, this problem has met with little interest in the physics and mathematics community. However, with the development of quantitative biomechanics (in the footsteps of scientists like Skalak and Fung), the mathematical development of exact elasticity, the physical modeling of growth, and computational advances, a theory of growth has emerged and the mathematical analysis of its consequences is finally possible. The purpose of this article is to provide an introduction to the problem of growth, its mathematical issues and the scientific challenges ahead of us.

The emphasis will be on the particular role played by mechanical quantities (such as stresses and strains) and their interaction with changes in geometry arising during growth. Based on these observations and simple mechanical systems in one dimension we will discuss different approaches to modeling macroscopic growth in continuum mechanics and show how to generalize the

classical theory of exact elasticity. The mathematical analysis of such a theory of growth enables us to understand the particular interaction between geometry and mechanics and helps us to identify particular mechanisms that can be used either in building specific material properties, in homeostatic regulation, or in embryonic development through instability-driven pattern formation.

2 One-dimensional Theory: Elasticity, Visco-elasticity, and Plasticity

We start with a simple conceptual framework by considering growth phenomena in one dimension. That is, we consider the growth of a (mostly) filamentary structure. This type of growth is found in many microbial systems such as filamentous bacteria and fungi but also in plants where stems, roots, and tendrils all display some aspects of one-dimensional growth. In size, these systems span at least 6 orders of magnitude from microns to meters. Biologists would be quick (and correct) to point out that growth in these systems is much more complex and involves structural details at the wall level which are necessary to describe any features related to growth. Here we choose to look at these systems as a mechanical continuum for which some features and time-evolution are dominated by its slenderness and hence can be modeled as elastic filaments.

Before reviewing the published plant growth models it is useful to recall the basic facts about Kelvin solids, Maxwell fluids, and Bingham fluids.

2.1 Kelvin Solids

A purely elastic material is one in which the response to applied stresses is instantaneous and reversible. A Kelvin solid is an elastic solid but one in which the response to the stress occurs over a finite time determined by the viscous characteristic of the "solid". In the simplest one dimensional case, one can write

$$\sigma = E\epsilon + \eta\frac{\pi\epsilon}{\pi t}, \tag{1}$$

where σ denotes the (applied) stress, and ϵ denotes strain. E is an elastic modulus and η a coefficient of viscosity. If $\eta = 0$ the equation reduces to the classical constitutive relation for an elastic solid. If $\eta \neq 0$, the basic idea can be illustrated by the case of a constant applied stress σ_0.

$$\sigma_0 = E\epsilon + \eta\frac{\pi\epsilon}{\pi t}. \tag{2}$$

The equation is easily solved for ϵ to give

$$\epsilon(t) = \frac{\sigma_0}{E}\left(1 - e^{-t/\tau_r}\right), \tag{3}$$

where

$$\tau_r = \eta/E, \tag{4}$$

is the "visco-elastic" relaxation time. Another standard exercise is to impose a periodic stress $\sigma(t) = \sigma_0 \cos(\omega t)$ which gives, in the limit $t \to \infty$

$$\epsilon(t) = \frac{\sigma_0 E}{E^2 + \eta^2 E^2} \cos(\omega t - \alpha), \tag{5}$$

which shows a phase lag α between the strain and the applied stress, where

$$\alpha = \arctan(\tau_r/\tau_\sigma), \tag{6}$$

represents the competition between the response time of the solid $\tau_r = \eta/E$ and the time scale of the applied stress $\tau_\sigma = 1/\omega$. One should also note that if the stress is suddenly turned off, the strain relaxes as

$$\epsilon \sim e^{-t/\tau_r}. \tag{7}$$

2.2 Maxwell Fluids

A Maxwell fluid is a viscous fluid with some elastic properties. The simplest model is one in which the *rate of strain* obeys the equation

$$\frac{\pi\epsilon}{\pi t} = \frac{1}{\eta}\sigma + \frac{1}{E}\frac{\pi\sigma}{\pi t}. \tag{8}$$

Three simple comments:

1. If the second term on the r.h.s is dropped one is left with the simplest fluid model
$$\frac{\pi\epsilon}{\pi t} = \frac{1}{\eta}\sigma, \tag{9}$$
in which rate of strain is proportional to the stress and the fluid exhibits irreversible flow.
2. If the first term on the r.h.s. is dropped one is left with
$$\frac{\pi\epsilon}{\pi t} = \frac{1}{E}\frac{\pi\sigma}{\pi t}, \tag{10}$$
which represents the elastic component of the material: after integrating both sides w.r.t. time one simply has the pure elastic response $\epsilon = \sigma/E$.
3. If the strain is turned off, the stress relaxes as
$$\sigma \sim e^{-t/\tau_r}, \tag{11}$$
where τ_r is as above - which should be contrasted with the equivalent strain relaxation of a Kelvin solid when the stress is turned off.

Basic features of the Maxwell model can be illustrated with an applied stress of the form

$$\sigma = \sigma_0 \left(1 - e^{-t/T}\right). \tag{12}$$

If T is small the stress ramps up rapidly; if T is large, the stress ramps up slowly. The rate of strain equation can be integrated explicitly to give

$$\epsilon(t) = \frac{\sigma_0}{\eta}\left(t + T\left(e^{-t/T} - 1\right)\right) + \frac{\sigma_0}{E}\left(1 - e^{-t/T}\right). \tag{13}$$

(i) For the case $0 < t < T$, with T small, *i.e.* for short times in the case of a rapidly ramped stress, one finds that

$$\epsilon \sim \frac{\sigma_0}{ET} t, \tag{14}$$

which shows that the strain follows the applied stress according the elastic part of the system, $\dot{\epsilon} \sim \dot{\sigma}/E$.
(ii) For $t \gg T$, one finds that

$$\epsilon \sim \frac{\sigma_0}{\eta} t, \tag{15}$$

which shows that the strain is dominated by the fluid component of the system, $\dot{\epsilon} \sim \sigma/\eta$.

2.3 Bingham Fluids

In a simple fluid there is (irreversible) flow in response to applied stress, however small. For non-Newtonian fluids such as paint, flow does not begin until a critical yield stress, σ^*, has been exceeded. This is the Bingham model which, in its simplest form, is expressed as

$$\frac{\pi\epsilon}{\pi t} = \frac{1}{\nu}[\sigma - \sigma^*], \tag{16}$$

where $[\sigma - \sigma^*] = \max(0, (\sigma - \sigma^*))$. This model of irreversible extension (flow) once a critical stress has been exceeded, has been the paradigm for most plant growth models. The Bingham model can easily be generalized to a Maxwell-Bingham type system represented by

$$\frac{\pi\epsilon}{\pi t} = \frac{1}{\nu}[\sigma - \sigma^*] + \frac{1}{E}\frac{\pi\sigma}{\pi t}. \tag{17}$$

In terms of terminology, the convention is (or should be!) to call a Maxwell fluid *visco-elastic*, reflecting the combination of irreversible flow generated by viscous stresses with an elastic component; and to term a Bingham fluid *plastic* - a much misused term which is (or should) be used to mean an irreversible deformation beyond a critical yield stress. A comparison of the strains produced by the different models is given in Fig.1).

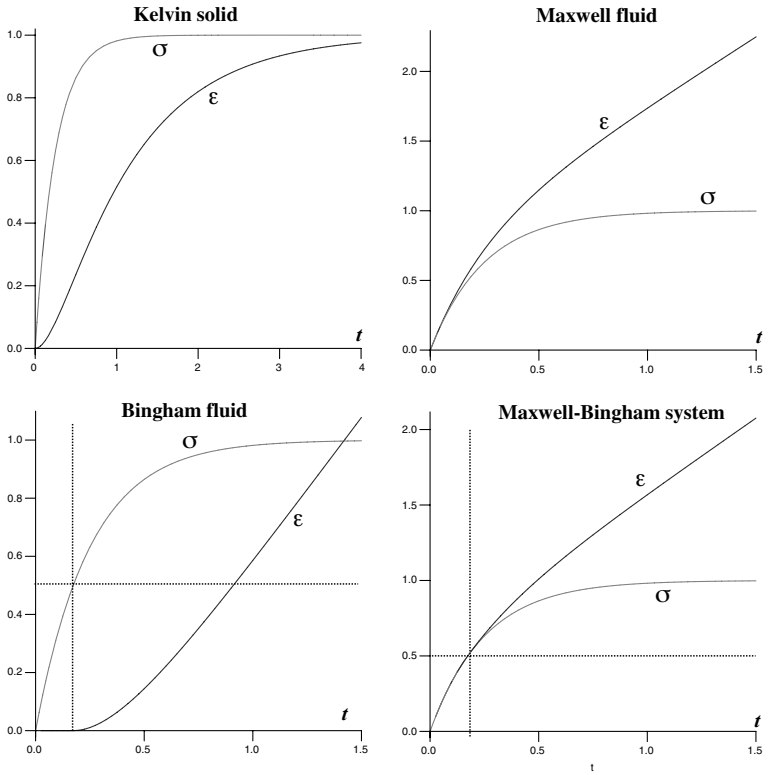


Fig. 1. Comparison of strains produced by a ramping of the stress for four different materials, $\nu = 1/4, \sigma^* = 1/2, E = \sigma_0 = \eta = 1, T = 1/4$.

3 One-dimensional Theory: Growth Models

3.1 Lockhart-Ortega-Cosgrove Model

Lockhart's model is one of the earliest quantitative models of plant cell growth [45]. Geometrically the plant cell is considered to be an axisymmetric cylinder of constant radius, but growing length. This growth in length is related to the increase in volume due to water entering the cell, and the irreversible length increase is, in turn, taken to be the result of the turgor pressure exerted on the cell wall (in fact, in this model, this is the pressure exerted on the end of cylinder which is treated as a flat cap). Lockhart's discussion begins with a seemingly simple statement about the elastic strain in the system which he defines as

$$\epsilon = \frac{l - l_0}{l_0}, \tag{18}$$

and is governed by a simple Hooke's law, namely

$$\epsilon = \frac{1}{E}\sigma. \tag{19}$$

Assuming a constant applied stress σ, time differentiation of the strain gives

$$\dot{\epsilon} = \frac{\dot{l}}{l_0} - \frac{l}{l_0}\frac{\dot{l_0}}{l_0} = 0, \tag{20}$$

and hence

$$\frac{\dot{l}}{l} = \frac{\dot{l_0}}{l_0}. \tag{21}$$

In terms of our own language, what does this mean? If we regard l as the current length and l_0 as the reference length, then if both are growing in time, a constant rate of strain can be maintained if the reference length extends elastically, at each instant, the same amount as the current length. Lockhart points out that the time scale of elastic equilibrium for plant cells (minutes) is much more rapid than the time scale of the irreversible extension (hours); hence the system is claimed to always be in a state of elastic equilibrium as it grows. Lockhart's argument proceeds in two parts. The increase in (current) length due to volume increase (due to osmosis) is expressed as

$$\frac{dl}{dt} = \frac{KA}{a}(\Delta\Pi - P) \tag{22}$$

where K is a water permeability constant, $A = 2\pi rl$ is the cylinder side wall area, $a = \pi r^2$ the cross-sectional area, $\Delta\Pi$ an osmotic pressure variable, and P the turgor pressure. This is re-expressed as

$$\frac{1}{l}\frac{dl}{dt} = \frac{2K}{r}(\Delta\Pi - P). \tag{23}$$

The left hand side $\dot{l}/l$ could be thought of as a "current configuration rate of strain". He then goes on to define the "irreversible wall extension" as

$$\frac{1}{l_0}\frac{dl_0}{dt} = \Phi\sigma, \tag{24}$$

where Φ characterizes the cell wall's "rate of irreversible flow" (Lockhart's terminology). A few comments: (i) the left hand side $\dot{l_0}/l_0$ could be thought of as a "reference configuration rate of strain" - from the point of view of a growth process it may indeed be appropriate to measure the growth in terms of the change in reference configuration length; (ii) the equation of motion represents a simple fluid flow (*i.e.* rate of strain $\propto$ stress); (iii) as Lockhart points out (24) could be modified to correspond to a Bingham type flow, namely

$$\frac{1}{l_0}\frac{dl_0}{dt} = \Phi[\sigma - \sigma^*]. \tag{25}$$

The stress σ is expressed, in the standard way, in terms of the (turgor) pressure on the end cap and the wall thickness, δ, *i.e.* $\sigma = \pi r^2 P / 2\pi r\delta$. The equation of motion (24) is then used to express P in terms of the extension, namely

$$P = \frac{2\delta}{r\Phi}\frac{1}{l_0}\frac{dl_0}{dt} = \frac{2\delta}{r\Phi}\frac{1}{l}\frac{dl}{dt}, \tag{26}$$

where (21) is used to obtain the last equality. This expression for P is then combined with (23) to give

$$\frac{1}{l}\frac{dl}{dt} = \frac{2rK\Delta\Pi\Phi}{4\delta K + r^2\Phi}, \tag{27}$$

which is essentially Lockhart's main result (introducing a Bingham type flow only modifies the equation slightly) which, in the end, relates the current rate of strain to a stress representing the interplay of osmotic and turgor pressures. The various way in which strain, and rates of strain, are defined - and then connected through the assumption of constant elastic strain - is not especially satisfactory. We also note (see below) that the equation does not depend on the elastic modulus of the system.

The assumed state of constant elastic equilibrium is a consequence of the assumed constancy of the stresses. This means that the constitutive relation (24) is simply that of a (simple) fluid. Ortega proposed that the effect of elasticity can be explicitly restored by replacing (24)or (25) by a Maxwell type relationship [58], namely

$$\frac{de}{dt} = \Phi[\sigma - \sigma^*] + \frac{1}{E}\frac{d\sigma}{dt}, \tag{28}$$

where de/dt is the elongation strain rate which Ortega defines as $de/dt = (1/l)dl/dt$. Ortega's analysis of his model yields some interesting extensions of Lockhart's model - which we shall not pursue here. A somewhat similar discussion/extension of Lockhart's model was also given by Cosgrove [11].

3.2 Goodwin Model

Goodwin [22] begins by summarizing the Lockhart-Cosgrove-Ortega model in the form

$$\frac{1}{V}\frac{dV}{dt} = \phi[P - Y] + \frac{1}{E}\frac{dP}{dt}, \tag{29}$$

where he describes Y as the yield threshold, ϕ as the extensibility coefficient of the wall, and E as the volumetric elastic modulus $E = VdP/dV$. Noting that the volumetric growth rate $(1/V)dV/dt$ is analogous to a rate of strain and P corresponds to a mechanical stress, he recasts the equation in the form

$$E\frac{d\epsilon}{dt} = \frac{1}{\tau}[\sigma - Y] + \frac{d\sigma}{dt}, \tag{30}$$

where ϵ is the strain and in the form of a Maxwell-Bingham constitutive relationship. He notes that when $\sigma > Y$, the strain ϵ is a combination of an elastic strain (proportional to the stress variations) and a growth strain Γ, where the rate of growth is expressed as

$$E\frac{d\Gamma}{dt} = \frac{1}{\tau}[\sigma - Y] \tag{31}$$

We see that (30) can be written as

$$\frac{d\epsilon}{dt} = \frac{d\Gamma}{dt} + \frac{1}{E}\frac{d\sigma}{dt}. \tag{32}$$

and as we will see this is the bridge to the three-dimensional growth model by Rodriguez *et al.* described in section 4.1. Goodwin defines the growth rate in terms of the reference length, namely

$$\frac{d\Gamma}{dt} = \frac{1}{l_0}\frac{dl_0}{dt}, \tag{33}$$

and hence the growth rate equation

$$\frac{1}{l_0}\frac{dl_0}{dt} = \frac{1}{\tau}[\epsilon - s], \tag{34}$$

where $s = Y/E$ and the stress σ is expressed in terms of the elastic strain, *i.e.* $\sigma = \epsilon E$, This latter assumption may not be a good physical model. The growth rate equation (34) is now a strain based model, *i.e.* irreversible extension if the elastic strain exceeds a critical *strain* threshold. Goodwin then argues that there could also be a contribution to the growth as a result of a change in the elastic modulus. Thus assuming the Hooke's law $\sigma = E\epsilon$ and differentiating both sides w.r.t. time under the assumption of constant stress gives

$$\frac{d\epsilon}{dt} = -\frac{\epsilon}{E}\frac{dE}{dt}, \tag{35}$$

If the elastic strain ϵ is taken to be $\epsilon = (l - l_0)/l_0$ and that the variations in ϵ is due to changes in l_0 (a proposition that might require a little more thought) then the above relationship can be cast in the form

$$\frac{1}{l_0}\frac{dl_0}{dt} = \frac{\epsilon}{1+\epsilon}\frac{1}{E}\frac{dE}{dt}. \tag{36}$$

If this is added to (34) one has the Goodwin growth rate model

$$\frac{1}{l_0}\frac{dl_0}{dt} = \frac{1}{\tau}[\epsilon - s] + \frac{\epsilon}{1+\epsilon}\frac{1}{E}\frac{dE}{dt}. \tag{37}$$

Two comments: (i) the recognition of a total rate of strain being decomposed in to a growth rate part *analogous to* a simple Bingham fluid flow, and an elastic (rate of strain) part is fundamental to the model and is, in fact, equivalent to the simplest form of Prandtl-Reuss equations for elasto-plastic deformation ; (ii) the separate consideration of the elastic stress variation (35) and then its addition to the growth rate component, could be interpreted as a statement of the widely different time scales of the growth and elastic processes.

3.3 Stein Model

The discussion of (biological) rod growth by A.A. Stein [71] is the most self-consistent of the models to date and, as is quickly apparent, a linearized version of the more general approach of Rodriguez *et. al.* (see section 4.1). His starting point is

$$\dot{\epsilon} = \dot{\epsilon_g} + \dot{\epsilon_e}, \tag{38}$$

which could be interpreted as the total rate of strain, $\dot{\epsilon}$ equals the growth rate $\dot{\epsilon_g}$ plus the elastic rate of strain $\dot{\epsilon_e}$. [1] He proposes that the growth rate takes the general form

$$\dot{\epsilon_g} = A + M\sigma, \tag{39}$$

and the elastic *strain*, ϵ_e, satisfies a Hookean relationship

$$\epsilon_e = K\sigma. \tag{40}$$

In the above three equations all variables are now taken to be tensorial. Given the above, the rest is straight forward. In component form (38) is thus

$$\dot{\epsilon}_{kl} = A_{kl} + M_{klmn}\sigma^{mn} + \frac{d}{dt}\left(M_{klmn}\sigma^{mn}\right). \tag{41}$$

He makes an intriguing aside that since the elastic deformations are so small, the type of time derivative is unimportant, so d/dt can be taken as differentiation w.r.t. time for fixed comoving coordinates, and that use of the Oldroyd derivative instead of this would only add negligibly small corrections. The stresses are assumed to satisfy the standard equilibrium equations

$$\frac{\pi\sigma^{kl}}{\pi x^l} + F^k = 0. \tag{42}$$

Stein first considers the case of a growing rectangular cylinder under pressure yielding simple equations for the growth of the cylinder length (no great surprises found here). It would appear that be defines his (total) rate of strain as $\dot{l}/l$ where l is the current length. He then goes on to tackle the more difficult case of a growing bending rod.

3.4 Growing Cosserat Rods

A rod [2, 10, 47] is represented by its centerline $\mathbf{r}(s)$ where s is a material parameter taken to be the arc length in a stress free configuration ($0 \leq s \leq L$) and two orthonormal vector fields $\mathbf{d}_1(s)$, $\mathbf{d}_2(s)$ representing the orientation of a material cross section at s.

[1] Care must be taken when defining "elastic rate of strain" - a point we will discuss later. (We comment that Stein refers to the various $\dot{\epsilon}$ as "velocity tensors of the [associated]...strains". As a pedantic point - which we will pursue later - we note that velocity gradients and rates of strain are only equivalent in infinitesimal elasticity theory.)

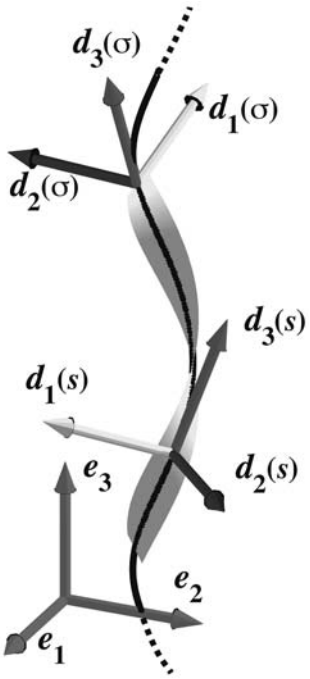


Fig. 2. The director basis represents the evolution of a local basis along the rod.

A local orthonormal basis is obtained (see Fig. 2) by defining $\mathbf{d}_3(s) = \mathbf{d}_1(s) \times \mathbf{d}_2(s)$ and a complete kinetic description is given by:

$$\mathbf{r}' = \mathbf{v}, \tag{43}$$

$$\mathbf{d}_i' = \mathbf{u} \times \mathbf{d}_i, \qquad i = 1, 2, 3, \tag{44}$$

$$\dot{\mathbf{d}}_i = \mathbf{w} \times \mathbf{d}_i \qquad i = 1, 2, 3, \tag{45}$$

where $(\,)'$ and $\dot{(\,)}$ denote the derivative with respect to s and t, and $\mathbf{u}$, $\mathbf{v}$ are the *strain* vectors and $\mathbf{w}$ is the *spin* vector. The components of a vector $\mathbf{a} = \mathsf{a}_1\mathbf{d_1} + \mathsf{a}_2\mathbf{d_2} + \mathsf{a}_3\mathbf{d_3}$ in the local basis are denoted by $\mathbf{a} = (\mathsf{a}_1, \mathsf{a}_2, \mathsf{a}_3)$ (following [2], we use the sans-serif fonts to denote the components of a vector in the local basis). The two first components represent transverse shearing while $\mathsf{v}_3 > 0$ is associated with stretching and compression. The two first components of the *curvature vector* $\mathbf{u}$, are associated with bending while u_3 represents twisting.

The stress acting at $\mathbf{r}(s)$ is given by a resultant force $\mathbf{N}(\mathbf{s})$ and resultant moment $\mathbf{m}(s)$. The balance of linear and angular momenta yields [2]

$$\mathbf{n}' + \mathbf{f} = \rho\mathcal{A}\ddot{\mathbf{r}}, \tag{46}$$

$$\mathbf{m}' + \mathbf{r}' \times \mathbf{n} + \mathbf{l} = \rho\left(I_2\mathbf{d_1} \times \ddot{\mathbf{d}}_\mathbf{1} + I_1\mathbf{d_2} \times \ddot{\mathbf{d}}_\mathbf{2}\right), \tag{47}$$

where $\mathbf{f}(s)$ and $\mathbf{l}(s)$ are the body force and couple per unit length applied on the cross section at s (these body forces and couple can be used to model different effects such as short and long range interactions between different part of the rod or can be the result of self-contact or contact with another body), $\mathcal{A}$ is the cross-section area, ρ the mass density, and $I_{1,2}$ are the principal moments of inertia of the cross section (corresponding to the directions $\mathbf{d}_{1,2}$).

To close the system, we assume that the resultant stresses are related to the strains. There are two important cases to distinguish.

Extensible and Shearable Rods

First, we consider the case where the rod is extensible and shearable and we assume that there exists a strain-energy density function $W = W(\mathbf{y}, \mathbf{z}, s)$ such that the constitutive relations for the resultant moment and force in the local basis are given by

$$\mathbf{m} = f(\mathbf{u} - \hat{\mathbf{u}}, \mathbf{v} - \hat{\mathbf{v}}, s) = \partial_{\mathbf{y}} W(\mathbf{u} - \hat{\mathbf{u}}, \mathbf{v} - \hat{\mathbf{v}}, s), \tag{48}$$

$$\mathbf{n} = g(\mathbf{u} - \hat{\mathbf{u}}, \mathbf{v} - \hat{\mathbf{v}}, s) = \partial_{\mathbf{z}} W(\mathbf{u} - \hat{\mathbf{u}}, \mathbf{v} - \hat{\mathbf{v}}, s), \tag{49}$$

where $\hat{\mathbf{v}}$, $\hat{\mathbf{u}}$ are the strains in the unstressed reference configuration ($\mathbf{m} = \mathbf{n} = 0$ when $\mathbf{u} = \hat{\mathbf{u}}$, $\mathbf{v} = \hat{\mathbf{v}}$). Typically, W is assumed to be continuously differentiable, convex, and coercive. The rod is *uniform* if its material properties do not change along its length (*i.e.* W has no explicit dependence on s) and the stress-free strains $\hat{\mathbf{v}}$, $\hat{\mathbf{u}}$ are independent of s.

Inextensible and Unshearable Rods

In the second case, we assume that the rod is inextensible and unshearable, that is we take $\mathbf{v} = \mathbf{d}_3$ and the material parameter s becomes the arc length. In that case, there is no constitutive relationship for the resultant force and the strain-energy density function is a function only of $(\mathbf{u} - \hat{\mathbf{u}})$, that is

$$\mathbf{m} = \partial_{\mathbf{y}} W(\mathbf{u} - \hat{\mathbf{u}}) = f(\mathbf{u} - \hat{\mathbf{u}}) \tag{50}$$

In the simplest (and most widely used) case the energy is

$$W_1 = K_1 u_1^2 + K_2(u_2 - \hat{u}_2)^2 + K_3(u_3 - \hat{u}_3)^2, \tag{51}$$

where $\hat{u}_2$ and $\hat{u}_3$ represent the intrinsic curvature and torsion that represent the shape of the filament when unloaded. Explicitly, the resultant moment is

$$\mathbf{m} = EI_1 \mathsf{u}_1 \mathbf{d}_1 + EI_2(u_2 - \hat{u}_2)\mathbf{d}_2 + \mu J(u_3 - \hat{u}_3)\mathbf{d}_3 \tag{52}$$

where E is the young modulus, μ is the shear modulus, and J is a parameter that depends on the cross-section shape (an explicit form for J and examples can be found in [29]). For a circular cross-section, these parameters are:

$$I_1 = I_2 = \frac{J}{2} = \frac{\pi R^4}{4}, \tag{53}$$

where R is the radius of the cross-section. The products EI_1 and EI_2 are usually called the *principal bending stiffnesses* of the rod, and μJ is the *torsional stiffness*.

The orthonormal frame $(\mathbf{d_1}, \mathbf{d_2}, \mathbf{d_3})$ is different from the Frenet-Serret frame defined by the triple (normal,binormal,tangent)=$(\boldsymbol{\nu}, \boldsymbol{\beta}, \boldsymbol{\tau})$. If we take $\mathsf{v}_1 = \mathsf{v}_2 = 0,\ \mathsf{v}_3 = 1$, then the vectors $(\mathbf{d_1}, \mathbf{d_2})$ lie in the normal plane to the axis and are related to the normal and binormal vectors by a rotation through an angle φ:

$$\mathbf{d}_1 = \boldsymbol{\nu} \cos\varphi + \boldsymbol{\beta} \sin\varphi \tag{54}$$

$$\mathbf{d}_2 = -\boldsymbol{\nu} \sin\varphi + \boldsymbol{\beta} \cos\varphi \tag{55}$$

This rotation implies that

$$\mathbf{u} = (\kappa \sin\varphi, \kappa \cos\varphi, \tau + \varphi') \tag{56}$$

where κ and τ are the usual Frenet curvature and torsion.

Growing Rods

We are now in a position to model growth in elastic rods. There are actually three different ways that this can be achieved. The first approach, which we refer to as *parameter variation* consists in considering families of rod solutions (typically static due to the slow time evolution of growth with respect to viscous damping in the rod) parametrized by one of the material parameters. For instance, in the growth of a tree, one may consider the length and width as two parameters that evolve in time. At each time, we increase the value of such parameters and recompute the static solution that match the boundary conditions. The second approach is *remodeling*. The idea is now to consider a separate evolution law for the material parameters that may depend on time and history of the material. This is fundamentally different from the previous approach since the material parameters may now be a local function of the position and their values depend on the evolution in time. The third approach is the *evolution of the natural configuration*. This is somewhat more subtle and will open the discussion to the general discussion of growth in three-dimensional nonlinear elasticity.

Growing Rods: Parameter Variation

In this quasi-static approach, each solution remains a solution of the classic Kirchhoff equations and growth is studied by considering the evolution of such solutions w.r.t. the parameters. The idea is to study the evolution of shape

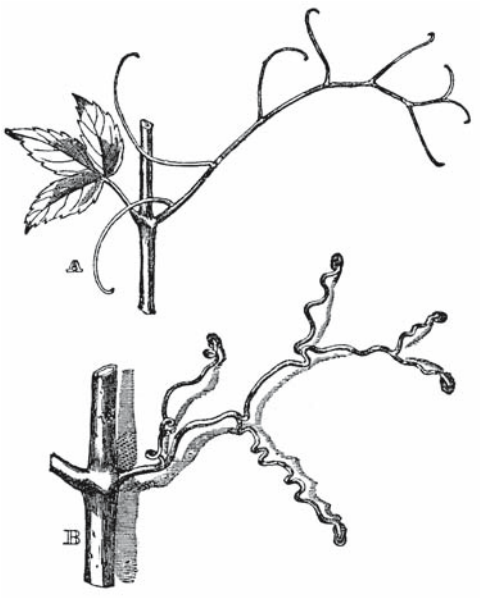


Fig. 3. Growth of a tendril. Once attached the tendril develops a perversion composed of two helical structure with opposite handedness Drawing from [12]. In the first stage (A), the tendrils are *circumnutating* until they find an attachment. In the second stage (B), the tendrils are attached and perversion sets in.

and change in shape through a bifurcation process mediated by a control parameter.

As a first example, consider the evolution of tendrils in plants [28, 50]. A tendril is a modified leaf that can be found at the extremities of some climbing plants and are used by the plants to achieve vertical growth by attaching itself to other supports. A tendril can be modeled as an elastic rod under tension. Once a tendril has grasped a support, it starts developing curvature by differential growth until it bifurcates to a shape made out of 2 (or more) helical structures with opposite handedness called a *perversion* (this is due to the fact that the original structure has no twist and neither ends are allowed to rotate–See Fig.3). These helical springs provide the climbing plant with a firm but elastic connection to its support [12]. The creation of these helical structures from a straight filament can be understood in terms of parameter variation. The tendril is modeled as an initially straight filament under tension. Its constant intrinsic curvature increases slowly in time and the filament is considered to be in static equilibrium at all time. The problem reduces to exploring the possibility of a bifurcation from a straight solution to a solution connecting asymptotically two helical structure of opposite handedness (See Fig. 4).

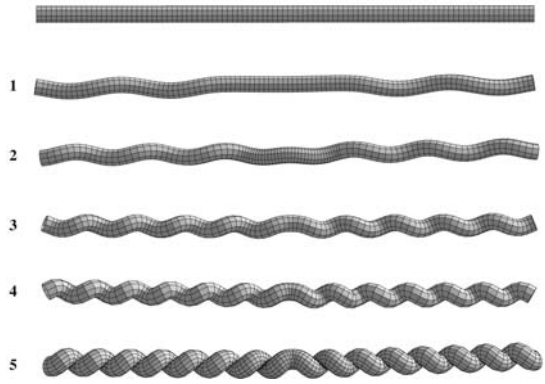


Fig. 4. Model of perversion. A straight rod under tension undergoes a bifurcation when its intrinsic curvature is increased (from top to bottom). Reproduced from [50].

As a second example, consider the growth of twining vines. Twining plants achieve vertical growth by revolving around supports of different sizes on which they exert a pressure. During growth, the growing tip waves around in a circular motion known as circumnutation until it finds an appropriate upright support and then start wrapping around it to extend upward. The tip of the vine keeps nutating and the vine pursue its climbing process by forming a spiral around the support. The growth process of twining plants raises many interesting mechanical questions already noted by 19th century botanists and further studied by Silk, Holbrook and co-workers [48,64,68–70]. Viewed as a growth problem, we can study the possible equilibria of a rod with intrinsic curvature and torsion in contact with a cylinder and with increasing length. Therefore the problem reduces to finding suitable solution with increasing length, taken as our control parameter. There exist many different regimes that can be studied from a bifurcation standpoint, in particular, one can determine the maximal pole radius around which a vine can grow, an interesting question raised by Charles Darwin (1888) (see [25] for details). Here we restrict our attention to the problem of finding a solution that corresponds to the correct mechanical behavior of the plant during growth. That is, the vine connect a helical solution to a hook like structure (termed the *anchor*). Such solutions were found and an example is given in Fig.5.

Another example of growth through parameter variation can be found in [76] where the growth of twisted circular ring with application to the growth of *B. subtilis* was considered (See below).

Growing Rods: Remodeling

In the previous examples, growth was passive. That is, it is modeled by the evolution of an outside control parameter without any feedback from the form to the material parameters. In many growth process, the evolution of the struc-

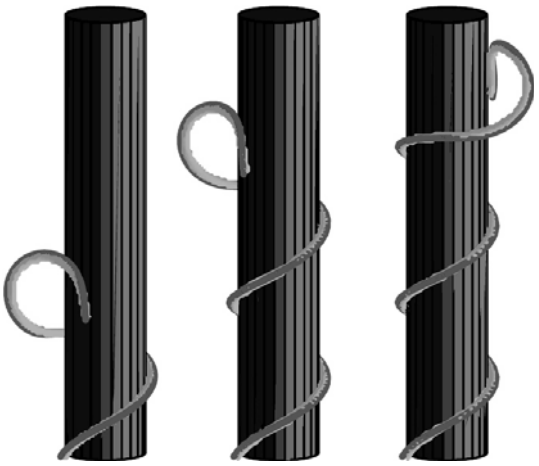


Fig. 5. A sequence of three-dimensional solutions to the attachment problem. Note the continuous, almost helical, solution, followed by the anchor that provides tension in the filament. Figure reproduced from [25].

ture directly influences the evolution of the material parameter. For instance, a branch of a tree can be trained to grow in a certain shape (which makes for beautiful alleys in French garden). At first, the elastic structure is loaded and stressed into a particular shape. As time passes, the structure remodels itself in such a way as to relieve the stresses and the structure permanently sets in, even in the absence of loading.

As a simple example, consider the case of an unshearable, inextensible planar rod under end compression [21]. As the rod buckles it takes a new shape. At this point it is assumed that the natural state of the rod will evolve towards this equilibrium shape. That is, its intrinsic curvature $\hat{\kappa} = \hat{u}_2$ evolves in time towards its actual curvature. A simple model for viscoelastic relaxation of curvature is

$$T\frac{\partial \hat{\kappa}}{\partial t} = \kappa - \hat{\kappa} \tag{57}$$

where T is a typical time corresponding to the viscoelastic response of the material. The important point to notice here is that, after buckling, the curvature changes at all points and when the intrinsic curvature relaxes it takes different values at all point. The process depends on the loading and the parameters but also on the history of the loading process which makes it fundamentally different from the previous modeling (through parameter variation). To emphasize this point, consider Fig. 6 where an initially straight rod was loaded with a given ramp and allowed to relax following the rule given.

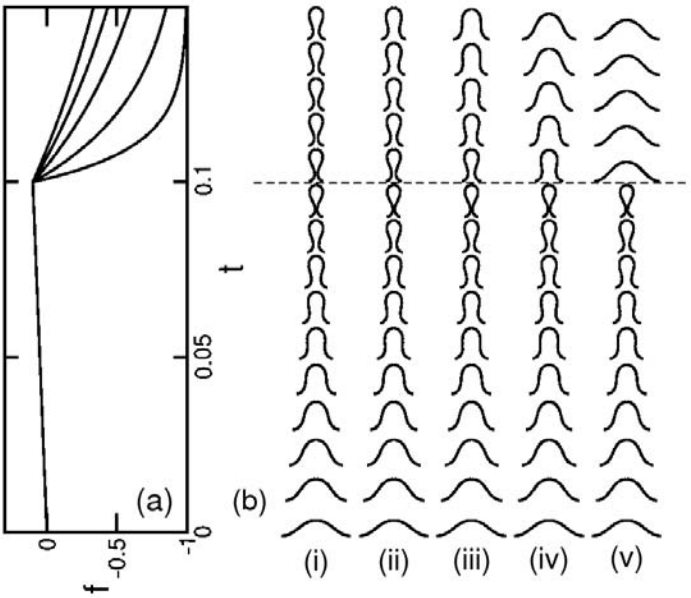


Fig. 6. Filaments trained with time-dependent forces. For each of the force-time profiles shown in (a) is the corresponding sequence of filament shapes (a) shown at intervals of $\Delta t = 0.01$. Dashed line indicates end of linear ramp [21]. As can be clearly seen in picture (i), the filament has remodeled into its shape and retains it after the load is removed. With longer unloading times ((ii) to (v)), the filament when unloaded, partially relaxes to its original shape.

Growing Rods: Evolution of the Natural Configuration

The two previous modeling approaches work for systems where the law for the growth evolution of the material can be described by changes in the material parameters. However, it is not suitable to describe other aspects of growth. For instance, consider a naturally straight untwisted rod in its unstressed configuration and allow it to increase in length. If the increase in length is uniform (independent of the material parameter), then it can be described by changing the length (a material parameter) as before. However, if growth is not uniform but depends on the position, stresses, or strains, the growth evolution cannot be simply described by a change in the material parameters. The essence of the problem comes from the fact that the reference configuration of the rod changes due to growth. For the purpose of this discussion consider an unshearable but extensible rod assume that growth only acts by changing the local element of length. The strain variable associated with a local change in length is $\mathsf{v}_3 = \lambda$. If the rod is not growing, a typical constitutive law for $\lambda = \lambda_\mathrm{e}$ is

$$\mathsf{n}_3 = \epsilon(\lambda_\mathrm{e} - 1) \tag{58}$$

with an extension modulus ϵ relating the tension in the rod to its elastic extension ($\lambda_e > 1$) or compression ($0 < \lambda_e < 1$). Now, the extension may also be created through growth. In the absence of tension, we introduce λ_g to describe the local extensional growth ($\lambda_g > 1$ for growth). In general when both growth and elasticity are combined, we split the extensional strain into

$$\lambda = \lambda_e \lambda_g \tag{59}$$

Since we have introduced a new strain variable λ_g a constitutive relationship for its evolution must be prescribed. This depends on the system being considered (see Section 4.1 for a discussion). Typically, an equation for the evolution of the growth rate as a function of the stress and material parameter will be specified

$$\dot{\lambda}_g = \lambda_g F(\mathbf{n}, s). \tag{60}$$

As an example of a local growth law, we consider the evolution of length and twist in a model for the growth of *Bacillus subtilis*.

3.5 Application to the Growth of Bacillus subtilis

The individual cells of the bacterial strain *Bacillus subtilis* are rod-shaped and typically of length $3 - 4\mu m$ and diameter $0.8\mu m$. Under certain circumstances they are found to grow into filaments consisting of the cells linked in tandem due to the failure of daughter cells, produced by growth and septation, to separate. As they elongate these filaments, which are immersed in a liquid environment (whose temperature and viscosity can be controlled), are observed to twist at uniform rate. The degree of twist and handedness can, in fact, be controlled experimentally and a wide range of states from left-handed to right-handed forms, can be produced. The actual twist state of the cells seems to be related to properties of the polymers which are inserted into the the cell wall during growth. As they elongate the filaments are observed to writhe and eventually deform into double-helical structures. These continue to grow and periodic repetition of this process results in macroscopic fibers (termed "macro-fibers") with a specific twist state and handedness. (A schematic representation of this dynamics is given in Fig. 7).

A striking feature of this iterated process is that at every stage of the self-assembly, the handedness of the helical structures that are created is the *same* (e.g. a right-handed double helix gives rise to a right-handed four-strand helix and so on). The nature of the environment does, however, influence certain aspects of the self-assembly. In a viscous environment the basic writhing instability leads to something of a "buckling" at the middle of the filament with the formation of a tight central loop; this is followed by a helical wind-up which starts at the base of this loop (Fig. 7B). By contrast, in a non-viscous medium, the instability causes the filament to fold over into a large loop closed by contact between the ends of the filament. This closure is then followed by

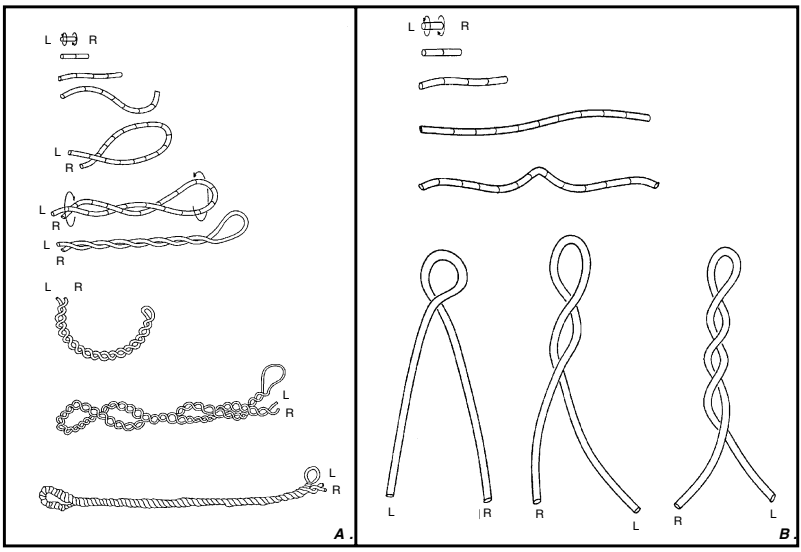


Fig. 7. The two basic looping of *Bacillus subtilis*: A. In non-viscous medium, B. In viscous medium (Picture courtesy of Neil Mendelson).

a helical wind-up starting at that point. In both cases this self-assembly conserves handedness and usually continues over long periods until macro-fibers, several millimeters long, are formed.

The dynamics of the self-assembly and the mechanical properties of the bacterial threads have been studied in great detail by Mendelson and coworkers over many years [51–53]. In addition to the fascinating questions of growth and form raised by this process, the macrofibers themselves offer the prospect of unusual bio-materials that can be mineralized and packed in ways that are are of practical biomedical and biotechnical use.

Mathematical modeling of Bacillus subtilis presents many challenges. With a little thought the handedness preserving nature of the basic instability can be explained in qualitative terms in which an "under-twisted" filament undergoes a writhing instability. Fundamental to this picture is the idea of an "intrinsic twist" associated with each cell and which drives the dynamics. This picture is, in a sense, quite universal and can also be seen in other filamentary structures ranging from the microscopic (e.g. DNA) to the macroscopic (*e.g.* telephone cables). The case of *Bacillus subtilis* is complicated by the fact that the cell filament is growing exponentially and that the ends of the filament are normally unconstrained. This freedom of the ends is a nontrivial feature of the basic instability that initiates the self-assembly. Once the filament has folded over (in either a big loop for non-viscous media or a tight central loop for viscous ones) the resulting self-contact effectively blocks free rotation of the filament. This changing of boundary conditions (one end is now effec-

tively fixed) provides the mechanism for the subsequent helical wind up of the strands.

In *Bacillus subtilis* (and DNA) where the twisting and supercoiling are of the *same* handedness. The key to this transition is the idea of an *intrinsic twist*; namely a natural state of the filament with a non-zero twist density (twist per unit length) as described in the Kirchhoff model by Eq.(52) with a non-vanishing parameter $\hat{u}_3$.

To understand the process, consider a rubber tube with right-handed intrinsic twist represented by drawing marker lines with a right-handed helical pitch. One end is twisted in the opposite direction to the marker lines until they are approximately straight. The twist density of the filament now appears to be zero whereas its natural state is one of non-zero twist density - as indicated by the original helical marker lines. Thus we now have a "twist deficit". To return to its natural state the tube must make up for this deficit by restoring twist. This can be achieved in two different ways. If one end is freed the tube winds sending a twist wave down the rod. Alternatively, if the ends are held but brought towards each other, the tube will relax by supercoiling with the *same* handedness as the intrinsic twist. This is the behavior observed in *Bacillus subtilis* and DNA. In the latter case the "intrinsic" twist corresponds to the right-handed helical architecture and in the former it is believed to be related to either the cell wall architecture or anisotropy.

In order to model *Bacillus subtilis* each bacterial cell is assumed to possess an "intrinsic" twist and to make up part of an elastic filament. Reproductive growth of all the cells in the bacterial filament results in an exponential growth of its length accompanied by a reduction of twist density.

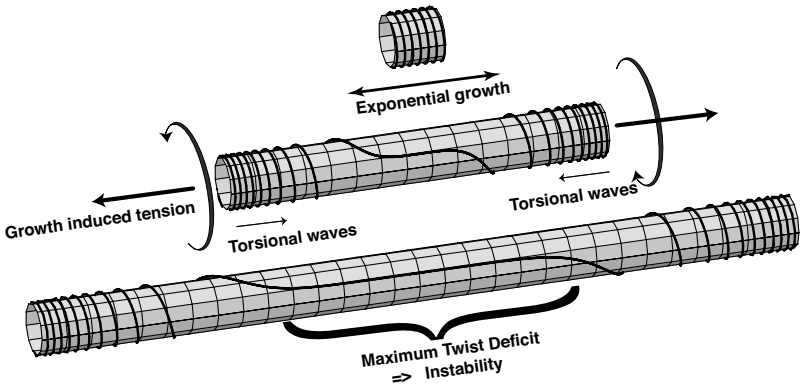


Fig. 8. Exponential growth and linear torsional wave competition, as the filament growth exponentially, the linear twist wave propagating from the end do not reach the middle of the filament where the twist deficit is maximal.

Using Kirchhoff's model as a starting point, we can study the dynamical properties of filaments and generalize it, as explained in the previous section, to include the effect of growth. Mathematical results on the stability of thin elastic rods [27] have enabled us to develop a complete picture of the mechanism of self-assembly in *Bacillus subtilis* [26] as well as computer simulation [41, 42]. Among other things the model gives quantitative predictions about the self-assembly geometry, such as the way the loop size scales with environmental conditions. The computer simulation of growing rod with intrinsic twist predicts the formation of looping in filaments remarkably similar to the ones found in the experiments (See Fig. 9)

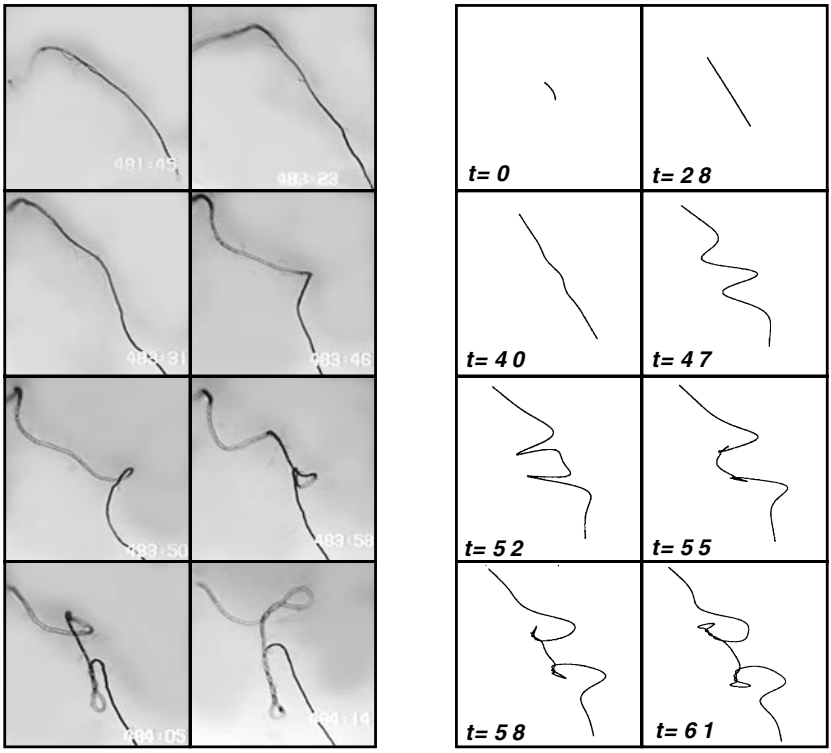


Fig. 9. Writhing Dynamics of *B. subtillis*. Left: experiments (Courtesy of Neil Mendelson), Right: Simulation of Kirchhoff rods with growth [41, 42].

A great deal of experimentation has shown that the twist state and helix hand of *B. Subtillis* macrofibers stem from the individual cell from which the fiber is derived. The information required to control macrofiber morphogenesis appears to reside in the growth plan of this cell and all its descendants. Intrinsic twist, a key feature of the dynamic model described here, is a logical

candidate for the mechanical information in the growth plan that dictates all subsequent growth and form.

Although our discussion here has focussed on the behavior of *B. subtillis*, filamentary structures are common in biological materials encompassing scales from molecular to organismal and we believe that the types of arguments forwarded here; namely the importance of axial growth, the decomposition of strain variables, buckling instability, twist-to-writhe conversion as a dynamical process and the special role played by intrinsic twist (or, presumably in other contexts, intrinsic curvature), may have quite general applicability.

4 Three-dimensional Growth

We now turn our attention to a general formulation of growth for three-dimensional nonlinear elastic body.

4.1 Basic Definitions of Morphoelasticity

Consider a continuous body with *reference configuration* $\mathcal{B}_0$. Let $\mathbf{X}$ denote the position vectors in $\mathcal{B}_0$. Now suppose the body is deformed to a new configuration, $\mathcal{B}_1$. We refer to $\mathcal{B}_1$ as the *current configuration* where the body is defined by $\mathbf{x} = \boldsymbol{\chi}(\mathbf{X}, t)$. The deformation gradient, $\mathbf{F}$(X,t)=Grad $\boldsymbol{\chi}$, relates a material segment in the reference configuration to the same segment in the current configuration. The key idea[2]introduced by Rodriguez, Hoger, and McCulloch [60] is to decompose the total deformation into a growth tensor $\mathbf{G}$(X,t) and an elastic tensor $\mathbf{A}$(X,t)

$$\mathbf{F}(\mathbf{X}, t) = \mathbf{A}(\mathbf{X}, t) \cdot \mathbf{G}(\mathbf{X}, t). \tag{61}$$

That is, the growth deformation may not result in a continuous change from point to point and may not be compatible. However, if we require continuity as the body grows, then an elastic deformation is introduced to maintain compatibility. As shown in Fig. 10, the growth tensor $\mathbf{G}$(X,t) maps $\mathcal{B}_0$ to the virtual configuration V_1 which is locally stress-free. The elastic deformation then maps V_1 to a grown stressed state $\mathcal{B}_1$ in order to maintain continuity of the body. The overall deformation gradient is the composition. Because of the need to ensure compatibility, the elastic deformation is introduced which in turn causes residual stress in the body.

[2] We also note that this decomposition was earlier introduced by Lee [44] who proposed representing elasto-plastic deformations in the form $\mathbf{F} = \mathbf{F}_e\mathbf{F}_p$, where $\mathbf{F}_e$ denotes the contribution to the total deformation as a result of elastic deformation, and $\mathbf{F}_p$ denotes the contribution to the total deformation due to plastic deformation. This has become a classical tool of elasto-plasticity. Note however, that there are still fundamental problems associated with such multiplicative decomposition (see for instance [55] or [77]).

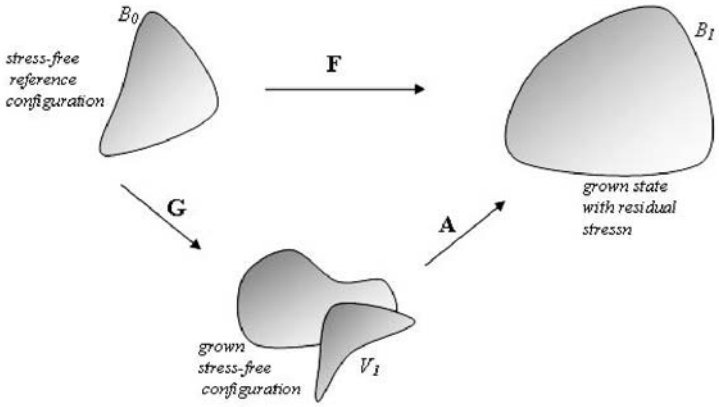


Fig. 10. The decomposition of finite growth. The deformation gradient $\mathbf{F}$ maps the reference configuration $B_0(\mathbf{X},t)$ into the current configuration $\mathcal{B}_1$. $\mathbf{F}$ can be represented as the product of a growth tensor $\mathbf{G}$ and an elastic deformation tensor $\mathbf{A}$. The intermediate configuration V_1 is a virtual state because $\mathbf{G}$ may not maintain continuity.

4.2 Strain Rate

We first recap some of the standard formalism and terminology. The tensor $\mathbf{F}$ is the *deformation gradient*, and

$$\mathbf{E} = \frac{1}{2}(\mathbf{F}^T\mathbf{F} - I), \tag{62}$$

is the Green (Lagrangian) strain tensor, and

$$\mathbf{D} = \mathbf{A} - I, \tag{63}$$

is the *displacement gradient*. Thus

$$\mathbf{E} = \frac{1}{2}(\mathbf{D} + \mathbf{D}^T + \mathbf{D}^T\mathbf{D}). \tag{64}$$

The linearization of this, $\mathbf{E} = \frac{1}{2}(\mathbf{D} + \mathbf{D}^T)$ is, of course, the familiar strain tensor of linear elasticity theory. If $\dot{(\,)}$ denotes differentiation w.r.t. time for fixed reference coordinate values, X, then one may show that the *rate of deformation*

$$\dot{\mathbf{F}} = \boldsymbol{\Gamma}\mathbf{F}, \tag{65}$$

where

$$\boldsymbol{\Gamma} = \operatorname{grad}\mathbf{v}(x,t), \tag{66}$$

is the Eulerian velocity gradient tensor, and hence

$$\dot{\mathbf{E}} = \mathbf{F}^T \Sigma \mathbf{F}, \tag{67}$$

where

$$\boldsymbol{\Sigma} = \frac{1}{2}(\boldsymbol{\Gamma} + \boldsymbol{\Gamma}^T), \tag{68}$$

is the *Eulerian strain rate tensor* [56]. Because $\boldsymbol{\Sigma} = \mathbf{F}^{-T}\dot{\mathbf{E}}\mathbf{F}^{-1}$ it is not a direct time derivative of $\mathbf{E}$, and thus cannot be called a (true) *rate of strain* tensor. However to first order

$$\dot{\mathbf{E}} = \mathbf{A}^T \boldsymbol{\Sigma} \mathbf{F} \simeq (I + \mathbf{D})^T \boldsymbol{\Sigma}(I + \mathbf{D}) \simeq \boldsymbol{\Sigma} + \text{h.o.t}, \tag{69}$$

and thus, for the linear theory, we can call $\boldsymbol{\Sigma}$, a rate of strain tensor.

Equation(61) is the starting point for our own discussions of the combination of elasticity and growth in a three-dimensional setting. By analogy with elasto-plasticity, we will refer to this approach as the theory of *morphoelasticity*. It is worth recalling how it was originally introduced [60]. For a system with density, ρ, the rate of growth of mass per unit volume, V, is

$$\dot{m} = \frac{d(\rho V)}{dt}, \tag{70}$$

while by basic conservation

$$\dot{m} = \frac{\pi \rho}{\pi t} + \text{div}(\rho \mathbf{v}_g), \tag{71}$$

where v_g is the *growth velocity vector*. It should be noted that v_g is defined in the Eulerian frame. For constant density these two equations combine to give

$$\frac{dV}{dt} = \text{div}\mathbf{v} = Tr\mathbf{D}_g, \tag{72}$$

where $\mathbf{D}_g$ is the *rate of growth tensor* - which is analogous to the rate of deformation tensor in classical continuum mechanics. Since $\mathbf{D}_g$ is Eulerian it has the advantage of not requiring a reference configuration but, it can be related to a Lagrangian *rate of growth stretch tensor*, $\dot{U}_g$, through

$$\mathbf{D}_g = \frac{1}{2}\left(\dot{\mathbf{U}}_g \mathbf{U}_g^{-1} + U_g^{-1}\dot{\mathbf{U}}_g\right). \tag{73}$$

Since $\dot{\mathbf{U}}_g$ is Lagrangian, the actual growth stretch tensor $\mathbf{U}_g$ is simply given by

$$\mathbf{U}_g = \int^t \dot{\mathbf{U}}_g dt. \tag{74}$$

The growth stretch tensor $\mathbf{U}_g$ is related to the growth deformation gradient tensor G by the right polar decomposition

$$\mathbf{G} = \mathbf{R}_g \mathbf{U}_g. \tag{75}$$

A general argument is that, without loss of generality, one can set $\mathbf{R}_g = I$ and work with $\mathbf{G} = \mathbf{U}_g$. [3]

[3] It appears to us that the reason given by many authors about how the rotation part of $\mathbf{G}$ can always be absorbed in the rotational part of $\mathbf{F}$, and that $\mathbf{G}$

4.3 Cauchy Stress and Equations of Motion

The forces distributed on a body $\mathcal{B}_1$ include a contact-force density $\mathbf{t}_n$ and a body-force density $\mathbf{b}$. In accordance with Euler's laws of motion, the balance of linear momentum is written as

$$\int_{\mathcal{B}_1} \rho(\mathbf{x},t)\mathbf{b}(\mathbf{x},t)\mathrm{d}v + \int_{\partial\mathcal{B}_1} \mathbf{t}_n \mathrm{d}a = \int_{\mathcal{B}_1} \rho(\mathbf{x},t)\dot{\mathbf{v}}(\mathbf{x},t)\mathrm{d}v. \tag{76}$$

Cauchy's theorem states that if $\mathbf{t}_n$ is continuous in $\mathbf{x}$, then $\mathbf{t}_n$ depends linearly on the unit normal $\mathbf{n}$. In other words, there exists a linear transformation $\mathbf{T}$ independent of $\mathbf{n}$ such that

$$\mathbf{t}_n = \mathbf{T}\mathbf{n}, \tag{77}$$

where $\mathbf{T}$ is referred to as the Cauchy stress tensor. Using (77) and applying the divergence theorem to (76) leads to

$$\int_{\mathcal{B}_1} \left(\rho(\mathbf{x},t) + \mathrm{div}\mathbf{T} - \rho(\mathbf{x},t)\dot{\mathbf{v}}(\mathbf{x},t)\right) \mathrm{d}v = 0. \tag{78}$$

Equation (78) is valid for any body $\mathcal{B}_1$. This leads to Cauchy's first law of motion,

$$\mathrm{div}(\mathbf{T}) + \rho\mathbf{b} = \rho\dot{\mathbf{v}}. \tag{79}$$

If the body is at rest, that is $\mathbf{v}(\mathbf{x},\mathrm{t})=\mathbf{0}$ for all $\mathbf{x}\in\mathcal{B}_1$, the equilibrium equation becomes

$$\mathrm{div}(\mathbf{T}) + \rho\mathbf{b} = 0. \tag{80}$$

Furthermore, if body forces are absent, (80) reduces to

$$\mathrm{div}(\mathbf{T}) = 0. \tag{81}$$

4.4 Strain-energy Functions

We assume that our material is hyperelastic. That is, there exists a strain-energy function $W = W(\mathbf{F})$ from which the stresses can be derived.

$$\mathbf{T} = J^{-1}\mathbf{F}\frac{\partial W}{\partial \mathbf{F}} - p\mathbf{1}, \tag{82}$$

where $J = det(\mathbf{F})$ is equal to one in the incompressible case and p is a Lagrange multiplier associated with the internal constraint of incompressibility ($p = 0$ in the compressible case). Many different general functional forms have been proposed for the the response of elastic materials under loads [6,62,65]. Here, we show some typical functions that have been proposed to model either elastomers or soft tissues. For the sake of simplicity, we restrict our attention

needs to be diagonal to ensure objectivity is not satisfactory and requires further discussions–see [77].

to homogeneous isotropic materials. The energy can be written in terms of the principal stretches λ_1, λ_2, λ_3 (the square roots of the principal values of $\mathbf{F}^{\mathrm{T}}\mathbf{F}$) or, equivalently for incompressible solids, in terms of the first two principal invariants of the Cauchy-Green strain tensors,

$$I_1 = \lambda_1^2 + \lambda_2^2 + \lambda_3^2, \quad I_2 = \lambda_2^2\lambda_3^2 + \lambda_3^2\lambda_1^2 + \lambda_1^2\lambda_2^2. \tag{83}$$

An essential property of many biological material is the strain-stiffening property which can be modeled either by algebraic power dependence (one-term Ogden model), by exponential behavior (as in the popular Fung model), or by limited chain extensibility (Gent model [19, 35, 38]). These three models can be written with a single parameter (ν, α, β, respectively) such that the classical neo-Hookean model is obtained in the limits $\nu \to 2$, $\alpha \to 0$, or $\beta \to 0$. Additionally, we also use the classical Mooney-Rivlin strain-energy density, often used to model elastomers.

Name	Definition	soft tissues	elastomers
neo-Hookean	$W_{\mathrm{nh}} = \frac{1}{2}(I_1 - 3)$		
Mooney-Rivlin	$W_{\mathrm{mr}} = \frac{(I_1 - 3) + \mu(I_2 - 3)}{2(1+\mu)}$		
1-term Ogden	$W_{\mathrm{og}} = \frac{2(\lambda_1^\nu + \lambda_2^\nu + \lambda_3^\nu - 3)}{\nu^2}$	$\nu \geq 9$	$\nu \approx 3$
Fung	$W_{\mathrm{fu}} = \frac{\exp \alpha(I_1 - 3) - 1}{\alpha}$	$3 < \alpha < 20$	
Gent	$W_{\mathrm{ge}} = \frac{-\log[1 - \beta(I_1 - 3)]}{\beta}$	$0.4 < \beta < 3$	$\beta < 0.05$

Table 1. A list of strain-energy functions. Note that the materials share the same infinitesimal shear modulus, which without loss of generality was taken equal to one. The limits $\mu \to 0$, $\alpha \to 0$, $\beta \to 0$, $\nu \to 2$ all lead to the neo-Hookean potential. References: 1-term Ogden [5, 66], Fung [13, 34] Gent [19, 20, 36, 37].

4.5 Constitutive Theory for G

There is an interesting discussion as to whether $\mathbf{G}$ or $\dot{\mathbf{G}} = \dot{\mathbf{U}}_g$ (the Lagrangian rate of growth tensor) can/should be a function of the Cauchy stress, namely

$$\mathbf{G} = f(\mathbf{T}), \quad \text{or} \quad \dot{\mathbf{G}} = g(\mathbf{T}). \tag{84}$$

Fung suggested that there might exist a growth equilibrium stress state $\bar{\mathbf{T}}$ at which the growth rate would be zero, *i.e.*

$$\dot{\mathbf{G}} = g(\mathbf{T} - \mathbf{R}\bar{\mathbf{T}}\mathbf{R}^T) \tag{85}$$

where the rotation tensor $\mathbf{R}$ comes from the polar decomposition and is required to ensure that both T and $\bar{T}$ are measured in the same frame of reference as required by objectivity.

There have been early on attempts to use the morphoelasticity formalism to model simple situations and understand the effect of growth and the feedback due to stress. These include the following cases:

Constant Growth

The simplest choice for $\mathbf{G}$ is to consider a constant tensor. A constant diagonal tensor $\mathbf{G}$ has been used in spherical geometry by Hoger and co-workers [9,43]. This case is interesting since analytical results can be obtained corresponding to small increments and explicit values of residual stress computed for growth without loading. In [4], the stability of such growing shells is considered.

Position Dependent Growth

Many growth processes depend on the location in the material. This effect is sometimes referred to as *differential growth* to indicate that some parts of a tissue grow faster than others. In morphoelasticity, it implies that $\mathbf{G}$ is a function of either $\mathbf{X}$ or $\mathbf{x}$. Both situations are of interest. In the first case, growth is a function of material points $\mathbf{X}$ in the reference configuration and this dependence assumes that the material is made out of points that grow at different rates and keep growing differentially as time goes by. In the second case, the ability of a tissue to grow depends on its location at any given time. This is the case, for instance, when cell reproduction depends on the availability of some nutrients that diffuse through the boundary. At any given time, the amount of nutrient may be described by the distance to the boundary as in the case in the growth of spheroids in tumor experiments [33]. The stability analysis of differentially growing shells was considered in [24].

Stress-dependence

It has been recognized experimentally and theoretically in many systems (such as aorta, muscles and bones) that one of the main biomechanical regulator of growth is stress [18,39,59,60,72–74]. It has even been suggested that stresses on cell walls play the role of a pacemaker for the collective regulation of tissue growth [67]. Accordingly, the growth rate tensor should be a function of the Cauchy stress tensor which could also vary according to the position of tissue elements in the reference configuration.

Most growth laws are of a phenomenological nature (See however the discussion in [18]) and currently there is no established theory of how they can be derived from biophysical principles An intriguing contribution appeared in [14] where it is assumed that growth can be associated with mechanical

accretive forces. In this context, the energy stored in the growth process can be transformed into mechanical elastic energy. By neglecting the possibility of energy sources and under suitable assumptions, a dissipation principle for the growth law can be derived, which in turn leads to a constitutive equation for the growth rate

$$\mathbf{G}^{-1}.\dot{\mathbf{G}} = \mathbf{M}_0 - W\mathbf{I} + A^T W_{\mathbf{A}} \tag{86}$$

where, as before, W is the free energy. We note that $W\mathbf{I} - A^T W_{\mathbf{A}}$ has the form of an Eshelby tensor as found in the theory of elasto-plasticity [15,49], and that for small deformations, $\mathbf{M}_0$ plays the role of a constant (homeostatic) stress. These constitutive relations (shown here in the simplest case) are typically nonlinear and deserve further exploration.

4.6 Cumulative Growth

Now consider a sequence of growth steps in which each step can be decomposed into a growth deformation and an elastic deformation (see Fig. 11).

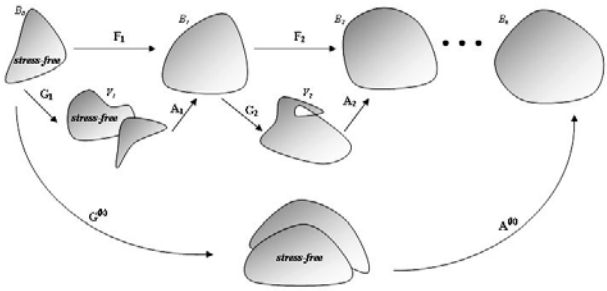


Fig. 11. Cumulative growth with k steps.

The cumulative deformation gradient is given by

$$\begin{aligned}\mathbf{F}^{(k)} &= \mathbf{F}_k \cdot \mathbf{F}_{k-1} ... \mathbf{F}_2 \cdot \mathbf{F}_1 \\ &= \mathbf{A}_k \cdot \mathbf{G}_k \cdot \mathbf{A}_{k-1} \cdot \mathbf{G}_{k-1} ... \mathbf{A}_2 \cdot \mathbf{G}_2 \cdot \mathbf{A}_1 \cdot \mathbf{G}_1.\end{aligned}$$

Assume the growth and elastic tensor commute, that is $\mathbf{A}_i \cdot \mathbf{G}_j = \mathbf{G}_j \cdot \mathbf{A}_i$, for all i, j. Then the elastic and growth tensors can be written as

$$\mathbf{A}^{(k)} = \mathbf{A}_k \cdot \mathbf{A}_{k-1} ... \mathbf{A}_2 \cdot \mathbf{A}_1, \qquad \mathbf{G}^{(k)} = \mathbf{G}_k \cdot \mathbf{G}_{k-1} ... \mathbf{G}_2 \cdot \mathbf{G}_1. \tag{87}$$

The stress in $\mathcal{B}_k$ is

$$\nabla_{x_k} \cdot (\mathbf{T}_k) = 0, \tag{88}$$

where

$$\mathbf{T}_k = \mathbf{A}^{(k)} \cdot \partial_{\mathbf{A}^{(k)}} W - p_k \mathbf{1}. \tag{89}$$

An example of a cumulative process of growth is given in Section 4.9.

4.7 A Simple Example of Homogeneous Growth

As a first example, we revisit the model of Rodriguez, *et al.* [60] where growth is a function of the stress tensor. Because the problem is homogeneous, residual stress is absent. A rectangular block of bone is subjected to compression along its longitudinal axis. It grows along the x and y directions as a linear function of the difference between the axial stress and a no-growth equilibrium stress state. The bone is assumed to have a Young's modulus of 18.4 GPa along the z-axis. According to the authors "...at each step the axial strain is adjusted so that the applied axial *force* remains constant." Recalling that stress = force/area, the calculation proceeds as follows. At each step the cross sectional area $S(t) = A_{xx}(t)A_{yy}(t) = \lambda_{xx}(t)\lambda_{xx}(t)/\lambda_z(t)$ is computed. In order to ensure that the axial force, $F = T_{zz}(t)S(t)$, where T_{zz} is a certain function of λ_z, remains constant, λ_z is reduced as necessary. This, in turn, reduces T_{zz} at the next step until it reaches $\bar{T}_{zz}$ at which point the computation stops. In this model the stresses are only as a result of loading so there is no residual stress.

More explicitly, the reference configuration is compressed along the longitudinal axis which results in the elastic deformation gradient

$$A = \mathrm{diag}\left(\frac{1}{\sqrt{\lambda_z}}, \frac{1}{\sqrt{\lambda_z}}, \lambda_z\right), \tag{90}$$

where λ_z is the stretch ratio corresponding to a 0.1% shortening ($\lambda_z = 0.999$). The strain is calculated from the Green strain tensor

$$\mathbf{E} = \frac{1}{2}(\mathbf{A}^T\mathbf{A} - \mathbf{I}), \tag{91}$$

and the longitudinal stress is found from

$$\mathbf{T} = \mathbf{A}^T\frac{\partial W}{\partial \mathbf{E}}\mathbf{A}, \tag{92}$$

where the strain energy is given by

$$W = c_1(E_{xx}^2 + E_{yy}^2 + E_{zz}^2). \tag{93}$$

The stress along the z-direction is then

$$\begin{aligned} t_3 &= 2c_1E_{zz}F_{zz}E_{zz} \\ &= c_1\lambda_z^2(\lambda_z^2 - 1), \end{aligned}$$

where $2c_1$ is the value of the Young's modulus along the z-direction. Following the elastic deformation, the tissue grows or resorbs along the x and y axes until equilibrium is restored. The equilibrium value for the longitudinal stress is given by t_3^*. The rate of growth in the x and y directions at time t is given by

$$\lambda_x(t) = k_x(t_3(t) - t_3^*), \tag{94}$$

$$\lambda_y(t) = k_y(t_3(t) - t_3^*), \tag{95}$$

where k_x and k_y are growth rate constants with equal values of -0.27 time^{-1}GPa^{-1} and the no-growth equilibrium longitudinal stress t_3^* is -4.5 MPa. The growth stretch ratios with respect to the reference configuration can be written as

$$\lambda_x(t + \mathrm{d}t) = \lambda_x(t) + \dot{\lambda}_x \mathrm{d}t, \tag{96}$$

$$\lambda_y(t + \mathrm{d}t) = \lambda_y(t) + \dot{\lambda}_y \mathrm{d}t, \tag{97}$$

where dt corresponds to one iteration. The growth deformation gradient is then

$$\mathbf{G}(t + \mathrm{d}t) = \mathrm{diag}\left(k_x(t_3(t) - t_3^*)\mathrm{d}t + G_{xx}(t), k_y(t_3(t) - t_3^*)\mathrm{d}t + G_{yy}(t), G_{zz}\right). \tag{98}$$

The axial stress is computed after each iteration and the tissue growth is then calculated. The applied axial load remains constant through each iteration and the axial stress is adjusted in order to maintain this force. Each growth deformation is assumed to be compatible and therefore no residual stress arises. Fig. 12 shows the axial stress and the growth stretch ratios (λ_x and λ_y) as functions of time. Following each iteration, the cross-sectional area increases which causes a decrease in the elastic stress and a decrease in the growth rate.

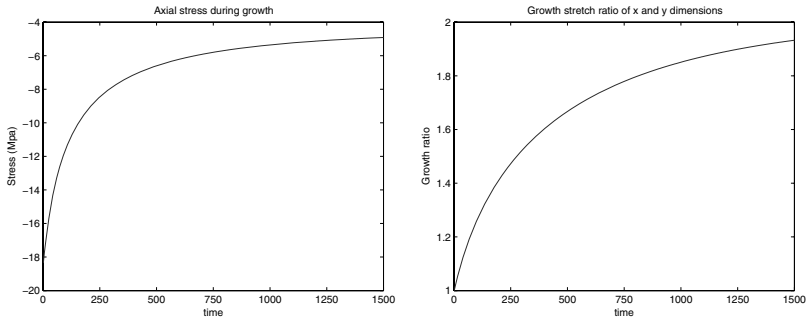


Fig. 12. A load leading to a 0.1% shortening on a block of bone is imposed. The block grows as a linear function of the difference between the loading stress and the no-growth equilibrium stress state. The left panel shows the axial stress as a function of time as it approaches the no-growth equilibrium stress. The right panel displays the growth ratios for the x and y directions.

The evolution of the axial stress can be written as a discrete mapping given by

$$t_3(t+\mathrm{d}t) = \frac{t_3(0)}{\left(\sqrt{\frac{t_3(0)}{t_3(t)}} + k_x(t_3(t) - t_3^*)\right)^2}. \tag{99}$$

The mapping can be converted to the following differential equation:

$$\frac{\mathrm{d}t_3(t)}{\mathrm{d}t} = \frac{t_3(0)t_3(t)}{\left(\sqrt{t_3(0)} + k_x\sqrt{t_3(t)}(t_3(t) - t_3^*)\right)^2} - t_3(t). \tag{100}$$

Fig. 13 shows the discrete mapping presented by Rodriguez *et al.*, as well as the numerical solution to Equation (100).

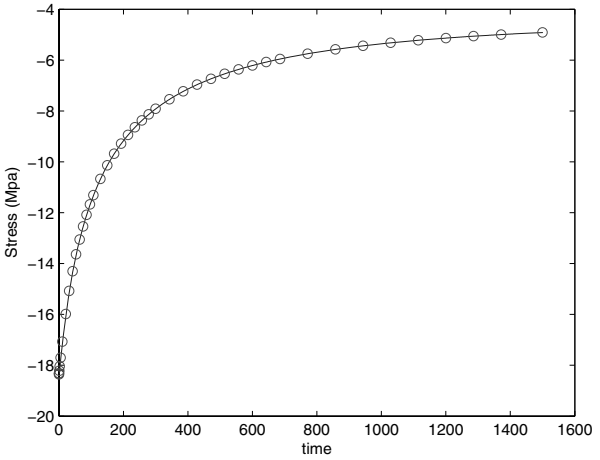


Fig. 13. Comparison between the continuous solution (solid curve) and the discrete solution (open circle) obtained by incremental computation.

To supplement the previous results, consider stress-dependent growth in the case of a cylindrical tube. The tube is initially subjected to compression along its longitudinal axis resulting in a 0.1% shortening ($\lambda_z = 0.999$). An elastic energy associated with a Neo-Hookean material is used. After axial compression, the cylinder is allowed to grow in which the following stress dependent growth deformation gradient is used:

$$\mathbf{G}(t+\mathrm{d}t) = \mathrm{diag}\left(k(t_3(t) - t_3^*)\mathrm{d}t + G_{rr}(t), k(t_3(t) - t_3^*)\mathrm{d}t + G_{\theta\theta}(t), G_{zz}\right). \tag{101}$$

where $k = -0.27\ \mathrm{time}^{-1}\mathrm{GPa}^{-1}$ and $t_3^* = -4.5$ MPa.

In the cylindrical case, the material response is not homogenous, and therefore the longitudinal force is now a function of the radial component. Therefore, consider the resultant load, $N = \int_a^b rt_3(r)dr$, and the resultant

longitudinal force $F_z = N \cdot Area$. In each iteration, the resultant axial stress is evaluated and the material is allowed to grow. The applied resultant axial force F_z is constant at each step and the the resultant longitudinal stress is adjusted accordingly to maintain the constant force. Fig. 14 shows the resultant axial stress as a function of time. The curve is approaching the equilibrium resultant stress, $N^* = \int_a^b rt_3^* dr$.

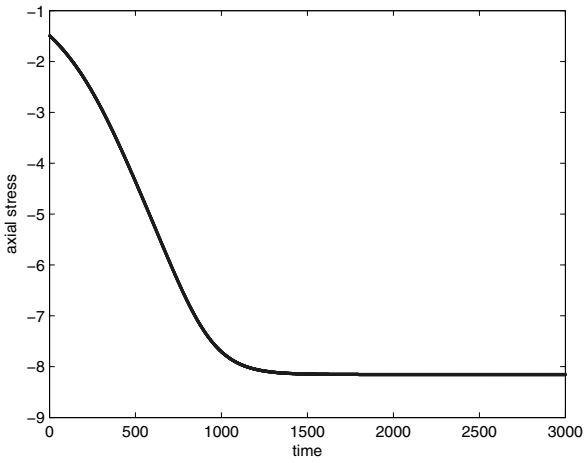


Fig. 14. Resultant longitudinal stress for an axially loaded cylinder. An initial five percent shortening was incurred. The cylinder was allowed to grow as the state of stress goes toward the predetermined equilibrium state.

We comment that the convenient thing about doing the calculation in discrete time steps is that one does not have to worry about separating elastic and growth time scales: one simply makes whatever elastic adjustments are necessary before implementing the next growth step.

4.8 Cylinder Growth: One Step Growth

We now consider a simple growth tensor to demonstrate how residual stress can arise from growth. Consider a cylindrical tube whose reference configuration has length L and internal and external radii A and B respectively. Therefore the tube is defined as

$$A \leq R \leq B, \qquad 0 \leq \Theta \leq 2\pi, \qquad 0 \leq Z \leq L. \tag{102}$$

Now let the cylinder undergo uniform circumferential growth or resorption which then results in an elastic deformation. The resulting deformation is given by

$$r = r(R), \qquad \theta = k\Theta, \qquad z = \lambda_z Z. \tag{103}$$

The deformation gradient in cylindrical coordinates is

$$\mathbf{F} = \mathrm{diag}(r', \frac{r}{R}, \lambda_z) \tag{104}$$

where $r' = \frac{dr}{dR}$. Assume there is no change in length so that $\lambda_z = 1$. The deformation tensor is decomposed as $\mathbf{F} = \mathbf{A}\cdot\mathbf{G}$. Assuming constant circumferential growth, the growth deformation gradient is

$$\mathbf{G} = \mathrm{diag}(1, k, 1), \tag{105}$$

where $\mathbf{G}$ maps the reference state $\mathcal{B}_0$ into the grown stress-free state V. In order to maintain continuity of the body, an elastic deformation

$$\mathbf{A} = \mathrm{diag}(\frac{1}{\alpha}, \alpha, 1), \tag{106}$$

maps the virtual stress-free state V to the final intact configuration $\mathcal{B}_1$. Note the incompressibility condition $\det\mathbf{A} = 1$ is used to express the three principal strain components in terms of a single variable α. Assuming the cylindrical tube is composed of a Fung material, the strain energy function is

$$W_{\mathrm{fu}} = \frac{c}{2}(e^Q - 1), \tag{107}$$

where c is a constant and Q is a function of the three principal strain values, that is $Q = 2b_1(\lambda_1^2 + \lambda_2^3 + \lambda_3^2)$. In the present case $\lambda_1 = 1/\alpha, \lambda_2 = \alpha$, and $\lambda_3 = 1$. Therefore, Q is given by

$$Q = 2b_2(\frac{1}{\alpha^2} + \alpha^2 - 2). \tag{108}$$

The incompressibility constraint is $\det(\mathbf{A}) = 1$, or equivalently $\det(\mathbf{F}\mathbf{G}^{-1}) = 1$. This implies $\det(\mathbf{F}) = \det\mathbf{G}$, so that

$$\frac{r'r}{R} = k. \tag{109}$$

The previous equation can be integrated to obtain

$$r = (a^2 + k(R^2 - A^2))^{1/2}. \tag{110}$$

The incompressibility constraint along with the relation $\mathbf{F} = \mathbf{A}\mathbf{G}$ provides an equation for the strain, $\alpha = r/(kR)$. The only non-vanishing equilibrium equation in (81) is

$$\frac{\partial t_1}{\partial r} + \frac{t_1 - t_2}{r} = 0. \tag{111}$$

The radial stress and hoop stress are denoted as $t_1 = T_{11}$ and $t_2 = T_{22}$, respectively. Using the stress-strain relationship in Equation (82), equations for the radial and hoop stress can be written as

$$t_1 = \frac{1}{\alpha}\frac{\partial W}{\partial \lambda_1} - p \tag{112}$$

$$t_2 = \alpha\frac{\partial W}{\partial \lambda_2} - p. \tag{113}$$

Substitute these values into (111) to obtain a closed equation for t_1:

$$\frac{\partial t_1}{\partial r} = \frac{\alpha}{r}\partial_\alpha \hat{W} \tag{114}$$

where $\hat{W} = W(\frac{1}{\alpha}, \alpha, 1)$. In terms of R,

$$\frac{\partial t_1}{\partial R} = \frac{\partial_\alpha \hat{W}}{k\alpha R}. \tag{115}$$

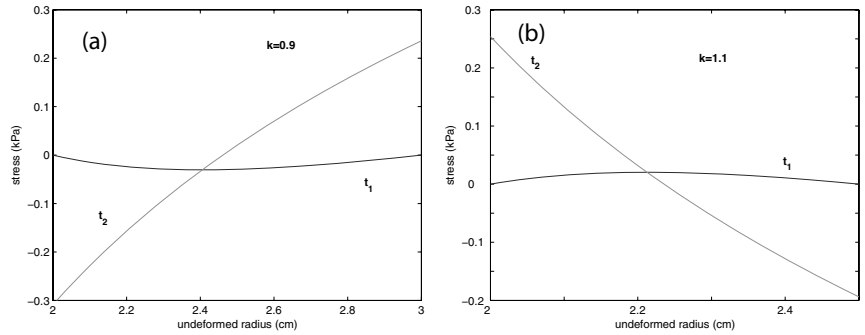


Fig. 15. Plots of residual stress vs. undeformed radius for a cylindrical tube following uniform circumferential resorption (a) and growth (b). The cylinder is unload which results in zero radial stress at the boundaries in both cases. When k=0.9 (resorption) the circumferential residual stress is in compression in the inner wall and in tension in the outer wall. When k=1.1 (growth) the circumferential residual stress is tensile in the inner layers and compressive in the outer layers.

Using the boundary conditions $t_1(A) = t_1(B) = 0$, integrate the last equation to obtain an equation for the radial stress

$$t_1(R) = \frac{1}{k}\int_A^R \frac{\partial_\alpha \hat{W}}{\alpha R} dR, \tag{116}$$

and now the hoop stress can be evaluated as $t_2 = t_1 + \frac{\alpha}{2}\partial_\alpha \hat{W}$. The hollow cylindrical tube model demonstrates how circumferential growth produces a transmural distribution of residual stress that would cause the cylinder to change shape when cut. Consider a tube with initial internal and external radii of 2.0 and 3.0 cm and $k = 0.9$ (resorption). Fig. 15(a) shows the radial

stress t_1 and the circumferential stress t_2. The radial stress is zero at the boundaries ($r = A$ and $r = B$) because the cylinder is unloaded. The circumferential stress is in compression in the inner layers and tension in the outer layers. The grown internal and external radii were 1.75 and 2.75 cm.

Now consider the growth case where $k = 1.1$. The equilibrium internal and external radii were 2.25 and 3.25 cm. Notice in Fig. 15(b) the graphs are reversed from the resorption case. The circumferential stress is in tension in the inner layers and compression in the outer layers. The longitudinal stress t_3 is nonzero in both the resorption and growth cases. The longitudinal stress will cause the cylinder to extend or shorten. However, the resultant stress is close to zero, and therefore the simplifying assumption of $\lambda_z = 1$ will not affect the circumferential residual stress. In the three-dimensional problem, we will later discuss how to circumvent the issue of a non-zero longitudinal stress on the ends of the cylinder.

4.9 Cylinder Growth: Modeling Incremental Growth

Consider once again a cylindrical tube but now assume the incremental growth tensor $\mathbf{G}_{inc}$ is a function of position. There is growth only along the z-axis, so that $\mathbf{G}_i = \text{diag}(1, 1, g(r_i))$. Because the incremental growth tensor is dependent on the current configuration, an implicit dependence on the stress tensor exists which must be computed at each iteration [23]. The growth function g_{inc} at the kth step along the z-axis is written as the product $g(R) = \prod_{i=1}^{k-1} g_{inc}(r_i)$ where r_i is the current configuration following the i-th deformation. The cumulative deformation gradient in cylindrical coordinates after i steps is

$$\mathbf{F}_i = \text{diag}(r_i', \frac{r_i}{R}, \lambda_{z_i}), \tag{117}$$

where the cylinder is extended uniformly to length $l_i = \lambda_{z_i} l_{i-1}$. Denote the internal and external radii in the initial configuration by A and B, and let $a_i = r_i(A)$ and $b_i = r_i(B)$ be the radii in the current configuration. The principal stretches of the elastic tensor $\mathbf{A}$ are

$$\lambda_{i1} = g_i(R)(\alpha_i \lambda_{z_i})^{-1}, \qquad \lambda_{i2} = \alpha_i = \frac{r_i}{R}, \qquad \lambda_{i3} = \frac{\lambda_{z_i}}{g_i(R)}, \tag{118}$$

where α_i is defined as the azimuthal principal stretch. Assume the elastic cylindrical tube is composed of a neo-Hookean material, that is $W_{\text{nh}} = \mu(\lambda_{i1}^2 + \lambda_{i2}^2 + \lambda_{i3}^2 - 3)$. The incompressibility condition $\det(\mathbf{A}_i) = 1$ implies $\det(\mathbf{F}_i) = \det(\mathbf{G}_i)$ so that

$$\lambda_{z_i} \frac{r_i' r_i}{R} = g(R). \tag{119}$$

Integrate the last equation to find

$$r_i(R) = \left(a_i^2 + \frac{2}{\lambda_{z_i}} \int_A^R \rho g(\rho) d\rho \right)^{1/2}. \tag{120}$$

The constitutive relationships for t_1 and t_2 are as follows:

$$t_1 = \lambda_{i1} \frac{\partial W}{\partial \lambda_{i1}} - p, \tag{121}$$

$$t_2 = \lambda_{i2} \frac{\partial W}{\partial \lambda_{i2}} - p. \tag{122}$$

The only nonvanishing equilibrium equation of $\nabla_{x_i} \cdot (\mathbf{T}_1) = 0$ is, as previously shown in (111),

$$\frac{\partial t_1}{\partial r} + \frac{t_1 - t_2}{r} = 0. \tag{123}$$

Rearrange the previous equation and use (121) and (122) to obtain

$$\frac{\partial t_1}{\partial r_i} = \frac{\alpha_i}{r_i} \partial_\alpha \hat{W} \tag{124}$$

where $\hat{W} = W(\frac{g(R)}{\alpha \lambda_{z_i}}, \alpha, \frac{\lambda_{z_i}}{g(R)})$. In terms of R,

$$\frac{\partial t_1}{\partial R} = \frac{R g(R) \alpha_i}{\lambda_{z_i} r_i^2} \partial_{\alpha_i} \hat{W} \tag{125}$$

where (120) and (118) are used to express $r_i = r_i(R)$ and $\alpha_i = \alpha_i(R)$. We would like the surface of the cylinder to be free of any traction,

$$t_1(a_i) = t_1(b_i) = 0, \quad 0 \le \theta \le 2\pi, \quad 0 \le z_i \le \lambda_{z_i} l_{i-1}, \tag{126}$$

and

$$t_3(0) = t_3(\lambda_{z_i} l_{i-1}) = 0, \quad 0 \le \theta \le 2\pi, \quad a_i \le r_i \le b_i. \tag{127}$$

However, $t_3 = t_3(R)$ and does not explicitly depend on z_i. Therefore (127) requires $t_3 = 0$, but t_3 is a function of R and is not equal to zero. Instead, at each iteration we may solve for λ_{z_i} so that a zero resultant load N is imposed [17, 30–32]

$$N_i(\lambda_{z_i}) = 2\pi \int_{a_i}^{b_i} r_i t_3(r_i, \lambda_{z_i}) dr_i = 0. \tag{128}$$

Therefore, the end conditions in (127) are replaced by the condition above. Using (120), (125), and (126), the longitudinal stretch λ_{z_i}, the deformed radius a_i, and the radial stress t_1 can be found at each iteration and the deformation is completely determined. At each stage the growth function along the z-axis can be computed for the next iterate,

$$g(R) = \prod_{i=1}^{k} g_{inc}(r_i). \tag{129}$$

First consider a linear incremental growth function $g_{inc}(r_i) = 1 + \mu(r_i - a_i)$ where no longitudinal growth occurs at the inner wall and the growth linearly increases toward the outer wall. Choose initial values $A = 1$, $B = 2$, and calculate μ at each iteration such that the volume increases by 1%. Next consider the growth function $g_{inc}(r_i) = 1 + \mu(b_i - r_i)$ where no longitudinal growth occurs at the outer wall and the growth linearly increases toward the inner wall. Fig. 16 shows the cumulative growth function in the current configuration.

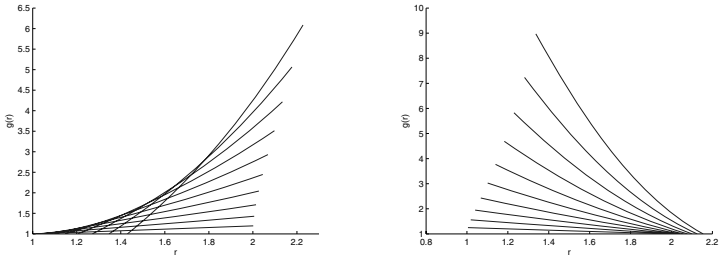


Fig. 16. The cumulative growth function viewed in the current frame $g = g(r)$. The volume increase in each step is 1%. The cumulative growth curve is plotted every ten steps.

Now consider changing μ and looking at how the longitudinal stress changes after one step. Fig. 17 shows $N(\lambda_z) = 2\pi \int_a^b r t_3(r, \lambda_z) dr$ as μ increases from zero to 1.5. In order to obtain a zero resultant load, we need to find $\lambda_{z_{crit}}$ such that $N(\lambda_z) = 0$. Once $\lambda_{z_{crit}}$ is found, Fig. 17 shows the longitudinal stress. When the inner layers of the cylinder grow faster than the outer layers, the inner layers are in compression and the outer layers are in tension as predicted. In contrast, when the outer layers of the cylinder grow faster than the inner layers, the inner layers are tensile and the outer layers are compressive.

4.10 Cylinder Growth: Embedded in an Elastic Medium

We can also nest two materials with different strain functions (i.e. a neo-Hookean material inside Fung material) or different growth factors. The boundary between the two materials, which we call C and c in the reference and current configurations respectively, is constrained so that no gaps form between materials.

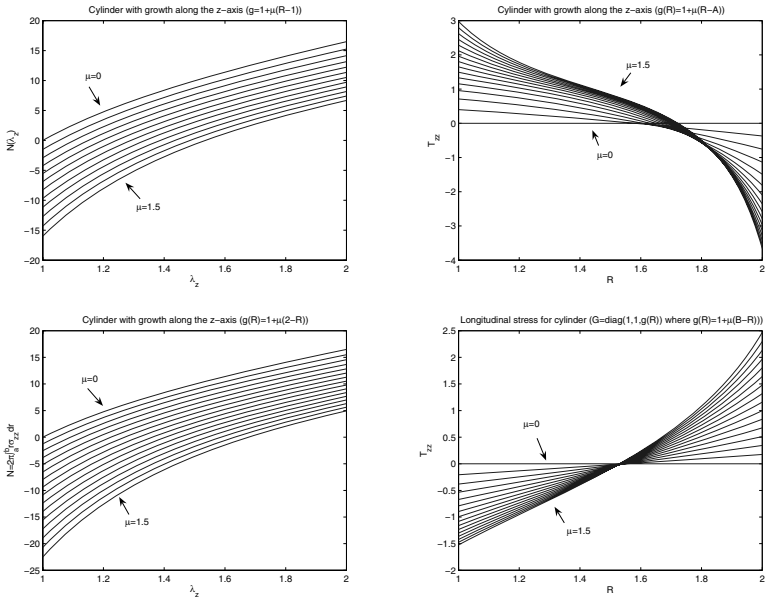


Fig. 17. Longitudinal stress along the cylinder radius (right). Note the compressive and tensile stress in different regions depending on which part of the cylinder grows faster.

As an example we look at two different materials. The inner material is neo-Hookean, and subject to shrinking in the radial direction ($\gamma_1 = 0.5$), while the outer material is Fung and is growing radially ($\gamma_1 = 2$). We set external pressure to zero, and set the initial boundary between the materials to halfway between the shell boundaries ($C = 1.5$).

Fig. 8 and Fig. 9 show the results for this configuration. As can be seen in the first figure, the deformation $r(R)$ is continuous, but the position of the boundary between layers is no longer halfway between the edges, rather it is closer to the inner deformed radius a. The second figure shows the stresses in the material due to growth, and there exist tensile and compressive stresses on all axes.

4.11 Cylinder Growth: with Twist

We now comment on growth models involving twist. The decomposition $\mathbf{F} = \mathbf{AG}$ is not unique and some thought needs to go into making a reasonable choice. Our model is that of a cylindrical rod that grows in length and exhibits growth induced twist, but maintains a constant radius. This would correspond to the (total) deformations

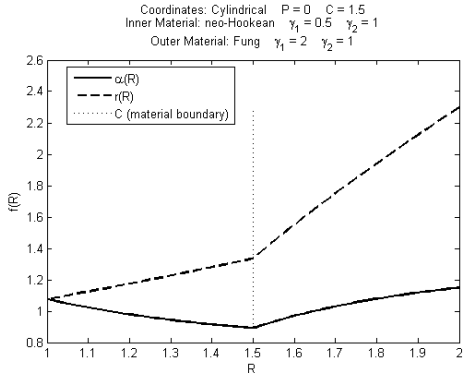


Fig. 18. The strain function $\alpha(R)$ and the deformation $r(R)$ for two nested materials (see text).

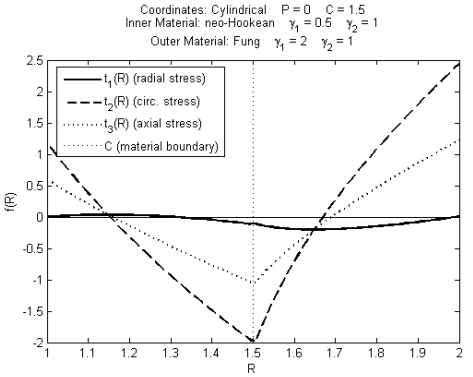


Fig. 19. The radial, circumferential and axial stresses for two nested materials (see text).

$$r = R \tag{130}$$

$$\theta = \Theta + z\tau = \Theta + \lambda Z\tau \tag{131}$$

$$z = \lambda Z. \tag{132}$$

Note that the torsion in the current configuration depends on the current height z. With these deformations, the (total) deformation gradient tensor is

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & r\lambda\tau \\ 0 & 0 & \lambda \end{pmatrix}. \tag{133}$$

For our growth tensor we choose

$$\mathbf{G} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & r\tau^* \\ 0 & 0 & \lambda \end{pmatrix}, \tag{134}$$

which corresponds to the growth-induced deformation

$$r = R \tag{135}$$
$$\theta = \Theta + Z\tau^* \tag{136}$$
$$z = \lambda Z, \tag{137}$$

where τ^* is a (fixed) twist determined by the growth process. Note that in the growth process we claim that the amount of torsion depends on the reference configuration height Z. The elastic deformation gradient tensor is easily determined to be

$$\mathbf{F} = \mathbf{A}\mathbf{G}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & r\left(\tau - \frac{\tau^*}{\lambda}\right) \\ 0 & 0 & 1 \end{pmatrix}. \tag{138}$$

With this form of $\mathbf{B}$, the Cauchy elastic stress tensor is

$$\mathbf{T}(\mathbf{F}) = \begin{pmatrix} T_{11} & 0 & 0 \\ 0 & T_{22} & T_{23} \\ 0 & T_{32} & T_{33} \end{pmatrix}, \tag{139}$$

where

$$T_{11} = P + \Phi + \Psi\left(2 + r^2\left(\tau - \frac{\tau^*}{\lambda}\right)^2\right) \tag{140}$$
$$T_{22} = P + \Phi + 2\Psi + (\Phi + \Psi)r^2\left(\tau - \frac{\tau^*}{\lambda}\right)^2 \tag{141}$$
$$T_{23} = (\Phi + \Psi)r\left(\tau - \frac{\tau^*}{\lambda}\right) \tag{142}$$
$$T_{32} = (\Phi + \Psi)r\left(\tau - \frac{\tau^*}{\lambda}\right) \tag{143}$$
$$T_{33} = P + \Phi + 2\Psi, \tag{144}$$

where P is the turgor pressure and $\Phi = 2\pi W/\pi I_1$, $\Psi = 2\pi W/\pi I_2$. For a Mooney-Rivlin material, Φ and Ψ are constant.

If we require that during growth the process is free of (residual) torsional stress, *i.e.* $T_{\theta z} = 0$, then the twist must obey

$$\tau = \frac{\tau^*}{\lambda}, \tag{145}$$

which also eliminates the elastic torsional strain, *i.e.* $\mathbf{F}_{\theta z} = 0$. If our growth model allows for (gradually) increasing length, *i.e.* λ increases with time, *then*

condition (145) shows that as the cylinder grows it gradually unwinds. Under this condition, all the other stress

$$T_{rr} = T_{\theta\theta} = T_{zz} = P + \Phi + 2\Psi, \tag{146}$$

only depend on the load (the turgor press P) and hence there is no residual stress at all.

5 Conclusions

In this article, we have reviewed different models to describe growth in biological systems. The common thread to most of these models is the decomposition of strain variables in elastic and growth components. The elastic component is connected to the stresses by the usual constitutive equation whereas the strain associated with growth requires a separate evolution law. These laws are not yet well-understood and much experimental and theoretical work is needed before a clear picture of how growth is related to stress emerges. Nevertheless, it is already possible to explore the consequences of growth such as the ability of a growing body to either build mechanical properties or undergo pattern formation through a buckling instability.

An area in which growth modeling will play an increasingly important role is in the modeling of tumor growth and cancer. We would just like to mention the existence of growing body of literature on various modeling aspects and data analysis of the problem that will drive the theory of growth (see [1, 3, 7, 8, 16, 33, 46, 54, 61, 63, 75]).

It is important to note that the description given here is not the only possible approach to modeling growth. Other interesting approaches to growth have been proposed either in terms of mixture theory, coupled with the evolution of natural configurations [40] or by focusing on the evolution of residual stress in the material [57].

Acknowledgement. This material is based upon work supported by the National Science Foundation under Grant No. DMS-0604704 (A. G.) and DMS-IGMS-0623989 (A.G. and M. T.). We are indebted to Martine Ben Amar for many interesting discussions and for her hospitality at the Ecole Normale Supérieure (Paris). This work was possible by the Centre National de la Recherche Scientifique (A. G. and M.T.).

References

1. D. Ambrosi and F. Mollica. The role of stress in the growth of a multicell spheroid. *J. Math. Biol.*, 48: 477–499, 2004.
2. S. S. Antman. *Nonlinear problems of elasticity.* Springer Verlag, New York, 1995.

3. R. P. Araujo and D. L. S. McElwain. New insights into vascular collapse and growth dynamics in solid tumors. *J. Theor. Biol.*, 228: 335–346, 2004.
4. M. Ben Amar and A. Goriely. Growth and instability in soft tissues. *J. Mech. Phys. Solids*, 53: 2284–2319, 2005.
5. D. K. Bogen and Th A. McMahon. Do cardiac aneurysms blow out? *Biophys. J.*, 27: 301–316, 1979.
6. M. C. Boyce and E. M. Arruda. Constitutive models for rubber elasticity: a review. *Rubber Chem. Technol.*, 73: 504–523, 2000.
7. H. Byrne and L. Preziosi. Modelling solid tumour growth using the theory of mixtures. *Math. Med. Biol.*, 20: 341–366, 2003.
8. C. Y. Chen, Byrne H. M., and J. R. King. The influence of growth-induced stress from the surrounding medium on the devlopment of multicell spheroids. *J. Math. Biol.*, 43: 191–220, 2001.
9. Y. Chen and A. Hoger. Constitutive functions of elastic materials in finite growth and deformation. *J. Elasticity*, 59: 175–193, 2000.
10. B. D. Coleman, E. H. Dill, M. Lembo, Z. Lu, and I. Tobias. On the dynamics of rods in the theory of Kirchhoff and Clebsch. *Arch. Rational Mech. Anal.*, 121: 339–359, 1993.
11. D. J. Cosgrove. Cell wall yield properties of growing tissue. evaluation by in vivo stress relaxation. *Plant Physiol.*, 78: 347–356, 1985.
12. Ch Darwin. *The Movements and Habits of Climbing Plants*. Appleton, New York, 1888.
13. A. Delfino, N. Stergiopulos, J. E. Moore, and J. J Meister. Residual strain effects on the stress field in a thick wall finite element model of the human carotid bifurcation. *J. Biomech.*, 30: 777–786, 1997.
14. A. DiCarlo and S. Quiligotti. Growth and balance. *Mech. Res. Comm.*, 29: 449–456, 2002.
15. M. Epstein and G. Maugin. Thermomechanics of volumetric growth in uniform bodies. *Int. J. Plasticity*, 16: 951–978, 2000.
16. H. Frieboes, X. Zheng, C. H Sun, B. Tromberg, R. Gatenby, and V. Cristini. An integrated computational/experimental model of tumor invasion. *Cancer Research*, 66: in press, 2006.
17. Y. B. Fu and R. W. Ogden. *Nonlinear Elasticity. Theory and applications*. Cambridge University Press, Cambridge, 2001.
18. Y. C Fung. Stress, strain, growth, and remodeling of living organisms. *Z. angew. Math. Phys.*, 46 (special issue): S469–S482, 1995.
19. A. Gent. A new constitutive relation for rubber. *Rubber Chem. and Technol.*, 69: 59–61, 1996.
20. A. N. Gent. Elastic instabilities in rubber. *Int. J. Non-Linear Mech.*, 40: 165–175, 1995.
21. R. Goldstein and A. Goriely. The morphoelasticity of tendrils. *Phys. Rev. E*, 74: #010901, 2006.
22. B. C. Goodwin and C. Briere. A mathematical model of cytoskeletal dynamics and morphogenesis in acetabularia. In D. Menzel, editor, *The Cytoskeleton of the Algae*, 219–233. CRC Press, Boca Raton, 1992.
23. A. Goriely and M. Ben Amar. On the definition and modeling of incremental, cumulative, and continuous growth laws in morphoelasticity. *Biomechanics and Modeling in Mechanobiology*, To be published, 2006.
24. A. Goriely and M. Ben Amar. Differential growth and instability in elastic shells. *Phys. Rev. Lett.*, 94: #198103, 2005.

25. A. Goriely and S. Neukirch. The mechanics of climbing and attachment in twining plants. *Phys. Rev. Lett.*, To be published: 1–4, 2006.
26. A. Goriely and M. Tabor. Nonliner dynamics of filaments. *Nonlinear Dynamics*, 21: 101–133, 2000.
27. A. Goriely and M. Tabor. New amplitude equations for thin elastic rods. *Phys. Rev. Lett.*, 77: 3537–3540, 1996.
28. A. Goriely and M. Tabor. Nonlinear dynamics of filaments IV: The spontaneous looping of twisted elastic rods. *Proc. Roy. Soc. London (A)*, 455: 3183–3202, 1998.
29. A. Goriely, M. Nizette, and M. Tabor. On the dynamics of elastic strips. *J. Nonlinear Science*, 11: 3–45, 2001.
30. D. M. Haughton and R. W. Ogden. On the incremental equations in non-linear elasticity-II. Bifurcation of pressurized spherical shells. *J. Mech. Phys. Solids*, 26: 111–138, 1978.
31. D. M. Haughton and R. W. Ogden. Bifurcation of inflated circular cylinders of elastic material under axial loading-I. Membrane theory for thin-walled tubes. *J. Mech. Phys. Solids*, 27: 489–512, 1979.
32. D. M. Haughton and A. Orr. On the eversion of compressible elastic cylinders. *Int. J. Solids Structures*, 34: 1893–1914, 1997.
33. G. Helmlinger, P. A. Netti, H. C. Lichtenbeld, R. J. Melder, and R. K. Jain. Solid stress inhibits the growth of multicellular tumor spheroids. *Nature Biotech.*, 15: 778–783, 1997.
34. G. A. Holzapfel, T. C. Gasser, and R. W. Ogden. A new constitutive framework for arterial wall mechanics and a compartaive study of material models. *J. Elasticity*, 61: 1–48, 2000.
35. C. 0. Horgan and G. Saccomandi. A molecular-statistical basis for the Gent constitutive model of rubber elasticity. *J. Elasticity*, 68: 167–176, 2002.
36. C. 0. Horgan and G. Saccomandi. Constitutive modeling of rubber-like and biological materials with limited chain extensibility. *Math. Mech. Solids*, 7: 353–371, 2002.
37. C. 0. Horgan and G. Saccomandi. A description of arterial wall mechanics using limiting chain extensibility constitutive models. *Biomechan. Model Mechanobiol.*, 1: 251–266, 2003.
38. C. 0. Horgan and G. Saccomandi. Constitutive models for compressible nonlinearly elastic materials with limiting chain extensibility. *J. Elasticity*, 77: 123–138, 2004.
39. F. H Hsu. The influences of mechanical loads on the form of a growing elastic body. *J. Biomech.*, 1: 303–311, 1968.
40. J. D. Humphrey and K. R. Rajagopal. A constrained mixture model for growth and remodeling of soft tissues. *Math. Models. Meth. Appl. Sci.*, 12: 407–430, 2002.
41. I. Klapper. Biological applications of the dynamics of twisted elastic rods. *J. Comp. Phys.*, 125: 325–337, 1996.
42. I. Klapper and M. Tabor. Dynamics of twist and writhe and the modeling of bacterial fibers. In J. Mesirov, K. Schuiten, and De Witt Summers, editors, *Mathematical Approaches to Biomolecular Structure and Dynamics*, 139–159. Springer, 1996.
43. S. M. Klisch, T. J. Van Dyke, and A. Hoger. A theory of volumetric growth for compressible elastic biological materials. *Math. Mech. Solids*, 6: 551–575, 2001.

44. E. H. Lee. Elastic-plastic deformation at finite strains. *J. Appl. Mech.*, 36: 1–8, 1969.
45. J. A. Lockhart. An analysis of irreversible plant cell elongation. *J. Theor. Biol.*, 8: 264–275, 1965.
46. B. D. MacArthur and C. C. Please. Residual stress generation and necrosis formation in multicell tumour spheroids. *J. Math. Biol.*, 49: 537–552, 2004.
47. J. H. Maddocks. Bifurcation theory, symmetry breaking and homogenization in contiunuum mechanics descriptions of DNA. In M. J. Givoli D. Grote and G. Papanicolaou, editors, *A Celebration of Mathematical Modeling: The Joseph B. Keller Anniversary Volume*, pages 113–136. Kluwer,, 2004.
48. A. A. Matista and W. K. Silk. An electronic device for continuous, in vivo measurement of forces exerted by twining vines. *Am. J. Bot.*, 84: 1164–1168, 1997.
49. G. A. Maugin. Pseudo-plasticity and pseudo-inhomogeneity effects in material mechanics. *J. Elasticity*, 71: 81–103, 2003.
50. T. McMillen and A. Goriely. Tendril perversion in intrinsically curved rods. *J. Nonlinear Science*, 12: 241–281, 2002.
51. N. H. Mendelson. Helical *Bacillus subtilis* macrofibers: morphogenesis of a bacterial multicellular macroorganism. *Proc. Natl. Acad. Sci. USA*, 75: 2472–2482, 1978.
52. N. H. Mendelson. Bacterial macrofibers: the morphogenesisof complex multicellular bacterial forms. *Sci. Progress Oxford*, 74: 425–441, 1990.
53. N. H. Mendelson. *Bacillus subtilis* macrofibres, colonies and bioconvection patterns use different strategies to achieve multicellular organization. *Environmental microbiol.*, 1: 471–478, 1999.
54. F. Michor, Y. Iwasa, and M. A. Nowak. Dynamics of cancer progression. *Nature Rev. Cancer*, 4: 197–205, 2004.
55. P. M. Naghdi. A critical review of the state of finite plasticity. *Zeitschrift fr Angewandte Mathematik und Physik*, 41: 315–394, 1990.
56. R. W. Ogden. *Non-linear elastic deformation*. Dover, New york, 1984.
57. R. W. Ogden and A. Guillou. *Growth in soft biological tissue and residual stress developments*. Preprint, 2005.
58. J. K. E. Ortega. Augmented growth equation for cell wall expansion. *Plant Physiol.*, 79: 318–320, 1985.
59. A. Rachev. Theoretical study of the effect of stress-dependent remodeling on arterial geometry under hypertensive conditions. *J. Biomech.*, 30: 819827., 1997.
60. E. K. Rodriguez, A. Hoger, and McCulloch A. Stress-dependent finite growth in soft elastic tissue. *J. Biomechanics*, 27: 455–467, 1994.
61. T. Roose, P. A. Netti, L. L. Munn, Y. Boucher, and R. K. Jain. Solid stress generated by spheroid growth estimated using a linear poroelasticity model. *Microvascular R.*, 66: 204–212, 2003.
62. M. S. Sacks. Biaxial mechanical evaluation of planar biological materials. *J. Elasticity*, 61: 199–246, 2000.
63. M. Sarntinoranont, F. Rooney, and M. Ferrari. Interstitial stress and fluid pressure within a growing tumor. *Annals biomed. Eng.*, 31: 327–335, 2003.
64. J. L. Scher, MN M. Holbrook, and W. K. Silk. Temporal and spatial patterns of twining force and lignification in *Ipomoea purpurea*. *Planta*, 213: 192–198, 2001.
65. A. P. S. Selvadurai. Deflections of a rubber membrane. *J. Mech. Phys. Solids*, 54: 1093–1119, 2006.

66. O. A. Shergold, N. A. Fleck, and D. Radford. The uniaxial stress versus strain response of pig skin and silicone rubber at low and high strain rates. *Int. J. of Impact Eng.*, 32: 1384–1402, 2006.
67. B. I. Shraiman. Mechanical feedback as a possible regulator of tissue growth. *Proc. Natl. Acad. Sci. USA*, 102: 3318–3323, 2005.
68. W. K. Silk. Growth rate patterns which maintains a helical tissue tube. *J. Theor. Biol.*, 138: 311–327, 1989.
69. W. K. Silk and N. M. Holbrook. The importance of frictional interaction in maintaining the stability of the twining habit. *Amer. J. Bot.*, 92: 1820–1826, 2005.
70. W. K. Silk and M. Hubbard. Axial forces and normal distributed loads in twining stems of morning glory. *J. Biomechanics*, 24: 599–606, 1991.
71. A. A. Stein. The deformation of a rod of growing biological material under longitudinal compression. *J. Appl. Math. Mech.*, 59: 139–146, 1995.
72. L. A. Taber. Biomechanics of growth, remodeling and morphogenesis. *Appl. Mech. Rev.*, 48: 487–545, 1995.
73. L. A. Taber. Biomechanical growth laws for muscle tissues. *J. Theor. Biol.*, 193: 201–213, 1998.
74. L. A. Taber and D. W. Eggers. Theoretical study of stress-modulated growth in the aorta. *J. Theor. Biol.*, 180: 343–357, 1996.
75. D. Wodarz and N. L. Komarova. *Computational biology of cancer: Lecture notes and mathematical modeling*. World Scientific, Singapore, 2005.
76. C. W. Wolgemuth, R. E. Goldstein, and T. R. Powers. Dynamic supercoiling bifurcation of growing elastic filaments. *Phys. D*, 190: 266–289, 2004.
77. H. Xiao, T. O. Bruhns, and A. Meyers. Elastoplasticity beyond small deformations. *Acta Mechanica*, 182: 31–111, 2006.

A Model of Pattern Coupled to Form in Metazoans

Frederick W. Cummings

Professor of Physics (emeritus), University of California Riverside. Present address: 136 Calumet Ave., San Anselmo, Ca., 94960
fredcmgs@berkeley.edu

Summary. A model of patterning in living systems is examined, one involving the sequential interaction of a pair of signaling pathways. The model of pattern is coupled to the changing shape of a (closed, thick) epithelial shape. Focus is on patterning that couples localized cell differentiation and epithelial shape changes. Aspects of the model are discussed in turn: the pattern, the epithelial sheet geometry, and the coupling of the latter two. Changes in the pattern give rise to a changing epithelial shape, and a changing epithelial shape in turn causes a change in the pattern, and so on, as the total a area increases. The model is intended to provide a simplest example of morphogenesis. An effort is made to reduce morphogenesis to its most elemental 'modules': pattern, shape and their interaction. This may be seen as a reversal from the usual historical progression (e.g., as in physics), i.e., first from phenomenological modeling, then only later, to its reductionist underpinnings: first came thermodynamics and later statistical mechanics. Post DNA biology has been for the past fifty years primarily, and certainly very fruitfully, focused on the genetic and molecular basis of development. The present work attempts to work from the genetic and molecular underpinnings toward a more phenomenological model of the origin of animals, but one simpler and more comprehensible. Mathematical modeling has rarely dealt with animal form, or with its coupling to pattern, and an attempt is made in this direction. A preliminary attempt is made to uncover possible molecular and genetic foundations of the present elemental model of interacting shape and pattern. Very important recent genetic findings suggest that the very most conserved regions of DNA are of relevance to providing a deeper understanding of the successive coupling of signaling pathways. Molecular bases for pattern, shape and their interactions are discussed, and possible connections to the model are tentatively proposed. The model provides positional specification for Stem cells as interstitial to the regions undergoing determination and differentiation. The role of stem cells in patterning and morphology is speculated to occur by way of the ubiquitous cytostolic-membrane-phosphoinositides.

Key words: Embryogenesis, epithelial shape changes, Fibonacci patterns, pattern model, Gauss curvature.

1 Introduction

Much of what was known by about 1980 of an organism's development from a fertilized egg came from a mixture of genetics and molecular biology, mostly from two organisms, C. elegans and Drosophila [47]. This groundbreaking work has been extended in recent times to vertebrates. What has become increasingly clear in recent years is that deciphering the DNA code is only the very first step in understanding the complex dynamics of the genome and its regulatory network. The simple canon of gene → mRNA → protein is no longer valid; the genetic regulatory system is very much more complex, with RNA, as just one example, playing a much more varied role as genome regulator. It may be argued that unraveling the human genetic regulatory 'code' may greatly exceed the human genome project in time and expenditure.

The past ten years or so has seen a resurgence of great interest in connections between development and evolution. These two areas of research have not for many years recognized overlapping interests. This rebirth is commonly known as 'evo-devo', colloquial for 'evolution and development'. One main tenet of the new discipline is the conservation of genetic and developmental commonality over some six hundred million years that underlies all the diversity of animal types. There was, on this view, a single ancestor of all multicellular metazoa, and most likely a dominant 'rule' or generality of cell-cell association discovered by evolution at the time of multicellular origins. Such cellular associations were initially into two epithelial sheets (e.g., cnidaria), but in relatively short time a third layer evolved, giving then three germ layers, endoderm, ectoderm and mesoderm, i.e., the bilateral animals. It is the claim that remnants of such a rule of cellular association are still discernable in spite of a half billion years of evolutionary elaboration and adaptation. The limited number of major phyla, in spite of numerous mass extinctions expected to have created a predator vacuum, is taken as one rationale for such an assumption. Further, researchers in evo-devo have uncovered remarkable conservation of many of the most crucial genetic mechanisms, and it is relevant to emphasize for present purposes that these very frequently involve the signaling pathways involved in early development (e.g., [9, 62]).

Certain features of animal body plans, such as bilateral symmetry, have been conserved since the early Cambrian or possibly earlier. In contrast, at the species level there has been a continuous accumulation of changes; for only one example, one notes the half million or so species of beetle. It is proposed that the genetic regulatory networks associated with development contain components, or domains, that differ in their evolutionary conservation. Relatively unchanging DNA domains act to perform essential upstream functions in building given body parts, while other domains have been repeatedly taken over during evolution to perform diverse developmental tasks. Then the highly flexible, individual cis-regulatory linkages have been left to regulate detailed phenotypic variation [21].

Essentially the same gene can make quite different structures. Evo-devo emphasizes the non-coding DNA that sits between genes, and the switches that activate this region, while playing down the earlier 'protein-centrist' perspective. While this new perspective may not be universal, it is clearly very important. One common example among many is the Pax6 gene (in mice) and the corresponding eyeless gene (fly). This gene apparently sits at the top of the regulatory chain of genes involved in the specification of all types of eyes; it simply gives an instruction "turn on the eye making process". The eyeless (fly) gene inserted into a mouse leg gives instruction to make a mouse eye at that location. Changes in bird plumage color often involve the same gene that causes red hair in humans. This sort of amazing genetic conservation across nearly all animals is one of the key findings of evo-devo. This is accomplished apparently by 'switches' which sit next to non-coding genes [9], and such promoters are turned on by 'transcription' factors (proteins). However, only two percent of the human genome, for example, gives rise proteins, while stretches of DNA in the remaining 98% of non-coding sequences inform genes when and where to be turned off or on. Promoters are regulatory genes that typically sit immediately before a gene and signal where transcription should begin. Enhancers on the other hand sit before, after and even within genes that they regulate. In some cases, the enhancers function millions of base pairs away from the genes they regulate.

Such master regulatory genes as promoters and enhancers effectively switch other genes on or off, and these in turn switch on yet other genes, and so on, in a cascade ending with final assembly of the animal. What is conspicuously missing in this genetic description is the subject of most interest at present, namely a specification of locale and sequence: how does a gene know to turn on 'here' and not 'there' in a tissue, now and not later. This last task is left to the 'patterning' process, one not extensively discussed by evo-devo.

The Hox genes are an important and ancient part of the highly conserved animal genetic repertoire. In bilateral animals, the Hox genes tell a given body part, a segment, which appendage it should grow. A Hox gene expressed in the head will tell head to grow antennae, while another Hox gene expressed in the body might tell the body to grow legs. Mutations may cause legs to appear where antennae normally belong, and so on. All animals have Hox genes, and all animals use their Hox genes to determine which appendages go where along an axis that runs from head to tail. Even cnidaria have Hox genes, and the number of Hox clusters varies with phyla, vertebrates having the most. A large number of genes that act in the same general way, it turns out, are as old as Hox genes, that is, they have also been around since the beginning of the bilaterians. The most parsimonious model for the evolution of Hox/ParaHox clusters has been given recently [10]. It is worth emphasizing here that the locus of an appendage around a segment (a leg, say) must still be specified by some further mechanism, the pattern mechanism.

The two main preoccupations, or projects, of Physics may be labeled 'Unification' and 'Conservation'. Newton (1687) first unified the earth physics of Galileo with that of the heavens. Maxwell (1860) unified the electric and magnetic fields and electromagnetic waves, and the unification project continues today with efforts to unify the three forces (e&m, weak and strong) with gravity. The conservation of energy (now more generally, mass-energy) became a powerful concept post Newton, as did the conservation of charge with Coulomb ($\sim$ 1880), followed by identification of any number of other important conserved quantities. It is most interesting to witness these two concepts, unification and conservation, albeit in somewhat different clothing, now taking center stage in biology as well.

It is desired to include the most basic elements of embryogenesis and morphogenesis. Such a parsimonious effort must be at best preliminary, excluding as it does a number of key developmental events, e.g., such as cell migration. However, it is helpful to remember that ultimately all models are wrong [13]. The origin of multicellular animals is a subject of primary interest, given that eukaryotic cells had reached a degree of high sophistication by the Precambrian. The present paper describes a patterning model intended to be congruent with evolutionary conservation. The model specifically requires coupling of pattern to epithelial tissue shape. The assumption is that there are discernable remnants of the earliest pattern formation of multicellulars, and that the many subsequent evolutionary changes are most selectively focused on features that are not involved with this most primitive conserved beginning, one assumed to involve the coupling of signaling pathways.

A number of important recent and relevant findings concerning genetic regulatory systems are briefly summarized in Section 2. These findings involve the many highly conserved genetic regulatory domains (chromatin and nucleosomes) spaced out along DNA. These conserved regulatory domains are those most closely coupled to the elements of the patterning model and its coupling to the epithelial shape changes.

Section 3 reviews and discusses a recent model of patterning, specifically those elements of patterning which couple to morphogenesis. Ultimately, the aim is to make contact of the model with the genetic regulatory results of Section 2. First steps in this direction are attempted. That the pattern part of the model necessarily involves the geometry of the epithelial sheet is clarified. The mathematical details of the model are in Appendix A. Examples of patterns on specified geometries are presented in fig. 2 and fig. 3, indicating the versatility of the model. Fig. 2 shows a number of steady state small amplitude solutions on a 'stretched' torus, or elongated donut. Fig. 3 shows, also in steady state and small amplitude, the most frequent plant patterns for the 'leaves' (or modified leaves such as petals) on a cylindrical stem, with the Fibonacci patterns most prominent. The mathematical details pursuant to the plant patterns of fig. 3 are included in Appendix B.

Section 4 discusses aspects of the geometry of epithelial tissues. Simple expressions for the two curvatures are given in terms of the apical and basal

areas and the sheet thickness at any point of the surface. Section 5 explores aspects of the numerical computation of an invaginating surface. The mean and Gauss curvatures were assumed to be functions of the two morphogens and their gradients. In particular, the famous Gauss-Bonnet theorem, which states that the area integral of the Gauss curvature over any closed surface must be an integer times 4π, is seen to provide important constraint to numerical computations. Of necessity, any surface invagination will involve a 'twisted' cell configuration, the 'twist' gene, and the important protein β-catenin. The particular choice for the functional dependence of the Gauss curvature as a function of pattern 'morphogens' used in the numerical integration of the invagination, starting from an original sphere in the steady state model, is shown to be most reasonable.

A model of a simplified closed system emerges, pattern affecting geometry, geometry in turn affecting pattern. The pattern affects the geometry, via the genome, and the geometry in turn affects the pattern, as growth takes place and the total area of the epithelial (middle) surface increases. The model describes a binary branching genetic tree at each growth cycle. A growth cycle consists of a time interval during which a given pair of active signaling pathways, each most active in separate spatial domains, grows to a maximum of excitation, followed by their rapid decay. Section 6 discusses a few more technical non biological aspects of numerical computation of the model in axial symmetry.

2 Aspects of the Genome

Even though all the cells of an organism carry the same genome, each type of cell must have access to only a few of the genes of the genome, with all others permanently unavailable to it. The operation of this system that assigns various identities to the cells is a central mystery of animal existence, but one about which much new knowledge has been gained recently. It is the present belief that these new finding will prove to be the genetic aspects most closely coupled to the present model of coupled signaling pathways.

Recently it has been discovered that embryonic cells are kept in a poised state, so that all of the genome's many developmental programs are in an inactive state, while at the same time their genes are ready to be read out when the cell is assigned to a particular developmental path. The developmental programs are initiated by master regulatory genes, and these control many lower genes. The master regulatory genes do not code for proteins in the usual sense of gene-RNA-protein, but perform their function by producing what are called transcription factors, proteins that seek out and bind to sites on the DNA, thus controlling lower-level target genes [6]. A key question concerns what controls the non coding master regulatory genes, the promoters and enhancers. The answer apparently lies, partly at least, in the chromatin. The DNA is looped around very many 'spools' of chromatin, wrapped just

1.65 times around each spool. The nucleosome, with ~ 200 base pairs, is the fundamental repeating unit of chromatin, and is wrapped around a core of proteins called histones. The chromatin on a given spool can either be tightly wound, in which case the DNA is inaccessible, or it can be wound loosely, when it is accessible to transcription factors that aim to copy a gene to produce a specified protein. Each spool of chromatin then in effect contains two control 'tags', an 'on' tag and an 'off' tag. For example, a triple methylation mark on lysine 4 of histone H3, designated as H3K4me3, is a hallmark of all active genes [39, 63]. Such loci cause other proteins to wrap the DNA loosely, while still other proteins act to wrap other tags tightly, preventing genetic readout. Mature cells are known to have long stretches of chromatin corresponding to chromatin regions of the DNA turned off. Much shorter sections of mature cells have the 'loose' tag, indicating that the local genes are allowed to be activated there. This is a simplistic description of a very complex biochemical situation, still in the process of being unraveled [1].

The bivalent chromatin domains occur at regions on the chromosome where some of the DNA is highly conserved, which implies that the same DNA sequence is found in very different species, or even different phyla. Being highly conserved implies that this domain is spared the usual mutations, and that this DNA plays some vital role in development. The reasonable speculation is that many of these domains are precisely those involved in cell signaling, tissue shape and growth. These sequences do not contain the usual genes that code for proteins ('coding genes'), but only master regulatory genes, meaning that these domains are conserved for some other reason.

The DNA not only encodes the instructions for making the organism in its sequence of base pairs, but also provides signposts for regulating its own readout. Such signposts allow regulatory factors to take up the positions along the DNA that are necessary for proper functioning. The nucleosome positioning is probably a key element in regulating access of other regulatory proteins to DNA. Such things as transcription factor binding, transcription initiation, and possibly even reorganization of the nucleosome positioning itself are likely crucially affected by the nucleosome positioning. Segal et al. [56] have shown the high probability that there is a genomic code for nucleosome positioning along the DNA.

When embryonic cells are investigated, the chromatin covering the regions of DNA where the master regulatory genes are situated are found to have both tags, indicating 'available to be read', and also for 'closed' or unreadable. These genes were neither accessible nor inaccessible, but in a state of suspended animation, *both* silenced and at the same time ready for readout. These ambiguous stretches of 'poised' chromatin are called 'bivalent'. The sense of this finding is that each cell of the embryo must not be committed to any fate, not for the time being. This bivalent or third state was found only in the embryonic cells. The master regulatory genes must be suppressed, but on the other hand, must be ready to activate a particular master regulatory gene when its fate is determined in some way. How this particular designa-

tion of specific master regulatory genes happens biochemically is a remaining mystery; but it is as if each chromatin domain contains a three way 'toggle', with an 'on', an 'off' and a 'bivalent' setting. Enhancer genes are even more mysterious, and scientists still don't know enough about them yet to even reliably distinguish them from other sequences.

The regions of cell determination will be addressed in later section, under the heading of "pattern", when the effect of neighboring cells of any given cell comes into play. The ultimate goal, and one certainly not achieved at present, is to provide a seamless connection between the genome and the coupled signaling system, the latter crucially involved with patterning and thus with neighboring cells.

A very interesting recent finding [3, 5, 64] involves the bivalent domains of embryonic stem cells. It was found that there are certain genes, most prominent among them (e.g.) oct4, sox2, nanog, and c-myc, which are particularly active in mammal stem cells [53]. Analogous genes with different interactions between them are presumed to be active in all other eukaryotic animal stem cells. The gene c-myc is long known to be involved in control of cell division [49]. These genes make transcription factors that act on each other's control sites in such a way as to form a sort of circuitry for controlling the stem cell's master regulatory genes. Their function in stem cells is, surprisingly, to keep the genes that effect cell determination inactive. Loss of activity in certain individual genes of this group initiate distinct differentiation programs. Several gene products are generally required in order to activate the cell's promoter sites, and seven such regulatory genes have been described. Specific functional roles have been suggested [35]. Some limited tie-in of these finding with the 'pattern' model of Section 3 is attempted, in an ongoing effort.

Emphasizing the complexity of the genome further, even epigenetic control of gene expression may occur, as first suggested by Waddington [29].

The above description of the genomic control system of a cell contains no mention of the influence of the neighbors of any specific cell under discussion. This omission is a main topic of interest below. It will be argued that the neighboring cells play a decisive role in determining the fate of a given cell, in a binary manner in each 'growth cycle'. The central role of a certain set of signaling pathways in patterning, tissue shape and cellular determination is seen as essential.

3 The Pattern Model

Focus is now on a particular cell, or perhaps a small group of similar cells. The pattern most generally includes the geometric shape of the tissues. However, a somewhat artificial separation will be made at first by use of the terms "pattern" and "shape" (or geometry) of a tissue. It is found easier to first treat the two aspects as if separate, before their interaction is introduced,

after which the distinction becomes blurred. In the ensuing, one affects the other, pattern determines shape, and shape determines pattern; a 'chicken and egg' situation.

One of the most fundamental questions of development involves patterning. How is it that an animal with the same genetic information in each cell comes to have different organs? This is the question to which 'pattern' is addressed. It is a question that should be treated not only from the point of view of the genetic regulatory system of a single cell, but also by recognizing the crucially important interaction of a cell with its neighboring cells. When it is asserted that certain proteins are produced at a particular site because certain genes (or gene network) were activated, as true a statement as this may be, the question of "why here at this locale, and not there" always looms. Often pattern models are of the seminal 'Turing' class [60], where two substances necessarily diffuse at different rates, and interact to produce a pattern due specifically to the nonlinear nature of the two interacting substances [38, 42, 43, 45]. Other well studied patterning agents include the diffusing maternal gene products in the case of long germ syncytial insects such as (esp.) the fly, when cell membranes do not form for numerous (13) cell division cycles. Dilão has recently given a review of reaction-diffusion models. The mass action law is introduced, and mass conservation introduces memory effects in biological development in patterns that can be explained as differences in initial conditions occurring during development [22]. Clearly the present model only applies to animals with epithelial cells, i.e., cells with plasma membranes; in the pre-cell state such as that of syncytial insects, other pattern mechanisms are expected to apply.

Beloussov [2] has emphasized an approach pointedly not based on reaction-diffusion schemes; the emphasis is in uncovering possible universal "rules of morphogenesis", and such a search is also a motivation of the present work [59]. While the motivation is the same as that of Beloussov, the present approach is quite different in that it is assumed that signaling pathways affect cell shape changes through interaction with the genome, and that the cell shape changes in turn affect the pattern, and so on.

The present pattern model consists of two interacting signaling pathways [16, 17], a different pair (perhaps) specified at each cycle of growth. The process of signal transduction allows a cell to receive messages from its environment, most often very local, as from neighboring cells, and transfer this signal from the membrane through the cytoplasm and into the nucleus. Here the signal contributes to the cell's response by altering the expression of various regulatory genes. Regardless of the signal's nature, the general logic of the transduction pathways that are triggered by ligands is roughly the same. The signaling protein (ligand) binds to a receptor, which consequently undergoes conformational change. Most commonly, phosphorylation by a kinase enzyme allows the receptor to recruit cytoplasmic signaling components, initiating a cascade of events resulting in changes in gene expression, and often other events as well. The well studied 'Wnt' pathway may act as a model in

many respects, although it has some important twists of its own which go beyond the simple description just given [13, 20, 46].

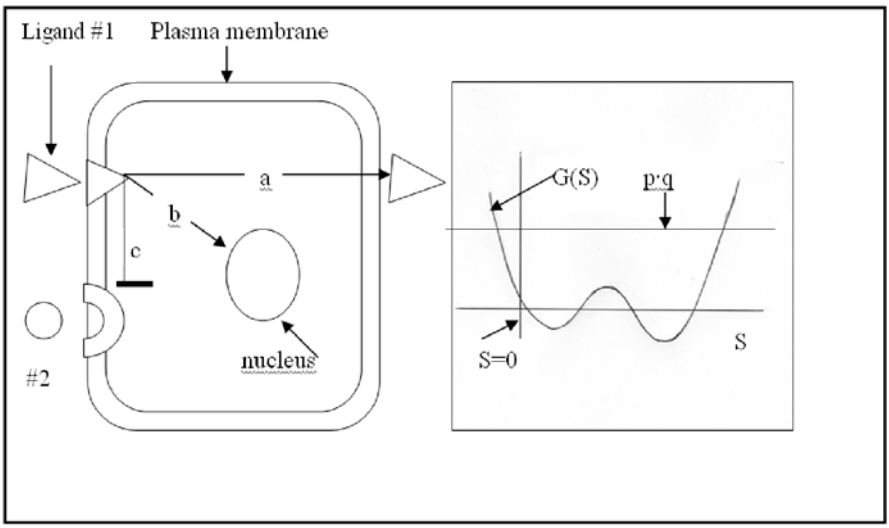


Fig. 1. Schematic of a cell shows (on left side of fig. 1) two receptors on the left associated with two different signaling pathways, both embedded in the plasma membrane. Also indicated are two extracellular ligands, labeled #1 and #2. It is assumed that only receptor #1 is activated by its ligand. There are three intracellular branches indicated leading from the activated receptor #1, labeled as a, b and c in the figure. Path a provides stimulated production of further ligand of type #1, as indicated by the ligand to the right of the cell. Path b is a path leading to activation of a transcription factor in the nucleus specific to signaling pathway #1. Activation of relevant genes is achieved via path b. Path c blocks both production of ligand #2, as well as activation of the transcription factor specific to signaling pathway #2. The paths a and c act relatively fast compared to the path b going to the nucleus. This path to the nucleus b is not considered explicitly in the math model, since such is specific to each animal. Fig. 1b shows a plot of $G(S)$ vs. S. Spontaneous activation of the pattern arises when there is a negative root of the equation $G(S) = p.q$, defined by eq. (21), and shown in fig. 1b. Shown is the case of two real roots and two complex roots. The two complex roots give rise to traveling waves, with a velocity equal to the imaginary part of the root divided by the wave number k. A wave damped in time occurs if the real part of the complex root is positive, otherwise a growing wave results.

The focus is first on a specific growth cycle, with one pair of signaling pathways active. Specific ligands exist for each type of receptor. The ligand densities for the two coupled pathways are denoted as L_1 and L_2. The notation R_1 and R_2 denotes the densities of the corresponding activated receptors, that is, the receptor plus its associated ligand. It is these latter entities which are

taken to be what we term the 'morphogens'. Activation of a receptor leads to activation of transcription factors in the nucleus, after upstream interactions with numerous intracellular cytoplasmic mediators unique to each pathway [25].

The literature on signaling can be a miasma of acronyms; but a reasonably simple underlying logic is beginning to appear. The pathways of interest here are those that interact in the particular way envisioned in the paragraph just below, ones that have a ligand captured by a receptor culminating in transport of a transcription protein into the nucleus. The interactions in the cytoplasm are specific to each of the two pathways at each time step, and since the interest here is in the basic events of embryogenesis, the result is both cellular differentiation and tissue shape change.

The two main elements of the model may be simply stated: 1) Activation of receptor '1' by its associated ligand leads to deactivation of receptor '2', and vice versa. This is indicated in fig. 1a by the arrow labeled *c*. The deactivation of (e.g.) receptor '2' denotes deactivation of the production of 'like' ligands of pathway '2'. Each active receptor (e.g., '1') produces a cascade that acts to deactivate the other type (e.g., '2') pathway. It is likely that this function is facilitated by a member of the GTPase family [26, 32]. The result is that each receptor R_1 or R_2 leads to production of ligand of like kind [18, 19]. This production of like ligand by its receptor activation occurs by way of the nucleus; this process has not as yet been empirically attributed to a specific reaction. A speculation concerning the activation of like ligand as well as the deactivation of the 'other' ligand will be given later in this section, one involving the ubiquitous membrane associated phosphoinositides.

The arrow in fig. 1a labeled *a* indicates the production of 'like' ligand. The pathway may be by way of the nucleus, or by way of a combination of pathways, one of which involves the nucleus and transcription factors, and one bypassing the nucleus. The pathway shown in fig. 1a indicates only a pathway bypassing the nucleus, for simplicity. Also, of course, activation of a given receptor culminates in activation of relevant master regulatory genes in the nucleus, and indicated in fig. 1a by the arrow labeled *b*. This provides for three critical functions to be performed by each activated receptor: (a) secretion of 'like' extracellular ligand by its receptor; (b) activation of those regulatory genes setting in motion a cascade of genes associated with the particular genetic pathway associated with each spatial region. It is assumed that transcription factors which control the actin cytoskeleton and thus cell shape will be activated; (c) deactivation of the pathway giving rise to secretion of the 'other' ligand.

These assumptions allow one to at once write equations for the four quantities involved in the patterning process, namely, L_1, L_2, R_1 and R_2 (Appendix A). The R_1, R_2 will play the role of the usual "morphogens" in the steady state example patterns of Section 4. The details of how the signal activates the relevant genes, or which genes are activated, is not addressed by the model, and is of course specific to each species at each growth cycle. The model aims

to focus attention on the elements of pattern and shape which have universal relevance.

Three events are labeled in fig. 1a. In fig. 1a the pathway labeled a denotes the stimulated emission of like ligand into the extracellular space upon capture of ligand '1' by its receptor (brown), whether such event involves bypassing the nucleus or not. The pathway b denotes the (perhaps relatively slower) activation of transcription factors in the nucleus following capture of ligand '1', and followed ultimately by specific protein production. There is no term in the math model dealing with this b path interaction. The pathway indicated as c denotes the (perhaps relatively fast) blockage following activation of receptor '1' of the secretory activity of the signaling pathway activated by ligand '2'. The same blocking of '1' occurs upon activation of the '2' signaling pathway. The two pathways act essentially in a mirror image fashion; what is true of pathway #1 is also true of pathway #2. Just how the stimulated secretion is facilitated, as well as how activation of one pathway acts to deactivate the other can be seen as a separate subject. However, the present speculation is that the phosphoinositides play a crucial role in both respects, stimulated emission and second ligand deactivation.

The model has a number of interesting properties. First among these is the fact that spontaneous pattern activation from zero amplitude does not depend on the form of the nonlinearities (which of course invariably enter), or even on their existence at all. Setting each of the four variables proportional to $\exp(i\mathbf{k}\cdot\mathbf{x}-st)$ in the linear version of the model in the usual way (e.g., [45]) leads to a fourth order equation to determine four roots of a function $F(s)$, or equivalently, $G(S)-pq$, where p and q are (constant) parameters of the model ([16]; Appendix A). Fig. 1b shows the function G, which has four analytically determined real roots. Under appropriate parameter values, one of the roots of $F(S) = G(S) - pq$ will be negative, one positive, and often two others will be imaginary. A solution with two imaginary roots is shown in fig. 1b. The maximum of the function G lies below the constant $p.q$ in this case. Such roots imply traveling wave solutions, waves traveling with a velocity given by the imaginary part of the root divided by the magnitude of $\mathbf{k} = |k|$. This is seen by substituting either of the two imaginary roots into the exponential just above, so that there is a solution of the form of a traveling wave. Such a wave will be growing if the real part is negative. Traveling waves have been observed during development (e.g., [30,31]); Cummings [15] has given a simple model of waves in Drosophila eye development based on signaling pathway interaction.

A second important property of the model is that there is no requirement that any quantity diffuse faster than any other. The activated receptors, or rather their density, define the 'morphogens' of the model. The active receptors are fixed in the cell plasma membrane, and are relatively immobile. The ligands themselves can be imagined to in general diffuse only a few cellular diameters at most in the extra-cellular space until they are captured by its receptor, and ligand may even most often originate from nearest neighbors of the receptor cells. Due to a relay effect, there is often the appearance of

long-range diffusion. Of especial interest is that the model is based squarely (if simply) on the interaction of two signaling pathways at each growth cycle, such pathways being well known to be crucially important in embryonic development (e.g., Nature reviews, 2006, 441, #7092, 423-530), at least when acting even independently.

The model leads in a natural way to a binary genetic decision at each 'growth cycle', or time step. Each region of (epithelial sheet) space is divided into relatively smaller and smaller pattern domains as growth takes place. At each time step, or growth cycle, only one pair of all possible pairs of signaling pathways is selected. A given signaling pair, along with the geometry of the epithelia, determines the complete pattern. The Wnt [62], Hedgehog, Notch and BMP are examples of highly conserved pathways [16,33] of a surprisingly small handful of candidates involved in embryogenesis, approximately seven. Signaling pathways have long been known to direct growth and patterning during embryogenesis. Mis-specification of cells towards stem-cell fates rather than their intended differentiated designation often result in tumorgenesis [57].

The amplitude (maximum size of the densities R_1 or R_2) of any particular pattern increases as tissue growth occurs, and as total area of the (middle surface) epithelial sheet increases. Numerical calculations show that solutions decay away quickly once the amplitudes reach a certain maximum value, as required by the model (and evident from eqs. (14) and (15) of Appendix A). The next allowed, more complex pattern of lower symmetry is then allowed to emerge. This next pattern is one of originally small amplitude, arising when the total tissue area becomes adequate by virtue of growth. The two to be differentiated spatial areas of the model are specified by the two scale parameters k_1 and k_2, which will in general be different at each growth cycle, and defined in terms of the model parameters below eq. (19) of Appendix A.

Last but not least, stem cells find a niche in the present pattern generator. Stem cells play such an important role in development, it is important that their patterning be specified. They lie between the two morphogens which provide a common region in which each morphogen is of too small amplitude density to effect differentiation.

Fig. 2 illustrates aspects of the steady state ($\partial/\partial t = 0$ in eqs. (12) and (13)) pattern model on a fixed cylindrical toroidal geometry, for several values of increasing total area. Each of the three figures shown is a thin torus, requiring periodic boundary conditions in two coordinates. Several cycles of growth are illustrated. The colors do not specify the different possible regions of cellular determination, but rather only mark off the regions between patterns where stem cells are expected to occur. For example, in fig. 2b four colors would be required to uniquely specify all determined regions, and distinguishing the dorsal from the ventral region. Eight different colors would otherwise be required in fig. 2c if each region possibly undergoing differentiation were shown. Interstitial stem cells regions are indicated in each case by a thin black line separating each (much larger) pattern region. The stem cell regions are

those interstitial regions where the amplitude of neither morphogen (R_1 or R_2) is sufficiently large to induce cellular determination.

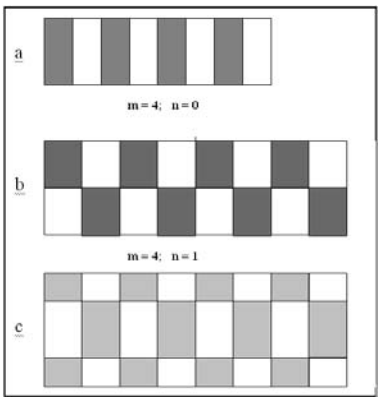


Fig. 2. A schematic of the small amplitude solution of the model of Appendix A is shown. The geometry is that of a torus, and each of the three figures 2a, 2b and 2c are shown in a side view of the outermost cylindrical surface. Two cylinders of the same lengths but slightly different radii, one surrounding the other, are imagined as joined at their ends. The model has been solved in the small amplitude case, and time independent. Fig. 2a shows patterning on the external cylinder after several growth cycles. The same colors indicate that, in this case, the same pair of signaling pathways was used for each cycle. Each pair of rectangular grey and white rectangles represents one growth cycle. The (thin) black lines separating the grey and white rectangles represent stem cell locations, interstitial to the larger rectangles, or differentiating regions. How many growth cycles have occurred depends on whether or not all possible stem cell regions underwent growth at each cycle. The numbers $m = 4$, $n = 0$ indicate four axial growth cycles, and no radial pattern has yet emerged. fig. 2b represents just one extra cycle of growth beyond fig. 2a. This pattern introduces bilateral symmetry; a most basic 'animal', with a through gut, and bilateral symmetry emerges [18,19]. Two horizontal stem cell regions are indicated by one horizontal line, in side view; the intersecting vertical lines are again stem cell locations. The bilateral symmetry can arise only when the radius of the cylinder grows to a large enough value, indicated by $m = 4$, $n = 1$, labeling this solution. Fig. 2c indicates an even larger radius, indicated by the next allowed solution $n = 2$. Here is seen intersection of two horizontal stem cell regions (on one side) with several (here seven) vertical stem cell regions. There will clearly only always be an odd number of ventral (lower) intersections, regardless of the value of axial length, m. This is in view of the fact that the model always requires that both patterns occur together (e.g., grey and white as in fig. 2a). This may be of interest in view of the fact that all species of centipede (> 3000) have only an odd number of leg pairs. This would occur if it were supposed that all the intersections provided positional cues for development of legs. The dorsal intersections may further provide positional cues for development of (e.g) gills, antennae, and wings in some more evolutionarily advanced animal than the centipede, i.e., one with a greater number of Hox clusters.

Fig. 3 illustrates the application of the equations of Appendix A to plant patterns, and intended to show the versatility of the pattern model. The famous Fibonacci as well as other well known patterns, decussate, distichous and whorled are simply obtained. The patterns are of the 'leaves' or 'florets' on a cylindrical stem, shown on a flattened rectangle representing one repeating period. It is suggested that a similar patterning mechanism may be at work to that of animals (two interacting signaling pathways), albeit no doubt with quite different bio-chemicals involved. The phosphoinositides are key biochemicals common to both animals and plants [11, 12], and are speculated to play an analogous role in both.

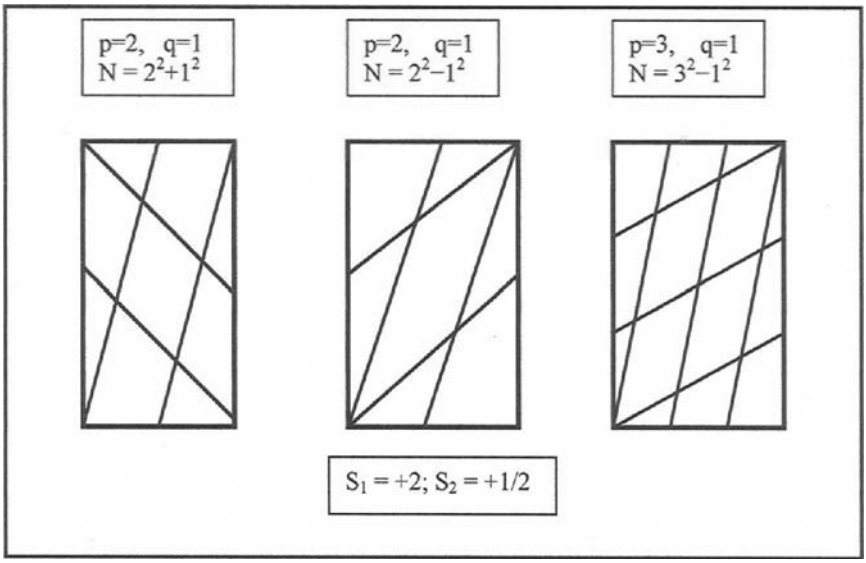


Fig. 3. Three Fibonacci plant patterns are shown. The number of leaves in the three cases shown are $N = 5$, $N = 3$ (middle) and $N = 8$. These are the most frequently seen Fibonacci patterns. The intersections of the two straight lines in each case indicate the position of the leaves. The first leaf is always at the origin in the lower left hand corner. The box below the middle $N = 3$ pattern gives the slopes of the two lines, both positive, when $N = 2^2 - 1^2$. The spiral pattern with $N = 4$ is interestingly not allowed. The model also gives the common decussate pattern ($p = q = 2$), the distichous pattern ($p = q = 1$), as well as the bijugate, trijugate, ... patterns, where $p > q > 0$ but p and q are not relatively primed. When $q = 0$, the slopes of the straight lines are zero and infinite, and the patterns are superposed whorls, with p florets on a level.

The fact that interactions with nearest neighbors are required by any pattern model of necessity introduces the geometry of (in general) curved surfaces. Mathematically this enters through the Laplace operator (Appendix A, eqs.

(12), (13)). The Laplacian enters all known candidates of patterning proposed to date, albeit often in disguised form. The Laplacian of the ligand density in some small area is proportional to the average of the given density at nearest neighbor sites minus the density at the site in question. This implies introduction of the second partial spatial derivative (at least in the limit of small site areas), and this insures that surface curvatures must enter the model. Tissue shape is discussed in more detail in Section 4.

An equivalent picture of the purely pattern aspect of the model is of interest, and serves to emphasizes the model's uniqueness. Appendix B gives a simplest steady state formulation of a quite differently formulated model, one consisting of two 'morphogens' of unspecified composition, where the two quantities try to avoid each other spatially at each growth cycle. The two quantities ('morphogens') form into separate spatial regions that avoid significant overlap, as in the model of Appendix A. The linear version of the model of Appendix B is essentially the same as the linear version of the model of Appendix A. Then two different views of the pattern 'algorithm' emerge: coupled signaling pathways, and two morphogens that avoid each other. In either case, the result is that there is always an interstitial region between the two pattern regions where both morphogens have very small amplitude, i.e., and thus not effective at producing differentiation there. Between the regions, and below some presumed threshold, is a spatial network for which neither morphogen has sufficient potency (i.e., active receptor density in the case of the model of Appendix A) to lead to cellular determination in this region [18,19]. A natural designation for the spatial distribution of stem cells is then given. (A more detailed definition of stem cells is given below: they are to some extent undifferentiated cells, as well as ones having a complement of three or more active genes). The model may now be seen as a tripartite patterning model, with new stem cells finding their niches at each growth cycle, always interspersed between two (much larger) regions slated to undergo determination [54]. It will be assumed that growth originates from these interstitial regions.

Growth is assumed to occur from the interstitial stem cell regions. As a pattern decays from its maximum amplitude, the epithelial (middle) area begins to increase due to increase in the stem cell region. The control of stem cell growth by the two signaling pathways is being relinquished as the pattern decays. This area increase continues until sufficient area allows the next pattern to begin to arise, in accordance with the model. As the pattern amplitude arises beyond a threshold, the set of active genes, the activity of which genes virtually defines the stem cells, are selectively turned off in the regions undergoing differentiation by the activity of the signaling pathways, again with binary differentiation occurring in different spatial regions. Stem cell growth is thus controlled by the pair of signaling pathways, and loss of such control clearly is most often a disaster for the animal. Further, a) the genes defining stem cells begin to return to activity in the stem cells as the pattern amplitude declines, while the new, growing pattern is simultaneously deactivating these same genes in the new pattern regions being established. In particular,

the genes specifying growth, such as c-myc, are being deactivated in the regions undergoing differentiation; and b) the stem cells retain the memory, or labeling by the particular binary genetic path taken in each region, resulting in (e.g.), liver stem cells, neural stem cells, etc., as development proceeds. Denoting that a specific cell has taken binary path G_n or g_n at the nth binary selection, a specified cell may, for example, be labeled $G_1G_2g_3G_4g_5\ldots$

In Appendix B, the mathematical details of the pattern capabilities of the linear steady state model are illustrated by application to patterns in plants. The famous Fibonacci patterns are obtained (fig. 3) as well as the common decussate, distichous, and whorled patterns for the leaf placement on a plant stem [14]. The observed fact of the very infrequent appearance of four leaves in a simple spiral before a repeating pattern occurs defies explanation in terms of natural selection, and in fact a 'four spiral' does not occur at all in either model above, i.e., neither in the linear version of Appendix A, or the model of Appendix B. Perhaps hints regarding a fundamental plant patterning mechanism is hidden in facts such as these, such acting as constraints on the otherwise pre-eminent role of natural selection.

Fig. 2 is a schematic of the solution of the linear (small amplitude) solutions of eqs. (18)-(19). Shape changes to the basic shapes are indicated as amplitudes grow as the total surface area changes. The figure emphasizes the tripartite nature of the model. A side view of the external surface of a toroidal (donut-like) geometry shows numerous growth cycles, several in fig. 2a, and only one further in each of fig. 2b and 2c. Fig. 2a indicates that the same pair of pathways are employed in each growth cycle, as indicated by the same two colors, grey and white. The pattern rectangles in each case represent the binary regions of cellular determination, while the very narrow black stripes separating these larger regions indicate stem cell regions. The integer m represents the number of axial modes which have been generated, while the integer n represents the angular modes. The simplest case would correspond to $m = 1$, and would simply show one grey exterior, with a white region (say) interior to the torus. The appearance of one type of determined region (e.g., grey) is always accompanied by its counterpart (e.g., white). The side view of fig. 1a represents a (thin) donut shape, (of a fixed volume, and elongated along its axis). Each region is separated by smaller circular stem cell regions.

The pattern corresponding to n=1 begins to emerge when the cylinder radius increases to sufficient value [18, 19]. Fig. 2b shows a case where bilateral symmetry has been attained, providing the most elementary through-gut, bilaterally symmetric 'animal', with anterior and posterior having been designated as well at the earliest growth cycle. Fig. 2c shows the next radial mode, the radius having increased further, and corresponding to integers $m = 4$ and $n = 2$.

Growth and subsequent development are considered to occur from the stem cell regions. In fig. 2a, the 'embryo' consists at first of the two patterned parts separated by a stem cell region, and corresponding to $m = 1$, $n = 0$ in the model. This growth pattern may continue along the axial direction for an

unspecified number of growth cycles, corresponding to some total area, before the $n = 1$ bilateral pattern may emerge, as the radius increases. It is presumed that the segment-like patterned pairs in fig. 2b are not distinguished from each other (i.e., by further Hox clusters, as animal complexity increases), and that only one Hox cluster is present, serving to delineate head from tail. A key element of the model is that the two pattern elements must always appear together, and adjacent to each other, i.e., grey adjoining white.

If the intersection of ventral stem cell regions may be considered as specifying the locations of limb placement, (the lower horizontal line of fig. 2c intersecting with the several vertical lines representing circular stem cell regions); then there will always be an odd number of limb pairs, independent of the number of axial 'segments', provided that a limb pair is formed at each of the ventral locations indicated in fig. 2c. This is a consequence of the requirement of the model that one determined region be always accompanied by a second region of different determination. Of note is the observation that of the more than 3,000 species of centipede, with leg pair number from 15 to 191, all have an odd number of leg pairs. This is not to be understood by appeal to natural selection, and indicates an ancient constraint involving pattern formation. Both observations: (a) the lack of the number 'four' in spiral plant patterning as well as: (b) the centipede leg pair number observation are encompassed by the present model.

The positioning of appendages such as wings and antennae may further be specified in more evolutionarily complex animals by the dorsal stem cell intersections of fig. 2c. Head, thorax and abdominal regions have in this case generally been distinguished by addition of further Hox clusters.

The model differs from most others in that two separated differentiated regions are formed at each growth cycle, without having to arbitrarily prescribe numerous thresholds to delineate one differentiated region from the other by arbitrary thresholds each patterning event. A stem cell region is designated at the termination of each growth cycle as separating the two regions. Further, boundaries sharpen between the two regions undergoing determination as growth proceeds, and the receptor density amplitudes rise, due to the presence of nonlinearities in the model. No sources are introduced in the present model; often in other models, sources are invoked as the origin of diffusing morphogens, and their position(s) are arbitrarily designated to produce the desired result; the sources are most often not specified by an auxiliary patterning mechanism.

Two morphogens reacting and diffusing as envisioned by the usual models (e.g., [45]) also leaves open the question of what terminates the process, or what then subsequently initiates the next cycle of reacting-diffusion, presumably consisting of different diffusing bio-chemicals than the previous pair. This latter question is one which must be addressed by any pattern model, including the present one. This remains an open question for all models, including the present. Suffice it to say, if only vaguely, that in the present model the following signaling pair is determined genetically by switches thrown as devel-

opment proceeds, cued somehow by action of the preceding pair of pathways; the fact that the present model is specifically closely coupled to the genes adds some entre to an explanation in terms of genetic switches. The binary decision made at each cycle calls for an explanation beyond "it's a genetic program".

The source of the ligand which binds a given receptor is usually left as an unknown in discussions of signaling pathways. The possibilities for ligand sources seem to be limited. For example, neighboring cells could emit ligand at random times and places; or at some steady rate; neither is envisioned as a viable mechanism. Or a 'source' cell could possibly be the emitter of ligand, in which case the question of placement in space and time of ligand emission arises. This would require specification of a second patterning model to specify the locale of the source(s). The present proposal of stimulated emission of ligand from a cell by stimulating 'like' ligand is the greatly to be preferred alternative to the above. This stimulated emission of "like" ligand is a phenomenon yet to be empirically observed, possibly largely due to experimental difficulty. One suggestion is that the ligand of the stimulated secretion may involve a previously stored membrane cache in anticipation of an appropriate signal, and is (in this restricted sense only) similar to nerve cell stimulated secretion; at least one might suspect that the two processes, secretion from nerve cell junctions and the ligand secretion envisioned in the present model, have a number of elements in common. At any rate, the early appearance in evolution of primitive nerve cells calls for a precursor.

In the embryo, each chromatin site of the DNA in effect holds two tags that are each in a 'poised' state; each potentially can be toggled 'on' (read) and the other site 'off' (don't read) in a permanent fashion upon the appropriate signal. It is proposed that each transcription factor of the two pathways acts at the same time at two locales of master regulatory switches. Pathway '1' by way of its transcription factor leads to activation ('on') at chromatin site #1 (say), and 'off' at chromatin site #2; pathway '2' leads to 'off' at chromatin site #1, and 'on' at site #2. The two different transcription factors each activate an 'on' ('read') tag of their target chromatin spool, at the same time activating the 'off', the 'no-read' tag at the corresponding chromatin site of the other. We remember that these two transcription factors are acting in disjoint spatial regions. This procedure occurs at each cycle of growth, with different (in general) signaling pathways generating two different transcription factors, leading to different cellular determination in the two regions.

Curiosity concerns the possible role of the four genes mentioned in Section 2 that are active in all stem cells. Such gene activity can be taken as a defining property of stem cells. At least three (sox2, oct4 and nanog in mammals, or homologues in other animals), and perhaps more, act as a sort of variable circuit, each associating with another's control sites [53]. There are transcription factors made by these genes, and present at the onset of each growth cycle. The proposal here is that these genes and their transcription factors play a central role, however, in the secretion of 'like' ligand required by the model,

as well as the concomitant prevention of secretion of ligand of second type in the specified region.

As the morphogen amplitudes rise from their initial very small amplitudes, these genes that are active in stem cells must be selectively deactivated in each of the dominant pattern regions, i.e., in those regions destined for determination due to rising morphogen values. This is because these active genes are taken to be the definition of stem cells.

It is proposed that the highly conserved, ubiquitous phosphoinositides (PI) will play a key role in elucidating both the "like ligand" stimulated secretion as well as the switching off of the activity of the "second" ligand secretion pathway in region #1. Phosphoinositides collectively refer to phosphorylated derivatives of phosphatylinositol, and have a pivotal role as precursors of signaling molecules, as well as regulating the actin cytoskeleton. Certain members of this family of PIs directly regulate the actin cytoskeleton by modulating the activity and targeting of actin regulatory proteins [58]. Phosphoinositide head-groups, which can be reversibly phosphorylated to generate seven species, allow phosphoinositides to play a fundamental role in controlling membrane-cytosol interfaces. Their function includes signal transduction at the cell surface as well as regulation of membrane traffic, interaction with the cytoskeleton, and permeability and transport functions of the plasma membrane [23]. Their importance to the present model would seem clear, and such a connection is under ongoing investigation.

A tentative speculation involves a partial explanation of the mechanism involved with secretion of ligand #1, and simultaneous prevention of secretion of ligand #2 (or vice-versa) via the four active stem cell genes mentioned. Let us call these active stem cell genes *a*, *b*, *c* and c-myc, assumed active in all stem calls, constituting in fact a definition of "stem cells", along with their status of being at least to some degree undifferentiated. The gene *c*-myc is thought to be associated with cell division and thus growth. Reversible phosphorylation of the inositol ring at the three positions (designated as '3', '4' and '5') results in the generation of seven PI species. Each of the seven has a unique subcellular distribution, and combined with their high turnover, makes them optimal mediators of signaling events, recruiting cytoskeletal or signaling components to the membrane. Stipulate first that the role of the three active genes (say, a, b, c) of a stem cell is to produce proteins which occupy the three inositol ring 'docking' positions, such a 'full house' indicating that no differentiation is to take place. Imagine that ligand #1 of the signaling pathway pair deactivates gene *a* (for the simplest example) in the stem cell nucleus. Then only the gene products of *b* and *c* will occupy sites (e.g., '3' and '4') on the inositol ring, leading to secretion of ligand #1, and loss of stem cell status due to the deactivation of (one of, in this case, *a*) the three genes. At the same time prevention of secretion of ligand #2 occurs, due to conformational change in the inosital ring upon occupation of ring sites by gene products of *b* and *c*, but not by the products of gene *a*. The gene *c*-myc will be expected to be deactivated also by any particular ligand and

its subsequent transcription factor activity. Then similarly, ligand #2 can be supposed to lead to deactivation of stem cell gene *b*, (predominately) in a different spatial region, leading to the occupation of two different inositol docking sites (e.g., '3' and '5') of the three possible inositol sites, and leading to secretion of ligand #2, prevention of ligand #1 secretion in this region, and loss of stem cell status in favor of cellular determination ala ligand #2. Yet another pair of different type of ligands may deactivate both *a* and *b*, or the pair *a* and *c*, and so on, giving a total of seven possible signaling pathways which may be involved in such a scheme.

At any rate, the observed activity of stem cell genes calls for a roll for them to fill. Here it is speculated that the filling of the three inositide head-group ring sites by three of the gene products (not including those of *c*-myc) provides a sensible possibility; their status as stem cells is preserved as long as the three head-group sites are filled. This must serve as a prediction, namely that the three sites on the inositol ring are occupied in a stem cell. It is further to be understood that stem cells get freshly identified or labeled by the genetic pathway taken at each binary genetic decision, so that one may then end up with 'tail' stem cells, 'liver' stem cells, and 'neural' stem cells, etc., denoting that the cells have taken one or the other genetic pathway.

4 Morphology

One of the most fascinating problems in developmental biology involves understanding how tissues are shaped to produce highly organized embryos. Model systems in various organisms have demonstrated the importance of polarized epithelial cells in the developing embryo. Gastrulation is the ubiquitous stage during which three different germ layers are set aside. It follows highly reproducible patterns of cell movements and rearrangements of epithelial cells in each organism. The result of gastrulation is to produce an 'inside' and 'outside' to the organism, an endoderm and ectoderm, a protective outer covering and a primitive gut. Gastrulation shares similarities in different organisms, and the realization has emerged that complex three-dimensional morphogenetic rearrangements take place from a very limited number of cell biological processes.

Morphological shaping may be accomplished by way of five principal mechanisms, including cellular shape changes. These five may be enumerated as: loss of epithelial polarity during epithelial to mesenchymal transition; apical constriction of epithelial cells; cell intercalation; cell migration and finally, polarized cell division. Pilot and Lecuit [50] have ably reviewed each of these processes.

The proper architecture of tissues demands that epithelial cells display an essential and typical apicobasal polarity. Epithelial cells have the ability to stick closely together, and form protective barriers and to regulate cell proliferation and differentiation; they are constrained in their ability to move

through the formation of strong adhesive attachments at cell junctions. The possibility of tissues to change shape appropriately during development is thus highly constrained. Epithelial cells can nevertheless engage in a large number of rearrangements, the differential remodeling occurring from three sources. These are the apical surface, the basolateral surface or the junctional area, these three membrane domains apparently each requiring their unique signal [28, 51]. Epithelial cells can then progress through numerous and complex morphogenetic pathways at the same time as maintaining the integrity of the epithelial tissue. Epithelial cell shape change is the focus of the morphological shape change here.

Mesenchymal cells play an important role in development of all bilateral animals. Their unpolarized nature endows them with the ability to migrate. Cell migration is essential during a number of processes such as neural crest development and mesoderm morphogenesis. Intercalation of mesenchymal cells into epithelia is an important developmental activity, one often leading to elongation of the anterior-posterior axis. Intercalation of mesenchymal cells into an existing epithelial sheet can occur in several ways, for example, by remodeling specific contacts in an ordered directional pattern so that they progressively exchange places with their neighboring cells. Cells can also intercalate by means of a process akin to cell migration [50].

Following the trajectory of individual epithelial cells is already a daunting task as they slip and slide in the plane of the epithelial sheet, so often it is most advantageous (as here) to follow small groups of cells, considered as an 'average' cell. Modeling the behavior of migrating mesenchymal cells presents an even greater challenge, one not attempted here. Suffice it to say that during migration mesenchymal cells most often follow paths of chemical gradients on epithelia sheets, or gradients of epithelial cell surface cues. We confine our attention here to the geometry of epithelial sheets, as these are most important in maintaining the organism's geometric integrity.

The Mean and Gauss curvatures of the middle surface bisecting the epithelial sheet are H and K respectively. These two functions are necessary to uniquely specify a surface. By surface is meant the middle surface bisecting the apicobasal axis of the epithelial surface. The curvatures are usually defined in terms of their two radii of curvature, or equivalently in terms of their inverse curvature coefficients κ_1 and κ_2. These definitions are $K = \kappa_1\kappa_2$, and $H = (\kappa_1 + \kappa_2)/2$ [24]. These quantities do not, however, immediately make sufficient contact with more relevant ones of interest such as apical area, basal area and height of a given cell. The two curvatures of the middle surface bisecting the sheet of thickness h can be most simply expressed in terms of three quantities: the sheet thickness h, and the (dimensionless) apical and the basal areas. These last two quantities are A and B and are the apical and basal areas divided by the area of a (small, square, imaginary) middle surface at each 'point' of the middle surface bisecting the epithelial sheet. Algebra alone is required to determine the Gauss and mean curvatures of the middle surface as a function of A, B and the sheet thickness h. Fig. 4 illustrates the three

quantities A, B, and h, as well as the middle area A_m for three representative 'cell' shapes. The Gauss curvature K is then expressed as [17–19]

$$K = \left(\frac{2}{h}\right)^2 \left(\frac{A+B}{2-1}\right) \quad . \tag{1}$$

The Mean curvature H is given simply as proportional to the apical-basal area difference by

$$H = \left(\frac{2}{h}\right) \left(\frac{A-B}{4}\right) \quad . \tag{2}$$

Both A and B satisfy $0 \leq (A, B) \leq 4$. However, there is a maximum curvature which may be achieved depending on h, the sheet thickness; if $A\,(\text{or B}) = 4$, then $B\,(\text{or A}) = 0$. The inequality enforced by their definitions, namely $K \leq H^2$, greatly restricts the possible domain of A and B in the A, B plane [17].

A key goal is to determine, or rather provide a convincing model, of the three quantities A, B and h, and thus K and H, in terms of the morphogen densities R_1 and R_2 and their derivatives at each point of the surface. Such specification will 'close' the model as far as being a closed set of coupled differential equations. The boundary condition of continuity of the morphogens and geometry is all that is further required to obtain solutions. The total area of the sheet is specified phenomenologically as an increasing function of time. The total area $A(t)$ is to be supplied "by hand" in the present formulation, as an increasing function. The locale of growth is not determined in the model, a failing that is under consideration at present.

Equations (75) and (76) may be deduced by noting the rhomboid shape of the sides of the unit cell in fig. 4. Three radii may be drawn corresponding to each of the two (in general) different faces at right angles to each other. These three radii for only one face will meet at a common point (the 'origin'); imagine radii running along the sides of the cells of fig. 4. Call these three radii R_{1a}, R_{1b} and R_1, the '1' denoting face "1". Here R_{1a} denotes the radius from the origin to the apical area A_a, R_{1b} denotes the radius from the same origin to the basal area A_b, and R_1 is the distance from the same origin to the (imaginary) middle area A_m. A similar construction is made for the orthogonal "2" face. The definitions of K and H in terms of their radii of curvature are [24] are

$$K = \frac{1}{R_1 R_2} \quad , \quad H = \frac{1}{2}\left(\frac{1}{R_1} + \frac{1}{R_2}\right) \quad . \tag{3}$$

Using the obvious definitions $a_1 a_2 = A_a$, $b_1 b_2 = A_b$ and similarly for A_m, along with the equations $a_1/R_{1a} = b_1/R_{1b}$ and $R_{1a} - R_{1b} = h$, together with similar equations for cell side "2", leads, after a little algebra, to eqs. (12) and (13). The two 'origins' of the two orthogonal sides lie on opposite sides of the epithelial surface for the case of fig. 4 where $K < 0$. The origins both lie below

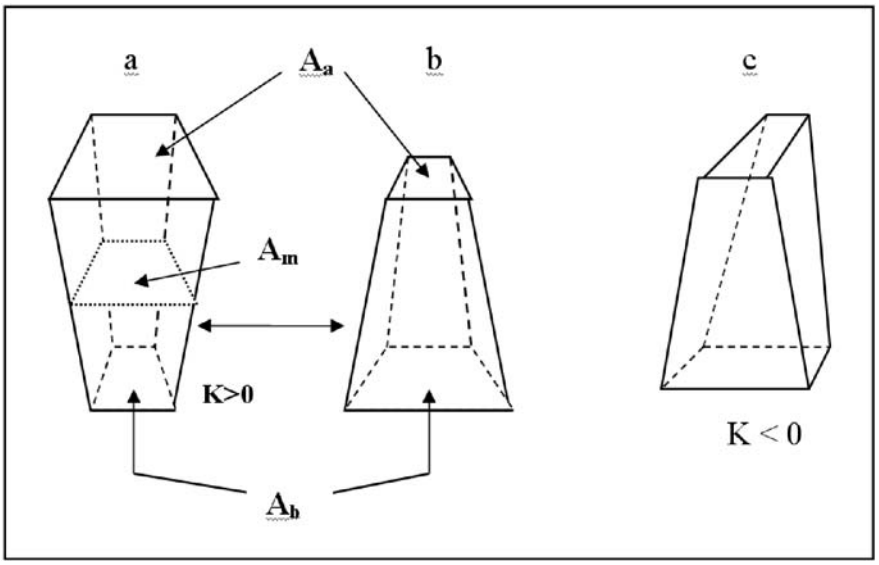


Fig. 4. Only geometrical changes that are considered in the present work are those involving cell shape changes. Consider an epithelium and an 'average' cell. After receiving the appropriate signal, the nucleus produces protein(s) which ultimately migrate to localized positions in the cell, and affect cell shape changes. Fig. 4a shows a situation of negative mean curvature H, and a positive Gauss curvature K; the apical (square) area is smaller than the basal area. Fig. 4b shows a case where H is positive, and again K is positive; the apical area is now larger than the basal area. If such a cell were joined to other similar ones, the result would be a sphere. A solid sphere with radius $= h$ results if the basal area is zero, and the ratio (apical area)/(middle area) $= 4$. Figures 4a and 4b represent situations usually envisioned. Invagination however requires shapes with negative Gauss curvature, and such a possible cell shape is shown labeled as $K < 0$. Such necessarily 'twisted' shapes have apical and basal area as rectangles oriented at right angles, as indicated. As K becomes increasingly more negative, the apical and basal areas both shrink. The limiting case of $A = B = 0$, when the rectangles become simply lines, would give the maximally negative surface, a saddle, with $H = 0$, and $K = -4/h^2$, h being the epithelial sheet thickness.

the cell in fig. 4a, when $H > 0$, and both above in the case of $H < 0$ of the middle cell.

Pattern is connected to geometry through the metric factor g which enters the Laplacian of the pattern (eq. (80) below). The metric factor g is determined by the Gauss curvature K via the exact Gauss equation

$$\nabla^2 \log_e(g) = -2K \ . \tag{4}$$

The key Gauss equation, eq. (77) has been expressed in special "conformal" coordinates. The coordinates have been chosen in this particular way for

convenience, as well as simplicity. The coordinates are always taken as embedded in the surface. The surface is covered everywhere by (infinitesimally) small squares whose sizes vary, so that the squared distance between any two nearby points designated by coordinates u and v is given by the Pythagoras-like formula

$$ds^2 = g.(du^2 + dv^2) \ . \tag{5}$$

The area of each small square is then given by

$$dA = g(u, v).(du\, dv) \ . \tag{6}$$

The function g carries the dimension of area, while the coordinates u and v are simply dimensionless position markers. Note that neither the 'metric' function g nor the coordinates have a physical interpretation, but rather only certain 'real' combinations do, such as dA and ds. (The metric of the space is more properly a diagonal 2×2 tensor, and here, one with a single diagonal element g). The total area, which is to be thought of as a increasing function of time, mimicking growth, is the integral over the coordinates u and v of eq. (79).

There is a small 'price' to pay for the adoption of the conformal coordinates, and the resultant single metric coefficient function g. However, these coordinates seem to be the most useful for surfaces which are changing shape continually. Consider the simplest case of a sphere of radius R. In this case, the (simplest) metric function is $g = R^2/\cosh^2(u)$, giving a bell shaped curve for g, maximum at $u = 0$ and quickly falling to zero at the 'poles', as $u \to \pm\infty$. The coordinate u runs from one pole to the other, and $-\infty < u < +\infty$, excluding the mathematical points at the two poles. In practice, since the slope (derivative) of the morphogen functions must be zero at the poles, the coordinate u may be generally taken in numerical work as $-10 < u < +10$. It then is to be remembered that the (infinitesimal) distance along the direction of u is $ds = \sqrt{g}\, du$. This simple example of the sphere serves to accentuate the differences from the usual spherical polar coordinates used for spherical shapes. The spherical polar coordinates begin to lose their meaning as invagination proceeds, however.

The invariant Laplacian (e.g., eqs. (12) and (13)) in these 'conformal' coordinates is

$$\nabla^2 = \frac{1}{g(u, v)} \left(\frac{\partial^2}{\partial u^2} + \frac{\partial^2}{\partial v^2} \right) \ , \tag{7}$$

displaying the metric function explicitly in the denominator. Any analytic transformation of coordinates will preserve the covariant nature of the coordinate system, and the form of eqs. (77) – (80).

The Gauss curvature K is assumed to be a function of both morphogen densities R_1 and R_2 and their derivatives, when a closed system is achieved in the steady state. The (normed) receptor densities are assumed to be invariants. Specific model forms for this functional dependence are discussed in Section 5, where it will be seen that there are a number of important constraints

on this choice. Once this specification of the Gauss curvature K is given in terms of the receptor densities, three coupled second order partial differential equations exist, which must then be solved in the steady state for the three variables: the active receptor densities R_1 and R_2, and the metric function g.

The present model supposes that the active receptors will act to change cell shape by a pathway to the genome, and not by a possibly more direct pathway.

The active receptors set in motion numerous intermediate cytoplasmic interactions, ones not specified in the model, and these ultimately activate transcription factors in the nucleus. The complex cascade of intermediate interactions is different for each signaling pathway, but the net result is in all cases activation of specific master regulatory genes via transcription factors [6]. This "canonical" pathway is expected to produce, among other things, at least three protein factors affecting the cell and tissue shape, i.e., factors most locally determining the three quantities h, A and B, or the Mean and Gauss curvatures by eqs. (4.1) and (4.2) at each point of the middle surface. Epithelial sheets are characterized by cells which show columnar or cuboidal shape, as well as pronounced apical-basal polarity. Tumors of epithelial origin lose these characteristics. Candidate proteins which affect the model factors A, B and h are expected to be factors that affect the actin cytoskeleton, the microfilaments, the microtubules, and the intermediate filaments. The actin cytoskeleton mediates a variety of essential biological functions in all eukaryotic cells, in addition to providing a structural framework around which cell shape and polarity coalesce [28, 32, 51]. The adhesion molecules such as E-cadherin certainly also play a key role in tissue shape. Cadherin based adherens junctions are found between polarized epithelial cells, and are also intimately linked to the actin cytoskeleton, and Rho and Rac are required for their assembly. There is now good evidence of the crucial effects of the Rho GTPases on the organization of the actin cytoskeleton in all eukaryotic cells. It is reasonable that Rho and Rac affect signaling pathways since adherens junctions are known to participate in signal transduction pathways that affect gene transcription.

Here again, the ubiquitous phosphoinositides (PI) are expected to play an essential role. The function of actin-cytoskeleton regulation by phosphoinositides has been reviewed by Maruta et al. [41]. Tapon and Hall [58] and also Di Paolo and De Camilli [23] have discussed the role of the small GTPases Rho, Rac and Cdc42 and their role in regulating the organization of the actin cytoskeleton. There is a connection of the actin cytoskeleton with the E-cadherin adhesion sites as well, as there is a projection of the microtubules into the sites of adhesion.

Three factors have been identified as suspects that affect cell shape most profoundly. These are the three gene products Scrib, Dlg and Lgl. Mutations in any of the genes which produce the three factors lead to loss of apical-basal polarity and overgrowth of epithelia. The Scrib and Dlg co-localize and overlap with Lgl in epithelia, and these three 'tumor suppressors' act together in a

common pathway to regulate both cell polarity and growth control [4, 34, 37]. Gene products encoded by tumor suppressor genes are of several different categories, such as cell surface receptors of the signaling pathways (e.g., Frizzled and Patched), cell adhesion molecules such as E-cadherin, cytoplasmic proteins such as PTEN and APC, and nuclear factors such as p53 [57]. The APC protein is involved in regulating the cellular cytoskeleton, and also involved with the efficient turnover of β-catenin in the Wnt pathway [40]. It is of most interest to note that most of the tumor suppressor gene products are involved in signal transduction pathways involved with cell growth or apoptosis. The interaction or relationship of the GTPases with the three proteins, Scrib, Dlg and Lgl is however not at present known. More specifically, the connection between the three proteins Scrib, Lgl and Dlg and changes in the quantities A, B and h are not at present known, but is knowledge to be desired.

Certain reasonable assumptions are made in the sample numerical work shown in fig. 5. Gastrulation or invagination occurs easily however if the mean curvature H, eq. (76), is taken as simply proportional to the morphogen difference $(R_1 - fR_2)$, where $f \sim 1$. The Gauss curvature K is then found from $K = H^2 - D^2$, when D^2 is taken as proportional to the squared gradient of the morphogen difference [16, 17]. The invariant D^2 is the square of the difference of the two curvature coefficients κ_1 and κ_2 divided by two. Much ignorance prevails at present concerning the connection between A, B and the three protein factors of the previous paragraph, or of the role of the phosphoinosotides. It is reasonably safe to say that the particular docking sites which are filled on the inositide ring, as well as the density of such sites filled, can determine the apical surface area of a cell. In the model calculations of fig. 4, the apical area A has been taken to be linear in R_1, while the basal area is assumed linear in R_2, so that the mean curvature in eq. (76) is linear in the difference of the two densities.

It remains also to determine the variation of h, the sheet thickness, with the model variables; the sheet thickness h is simply held constant in the illustrative numerical work of fig. 5. It seems reasonable that the thickness may be affected by the adhesion molecules; when adhesion is strong the cells squeeze together tightly, and the cells may be expected to elongate, and when the cell adhesion is weaker the cells decrease in height. This is consistent with cells that are dividing to temporarily lose their adhesion and form into a more rounded shape, with normal adhesion not being reestablished in cancerous division. Further thought regarding the possible dependence of the height h on the active receptor densities is required.

Farge found that compressing the Drosophila embryo induced β-catenin to move into the cell nucleus, where it activates the transcription factors Tcf. Interestingly, a gene called *twist*, know to be involved in dorsal-ventral patterning, is normally expressed only in the ventral region of the embryo. But within eight minutes of artificial compression, *twist* expression had spread to encompass the entire embryo. This suggests that a pathway that responds to mechanical stress could affect patterning in the embryo. How is mechanical

force translated into a change in gene expression? [27, 55]. One protein that may couple these events is β-catenin, which is involved in both cell adhesion and gene activation, and is a key element in the Wnt signaling pathway. Inhibiting β-catenin, on the other hand, suppressed the stress-activating expression of *twist*. This is an artificial, external compression that does not occur during embryogenesis, so that it is important to find if such does occur in actual embryos not undergoing such artificial stresses.

Farge found that in certain cells in the invaginating anterior mesoderm and foregut, increasing compression of these cells coincided with increasing expression of the *twist* gene. Further, *twist* expression required β-catenin activity in the nucleus. So it seems that β-catenin accumulation in the nucleus may be caused not only by intracellular signaling mediated by ligands from neighboring cells, as in the canonical Wnt pathway, but possibly also simply by relocation of β-catenin from near the cell surface to the cytoplasm and then into the nucleus. This would mean that β-catenin function may respond to two different stimuli, mechanical deformation and intercellular signaling.

A question is whether β-catenin is a key player in all developing tissues. Many targets other than *twist* are regulated by β-catenin. It is possible that all invaginations involve β-catenin and twist. It would not be terribly surprising if the Wnt pathway, or a homologue, is very frequently one of the two signaling pathways, considering its ancient ancestry [7, 62], and its well known interaction with the cell adhesive module, c-myc and the protein APC. Wnt transcriptional enhancers are also known to integrate input from multiple signaling pathways. For example, inputs from both Wnt and Dpp pathways are required in Drosophila midgut development [52], and similarly the Drosophila EGF receptor and Wnt apparently interact at the level of target gene promoters [48]. The large number of different Wnt ligands and receptors (Fz) may argue for a radiation near the Cambrian from a single original Wnt/Fz ligand and receptor.

Perhaps the most completely understood example of tissue shape deformations is the mesoderm invagination that initiates Drosophila embryo gastrulation. In this case the *twist* and *snail* ventral genes are required for proper mesoderm invagination. Both genes are required for the simultaneous cell apical constrictions that induce the bending force necessary for the mesoderm invagination [8]. If mechanical deformation is what gives rise to β-catenin migration from the cell surface into the nucleus, then what gives rise to the mechanical deformation at that locale in the first place; just *there* and nowhere else? In the present model, pattern emerges at each growth cycle spontaneously from new tissue.

The present model proposes that geometric deformation changes pattern, by way of change in the g, the metric (eq. (79)) that occurs in the Laplacian (eq. (80)). Geometrical changes affect pattern, which in turn affect the genome; on the basis of the model, geometrical deformation would be expected to lead to gene activation due to altered pattern. Migration of β-catenin, from the cell membrane, rather than via the Wnt pathway, into the nucleus is then

an unnecessary explanation from the point of view of the model. Which is not to say incorrect; perhaps there is another, more direct connection between geometry and gene activation than presently proposed via signaling pathways, a connection between geometry and gene activation not yet uncovered.

Again the question of making a smooth transition from one pair of signaling pathways to the next comes up. The model supposes that the initial values of the geometry factor g, as well as the Gauss and mean curvatures K and H at the beginning of the next growth cycle will be those that existed at the close of the previous growth cycle, that is those values corresponding to the maximum amplitudes of the receptor densities of the previous cycle.

5 Notes on Sample Numerical Results

Fig. 5 shows numerical results of the model, for two total areas, as a ratio A/A_0. Until the sphere (blastoderm) reaches a critical area A_0, no solution exists for the pattern, and the morphogens remain zero until that point. The fact that the pattern cannot appear until a critical area is achieved is an important fact due to the model's structure. As area increases beyond A_0, the dimensionless morphogens $\phi_1 \equiv R_1/R_0$ and $\phi_2 \equiv fR_2/R_0$ ($f \approx 1$) increase, with their respective maxima at the two 'poles' on the axis. Here R_0 is a normalizing density, such as the maximum allowed. The right column of fig. 5 shows only the middle surface through the epithelia of height h, where h has been taken as $h^2/A_0 = 0.1$.

The Mean curvature has been taken to be of the form

$$H = \sqrt{\frac{4\pi}{A}}\,(1 + \lambda_1(\phi_1 - \phi_2)) \quad . \tag{8}$$

This says that the apical area is increased as the morphogen density R_1 increases, and decreased as morphogen density R_2 is increased. The basal areas change in an inverse relation; the basal area decreases as the apical area increases, etc. Clearly more nonlinear terms can easily be added to this expression. The Gauss curvature is now taken as

$$K = H^2 - \lambda_2^2\,(\nabla(\phi_1 - \phi_2))^2 \quad . \tag{9}$$

This means that the two principal curvatures are taken as

$$\kappa_1 = H + \lambda_2|\nabla(\phi_1 - \phi_2)| \ , \quad \text{and} \quad \kappa_2 = H - \lambda_2|\nabla(\phi_1 - \phi_2)| \ . \tag{10}$$

The apical and basal areas are then given by

$$A = \left(\frac{1+hH}{2}\right)^2 - \left(\frac{hD}{2}\right)^2 \ , \quad \text{and} \quad B = \left(\frac{1-hH}{2}\right)^2 - \left(\frac{hD}{2}\right)^2 \ , \tag{11}$$

displaying the constraint $(h/2)^2 D^2 \leq 1$. It is noted that the invariant squared gradient entering eq. (9) is simply (e.g.) $(\mathrm{grad}(L))^2 = [(\partial L/\partial u)^2 +$

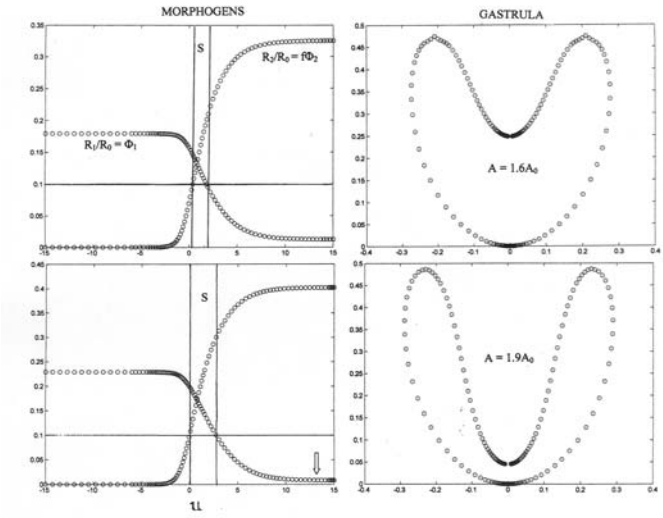


Fig. 5. Shown are numerical solutions of the model in steady state, with pattern coupled to cell shape change. Three coupled second order ordinary differential equations are integrated in axial symmetry. Only one coordinate (u) is needed to specify position on the invaginating sphere. The left column shows the morphogen R_1 and R_2 as a function of the coordinate u, for two different total areas A/A_0. No pattern emerges until $A > A_0$; an animal must have a minimum total epithelial surface area A_0, or corresponding cell number. The coordinate u may be viewed as running from ~ -10 at the exterior pole, up to $\sim +10$ at the 'pole' in the interior of the invagination, both on the symmetry axis. As total area increases, the amplitudes R_1 and R_2 increase, up to a maximum for each area. A small arrow indicates the (positive definite) R_1 decreasing toward zero in the region where R_2 is maximum, indicating an approaching cutoff. The right column shows the corresponding (middle) surfaces. From a point at the maximum height shown on the 'gastrula' to some distance into the invagination, the Gauss curvature must be negative, and cell shapes then corresponds to a 'twisted' shape shown in fig. 4c. Only a handful of parameters are required in the simplest model: (a) the two inverse lengths k_1 and k_2, (b) the ratio $f = \alpha/\beta$, (c) a parameter c entering the $N.L.$ terms in eqs. (12) and (13), and (d) the two parameters entering the curvatures H and K, namely λ_1 and λ_2. The parameter A/A_0 is taken to be a slowly increasing function of time, mimicking growth. The interesting numerical problem of non-axial symmetry has not been attempted as yet.

$(\partial L/\partial v)^2)]/g(u,v)$ in these special coordinates, showing the dimensions of 1/(area) explicitly.

The two points on the axis shown on the right side in fig. 5 must have the 'sphere value' $K = H^2$, and the morphogens must have zero derivatives there. This is of course a main motivation for choosing the gradient of morphogens in eq. (9). A second reason is that the twisted shape shown in fig. 5 has a rectangular apical surface perpendicular to a rectangular basal area, indicating that a direction (i.e., a gradient) must be available to specify the orientation. Further, the squared gradient is an invariant, required by any equation for the invariant Gauss curvature [24]; it also at once has the correct dimensions for Gauss curvature of 1/area if λ_2 is a dimensionless number (taken about unity in numerical work). In the model, the R's, the active receptors densities, are assumed to be normed scalar quantities, and the ϕ's (namely, the receptor densities divided by the density R_0) are these normed densities. All equations as present must be independent of coordinates, 'covariant', since coordinates are always a free choice of the observer. The parameter λ_2 was taken to be 0.7 in the numerical work shown in fig. 5, and λ_1 was taken as 7.0.

In eq. (8) for the mean curvature H, a factor $\sim 1/\sqrt{\text{Area}}$ is necessary so that, as long as the morphogen densities are zero, the answer for a sphere must be that $H = 1/R$ and $K = 1/R^2$, with R the (increasing) sphere radius. The squared gradient factor in K in eq. (9) provides for the correct dimension for both H and K after invagination begins, and gradients begin to appear. The eq. (9) was then used along with eq. (8) in eq. (77) to give a solution for the metric factor g, which in turn was used in the Laplacian (eq. (80)) (which entered eqs. (12) and (13)), thus coupling surface shape to pattern.

Equations (12) – (18) plus eq. (77) in steady state ($\partial/\partial t = 0$) now form a closed set of three coupled second order partial differential equations. The numerical solution has to date only been attempted in the case of axial symmetry, and one example (one parameter set) is shown by fig. 5. The stem cell region is shown as occurring below a threshold shown as a horizontal line. The numerical solution for the more difficult problem when axial symmetry is broken should yield interesting new insights. This more difficult computational problem is not being attempted by the present author.

Several comments regarding the invagination of closed surfaces are interesting, especially regarding use as constraints in numerical work. Consider the surface shown in the right side of fig. 5. Restrict attention to axially symmetric closed surfaces, and suppose that the surface is described for the moment by two 'exterior' coordinates ρ and z, where ρ measures the distance from the axis at any height z along the axis. At the point at the very top of the figure for the blastula in fig. 5, at a point which describes a circle, a flat ($K = 0$) circular disc may be substituted for the entire invaginating part of the surface. This can occur here without introducing a discontinuity, since at this point, the derivative $(dz/d\rho)$ is zero just as the surface begins to 'curve over' to begin invagination, and $K = 0$ around this circle. The amazing Gauss-Bonnet theorem of surfaces says that the total area integral of the Gauss curvature

K for any surface which is topologically equivalent to a sphere, such as the invaginating surface here considered, has a value of 4π [24]. This means that the surface integral of K over the non-invaginated part alone is 4π, i.e., the 'ectoderm' area up to the maximum points (a circle) of fig. 5. The conclusion is then that the *surface integral of K over the invaginated part of the surface must be zero, regardless of its shape!* This is non-intuitive – but correct. This then requires that the Gauss curvature becomes negative in the invaginated region near the dorsal lip, only to again become positive as the axis is approached, when $H < 0$, and $K > 0$.

The fact that the Gauss curvature must be negative near the beginning of invagination means that the cell shape in the region of the gastrula lip must take on the 'twisted' shape indicated in fig. 5c. Usual consideration of cell shape changes in development include the changes in apical and basal areas shown in fig. 4a, and 4b, but do not consider the necessity of including the 'twisted' shape shown in fig. 4c. In fig. 4c, the rectangular apical and basal areas of the cell in the farthest right picture are always at right angles to each other. Negative K implies 'twisted' shape, and is a necessity during invagination. The elongated rectangular apical length of a 'cell' shape (fig. 4c) of negative K points into the invagination, while the elongated basal length is oriented at right angles to the apical area, i.e., along the circumference of the invagination. It is interesting to speculate concerning a connection between the 'twist' gene of Farge [27] and the required 'twisted' cell shape in this region. It is suggested by this result that the Wnt pathway is always one of the pathways involved in invagination during embryogenesis.

Consider next the case of the torus or 'donut' topology. As the simplest model of an animal with a through gut, we imagine a surface topologically equivalent to a 'donut'. In fig. 2, a 'donut' topology is shown in side view, an ordinary 'donut' or torus stretched along its axis. The Gauss-Bonnet integral is now equal to zero. By inserting a flat disc with $K = 0$ at **both** ends of the surface at the two circles of $K = 0$, it is seen that the surface integral of K over the interior (endoderm) of the 'animal' must be -4π, as the integral over the exterior (ectoderm) surface must be $+4\pi$. This is a result independent of the shape of the interior or exterior surfaces, independent of the number of wiggles along the axis, etc.

6 Summary and Outlook

A model of pattern coupled to form has been discussed. It is a model attempting to hold out the promise, or the possibility, that the human mind may overarch some of the vast biochemical complexity of the actual situation, and comprehend in some sense many of the most elementary facets of embryogenesis. Certainly any attempt to couple all molecules of a realistic system, even for a restricted module, is most likely to escape comprehension. A (paraphrased) quote of E. Wigner capsulizes the frequent dilemma of very complex

computer computations: "The computer seems to understand the answer, but I don't". Use of the word 'comprehend' must be used timorously in any case.

Two interacting signaling pathways provide patterning at each time interval, with a second pair taking over in a continuous manner where the previous pair leaves off. The phosphoinositides, a seven faceted biochemical membrane entity facilitates both the secreted ligand as well as the actin cytoskeleton remodeling aspects of the patterning.

In the sense of 'modules', the present model is minimalist of any attempt to reduce the subject to modules. Unfortunately, a very good analogy cannot be made to phenomenological thermodynamics and its corresponding 'reductionist' statistical mechanics, only partly because the latter discipline is so vastly less complex than its biological 'reductionist' counterpart. And of course, thermodynamics came in time well before its reductionist underpinnings in terms of molecules, while the present work attempts to reverse that time order. There has been here an attempt however to provide a first sketch of possible molecular underpinnings of the model, an effort certainly to be viewed as preliminary, if not wrong, or a work in progress.

Elucidation of the molecular underpinnings of patterning, epithelial geometry and their interactions must then await further and ongoing empirical investigations. Several 'modules' will possibly emerge that could make contact with the present model: one specifically oriented to the patterning aspect of the model, another directed to the epithelial, tissue geometrical changes, and yet another specializing in the interactions of the pattern with geometry due to cell shape changes. These modules may be in turn deconstructed into smaller ones. The present tissue shape changes are limited to those arising from cell shape changes only, and not, for example, due to intercalation. It is expected that intercalation per se leads primarily to growth, but does not appreciably affect the tissue geometry without an accompanying cell shape change.

A suggestion is that there is a 'sub-code' of the genetic code, one that is very evolutionarily conserved, and one that acts at each growth cycle to 'pick out' the appropriate interacting signaling pathway pair from the possibilities. Such a code would most likely involve the several active (four?) genes that remain active in all stem cells, providing their very definition. It is important to note that just how the animal picks out a subsequent signaling pair is not at hand, but involves the genome code in some unknown way.

A further suggestion is that the ubiquitous Wnt pathway and homologues will be the most frequent of the pair of interacting pathways. This is in view of its antiquity [62], and its interaction with so many key elements of development, such as proteins E-cadherin, Gsk-3, APC and the gene c-myc. The frequency with which the Wnt pathway pops up in early development makes it almost seem as if it may always, or very often at any rate, be one of the pair of signaling pathways.

It is to be emphasized that the principal components of the model have not as yet been discovered empirically. The two main aspects are: (1) produc-

tion of new extracellular ligand is stimulated by activation of its receptor by a like ligand, and (2) activation of one signaling pathway of a pair acts to deactivate its coupled partner. These two events must serve as principal predictions of the model. It seems very likely that phosphoinositides will play a key role in patterning and shape change, and details of this connection need further investigation. Through interactions mediated by their headgroups, which can be reversibly phosphorylated to generate seven species, phosphoinositides (PI) play a fundamental role in controlling membrane-cytosol interfaces. Their function includes signal transduction at the cell surface as well as regulation of membrane traffic, interaction with the cytoskeleton, and permeability and transport functions of the plasma membrane [23]. Then the two principal elements of the model are seen as the interaction of two signaling pathways, along with the essential facilitation of secretion and actin reorganization via the seven PI species.

Any model which involves a Laplacian, or interactions of a cell with nearest neighbor cells, necessarily contains the metric of the surface, and thus the information concerning surface geometry. It is a further present assumption that this metric (or metric tensor) can be found from having the Gauss curvature be a function of the two morphogens and their gradients at each growth cycle. An effective functional form of the Gauss curvature is surprisingly simple, at least in the case of the simplest 'gastrula' development presented, and perhaps holds some promise of there being a universal functional form.

The most difficult area for any model is in making a smooth transition from one growth cycle to the next, and research continues on this area. The present picture of this transition is a tentative one. After one pair of receptor densities reaches maximum amplitude, a succeeding pair begins to become active, at first at low amplitude, when the appropriate area becomes available. The geometry g, as well as K and H are then taken as proceeding to their next values by starting from their existing values, corresponding to the receptor density maxima functions of the preceding growth cycle; these functions over the surface are taken as the initial values for the following changes of g, K and H. As the succeeding pair receptor density amplitude begins to grow, the previous pair decays, as the total area increases. Growth proceeds from the stem cell regions.

"There seems to be things that I can almost get hold of, and think about; but when I am just on the point of seizing them, they start away, like slippery things."

(Nathaniel Hawthorne, $\sim$ 1862)

Acknowledgement. I wish to thank Lev Beloussov and Alex Spirov for helpful conversations.

Appendix A

This appendix provides the mathematics of the model described in the text, in Section 3. Attention is focused on a small cluster of cells, $\sim$ five-ten, when use of such terms as 'ligand density' and 'receptor density' has meaning. The cells are to be thought of as being part of a closed epithelial surface, so that the densities of the model have dimensions of 'number/area'. Variation of the morphogens (the R's or L's) along the apical-basal cell direction of a cell is not considered, or rather thought of as being an averaged value in this dimension.

First of all, each such a 'cell', or rather cell cluster, produces an increase in ligand density according to the level of receptor activation of like kind. Morphogen R_1, an activated receptor, stimulates production of ligand L_1, otherwise the process would be limited to a purely local process in the absence of "like-ligand" production, with the particular cell in question then acting as a 'sink'.

The second key element in the model is that activation of a particular pathway acts to inactivate the other; as R_1 increases, the level of ligand production L_2 is decreased, and similarly for R_2. This is the culmination of a complex cascading biochemical pathway, as also is the production of like ligand, but these 'modules' are caricatured simply as below.

The equations representing such a process are then able to be written at once, and are

$$\frac{\partial L_1}{\partial t} = D_1 \nabla^2 L_1 + \alpha R_1 - \beta R_2 + NL \ , \tag{12}$$

$$\frac{\partial L_2}{\partial t} = D_2 \nabla^2 L_2 + \beta R_2 - \alpha R_1 + NL \ , \tag{13}$$

$$\frac{\partial R_1}{\partial t} = C_1 \bar{R}_1 L_1 - \mu R_1 \ , \tag{14}$$

$$\frac{\partial R_2}{\partial t} = C_2 \bar{R}_2 L_2 - \nu R_2 \ . \tag{15}$$

The first two terms in eqs. (12), (13) represent in the usual way diffusion of the ligands in the extracellular space. All parameters in the model (e.g., α, β, D_1, C_1, μ, ν) have positive values, as do also, of course, the densities L_1, L_2, R_1 and R_2.

The terms αR_1 in eq. (12) and βR_2 in eq. (13) represent the production of 'like' ligand by the corresponding activated receptor. These same terms are used with minus signs in the same equations to represent the fact that activation of receptors of density R_2 deactivate or turn off production of free ligands of density L_1, and vice versa. A region of high activation of one morphogen then implies low activation of the second. The term NL on the r/h/s of eqs. (12) and (13) indicate that there are expected to be nonlinearities; saturation effects set in for large enough values of either active receptor density. These NL terms will be expected to play a role in sharpening the boundaries between the two regions undergoing differentiation.

The transmembrane receptors, which reside in the lateral cell plasma membrane, are relatively immobile. The respective activated densities decay at rates μ and ν, and this 'decay' returns the receptors to their inactive state. Two first terms on the right side of eqs.(14), (15) say that there is a positive rate of change of R_1 or R_2 proportional to both the density of empty receptor sites ($\bar{R}_1$, $\bar{R}_2$) and also to the density of free ligands at the particular local cell site. The density of empty sites may be obtained from the expression

$$R_1 + \bar{R}_1 = R_0 + \eta R_1 \ , \quad (R_0 = \text{ const.}) \ , \tag{16}$$

where the last term on the r/h/s expresses the possibility that the total number of receptors of each type (e.g., '1') increases with activation of that same type receptor, and new (empty) receptors are thus added. Then the empty receptor site density is

$$\bar{R}_1 = R_0 - (1 - \eta) R_1 \tag{17}$$

and similarly for type '2'. The values $\eta = 0$ implies that there is no receptor augmentation $\sim R_1$, while $\eta \sim 1$ implies either that there is a new empty receptor created for (almost) every one occupied, or that there are very many more empty sites than occupied ones. When eq. (17) and the analogous equation for type '2' is used in eqs. (14) and (15) to eliminate the unoccupied site densities, the model then comprises four coupled first order differential equations for four unknowns. The coupling from epithelial shape to morphogen is discussed elsewhere ([16, 17], and also in Section 4.

The small amplitude, time independent ($\partial/\partial t = 0$) version of eqs. (12) – (17) are simply the Helmholtz and Laplace equations

$$\nabla^2 (R_1 - f R_2) + k^2 (R_1 - f R_2) = 0 \ , \tag{18}$$

$$\nabla^2 \left(\frac{R_1}{k_1^2} + \frac{f R_2}{k_2^2} \right) = 0 \ . \tag{19}$$

The definitions $k_2 = k_1^2 + k_2^2$, $f = \beta/\alpha$, $k_1^2 = \alpha C_1 R_0/(D_1 \mu)$, and $k_2^2 = \beta C_2 R_0/(D_2 \nu)$ have been used. The k's are the two inverse lengths of the model. Several forms may serve to model the $N.L$ terms on the r/h/s of eqs. (12) and (13). Such nonlinearities may simply serve as saturation to 'turn off' the continued rise of the linear terms on the right side in eqs. (12) and (13). The simplest, and the one used in present simulations is to let

$$R_1 - \left(\frac{\beta}{\alpha} \right) R_2 \rightarrow \frac{R_1 - (\beta/\alpha) R_2}{1 + ((R_1 - (\beta/\alpha) R_2)/c)^n} \ .$$

The constant c is ~ 1, and $n = 2$ in present numerical work (fig. 5 and Section 5). The highly nonlinear and exact Gauss equation, eq. (77), carries the dominant nonlinearity of the pattern coupling pattern to surface shape. All models must involve the Laplacian as in eqs. (12) and (13), and this operator involves the surface geometry. Besides the built in nonlinearity due to the log term of the Gauss equation, the assumed linear relation of the mean

curvature H may easily and sensibly be made nonlinear. Other forms of the nonlinear terms of the pattern equations, the *N.L.* of eqs. (12) and (13), lead to somewhat different surface shapes for the invagination.

The 'morphogens' of the present work are taken as

$$\phi_1 = \frac{R_1}{R_0} \ , \quad \text{and} \quad \phi_2 = \left(\frac{\beta}{\alpha}\right)\frac{R_2}{R_0} \equiv \frac{fR_2}{R_0} \ . \tag{20}$$

The very common Wnt pathway, for example, has two known modes of action, one that bypasses the nucleus, and a second 'canonical' pathway leading to gene activation via stabilization of nuclear β-catenin. The model predicts that the 'non-canonical' path bypassing the nucleus acts (at least in part) to release stored Wnt ligand relatively rapidly. Other pathways involved with patterning also act in a similar way.

When each of the four variables of the linear version of eqs. (12) – (17) is taken as proportional to $\exp(i\mathbf{k}\cdot\mathbf{x} - st)$, then a fourth order equation is obtained for the allowed solutions for s. The parameter values are sought which give one solution of $F(s) = 0$ for a negative value of s. A function $G(S)$ is determined by

$$G(S) = \left[(S - x^2 d)(S - m) - q\right] . \left[\left(S - \frac{x^2}{d}\right)\left(S - \frac{1}{m}\right) - p\right] \ , \tag{21}$$

and the four roots of $F(S)$ are determined by the equation $G(S) = p.q$. The four roots of G are all real. The following dimensionless definitions apply in eq. (21) [16]:

$$\text{(a)}\, S = s\sqrt{\mu\nu} \ ; \quad \text{(b)}\, x^2 = k^2\sqrt{\frac{D_1 D_2}{\mu\nu}} \ ; \quad \text{(c)}\, d = \sqrt{\frac{D_1}{D_2}} \ ;$$

$$\text{(d)}\, m = \sqrt{\frac{\mu}{\nu}} \ ; \quad \text{(e)}\, p = \beta C_2 R_0 \mu\nu \ ; \quad \text{(f)}\, q = \frac{\alpha C_1 R_0}{\mu\nu} \ . \tag{22}$$

When $x^2 < \{q/(md) + p(md)\}$ then a negative root of $F(s)$ appears, and the system becomes exponentially unstable. The inverse of k is a length, and there exists a minimum length for which pattern formation can occur, corresponding to minimum size of an animal. It is easy to check that when $d = m = 1$ and $p = q$ there are four real roots, when there is no possibility of traveling waves. As an example of a parameter set allowing traveling waves, the parameter set $x = 2$, $d = 4$, $m = 6$, $p = q = 10$ gives two imaginary roots. Fig. 1b shows an example of the function $G(S)$ and the constant pq for the case of two real roots and two complex roots, when traveling waves can occur; the maxima of $G(S)$ then lie below the constant $p.q$.

Appendix B

The aim is to give an alternate point of view of the model of Appendix A, as well as give a second example of the patterning aspects of the model of eqs.

(12) – (17). There, the model was linked to the interaction of two signaling pathways. In this appendix, it is shown, at least in the linear regime and in steady state, when the amplitudes are small, that another way to view the model is that there are two 'morphogens', each of which attempts to avoid the other. In a region where the overlap between the two substances is smallest, a third region exists, a 'niche' associated with 'stem cells', from which growth is assumed to emanate.

Introduce two morphogens ϕ_1 and ϕ_2. These could in general represent any two different chemicals, etc., called 'morphogens'. In the absence of interaction, assume that the two 'morphogens' will spread themselves as smoothly as possible over the available space. This is described mathematically by saying that the integral over the entire spatial region of the sum of the squares of the gradients of both morphogens is a minimum. With no interaction, this tells us that each morphogen separately obeys the Laplace equation. Now a term is added to the integral so that it is ('energetically') unfavorable to have the two morphogens overlap; this function is called J, and is a rapidly rising function of the squared difference of the morphogens; the greater is the extent that the morphogens do not overlap, the lower the 'energy'. An example of a particular J is

$$J\left((\phi_1 - \phi_2)^2\right) = a\left(1 - \exp\left(-b\left((\phi_1 - \phi_2)^2\right)\right)\right) \quad . \tag{23}$$

The integral to be minimized is then given by

$$I = \int \left(D_1(\nabla\phi_1)^2 + D_2(\nabla\phi_2)^2 - J((\phi_1 - \phi_2)^2)\right) d(\text{Area}) \quad . \tag{24}$$

Minimization of this expression I provides two Euler-Lagrange equations. In the case that the specific example for J of eq. (23) is used, these two equations, in their linear manifestation, are

$$\nabla^2(\phi_1 - \phi_2) + k^2(\phi_1 - \phi_2) = 0 \quad , \tag{25}$$

$$\nabla^2(D_1\phi_1 + D_2\phi_2) = 0 \quad , \tag{26}$$

where $k^2 = k_1^2 + k_2^2 = ((ab/D_1) + (ab/D_2)) = \text{constant}$.

Equations (25) and (26) are (except for change of notation) essentially the same as the linear version of eqs. (12) – (17), namely eqs. (18) and (19). The interpretation is quite different in the two cases however. Note that the pattern overlap decreases as the area grows if the nonlinear terms are retained. The Laplacian operator is the second derivative w/r/t the coordinates divided by the metric function g(in conformal coordinates), giving in an apparent way the connection of the pattern to the geometry of the epithelial surface.

Whether using the linearized version of eqs. (12) – (18), namely eqs. (18) and (19), with its implied connection to coupled signaling pathways, or the version of these given by this Appendix B, it is noted that all of the most frequent plant patterns for the leaves on a stem are obtained, including the

famous Fibonacci patterns, as well as the decussate and whorls [14]. This illustrates the flexibility of the patterning aspects of the model. A point of interest in the case of plant patterns is the very restricted appearance of the pattern with four leaves on the stem in a simple spiral, for which, as in the case of the invariant odd leg pairs of centipedes, no argument from natural selection exists. The suggestion is that plants may use somewhat similar patterning mechanisms, coupled signaling pathways, in spite of undoubtedly employing quite different bio-chemicals.

In this case of plant patterns, the linear version eqs. (18) and (19) have the solutions

$$\phi_1 = C + (k_1^2 D)\left[\cos\left(2\pi\left(\frac{px}{x_0} - \frac{qy}{y_0}\right)\right) + \cos\left(2\pi\left(\frac{qx}{x_0} \pm \frac{py}{y_0}\right)\right)\right] \ , \quad (27)$$

$$\phi_2 = C - (k_2^2 D)\left[\cos\left(2\pi\left(\frac{px}{x_0} - \frac{qy}{y_0}\right)\right) + \cos\left(2\pi\left(\frac{qx}{x_0} \pm \frac{py}{y_0}\right)\right)\right] \ . \quad (28)$$

Here p and q are integers which specify a particular pattern of interest, and C and D are positive constants, $k_2^2 D < C$. These equations satisfy periodic boundary conditions; ϕ_1 and ϕ_2 each have the same value at the points on the four corners of the rectangle ($x = 0$, $y = 0$); ($x = x_0$, $y = 0$); ($x = x_0$, $y = y_0$), and ($x = 0$, $y = y_0$). The convention adopted is that ϕ_1 takes on its maxima at these four corner points, where ϕ_2 takes on its minima. The rectangle represents one periodicity of the pattern on a cylindrical stem. The allowed positions of "florets" or leaves in a pattern is given by the simultaneous maxima of ϕ_1 and minima of ϕ_2, starting with the first leaves at the four corners.

The parameter k^2 satisfies the equation $k^2 A = (p^2 + q^2)(x_0/y_0 + y_0/x_0)$. Here $A = x_0 y_0$, the area of the rectangle. As the area A increases, the integers p or q will increase, often giving a new (p, q) pattern when the area is sufficient. However, if the rectangle is such that $y_0 \gg x_0$, then the expression for $k^2 A$ can be independent of p, q and increase only with y_0, while keeping the radius of the stem, $x_0/(2\pi)$ fixed; keeping the radius fixed, the leafs can then get far from each other along the stem while maintaining the same pattern, (p, q), as is often observed. The other limit, $x_0 \gg y_0$ on the other hand suggests flower arrangements, where the pattern can remain constant as the radius of the head increases. Daisies often have $N = 8$, 13 or 21 petals of different species.

The maxima of ϕ_1, and the minima of ϕ_2 are then obtained by setting

$$\left(\frac{px}{x_0} - \frac{qy}{y_0}\right) = i \ , \quad \text{and} \quad \left(\frac{qx}{x_0} \pm \frac{py}{y_0}\right) = j \ , \quad (29)$$

where i and j are integers running over values covering the rectangle, designating positions of leaves. The intersections of the two straight lines defined by eq. (29) gives the positions of the leaves. They lines have slopes p/q and $\pm q/p$.

When the integer pair (p, q) specifying a given pattern satisfies $p > q > 0$, and the pair is relatively primed, the number of leaves in a given pattern is

$$N = p^2 \pm q^2 \ , \quad (p > q > 0) \ . \tag{30}$$

Equation (30) is easily obtained by eliminating y/y_0 from eq. (29), giving the equation $(p^2 \pm q^2)(x/x_0) = (pi \pm qj)$; clearly the positive right hand side includes those i's and j's such that the integer $(piqj)$ runs from over the integers from zero to N, the last being the number corresponding to $x = x_0$.

Equation (30) holds in the case that there is only one floret on a single stem level (where $y =$ constant). This case encompasses the famous Fibonacci spirals, with numbers

$$N = 1^2 + 1^2 = 2 \ ; \ N = 2^2 - 1^2 = 3 \ ; \ N = 2^2 + 1^2 = 5 \ ;$$
$$N = 3^2 - 1^2 = 8 \ ; \ N = 3^2 + 2^2 = 13 \ ; \ N = 5^2 - 2^2 = 21 \ ; \ \ldots$$

Each pattern is generated in this case from the previous by adding only one leaf to each existing row of minimum slope. The pattern with $N = 4$ is conspicuously absent; it represents only a small percent of observations compared to the common $N = 2$, 3, or 5 patterns observed [36,44]. An adaptation argument for such an omission is not available.

Fig. 3 shows three examples, the cases when $N = 3$, 5 and 8. Each one shows two straight lines intersecting, and the intersections indicate the leaf positions in each case. In the $N = 3$ case (middle), the slopes of the two lines are given below.

One common simple pattern is the distichous pattern. This is the $N = 1^2 + 1^2$ pattern, which does not satisfy the $p > q$ criteria. Here, there is one leaf on a level, followed by a higher (or lower) leaf rotated by 180 degrees. Compound leaves often show this pattern. Another very commonly occurring non-Fibonacci pattern is the decussate pattern. This is the simplest alternating whorl pattern. This occurs when $p = q = 2$; higher order alternating whorls occur when $p = q$, and the number N of leaves in a pattern is given by $N = p + q = 2p$. The decussate pattern has two leaves on each level separated by 180 degrees, with the next higher pair rotated by 90 degrees relative to the first, for a total of $N = 4$.

Whenever there are J florets or leaves on a level, then the number N becomes, instead of eq. (30), $N = (p^2 \pm q^2)/J$. The bijugate, trijugate, ... spirals arise when $p > q$, and p and q are not relatively primed. Here there are two, three, ..., J leaves on a level. For example, the bijugate spiral case of $N = 10$ has $N = (4^2 + 2^2)/2 = 2(2^2 + 1^2)$; it is essentially the Fibonacci pattern of $N = 5$, except that there are two leaves on a level rotated with respect to each other by 180 degrees. In the trijugate case, the leaves would be rotated by 120 degrees on one level, and so on.

References

1. Becker, P.: Gene regulation: a finger on the mark. Nature (news and views), **442**, 31–32 (2006)
2. Beloussov, L.: Morphomechanics: goals, basic experiments and models, Int'l. J. Dev. Biol. **50**, 81–92 (2006)
3. Bernstein, B., Mikkelsen, T., Xie, X., Kamal, M, Huebert, D., Lander, E.: A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell **125(2)**, 315–326 (2006)
4. Bilder, D., Li, M., Perrimon, N.: Cooperative regulation of cell polarity and growth by Drosophila tumor suppressors. Science **289**, 113–116 (2000)
5. Boyer, L., Johnstone, S., Zucker, J., Kumar, R., Jenner, R., Melton, D., Young, R. A.: Core transcriptional regulatory circuitry in human embryonic stem cells. Cell, **122**, 1–10 (2005)
6. Brivanlou, A., Darnell, J.: Signal transduction and the control of gene expression. Science, **95**, 813–820 (2002)
7. Broun, M., Gee, L., Reinhardt, B., Bode, H.: Formation of the head organizer in hydra involves the canonical Wnt pathway. Development, **132**, 2907–2916 (2005)
8. Brouzes, E., Farge, E.: Interplay of mechanical deformation and patterned gene expression in developing embryos. Curr. Opinion in Genetics & Devel., **14**, 367–374 (2004)
9. Carroll, S., Grenier, J., Weatherbee, S.: From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design. Blackwell Sci. (2001)
10. Chourrout, D., Delsuc, F., Chourrout, P., Edvardsen, R., Rentzsch, F., Renfer, E., Jensen, M., Zhu, B., deJong, P., Steele, R., Technau, U.: Minimal ProtoHox cluster inferred from bilaterian and cnidarian cluster. Nature, **442**, 684–687 (2006)
11. Cote, G., Crain, R.: Biochemistry of phosphoinositides. Annu Rev. Plant Physiol., **44**, 333–356 (1993)
12. Cote, G., Crain, R.: Why do plants have phosphoinositides? BioEssays, **16**, 39–46 (1994)
13. Cox, R., Peifer, M.: Wingless signaling: The inconvenient complexities of life. Current Biology, **8**, R140–R144 (1998)
14. Cummings, F., Strickland, J.: A model of phyllotaxis. J. Theor. Biol., **192**, 531–544 (1998)
15. Cummings, F.: Waves of pattern formation and signal pathways. J. Theor. Biol., **196**, 27–31 (1999)
16. Cummings, F.: A model of morphogenesis. Physica A, **339**, 531–547 (2004)
17. Cummings, F.: Interaction of morphogens with geometry. Physica A, **355**, 427–438 (2005)
18. Cummings, F.: On the origin of pattern and form in early metazoans. Int. J. Dev. Biol., **50**, 193–208 (2006)
19. Cummings, F.: On the origin of metazoans. In: Mondaini, R.P., and Dilão, R. (eds) Proceedings of the International Symposium on Mathematical and Computational Biology, BIOMAT 2005. World Scientific (2006)
20. Davidson, G., Wu, W., Shen, J., Bilic, J., Fenger, U., Stannek, P., Glinka, A., Niehrs, C.: (2005) Casein kinase 1γ couples Wnt receptor activation to cytoplasmic signal transduction. Nature, **438**, 867–872 (2005)

21. Davidson, E.: Genomic Regulatory Systems. Acad. Press, San Diego, Ca. (2001)
22. Dilão, R. The reaction diffusion approach to morphogenesis. In: Mondaini, R.P., and Dilão, R. (eds) Proceedings of the International Symposium on Mathematical and Computational Biology, BIOMAT 2005. World Scientific (2006)
23. Di Paolo, G., De Camilli, P.: Phosphoinosotides in cell regulation and membrane dynamics. Nature (review), **443**, 651–657 (2006)
24. DoCarmo, M.: Differential Geometry of Curves and Surfaces. Prentice Hall, Englewood Cliffs, N.J. (1976)
25. Downward, J.: The ins and outs of signaling (news and views feature). Nature, **411**, 759–473 (2001)
26. Etienne-Manneville, S., Hall, A.: Rho GTPases in cell biology. Nature (review), **420**, 629–635 (2002)
27. Farge, E.: Current Biology, **13**, 1365–1377 (2003)
28. Fehon, R.: Polarity bites. Nature (news and views), **442**, 519-520 (2006)
29. Fraga, M. F., et. al.: Proc. Nat. Acad. Sci, USA, **102**, 10604–10609 (2005)
30. Gerhart, J., Kirschner, M.: Cells, Embryos and Evolution; Toward a Cellular and Developmental Understanding of Phenotypic Variation and Evolutionary Adaptability. Blackwell Science, Oxford (1997)
31. Gordon, R., Bjorklund, N.: How to observe surface contraction waves in axolotl embryos. Int. J. Dev. Biol., **40(4)**, 913–914 (1996)
32. Hall, A.: Rho GTPases and the actin cytoskeleton. Science, **279**, 509–514 (1998)
33. Harland, R.: A twist on embryonic signaling. Nature (news and views) **410**, 423–424 (2001)
34. Humbert, P., Russell, S., Richardson, H.: Dlg, Scribble and Lgl in cell proliferation and cancer. BioEssays, **25**, 542–553 (2003)
35. Ivanova, N., Dobrin, R., Lu, R., Kotenko, I., Levorse, J., DeCosta, C., Schafer, X. Lun, Y., Lemischka, I.: Dissecting self-renewal in stem cells with RNA interference. Nature, **442/3** 533–538 (2006)
36. Jean, R.: Phyllotaxis. Cambridge Univ. Press (1994)
37. Johnson, K., Wodarz, A.: A genetic hierarchy controlling cell polarity. Nature Cell Biol. **5**, 12–14 (2003)
38. Koch, A., Meinhardt, H.: Biological pattern formation: from basic mechanism to complex structures. Rev. Modern Phys., **66**, 1481–1507 (1994)
39. Li, H., Ilin, S., Wang, W., Duncan, E., Wysocka, J., Allis, C., Patel, D. J.: Molecular basis for site specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. Nature, **442**, 91–95 (2006)
40. Livingston, D.: Chromosome defects in the colon. Nature (news and views), **410**, 536–537 (2001)
41. Maruta, H., He, H., Tikoo, A., Vuong, T., Nur-E-Kamal, MSA.: G proteins, phosphoinositides, and actin-cytoskeleton in the control of cancer growth. Microscopy Research and Technique, **47**, 61–66 (1999)
42. Meinhardt, H.: Models of Biological Pattern Formation. Acad. Press, New York (1982)
43. Meinhardt, H.: Interview with Hans Meinhardt. Int. J. Dev. Biol., **50**, 1–4 (2006)
44. Mitchison, G.: Phyllotaxis and the Fibonacci series. Science, **196**, 270–275 (1972)

45. Murray, J.: Mathematical Biology. Springer-Verlag, Berlin, New York (1990)
46. Nusse, R.: Relays at the membrane. Nature (news and views), **438**, 747–748 (2005)
47. Nusslein-Volhard, C.: From cell to Organism; how genetics unlocked the mysteries of development. Kales Press (2006)
48. O'Keefe, L., Dougan, S., Gabay, L., Raz, E., Shilo, B-Z., DiNardo, S.: Spitz and Wingless, emanating from distinct borders, cooperate to establish cell fate across the Engrailed domain in the Drosophila epidermis. Development **124**, 4837–4845 (1997)
49. Pelengaris, S., Khan, M., Evan, G.: c-myc: more than just a matter of life and death. Nature Reviews Cancer, **2**, 764–776 (2002)
50. Pilot, F., Lecuit, T.: Compartmentalized Morphogenesis in Epithelia: From Cell to Tissue Shape. Dev. Dynamics, **232**, 685–694 (2005)
51. Pilot, F., Philippe, J.-M., Lemmers, C., Lecuit, T.: Spatial control of actin organization at adherens junctions by the synaptotagmin-like protein Btsz. Nature, **442**, 580–584 (2006)
52. Riese, J., Yu, X., Munnerlyn, A., Eresh, S., Hsu, S-C., Grosschedl, R., Bienz, M.: LEF1, a nuclear factor coordinating signalling inputs from wingless and decapentaplegic. Cell **88**, 777–787 (1997)
53. Rodda, D., Chew, J-L., Lim, L-H., LOh, Y-H., Wang, B., Ng, H-H., Robson, P.: Transcriptional regulation of NANOG, by OCT4 and SOX2. J. Biol. Chem. **280**, 24731–24737 (2005)
54. Scadden, D.: The stem cell niche as an entity of action. Nature (Insight review), **441**, 1075–109 (2006)
55. Scott, I., Didier, Y., Stainier, R.: Twisting the body into shape. Nature (news and views), **425**, 461–466 (2003)
56. Segal, E., Fondufe-Mittendorf, Y., Chen, L., Thaström, A., Field, Y, Moore, I, Wang, J-P., Widom, J.: A genomic code for nucleosome positioning. Nature, **442**, 772–778 (2006)
57. Taipale, J., Beachy, P.: The Hedgehog and Wnt signalling pathways in cancer. Nature, **411**, 349–353 (2001)
58. Tapon, N., Hall, A.: Rho, Rac and Cdc42 GTPases regulate the organization of the actin cytoskeleton. Current Opinion in Cell Biology **9**, 86–92 (1997)
59. Tsikolia, N.: The role and limits of a gradient based explanation of morphogenesis: a theoretical consideration. Int'l. J. Dev. Biol. **50**, 333–340 (2006)
60. Turing, A. The chemical basis of morphogenesis. Phil. Trans. R. Soc. London B, 37–72 (1952)
61. Wallace, A.: The emerging conceptual framework of evolutionary developmental biology. Nature, **415**, 757–764 (2002)
62. Wikramanayake, M., Hong, M., Lee, P., Pang, K., Byrum, C., Bince, J., Xu, R., Martindale, M.: An ancient role for nuclear β-catenin in the evolution of axial polarity and germ layer segregation. Nature, **426**, 446–449 (2003)
63. Wysocka, J., Swigut, T., Xiao, H., Milne, T., Kwon, S., Landry, J., Kauer, M., Tackett, A., Chait, B., Badenhorst, P., Wu, C., Allis, C.D.: A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature, **442**, 86–89 (2006)
64. Yamanaka, S.: as reported by Check, E., Nature 'news', **442**, 11 (2006)

Mathematical Modeling of HIV-1 Infection and Drug Therapy

Libin Rong[1], Zhilan Feng[1], and Alan S. Perelson[2]

[1] Department of Mathematics, Purdue University, West Lafayette, IN 47907, USA
rong@math.purdue.edu, zfeng@math.purdue.edu
[2] Theoretical Biology and Biophysics, Theoretical Division, Los Alamos National Laboratory, MS-K710, Los Alamos, NM 87545, USA asp@lanl.gov

Summary. Mathematical models have made significant contributions to our understanding of HIV-1 dynamics. Many important features of the interaction between virus particles and cells have been determined by fitting mathematical models to experimental data. In this chapter, we begin with a brief review of some basic models used to study viral infection and estimate parameters that govern viral production and viral clearance. We then discuss recent developments in the modeling of HIV-1 dynamics and antiretroviral response. We focus on the impact of various classes of antiretroviral drugs that target different stages of the viral replication cycle. We also discuss how drug treatment affects the emergence of drug resistance during treatment, and how a low level of virus and latently infected cells can persist in infected individuals for a prolonged period of time despite an apparently effective antiretroviral therapy.

Key words: Human immunodeficiency, antiretroviral therapy, viral equilibrium, viral load decays, drug resistance, virus dynamics.

1 Introduction

Since the discovery of the human immunodeficiency virus subtype 1 (HIV-1) in the early 1980s, the disease has spread in successive waves to most regions of the world. HIV has infected more than 60 million people and over a third of them died subsequently [22]. Considerable scientific effort has been devoted to the understanding of viral pathogenesis, host/virus interactions, immune response to infection and antiretroviral (ARV) therapy.

Important insights into the host-pathogen interaction in HIV-1 infection have been derived from the development of a quantitative assay for virus particles in plasma [79] and the use of mathematical models to interpret experimental results [40, 73, 78, 102]. Consistent detection of a relatively stable level of viral RNA in plasma over a prolonged period of chronic infection indicates

that viral production and clearance are in equilibrium. However, by simply measuring the equilibrium level of virus one can not tell whether the virus is produced quickly or slowly. In 1995, Ho, Perelson and colleagues [40] and Shaw, Nowak and colleagues [102] realized that perturbations of this quasi-steady state might provide insights into the dynamics of viral production and clearance *in vivo*. Following the initiation of treatment with potent inhibitors of HIV-1 protease or reverse transcriptase, there was a rapid decline in viral load [40,102]. Perelson et al. [78] used a simple mathematical model to analyze a set of viral load data collected from infected patients after the administration of ritonavir, a protease inhibitor, and estimated virion clearance rate, infected cell life-span, and viral generation time. Model results demonstrated that the rapid decline in viral load after therapy represented a composite of the turnover of free virus and productively infected cells. The clearance of free virus from plasma occurred very quickly ($t_{1/2} \leq 0.24 \pm 0.06$ day), and the half-life of the cells that produce almost all of the plasma virus was also very short ($t_{1/2} \leq 1.55 \pm 0.57$ day) [78].

The short half-life of free virus implies that half of the entire plasma virus population is replaced every 6 hours or less. Consequently, a large amount of virus is produced per day. It was estimated that more than 10^{10} virions were produced daily in the average mid-stage HIV-1-infected untreated patient [78].

When HIV-1 replicates, its RNA genome must be reverse transcribed into DNA. However, this copying process is highly error-prone, leading to the realization that viral genomes with every possible single point mutation could arise daily [13]. High replication and high mutability of HIV-1 have greatly advanced understanding the evolution of drug resistance mutations in the presence of ARV therapy. When drugs are administered individually, treatment often fails to achieve successful viral suppression (virological failure) because of the rapid emergence of drug resistant virus variants.

The advent of potent combination antiretroviral therapy has contributed to a substantial reduction in the incidence of HIV-1-related morbidity and mortality. Antiretroviral therapy using a combination of three or more drugs from two or more classes (for example, two nucleoside reverse transcriptase inhibitors (NRTI) combined with either a protease inhibitor (PI) or an non-nucleoside reverse transcriptase inhibitor (NNRTI)) has proved extremely effective in suppressing the plasma viral load of most HIV-1-infected patients below the limit of viral detection (50 RNA copies ml^{-1}) by standard assay to date [14]. Since viral replication is directly linked to CD4^+ depletion and disease progression [56], the viral decline in the presence of combination therapy has profound clinical significance.

Quantitative analysis of viral decay following the initiation of highly active antiretroviral therapy (HAART) has suggested that the plasma viral load declines in at least three distinct phases (see reviews in [4, 28, 45, 96]). The first-phase decline, in which the plasma level of viral RNA was observed to drop by one to two orders during the first 2 weeks of treatment, reflects rapid viral clearance and turnover of productively infected T lymphocytes

with a half-life of less than one day [78]. This is followed by a slower second-phase viral decline with a half-life of 1-4 weeks [73]. The nature of the cellular reservoir that is responsible for the second-phase viral decay remains obscure and several cell populations might contribute. One possibility is the existence of a long-lived population of productively infected cells, such as infected macrophages, which are less susceptible to the viral cytopathic effects than are $CD4^+$ lymphoblasts [41]. Other sources that contribute to plasma virus include infected $CD4^+$ T cells in a latent state [107], which can be activated to produce virus when encountering specific antigens, and release of the virus trapped in tissue reservoirs, for example, on the surface of follicular dendritic cells (FDCs) [36, 37]. Fitting viral load and infected cell declines suggested that the loss of long-lived infected cells was a major contributor to the second phase, whereas the activation of latently infected lymphocytes was only a minor source [73].

After several months of HAART, many treated patients attain levels of plasma HIV-1 RNA that is below 50 copies/ml, the limit that can be detected by current standard assays. However, this does not imply that the treatment has completely suppressed viral replication. In fact, even in patients with suppression of virus below detection for many years, a low level of virus may still persist in plasma, which can be detected by supersensitive assays [17, 20, 71]. Generally, this phase with HIV-1 RNA below 50 copies/ml corresponds to the third phase of viral decay. The viral persistence despite prolonged treatment is possibly due to the slow intrinsic decay characteristics of a long-lived latent reservoir of HIV-1, such as resting memory $CD4^+$ T lymphocytes [9], or the inability of current drug regimens to completely suppress viral replication.

In this chapter, we first briefly review several mathematical models, formulated as ordinary differential equations or delay differential equations, that have been used to study HIV-1 dynamics. These studies demonstrate how simple models, combined with parameter estimation techniques, can help understand important features of HIV-1 infection and replication. We then present two age-structured models that can be used to study the influence of antiretroviral drugs on the evolution of HIV-1. These models incorporate the age of infection as well as combination therapies involving reverse transcriptase (RT), protease, and entry/fusion inhibitors. We focus on the impact of various classes of inhibitors that target different stages of the viral replication cycle. An important advantage of using the age-structured models is that they provide a more accurate description of the functional dependence of the reproductive ratio on the efficacy of RT inhibitors than that obtained previously from non-age-structured models. This may have significant implications in predicting the effects of drug therapy. We also employ a two-strain model to study the mechanism underlying the emergence of drug resistant variants during treatment. Analytical results show that drug resistance is more likely to arise for intermediate levels of treatment effectiveness, at which the reproductive ratios of both the sensitive and drug resistant strains are close. Antiviral

response is further investigated when we incorporate realistic pharmacokinetics and adherence behavior. Simulations suggest that perfect adherence to regimen protocol will well suppress the viral load of the sensitive strain while drug resistant variants develop slowly. However, intermediate levels of adherence may result in the dominance of the drug resistant virus. When the level of adherence is low, the failure of suppression of the wild-type virus will be observed, accompanied by a relatively slow increase of the drug resistant viral load. Finally, we show how a mathematical model can be used to explain the persistence of a low level of virus and latently infected cells in infected individuals on effective antiretroviral treatment over a prolonged period of time.

2 Basic Model of Virus Infection

After the first few months of HIV-1 infection, plasma virus attains a set-point or constant level for years [38]. In order to maintain a relatively stable viral load at a constant level, viral production and clearance in infected individuals must be in equilibrium. Although quantitative assays like RT-PCR can quantitate HIV RNA in plasma throughout the course of infection, we still can not discriminate the scenarios in which virus is produced quickly or slowly. The first studies that provided information on rates of viral production and clearance were carried out to perturb the viral equilibrium by using the protease inhibitors ritonavir [40] or saquinavir [102], as well as the RT inhibitor nevirapine [102]. If viral replication is completely suppressed by inhibitors, then the viral decay reveals the clearance rate of free virus. If it is not completely suppressed, then the rate of viral decay will depend on not only the clearance rate, but also the death rate of productively infected cells, as well as the effectiveness of the drug treatment.

A basic model that has been used to study HIV-1 dynamics can be described by the following equations [78]:

$$\begin{aligned} \frac{d}{dt}T(t) &= \lambda - dT - kVT, \\ \frac{d}{dt}T^*(t) &= kVT - \delta T^*, \\ \frac{d}{dt}V(t) &= N\delta T^* - cV, \end{aligned} \tag{1}$$

where $T(t), T^*(t)$ and $V(t)$ denote the concentrations of uninfected $CD4^+$ T cells, productively infected cells, and free virus at time t, respectively. λ represents the recruitment rate of uninfected T cells, d is the per capita death rate of uninfected cells, k is the rate constant at which uninfected cells are infected by free virus. Here the infection is modeled by a simple but commonly

used method, i.e., a "mass action" term kVT. δ is the per capita death rate of infected cells, N (burst size) is the total number of virus particles produced by a productively infected cell during its lifetime and c is the clearance rate of virus. Therefore, $N\delta$, which is N divided by the cell life-span, $1/\delta$, gives the per capita viral production rate.

Two classes of ARV drugs are often used to perturb the system. One class is reverse transcriptase inhibitors (RTIs), which can effectively block the infection of target T cells by free virus; the other is protease inhibitors (PIs), which prevent HIV-1 protease from cleaving the HIV polyprotein into functional units, causing infected cells to produce immature virus particles that are non-infectious. In the presence of these two inhibitors, the model equations (1) are modified to become:

$$\begin{aligned}
\frac{d}{dt}T(t) &= \lambda - dT - (1-\epsilon_{RT})kV_I T,\\
\frac{d}{dt}T^*(t) &= (1-\epsilon_{RT})kV_I T - \delta T^*,\\
\frac{d}{dt}V_I(t) &= (1-\epsilon_{PI})N\delta T^* - cV_I,\\
\frac{d}{dt}V_{NI}(t) &= \epsilon_{PI}N\delta T^* - cV_{NI},
\end{aligned} \tag{2}$$

where ϵ_{RT} and ϵ_{PI} ($0 \le \epsilon_{RT},\ \epsilon_{PI} \le 1$) are the efficacies of RT and PI, V_I and V_{NI} are the concentrations of infectious and non-infectious virus, respectively. $V = V_I + V_{NI}$ is the total amount of virus.

If we assume that a 100% effective PI ($\epsilon_{PI} = 1$) is given to an infected individual at quasi steady state with initial viral load, V_0, and that the uninfected target cells remain approximately at a constant level, T_0, over the time period of interest, then the viral load at time t can be solved from model (2) [78]:

$$V(t) = V_0 e^{-ct} + \frac{cV_0}{c-\delta}\left[\frac{c}{c-\delta}(e^{-\delta t} - e^{-ct}) - \delta t e^{-ct}\right]. \tag{3}$$

Using nonlinear regression analysis, estimates of the parameters, c and δ, were obtained by fitting equation (3) to the plasma HIV-1 RNA data [78]. These estimates give upper bounds for the half-life of the plasma virus ($t_{1/2} = \ln 2/c$) and the half-life of productively infected cells ($t_{1/2} = \ln 2/\delta$) because therapy in reality is not 100% effective and additional viral clearance and/or loss of productively infected cells is required to account for the residual viral replication.

In clinic, when multiple ARV drugs are simultaneously given to HIV-1 patients, plasma viral load decays with an initial rapid exponential decline of nearly 2 logs, followed by a slower exponential decline. To explain the second-phase viral decline, a new model was introduced [73], which postulated the

existence of other sources that could also contribute to plasma virus, such as long-lived productively infected cells and activation of latently infected cells. The model was described by the following equations:

$$\begin{aligned}
\frac{d}{dt}T^*(t) &= kVT + aL - \delta T^*,\\
\frac{d}{dt}L(t) &= fkVT - \mu_L L,\\
\frac{d}{dt}M^*(t) &= k_M VM - \mu_M M^*,\\
\frac{d}{dt}V(t) &= N\delta T^* + pM^* - cV,
\end{aligned} \tag{4}$$

where cells, M, which upon infection with a rate, k_M, become long-lived infected cells, M^*, which produce virus at a rate p and are lost with a rate μ_M. The model also contains latently infected cells, L, which are generated with a rate fk, smaller by a factor f, die with a rate δ_L, and are activated into productively infected cells with a rate a, giving a total rate constant of loss $\mu_L = a + \delta_L$.

Assuming that both RT inhibitor and protease inhibitor are 100% effective, the viral level after drug therapy can be solved similarly as in (3) [73]:

$$V(t) = V_0\left[Ae^{-\delta t} + Be^{-\mu_L t} + Ce^{-\mu_M t} + (1 - A - B - C)e^{-ct}\right],$$

where

$$A = \frac{NkT_0}{c-\delta}(1 - \frac{af}{\delta - \mu_L}), B = \frac{af\delta NkT_0}{\mu_L(\delta-\mu_L)(c-\mu_L)}, C = \frac{c - NkT_0(1+\frac{af}{\mu_L})}{c-\mu_M},$$

and the level of infected cells in blood is given by

$$I(t) = T^*(t) + L(t) = \frac{kV_0T_0}{\delta}\left[(1 - \frac{af}{\delta-\mu_L})e^{-\delta t} + \frac{f\delta}{\mu_L}(1 + \frac{a}{\delta-\mu_L})e^{-\mu_L t}\right].$$

Simultaneously fitting $V(t)$ and $I(t)$ to the plasma virus and peripheral blood mononuclear cells (PBMC) infectivity data shows that the loss of long-lived infected cells ($t_{1/2}$ of 1-4 weeks) is a major contributor to the second phase of viral decay, whereas the activation of latently infected cells is only a minor source [73].

3 Delay Models

Upon the onset of drug therapy, viral load remains at the pretreatment level for a short period of time (hours to a few days), called the shoulder period,

before the first rapid exponential decline [78]. Two factors might contribute to this shoulder phase: one is the intracellular delay, which arises from the time required for infected cells to produce progeny virus after infection; the other is the pharmacological delay, which is the time interval from the administration of the drug to the beginning of drug action within cells [19, 78]. These delays need to be considered to accurately estimate parameters that govern the kinetics of viral infection and cell death *in vivo* [10, 19, 35, 59, 63–65].

3.1 Intracellular Delay

To characterize the time between the initial viral entry into a target cell and subsequent viral production, an intracellular delay was first introduced by Herz et al. [35] to analyze the clinical data. The modified basic model incorporating a fixed intracellular delay, τ, is

$$\begin{aligned}
\frac{d}{dt}T(t) &= \lambda - dT - k(t)VT,\\
\frac{d}{dt}T^*(t) &= k(t-\tau)V(t-\tau)T(t-\tau)e^{-m\tau} - \delta T^*,\\
\frac{d}{dt}V(t) &= N\delta T^* - cV,
\end{aligned} \tag{5}$$

where m is the death rate of cells between initial infection and the beginning of viral production τ time units later. If the death is assumed to be a first order process, then the probability a cell infected at time t is alive at time $t+\tau$ is $e^{-m\tau}$.

If we consider monotherapy with an RT inhibitor that has 100% effectiveness and the system is assumed to be in pretreatment steady state with

$$T(0) = T_0 = \frac{c}{kN}e^{m\tau}, \quad T^*(0) = T_0^* = \frac{\lambda}{\delta}e^{-m\tau} - \frac{dc}{kN\delta}, \quad V(0) = V_0 = \frac{N\delta T_0^*}{c},$$

where k is the constant infection rate, then model (5) can be solved for $0 \le t < \tau$ to obtain

$$T(t) = \frac{\lambda}{d} + (T_0 - \frac{\lambda}{d})e^{-dt}, \quad T^*(t) = T_0^*, \quad V(t) = V_0,$$

and for $t \ge \tau$, $T(t)$ remains the same as above and

$$T^*(t) = T_0^* e^{-\delta(t-\tau)}, \quad V(t) = \frac{V_0}{c-\delta}\left[ce^{-\delta(t-\tau)} - \delta e^{-c(t-\tau)}\right].$$

Considering monotherapy with a protease inhibitor, model (5) changes to

$$\begin{aligned}
\frac{d}{dt}T^*(t) &= kT_0V_I(t-\tau)e^{-m\tau} - \delta T^*,\\
\frac{d}{dt}V_I(t) &= (1-\epsilon_{PI})N\delta T^* - cV_I,\\
\frac{d}{dt}V_{NI}(t) &= \epsilon_{PI}N\delta T^* - cV_{NI}.
\end{aligned} \tag{6}$$

If the protease inhibitor is assumed to be 100% effective, then the viral load can be solved for $0 \le t < \tau$ to get $V(t) = V_0$, and for $t \ge \tau$

$$V(t) = V_0e^{-c(t-\tau)} + \frac{cV_0}{c-\delta}\left[\frac{c}{c-\delta}[e^{-\delta(t-\tau)} - e^{-c(t-\tau)}] - \delta(t-\tau)e^{-c(t-\tau)}\right],$$

where $V(t)$ is the total amount of infectious and non-infectious virus.

In [64], Nelson and Perelson analyzed model (6) allowing for imperfect drug treatment. They provided an analytical expression, $\lambda_d \sim -\delta\epsilon_{PI}C(\epsilon_{PI}, \delta, \tau)$, where λ_d is the dominant eigenvalue that determines the rate of viral decay and $C(\epsilon_{PI}, \delta, \tau) = 1/(1 + (1-\epsilon_{PI})\delta\tau)$. This result explains why there was no change in the estimate of δ with delay models in [35]. The delay effect canceled out due to the assumption of 100% effectiveness ($\epsilon_{PI} = 1$). Combining the intracellular delay with less than perfect antiretroviral treatment resulted in a significant increase in the estimated value for δ compared with the case of perfect drug therapy when model (6) was fitted to clinical data [64].

Assuming a fixed intracellular delay for every infected cell, however, may not be biologically reasonable. Conversion of a newly infected cell into a productively infected one is a multi-step process, and thus we cannot expect that all infected cells finish all these processes in the same time. Mittler et al. [60] used a gamma distribution to describe a continuous time delay between viral infection and production. The model was

$$\begin{aligned}
\frac{d}{dt}T^*(t) &= \int_0^\infty kT_0f(t')V_I(t-t')e^{-mt'}dt' - \delta T^*,\\
\frac{d}{dt}V_I(t) &= [1-h(t)](1-\epsilon_{PI})pT^* - cV_I,\\
\frac{d}{dt}V_{NI}(t) &= h(t)\epsilon_{PI}pT^* - cV_{NI},
\end{aligned} \tag{7}$$

where t' is the delay variable, p is viral production rate. $h(t)$ is a Heaviside function employed to account for the pharmacological delay of a protease inhibitor, which will be further discussed in Section 3.2. The delay kernel $f(t')$ was assumed to be a gamma distribution,

$$f(t') = \frac{t'^{n-1}}{(n-1)!b^n}e^{-t'/b},$$

with mean, $\tau = nb$, variance, nb^2, and peak, $(n-1)b$.

For a completely effective protease inhibitor, an analytical solution for the viral load can be obtained through the repeated application of standard techniques [60],

$$V(t) = V_I(0)\Bigg\{e^{-ct} + \left[G_n ct - \frac{c}{\delta - c}H_n + \frac{c\hat{b}}{1-c\hat{b}}\sum_{\hat{k}=0}^{n-1}\frac{1 - G_{n-\hat{k}} + H_{n-\hat{k}}}{(1-c\hat{b})^{\hat{k}}}\right]e^{-ct}$$
$$+ \frac{c}{\delta - c}H_n e^{-\delta t} - \frac{c\hat{b}}{1-c\hat{b}}\sum_{\hat{k}=0}^{n-1}\frac{1 - G_{n-\hat{k}} + H_{n-\hat{k}}}{(1-c\hat{b})^{\hat{k}}}\sum_{l=0}^{\hat{k}}\frac{(1-c\hat{b})^l t^l}{l!\tau^l}e^{-t/\hat{b}}\Bigg\},$$

where

$$\hat{b} = \frac{b}{1+mb}, \quad G_i = \frac{\delta}{\delta - c}(1-c\hat{b})^{-i}, \quad H_i = \frac{\delta}{\delta - c}(1-\delta\hat{b})^{-i}.$$

By data fitting, they found that the previously estimated values for c were underestimates. However, they did not find any changes in the estimate of δ due to the assumption of a perfect drug [65].

3.2 Drug Pharmacokinetics

Despite the progress in the estimate of parameters with the incorporation of intracellular delay, these models do not include drug pharmacokinetics or recognize the drug dependence of intracellular delay. Drug concentrations continuously vary due to drug absorption, distribution, and metabolism in the body, resulting in a continuously time-varying drug efficacy. Recently, a model that combines drug pharmacokinetics and intracellular delay was developed by Dixit and Perelson [19] to study HIV-1 dynamics under therapy.

For RT inhibitor monotherapy, the model is given as follows:

$$\begin{aligned}
\frac{d}{dt}T(t) &= \lambda - dT(t) - (1-\epsilon_{RT}(t))kV(t-\tau_1)T(t-\tau_1)e^{-m\tau_1},\\
\frac{d}{dt}T^*(t) &= (1-\epsilon_{RT}(t-\tau_2))kV(t-\tau)T(t-\tau)e^{-m\tau} - \delta T^*, \qquad (8)\\
\frac{d}{dt}V(t) &= N\delta T^* - cV,
\end{aligned}$$

where τ_1 and τ_2 were defined to be the durations of the phases before and after the stage in the replication cycle affected by drug action. $\tau = \tau_1 + \tau_2$ is the viral replication time.

For protease inhibitors, the corresponding model is

$$\begin{aligned}
\frac{d}{dt}T(t) &= \lambda - dT(t) - kV_I(t)T(t),\\
\frac{d}{dt}T^*(t) &= kV_I(t-\tau)T(t-\tau)e^{-m\tau} - \delta T^*(t),\\
\frac{d}{dt}V_I(t) &= (1-\epsilon_{PI}(t-\tau_2))N\delta T^*(t) - cV_I(t),\\
\frac{d}{dt}V_{NI}(t) &= \epsilon_{PI}(t-\tau_2)N\delta T^*(t) - cV_{NI}(t).
\end{aligned} \tag{9}$$

A two compartment pharmacokinetics model including blood and cells was used to determine time-varying drug efficacies of RT inhibitor and protease inhibitor, $\epsilon_{RT}(t)$ and $\epsilon_{PI}(t)$, which will be discussed in more details in Section 5.3.

Model calculations showed that viral load decay in HIV-infected patients under monotherapy could display remarkably complex patterns depending on the relative magnitudes of the pharmacokinetics, intracellular, and intrinsic viral dynamic time-scales [19]. The commonly assumed exponential decay after therapy is only a special case.

Model (9) was also applied to analyze the plasma virus data from 5 patients under ritonavir (PI) monotherapy and obtain estimates of the intracellular delay, τ, and the *in vivo* efficacy of ritonavir [18]. Assuming $\delta = 1$ day^{-1}, they found $\tau \sim 1$ day, which is in agreement with previous estimates of the viral replication time. The average *in vivo* efficacy of ritonavir was estimated to be 0.65, which is significantly lower than the value of ~ 0.9 obtained from *in vitro* studies.

4 Age-structured Models

Although delay models have a distinct advantage over non-delay models in their ability to correctly account for the influence of intracellular delay and pharmacological delay on drug action, they employ several approximations. First, they assume that the viral production rate is zero in a newly infected cells and increases instantaneously to $N\delta$ after a time delay τ. Second, the death rate of infected cells, δ, is assumed to be constant. Recently, age-structured models have received increasing interest due to the incorporation of age-of-infection, and thereby their greater flexibility in modeling viral production and mortality of infected cells [31,62]. Nelson et al. [62] considered an age-structured model that allowed both the viral production rate, $p(a)$, and the death rate of infected T cells, $\delta(a)$, to be age-dependent. For a specific form of functions $p(a)$ and $\delta(a)$, they performed a local stability analysis of the nontrivial equilibrium point. Numerical simulations illustrated that the time to reach the peak viral level depended not only on the initial conditions

but also on the speed at which viral production achieves its maximum value. Based on this age-structured model, Gilchrist et al. [31] used the various life history trade-offs between viral production and clearance of infected cells to derive the within-host relative viral fitness.

Various classes of antiretroviral drugs are used to treat HIV-infection and they target different stages of the viral life cycle. Age-structured models can be employed to study the impact of these drugs on HIV-1 dynamics. Kirschner and Webb [46] proposed a model that incorporated age structure of infected cells to account for the mechanism of AZT (zidovudine) treatment. Feng and Rong [24] formulated an age-structured model for HIV-1 infection incorporating the combination of RT inhibitors and protease inhibitors to study the antiretroviral responses. They found in the age-structured model a different functional dependence of the reproductive ratio $\mathcal{R}$ on ϵ_{RT} than that found previously in non-age-structured models, which has significant implications in predicting the effect of drug treatment [88].

Before we present age-structured models incorporating drug effects, we give a brief review of the viral replication cycle of HIV-1. HIV infection begins by the attachment of a virus to a $CD4^+$ T cell. If an entry inhibitor is used, HIV is unable to bind to the surface of the T cell and gain entry into the cell. Inside the cell the HIV-1 enzyme reverse transcriptase (RT) makes a DNA copy of the virus' RNA genome. During this process, if an RT inhibitor is present then the viral genome will not be copied into DNA and therefore the host cell will not produce new virus. When the virus replicates, its DNA is read out to produce viral proteins. A large polyprotein is made, and a viral protease is needed to cut the long polypeptide chain into individual components that are needed to produce infectious virus particles. If the HIV-1 protease is inhibited, the newly produced virus will be noninfectious.

From the above description of the HIV life cycle and the roles of various inhibitors it is clear that the infection age of an infected cell can be important for the study of HIV dynamics under the influence of ARV treatment. In this section, we present two age-structured models, which extend the existing age-structured models [31, 46, 62] by incorporating combination therapy. The first model includes therapy with a combination of an RT inhibitor and a protease inhibitor, while the second model includes an entry/fusion inhibitor and a protease inhibitor. Stability result is obtained for a general form of both the viral production rate and the mortality rate of infected cells. The stability of the infection-free or the infected steady state depends on the reproductive ratio $\mathcal{R}$ being smaller or greater than 1. The formulation of this reproductive ratio also provides an appropriate measure for the within-host viral fitness, which can be used to explore the optimal viral production rate for which $\mathcal{R}$ is maximized [24].

4.1 The Model with RT and Protease Inhibitors

We begin with the age-structured model of HIV-1 infection (without drug treatment) given in [62]:

$$\begin{aligned}
&\frac{d}{dt}T(t) = \lambda - dT - kVT,\\
&\frac{\partial}{\partial t}T^*(a,t) + \frac{\partial}{\partial a}T^*(a,t) = -\delta(a)T^*(a,t),\\
&\frac{d}{dt}V(t) = \int_0^\infty p(a)T^*(a,t)da - cV,\\
&T^*(0,t) = kVT,
\end{aligned} \tag{10}$$

where $T^*(a,t)$ denotes the concentration of infected cells with infection age a (i.e. the time that has elapsed since an HIV virion has penetrated the cell) at time t. $\delta(a)$ is the age-dependent per capita death rate of infected cells and $p(a)$ is the viral production rate of an infected cell with age a.

The functional forms of $p(a)$ and $\delta(a)$ need to be determined experimentally [39, 62]. In [62], the authors chose the following function for the production rate

$$p(a) = \begin{cases} p^*\left(1 - e^{-\theta(a-a_1)}\right) & \text{if } a \geq a_1,\\ 0 & \text{else,} \end{cases} \tag{11}$$

where θ determines how quickly $p(a)$ reaches the saturation level p^*, and a_1 is the age at which reverse transcription is completed.

Model Formulation

To account for the effect of RT inhibitors we divide the class of infected cells, $T^*(a,t)$, into two subclasses: $T^*_{preRT}(a,t)$ and $T^*_{postRT}(a,t)$. $T^*_{preRT}(a,t)$ represents the density of cells that have been "infected" by an HIV virus but in which reverse transcription has not been completed at infection age a. An RT inhibitor could allow a preRT cell to revert back to an uninfected cell (because if reverse transcription fails to complete, cellular nucleases will degrade the HIV RNA that entered the cell) or reduce the probability that a preRT cell progresses to the postRT state [21]. $T^*_{postRT}(a,t)$ represents the density of infected cells that have progressed to the postRT phase at infection age a. The densities of the preRT and postRT cells are related by a function $\beta(a)$ $(0 \leq \beta(a) \leq 1)$ that describes the proportion of infected cells that have not completed reverse transcription, i.e.,

$$T^*_{preRT}(a,t) = \beta(a)T^*(a,t), \qquad T^*_{postRT}(a,t) = (1-\beta(a))T^*(a,t). \tag{12}$$

We assume that $\beta(a) \in L^1[0,\infty)$ is a non-increasing function with the following properties:

$$0 \leq \beta(a) \leq 1;\ \beta(0) = 1;\ \beta(a) = 0 \text{ for } a \geq a_1;\ \beta'(a) \leq 0 \ \ a.e.$$

where a_1 has the same meaning as mentioned above.

With the effect of protease inhibitor, new infectious virus particles are produced at the rate

$$\int_0^\infty (1 - \epsilon_{PI}) p(a) T^*_{postRT}(a,t) da.$$

Let $\eta(\epsilon_{RT})$ denote the rate at which preRT cells revert to the uninfected state due to the failure of reverse transcription. The rate at which preRT cells of all ages become uninfected is then given by

$$\int_0^\infty \eta(\epsilon_{RT}) T^*_{preRT}(a,t) da.$$

The reversion rate $\eta(\epsilon_{RT})$ is an increasing function of drug efficacy ϵ_{RT}. In the absence of drug therapy, we assume there are no infected cells going back to the uninfected class, i.e., $\eta(0) = 0$. As the limit case, when the RT inhibitor is 100% effective, $\eta(\epsilon_{RT})$ should be very large. We shall discuss the functional form of $\eta(\epsilon_{RT})$ more when we compare the treatment effectiveness of different drug combinations in Section 4.3. All analytical results are obtained for a general reversion rate function.

Incorporating these drugs in model (10), we have:

$$\begin{aligned}
&\frac{d}{dt}T(t) = \lambda - dT - kV_I T + \int_0^\infty \eta(\epsilon_{RT}) T^*_{preRT}(a,t) da, \\
&\frac{\partial}{\partial t}T^*(a,t) + \frac{\partial}{\partial a}T^*(a,t) = -\delta(a) T^*(a,t) - \eta(\epsilon_{RT}) T^*_{preRT}(a,t) da, \\
&\frac{d}{dt}V_I(t) = \int_0^\infty (1 - \epsilon_{PI}) p(a) T^*_{postRT}(a,t) da - cV_I, \\
&\frac{d}{dt}V_{NI}(t) = \int_0^\infty \epsilon_{PI} p(a) T^*_{postRT}(a,t) da - cV_{NI}, \\
&T^*(0,t) = kV_I T.
\end{aligned} \tag{13}$$

Notice that the variable V_{NI} does not appear in equations for other variables. Thus, we can ignore the V_{NI} equation when studying the dynamics of infection. Using the relation (12) we have the following system:

$$\begin{aligned}
&\frac{d}{dt}T(t) = \lambda - dT - kV_I T + \int_0^\infty \eta(\epsilon_{RT})\beta(a)T^*(a,t)da,\\
&\frac{\partial}{\partial t}T^*(a,t) + \frac{\partial}{\partial a}T^*(a,t) = -\delta(a)T^*(a,t) - \eta(\epsilon_{RT})\beta(a)T^*(a,t),\\
&\frac{d}{dt}V_I(t) = \int_0^\infty (1-\epsilon_{PI})(1-\beta(a))p(a)T^*(a,t)da - cV_I,\\
&T^*(0,t) = kV_I T.
\end{aligned} \tag{14}$$

Because we are interested in the effect of combination therapy on virus dynamics, we assume that HIV-1 patients are initially at steady state and the combination of drugs is administered at time 0. We choose the initial conditions to be $T(0) = T_0$, $V_I(0) = V_{I0}$, $V_{NI}(0) = 0$ and $T^*(a,0) = T_0^*(a)$, where T_0 and V_{I0} are the steady state levels of target cells and infectious virus, respectively. $T_0^*(a)$ is the age distribution of infected cells at the initial time $t = 0$ and $\int_0^\infty T_0^*(a)da$ represents the steady state level of infected cells before the onset of drug therapy.

Following the analysis given in [88], we can show that in order to study the steady states of model (14), it suffices to study the following system:

$$\begin{aligned}
&\frac{d}{dt}T(t) = \lambda - dT(t) - kV_I(t)T(t) + \int_0^\infty kK_1(a)V_I(t-a)T(t-a)da,\\
&\frac{d}{dt}V_I(t) = \int_0^\infty kK_2(a)V_I(t-a)T(t-a)da - cV_I,
\end{aligned} \tag{15}$$

where

$$\begin{aligned}
K_1(a) &= \eta(\epsilon_{RT})\beta(a)e^{-\int_0^a(\delta(s)+\eta(\epsilon_{RT})\beta(s))ds},\\
K_2(a) &= (1-\epsilon_{PI})(1-\beta(a))p(a)e^{-\int_0^a(\delta(s)+\eta(\epsilon_{RT})\beta(s))ds}.
\end{aligned} \tag{16}$$

$K_2(a)$ is the product of the age-specific survival probability of an infected cell and the rate at which infectious virus particles are produced by an infected cell of age a. Thus, the integral of $K_2(a)$ over all ages, i.e.,

$$\mathcal{K}_2 = \int_0^\infty K_2(a)da,$$

gives the total number of infectious virus particles produced by one infected cell over its lifespan. For convenience, we call $\mathcal{K}_2$ the infectious virus burst size.

The Steady States and Stability

System (15) has two steady states: the infection-free steady state

$$\bar{E} = (\bar{T}, \bar{V}_I) = (\frac{\lambda}{d}, 0),$$

and the infected steady state

$$E^\diamond = (T^\diamond, V_I^\diamond) = \left(\frac{c}{k\mathcal{K}_2}, \frac{\lambda k\mathcal{K}_2 - dc}{kc(1-\mathcal{K}_1)}\right) \tag{17}$$

where $\mathcal{K}_1 = \int_0^\infty K_1(a)da$. Noticing that $\mathcal{K}_1 < 1$, we have $V_I^\diamond > 0$ if and only if $\lambda k\mathcal{K}_2 - dc > 0$, or $\mathcal{R}_1 > 1$, where

$$\mathcal{R}_1 = \frac{\lambda k\mathcal{K}_2}{dc}. \tag{18}$$

Clearly, the infected steady state (17) is feasible if and only if $\mathcal{R}_1 > 1$. Notice that λ/d is the cell density in the absence of infection, k and c are the cell infection and viral clearance rate, respectively. Recall that $\mathcal{K}_2$, the infectious virus burst size, gives the number of infectious virus particles produced by one infected cell over its lifespan. Therefore, $\mathcal{R}_1$ gives the reproductive ratio of the virus under the impact of drug treatment.

By analyzing the characteristic equation of system (15), we can show that the infection-free steady state $\bar{E}$ is locally asymptotically stable when $\mathcal{R}_1 < 1$, and it is unstable when $\mathcal{R}_1 > 1$; the infected steady state $E^\diamond$ is locally asymptotically stable whenever it exists, i.e., $\mathcal{R}_1 > 1$ [88].

4.2 The Model with Entry/Fusion and Protease Inhibitors

Significant progress in drug development has been made since the discovery of RT inhibitors and protease inhibitors. Recently, a new class of drugs, entry/fusion inhibitors, have been introduced [22,33] and became available with the FDA approval of enfuvirtide (Fuzeon) in 2003. These compounds can block the fusion of viral envelope to the target cell membrane and interfere with continued infection.

In this section we present an age-structured model that accounts for the effects of both an entry inhibitor and a protease inhibitor. The model can be described by the following equations

$$\begin{aligned}
&\frac{d}{dt}T(t) = \lambda - dT - (1-\epsilon_{EI})kV_IT,\\
&\frac{\partial}{\partial t}T^*(a,t) + \frac{\partial}{\partial a}T^*(a,t) = -\delta(a)T^*(a,t),\\
&\frac{d}{dt}V_I(t) = \int_0^\infty (1-\epsilon_{PI})(1-\beta(a))p(a)T^*(a,t)da - cV_I,\\
&\frac{d}{dt}V_{NI}(t) = \int_0^\infty \epsilon_{PI}(1-\beta(a))p(a)T^*(a,t)da - cV_{NI},\\
&T^*(0,t) = (1-\epsilon_{EI})kV_IT,
\end{aligned} \tag{19}$$

where ϵ_{EI} represents the efficacy of entry inhibitor and the other parameters and variables have the same meaning as in model (14).

The following limiting system is used to study the steady states of (19) [88]:

$$\begin{aligned}\frac{d}{dt}T(t) &= \lambda - dT(t) - (1-\epsilon_{EI})kV_I(t)T(t),\\ \frac{d}{dt}V_I(t) &= \int_0^\infty (1-\epsilon_{EI})kK_3(a)V_I(t-a)T(t-a)da - cV_I,\end{aligned} \tag{20}$$

where

$$K_3(a) = (1-\epsilon_{PI})(1-\beta(a))p(a)e^{-\int_0^a \delta(s)ds}.$$

System (20) has two steady states: the noninfected steady state

$$\bar{E} = (\bar{T}, \bar{V}_I) = (\frac{\lambda}{d}, 0),$$

and the infected steady state

$$E^\circ = (T^\circ, V_I^\circ) = \left(\frac{c}{k(1-\epsilon_{EI})\mathcal{K}_3}, \frac{\lambda k(1-\epsilon_{EI})\mathcal{K}_3 - dc}{kc(1-\epsilon_{EI})}\right)$$

where

$$\mathcal{K}_3 = \int_0^\infty K_3(a)da.$$

Clearly, $V_I^\circ > 0$ if and only if $\lambda k(1-\epsilon_{EI})\mathcal{K}_3 - dc > 0$, or $\mathcal{R}_2 > 1$, where

$$\mathcal{R}_2 = \frac{\lambda k(1-\epsilon_{EI})\mathcal{K}_3}{dc} \tag{21}$$

is the reproductive ratio for the model (19). Hence, the infected steady state E° exists if and only if $\mathcal{R}_2 > 1$.

Using similar arguments [88], we can show that the noninfected steady state is locally asymptotically stable if $\mathcal{R}_2 < 1$, and it is unstable if $\mathcal{R}_2 > 1$; the infected steady state E° is locally asymptotically stable whenever it exists.

4.3 Comparison of Different Drug Combinations

In this section, we provide numerical simulations to compare the effectiveness of different drug combinations. We choose $p(a)$ as in (11) and $\delta(a)$ to be constant $\delta = 1$ day^{-1} [54]. $\beta(a)$ is assumed to be the following function

$$\beta(a) = \begin{cases} 1, & 0 \le a < a_1,\\ 0, & a \ge a_1,\end{cases} \tag{22}$$

with $a_1 = 0.25$ days [46]. We use $T(0) = 10^6$ ml^{-1} [75] and $V(0) = 10^{-6}$ ml^{-1} [94] in the basic model (1) to get the following steady state values:

$T = 3.8333 \times 10^5$ ml^{-1}, $T^* = 6.1675 \times 10^3$ ml^{-1}, $V = 6.7038 \times 10^5$ ml^{-1}, which will be used as the initial values of our models (14) and (19).

The reversion rate function, $\eta(\epsilon_{RT})$, remains to be determined. We know $\eta(0) = 0$ and when $\epsilon_{RT} \to 1$, $\eta(\epsilon_{RT})$ should be sufficiently large such that all the preRT cells will revert back to the uninfected class. In our simulation, we assume the reversion rate function takes the form: $\eta(\epsilon_{RT}) = -\rho \ln(1 - \epsilon_{RT})$, where the constant ρ controls the steepness of the function. From the basic model (1) in which there are only short-lived infected cells, the viral level will be theoretically suppressed to be below the limit of viral detection (50 RNA copies ml^{-1} in the blood) in 10.2 days if RTIs are assumed to be 100% effective (we choose the initial viral load to be $V_0 = 6.7038 \times 10^5$ ml^{-1} and assume the same model parameters)[1]. In model (14), under the same initial conditions and parameters, if we choose $\rho = 2$ day^{-1}, then the viral load can reach the same limit in 10.2 days when the drug efficacy of RTIs is very close to 1. Therefore, we first use the value $\rho = 2$ day^{-1} in our simulation to study the effect of RT inhibitors on HIV-1 dynamics. The abilities of RTIs with different ρ to suppress viral load will be discussed later.

Figures 5 and 6 show numerical simulations of model (14) and (19), respectively. The efficacy of the protease inhibitor is fixed at $\epsilon_{PI} = 0.50$. Figures 5(a,b) and (c,d) are for different values of ϵ_{RT} that increases from $\epsilon_{RT} = 0.2$ (Figure 5(a,b)) to $\epsilon_{RT} = 0.5$ (Figure 5(c,d)). We observe that, when ϵ_{RT} is increased, the infection level at which the system stabilizes is decreased as expected. When ϵ_{RT} is greater than a threshold value ($\epsilon_{RT} = 0.41$, which can be observed in Figure 8(c)), the virus population will die out. Figure 6 shows a similar qualitative behavior of the viral load although the efficacy of the entry inhibitor has a different threshold value, $\epsilon_{EI} = 0.23$ (Figure 8(c)). The different threshold values for viral eradication indicate that the entry inhibitor appears more effective than the RT inhibitor under given conditions. However, this comparison of effectiveness depends heavily on the choice of parameter ρ. If ρ is increased to 5, then RTIs can suppress viral load more effectively than EIs (see more discussion in Figure 8).

We compare the effectiveness of RTIs and EIs in more details with Figure 8. With the functions $p(a)$ and $\beta(a)$ given in (11) and (22), we have the following reproductive ratios

$$\mathcal{R}_1 = e^{-a_1 \eta(\epsilon_{RT})} M_0, \quad \mathcal{R}_2 = (1 - \epsilon_{EI}) M_0, \tag{23}$$

where

$$M_0 = \frac{\lambda k \theta}{cd\delta(\theta + \delta)} (1 - \epsilon_{PI}) p^* e^{-\delta a_1}.$$

Let $V_I^{(1)}$ and $V_I^{(2)}$ denote the viral steady states of models (14) and (19), respectively. Then

[1] In reality, the time to reach this limit is much longer, probably due to the existence of long-lived infected cells and latently infected cells [72, 73].

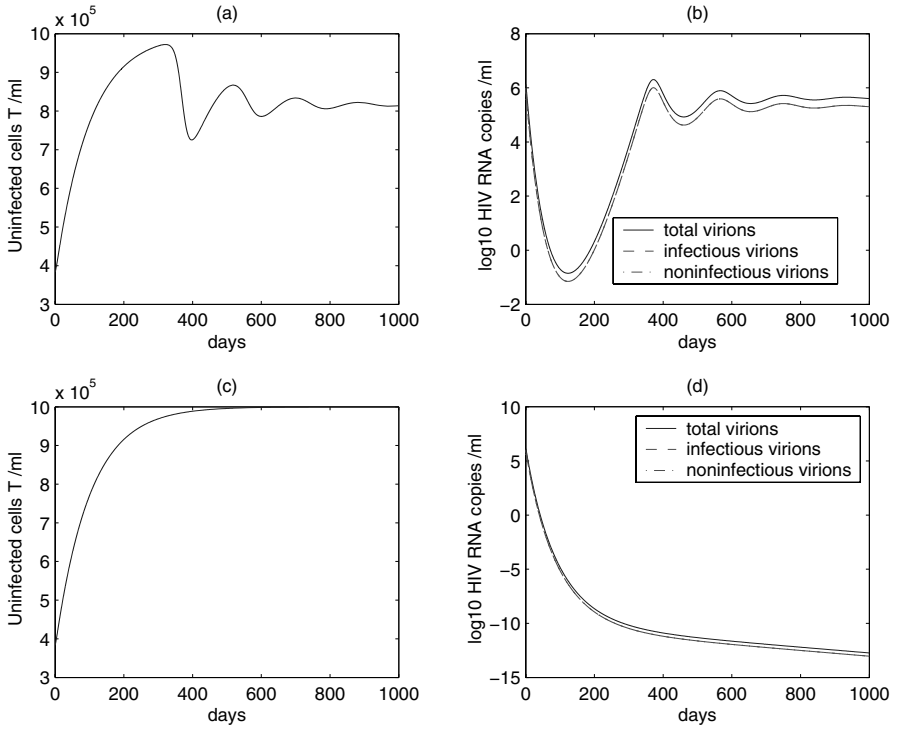


Fig. 1. Simulation of model (14) with $\epsilon_{PI} = 0.50$. The upper panel: $\epsilon_{RT} = 0.20$; the lower panel: $\epsilon_{RT} = 0.50$. The other parameters for each panel are the same [19, 62]: $\lambda = 10^4$ ml^{-1} day^{-1}, $d = 0.01$ day^{-1}, $c = 23$ day^{-1}, $k = 2.4 \times 10^{-8}$ ml day^{-1}, $\delta = 1$ day^{-1}, $p^* = 6.4201 \times 10^3$ day^{-1}, $\theta = 1$, $a_{max} = 10$ days. The reproductive numbers of the upper and lower panel are 1.1666 and 0.9223, respectively. The upper panel shows that the virus population stabilizes at a steady state and uninfected T cells concentration remains at 800 μl^{-1}, the lower panel shows that the virus dies out and the T cell count reaches 1000 μl^{-1}.

$$V_I^{(1)} = \frac{d(\mathcal{R}_1 - 1)}{k(1 - \mathcal{K}_1)}, \quad V_I^{(2)} = \frac{d(\mathcal{R}_2 - 1)}{k(1 - \epsilon_{EI})},$$

where

$$\mathcal{K}_1 = \frac{\eta(\epsilon_{RT})}{\delta + \eta(\epsilon_{RT})}(1 - e^{-(\delta + \eta(\epsilon_{RT}))a_1}).$$

If we assume the reversion rate takes the form $\eta(\epsilon_{RT}) = -\rho \ln(1 - \epsilon_{RT})$, then $\mathcal{R}_1$ can be simplified and equation (23) reduces to

$$\mathcal{R}_1 = (1 - \epsilon_{RT})^{a_1 \rho} M_0, \quad \mathcal{R}_2 = (1 - \epsilon_{EI}) M_0. \tag{24}$$

In Figure 8(c), we let $\epsilon_{PI} = 0.50$ and plot $\mathcal{R}_1(\epsilon_{RT})$ and $\mathcal{R}_2(\epsilon_{EI})$ as functions of ϵ_{RT} and ϵ_{EI}, respectively. We observe that there is a threshold value,

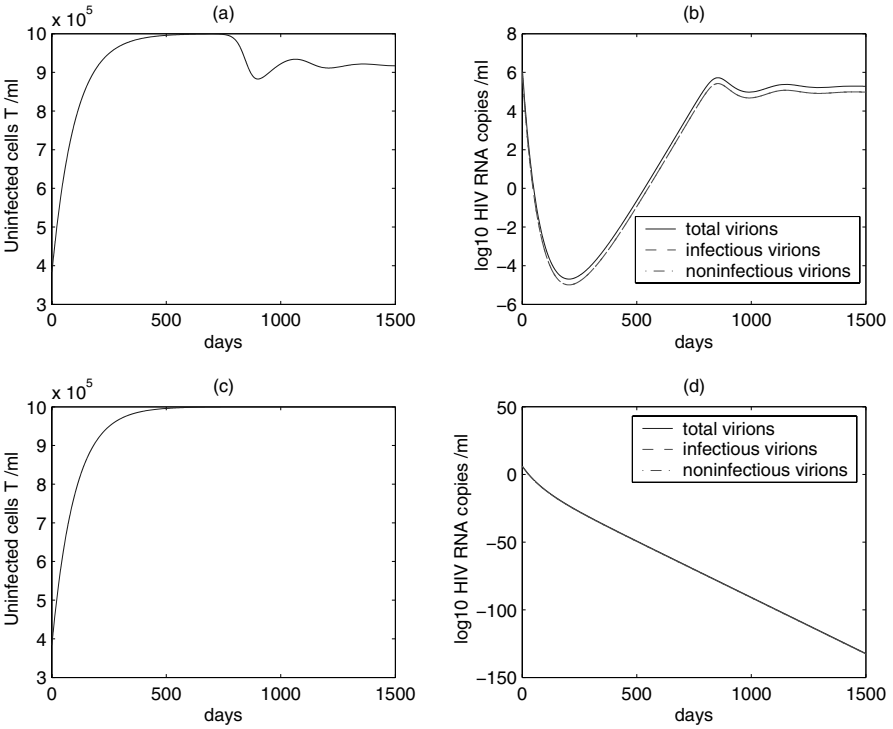


Fig. 2. Simulation of model (19) with $\epsilon_{PI} = 0.50$. The upper panel: $\epsilon_{EI} = 0.20$; the lower panel: $\epsilon_{EI} = 0.50$. The other parameters are the same as those in Figure 5. The reproductive numbers of the upper and lower panel are 1.0435 and 0.6522, respectively. The upper panel shows that the virus population stabilizes at a lower steady state than in Figure 5(b) (the graphs do not show this clearly but numerical values show the difference) and the uninfected T cell concentration remains more than 900 μl^{-1}. The lower panel shows that the virus dies out and the T cell count reaches 1000 μl^{-1}. This suggests that the entry inhibitor might be more effective than the RT inhibitor in the given conditions.

$\epsilon_{RT} = 0.41$, such that $\mathcal{R}_1 > 1$ when $\epsilon_{RT} < 0.41$ and $\mathcal{R}_1 < 1$ when $\epsilon_{RT} > 0.41$. By comparison, the threshold value of ϵ_{EI} is 0.23, which suggests that the virus population will die out when $\epsilon_{EI} > 0.23$.

In Figure 8(d), $V_I^{(1)}(\epsilon_{RT})$ and $V_I^{(2)}(\epsilon_{EI})$ are plotted as functions of ϵ_{RT} and ϵ_{EI}, respectively. Given the same efficacy, $V_I^{(2)}(\epsilon_{EI})$ is less than $V_I^{(1)}(\epsilon_{RT})$, which indicates that the entry inhibitor appears more effective in reducing the viral load than the RT inhibitor in this scenario ($\rho = 2$). In fact, more information can be obtained by looking at the slopes of $\mathcal{R}_1(\epsilon_{RT})$ and $\mathcal{R}_2(\epsilon_{EI})$ in Figure 8(c) since the slope characterizes the effectiveness of drug treatment in infection control when drug efficacy is increased. From (24) we see that $\mathcal{R}_1$ decreases nonlinearly as ϵ_{RT} increases and the decay rate is $(1 - \epsilon_{RT})^{a_1\rho}$,

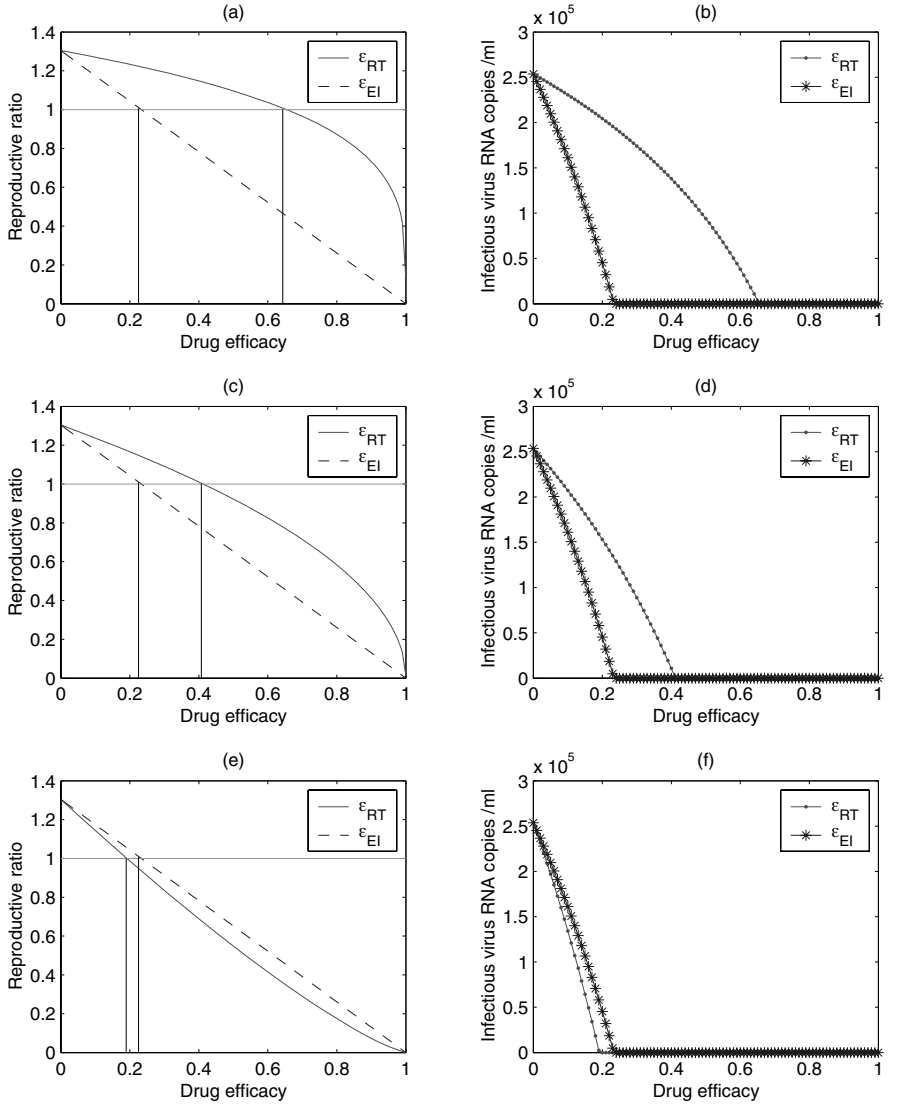


Fig. 3. Comparison of the two combination treatments with fixed drug efficacy $\epsilon_{PI} = 0.50$. Left column: reproductive numbers $\mathcal{R}_1$ and $\mathcal{R}_2$ as the function of ϵ_{RT} and ϵ_{EI}, respectively. If $\epsilon_{EI} > 0.23$ then $\mathcal{R}_2 < 1$ and hence virus will die out. Right column: Steady state V_I of model (14) and (19) as the function of ϵ_{RT} and ϵ_{EI}, respectively. The upper panel: $\rho = 1$, the threshold for $\mathcal{R}_1 < 1$ is $\epsilon_{RT} > 0.65$; the middle panel: $\rho = 2$, the threshold for ϵ_{RT} is 0.41; the bottom panel: $\rho = 5$, the threshold for ϵ_{RT} is 0.19.

while $\mathcal{R}_2$ decreases linearly with the decay rate $(1 - \epsilon_{EI})$ as ϵ_{EI} increases. This implies that the effectiveness of RTIs depends heavily upon the reversion constant ρ. In our simulation, we choose $\rho = 2$ day^{-1} and find that the entry inhibitor is more likely able to eliminate the virus population than the RT inhibitor when the efficacy is increased by the same percentage (see Figures 5, 6 and 8(c,d)). However, when ρ is chosen to be greater than $1/a_1$ we obtain the contrary result (Figure 8(e,f)).

5 Drug Resistance

Treating HIV-infected patients with a combination of several ARV drugs usually contributes to a substantial decline in viral load and an increase in CD4$^+$ T cells. However, continuing viral replication in the presence of therapy can lead to the emergence of drug resistance, which subsequently results in incomplete viral suppression and a greater risk of disease progression.

The question of why drug-resistant strains of HIV-1 emerge during treatment is of great interest. Significant progress has been made both in the genotypic and phenotypic characterization of drug-resistant virus variants. Phenotypic characteristics are conferred by the genotype of mutants with genetic variations and mutations. Most of our knowledge regarding the resistance to nucleoside analogues comes from the genotypic analysis of HIV isolates from patients receiving prolonged therapy with these drugs. Mutations at several codons of reverse transcriptase have been associated with the resistant phenotype [43, 50, 85]. Mutations detected during the clinical use of nevirapine have confirmed that as few as one amino acid change can generate resistance *in vivo* during therapy [34, 87]. Exploring the association between genotypic and phenotypic characteristics will provide information that can be used to understand the emergence of drug resistance and better predict treatment outcomes.

Insights into HIV drug resistance have been obtained from mathematical modeling of antiretroviral responses and the evolution of mutant virus. Kirschner and Webb [47] studied drug resistance during treatment of HIV infection with a single drug and compared the treatment outcomes when drug therapy was initiated at different CD4$^+$ T cell levels. McLean and Nowak [55] showed that the competition between drug-resistant and wild-type strains determines which type of virus will eventually dominate the virus population during the course of AZT treatment. Ribeiro et al. [84] calculated the frequency of drug-resistant mutant virus before the initiation of therapy, and suggested that drug resistant virus can pre-exist in patients and then be selected when therapy is started. Nowak et al. [68] discussed a two-strain model and compared the model results with data on drug resistance development in patients treated with nevirapine. The role of an immune response in the emergence of drug resistance was investigated in [91, 103]. Clinical data on the evolution of drug-resistant mutants for two RT inhibitors, lamivudine and

zidovudine, was explained in detail by a mathematical model that incorporated empirical evidence on the mutation frequency of various HIV-1 drug resistance mutations [97]. Murray and Perelson [67] showed how the quasi-species nature of HIV can influence the development of resistance to AZT and the maintenance of drug resistance clones after cessation of therapy. Kepler and Perelson [44] showed how drug concentration heterogeneity facilitates the evolution of drug resistance.

In the clinic, HIV resistance can result from the transmission of drug-resistant mutants to susceptible individuals [5, 6] or from the acquisition of mutations generated during treatment. It is important to distinguish between scenarios in which drug-resistant virus preexists before the onset of therapy or in which they are produced by residual virus replication during therapy, because each process requires different drug regimens to maximize the clinical benefits [8]. The calculation of probabilities of both processes suggests that under a wide range of conditions, treatment failure is most likely due to the preexistence of drug-resistant virus before therapy [83]. Provided that drug-resistant virus preexists, Bonhoeffer and Nowak [8] showed that a more efficient therapy would lead to a greater initial reduction of virus, but would also cause a faster rise of resistance mutations.

In this section we investigate analytically the mechanisms underlying the emergence of drug resistance during therapy. A two-strain model is employed to study the effect of ARV drugs on the evolution of drug-resistant HIV mutants. Although HIV resistance is not an all-or-nothing phenomenon and generally mutations accumulate to provide more resistance to drug therapy [11], even a single mutation can confer a significant degree of resistance to a drug or an entire class of drugs [8, 50, 61, 95]. For example, the M184V mutation in reverse transcriptase can result in complete resistance to lamivudine [11]. In this model, we assume that the drug sensitive and resistant strains differ by a single mutation. The model can be extended to include multiple resistant strains in which two or more point mutations are required [84, 97]. We will also discuss the evolution of resistant strains that require multiple mutations with the present model.

5.1 The Pretreatment Two-strain Model

Because of an exceptionally high viral replication rate and highly error-prone reverse transcription, the probability that mutations occur is quite high (the average number of changes per genome is 0.3 per replication cycle, which implies that after reverse transcription about 22% of infected cells should carry proviral genomes with one mutation [74,77]). As a single mutation or a number of mutation combinations can result in drug resistance, there is a reasonable chance that resistant HIV variants preexist even before the initiation of therapy [83]. Here we use a two-strain model including both wild-type, i.e., drug sensitive, and drug resistant strains to study the viral load of each strain in the absence of ARV drugs. The model is

$$\begin{aligned}
\frac{d}{dt}T(t) &= \lambda - dT - k_s V_s T - k_r V_r T,\\
\frac{d}{dt}T_s(t) &= (1-\mu)k_s V_s T - \delta T_s,\\
\frac{d}{dt}V_s(t) &= N_s \delta T_s - cV_s,\\
\frac{d}{dt}T_r(t) &= \mu k_s V_s T + k_r V_r T - \delta T_r,\\
\frac{d}{dt}V_r(t) &= N_r \delta T_r - cV_r,
\end{aligned} \tag{25}$$

where the subscripts s and r represent "sensitive-strain" and "resistant-strain", respectively. μ $(0 \leq \mu < 1)$ is the rate at which cells infected by sensitive virus mutate and become drug resistant during the process of reverse transcription of viral RNA into proviral DNA. Here we have ignored the backward mutation from resistant to sensitive strain since the wild-type virus dominates the population before the initiation of therapy [8, 68, 91].

We remark that several two-strain models have been studied in the literature [7, 8, 55, 68, 84]. However, many of them focus on the question that whether drug-resistant mutants preexist before the therapy and are selected in the pressure of drugs or if they are produced only in the course of treatment. Here we aim to study in depth the virus dynamics as well as the development of drug-resistant strains. Particularly, we attempt to address the following questions: Under what condition does the wild-type strain dominate in the absence of therapy? How does the resistant strain develop when ARV drugs are used? Is it possible to eradicate both strains? How soon will the resistance appear if a small number of drug doses are missed? What is the situation if more doses are missed?

Model Parameters

Since mutant strains are often associated with the changes of highly conserved amino-acid residues that are believed to be important for enzyme function, many resistant mutants display some extent of resistance-associated loss of fitness when compared with drug sensitive strains [12]. Recent evidence also shows that the emergence of drug resistance reduces the inherent replicative capacity of resistant strains although it increases the ability of HIV to replicate in the presence of drugs [3]. However, direct experimental measurements to date have not determined whether the replicative defect is due to impaired infectivity of HIV or reduced viral production [93]. Thus, we assume that both the infection rate and burst size of resistant strain are less than those of the wild-type strain, i.e., $k_r < k_s$ and $N_r < N_s$. We choose $k_s = 2.4 \times 10^{-8}$ ml day^{-1}, $k_r = 2.0 \times 10^{-8}$ ml day^{-1}, $N_s = 3000$ and $N_r = 2000$ in our simulation.

The mutation rate from the wild-type strain to a drug-resistant strain is $\mu = 3 \times 10^{-5}$ [53]. This mutation rate applies only when the wild-type and the

mutant strains differ by a single point mutation. In the case of two or more point mutations, the probability of mutation directly from the wild-type to a resistant strain will be much smaller (an approximation is μ^n where n is the number of point mutations they differ by). We discuss the evolution of resistant strains with multiple mutations later.

The other parameters are chosen to be the same as those in Section 4.3.

The Steady States and Stability

Let $\bar{E} = (\bar{T}, \bar{T}_s, \bar{V}_s, \bar{T}_r, \bar{V}_r)$ denote a constant solution (steady state) of model (25). There are three possible steady states: the infection-free steady state

$$E_0 = (\frac{\lambda}{d}, 0, 0, 0, 0), \tag{26}$$

the boundary steady state (only the resistant strain is present)

$$E_r = \left(\frac{c}{k_r N_r}, 0, 0, (\mathcal{R}_r - 1)\frac{dc}{k_r N_r \delta}, (\mathcal{R}_r - 1)\frac{d}{k_r}\right), \tag{27}$$

and the interior steady state (coexistence of both the wild-type and resistant strains)

$$E_c = \left(\frac{c}{(1-\mu)k_s N_s}, \tilde{T}_s, \frac{N_s\delta}{c}\tilde{T}_s, \tilde{T}_r, \frac{N_r\delta}{c}\tilde{T}_r\right) \tag{28}$$

where

$$\tilde{T}_s = \frac{[(1-\mu)\sigma - 1][(1-\mu)\mathcal{R}_s - 1]\lambda}{(\sigma-1)(1-\mu)\mathcal{R}_s\delta}, \quad \tilde{T}_r = \frac{[(1-\mu)\mathcal{R}_s - 1]\mu\sigma\lambda}{(\sigma-1)(1-\mu)\mathcal{R}_s\delta},$$

$\sigma = \mathcal{R}_s/\mathcal{R}_r$, and $\mathcal{R}_s$ and $\mathcal{R}_r$ are the basic reproductive ratios of the sensitive strain and the resistant strain, respectively, which are given by

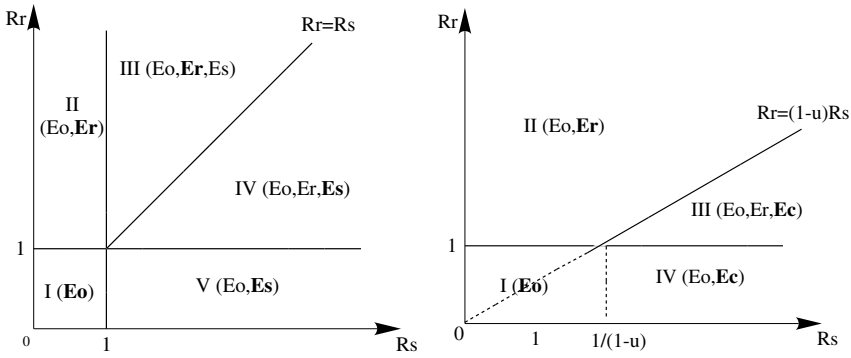
$$\mathcal{R}_s = (k_s N_s \lambda)/(dc), \quad \mathcal{R}_r = (k_r N_r \lambda)/(dc). \tag{29}$$

Notice that in a special case $\mu = 0$ (i.e., there is no mutation), the interior steady state E_c reduces to another boundary steady state E_s (at which only the wild-type strain is present):

$$E_s = \left(\frac{c}{k_s N_s}, (\mathcal{R}_s - 1)\frac{dc}{k_s N_s \delta}, (\mathcal{R}_s - 1)\frac{d}{k_s}, 0, 0\right), \tag{30}$$

and the other steady states (E_0 and E_r) remain the same.

It is clear that E_r exists if and only if $\mathcal{R}_r > 1$, and E_c exists if and only if $\mathcal{R}_s > 1/(1-\mu)$ and $\sigma > 1/(1-\mu)$. In fact, we can prove that these existence conditions also provide threshold conditions for the stability of steady states [89]. All the stability results for the pretreatment model (25) are summarized in Figure 2.



(a) $\mu = 0$ (b) $0 < \mu < 1$

Fig. 4. Steady states of model (25) and stability regions (the steady state in bold type is stable in that region).

From the infected steady states (E_r and E_c), we observe that drug-resistant mutants will always be present as long as mutation from the wild-type to the mutant strain is possible (i.e., $\mu \neq 0$), and that the wild-type strain cannot persist alone. Considering the assumptions $k_r < k_s$ and $N_r < N_s$, we have $\mathcal{R}_s > \mathcal{R}_r > 1$ in the absence of treatment (with the assumed parameters). Hence, σ is greater than 1 but not close to 1. Therefore, from the formula for $\tilde{V}_r$ (see (28)) we know that the viral load of the resistant strain is very low (e.g., under the detection limit with the assumed parameter values) due to the low mutation rate μ despite the coexistence of both strains. The more point mutations needed to confer drug resistance for the mutant strain, the lower the viral load of the resistant strain in the steady state. This explains why the wild-type virus dominates the virus population before the initiation of treatment.

Using initial conditions $T(0) = 10^6$ ml^{-1} and $V_s(0) = 10^{-6}$ ml^{-1} in (25) gives the steady states of the pretreatment model: $T = 3.8334 \times 10^5$ ml^{-1}, $T_s = 6.1660 \times 10^3$ ml^{-1}, $V_s = 6.7022 \times 10^5$ ml^{-1}, $T_r = 0.5550$ ml^{-1}, $V_r = 48.2600$ ml^{-1}. The viral load of the resistant strain is much smaller than that of the wild-type strain in the absence of therapy. These steady state values will be used as the initial conditions to perform simulations of virus dynamics during therapy.

5.2 Effect of Antiretroviral Therapy

The preceding discussion suggests that resistant strains exist before the initiation of antiretroviral therapy. However, the viral level of resistant mutants is relatively low and wild-type virus dominates the population. Given a small mutation rate μ, the steady state of mutant virus can be approximated as:

$$\tilde{V}_r = \frac{d(\mathcal{R}_s - 1)}{k_r(\sigma - 1)}\mu, \tag{31}$$

which is proportional to the mutation rate. This implies that drug resistant mutants are less likely to arise if they differ from wild type by two or more points mutations. In this section, we analytically study the mechanisms underlying the emergence of resistant strains during therapy.

Let ϵ^s_{RT}, ϵ^r_{RT} be the efficacies of RTIs and ϵ^s_{PI}, ϵ^r_{PI} be the efficacies of PIs for the sensitive and resistant strain, respectively. We incorporate the effect of drugs in the two-strain model (25) and obtain the following equations:

$$\begin{aligned}
\frac{d}{dt}T(t) &= \lambda - dT - k_s(1-\epsilon^s_{RT})V_sT - k_r(1-\epsilon^r_{RT})V_rT,\\
\frac{d}{dt}T_s(t) &= (1-\mu)k_s(1-\epsilon^s_{RT})V_sT - \delta T_s,\\
\frac{d}{dt}V_s(t) &= N_s(1-\epsilon^s_{PI})\delta T_s - cV_s,\\
\frac{d}{dt}T_r(t) &= \mu k_s(1-\epsilon^s_{RT})V_sT + k_r(1-\epsilon^r_{RT})V_rT - \delta T_r,\\
\frac{d}{dt}V_r(t) &= N_r(1-\epsilon^r_{PI})\delta T_r - cV_r.
\end{aligned} \tag{32}$$

Here V_s and V_r represent infectious wild-type virus and infectious resistant virus, respectively. We have left out two equations that represent the non-infectious virus of both strains since they can be decoupled from model (32).

Provided that all the efficacies are constant, the incorporation of drug effects will not affect our analysis of the steady states of the pretreatment model (25). The new reproductive ratios in the presence of drug therapy can be expressed as:

$$\mathcal{R}'_s = (1-\epsilon^s_{RT})(1-\epsilon^s_{PI})\mathcal{R}_s, \quad \mathcal{R}'_r = (1-\epsilon^r_{RT})(1-\epsilon^r_{PI})\mathcal{R}_r, \tag{33}$$

where $\mathcal{R}_s$ and $\mathcal{R}_r$ are given in (29). We can also define the ratio σ' as $\mathcal{R}'_s/\mathcal{R}'_r$.

Because of the reduced viral fitness of the resistant strain compared with the wild-type strain, $\mathcal{R}_s > \mathcal{R}_r > 1$ and hence the wild-type strain dominates the virus population before therapy. After the initiation of treatment, the reproductive ratios of both strains begin to decrease. However, we have $\epsilon^s_{RT} > \epsilon^r_{RT}$ and $\epsilon^s_{PI} > \epsilon^r_{PI}$ as the wild type is more susceptible to drugs, thus $\mathcal{R}'_s$ decreases faster than $\mathcal{R}'_r$. For convenience, we define an overall treatment effect for each strain, i.e.,

$$\epsilon_s = 1-(1-\epsilon^s_{RT})(1-\epsilon^s_{PI}), \quad \epsilon_r = 1-(1-\epsilon^r_{RT})(1-\epsilon^r_{PI}). \tag{34}$$

We begin with the assumption $\epsilon_r = \alpha\epsilon_s$, where α $(0 < \alpha < 1)$ represents the resistance level of the mutant virus. A larger α indicates that the resistant

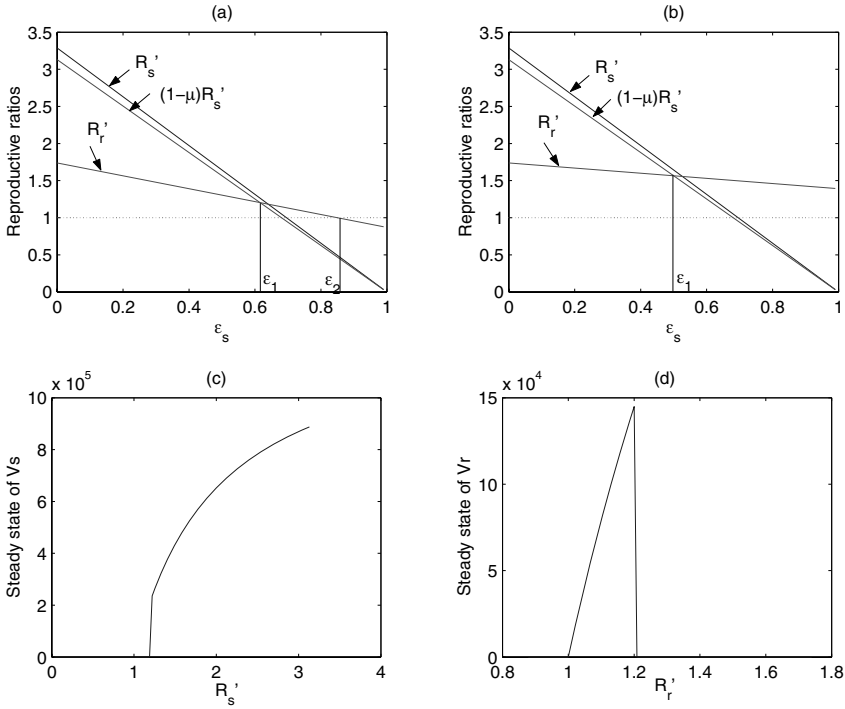


Fig. 5. The upper panel: the reproductive ratio of each strain as a function of the overall drug efficacy for the sensitive strain ϵ_s (see (33) and (34)). (a) $\alpha = 0.5$; (b) $\alpha = 0.2$. ϵ_r is reduced by a factor α, i.e., $\epsilon_r = \alpha\epsilon_s$. For the ease of illustration, we intentionally enlarge the difference between two lines $\mathcal{R}_s'$ and $(1-\mu)\mathcal{R}_s'$. The lower panel: steady states of the wild-type and resistant virus as the function of reproductive ratios, $\mathcal{R}_s'$ and $\mathcal{R}_r'$, respectively. α is chosen to be 0.5. We observe that the wild-type virus can be completely suppressed even when the reproductive ratio $\mathcal{R}_s'$ is greater than 1. The resistant virus dies out only when $\mathcal{R}_r' < 1$, and remains at a very low level (not clearly shown in the figure due to the magnitude of the vertical axis) when $\mathcal{R}_r' > 1.21$ (corresponding to a low drug efficacy).

strain is less resistant to the drug used. Estimates of drug efficacies for the wild-type and resistant strains will be provided in next section. Here we aim to obtain some explicit conditions for the emergence of drug resistance by assuming a fixed drug efficacy.

When the mutant strain is less resistant to the therapy (corresponding to a larger α, e.g., $\alpha = 0.5$), a typical scenario of the variation of reproductive numbers $\mathcal{R}_s'$ and $\mathcal{R}_r'$ is depicted in Figure 3(a). We observe that both $\mathcal{R}_s'$ and $\mathcal{R}_r'$ decrease as the drug efficacy ϵ_s increases. There are two threshold values for ϵ_s. One is ϵ_1, the intersection of $(1-\mu)\mathcal{R}_s'$ and $\mathcal{R}_r'$, which is given by

$$\epsilon_1 = \frac{(1-\mu)\mathcal{R}_s - \mathcal{R}_r}{(1-\mu)\mathcal{R}_s - \alpha\mathcal{R}_r}, \tag{35}$$

and the other one is ϵ_2 with the expression

$$\epsilon_2 = \frac{\mathcal{R}_r - 1}{\alpha\mathcal{R}_r}. \tag{36}$$

If the drug efficacy ϵ_s is less than ϵ_1 then $(1-\mu)\mathcal{R}'_s > \mathcal{R}'_r > 1$. Thus, both strains will coexist (similar to the pretreatment case), and the treatment fails primarily due to the wild-type virus. It is interesting to explore the variation of the viral steady state levels for both strains. When ϵ_s is small, the viral load of the resistant strain is very low (see Eq. (31)) compared with the wild-type virus although both strains coexist. As the drug efficacy increases, the steady state level of the resistant strain increases very slowly (see also Figure 4(b)), whereas the wild type decreases quickly (Figure 4(c)). When ϵ_s approaches ϵ_1, we know that $\sigma' = \mathcal{R}'_s/\mathcal{R}'_r$ increases and approaches $1/(1-\mu)$. From the formula for $\tilde{V}_r$ given in (28), a direct calculation shows that the steady state of the resistant virus in the presence of therapy increases substantially to $(\mathcal{R}'_r - 1)d/(k_r\epsilon^r_{RT})$. Coincidentally, this maximum value is identical to the steady state of the drug-resistant virus when only the mutant strain exists (see Eq. (27)). The steady state of the wild-type virus decreases rapidly to 0 as ϵ_s approaches ϵ_1 (see Figure 4(c)).

For intermediate values of ϵ_s in the interval (ϵ_1, ϵ_2), we know that $\mathcal{R}'_r > 1$ and $\mathcal{R}'_r > (1-\mu)\mathcal{R}'_s$. Thus, only the resistant virus will persist (see Figure 2(b) and Figure 4(a)). When the drug efficacy increases from ϵ_1 to ϵ_2, the steady state level of the drug-resistant virus decreases from the maximum value to 0 (Figure 4(a)).

It is important to note that under drug therapy ($\epsilon_s > \epsilon_1$) the wild-type virus can be suppressed even when the reproductive ratio of the wild-type strain is greater than 1 (see Figure 3(a,c) and Figure 4(c)). This is not surprising because the resistant and sensitive strains compete for the exact same resources—uninfected T cells, hence the resistant strain that becomes more fit as $\epsilon_s > \epsilon_1$ will outcompete the sensitive one according to the competitive exclusion principle. In Figure 3(c), we plot the steady state of the wild-type virus, V_s, as a function of the reproductive ratio $\mathcal{R}'_s$. We observe that, as the drug efficacy increases, the steady state of V_s decreases to 0 even when $\mathcal{R}'_s$ is greater than 1. Figure 3(d) is for the steady state of the resistant virus. The resistant virus dies out only if $\mathcal{R}'_r < 1$. When $\mathcal{R}'_r > 1.21$, the steady state of V_r remains at a very low level, which is not clearly shown in the figure due to the magnitude of the vertical axis.

Finally, for a large value of drug efficacy, $\epsilon_s > \epsilon_2$, we have that $\mathcal{R}'_r < 1$ and $\mathcal{R}'_s < 1/(1-\mu)$. Thus, both strains are predicted to be eradicated by the treatment (Figure 4(a,c,d)).

In Figure 3(b), we use a smaller α ($\alpha = 0.2$) to illustrate another case in which the mutant strain is more resistant to the drug than that in Figure 3(a). We observe that the new reproductive ratio of the resistant strain

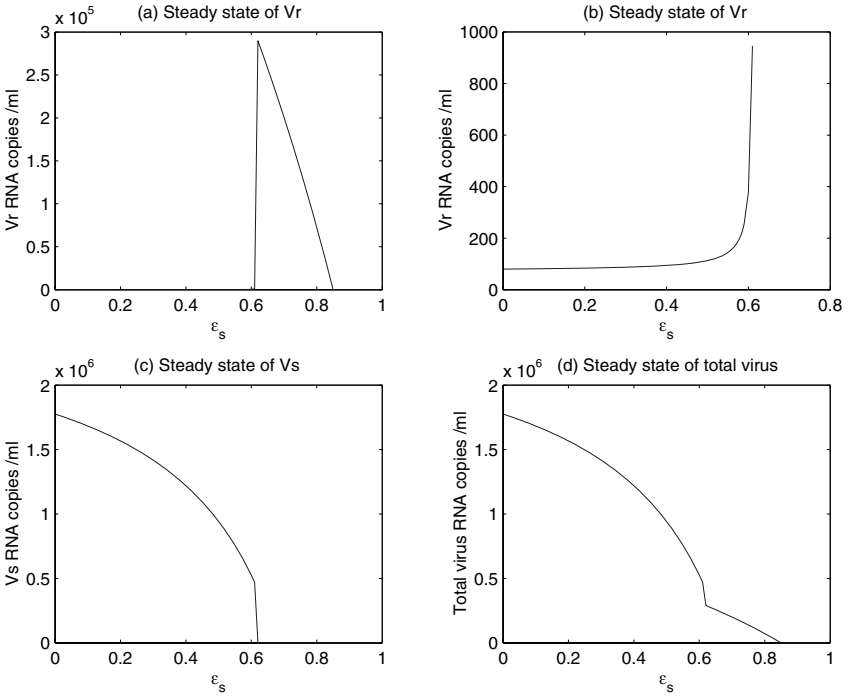


Fig. 6. $\alpha = 0.5$. Steady state of virus of model (32) as a function of the overall drug efficacy for the sensitive strain ϵ_s. For simplicity, we choose $\epsilon_{PI}^s = \epsilon_{PI}^r = 0$. Thus, $\epsilon_{RT}^s = \epsilon_s$, $\epsilon_{RT}^r = \epsilon_r$. We also assume that $\epsilon_r = \alpha\epsilon_s$. (a) Steady state of drug-resistant virus as a function of ϵ_s. (b) The change of steady state of resistant virus is illustrated when the drug efficacy is less than 0.61 (the threshold value for both strains to coexist). As ϵ_s increases, the steady state increases slowly at first, but it increases substantially to the maximum value when $\epsilon_s \to 0.61$. (c) Steady state of the wild-type virus as a function of ϵ_s. The wild-type viral level decreases when the drug efficacy increases, and declines significantly to 0 when $\epsilon_s \to 0.61$. (d) Total virus population as the function of drug efficacy. Although the drug-resistant strains undergo a rapid increase during therapy, the total virus keeps decreasing as the drug efficacy increases. When $\epsilon_s > 0.85$, both strains of virus will be eradicated.

will never decrease to below 1 even when the drug efficacy ϵ_s increases to 1. Consequently, there is only one threshold value ϵ_1, which is about 0.49. When $\epsilon_s < 0.49$, both strains coexist (Figure 7(a,c)), and the wild-type virus dominates the population if ϵ_s is not close to 0.49. The steady state level of the resistant virus undergoes a substantial increase when ϵ_s approaches 0.49 (Figure 7(b)). As $\epsilon_s > 0.49$, only the resistant strain persists.

From Figure 4(d) and Figure 7(d), we also observe that the steady state of the total virus (non-infectious virus is not counted) decreases as the drug efficacy increases. The treatment success shown in Figure 4(d) is because

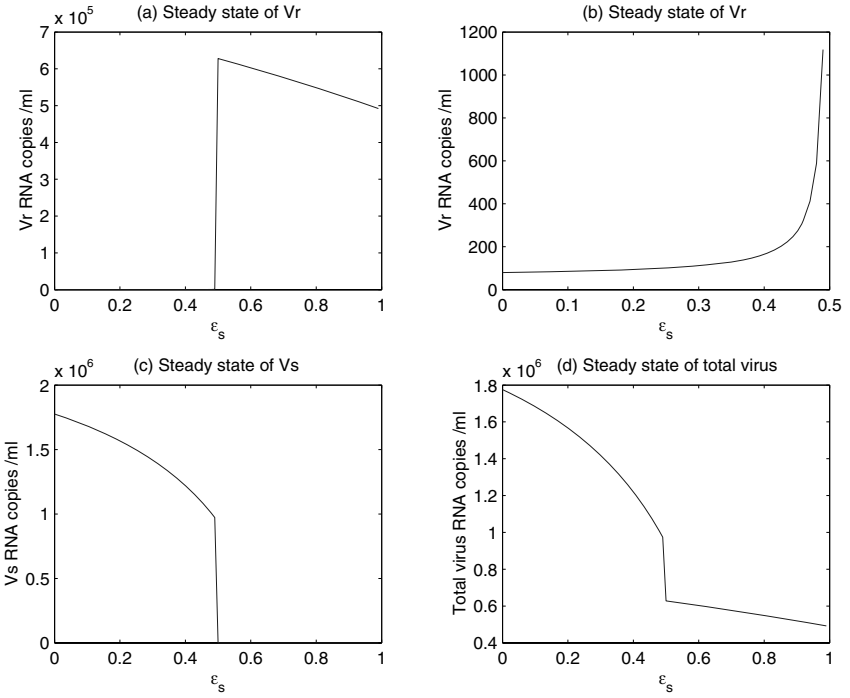


Fig. 7. Similar to Figure 4 except $\alpha = 0.2$. When the drug efficacy approaches the threshold value 0.49, the resistant strain undergoes a great increase. Even when the drug is 100% effective against the wild-type virus, the (resistant) virus still persists.

we have assumed that the resistant strain is still very susceptible to the drug regimen. When it is less susceptible to the drug, the virus can not be eradicated even with a highly effective treatment (see Figure 7(d)).

5.3 Time-varying Drug Efficacy and Effect of Adherence

Up to now, we have studied analytically the viral level in the presence of antiretroviral treatment, particularly the development of drug-resistance strains. In the analysis, drug efficacies of RTIs and PIs were assumed to be constant for both strains. This assumption may not be realistic since drug concentrations in the blood and in cells continuously vary when the drug is absorbed, distributed, and metabolized in the body. To estimate this time-varying drug efficacy, the pharmacokinetics and pharmacodynamic processes that determine drug action need to be considered [19].

Another important factor affecting drug efficacy is the level of adherence to regimen protocols. In clinical practice, it is widely believed that the level of compliance with prescribed ARV regimens is one of the crucial determinants of a successful treatment [15, 29, 61]. Much evidence shows that suboptimal

adherence is associated with the emergence of drug resistance [2, 29, 90, 101], viral rebound, and consequently an increased risk of transmitting drug resistant virus [1, 16, 100]. Richman [86] postulated that the relationship between drug resistance and antiretroviral activity was a "bell-shaped" curve - that is, drug resistance is more possible to appear with moderate levels of adherence than with perfect or low levels of adherence to highly potent treatment. With perfect adherence there might be little viral replication, whereas with poor adherence there might be insufficient drug pressure to select resistant variants. Wahl and Nowak [101] analyzed the outcome of therapy as a function of the degree of adherence to drug regimen and determined the conditions under which a resistant strain dominates.

In this section, we first adopt a pharmacokinetic model developed recently by Dixit and Perelson [19] to estimate the drug efficacies of tenofovir disoproxil fumarate (NRTI) and ritonavir (PI) for both strains of virus. We then discuss how these time-varying drug efficacies, due both to dosing regimens and different patterns of non-adherence, affect the antiviral responses, particularly the emergence of drug resistance.

Models for Drug Efficacy and Adherence

There are many existing models that use drug concentrations to estimate the efficacy of antiviral treatment [18, 19, 42, 101, 105]. Huang et al. [42] used a standard pharmacokinetic one-compartment model to estimate the drug efficacy and studied how drug pharmacokinetics affects antiviral response. However, it is the intracellular concentration rather than plasma concentration of drugs that determines drug effectiveness. Thus, here we will employ a two-compartment pharmacokinetic model developed by Dixit and Perelson [19] to estimate the efficacies of two ARV drugs, tenofovir disoproxil fumarate (DF) and ritonavir, for both the wild-type and resistant strains.

We first describe the two-compartment model briefly as follows (refer to [19] for a detailed description of the model formulation).

The instantaneous drug efficacy $\epsilon(t)$ can be estimated by a simple function [18, 30]

$$\epsilon(t) = \frac{C_c(t)}{IC_{50} + C_c(t)}, \tag{37}$$

where $C_c(t)$ is the intracellular concentration of the drug used, IC_{50} is a phenotype marker representing the intracellular concentration of drug needed to inhibit the viral replication by 50%.

If multiple doses of a drug are administered, then the concentration of drug in the blood is given by

$$C_b(t) = \frac{FDk_a e^{-k_e t}}{V_d(k_e - k_a)(e^{k_a I_d} - 1)} \left[1 - e^{(k_e - k_a)t}(1 - e^{N_d k_a I_d}) + \frac{(e^{k_e I_d} - e^{k_a I_d})(e^{(N_d - 1)k_e I_d} - 1)}{e^{k_e I_d} - 1} - e^{((N_d - 1)k_e + k_a)I_d} \right], \tag{38}$$

where F is the bioavailability of drug, D is the mass of drug administered in one dose, V_d is the volume of distribution, k_a and k_e are pharmacokinetic parameters that can be estimated from experiments. I_d is the dosing interval and N_d is the number of doses until time t.

For PIs, the intracellular concentration, C_c, can be derived directly according to the drug transport from the blood into the cell compartment:

$$\frac{dC_c}{dt} = k_{acell}C_x - k_{ecell}C_c, \tag{39}$$

where

$$C_x = \begin{cases} (1-f_b)HC_b - C_c & \text{if } (1-f_b)HC_b - C_c > 0, \\ 0 & \text{else.} \end{cases} \tag{40}$$

In the above equations, drug absorption is being driven by an effective concentration gradient. H quantifies the drug partitioning effect of the cell membrane, f_b denotes the fraction of drug that can not be transported into cells, k_{acell} and k_{ecell} represent the cellular absorbtion and elimination, respectively. See [19] for further details.

For RTIs, the drug action is more complicated. They need to be phosphorylated to their active forms in cells. Dixit and Perelson [19] used the following equations to model the phosphorylation of tenofovir DF:

$$\begin{aligned} \frac{dC_c}{dt} &= k_{acell}C_x - k_{ecell}C_c - k_{1f}C_c + k_{1b}C_{cp}, \\ \frac{dC_{cp}}{dt} &= -k_{ecell}C_{cp} + k_{1f}C_c - k_{1b}C_{cp} - k_{2f}C_{cp} + k_{2b}C_{cpp}, \\ \frac{dC_{cpp}}{dt} &= -k_{ecell}C_{cpp} + k_{2f}C_{cp} - k_{2b}C_{cpp}, \end{aligned} \tag{41}$$

where C_c, C_{cp} and C_{cpp} represent the respective intracellular concentrations of the native (monophosphorylated), diphosphorylated and triphosphorylated forms of the drug, C_x is given in (40), k_{1f}, k_{1b}, k_{2f} and k_{2b} characterize the phosphorylation reactions among C_c, C_{cp} and C_{cpp}.

Solving the equations in (41) with the initial condition

$$C_c(0) = C_{cp}(0) = C_{cpp}(0) = 0$$

and substituting C_{cpp} for C_c in (37), we obtain the time-dependent efficacy of the RT inhibitor, which is plotted in Figure 9(a). Similarly, solving the equation (39) with initial condition $C_c(0) = 0$ and plugging the solution into (37), we get the efficacy of the protease inhibitor, which is plotted in Figure 9(b).

In Figure 9, we also estimate the drug efficacy for the resistant strain with perfect/suboptimal drug adherence. We assume that the drug resistance increases the virus IC_{50} 10-fold [50]. The efficacy for the resistant strain is plotted in Figure 9(a,b) for comparison with that of the wild-type strain. Figure 9(c,d) plots the drug efficacy when every other dose of tenofovir DF and ritonavir is missed. It is clear that when a dose of drug is missed, the drug efficacy decreases to a low level before the next dose is administered.

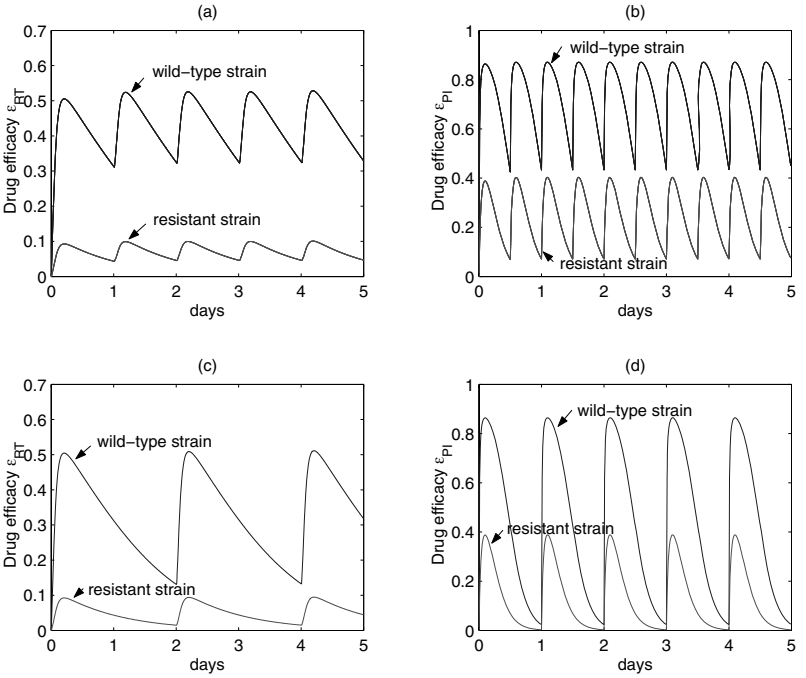


Fig. 8. (a) Drug efficacy of tenofovir DF (RTI) for both strains. The parameters are [19]: $D = 300$ mg, $I_d = 1$ day, $F = 0.39$, $V_d = 87500$ ml, $k_a = 14.64$ day^{-1}, $k_e = 9.6$ day^{-1}, $H = 1800$, $f_b = 0.07$, $k_{1f} = 9.6$ day^{-1}, $k_{1b} = 30.3$ day^{-1}, $k_{2f} = 270.7$ day^{-1}, $k_{2b} = 95.5$ day^{-1}, $k_{acell} = 24000$ day^{-1}, $k_{ecell} = 1.1$ day^{-1}, IC_{50} is 0.54 mg ml^{-1} for wild-type strain. (b) Drug efficacy of ritonavir (PI) for both strains. The parameters are [19]: $D = 600$ mg, $I_d = 0.5$ days, $F = 1$, $V_d = 28000$ ml, $k_a = 14.64$ day^{-1}, $k_e = 6.86$ day^{-1}, $H = 0.052$, $f_b = 0.99$, IC_{50} is 9×10^{-7} mg ml^{-1} for wild-type strain. (c) Every other dose of tenofovir DF is missed, i.e., $I_d = 2$ days. (d) Every other dose of ritonavir is missed, i.e., $I_d = 1$ day. We assume that drug resistance increases the IC_{50} value 10-fold.

Effects on Virus Dynamics

We conduct numerical simulations to study the time evolution of uninfected T cells and two strains of virus based on the time-varying drug efficacies obtained in previous section.

We first consider the scenario of perfect adherence to the prescribed dose levels and dosing times. As a single point mutation can confer different levels of resistance to the drug [51], we consider two cases in which drug resistance increases the baseline IC_{50} by different amounts. In Case 1, drug resistance is assumed to increase the IC_{50} value 10-fold; Case 2 deals with a 15-fold increase of IC_{50} for the resistant strain. Figure 10(b) shows that the viral load of the wild-type strain can be suppressed very well for each case if HIV-

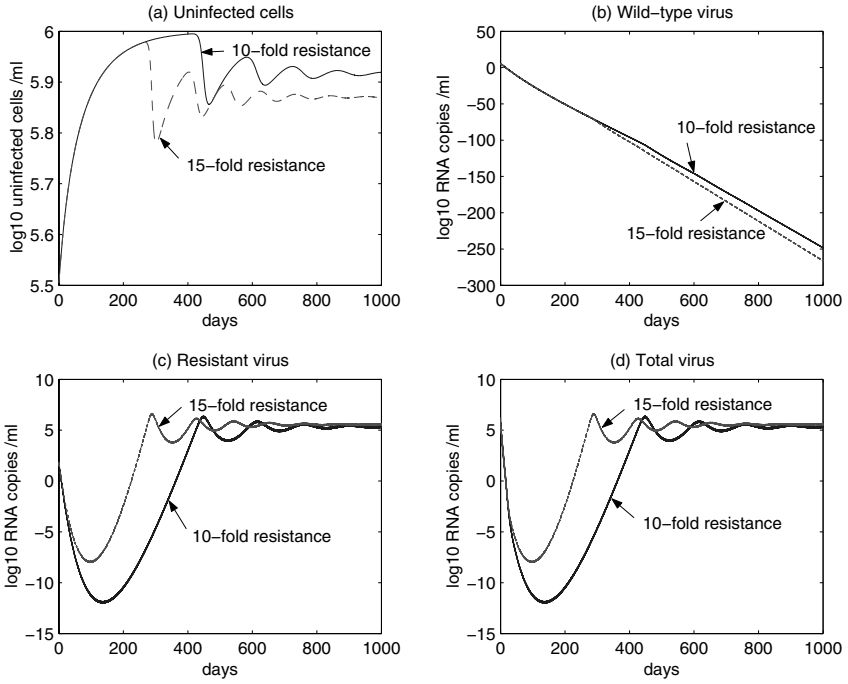


Fig. 9. Dynamics with perfect adherence. The solid line represents the dynamics of T cells and virus when drug resistance increases the baseline IC_{50} 10-fold. The dashed line is for 15-fold resistance. In both cases, wild-type virus will be eradicated due to the perfect adherence to ART. The higher resistance level of the mutant strain, the more quickly drug-resistant virus will emerge and dominate. The viral peak and steady state of drug-resistant virus are also increased when the resistance level increases. All the variables converge to steady states with very frequent oscillations because of the time-dependent drug efficacy.

infected individuals can adhere to the treatment perfectly. Drug resistance appears about one year after the initiation of therapy in Case 1, while the emergence of resistance in Case 2 is much earlier (see Figure 10(c)). We can also observe that if the mutant virus is more resistant to drugs, then the viral peak and the final steady state level are higher than that of the strain with a lower resistance level. In fact, if drug resistance increases the IC_{50} value only 5-fold or less (not shown in the figure), then both strains will be suppressed if perfect adherence is followed. From Figure 10(a), we observe that during therapy the number of uninfected T cells in both cases will oscillate to a steady state lower than the normal $CD4^+$ T level in uninfected individuals (1000 cells/μl). Case 2 (15-fold resistance) has a lower steady state level of uninfected T cells than that of Case 1 (10-fold resistance).

In Figure 11, we simulate the viral response under imperfect adherence. We still consider two cases. In the first case, the individual misses every other

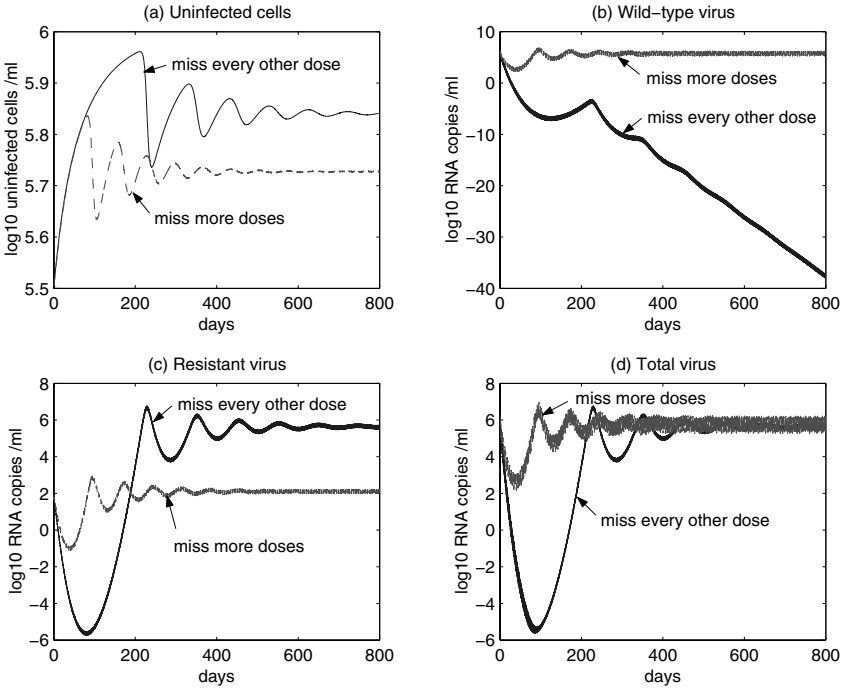


Fig. 10. Dynamics with suboptimal adherence. Drug resistance is assumed to increase the virus IC_{50} 10-fold. The solid line represents the dynamics of T cells and virus when every other dose of both drugs is missed. Wild-type virus is eradicated and drug resistance arises quickly and substantially. The dashed line illustrates the situation when more doses are missed (see the text for description). The treatment can not suppress the wild-type strain and the drug-resistant virus increases very slowly and remains at a low level.

dose of the RT inhibitor and the protease inhibitor. In the second case, the individual misses more doses: e.g., one dose of RT inhibitor is taken and followed by two missed doses; and one dose of protease inhibitor is taken and followed by three missed doses. In both cases, drug resistance is assumed to increase the baseline IC_{50} 10-fold. Figure 11(b) illustrates that the wild-type virus can still be suppressed well if every other dose of both drugs is missed. However, if more doses are missed, the wild-type virus can not be eradicated. Figure 11(c) shows the emergence of the drug-resistant virus during treatment. When every other dose is missed, resistant virus emerges quickly and substantially after the therapy compared with the case of perfect adherence (Figure 10(c)). If more doses are missed, drug resistant virus arises very slowly and the viral load is kept at a low level for several years. The wild-type virus out-competes the resistant virus. This is similar to the case discussed in previous sections, in which drug treatment is less effective and the viral load of the resistant strain

remains at a low level although both strains coexist. The dynamics of uninfected T cells are illustrated in Figure 11(a). When more doses are missed, the number of uninfected T cells converges to a lower steady state level (about 640 cells/μl). It should be mentioned that all the curves undergo frequent oscillations before converging to the steady states. When more doses are missed, the oscillation of the total viral load can be seen clearly from Figure 11(d).

6 Viral and Latent Reservoir Persistence

In many patients, despite prolonged treatment, a low level of virus remains stable. The factors influencing this persistence and their relative contributions have not been elucidated. One possibility is that antiretroviral therapy is not fully suppressive and HIV continues to replicate, particularly in some sites, such as the brain and testes, where drugs have poor penetration (drug sanctuaries). Another possibility is that although therapy is fully suppressive, HIV-1 establishes a state of latent infection in resting memory $CD4^+$ T cells [9], and continued viremia is caused by release of virus due to activation of these latently infected cells. Alternatively, both contribute to the persistence — the latent reservoir releases virus that fuels the ongoing viral replication while ongoing viral replication replenishes the latent reservoir.

The decay characteristics of the latent reservoir in resting memory $CD4^+$ T cells remain controversial. In a cohort initiating therapy during chronic infection, Finzi et al. [27] estimated the half-life of replication-competent HIV-1 in the pool of latently infected cells to be 44 months. Using a similar technique, Ramratnam et al. [81] obtained a half-life of 6 months in a cohort that included both chronically and acutely infected patients. Recently, studies by Strain et al. [98, 99] have suggested that the decay of latently infected cells is decelerating during treatment. They analyzed the decay kinetics of HIV-1 DNA and cell-associated infectivity, and estimated that the latent reservoir had a median half-life of 20 weeks during the first year of therapy. However, the decay rate declined significantly to a median half-life of 70 weeks. Furthermore, the deceleration appeared to continue.

The decelerating decay of the latent reservoir could be explained by the heterogeneity in the activation rate of latently infected cells [66, 98]. Cells specific for frequently encountered antigens may be preferentially activated and quickly cleared from the reservoir during the first year, whereas cells specific for rarely encountered antigens might persist without activation or be activated slowly in subsequent years.

A mathematical model was employed by Muller et al. [66] to study viral production by the activation of latently infected cells after prolonged fully suppressive therapy. They considered latently infected cells, L, which die at rate $\delta_L L$ and are activated at rate αL into the productively infected class, T^*. The model was

$$\frac{d}{dt}L(t) = -(\alpha + \delta_L)L,$$

$$\frac{d}{dt}T^*(t) = \alpha L - \delta T^*, \tag{42}$$

$$\frac{d}{dt}V(t) = pT^* - cV,$$

Since the dynamics of latently infected cells are much slower than those of productively infected cells, the latter can be assumed to be a quasi-steady state. Setting the second equation to zero, one has $T^* = (\alpha/\delta)L(t)$. Similarly, as virus particles turn over much faster than productively infected cells, the virus level will follow the latter as a constant ratio: $V(t) = (p/c)T^*(t)$. Therefore, the virus maintained by the activation of latently infected cells is

$$V_L(t) = \frac{p\alpha}{c\delta}L_0 e^{-(\alpha+\delta_L)t}.$$

Using a continuous range of activation rates instead of a single fixed value, the viral load equation changes to

$$V_L(t) = \frac{pe^{-\delta_L t}}{c\delta}\int_0^{\alpha_{max}} \alpha L_0(\alpha)e^{-\alpha t}d\alpha,$$

where $L_0(\alpha)$ represents the initial distribution of the latently infected cell pool with respect to the activation rate, which has the maximum value α_{max}.

To obtain the total virus level from the initiation of therapy, they also incorporated the initial fast-producing cell populations into the model. From the results given in [73], they obtained the total virus

$$V(t) = V_L(t) + \frac{p}{c}T_0^* e^{-\delta t} + \frac{p_M}{c}M_0 e^{-\delta_M t}, \tag{43}$$

where T_0^* and M_0 represent the initial populations of productively infected cells and long-lived infected cells, respectively. p_M and δ_M are the virus production rate and death rate, respectively, of the latter cells.

The graph of overall virus production according to (43) is a triphasic decline in which the third phase reflects the long tail of the decelerating production by activated latently infected cells [66]. The population of latently infected cells will gradually shift toward cells specific for increasingly rare antigens. Therefore, the progressively decelerating clearance of latently infected cells could be used to explain the observed persistence and long-term dynamics of HIV-1 during effective antiretroviral treatment.

7 Discussion

In this chapter, we have tried to show how mathematical models, in conjunction with clinical data from HIV-1 infected patients, can help understand the

virus infection, emergence of drug resistance, as well as viral persistence during treatment. HIV takes 10 years, on average, to advance from initial infection to full-blown AIDS. The viral load in plasma generally remains unchanged in patients with chronic HIV-1 infection, which suggested that viral replication rate was very slow and the asymptomatic phase of infection was a period of true latency. However, this has turned out not to be correct. When ritonavir, one of the first HIV-1 protease inhibitors, was used to perturb the quasi-steady state of the system, the plasma level of HIV-1 DNA was observed to drop by one to two orders during the first 2 weeks of therapy [40,102]. This result indicated that HIV is not a "slow virus" and it might be replicating very rapidly. Mathematical modeling, combined with data-fitting techniques, has provided quantitative information on virus production and clearance [73, 78]. HIV-1 is cleared from chronically infected patients at a rapid rate, with the half-life of virus in plasma estimated to be 6 hours or less, and more than 10^{10} virus particles, on average, are produced in an untreated patient per day. These findings are crucial in understanding the driving forces underlying disease progression. Rapid viral turnover provides enormous potential for HIV-1 to evolve in response to selective pressures exerted either by the immune system or by drug treatment [104].

More complex models have been developed to include more characteristics of HIV-1 infection and drug treatment. Delay models presented in Section 3 incorporating drug pharmacokinetics and intracellular delay were used to interpret the plasma virus data of infected individuals receiving antiretroviral drugs. Model calculations have shown that estimates of parameters, such as viral clearance rate, c, and death rate of productively infected cells, δ, can be further refined [35, 59, 64, 65], and that viral load decay can display more complex patterns other than the commonly assumed exponential decay [19]. Age-structured models proposed in Section 4 incorporating age-of-infection allow us to model antiretroviral drugs more explicitly since different classes of inhibitors target different stages of the viral life cycle. The effectiveness of an RT inhibitor was proved to depend heavily on the reversion rate, $\eta(\epsilon_{RT})$, at which cells that have not competed reverse transcription (preRT cells) revert back to an uninfected state because of the inhibitor. In fact, the reproductive ratio in the presence of an RT inhibitor is proportional to the factor $e^{-a_1\eta(\epsilon_{RT})}$ (a_1 is the age at which reverse transcription is complete), whereas the reproductive ratio in the presence of an entry inhibitor is proportional to the factor $(1-\epsilon_{EI})$. Thus we can compare the effectiveness of RT and entry inhibitors, although the comparison depends on a_1 and the functional form of $\eta(\epsilon_{RT})$. These findings have significant implications in predicting the effects of drug treatment.

The emergence of drug resistance can occur rapidly and compromise the benefits observed with antiretroviral treatment. Resistance results from the selection of mutations in the genes coding for HIV-1 reverse transcriptase and protease. Many models in the literature studied the likelihood of mutant variants preexisting before treatment as opposed to them being produced during

therapy. We have employed a two-strain model in Section 5 to investigate the mechanistic basis of the emergence of resistant strains in the course of treatment. We showed that notwithstanding the coexistence of both wild-type and mutant strains the viral level of mutant strain is very low compared with wild-type virus before treatment or when the treatment is poorly effective. Drug resistance is more likely to arise for intermediate levels of treatment effectiveness. The relationship between drug resistance and antiretroviral activity confirms that the level of adherence with ARV regimens is a key principle of effective ARV therapy. Suboptimal adherence is associated with a high risk of developing clinically significant HIV resistance. Simulations of the two-strain model have suggested that the wild-type virus and the mutant variants with a low level of resistance (for example, 5-fold drug resistance) will be suppressed well even if a small number of doses are missed. However, the mutant strain with a high level of resistance will flourish quickly after the initiation of treatment. If more doses are missed, resistant strains evolve slowly while the wild-type virus dominates the virus population.

We realize that the patterns of adherence discussed here are not very realistic. In fact, quantifying the real-world patterns of adherence and their influence on the effect of antiretroviral treatment is far from straightforward [25]. Even for the same fraction of the prescribed doses that are taken, different adherence patterns (for example, different block sizes) can induce different treatment outcomes [42,101]. Future work is required to combine pharmacokinetics with careful modeling of more realistic adherence patterns to better predict the antiretroviral responses, particularly the evolution of drug resistance. This might help provide a framework to improve the treatment benefits through structured treatment interruptions (STIs) [16,32,52,70]. The observation that drug-resistant virus declines to a low level after HIV-1-infected patients discontinue antiretroviral treatment for a period leads to the hypothesis that STIs could be served as a new treatment protocol to achieve similar clinical benefits while allowing patients drug holidays [15,58]. Although, more attention needs to be paid when designing treatment schedule through STIs. Some patients undergo substantial viral rebound within days during STIs [26], leading many clinicians to feel that STI is a doomed strategy. A review of different responses to STIs during therapy can be found in [1].

Although models, presented here and in many papers in the literature, predict that virus could be eradicated if the antiretroviral treatment is potent enough, there are many viral and host factors that might hamper treatment success. Even with effective ARV drugs, HIV-1 may replicate in body sites lacking adequate drug exposure to ARV drugs for the selection of drug-resistant mutants [106]. Kepler and Perelson [44] showed that there is a relatively narrow window of drug concentrations that favors the evolution of resistant virus if one considers the body as a single compartment. However, the window of opportunity for the generation of resistance is significantly widened if spatial heterogeneity is taken into account. In some regions, such as the brain and testes, the drug concentration may be low, which enables re-

sistant mutants to be generated more easily. Therefore, the results presented here might underestimate the range of drug concentrations that allow the emergence of drug resistance.

Another challenge to the long-term control or eradication of HIV-1 in infected patients receiving potent antiretroviral treatment is the persistence of HIV reservoirs, including latently infected resting $CD4^+$ T cells, for a prolonged period of time (see reviews in [4, 45, 80]). Estimates of the half-life of virus in resting memory $CD4^+$ T cells suggest that the decay of the latent reservoir is very slow [27, 81, 92, 106]. Moreover, the decay rate seems to decelerate as the treatment continues [98]. Therefore, the natural turn-over of the latent reservoir is not sufficient to achieve eradication of HIV in infected individuals receiving HAART alone.

Purging the latent reservoir would require not only potent antiretroviral treatment, but also the activation of quiescent cells, which causes latent virus to be released. Despite detectable improvements in antiviral potency, treatment intensification could not completely suppress viral replication, as indicated by continuing intermittent viremia (viral blips) in some individuals [82]. A combination of treatment intensification (didanosine and hydroxyurea) and immune stimulation (OKT3 followed by interleukin-2) has provided some evidence that residual replication can further be suppressed, although sustained remission was not achieved [48].

Further understanding of residual replication, viral latency and potential sanctuary sites, will be crucial in the rational design of regimen protocols in the future. Mathematical modeling and analyses of changes in HIV-1 RNA have provided many insights into the biological events underlying virus infection and drug therapy. We believe that modeling will continue to play an important role in helping us understand HIV-1 pathogenesis and guide treatment strategies aimed at sustained HIV-1 remission and/or viral eradication.

Acknowledgement. Portions of this work were performed under the auspices of the U.S. Department of Energy under contract DE-AC52-06NA25396. This work was supported by NSF grant DMS-0314575 and James S. McDonnell Foundation 21^{st} Century Science Initiative (ZF), and NIH grants AI28433 and RR06555 (ASP).

References

1. Bajaria, S., Webb, G., Kirschner, D., 2004. Predicting differential responses to structured treatment interruptions during HAART. Bull. Math. Biol. 66, 1093-1118.
2. Bangsberg, D., et al., 2001. Non-adherence to highly active antiretroviral therapy predicts progression to AIDS. AIDS 15, 1181-1183.
3. Barbour, J., et al., 2002. Evolution of phenotypic drug susceptibility and viral replication capacity during long-term virologic failure of protease inhibitor therapy in human immunodeficiency virus-infected adults. J. Virol. 76, 11104-11112.

4. Blankson, J., Persaud, D., Siliciano, R., 2002. The challenge of viral reservoirs in HIV-1 infection. Annu. Rev. Med. 53, 557-593.
5. Blower, S., Aschenbach, A., Gershengorn, H., Kahn, J., 2001. Predicting the unpredictable: transmission of drug-resistant HIV. Nature Med. 7, 1016-1020.
6. Blower, S., Aschenbach, A., Kahn, J., 2003. Predicting the transmission of drug-resistant HIV: comparing theory with data. Lancet Infect. Dis. 3, 10-11.
7. Bonhoeffer, S., May, R., Shaw, G., Nowak, M., 1997. Virus dynamics and drug therapy. Proc. Natl. Acad. Sci. USA 94, 6971-6976.
8. Bonhoeffer, S., Nowak, M., 1997. Pre-existence and emergence of drug resistance in HIV-1 infection. Proc. R. Soc. Lond. B 264, 631-637.
9. Chun, T., et al., 1997. Quantitation of latent tissue reservoirs and total body load in HIV-1 infection. Nature 387, 183-188.
10. Ciupe, S., Bivort, B., Bortz, D., Nelson, P., 2006. Estimates of kinetic parameters from HIV patient data during primary infection through the eyes of three different models. Math. Biosci. 200, 1-27.
11. Clavel, F., Hance, A., 2004. HIV drug resistance. N. Engl. J. Med. 350, 1023-1035.
12. Clavel, F., Race, E., Mammano, F., 2000. HIV drug resistance and viral fitness. Adv. Pharmacol. 49, 41-66.
13. Coffin, J., 1995. HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. Science 267, 483-489.
14. Collier, A., et al., 1996. Treatment of human immunodeficiency virus infection with saquinavir, zidovudine, and zalcitabine. N. Engl. J. Med. 334, 1011-1017.
15. Deeks, S., 2003. Treatment of antiretroviral-drug-resistant HIV-1 infection. Lancet 362, 2002-2011.
16. Deeks, S., et al., 2003. Persistence of drug-resistant HIV-1 after a structured treatment interruption and its impact on treatment response. AIDS 17, 361-370.
17. Di Mascio, M., et al., 2003. In a subset of subjects on highly active antiretroviral therapy, human immunodeficiency virus type 1 RNA in plasma decays from 50 to <5 copies per milliliter, with a half-life of 6 months. J. Virol. 77, 2271-2275.
18. Dixit, N., Markowitz, M., Ho, D., Perelson, A., 2004. Estimates of intracellular delay and average drug efficacy from viral load data of HIV-infected individuals under antiretroviral therapy. Antiviral Therapy 9, 237-246.
19. Dixit, N., Perelson, A., 2004. Complex patterns of viral load decay under antiretroviral therapy: influence of pharmacokinetics and intracellular delay. J. Theor. Biol. 226, 95-109.
20. Dornadula, G., et al., 1999. Residual HIV-1 RNA in blood plasma of patients taking suppressive highly active antiretroviral therapy. JAMA 282, 1627-1632.
21. Essunger, P., Perelson, A., 1994. Modeling HIV infection of $CD4^+$ T-cell subpopulations. J. Theor. Biol. 170, 367-391.
22. Fauci, A., 2003. HIV and AIDS: 20 years of science. Nature Med. 9, 839-843.
23. Feng, Z., Iannelli, M., Milner, F., 2002. A two-strain Tuberculosis model with age of infection. SIAM J. Appl. Math. 62, 1634-1656.
24. Feng, Z., Rong, L., 2006. The influence of anti-viral drug therapy on the evolution of HIV-1 pathogens, in: Disease Evolution: Models, Concepts, and Data Analyses, Feng, Z., Dieckmann, U. and Levin, S. (eds.). American Mathematical Society, 261-279.
25. Ferguson, N., et al., 2005. Adherence to antiretroviral therapy and its impact on clinical outcome in HIV-infected patients. J. R. Soc. Interface 2, 349-363.

26. Fischer, M., et al., 2003. HIV RNA in plasma rebounds within days during structured treatment interruptions. AIDS 17, 195-199.
27. Finzi, D., et al., 1999. Latent infection of $CD4^+$ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. Nature Med. 5, 512-517.
28. Finzi, D., Siliciano, R., 1998. Viral dynamics in HIV-1 infection. Cell 93, 665-671.
29. Friedland, G., Williams, A., 1999. Attaining higher goals in HIV treatment: the central importance of adherence. AIDS 13(Suppl.1), S61-S72.
30. Gabrielson, J., Weiner, D., 2000. Pharmacokinetic and Pharmacodynamic Data Analysis: Concepts and Applications. Swedish Pharmaceutical Press, Stockholm.
31. Gilchrist, M., Coombs, D., Perelson, A., 2004. Optimizing within-host viral fitness: infected cell lifespan and virion production rate. J. Theor. Biol. 229, 281-288.
32. Gulick, R., 2002. Structured treatment interruption in patients infected with HIV. Drugs 62, 245-253.
33. Gulick, R., 2003. New antiretroviral drugs. Clin. Microbiol. Infect. 9, 186-193.
34. Havlir, D., Eastman, S., Gamst, A., Richman, D., 1996. Nevirapine-resistant human immunodeficiency virus: kinetics of replication and estimated prevalence in untreated patients. J. Virol. 70, 7894-7899.
35. Herz, A., Bonhoeffer, S., Anderson, R., May, R., Nowak, M., 1996. Viral dynamics in vivo: limitations on estimates of intracellular delay and virus decay. Proc. Natl. Acad. Sci. USA 93, 7247-7251.
36. Hlavacek, W., Stilianakis, N., Notermans, D., Danner, S., Perelson, A., 2000. Influence of follicular dendritic cells on decay of HIV during antiretroviral therapy. Proc. Natl. Acad. Sci. USA 97, 10966-10971.
37. Hlavacek, W., Wofsy, C., Perelson, A., 1999. Dissociation of HIV-1 from follicular dendritic cells during HAART: mathematical analysis. Proc. Natl. Acad. Sci. USA 96, 14681-14686.
38. Ho, D., 1996. Viral counts count in HIV infection. Science 272, 1124-1125.
39. Ho, D., Huang, Y., 2002. The HIV-1 vaccine race. Cell 110, 135-138.
40. Ho, D., Neumann, A., Perelson, A., Chen, W., Leonard, J., Markowitz, M., 1995. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. Nature 373, 123-126.
41. Ho, D., Rota, T., Hirsch, M., 1986. Infection of monocyte/macrophages by human T lymphotropic virus type III. J. Clin. Invest. 77, 1712-1715.
42. Huang, Y., Rosenkranz, S., Wu, H., 2003. Modeling HIV dynamics and antiviral response with consideration of time-varying drug exposures, adherence and phenotypic sensitivity. Math. Biosci. 184, 165-186.
43. De Jong, M., et al., 1996. Host-parasite dynamics and outgrowth of virus containing a single K70R amino acid change in reverse transcriptase are responsible for the loss of human immunodeficiency virus type 1 RNA load suppression by zidovudine. Proc. Natl. Acad. Sci. USA 93, 5501-5506.
44. Kepler, T., Perelson, A., 1998. Drug concentration heterogeneity facilitates the evolution of drug resistance. Proc. Natl. Acad. Sci. USA 95, 11514-11519.
45. Kim, H., Perelson, A., 2006. Dynamic characteristics of HIV-1 reservoirs. Curr. Opinion HIV and AIDS 1, 152-156.
46. Kirschner, D., Webb, G., 1996. A model for treatment strategy in the chemotherapy of AIDS. Bull. Math. Biol. 58, 367-390.

47. Kirschner, D., Webb, G., 1997. Understanding drug resistance for montherapy treatment of HIV infection. Bull. Math. Biol. 59, 763-786.
48. Kulkosky, J., et al., 2002. Intensification and stimulation therapy for human immunodeficiency virus type 1 reservoirs in infected persons receiving virally suppressive highly active antiretroviral therapy. J. Infect. Dis. 186, 1403-1411.
49. Larder, B., Darby, G., Richman, D., 1989. HIV with reduced sensitivity to zidovudine (AZT) isolated during prolonged therapy. Science 243, 1731-1734.
50. Larder, B., Kemp, S., 1989. Multiple mutations in HIV-1 reverse transcriptase confer high-level resistance to zidovudine (AZT). Science 246, 1155-1158.
51. Larder, B., 1996. Nucleosides and foscarnet-mechanisms, Antiviral Drug Resistance, D. Richman (ed), John Viley and Sons Ltd.
52. Lori, F., Maserati, R., Foli, A., Seminari, E., Timpone, J., Lisziewicz, J., 2000. Structured treatment interruptions to control HIV-1 infection. Lancet 355, 287-288.
53. Mansky, L., Temin, H., 1995. Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. J. Virol. 69, 5087-5094.
54. Markowitz, M., et al., 2003. A novel antiviral intervention results in more accurate assessment of human immunodeficiency virus type 1 replication dynamics and T-cell decay in vivo. J. Virol. 77, 5037-5038.
55. McLean, A., Nowak, M., 1992. Competition between zidovudine-sensitive and zidovudine-resistant strains of HIV. AIDS 6, 71-79.
56. Mellors, J., et al., 1996. Prognosis in HIV-1 infection predicted by the quantity of virus in plasma. Science 272, 1167-1170.
57. Miller, R., 1971. Nonlinear integral equations. W. A. Benjamin Inc., New York.
58. Miller, V., et al., 2000. Virological and immunological effects of treatment interruptions in HIV-1 infected patients with treatment failure. AIDS 14, 2857-2867.
59. Mittler, J., Markowitz, M., Ho, D., Perelson, A., 1999. Refined estimates for HIV-1 clearance rate and intracellular delay. AIDS 13, 1415-1417.
60. Mittler, J., Sulzer, B., Neumann, A, Perelson, A., 1998. Influence of delayed viral production on viral dynamics in HIV-1 infected patients. Math. Biosci. 152, 143-163.
61. Mugavero, M., Hicks, C., 2004. HIV resistance and the effectiveness of combination antiretroviral treatment. Drug Discovery Today: Therapeutic Strategies 1, 529-535.
62. Nelson, P., Gilchrist, M., Coombs, D., Hyman, J., Perelson, A., 2004. An age-structured model of HIV infection that allows for variations in the production rate of viral particles and the death rate of productively infected cells. Mathematical Biosciences and Engineering 1, 267-288.
63. Nelson, P., Mittler, J., Perelson, A., 2001. Effect of drug efficacy and the eclipse phase of the viral life cycle on estimates of HIV viral dynamic parameters. JAIDS 26, 405-412.
64. Nelson, P., Murray, J., Perelson, A., 2000. A model of HIV-1 pathogenesis that includes an intracellular delay. Math. Biosci. 163, 201-215.
65. Nelson, P., Perelson, A., 2002. Mathematical analysis of delay differential equation models of HIV-1 infection. Math. Biosci. 179, 73-94.
66. Muller, V., Vigueras-Gomez, J. and Bonhoeffer, S., 2002. Decelerating decay of latently infected cells during prolonged therapy for human immunodeficiency virus type 1 infection. J. Virol. 76, 8963-8965.

67. Murray, J., Perelson, A., 2005. Human immunodeficiency virus: Quasi-species and drug resistance. Multiscale Modeling and Simulation 3, 300-311.
68. Nowak, M., Bonhoeffer, S., Shaw, G., May, R., 1997. Anti-viral drug treatment: dynamics of resistance in free virus and infected cell populations. J. Theor. Biol. 184, 203-217.
69. Nowak, M., May, R., 2000. Virus dynamics: Mathematical Principles of Immunology and Virology. Oxford University Press.
70. Ortiz, G., et al., 2001. Structured antiretroviral treatment interruptions in chronically HIV-1-infected subjects. Proc. Natl. Acad. Sci. USA 98, 13288-13293.
71. Palmer, S., et al., 2003. New real-time reverse transcriptase-initiated PCR assay with single-copy sensitivity for human immunodeficiency virus type 1 RNA in plasma. J. Clin. Microbiol. 41, 4531-4536.
72. Perelson, A., 2002. Modelling viral and immune system dynamics. Nature Rev. Immunol. 2, 28-36.
73. Perelson, A., et al., 1997. Decay characteristics of HIV-1-infected compartments during combination therapy. Nature 387, 188-191.
74. Perelson, A., Essunger, P., Ho, D., 1997. Dynamics of HIV-1 and $CD4^+$ lymphocytes in vivo. AIDS 11(Suppl. A), S17-S24.
75. Perelson, A., Kirschner, D., De Boer, R., 1993. Dynamics of HIV infection of $CD4^+$ T cells. Math. Biosci. 114, 81-125.
76. Perelson, A., Nelson, P., 1999. Mathematical analysis of HIV-1 dynamics in vivo. SIAM Rev. 41, 3-44.
77. Perelson, A., Nelson, P., 2002. Modeling viral infections, Proceedings of Symposia in Applied Mathematics 59, 139-172.
78. Perelson, A., Neumann, A., Markowitz, M., Leonard, J., Ho, D., 1996. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. Science 271, 1582-1586.
79. Piatak, M., et al., 1993. High levels of HIV-1 in plasma during all stages of infection determined by competitive PCR. Science 259, 1749-1754.
80. Pierson, T., McArthur, J., Siliciano, R., 2000. Reservoirs for HIV-1: mechanisms for viral persistence in the presence of antiviral immune responses and antiretroviral therapy. Annu. Rev. Immunol. 18, 665-708.
81. Ramratnam, B., et al., 2000. The decay of the latent reservoir of replication-competent HIV-1 is inversely correlated with the extent of residual viral replication during prolonged anti-retroviral therapy. Nature Med. 6, 82-85.
82. Ramratnam, B., et al., 2004. Intensification of antiretroviral therapy accelerates the decay of the HIV-1 latent reservoir and decreases, but does not eliminate, ongoing virus replication. JAIDS 35, 33-37.
83. Ribeiro, R., Bonhoeffer, S., 2000. Production of resistant HIV mutants during antiretroviral therapy. Proc. Natl. Acad. Sci. USA 97, 7681-7686.
84. Ribeiro, R., Bonhoeffer, S., Nowak, M., 1998. The frequency of resistant mutant virus before antiviral therapy. AIDS 12, 461-465.
85. Richman, D., 1992. Selection of zidovudine-resistant variants of human immunodeficiency virus by therapy. Current Topics in Microbiology and Immunology 176, 131-142.
86. Richman, D., 1996. The implications of drug resistance for strategies of combination antiviral chemotherapy. Antiviral Res. 29, 31-33.
87. Richman, D., et al., 1994. Nevirapine resistant mutations of human immunodeficiency virus type 1 selected during therapy. J. Virol. 68, 1660-1666.

88. Rong, L., Feng, Z., Perelson, A., 2006. Mathematical analysis of age-structured HIV-1 dynamics with combination antiretroviral therapy, submitted.
89. Rong, L., Feng, Z., Perelson, A., 2006. Emergence of HIV-1 drug resistance during antiretroviral treatment, submitted.
90. Sethi, A., et al., 2003. Association between adherence to antiretroviral therapy and human immunodeficiency virus drug resistance. Clin. Infect. Dis. 37, 1112-1118.
91. Shiri, T., Garira, W., Musekwa, S., 2005. A two-strain HIV-1 mathematical model to assess the effects of chemotherapy on disease parameters. Mathematical Biosciences and Engineering 2, 811-832.
92. Siliciano, J., et al., 2003. Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting $CD4^+$ T cells. Nature Med. 9, 727-728.
93. Snedecor, S., 2003. Comparison of three kinetic models of HIV-1 infection: implications for optimization of treatment. J. Theor. Biol. 221, 519-541.
94. Stafford, M., Corey, L., Cao, Y., Daar, E., Ho, D., Perelson, A., 2000. Modeling plasma virus concentration during primary HIV infection. J. Theor. Biol. 203, 285-301.
95. St Clair, M., et al., 1991. Resistance to ddI and sensitivity to AZT induced by a mutation in HIV-1 reverse transcriptase. Science 253, 1557-1559.
96. Stevenson, M., 2003. HIV-1 pathogenesis. Nature Med. 9, 853-860.
97. Stilianakis, N., Boucher, C., De Jong, M., Van Leeuwen, R., Schuurman, R., De Boer, R., 1997. Clinical data sets of human immunodeficiency virus type 1 reverse transcriptase-resistant mutants explained by a mathematical model. J. Virol. 71, 161-168.
98. Strain, M., et al., 2003. Heterogeneous clearance rates of long-lived lymphocytes infected with HIV: intrinsic stability predicts lifelong persistence. Proc. Natl. Acad. Sci. USA 100, 4819-4824.
99. Strain, M., et al., 2005. Effect of treatment, during primary infection, on establishment and clearance of cellular reservoirs of HIV-1. J. Infect. Dis. 191, 1410-1418.
100. Tesoriero, J., et al., 2003. Stability of adherence to highly active antiretroviral therapy over time among clients enrolled in the treatment adherence demonstration project. JAIDS 33, 484-493.
101. Wahl, L., Nowak, M., 2000. Adherence and drug resistance: predictions for therapy outcome. Proc. R. Soc. Lond. B 267, 835-843.
102. Wei, X., et al., 1995. Viral dynamics in human-immunodeficiency-virus type-1 infection. Nature 373, 117-122.
103. Wodarz, D., Lloyd, A., 2004. Immune responses and the emergence of drug-resistant virus strains in vivo. Proc. R. Soc. Lond. B 271, 1101-1109.
104. Wodarz, D., Nowak, M., 1998. Mathematical models of virus dynamics and resistance. J. HIV Therapy 3, 36-41.
105. Wu, H., et al., 2005. Modeling long-term HIV dynamics and antiretroviral response. JAIDS 39, 272-283.
106. Zhang, L., et al., 1999. Quantifying residual HIV-1 replication in patients receiving combination antiretroviral therapy. N. Engl. J. Med. 340, 1605-1613.
107. Zhang, Z., et al., 1999. Sexual transmission and propagation of SIV and HIV in resting and activated $CD4^+$ T cells. Science 286, 1353-1357.

Overcoming the Key Challenges in De Novo Protein Design: Enhancing Computational Efficiency and Incorporating True Backbone Flexibility

Christodoulos A. Floudas[1], Ho Ki Fung[1], Dimitrios Morikis[2], Martin S. Taylor[3], and Li Zhang[4]

[1] Department of Chemical Engineering, Princeton University, Princeton, NJ 08544-5263 floudas@titan.princeton.edu
[2] Department of Bioengineering, University of California, Riverside, CA 92521
[3] Johns Hopkins University School of Medicine, Baltimore, MD 21205
[4] Department of Chemistry, University of California, Riverside, CA 92521

Key words: De novo protein design, true protein backbone flexibility, weighted average forcefield, bin variables, fold specificity stage.

1 Introduction

De novo protein design is initiated with a postulated or known flexible three-dimensional protein structure and aims at identifying amino acid sequences compatible with such a structure. The problem was first denoted as the "inverse folding problem" [4, 5] since protein design has intimate links to the well-known protein folding problem [6]. While the protein folding problem aims at determining the single structure for a sequence, the de novo protein design problem exhibits a high level of degeneracy; that is, a large number of sequences are always found to share a common fold, although the sequences will vary with respect to properties such as activity and stability.

Traditionally, protein design was performed using experimental techniques like rational design, mutagenesis, and directed evolution [2]. Though capable of producing good results, they all entail the major drawback of being able to screen only a highly restricted number of mutants. It was estimated that the maximum size of amino acid sequence search space these experimental approaches can handle is around $10^3 - 10^6$ [13]. In comparison, the number of sequences that computational de novo design methods can search through is significantly larger [106]. For instance, [42] reported their redesign of the

74 core residues of the catalytic antibody (PDB code: 1HKL), which corresponded to a rotamer search space of 4.7×10^{128}, an unimaginable size for experimentalists.

The search method that [42] used in their work is called dead-end-elimination (DEE) [15], which is arguably the most common deterministic type framework so far for computational protein design. It operates on the systematic use of energy comparison equations that eliminate rotamers that cannot be compatible with the global energy minimum conformation. Its popularity originates from the consistent convergence to the globally optimal solution for a mutation set of daunting complexity. In addition to DEE, there is another deterministic method for de novo protein design called the self-consistent mean field (SCMF) theory [74]. SCMF tests an ensemble of amino acid/rotamer combinations at each position of a fixed template, with each rotamer in the ensemble given the same Boltzmann probability. The rotamer Boltzmann probabilities for all other positions are then computed to obtain a weighted average energy, which is used to recalculate the Boltzmann probability for each rotamer at each position. Thus, the process is iterative and will terminate when the Boltzmann probability converges to a certain value. The main disadvantage of SCMF is that though deterministic in nature, it does not guarantee to yield the global minimum in energy [74]. Besides the deterministic category, there is a family of design methods that is stochastic in nature; the Monte-Carlo approach and genetic algorithms belong to this family. In Monte-Carlo methods, a mutation is performed at a certain position in the sequence and the Boltzmann probability calculated from the energies before and after the mutation, as well as temperature is compared to a random number. The mutation is allowed if the Boltzmann probability is higher than the random number, and rejected otherwise. The Baker group's protein design computer program, RosettaDesign, was based on the Monte-Carlo optimization algorithm [45, 46, 48]. On the other hand, genetic algorithms operate by generating a multitude of random amino acid sequences and exchanging them for a fixed template. Sequences with low energies form hybrids with other sequences while those with high energies are eliminated in an iterative process which only terminates when a converged solution is attained [77]. [14] applied a two-stage combination of Monte Carlo and genetic algorithms to design the hydrophobic core of protein 434cro. Being stochastic in nature, Monte-Carlo methods and genetic algorithms share the common disadvantage of lacking consistency in the final solution. It should be noted that instead of searching for the sequence with the lowest energy, de novo protein design can also be performed using site-specific probabilities for amino acids or rotamers. The statistical computationally assisted design strategy (*scads*) employed by [44, 70, 71] was based on an optimization model which maximizes an objective function of entropy based on these site-specific probabilities subject to certain constraints, which can be derived from the energy minimization formulation.

Initial computational protein design attempts were all focussed on the core, as protein folding is believed to be primarily driven by hydrophobic collapse and hence a well-folded stable core is required for de novo designed proteins [73]. There were quite a number of reported successes on hydrophobic core redesign [7, 11, 55]. As time went by the design scope also encompassed the surface and boundary regions, with numerous examples of experimentally validated successes. They include the full sequence design of a zinc finger domain [63] and the WW motif [60], and the designs of the active site of α-lytic protease [78], the catalytic site of superoxide dismutase [52], the B1 domain of protein G, the lambda repressor, the sperm whale myoglobin [12], a protein interface that prevents fibril formation [103], novel receptors and sensor proteins [54], periplasmic binding proteins [96], biologically active enzymes [95], a calcium binding protein [58], ferritin-like proteins with novel hydrophobic cavities [97], novel peptide inhibitors [1, 2, 100], and other therapeutic proteins [102]. De novo design was successfully employed for enhancing protein binding selectivity [99, 104], modulating protein-protein interaction specificity [51, 72], promoting stability of proteins [50, 65, 101], altering protein folding kinetics [49, 68, 69, 98], conferring novel binding sites or properties onto the template [56, 57], and locking proteins into certain useful conformations [59, 66].

Despite all these achievements, two major long-standing challenges are limiting the full application of de novo protein design: (1) the NP-hard nature of the problem in terms of computational efficiency [83, 85], and (2) the incorporation of protein backbone flexibility into the design model. The first challenge means required computational time scales exponentially with the number of design positions on the template, which makes the full sequence design of proteins of practical size (i.e., $100 - 300$ residues) extremely difficult. It commands a high efficiency level for any de novo protein design algorithm, especilly those of the deterministic type. The second issue is more subtle, and it can be dated back to the time of early computational protein design efforts which were mostly based on the rigid backbone premise, with the three-dimensional coordinates of all atoms fixed on the template [10]. This assumption was later proved to be dubious, as [61]demonstrated that backbone flexibility can allow residues that would not have been permissable had a rigid template been considered. [79] also provided convincing evidence for the superiority of backbone flexibility in their successful design of metal binding sites in the Gβ1 protein. They noted that elements of their design would be overlooked using a single, averaged template because the required conformations would have been deemed infeasible. On another occasion, backbone flexibility was also found to be of fundamental importance to obtaining stable folds [67]. So far different approaches have been proposed on how to incorporate backbone flexibility into de novo design models. In one approach atomic radii in calculating the van der Waals potential were scaled down, typically by five to ten per cent, to allow for small overlap between atoms during backbone movements [11, 47]. However, this method has the intrinsic

disadvantages of overestimation of attractive forces and possible hydrophobic core overpacking. In another approach either a fixed set of rotamers was considered or some super-secondary-structure parameters were changed to adjust the relative orientation and distances between secondary structures [64]. [14] also constructed ensemble of random structures from the template and solved each structure in the ensemble for the low energy sequence using the fixed backbone assumption. These two similar methods only take into account either a chosen subset or a random subset of all possible conformations of the protein template. Lately, backbone flexibility was treated by the Baker's group by iterating between the sequence space and the structure space until the algorithm converged to a final solution [46, 86]. Nevertheless, their approach only addressed backbone flexibility indirectly by movements along the structure space during iteration. To conclude, none of the existing approaches exactly observes true backbone flexibility, which is defined by lower and upper bounds on the distances between coordinates (e.g., alpha carbons or side chain centroids) on the template, as well as the backbone dihedral angles [88], except for a recently published one [1, 2, 85, 89, 106, 115].

In this chapter, our original two-stage framework for computational protein design proposed by [1, 2] will be outlined. It will then be followed by the presentation of our efforts to overcome the two key challenges in de novo design, namely algorithmic efficiency and template flexibility incorporation, by formulating new optimization models that are both computationally less expensive and compatible with true backbone flexibility. The chapter will be concluded by a demonstration about how the new models were applied to the design of: (i) compstatin, a synthetic 13-residue cyclic peptide that binds to complement protein 3 (C3) and inhibits the activation of the complement system (part of innate immunity); (ii) human β defensin-2, a 41-residue cationic peptide in the immune system; and (iii) a potential peptide-drug candidate derived from the C-terminal sequence of the C3a fragment of C3. Peptides (i) and (iii) have the potential to become therapeutics against inappropriate complement system activation, which is present in nearly every autoimmune disease and other pathological situations.

2 The Original De Novo Protein Design Framework

Our original de novo protein design methodology was first proposed by [1, 2]. It is a two-stage framework that not only selects and ranks amino acid sequences for a particular fold but also validates the specificity to the fold for these selected sequences. The sequence selection phase relies on a novel integer linear programming (ILP) model with several important constraint modifications that improve the tractability of the problem and enhance its deterministic convergence to the global minimum. In addition, a rank-ordered list of low lying energy sequences are identified along with the global minimum energy sequence. Once such a subset of sequences have been identified, the

fold validation stage is employed to verify the stabilities and specificities of the designed sequences through a deterministic global optimization approach that allows for backbone flexibility. The selection of the best designed sequences is based on rigorous quantification of energy based probabilities. In the following, we will discuss the two stages in detail.

2.1 Stage One: *In Silico* Sequence Selection

The original form of the sequence selection optimization model employed by [1, 2] takes the form:

$$\begin{aligned} \min_{y_i^j, y_k^l} \ & \sum_{i=1}^{n} \sum_{j=1}^{m_i} \sum_{k=i+1}^{n} \sum_{l=1}^{m_k} E_{ik}^{jl}(x_i, x_k) y_i^j y_k^l \\ \text{subject to} \quad & \sum_{j=1}^{m_i} y_i^j \ = \ 1 \ \ \forall \ \ i \\ & y_i^j \ , \ y_k^l \ = \ 0 - 1 \ \ \forall \ \ i, j, k, l \end{aligned} \tag{1}$$

Note that this formulation resembles the quadratic assignment problem (QAP). It differs, however, in the set of constraints. Set $i = 1, \ldots, n$ defines the number of residue positions along the backbone. At each position i there can be a set of mutations represented by $j\{i\} = 1, \ldots, m_i$, where, for the general case $m_i = 20 \forall i$. The equivalent sets $k \equiv i$ and $l \equiv j$ are defined, and $k > i$ is required to represent all unique pairwise interactions. Binary variables y_i^j and y_k^l are introduced to indicate the possible mutations at a given position. Specifically, variable y_i^j will be one if position i is occupied by amino acid j, and zero otherwise. Similarly, variable y_k^l will assume the value of one if position k is taken by amino acid l, and the value of zero otherwise. The composition constraints in the formulation require that there is exactly one type of amino acid at each position.

Energy parameter E_{ik}^{jl} indicates the pairwise interaction between the amino acid j at position i and the amino acid l at position k. It only contributes to the objective function of total energy of the system if both y_i^j and y_k^l are one, i.e., positions i and k are taken by amino acid j and l respectively. These energy parameters were empirically derived based on solving a linear programming parameter estimation problem subject to constraints which were in turn constructed by requiring the energies of a large number of low-energy decoys to be larger than the corresponding native protein conformation for each member of a set of proteins [19]. The resulting potential, which contains $1,680$ energy parameters for different amino acid pairs and distance bins, was shown to rank the native fold as the lowest in energy in more proteins tested than other forcefields and also yield higher Z-score [17–19]. In the research work outlined in this paper, the high resolution C^{α}-C^{α} forcefield and the high resolution centroid-centroid forcefield were employed for solving the sequence selection models [87]. Being a significant upgrade from the forcefield developed

by [19], these forcefields were derived from a large training set of 1,250 proteins based on an average of 800 high resolution decoys for each protein. The high quality decoys, which possessed close structural resemblance to the native conformations, were in turn constructed using a novel generation method [87].

In the original formulation (1), bilinear terms in the objective function of the original formulation were linearized using an equivalent representation [22]:

$$\begin{aligned} \min_{y_i^j, y_k^l} & \sum_{i=1}^{n} \sum_{j=1}^{m_i} \sum_{k=i+1}^{n} \sum_{l=1}^{m_k} E_{ik}^{jl}(x_i, x_k) w_{ik}^{jl} \\ \text{subject to} \quad & \sum_{j=1}^{m_i} y_i^j = 1 \;\forall\; i \\ & y_i^j + y_k^l - 1 \leq w_{ik}^{jl} \leq y_i^j \;\forall\; i,j,k,l \\ & 0 \leq w_{ik}^{jl} \leq y_k^l \;\forall\; i,j,k,l \\ & y_i^j \;,\; y_k^l = 0-1 \;\forall\; i,j,k,l \end{aligned} \tag{F1}$$

As indicated in the formula above, the nonconvex bilinear terms were transformed into a new set of linear variables, w_{ik}^{jl}, with the addition of the four sets of constraints to reproduce the characteristics of the original formulation. For example, for a given i, j, k, l combination, the four constraints require w_{ik}^{jl} to be zero when either y_i^j or y_k^l is equal (or when both are equal to zero). If both y_i^j and y_k^l are equal to one then w_{ik}^{jl} is also enforced to be one. The solution of the integer linear programming problem (ILP) can be accomplished rigorously using branch and bound techniques [20, 22] making convergence to the global minimum energy sequence consistent and reliable.

Performance of the branch and bound algorithm can be significantly enhanced through the use of the reformulation linearization techniques (RLTs). The basic strategy is to multiply appropriate constraints by bounded nonnegative factors (such as the reformulated variables) and introduce the products of the original variables by new variables in order to derive higher-dimensional lower and tighter bounding linear programming (LP) relaxations for the original problem [21]. In this case RLTs are introduced by multiplying the composition constraints by the binary variables y_k^l to produce:

$$\begin{aligned} y_k^l \sum_{j=1}^{m_i} y_i^j &= y_k^l \;\forall\; i,k,l \qquad \text{or:} \\ \sum_{j=1}^{m_i} w_{ik}^{jl} &= y_k^l \;\forall\; i,k,l \end{aligned} \tag{2}$$

In summary, the whole model for *in silico* sequence selection in [1, 2]'s de novo protein design framework is:

$$\begin{aligned} \min_{y_i^j, y_k^l} & \sum_{i=1}^{n} \sum_{j=1}^{m_i} \sum_{k=i+1}^{n} \sum_{l=1}^{m_k} E_{ik}^{jl}(x_i, x_k) w_{ik}^{jl} \\ \text{subject to} \quad & \sum_{j=1}^{m_i} y_i^j = 1 \; \forall \; i \\ & y_i^j + y_k^l - 1 \leq w_{ik}^{jl} \leq y_i^j \; \forall \; i, j, k, l \\ & 0 \leq w_{ik}^{jl} \leq y_k^l \; \forall \; i, j, k, l \\ & \sum_{j=1}^{m_i} w_{ik}^{jl} = y_k^l \; \forall \; i, k, l \\ & y_i^j \; , \; y_k^l = 0 - 1 \; \forall \; i, j, k, l \end{aligned} \tag{F2}$$

Incorporation of True Backbone Flexibility in the Sequence Selection Stage

It should be highlighted that [1,2] incorporated true protein backbone flexibility explicitly into the *in silico* sequence selection model. Rather than being a continuous function, the dependence on C^{α}-C^{α} or centroid-centroid distance of the energy parameter $E_{ik}^{jl}(x_i, x_k)$ in the objective function of (F2) is discretized into bins. In both the force field of [87] and that of [19], the distance bins are classified in a way as shown in Table 1. Bin 1 is for any distance between $l_{beg}(1) = 3.0\rho A$ and $l_{end}(1) = 4.0\rho A$, bin 2 for any distance between $l_{beg}(2) = 4.0\rho A$ and $l_{end}(2) = 5.0\rho A$, and so forth. Given a certain pair of amino acids, any distance between the alpha carbon of the two amino acids falling into the range bounded by the upper bound $l_{end}(n)$ and lower bound $l_{beg}(n)$ will belong to the same distance bin n, thus giving the same energy value. In this way the energy function is insensitive to backbone motion up to a certain degree. With the bin sizes varying between 0.5 and $1\rho A$, this discretization of the force field allows backbone flexibility of the same order of magnitude.

Table 1. Distance bin classification in the high resolution force field developed by [87] for the sequence selection of de novo protein design.

Distance Bin d	$l_{beg}(d)[\rho A]$	$l_{mid}(d)[\rho A]$	$l_{end}(d)[\rho A]$
Bin 1	3.0	3.5	4.0
Bin 2	4.0	4.5	5.0
Bin 3	5.0	5.25	5.5
Bin 4	5.5	5.75	6.0
Bin 5	6.0	6.25	6.5
Bin 6	6.5	6.75	7.0
Bin 7	7.0	7.5	8.0
Bin 8	8.0	8.5	9.0

2.2 Stage Two: Fold Specificity

Once a set of low lying energy sequences have been identified via the sequence selection procedure, the fold stability and specificity validation stage is used to identify the most optimal sequences according to a rigorous quantification of conformational probabilities. The foundation of the approach is grounded on the development of conformational ensembles for the selected sequences under two sets of conditions. In the first circumstance the structure is constrained to vary, with some imposed fluctuations, around the template structure. In the second condition, a free folding calculation is performed for which only a limited number of restraints (e.g., the disulfide bridge constraints) are likely to be incorporated and with the underlying template structure not being enforced. In terms of practical considerations, the distance constraints introduced for the template constrained simulation can be based on the structural boundaries defined by the NMR ensemble, or simply by allowing some deviation from a subset of distances provided by the structural template, and hence they allow for a flexible template on the backbone.

The formulations for the folding calculations are reminiscent of structure prediction problems in protein folding [23]. In particular, a novel constrained global optimization problem first introduced for structure prediction using NMR data [24], and later employed in a generic framework for the structure prediction of proteins [25] is employed. The global minimization of a detailed atomistic energy forcefield E_{ff} is performed over the set of independent dihedral angles, ϕ, which can be used to describe any possible configuration of the system. The bounds on these variables are enforced by simple box constraints. Finally, a set of distance constraints, E_l^{dis} $l = 1, \ldots, N$, which are nonconvex in the internal coordinate system, can be used to constrain the system. The formulation is represented by the following set of equations:

$$\begin{aligned} \min_{\phi} \quad & E_{ff} \\ \text{subject to} \quad & E_j^{dis}(\phi) \;\leq\; E_j^{ref} \quad j = 1, \ldots, N \\ & \phi_i^L \;\leq\; \phi_i \;\leq\; \phi_i^U \;\; i = 1, \ldots, N_\phi \end{aligned} \tag{3}$$

Here, $i = 1, \ldots, N_\phi$ corresponds to the set of dihedral angles, ϕ_i, with ϕ_i^L and ϕ_i^U representing lower and upper bounds on these dihedral angles. In general, the lower and upper bounds for these variables are set to -π and π. E_j^{ref} are reference parameters for the distance constraints, which assume the form of typical square well potential for both upper and lower distance violations. The set of constraints are completely general, and can represent the full combination of distance constraints or smaller subsets of the defined restraints. The forcefield energy function, E_{ff} can take on a number of forms, although the work performed by [1, 2] employed the ECEPP/3 model [26].

Formulation (3) represents a general nonconvex constrained global optimization problem, a class of problems for which several methods have been

developed. In the work presented by [1,2], the formulations were solved via the αBB deterministic global optimization approach, a branch and bound method applicable to the identification of the global minimum of nonlinear optimization problems with twice-differentiable functions [23,24,27–30,75]. In the αBB approach, a converging sequence of upper and lower bounds is generated. The upper bounds on the global minimum are obtained by local minimizations of the original nonconvex problem, while the lower bounds belong to the set of solutions of the convex lower bounding problems that are constructed by augmenting the objective and constraint functions through the addition of separable quadratic terms.

As the final step in the second stage of the de novo protein design framework, the relative probability for template specificity, p_{temp}, is found by summing the statistical weights for those conformers from the free folding simulation that resemble the template structure (denote as set *temp*), and dividing this sum by the summation of statistical weights for all conformers from the free folding simulation (denote as set *total*):

$$p_{temp} = \frac{\sum_{i \in temp} \exp[-\beta E_i]}{\sum_{i \in total} \exp[-\beta E_i]} \tag{4}$$

where $\exp[-\beta E_i]$ is the statistical weight for conformer i.

Incorporation of True Backbone Flexibility in the Fold Specificity Stage

In the second stage of the de novo design framework, backbone flexibility is explicitly included by treating C^{α}-C^{α} distances and dihedral angles as bounded continuous variables in the template-constrained structure prediction calculations. In particular, [1, 2] chose to set the lower and upper bounds to be $\pm 10\%$ of those in the template for the C^{α}-C^{α} distances and $\pm 10^{o}$ around the template for the phi and psi angles.

3 Improvement of the De Novo Protein Design Framework

We have made a few endeavors in improving the original framework proposed by [1, 2] in response to the two key challenges in de novo protein design, which are: (1) increasing algorithmic efficiency of the NP-hard de novo design model, and (2) incorporating backbone flexibility [85, 89]. These efforts are summarized in this section.

3.1 Enhancing Computational Efficiency

Our efforts in this area were focused on both the *in silico* sequence selection stage and the fold specificity stage.

In Silico Sequence Selection Stage

Algorithmic improvement of the sequence selection model (F2) was initiated by noting that the formulation is of $O(n^2)$, which means the number of linear constraints scales with n^2, where n is the number of binary variables [85]. For instance, if all 20 amino acids are considered for each position in a 40-residue peptide, then n equals $40 \times 20 = 800$. The number of variables w_{ik}^{jl} will be $400 \times 820 = 328,000$, and hence number of linear constraints is simply $4 \times 328,000 = 1,312,000$, which is $\sim |n|^2$. Performance of the model was compared to three other $O(n)$ formulations, which were derived for QAPs and proved to be equivalent to (F2) [85]. These $O(n)$ formulations are:

$$\begin{aligned}
&\min_{y_i^j, \zeta_i^j} && \textstyle\sum_{i=1}^{n} \sum_{j=1}^{m_i} D_i^{j^-} y_i^j \; + \; \zeta_i^j \\
&\textit{subject to} \\
&\zeta_i^j \geq \textstyle\sum_{k=i+1}^{n} \sum_{l=1}^{m_k} E_{ik}^{jl}(x_i, x_k) y_k^l \; - \; D_i^{j^-} y_i^j \; - \; D_i^{j^+}(1 \; - \; y_i^j) \;\; \forall i, j \\
&\textstyle\sum_{j=1}^{m_i} y_i^j = 1 \;\; \forall i \\
&\zeta_i^j \; \geq \; 0 \;\; \forall i, j \qquad \text{(F3)} \\
&y_i^j, y_k^l = 0 - 1 \;\; \forall i, j, k, l \\
&D_i^{j-} \; = - \textstyle\sum_{k=i+1}^{n} \max_{1 \leq l \leq m_k} |\min\{0, E_{ik}^{jl}(x_i, x_k)\}| \\
&D_i^{j+} \; = \textstyle\sum_{k=i+1}^{n} \max_{1 \leq l \leq m_k} \max\{0, E_{ik}^{jl}(x_i, x_k)\}
\end{aligned}$$

and

$$\begin{aligned}
&\min_{y_i^j, \zeta_i^j} && \textstyle\sum_{i=1}^{n} \sum_{j=1}^{m_i} \left(\sum_{k=i+1}^{n} \sum_{l=1}^{m_k} E_{ik}^{jl}(x_i, x_k) y_k^l - B_i^{j^+}(1 - y_i^j) + \zeta_i^j \right) \\
&\textit{subject to} \\
&\zeta_i^j \; \geq \; - \textstyle\sum_{k=i+1}^{n} \sum_{l=1}^{m_k} E_{ik}^{jl}(x_i, x_k) y_k^l \; + \; B_i^{j^-} y_i^j \; + \; B_i^{j^+}(1 \; - \; y_i^j) \;\; \forall i, j \\
&\textstyle\sum_{j=1}^{m_i} y_i^j = 1 \;\; \forall i \\
&\zeta_i^j \; \geq \; 0 \;\; \forall i, j \qquad \text{(F4)} \\
&y_i^j, y_k^l = 0 - 1 \;\; \forall i, j, k, l \\
&B_i^{j-} \; = - \textstyle\sum_{k=i+1}^{n} \max_{1 \leq l \leq m_k} |\min\{0, E_{ik}^{jl}(x_i, x_k)\}| \\
&B_i^{j+} \; = \textstyle\sum_{k=i+1}^{n} \max_{1 \leq l \leq m_k} \max\{0, E_{ik}^{jl}(x_i, x_k)\}
\end{aligned}$$

They were proposed by [81,82] (and are modifications of the initial technique by [105]) for a general linearly constrained quadratic 0–1 programming problem. Both formulations (F3) and (F4) are promising in terms of computational efficiency because compared to the original binary quadratic integer problem of n variables, the number of auxiliary linear constraints is reduced to n,

whereas the number of new continuous variables ζ_i^j introduced is n versus the n^2 binary variables w_{ik}^{jl} in (F2).

In the third O(n) model the number of new variables is increased to $2n$, and some of the upper bounding linear constraints are kept [80, 81]:

$$\begin{aligned}
\min_{s_i^j, y_i^j, \zeta_i^j} \quad & \textstyle\sum_{i=1}^{n} \sum_{j=1}^{m_i} s_i^j \; - \; M_i^{j-} y_i^j \\
\text{subject to } & \left[\textstyle\sum_{k=i+1}^{n} \sum_{l=1}^{m_k} E_{ik}^{jl}(x_i, x_k) y_k^l\right] \; - \; \zeta_i^j \; - \; s_i^j \; + \; M_i^{j-} \; \leq \; 0 \;\; \forall \;\; i, j \\
& \zeta_i^j \; \leq \; M_i^j (1 \; - \; y_i^j) \;\; \forall \;\; i, j \\
& \textstyle\sum_{j=1}^{m_i} y_i^j \; = \; 1 \; \forall \; i \qquad \text{(F5)} \\
& M_i^{j-} \; = \textstyle\sum_{k=i+1}^{n} \max_{1 \leq l \leq m_k} |\min\{0, E_{ik}^{jl}(x_i, x_k)\}| \\
& M_i^{j+} \; = \textstyle\sum_{k=i+1}^{n} \max_{1 \leq l \leq m_k} \max\{0, E_{ik}^{jl}(x_i, x_k)\} \\
& M_i^j \; = \; M_i^{j-} \; + \; M_i^{j+} \;\; \forall \;\; i, j \\
& \zeta_i^j \; \geq \; 0, s_i^j \; \geq \; 0, y_i^j \; = \; 0 - 1 \;\; \forall \;\; i, j
\end{aligned}$$

In addition to these O(n) equations, new equivalent O(n^2) models were formulated using different combinations of three novel elements for algorithmic improvement: (i) conversion of the equality RLT constraints into inequality constraints, (ii) addition of triangle inequalities, and (iii) execution of a preprocessing step using one iteration of the Dead-End Elimination theorem before solving the sequence selection optimization model [85]. Since the first two components lead to superfluous equations which do not affect the feasibility region of the original problem ((F1) or (F2)), and preprocessing simplifies the problem by eliminating the binary variables that are unnecessary, implementation of any combination of the three certainly does not affect the objective function value. The characteristics of the three elements are:

- RLT with inequalities
  The rationale for this technique is to relax the RLT constraints, which are supposed to be crucial in speeding up the branch and bound algorithm in the original $O(n^2)$ formulation that [2] proposed, by changing the equality in the equation to "less than or equal to," making the RLT equation look like:

$$\sum_{j=1}^{m_i} w_{ik}^{jl} \; \leq \; y_k^l \;\; \forall \;\; i, k, l \qquad (5)$$

  Considering that equality is equivalent to both "larger than or equal to" and "less than or equal to," implementing only the latter will probably lead to a problem that is easier and faster to solve.
- Addition of triangle inequalities
  Valid triangle inequalities as shown below were added to (F2) in an attempt to hasten convergence to the global energy minimum solution. Similar to

RLTs, they are supposed to enhance the algorithm by providing tighter lower bounds to the original problem.

$$(y_i^j - y_k^l)(y_i^j - y_m^p) \geq 0 \ \forall i < k < m, j, l, p \quad \text{or:}$$
$$y_i^j - w_{im}^{jp} - w_{ik}^{jl} + w_{km}^{lp} \geq 0 \ \forall i < k < m, j, l, p \tag{6}$$

$$2 - (y_i^j - y_k^l)^2 - (y_i^j - y_m^p)^2 - (y_k^l - y_m^p)^2 \geq 0$$
$$\forall i < k < m, j, l, p \quad \text{or:}$$
$$w_{ik}^{jl} + w_{im}^{jp} + w_{km}^{lp} - y_i^j - y_k^l - y_m^p + 1 \geq 0 \tag{7}$$
$$\forall i < k < m, j, l, p$$

where linear binary variables y_i^j and w_{ik}^{jl} were defined in the same way as before, whereas indices m and p were aliases of position sets i and k and amino acid sets j and l respectively.

An additional subtlety to consider in applying triangle inequalities is the total number of inequalities to impose, which supposedly has an optimal value giving the best computational efficiency [85]. In view of this, triangle inequalities were to be applied only if the sum of the pairwise energy triplets, namely $S_{ikm}^{jlp} = E_{ik}^{jl} + E_{im}^{jp} + E_{km}^{lp}$ was less than a certain cutoff value. Both cases of no cutoff and cutoff value of -40 were tried in the formulation comparison studies.

- Preprocessing

The way preprocessing delivers improvement in computational efficiency is usually by means of reducing the problem size by eliminating some of the variables. In mathematical terms the preprocessing step can be stated as follows:

$$\text{If } \exists\, \widetilde{j} \neq j \text{ s. t. } \textstyle\sum_{k,k>i} \min_l [E_{ik}^{jl} - E_{ik}^{\widetilde{j}l}] > 0$$
$$\text{then } y_i^j = 0 \tag{8}$$

The original idea of the above came from the Dead-End Elimination (DEE) criterion [8, 42, 43, 53]:

$$E(i_a) - E(i_b) + \sum_{k \neq i} \min_c [E(i_a, k_c) - E(i_b, k_c)] > 0 \tag{9}$$

which states that rotamer i_a at position i can be pruned if its energy contributation is always lowered by substituting with an alternative rotamer i_b. In our sequence selection model we only consider amino acids instead of rotamers for each position on the design template. Nevertheless, the DEE criterion is still applicable. Since in [1, 2]'s model the total energy only takes into account pairwise amino acid interactions but not each amino acid itself, the energies of the rotamers i_a and i_b themselves (i.e., $E(i_a)$ and $E(i_b)$) can be cancelled, yielding equation (8) which is in a form incorporable into formulation (F2).

With the three algorithmic enhancing components mentioned above, a list of $O(n^2)$ formulations were generated:

$$\begin{aligned} \min_{y_i^j, y_k^l} \ & \textstyle\sum_{i=1}^{n} \sum_{j=1}^{m_i} \sum_{k=i+1}^{n} \sum_{l=1}^{m_k} E_{ik}^{jl}(x_i, x_k) w_{ik}^{jl} \\ \text{subject to} \ & \textstyle\sum_{j=1}^{m_i} y_i^j \ = \ 1 \ \forall \ i \\ & y_i^j \ + \ y_k^l \ - \ 1 \ \leq \ w_{ik}^{jl} \ \leq \ y_i^j \ \forall \ i, j, k, l \\ & 0 \ \leq \ w_{ik}^{jl} \ \leq \ y_k^l \ \forall \ \ i, j, k, l \\ & \textstyle\sum_{j=1}^{m_i} w_{ik}^{jl} \ \leq \ y_k^l \ \ \forall \ \ i, k, l \\ & y_i^j \ , \ y_k^l \ = \ 0 - 1 \ \ \forall \ \ i, j, k, l \end{aligned} \tag{F6}$$

$$\begin{aligned} \min_{y_i^j, y_k^l} \ & \textstyle\sum_{i=1}^{n} \sum_{j=1}^{m_i} \sum_{k=i+1}^{n} \sum_{l=1}^{m_k} E_{ik}^{jl}(x_i, x_k) w_{ik}^{jl} \\ \text{subject to} \ & \textstyle\sum_{j=1}^{m_i} y_i^j \ = \ 1 \ \forall \ i \\ & y_i^j \ + \ y_k^l \ - \ 1 \ \leq \ w_{ik}^{jl} \ \leq \ y_i^j \ \forall \ i, j, k, l \\ & 0 \ \leq \ w_{ik}^{jl} \ \leq \ y_k^l \ \forall \ \ i, j, k, l \\ & \textstyle\sum_{j=1}^{m_i} w_{ik}^{jl} \ \leq \ y_k^l \ \ \forall \ \ i, k, l \\ & y_i^j \ - \ w_{im}^{jp} \ - \ w_{ik}^{jl} \ + \ w_{km}^{lp} \ \geq \ 0 \\ & \forall i < k < m, j, l, p \ \text{ s. t. } \ S_{ikm}^{jlp} \ = \ E_{ik}^{jl} \ + \ E_{im}^{jp} \ + \ E_{km}^{lp} \leq \text{cutoff} \\ & w_{ik}^{jl} \ + \ w_{im}^{jp} \ + \ w_{km}^{lp} \ - \ y_i^j \ - \ y_k^l \ - \ y_m^p \ + \ 1 \ \geq \ 0 \\ & \forall i < k < m, j, l, p \ \text{ s. t. } \ S_{ikm}^{jlp} \ = \ E_{ik}^{jl} \ + \ E_{im}^{jp} \ + \ E_{km}^{lp} \leq \text{cutoff} \\ & y_i^j \ , \ y_k^l \ = \ 0 - 1 \ \ \forall \ \ i, j, k, l \end{aligned} \tag{F7}$$

$$\begin{aligned} & \text{Preprocessing: If } \exists \, \widetilde{j} \neq j \text{ s. t. } \textstyle\sum_{k, k>i} \min_l [E_{ik}^{jl} \ - \ E_{ik}^{\widetilde{j}l}] \ > \ 0 \\ & \qquad \text{then } y_i^j \ = \ 0 \\ \min_{y_i^j, y_k^l} \ & \textstyle\sum_{i=1}^{n} \sum_{j=1}^{m_i} \sum_{k=i+1}^{n} \sum_{l=1}^{m_k} E_{ik}^{jl}(x_i, x_k) w_{ik}^{jl} \\ \text{subject to} \ & \textstyle\sum_{j=1}^{m_i} y_i^j \ = \ 1 \ \forall \ i \\ & y_i^j \ + \ y_k^l \ - \ 1 \ \leq \ w_{ik}^{jl} \ \leq \ y_i^j \ \forall \ i, j, k, l \\ & 0 \ \leq \ w_{ik}^{jl} \ \leq \ y_k^l \ \forall \ \ i, j, k, l \\ & \textstyle\sum_{j=1}^{m_i} w_{ik}^{jl} \ \leq \ y_k^l \ \ \forall \ \ i, k, l \\ & y_i^j \ , \ y_k^l \ = \ 0 - 1 \ \ \forall \ \ i, j, k, l \end{aligned} \tag{F8}$$

$$
\begin{aligned}
\text{Preprocessing:} \quad & \text{If } \exists\, \widetilde{j} \neq j \text{ s. t. } \textstyle\sum_{k,k>i} \min_l [E_{ik}^{jl} - E_{ik}^{\widetilde{j}l}] > 0 \\
& \text{then } y_i^j = 0 \\
\min_{y_i^j, y_k^l} \quad & \textstyle\sum_{i=1}^{n} \sum_{j=1}^{m_i} \sum_{k=i+1}^{n} \sum_{l=1}^{m_k} E_{ik}^{jl}(x_i, x_k) w_{ik}^{jl} \\
\text{subject to} \quad & \textstyle\sum_{j=1}^{m_i} y_i^j = 1 \;\forall\, i \\
& y_i^j + y_k^l - 1 \leq w_{ik}^{jl} \leq y_i^j \;\forall\, i,j,k,l \\
& 0 \leq w_{ik}^{jl} \leq y_k^l \;\forall\; i,j,k,l \\
& \textstyle\sum_{j=1}^{m_i} w_{ik}^{jl} \leq y_k^l \;\; \forall \;\; i,k,l \qquad \text{(F9)} \\
& y_i^j - w_{im}^{jp} - w_{ik}^{jl} + w_{km}^{lp} \geq 0 \\
\forall i < k < m, j, l, p \text{ s. t. } & S_{ikm}^{jlp} = E_{ik}^{jl} + E_{im}^{jp} + E_{km}^{lp} \leq \text{cutoff} \\
& w_{ik}^{jl} + w_{im}^{jp} + w_{km}^{lp} - y_i^j - y_k^l - y_m^p + 1 \geq 0 \\
\forall i < k < m, j, l, p \text{ s. t. } & S_{ikm}^{jlp} = E_{ik}^{jl} + E_{im}^{jp} + E_{km}^{lp} \leq \text{cutoff} \\
& y_i^j \;,\; y_k^l = 0 - 1 \;\; \forall \;\; i,j,k,l
\end{aligned}
$$

Formulation (F6) is just the original model (F2) with the equality in the RLT constraints changed to "$\leq$". (F7) is (F6) with the addition of triangle inequalities. (F8) is (F6) with preprocessing, whereas (F9) is (F6) with both triangle inequalities and preprocessing. Using the forcefield developed by [19] for the pairwise energy parameters, both cases with no cutoff and with cutoff value of -40 were attempted in imposing the triangle inequalities for formulation (F7). This is to confirm our speculation that a smaller subset rather than the full set of triangle inequalities are needed to speed up the algorithm. For all the other formulations that possess triangle inequalities, only a cutoff value of -40 was attempted.

Finally, the counterparts of formulations (F7), (F8), and (F9) with the equality RLT constraints were also included in the formulation comparison studies. They are:

$$
\begin{aligned}
\min_{y_i^j, y_k^l} \quad & \textstyle\sum_{i=1}^{n} \sum_{j=1}^{m_i} \sum_{k=i+1}^{n} \sum_{l=1}^{m_k} E_{ik}^{jl}(x_i, x_k) w_{ik}^{jl} \\
\text{subject to} \quad & \textstyle\sum_{j=1}^{m_i} y_i^j \;=\; 1 \;\forall\; i \\
& y_i^j \;+\; y_k^l \;-\; 1 \;\leq\; w_{ik}^{jl} \;\leq\; y_i^j \;\forall\; i,j,k,l \\
& 0 \;\leq\; w_{ik}^{jl} \;\leq\; y_k^l \;\forall\; i,j,k,l \\
& \textstyle\sum_{j=1}^{m_i} w_{ik}^{jl} \;=\; y_k^l \;\;\forall\;\; i,k,l \\
& y_i^j \;-\; w_{im}^{jp} \;-\; w_{ik}^{jl} \;+\; w_{km}^{lp} \;\geq\; 0 \\
& \forall i < k < m, j, l, p \;\text{ s. t. }\; S_{ikm}^{jlp} \;=\; E_{ik}^{jl} \;+\; E_{im}^{jp} \;+\; E_{km}^{lp} \leq \text{cutoff} \\
& w_{ik}^{jl} \;+\; w_{im}^{jp} \;+\; w_{km}^{lp} \;-\; y_i^j \;-\; y_k^l \;-\; y_m^p \;+\; 1 \;\geq\; 0 \\
& \forall i < k < m, j, l, p \;\text{ s. t. }\; S_{ikm}^{jlp} \;=\; E_{ik}^{jl} \;+\; E_{im}^{jp} \;+\; E_{km}^{lp} \leq \text{cutoff} \\
& y_i^j \;,\; y_k^l \;=\; 0 - 1 \;\;\forall\;\; i,j,k,l
\end{aligned}
\tag{F10}
$$

$$
\begin{aligned}
& \text{Preprocessing: If } \exists\, \widetilde{j} \neq j \text{ s. t. } \textstyle\sum_{k,k>i} \min_l [E_{ik}^{jl} \;-\; E_{ik}^{\widetilde{j}l}] \;>\; 0 \\
& \qquad \text{then } y_i^j \;=\; 0 \\
\min_{y_i^j, y_k^l} \quad & \textstyle\sum_{i=1}^{n} \sum_{j=1}^{m_i} \sum_{k=i+1}^{n} \sum_{l=1}^{m_k} E_{ik}^{jl}(x_i, x_k) w_{ik}^{jl} \\
\text{subject to} \quad & \textstyle\sum_{j=1}^{m_i} y_i^j \;=\; 1 \;\forall\; i \\
& y_i^j \;+\; y_k^l \;-\; 1 \;\leq\; w_{ik}^{jl} \;\leq\; y_i^j \;\forall\; i,j,k,l \\
& 0 \;\leq\; w_{ik}^{jl} \;\leq\; y_k^l \;\forall\; i,j,k,l \\
& \textstyle\sum_{j=1}^{m_i} w_{ik}^{jl} \;=\; y_k^l \;\;\forall\;\; i,k,l \\
& y_i^j \;,\; y_k^l \;=\; 0 - 1 \;\;\forall\;\; i,j,k,l
\end{aligned}
\tag{F11}
$$

$$
\begin{aligned}
\text{Preprocessing:} \quad & \text{If } \exists\, \widetilde{j} \neq j \text{ s. t. } \textstyle\sum_{k,k>i} \min_l [E^{jl}_{ik} - E^{\widetilde{j}l}_{ik}] > 0 \\
& \text{then } y^j_i = 0 \\
\min_{y^j_i, y^l_k} \quad & \textstyle\sum_{i=1}^{n} \sum_{j=1}^{m_i} \sum_{k=i+1}^{n} \sum_{l=1}^{m_k} E^{jl}_{ik}(x_i, x_k) w^{jl}_{ik} \\
\text{subject to} \quad & \textstyle\sum_{j=1}^{m_i} y^j_i = 1 \;\forall\, i \\
& y^j_i + y^l_k - 1 \leq w^{jl}_{ik} \leq y^j_i \;\forall\, i,j,k,l \\
& 0 \leq w^{jl}_{ik} \leq y^l_k \;\forall\; i,j,k,l \\
& \textstyle\sum_{j=1}^{m_i} w^{jl}_{ik} = y^l_k \;\;\forall\;\; i,k,l \qquad \text{(F12)} \\
& y^j_i - w^{jp}_{im} - w^{jl}_{ik} + w^{lp}_{km} \geq 0 \\
\forall i < k < m, j, l, p \text{ s. t. } & S^{jlp}_{ikm} = E^{jl}_{ik} + E^{jp}_{im} + E^{lp}_{km} \leq \text{cutoff} \\
& w^{jl}_{ik} + w^{jp}_{im} + w^{lp}_{km} - y^j_i - y^l_k - y^p_m + 1 \geq 0 \\
\forall i < k < m, j, l, p \text{ s. t. } & S^{jlp}_{ikm} = E^{jl}_{ik} + E^{jp}_{im} + E^{lp}_{km} \leq \text{cutoff} \\
& y^j_i \,,\; y^l_k = 0 - 1 \;\;\forall\;\; i,j,k,l
\end{aligned}
$$

In total we compared the computational performance of 12 equivalent formulations for *in silico* sequence search. The comparion was executed on five different test problems of various complexity level for the sequence selection for human β defensin 2 (PDB code: 1FD3) [85], with the results shown in Table 2.

Apparently, despite having a lot fewer linear constraints, all O(n) formulations performed poorly compared to the O(n^2) models. We suspect this is because in any O(n) model, pairwise interaction variables like w^{jl}_{ik} do not exist any more after linearization. Consequently, RLT type constraints, which are powerful algorithmic components, cannot be implemented and this leads to a huge deterioration of the performance of the branch-and-bound algorithm. It should be highlighted that (F11), which is the original sequence selection model (F2) plus preprocessing, gave the best results in this study.

Our next effort of algorithmic improvement on the sequence selection stage was based on the idea of replacing the inequalities in (F2) with equalities by somehow changing the declaration of some variables [89]. First, it was noticed that the RLT constraints $\sum_{j=1}^{m_i} w^{jl}_{ik} = y^l_k \;\;\forall\;\; i,k,l$ could have been written as two separate equations, namely $\sum_{j=1}^{m_i} w^{jl}_{ik} = y^l_k \;\;\forall\;\; i,k > i,l$ and $\sum_{l=1}^{m_k} w^{jl}_{ik} = y^j_i \;\;\forall\;\; i,k > i,j$. By doing so it becomes more apparent that the inequalities $w^{jl}_{ik} \leq y^j_i \;\;\forall\;\; i,j,k,l$ and $w^{jl}_{ik} \leq y^l_k \;\;\forall\;\; i,j,k,l$ are already implied by the RLTs, due to the fact that w^{jl}_{ik} is positive. The inequalities $y^j_i + y^l_k - 1 \leq w^{jl}_{ik} \;\;\forall\;\; i,j,k,l$ is active when both y^j_i and y^l_k are one, which will turn variable w^{jl}_{ik} into one. Had w^{jl}_{ik} been declared as binary variables instead of a continuous variables, these inequalities would have been made superfluous also. This can be shown as below:

Table 2. Comparison of CPU times in seconds to obtain one global energy minimum solution among the proposed formulations. Solutions were obtained with CPLEX 8.0 solver enabled with branch and bound algorithm on a single Intel Pentium IV 3.2GHz processor.

Test problem	Sequence search space	Formulations						
		(F1)[a]	(F2)[b]	(F3)[c]	(F4)[d]	(F5)[e]	(F6)[f]	(F7)[g]
1	1.3×10^8	0.30	0.14	0.05	0.04	0.05	0.15	0.23*, 0.21⋆
2	1.0×10^{13}	34874	1.93	12.80	65.04	13.23	2.16	44.02*, 3.01⋆
3	3.3×10^{19}	70.14% gap†	3.01	137.85	2052.2	278.0	3.22	64.39*, 2.87⋆
4	1.7×10^{31}	-	38.14	-	-	-	31.67	- , 29.06⋆
5	3.4×10^{45}	-	74713	-	-	-	30006	- , 65575⋆

Test problem	Sequence search space	Formulations				
		(F8)[h]	(F9)[i] cutoff for tri. ineq.=-40	(F10)[j] cutoff for tri. ineq.=-40	(F11)[k]	(F12)[l] cutoff for tri. ineq.=-40
1	1.3×10^8	0.16	0.11	0.16	0.17	0.11
2	1.0×10^{13}	2.15	2.26	2.01	2.52	2.10
3	3.3×10^{19}	2.94	3.31	3.03	3.43	3.04
4	1.7×10^{31}	31.08	35.48	35.92	25.00	36.15
5	3.4×10^{45}	32657	52276	61872	24388	57569

[a] Original $O(n^2)$ formulation proposed by [1, 2] without RLT constraints.
[b] Base case: original $O(n^2)$ formulation proposed by [1, 2].
[c, d, e] $O(n)$ formulations.
[f] Original $O(n^2)$ formulation with inequality RLT constraints.
[g] Original $O(n^2)$ formulation with inequality RLT constraints and triangle inequalities.
[h] Original $O(n^2)$ formulation with inequality RLT constraints and preprocessing.
[i] Original $O(n^2)$ formulation with inequality RLT constraints and triangle inequalities and preprocessing.
[j] Original $O(n^2)$ formulation with triangle inequalities.
[k] Original $O(n^2)$ formulation with preprocessing.
[l] Original $O(n^2)$ formulation with triangle inequalities and preprocessing.
† Integrality gap obtained after 100,000 sec. CPU time.
* No cutoff for triangle inequalities.
⋆ Cutoff = −40 for triangle inequalities.

Property 3.1 *With y_i^j, y_k^l, and w_{ik}^{jl} declared as binary variables, the following set of equations:*

$$\begin{aligned} \textstyle\sum_{j=1}^{m_i} y_i^j &= 1 \;\forall\; i \\ \textstyle\sum_{j=1}^{m_i} w_{ik}^{jl} &= y_k^l \quad \forall \quad i,k > i,l \\ \textstyle\sum_{l=1}^{m_k} w_{ik}^{jl} &= y_i^j \quad \forall \quad i,k > i,j \end{aligned}$$

already implies that if y_i^j and y_k^l are one, then w_{ik}^{jl} has to be one.

Proof.

- first notice that $\sum_{l=1}^{m_k}\sum_{j=1}^{m_i} w_{ik}^{jl} = \sum_{l=1}^{m_k} y_k^l = \sum_{j=1}^{m_i}\sum_{l=1}^{m_k} w_{ik}^{jl} = \sum_{j=1}^{m_i} y_i^j = 1 \;\forall\; i,k > i$
- expansion on $\sum_{j=1}^{m_i}\sum_{l=1}^{m_k} w_{ik}^{jl}$ gives $\sum_{j=1}^{m_i}\sum_{l=1}^{m_k} w_{ik}^{jl} = \sum_{j:y_i^j=0}\sum_{l:y_k^l=0} w_{ik}^{jl} + \sum_{j:y_i^j=0} w_{ik}^{jl}|_{l:y_k^l=1} + \sum_{l:y_k^l=0} w_{ik}^{jl}|_{j:y_i^j=1} + w_{ik}^{jl}|_{j:y_i^j=1,l:y_k^l=1} = 1 \;\forall\; i,k > i$
- $\sum_{j:y_i^j=0}\sum_{l=1}^{m_k} w_{ik}^{jl} = \sum_{j:y_i^j=0} y_i^j = 0$, and $\sum_{l:y_k^l=0}\sum_{j=1}^{m_i} w_{ik}^{jl} = \sum_{l:y_k^l=0} y_k^l = 0$. Obviously $\sum_{j:y_i^j=0}\sum_{l:y_k^l=0} w_{ik}^{jl}$ is also zero.
- $\sum_{j=1}^{m_i}\sum_{l=1}^{m_k} w_{ik}^{jl} = \sum_{j:y_i^j=0}\sum_{l:y_k^l=0} w_{ik}^{jl} + \sum_{j:y_i^j=0} w_{ik}^{jl}|_{l:y_k^l=1} + \sum_{l:y_k^l=0} w_{ik}^{jl}|_{j:y_i^j=1} + w_{ik}^{jl}|_{j:y_i^j=1,l:y_k^l=1} = 1 \;\forall\; i,k > i$
- it follows that $w_{ik}^{jl}|_{j:y_i^j=1,l:y_k^l=1} = 1 \;\forall\; i,k > i,j,l$

By taking out all the superfluous constraints aforementioned, and by declaring w_{ik}^{jl} as binary variables, a novel formulation, which is totally equivalent to model (F2), is obtained:

$$\begin{aligned} \min_{y_i^j, y_k^l} \; & \textstyle\sum_{i=1}^{n}\sum_{j=1}^{m_i}\sum_{k=i+1}^{n}\sum_{l=1}^{m_k} E_{ik}^{jl}(x_i,x_k) w_{ik}^{jl} \\ \text{subject to} \quad & \textstyle\sum_{j=1}^{m_i} y_i^j = 1 \;\forall\; i \\ & \textstyle\sum_{j=1}^{m_i} w_{ik}^{jl} = y_k^l \quad \forall \quad i,k > i,l \\ & \textstyle\sum_{l=1}^{m_k} w_{ik}^{jl} = y_i^j \quad \forall \quad i,k > i,j \\ & y_i^j \;,\; y_k^l \;\; w_{ik}^{jl} = 0-1 \quad \forall \quad i,j,k > i,l \end{aligned} \tag{F13}$$

Computational performance of (F13) was compared to that of (F2) by recording their CPU times to solve two benchmark sequence selection problems [89]. The first problem was exactly the same as test problem 5 in the formulation comparison study just mentioned (Table 2). Its mutation set was generated by fixing the native CYS at position 8, 15, 20, 30, 37, and 38 on the human β defensin design template, and allowing all other positions to choose from any of the 20 amino acids. This corresponded to a sequence search space of $20^{35} = 3.4 \times 10^{45}$. The second test problem for this study imposed 49

linear biological constraints in addition to the basic sequence selection models of (F2) and (F13). These biological constraints were aimed at improving the quality of the solutions by ensuring that all sequences observe certain properties which are important to molecular function. They include charge constraints:

$$\begin{aligned} 0 \le \textstyle\sum_i y_i^{Arg} + \sum_i y_i^{Lys} - \sum_i y_i^{Asp} - \sum_i y_i^{Glu} \le 3 \ \forall 5 \le i \le 10 \\ 5 \le \textstyle\sum_i y_i^{Arg} + \sum_i y_i^{Lys} \le 10 \ \forall i \\ 0 \le \textstyle\sum_i y_i^{Asp} + \sum_i y_i^{Glu} \le 2 \ \forall i \\ 4 \le \textstyle\sum_i y_i^{Arg} + \sum_i y_i^{Lys} - \sum_i y_i^{Asp} - \sum_i y_i^{Glu} \le 9 \ \forall i \end{aligned} \tag{10}$$

and constraints which place bounds on the occurrence of each amino acid in each sequence:

$$\begin{array}{llll} 0 \le \sum_i y_i^{Ala} \le 3 \, \forall i & 0 \le \sum_i y_i^{Gln} \le 3 \, \forall i \\ 0 \le \sum_i y_i^{Leu} \le 4 \, \forall i & 0 \le \sum_i y_i^{Ser} \le 6 \, \forall i \\ 1 \le \sum_i y_i^{Arg} \le 9 \, \forall i & 0 \le \sum_i y_i^{Glu} \le 3 \, \forall i \\ 0 \le \sum_i y_i^{Lys} \le 7 \, \forall i & 0 \le \sum_i y_i^{Thr} \le 4 \, \forall i \\ 0 \le \sum_i y_i^{Asn} \le 6 \, \forall i & \sum_i y_i^{Gly} \le 6 \, \forall i \\ 0 \le \sum_i y_i^{Met} \le 3 \, \forall i & 0 \le \sum_i y_i^{Trp} \le 2 \, \forall i \\ 0 \le \sum_i y_i^{Asp} \le 2 \, \forall i & 0 \le \sum_i y_i^{His} \le 4 \, \forall i \\ 0 \le \sum_i y_i^{Phe} \le 4 \, \forall i & 0 \le \sum_i y_i^{Tyr} \le 4 \, \forall i \\ \sum_i y_i^{Cys} = 6 \ \forall i & 0 \le \sum_i y_i^{Ile} \le 6 \, \forall i \\ \sum_i y_i^{Pro} \le 5 \ \forall i & 0 \le \sum_i y_i^{Val} \le 6 \, \forall i \end{array} \tag{11}$$

and constraints that restrict β strands to have at least two hydrophobic residues to ensure enough hydrophobic interaction for stability purpose:

$$\begin{aligned} \textstyle\sum_i y_i^{Cys} + \sum_i y_i^{Ile} + \sum_i y_i^{Leu} + \sum_i y_i^{Met} + \sum_i y_i^{Phe} + \\ \textstyle\sum_i y_i^{Trp} + \sum_i y_i^{Tyr} + \sum_i y_i^{Val} + \sum_i y_i^{Ala} \ge 2 \ \forall 14 \le i \le 16 \\ \textstyle\sum_i y_i^{Cys} + \sum_i y_i^{Ile} + \sum_i y_i^{Leu} + \sum_i y_i^{Met} + \sum_i y_i^{Phe} + \\ \textstyle\sum_i y_i^{Trp} + \sum_i y_i^{Tyr} + \sum_i y_i^{Val} + \sum_i y_i^{Ala} \ge 2 \ \forall 25 \le i \le 28 \end{aligned} \tag{12}$$

Lastly, the number of mutations on each solution sequence is permitted to be ten at maximum by the following equation:

$$\textstyle\sum_{i=1}^{n} \sum_{j=1, j \neq \text{native residues}}^{m_i} y_i^j \le 10 \tag{13}$$

These conserved properties were elucidated by running a sequence alignment tool like PSI-BLAST, which was created by the National Center for Biotechnology Information (NCBI) of the National Institute of Health, on the

human β defensin homologs. The mutation set was derived from Solvent Accessible Surface Area (SASA) patterning, and it corresponded to a sequence search space of 6.4×10^{37} in this case. In both problems the forcefield developed by [19] was employed for the energy parameters in the model.

The CPU times it took for the two formulations to converge to the global optimal solution for the two problems are tabulated in Table 3 [89]. As shown by the comparison, (F13) outperformed (F2) by a large margin. Its CPU times to solve the first and second problem were 82-fold and 327-fold shorter respectively.

Table 3. Comparison of computational performance of two different formulations for the sequence selection for a human β defensin (PDB code:1FD3) design template.

First Problem			
	Number of	CPU times† [sec]	
Problem complexity	biological constraints	Formulation (F2)	Formulation (F13)
3.4×10^{45}	none	53,263	649
Second Problem			
	Number of	CPU times [sec]	
Problem complexity	biological constraints	Formulation (F2)	Formulation (F13)
6.4×10^{37}	49	4,578	14

† Generated using CPLEX 9.0 on one single Pentium IV 3.2 GHz processor.

The discrepany between the CPU times taken by (F2) to solve the sequence selection problem with complexity 3.4×10^{45} as shown in Table 2 and Table 3 is suspected to be caused by the different versions of CPLEX solver used.

To summarize this subsection, (F13) constituted the best we have developed so far for the sequence selection model for a design template with single structure. We also formulated models which handle explicitly the case in which a design template exhibits multiple structures [89]. These novel formulations can be found in the subsection that follows which outlines the incorporation of higher degree of true backbone flexibility.

Fold Specificity Stage

The second stage is supposed to provide a more rigorous assessment of the specificity of the low energy sequences within the context of the flexible template. The ASTRO-FOLD method, which is what [1, 2] used for the second stage, performs rigorous protein folding calculations to generate two sets of conformational ensembles: one in which the protein is constrained to a region around the backbone and the other in which the protein is allowed to fold freely. The relative probability of specificity for the protein to assume the target fold is then calculated from the RMSD and energy of these two ensembles

based on the Boltzmann distribution. For rigorous ensemble generation, this method requires that a large number of free-folding calculations be performed, which is computationally expensive.

Approximate Method for Fold Specificity Calculation

As the sequence selection model (F13) can now effectively consider proteins that are reaching the limits of the ASTRO-FOLD method, a new second stage method was developed to handle these large proteins [115]. The approximate method for fold specificity calculation is outlined in Figure 1. First, a flexible template is defined based on the upper and lower bounds on both the distances between alpha carbons and the phi and psi angles between residues. An ensemble of hundreds of random structures is then generated (conformers) within the confines of the flexible template using the CYANA 2.1 software package for NMR structure refinement [90, 91]. CYANA 2.1 is then used to perform annealing calculations that simulate a rapid heating of the protein followed by a slow cooling in which high temperature torsion dynamics and annealing torsion dynamics are performed. Violations of Van der Waals radii and of the flexible template are minimized, minimizing the energy of the target structures. Hundreds of these structures are generated within the confines of the flexible template.

For each structure in the ensemble, local minimizations are then performed by the TINKER [92] package as directed by gradients in the fully atomistic force field AMBER [93]. AMBER is used to evaluate the potential energy of the structure. These ensembles are generated for the native sequence of the fold and for each candidate mutant sequence. The specificity of each mutant sequence to the target fold is then calculated relative to the native sequence using the Boltzmann distribution from statistical mechanics. Both the predicted energy of each conformer and its RMSD from the template structure are used in this calculation.

Ensemble Generation Using CYANA 2.1

The simulated annealing process in CYANA 2.1 [90, 91] starts by generating random structures for each amino acid sequence within the confines of the flexible template. A series of minimizations are then performed to reduce the number of strong overlaps between atoms as defined by their Van der Waal radii. CYANA then simulates a sharp increase in temperature, greatly increasing the degrees of freedom in the protein in a molecular and torsion dynamics simulation that allows the shape of the protein to change. The protein is then slowly cooled down, or annealed, as these calculations are performed. A final energy minimization is performed over all atoms to give the output structure. This process has been used previously for protein decoy generation using an older version of the software, DYANA [19]. CYANA 2.1

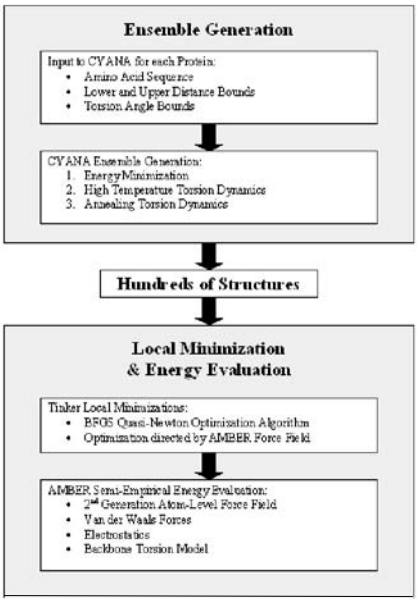


Fig. 1. Workflow for the novel approximate method for fold validation.

has a number of significant improvements over DYANA. These are outlined below:

- The simulated annealing has been improved in order to achieve better convergence of the structure calculations
- The standard amino acid residue library has been revised and van der Waals radii have been optimized to produce structures with better packing interactions.
- Improved angle restraints.

Local Minimizations and Energy Evaluation

The TINKER package of software tools for molecular design [92] performs a series of local minimizations of the structures in the ensemble. These minimizations use the BFGS Quasi-Newton optimization algorithm and are guided gradients in the force field of interest. In these calculations, the AMBER force field [93] is used. Following a set of minimizations, the potential energy of the structure is evaluated using the AMBER force field.

Analysis and Selection of Results

A method similar to that used by [1, 2] for ensemble comparison in fold validation was employed to give a relative ranking for specificity [115]. First, the mean and standard deviation of both RMSD and AMBER energies were found for the native sequence. Upper bounds on both RMSD and energy were then established; for RMSD the upper bound was selected as one and a half standard deviations above the mean, in the energy the upper bound was selected as two standard deviations from the mean. A structure is considered to make a contribution to the ensemble only if its energy and RMSD both fall under these upper bounds. This is illustrated in Figure 2.

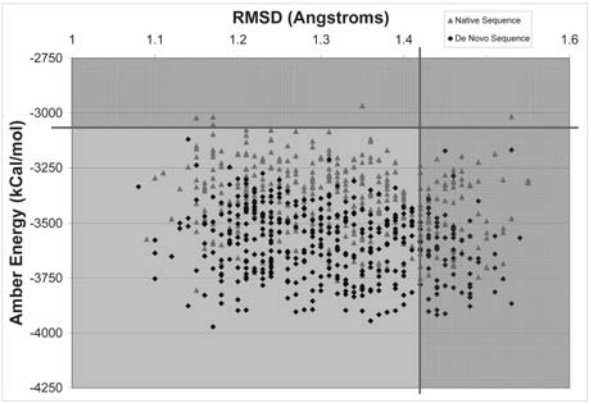


Fig. 2. Illustration of upper bounds on RMSD and AMBER energy. Red lines indicate upper bounds. Data points in the red-shaded regions are not considered in further calculations.

To calculate the relative factor for specificity, define the set native as the set of all data points from the native sequence that are below both upper bounds, and set novel as the set of all data points from the novel sequence that meet the same criterion. The factor for specificity $f_{specificity}$ is then calculated using Boltzmann probabilities as shown in the the following equation:

$$f_{specificity} = \frac{\sum_{i \in novel} \exp[-\beta E_i]}{\sum_{i \in native} \exp[-\beta E_i]} \tag{14}$$

where $\beta = \frac{1}{k_B T}$.

Incorporation of True Backbone Flexibility

Like ASTRO-FOLD, CYANA can pick any continuous values for the dihedral angles and C^{α}-C^{α} distances between preset bounds when it does the simulated annealing calculations. Hence the desirable feature of compatibility with true backbone flexibility [88] is reserved. The bounds are input by the user to the program, and they can be based on his or her observation about the flexible design template(s). The outcoming protein conformations can thus have any possible combination of continuous angle and distance values between the bounds.

3.2 Incorporating Higher Degree of True Protein Backbone Flexibility

Allowing true backbone flexibility by the use of distance bins, sequence selection models (F2) and (F13) can also be applied to highly flexible templates where all the distances between any position pair i and k fall into the same distance bin. However, in the general case, the same distance in different template structures will vary. To the best of our knowledge, there is currently no de novo protein design model in open literature that explicitly deals with such kind of high degree backbone flexibility. In view of this, we have developed novel sequence selection formulations to fill this void [89]. This subsection presents these new models which handle explicitly highly flexible design templates that have multiple crystal or NMR structures.

In our derivation two different approaches were applied: one uses a weighted average forcefield with the weights given by the occurrence frequencies of the C^{α}-C^{α} or centroid-centroid distance between a position pair i and k falling into different distance bins, and the other allows the possibility of spanning all the distance bins that the C^{α}-C^{α} or centroid-centroid distance between i and k covers by the use of binary distance bin variables.

Formulation Using a Weighted Average Forcefield

This approach is relatively simple and straightforward to follow from the model for single template structure. In the case when there is only one structure, the energy parameter $E_{ik}^{jl}(x_i, x_k)$ in the objective function can be immediately determined by the coordinates of the two C^{α} or centroid positions, i.e., x_i and x_k, as well as the amino acid at each of those two positions. There is no ambiguity as to which distance bin d it belongs to. In the case of multiple structures, the term $E_{ik}^{jl}(x_i, x_k)$ can be replaced by a weighted average energy term, $\sum_{d=1}^{b_m} E_{ik}^{jl}(x_i, x_k) wt(x_i, x_k, d)$, where the weights $wt(x_i, x_k, d)$ are given by:

$$wt(x_i, x_k, d) \;=\; \frac{\#\text{ of structures where dist.}(x_i, x_k)\text{ falls into bin } d}{\text{total }\#\text{ of template structures}} \;\forall\; i, k, d$$

The idea can also be examined this way: the distance between x_i and x_k is now replaced by a weighted average distance over all the structures, with the weights given by the above formula. The energy parameters $E_{ik}^{jl}(x_i, x_k)$ can be found using this weighted average distance and simple table lookup in the corresponding forcefield. For instance, in compstatin (PDB code: 1A1P), a synthetic 13-residue peptide and a pharmaceutical candidate that interferes with complement activation with its details quoted in the case studies section that immediately follows, the distribution for the distance between the alpha carbon of the first residue and the third residue is as follows: bin 4 (5.5 to 6.0 ρA): 1 structure; bin 5 (6.0 to 6.5 ρA): 9 structures; bin 6 (6.5 to 7.0 ρA): 10 structures; and bin 7 (7.0 to 8.0 ρA): 1 structure. Data for the 21 structures were deposited in the Protein Data Bank for compstatin. Therefore, $wt(x_1, x_3, 4) = \frac{1}{21} = 0.0476$, $wt(x_1, x_3, 5) = \frac{9}{21} = 0.429$, $wt(x_1, x_3, 6) = \frac{10}{21} = 0.476$, $wt(x_1, x_3, 7) = \frac{1}{21} = 0.0476$, and $wt(x_1, x_3, d) = 0\ \forall d \neq 4, 5, 6, 7$. It should be noticed that in the case of the force field used for generating results presented in this paper [87], the sum of the weights $wt(x_i, x_k, d)$ over the distance bin set $d = 1, \ldots, b_m = 8$ does not equal to one. This is simply because the distance bins only cover the range of 3 to 9 ρA in the force field. Had the bins covered the full positive distance range, the weights would have added up to one.

All the other components in formulation (F13) for single template structure can be kept for this new weighted average forcefield formulation. Therefore in summary, the novel weighted average forcefield formulation for designing proteins into multiple highly flexible templates takes the form [115]:

$$\begin{aligned} \min_{y_i^j, y_k^l} \ & \sum_{i=1}^{n} \sum_{j=1}^{m_i} \sum_{k=i+1}^{n} \sum_{l=1}^{m_k} \sum_{d=1}^{b_m} E_{ik}^{jl}(x_i, x_k) wt(x_i, x_k, d) w_{ik}^{jl} \\ \text{subject to} \quad & \sum_{j=1}^{m_i} y_i^j = 1 \ \forall\ i \\ & \sum_{j=1}^{m_i} w_{ik}^{jl} = y_k^l \ \ \forall \ \ i, k > i, l \\ & \sum_{l=1}^{m_k} w_{ik}^{jl} = y_i^j \ \ \forall \ \ i, k > i, j \\ & y_i^j \ , \ y_k^l \ , \ w_{ik}^{jl} = 0 - 1 \ \ \forall \ \ i, j, k, l \end{aligned} \tag{F14}$$

Like (F2) or (F13), it is an integer linear programming (ILP) model.

Formulation Using Binary Distance Bin Variables

Another more elegant and advanced approach to incorporate distance information from multiple structures is by using a binary distance bin variable, b_{ikd}, which assumes the value of one if the distance between x_i and x_k falls into distance bin d and the value of zero otherwise. A parameter, $disbin(x_i, x_k, d)$, which will be used in the derivation of the constraints, needs to be defined:

$disbin(x_i, x_k, d)$
= 1 if the distance between x_i and x_k in ANY of the template structures

falls into bin d
$= 0$ otherwise $\forall \;\; i, k > i, d$

Hence for the first and third residue of compstatin, parameter $disbin(x_1, x_3, d)$ equals one for $d = 4, 5, 6, 7$ and zero for other distance bins. With this new parameter, constraints $\sum_{d:disbin(x_i,x_k,d)=1} b_{ikd} = 1 \; \forall i, k > i$ can be imposed. This constraint essentially lets the energy minimization model free to pick only one of the distance bins that all the structures cover. Thus in the same example of compstatin, the constraint $\sum_{d=4,5,6,7} b_{13d} \;=\; 1$ is to be added. Since only the distance bin d with b_{ikd} assigned to be 1 will contribute to the total energy of the protein, when deriving this new formulation, the term $E^{jl}_{ik}(x_i, x_k)$ in the objective function of formulation (F13) can be replaced by $\sum_{d:disbin(x_i,x_k,d)=1} E^{jl}_{ik}(x_i, x_k) b_{ikd}$, leading to a new model that looks like:

$$\min_{y^j_i, y^l_k} \sum_{i=1}^{n} \sum_{j=1}^{m_i} \sum_{k=i+1}^{n} \sum_{l=1}^{m_k} \sum_{d:disbin(x_i,x_k,d)=1} E^{jl}_{ik}(x_i, x_k) b_{ikd} w^{jl}_{ik}$$

subject to

$$\begin{aligned}
\textstyle\sum_{j=1}^{m_i} y^j_i \;&=\; 1 \; \forall \; i \\
\textstyle\sum_{j=1}^{m_i} w^{jl}_{ik} \;&=\; y^l_k \;\; \forall \;\; i, k > i, l \\
\textstyle\sum_{l=1}^{m_k} w^{jl}_{ik} \;&=\; y^j_i \;\; \forall \;\; i, k > i, j \\
\textstyle\sum_{d:disbin(x_i,x_k,d)=1} b_{ikd} \;&=\; 1 \;\; \forall \;\; i, k > i \\
y^j_i \;,\; y^l_k \;,\; w^{jl}_{ik} \;,\; b_{ikd} \;&=\; 0 - 1 \;\; \forall \;\; i, j, k > i, l, d
\end{aligned} \tag{15}$$

Formulation (15) is non-convex because of the bilinear term $b_{ikd} w^{jl}_{ik}$ in the objective function. The formulation could have been linearized in the same way as for formulation (1), i.e., by using a positive continuous variable $z^{jl}_{ikd} = b_{ikd} w^{jl}_{ik}$ with the addition of four sets of inequalities to reproduce the original characteristics:

$$\begin{aligned}
b_{ikd} \;+\; w^{jl}_{ik} \;-\; 1 \;\leq\; z^{jl}_{ikd} \;\leq\; b_{ikd} \; \forall \; i, j, k > i, l, d \\
0 \;\leq\; z^{jl}_{ikd} \;\leq\; w^{jl}_{ik} \; \forall \;\; i, j, k > i, l, d
\end{aligned} \tag{16}$$

However, based on the observation about the superior computational performance of formulation (F13), linearization was done by declaring $z^{jl}_{ikd} = b_{ikd} w^{jl}_{ik}$ as a binary variable and using the RLT equations:

$$\begin{aligned}
w^{jl}_{ik} \sum_{d:disbin(x_i,x_k,d)=1} b_{ikd} \;&=\; 1 \;\; \forall \;\; i, j, k > i, l \;\; \text{or :} \\
\sum_{d:disbin(x_i,x_k,d)=1} z^{jl}_{ikd} \;&=\; w^{jl}_{ik} \;\; \forall \;\; i, j, k > i, l \\
w^{jl}_{ik} \;,\; b_{ikd} \;,\; z^{jl}_{ikd} \;&=\; 0 - 1 \;\; \forall \;\; i, j, k > i, l, d
\end{aligned} \tag{17}$$

This RLT equation already implies $z^{jl}_{ikd} \leq w^{jl}_{ik} \;\; \forall \;\; i, j, k > i, l, d$, and declaring z^{jl}_{ikd} as a binary variable means $z^{jl}_{ikd} \geq 0 \;\; \forall \;\; i, j, k > i, l, d$. Both of these equations can thus be dropped from the formulation.

Constraints on Distance Bin Variables

Since two alpha carbons are now free to pick any distance bin that the template structures cover, additional constraints have to be imposed on their distance bin variables to avoid physically meaningless results. This requirement can be proved necessary by considering the case shown in Figure 1, in which x_i, x_k, and x_p are three distinct positions and both the distance between x_i and x_k and that between x_k and x_p select to be in bin 1 in the energy minimization model. Assume the average distance between x_i and x_p over all the template structures is $10\rho A$. The selections will result in no overlap between the two shaded regions, each of which corresponds to the area where position x_k can possibly be. The constraints on the distance bin variables are supposed to ensure there is some kind of consistency about the possible location of any alpha carbon throughout the distance bin selection process.

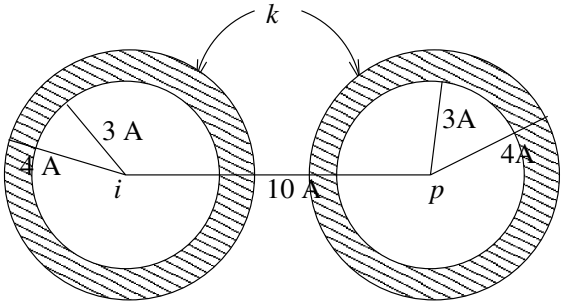


Fig. 3. No overlap between the shaded regions where position x_k can possibly be.

There are two cases in which the constraints should come into play. The first case is illustrated by Figure 2, where the areas corresponding to the binary variables b_{ikd} and $b_{kpd'}$ do not overlap. The condition for no overlap is: $l_{mid}(d) < dis(i,p) - l_{mid}(d')$, where $dis(i,p)$ is the average distance between x_i and x_p over all template structures. Since if both variables are one there will be no consistency about the location of position x_k, the necessary constraint is: $b_{ikd} + b_{kpd'} \leq 1$. However, it should be highlighted that infeasibility problem may occur to the model if only the no overlap condition is used for checking constraint applicability. This is because an average value has been used for $dis(i,p)$, with which the areas corresponding to the non-zero $disbin(x_i, x_k, d)$'s and that corresponding to the variable $b_{kpd'}$ may not overlap at all. Hence an additional condition has to be imposed besides the no overlap criterion: $\sum_{d"=d+1}^{b_m} disbin(x_i, x_k, d") \geq 1$. It means that there is at least one non-zero $disbin(x_i, x_k, d")$ for $d" > d$, whose area may overlap with that corresponding to $b_{kpd'}$. The model can thus select any of these bins for the distance between x_i and x_k and avoid the problem of infeasibility.

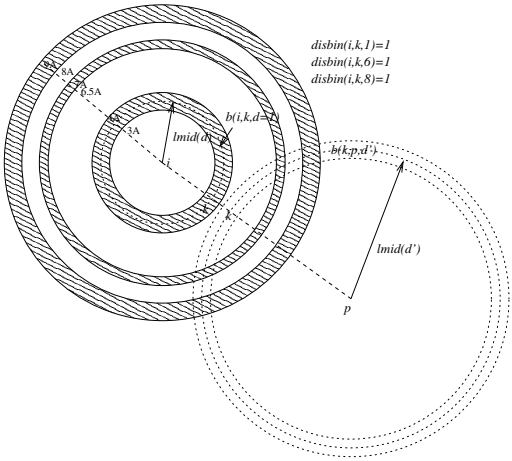


Fig. 4. First case in which constraints on distance bin variables are applicable: no overlap between the areas corresponding to the binary variables $b(x_i, x_k, d)$ and $b(x_k, x_p, d')$ because $l_{mid}(d) < dis(i,p) - l_{mid}(d')$. Condition $\sum_{d"=d+1}^{b_m} disbin(x_i, x_k, d") \geq 1$ has to hold to avoid infeasibility.

The second case in which the constraints are applicable is shown in Figure 3. This case differs from the first one in its no overlap condition, which can be expressed as the equation $l_{mid}(d') > dis(i,p) + l_{mid}(d)$. Again, to get around the problem of infeasibility, the constraints are only to be applied when $\sum_{d"=d+1}^{b_m} disbin(x_i, x_k, d") \geq 1$, in addition to the no overlap criterion.

In summary, the constraints on binary distance bin variables take the form of:

$$\begin{gathered} b_{ikd} + b_{kpd'} \leq 1 \\ \text{if } (l_{mid}(d') < dis(i,p) - l_{mid}(d) \text{ or } l_{mid}(d') > dis(i,p) + l_{mid}(d)) \\ \text{and } \sum_{d"=d+1}^{b_m} disbin(x_i, x_k, d") \geq 1 \text{ and } disbin(x_i, x_k, d) = 1 \\ \text{and } disbin(x_k, x_p, d') = 1 \ \forall \ i, k > i, p, d, d', i \neq k \neq p \end{gathered} \tag{18}$$

and the novel formulation for designing proteins into a template with multiple structures, by using binary distance bin variables and their constraints, is [115]:

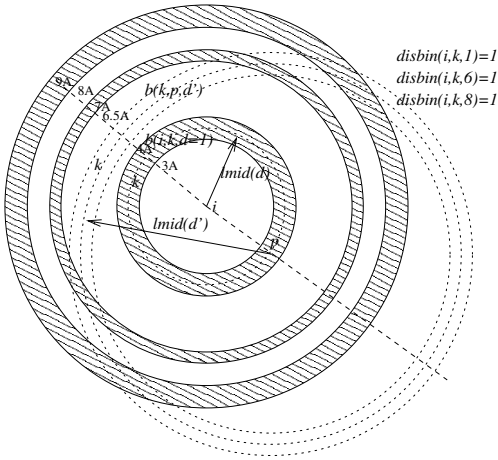


Fig. 5. Second case in which constraints on distance bin variables are applicable: no overlap between the areas corresponding to the binary variables $b(x_i, x_k, d)$ and $b(x_k, x_p, d')$ because $l_{mid}(d') > dis(i,p) + l_{mid}(d)$. Condition $\sum_{d"=d+1}^{b_m} disbin(x_i, x_k, d") \geq 1$ has to hold to avoid infeasibility.

$$\min_{y_i^j, y_k^l} \sum_{i=1}^{n} \sum_{j=1}^{m_i} \sum_{k=i+1}^{n} \sum_{l=1}^{m_k} \sum_{d:disbin(x_i,x_k,d)=1} E_{ik}^{jl}(x_i, x_k) z_{ikd}^{jl}$$

subject to

$$\begin{gathered}
\textstyle\sum_{j=1}^{m_i} y_i^j = 1 \;\forall\; i \\
\textstyle\sum_{j=1}^{m_i} w_{ik}^{jl} = y_k^l \;\;\forall\;\; i, k > i, l \\
\textstyle\sum_{l=1}^{m_k} w_{ik}^{jl} = y_i^j \;\;\forall\;\; i, k > i, j \\
\textstyle\sum_{d:disbin(x_i,x_k,d)=1} b_{ikd} = 1 \;\;\forall\;\; i, k > i \\
b_{ikd} + w_{ik}^{jl} - 1 \leq z_{ikd}^{jl} \leq b_{ikd} \;\forall\; i, j, k > i, l, d \\
\textstyle\sum_{d:disbin(x_i,x_k,d)=1} z_{ikd}^{jl} = w_{ik}^{jl} \;\;\forall\;\; i, j, k > i, l \qquad \text{(F15)} \\
b_{ikd} + b_{kpd'} \leq 1 \\
\text{if } (l_{mid}(d') < dis(i,p) - l_{mid}(d) \;\text{ or }\; l_{mid}(d') > dis(i,p) + l_{mid}(d)) \\
\text{and } \textstyle\sum_{d"=d+1}^{b_m} disbin(x_i, x_k, d") \geq 1 \text{ and } disbin(x_i, x_k, d) = 1 \\
\text{and } disbin(x_k, x_p, d') = 1 \;\forall\; i, k > i, p, d, d', i \neq k \neq p \\
y_i^j \;,\; y_k^l \;,\; w_{ik}^{jl} \;,\; b_{ikd} \;,\; b_{kpd'} \;,\; z_{ikd}^{jl} = 0 - 1 \\
\forall\;\; i, j, k > i, l, p \neq k \neq i, d, d'
\end{gathered}$$

In the next section, application of the new sequence selection models (i.e., (F13 for single template structures, or (F14) and (F15) for multiple template structures) and the approximate fold specificity method for de novo design will be illustrated. The illustration is aided with three redesign examples: compstatin, a 14-residue short constrained peptide that inhibits the viral surface glycoprotein HIV-1 gp41, and Complement 3a.

4 Case Studies

4.1 Compstatin

Compstatin (PDB code: 1A1P) is a synthetic 13-residue cyclic peptide that inhibits the cleavage of C3 to C3a and C3b in the human complement system and thus hinders complement activation. It is a novel drug candidate identified through the screening of a phage-displayed random peptide library with C3b, a proteolytically activated form of complement protein C3, and was later truncated to its present 13-residue form without loss of activity [32]. Although complement activation is part of normal inflammatory response, inappropriate complement activation can cause host-cell damage, which is the case in more than 25 pathological conditions, including autoimmune diseases, stroke, heart attack, Alzheimer's disease, and burn injuries [31].

Compstatin has shown highly promising results in numerous pre-clinically relevant trials conducted recently. Compstatin blocked the cleavage of C3 to the pro-inflammatory peptide C3a and the opsonin C3b in hemolytic assays and in human normal serum [32, 33], prevented heparine/protamine-induced complement activation in baboons in a situation resembling heart surgery [36], inhibited complement activation during the contact of blood with biomaterial in a model of extra-corporeal circulation [37], increased the lifetime of survival of porcine kidneys perfused with human blood in a hyper-acute rejection xenotransplantation model [38], blocked the E coli -induced oxidative burst of granulocytes and monocytes [39], and inhibited complement activation by cell line SH-SY5Y [40]. Compstatin was stable in biotranformation studies in vitro in human blood, normal human plasma and serum, with increased stability upon N-terminal acetylation [33]. Compstatin showed little or low toxicity and no adverse effects when these were measured [36–38]. Finally, compstatin showed species-specificity and is active only with human and primate C3 [41].

De novo design on compstatin is aimed at acquiring the sequences for the best inhibitors to C3. Earlier research done by [1, 2] based on one of the 21 NMR structures of compstatin found sequences for novel inhibitors that are up to 16 times more potent than the native compstatin [106]. This time our work will be based on the flexible template of all 21 structures. Works performed by [32–34], and [35], and a recent one by [84] revealed some sequence-function relationships for compstatin (refer to Table 4). The predictions from our de novo design are in excellent correlation with experimental activity data.

The Flexible Templates for Compstatin

The flexible template for compstatin as defined by its 21 NMR structures available from the PDB is shown in Figure 6. The structures in this case do not deviate from each other by too much. Each of them corresponds to

Table 4. Sequence and experimental relative activity of compstatin analogs with improved activity that were identified by rational design, experimental combinatorial design, and the novel in silico de novo protein design approach. Boldface is used to indicate that amino acids were fixed. Brackets indicate the disulfide bridge. Relative complement inhibitory activity is derived from IC_{50} measurements.

Peptide	Sequence	Relative activity	Ref
compstatin	$I[\mathbf{CV}V\mathbf{QDWG}HHR\mathbf{C}]T-NH_2$	1	[32]
Ac-compstatin	$Ac-I[\mathbf{CV}V\mathbf{QDWG}HHR\mathbf{C}]T-NH_2$	3	[33]
Ac-H9A	$Ac-I[\mathbf{CV}V\mathbf{QDWG}AHR\mathbf{C}]T-NH_2$	4	[34]
Ac-I1L/H9W/T13G	$Ac-L[\mathbf{CV}V\mathbf{QDWG}WHR\mathbf{C}]G-NH_2$	4	[35]
Ac-I1V/V4Y/H9F/T13V	$Ac-V[\mathbf{CV}Y\mathbf{QDWG}FHR\mathbf{C}]V-NH_2$	6	[1]
Ac-I1V/V4Y/H9A/T13V	$Ac-V[\mathbf{CV}Y\mathbf{QDWG}AHR\mathbf{C}]V-NH_2$	9	[1]
Ac-V4Y/H9F/T13V	$Ac-I[\mathbf{CV}Y\mathbf{QDWG}FHR\mathbf{C}]V-NH_2$	11	[1]
Ac-V4Y/H9A/T13V	$Ac-I[\mathbf{CV}Y\mathbf{QDWG}AHR\mathbf{C}]V-NH_2$	14	[1]
Ac-V4Y/H9A	$Ac-I[\mathbf{CV}Y\mathbf{QDWG}AHR\mathbf{C}]T-NH_2$	16	[1]
Ac-V4W/H9A	$Ac-I[\mathbf{CV}W\mathbf{QDWG}AHR\mathbf{C}]T-NH_2$	45	[84]

the native sequence of Ile^1-Cys^2-Val^3-Val^4-Gln^5-Asp^6-Trp^7-Gly^8-His^9-His^{10}-Arg^{11}-Cys^{12}-Thr^{13}-NH_2, with a disulfide bond connecting the two cysteines at positions 2 and 12. Similar to several other peptides relevant to immune system function, compstatin, when free in solution, adopts a β-turn structure, which is located across Gln^5-Asp^6-Trp^7-Gly^8. The β-turn is considered to be important for the structural stability of free compstatin and possibly for functional recognition through specific side chain interactions with C3 [33].

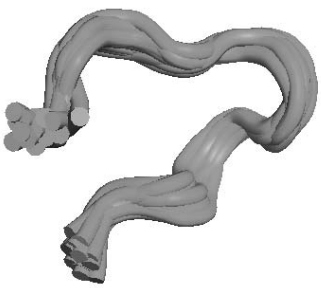


Fig. 6. Flexible template of compstatin for de novo protein design as illustrated by overlapping its 21 NMR structures available from the Protein Data Bank.

Mutation Set

Since the disulfide bridge was found to be essential for aiding in the formation of the hydrophobic cluster and prohibiting the termini from drifting apart, both residues Cys^2 and Cys^{12} were maintained. In addition, because the structure of the type-I β turn was not found to be a sufficient condition for activity, the turn residues were fixed to be those of the parent compstatin sequence; namely Gln^5-Asp^6-Trp^7-Gly^8. In fact, when stronger type I β sequences were constructed, which was supported by NMR data indicating that these sequences provided higher β turn populations than compstatin, some of these sequences resulted in lower or no activity [34]. Therefore, the further optimization of the turn residues, which would likely be a consequence of the computational peptide design procedure, may not enhance compstatin activity. This is especially true for Trp^7, which was found to be a likely candidate for direct interaction with C3. For similar reasons, Val^3 was maintained throughout the computational experiments.

After designing the compstatin system to be consistent with those features found to be essential for compstatin activity, six residue positions were selected to be optimized. Of these six residues, positions 1, 4, and 13 have been shown to be structurally involved in the formation of a hydrophobic cluster consisting of residues at positions 1, 2, 3, 4, 12, and 13, a necessary but not sufficient component for compstatin binding and activity. The remaining residues, namely those at positions 9, 10 and 11, span the three positions between the turn residues and the C-terminal cysteine. For the wild type sequence these positions are populated by positively charged residues, with a total charge of +2 coming from two histidine residues (0.5 charge each at pH 6-7) and one arginine residue.

Based on the structural and functional characteristics of those residues involved in the hydrophobic cluster, positions 1, 4 and 13 were allowed to select only from the hydrophobic amino acid set (A,F,I,L,M,W,V,Y). In addition, this set included threonine for position 13 to allow for the selection of the wild type residue at this position. For positions 9, 10, and 11, all residues were allowed, excluding cysteine and tryptophan. This mutation set leads to a problem with complexity 3.0×10^6. With both the forcefield developed by [19] and the one by [87], 500 sequences were generated for each of the formulation using a weighted average forcefield and the formulation using binary distance bin variables.

Results

The percentage occurrence of the selected amino acids at each of the 6 varied positions in the 500 sequence solutions is tabulated in Table 5.

The results show that for both forcefields, the formulation using weighted average energy and the formulation using binary distance bin variables gave

Table 5. Sequence selection results for de novo protein design with the flexible template of compstatin (21 NMR structures) using two different formulations. Selected amino acids with less than 5% occurrence are not listed.

Formulation using a weighted average forcefield			
Varied position	Native residue	Selections by the model	
		High Resolution forcefield [87]	LKF forcefield [19]
1	I	W, A, F, V	A, Y, V
4	V	F, W, Y	I, L, V, Y, W
9	H	T, H, K, R, F, I	A, N, P, S
10	H	E, F, H, V, N, T, A	Y, F, H
11	R	A, F	A, E, D, N
13	T	A, W	A, Y, V
Formulation using binary distance bin variables			
Varied position	Native residue	Selections by the model	
		High Resolution forcefield [87]	LKF forcefield [19]
1	I	F, W, V, A	Y, A, L, F
4	V	F, W, Y	I, Y, V, L, W
9	H	I, T, F, H, R, M, L	P, A, S, V, N
10	H	F, T, V, E	Y, F, H, V
11	R	A, F, V	A, N, E, D
13	T	W, A, F	A, Y, M, V

very similar results in the case of compstatin. This may be due to the slight deviation from each other among the 21 structures of the compstatin template. The two forcefields suggested slightly different predictions for each varied position. However, both forcefields suggested tryptophan (W) for position 4, which is in agreement with the experimental finding by [84]. These data proposed tryptophan or fused-ring non-natural amino acids at position 4 would contribute to high inhibitory activity of the peptide.

4.2 Human β Defensin-2

Human β Defensin-2 (hβD-2) is a cysteine-rich 41-residue cationic peptide found in the human immune system. It belongs to the class of small, cationic peptides known as defensins. hβD-2 is crucial to innate immunity [116]. It possesses antimicrobial property derived from the electrostatic force between the positive charge on the defensin molecule and the negative charge of the anionic head group of the microbe's membrane lipids. This electrostatic force disrupts the microbe's cell membrane and thus kills the cell [116].

De novo design on human β defensin-2 is carried out with the purpose of synthesizing better and more potent immunopeptides.

The Flexible Template for hβD-2

The design template for the de novo design of hβD-2 corresponds to the X-ray crystal structure elucidated by [116](PDB code: 1FD3) at a resolution of $1.35\rho A$ (Figure 7). Like other human β-defensins, hβD-2 has an N-terminus α-helix located at Pro^{5}-Lys^{10} which is held against the β-sheet by a S-S bond between Cys^{8} and Cys^{37}. Two other S-S bonds that stabilize the β-sheet are located at Cys^{15}-Cys^{30} and Cys^{20}-Cys^{38}. The β-sheet is made up of three anti-parallel β-strands held together by hydrophobic interaction. Though HβD-2 possesses an octameric tertiary structure constituted by eight identical chains: chain A, B, C, D, E, F, G, and H, each of which has the natural sequence of GIGDPVTCLKSGAICHPVFCPRRYKQIGTCGLPGTKCCKKP [117], only chain A was used in this work.

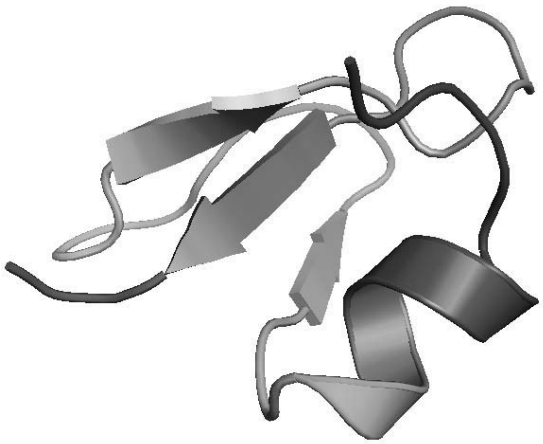


Fig. 7. Structure of human β defensin-2 (chain A). Its secondary structure consists of a β-sheet made up of three β-strands and an α-helix.

Mutation Set

SASA patterning was applied to restrict the sequence search space for the de novo design of hβD-2. The 41 positions in hβD-2 are classified into the core, surface, and intermediate categories which determine the mutation set for each position. The native residue for each position is also included in its mutation set. Proline is excluded from the list for surface and intermediate positions to avoid unnecessary rigidity imposed on the backbone, except for the case when the native residue for a position is proline. The mutation set for human β defensin-2 is tabulated in Table 6.

Complexity of the problem amounts to 6.40×10^{37}.

Table 6. Mutation set of human β defensin-2 given by SASA patterning.

Position	Native residue	Side-chain accessibility	Position type	Varied position?	Allowed mutations
1	G	139.6%	surface	√	R,N,D,Q,E,G,H,K,S,T
2	I	20.7%	intermediate	√	all except C and P
3	G	69.0%	surface	√	R,N,D,Q,E,G,H,K,S,T
4	D	52.5%	surface	√	R,N,D,Q,E,G,H,K,S,T
5	P	52.1%	surface	√	R,N,D,Q,E,G,H,K,P,S,T
6	V	99.9%	surface	√	R,N,D,Q,E,G,H,K,S,T,V
7	T	54.9%	surface	√	R,N,D,Q,E,G,H,K,S,T
8	C	0.0%	buried	×	none
9	L	64.5%	surface	√	R,N,D,Q,E,G,H,K,S,T,L
10	K	94.2%	surface	√	R,N,D,Q,E,G,H,K,S,T
11	S	52.2%	surface	√	R,N,D,Q,E,G,H,K,S,T
12	G	97.3%	surface	√	R,N,D,Q,E,G,H,K,S,T
13	A	1.8%	buried	√	A,I,L,M,F,Y,W,V
14	I	49.6%	intermediate	√	all except C and P
15	C	18.9%	buried	×	none
16	H	24.9%	intermediate	√	all except C and P
17	P	66.4%	surface	√	R,N,D,Q,E,G,H,K,S,T,P
18	V	79.0%	surface	√	R,N,D,Q,E,G,H,K,S,T,V
19	F	69.1%	surface	√	R,N,D,Q,E,G,H,K,S,T,F
20	C	10.7%	buried	×	none
21	P	32.0%	intermediate	√	all except C
22	R	92.2%	surface	√	R,N,D,Q,E,G,H,K,S,T
23	R	84.6%	surface	√	R,N,D,Q,E,G,H,K,S,T
24	Y	24.7%	intermediate	√	all except C and P
25	K	82.3%	surface	√	R,N,D,Q,E,G,H,K,S,T
26	Q	46.7%	intermediate	√	all except C and P
27	I	42.1%	intermediate	√	all except C and P
28	G	45.8%	intermediate	√	all except C and P
29	T	54.1%	surface	√	R,N,D,Q,E,G,H,K,S,T
30	C	2.6%	buried	×	none
31	G	60.3%	surface	√	R,N,D,Q,E,G,H,K,S,T
32	L	87.1%	surface	√	R,N,D,Q,E,G,H,K,S,T,L
33	P	86.1%	surface	√	R,N,D,Q,E,G,H,K,S,T,P
34	G	96.5%	surface	√	R,N,D,Q,E,G,H,K,S,T
35	T	13.9%	buried	√	A,I,L,M,F,Y,W,V,T
36	K	33.2%	intermediate	√	all except C and P
37	C	0.0%	buried	×	none
38	C	0.0%	buried	×	none
39	K	45.8%	intermediate	√	all except C and P
40	K	61.2%	surface	√	R,N,D,Q,E,G,H,K,S,T
41	P	58.5%	surface	√	R,N,D,Q,E,G,H,K,S,T,P

Biological Constraints

Homology search using PSI-BLAST was executed to determine the conserved properties among hβD-2 homologs. They include charge characteristics and occurrence frequency for each amino acid as tabulated in Table 7 and Table 8.

Table 7. Charge frequencies of the top 97 homologs of human beta defensin-2 from PSI-BLAST.

	Lower Bound	Upper Bound
Net charge on α-helix	0	+3
Total positive charges	5	10
Total negative charges	0	2
Total net charges	+4	+9

Table 8. Occurrence of each amino acid in the top 97 human β defensin-2 homologs from PSI-BLAST.

Amino Acid	Lower Bound	Upper Bound
Ala	0	3
Arg	1	9
Asn	0	6
Asp	0	2
Cys	4	7
Gln	0	3
Glu	0	3
Gly	3	7
His	0	4
Ile	0	6
Leu	0	4
Lys	0	7
Met	0	3
Phe	0	4
Pro	0	5
Ser	0	6
Thr	0	4
Trp	0	1
Tyr	0	4
Val	0	6

These conserved properties were translated into biological constraints (10) and (11). Along with the hydrophobic content constraints (12) and equation

(13) which limits the number of mutations to be no more than ten, 100 sequences were generated using the sequence selection model for single structure (F13). Fold specificity of the sequences was calculated using the approximate fold validation method.

Results

The top 10 out of the 100 sequences ranked by their fold specificities are listed in Table 4.2. The ten or less mutations are preferred to occur at positions 3, 4, 13, 14, 16, 17, 23, 26, 28, 36, 39, with the mutations being: G3T, D4R, A13F, I14V, H16V, P17H, R23(S/D/E), Q26F, G28F, K36(V/F), and K39A.

Fold specificity rank	Varied Positions																			
	3	4	12	13	14	16	17	21	22	23	25	26	28	31	32	34	36	39	40
Native	G	D	G	A	I	H	P	P	R	R	K	Q	G	G	L	G	K	K	K
1	**T**	**R**	G	A	**V**	**V**	**H**	P	**N**	R	K	**F**	**F**	G	L	G	**F**	K	**E**
2	**T**	D	G	**F**	I	**V**	**H**	P	R	**S**	K	**F**	**F**	**R**	L	G	**V**	**A**	K
3	**T**	**R**	G	A	**V**	**V**	**H**	P	**N**	R	K	**F**	**F**	G	L	G	**F**	K	**N**
4	**T**	**R**	G	A	**V**	**V**	**H**	P	R	**D**	K	**F**	**F**	G	L	G	**V**	**F**	K
5	**T**	**R**	G	**F**	I	**V**	**H**	P	**N**	R	K	**F**	**F**	G	L	G	**V**	K	**N**
6	**T**	**R**	G	**V**	I	**V**	**H**	P	R	**E**	K	**F**	**F**	G	L	G	**V**	**F**	K
7	**T**	D	G	**F**	I	**V**	**H**	P	R	**S**	K	**F**	**F**	G	L	**R**	**V**	**A**	K
8	**T**	**R**	G	A	**V**	**V**	**H**	P	R	**E**	K	**F**	**F**	G	L	G	**V**	**F**	K
9	**T**	**R**	G	A	**V**	**V**	**H**	P	R	**E**	K	**F**	**F**	G	L	G	**F**	**Y**	K
10	**T**	**R**	G	A	**F**	**V**	**H**	P	R	**E**	K	**F**	**F**	G	L	G	**V**	**Y**	K

Table 9. Top 10 out of 100 human β defensin-2 sequences selected in stage one ranked by their fold specificities. Mutations are indicated in bold font.

4.3 Complement 3a

Complement 3a (C3a) is a 77-residue small cationic peptide derived from the cleavage of the amino-terminus of the α-chain of complement component C3. It is a potent mediator which controls the pro-inflammatory activities of the complement system, considering that only micro-molar or sub-micro-molar concentrations of C3a are sufficient to elicit its biological effects [62]. Having small molecular size and high potency, C3a proves to be a strong candidate as a superior therapeutic agent with possible routes of administration other than injection, high cost-effectiveness and good pharmacokinetics.

The Flexible Templates for C3a

Three different sets of design templates were used for the de novo design of C3a [115]. One corresponds to the single template structure elucidated by [76],

and the other two were generated using molecular dynamics simulations, one with the generalized Born implicit solvation model and the other with explicit water molecules. Details for the first 12 residues are absent in [76]'s structure, so the de novo design using such template has to be done as if the protein were only 65-residue long. The initial structure for both MD simulations was constructed as a composite of the C3a domain of the newly determined crystal structure of C3 by [94] (PDB code: 2A73) for residues Val^{1}-Ala^{69} and [76]'s crystal structure for residues Ser^{70}-Arg^{76}. In both runs, a PDB structure file was produced every nanosecond for a total of 10 ns along the MD trajectory.

Alignment of the three different sets of structures (see Figure 8) indicates high similarity between the crystal structure and the MD structures generated with explicit water molecules, as well as significant deviation of the MD structures generated with implicit water solvent from the other two template sets.

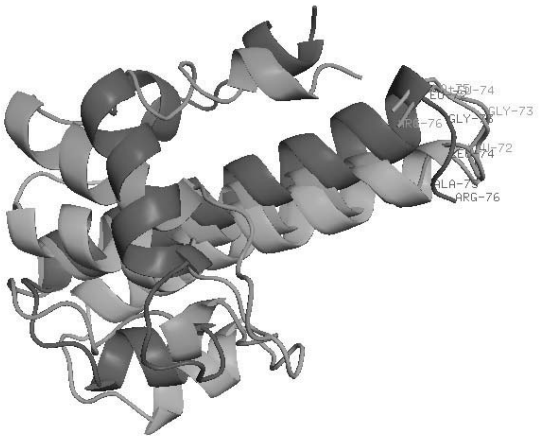


Fig. 8. Structural alignment of crystal structure of C3a [76](cyan), average of the structures from MD simulation with generalized Born implicit solvent model (magenta), and average of the structures from MD simulation with explicit water molecules (green).

The De Novo Design

In silico Sequence Selection Stage

Different models were employed for sequence selection, depending on the nature of the design template set [115]. The basic model for single structure (F13) suffices to perform the job for [76]'s crystal structure, which is singular. For the two other sets of MD templates, both the weighted average

forcefield formulation (F14) and the binary distance bin variable formulation (F15), which were tailored for handling templates with multiple structures, were utilized here for sequence selection.

Since the file for [76]'s crystal structure lacks side-chain information, only the high resolution C^{α}-C^{α} forcefield [87] was used for sequence selection based on such template. Both the high resolution C^{α}-C^{α} forcefield and centroid-centroid forcefield were employed for the other two MD template sets.

The current binding model of C3a to its receptor C3aR suggests interaction between the carboxy-terminus (residues 69 to 77) of the protein and the extracellular loop connecting helix 4 and helix 5 of the 7-helix transmembrane C3aR receptor. [62] did extensive sequence-activity study on C3a analogs based on mutations of the positions from 63 through 72. Residues 73 to 77, which have a wild-type sequence of LGLAR, were highly conserved among the C3a molecules obtained from different animal species [62] and considered to constitute the primary binding site of C3a to C3aR. Therefore they were not mutated in [62]'s work. Taking into account of their importance and the availability of knowledge about the effect of their mutations , residues 63 to 72 were chosen as the target of the de novo design presented here.

Positions 63 to 72 have a natural sequence of LRRQHARASH. As previously mentioned, positions 63 to 69 adopt an α-helical structure. Being highly conserved in the C3a consensus sequence like the carboxy-terminal segment LGLAR, the arginine at position 69 was deemed vital for functionality and thus treated as a non-varied position in stage one [62]. The mutation set for other positions is summarized in Table 10. It was devised based on the nature of the native residues, as well as the suggestions from [62]. In particular, the inclusion of the hydrophobic set for positions 63 and 64 was in agreement with the speculation of [62] that hydrophobic residues with bulky groups (e.g., tryptophan) at those positions were crucial for enhanced potency in C3a analogs.

Table 10. Mutation set of *in silico* sequence selection of C3a.

Positions	Native residue	Allowed mutations
63	L	A,I,L,M,F,Y,W,V
64	R	all except C and P
65	R	R,N,D,Q,E,G,H,K,S,T
66	Q	R,N,D,Q,E,G,H,K,S,T
67	H	R,N,D,Q,E,G,H,K,S,T
68	A	all except C and P
70	A	R,N,D,Q,E,G,H,K,S,T
71	S	R,N,D,Q,E,G,H,K,S,T
72	H	R,N,D,Q,E,G,H,K,S,T

As an effort to improve the quality of the sequences from stage one, linear constraints were added to the models to enforce the portion of position 63 to position 69 to possess the native charge of +3. These charge constraints take the form of:

$$\sum_i y_i^{Arg} + \sum_i y_i^{Lys} - \sum_i y_i^{Asp} - \sum_i y_i^{Glu} = 3 \quad \forall\, 63 \leq i \leq 69 \quad (19)$$

With the mutation set and the charge constraints mentioned above, 500 sequence solutions were generated using the high resolution C^{α}-C^{α} forcefield based on the singular template. For the two other sets of MD templates, 500 sequences were produced for each combination of the two high resolution forcefields and the two sequence selection formulations for flexible templates with multiple structures. All these sequences were processed through stage two for the fold specificity calculation.

Fold Specificity Stage

For the de novo design of C3a, the angle bounds and distance bounds input to the CYANA 2.1 package were $\pm 10^o$ around the template and $\pm 10\%$ of those in the template respectively for the sequences from [76]'s crystal structure, and the maximum and minimum of the corresponding angles and distances across all template structures for the sequences from the MD flexible templates [115]. 500 low energy conformations were generated for each sequence by the simulated annealing algorithm in CYANA and their energies were minimized further by the TINKER program. Finally, the fold specificity for each sequence was computed using formula (14), and the sequences were then ranked according to their specificities.

Results and Discussion

Since the results from stage one are only from a distance-dependent residue-residue based forcefield, whereas those from stage two are derived from the full-atomistic forcefield AMBER, the latter should be focused for data analysis. For the crystal structure template, the templates from MD simulations with implicit solvent, and the templates from MD simulations with explicit water molecules, the favored residues in the top 20 sequences ranked according to their fold specificities are tabulated in Table 11, Table 12, and Table 13 respectively.

Results Based on the Single Template from X-ray Crystallography

Amino acid selections based on fold specificity in this set of results are:
-(F/Y/W)^{63}-R^{64}-T^{65}-R^{66}-(T/Q)^{67}-(W/Y/Q)^{68}-R^{69}-(T/E)^{70}-(E/T/R)^{71}
-(E/N/T/Q)^{72}-, with multiple selections in the same position indicated by following their order of preference. As [76]'s structure file lacks side-chain information, results could only be generated using the high resolution C^{α}-C^{α} forcefield.

Results Based on the Templates from MD Simulations with Implicit Solvent

- High resolution C^{α}-C^{α} forcefield
  For the weighted average forcefield model, results indicate a pattern of: -$(F/L/W)^{63}$-$(F/I/Y)^{64}$-$(R/T)^{65}$-$(T/N)^{66}$-R^{67}-$(F/Y)^{68}$-R^{69}-E^{70}-$(T/Q/N)^{71}$-E^{72}-, which is quite different from the one shown by the results generated based on [76]'s crystal structure; the difference is attributed to the structural deviation between the two sets of templates as shown in Figure 8.
  With the binary distance bin variable formulation, the pattern becomes: -$(F/L)^{63}$-$(F/I/Y)^{64}$-R^{65}-$(T/N)^{66}$-R^{67}-$(F/Y)^{68}$-R^{69}-E^{70}-$(T/Q/N)^{71}$-E^{72}-, which is almost exactly the same as that from the weighted average formulation.
- High resolution centroid-centroid forcefield
  Results generated using the weighted average forcefield formulation take the form of: -W^{63}-W^{64}-$(R/Q)^{65}$-R^{66}-$(N/Q/S/R)^{67}$-W^{68}-R^{69}-$(D/E)^{70}$-$(N/E/Q)^{71}$-$(Q/R/S/N)^{72}$-. By comparing them with those corresponding to C^{α}-C^{α} forcefield, apparently the two forcefields lead to significantly different preferences. It should be highlighted that the centroid-centroid forcefield predicts TRP for positions 63 and 64, which is in agreement with what [62] proposed for synthesizing super-potent C3a analogs.
  With the binary distance bin variable formulation for stage one, the fold specificity preferences become: -W^{63}-W^{64}-$(R/Q)^{65}$-$(R/N)^{66}$-$(Q/R/T/S)^{67}$-W^{68}-R^{69}-$(E/D)^{70}$-$(E/N/Q)^{71}$-$(R/Q)^{72}$-. Again this shows high level of agreement with the weighted average formulation results, and [62]'s suggestion of TRP for positions 63 and 64 is in place.

Results Based on the Templates from MD Simulations with Explicit Water Molecules

- High resolution C^{α}-C^{α} forcefield
  Using the weighted average forcefield model, fold specificity results are: -F^{63}-R^{64}-$(T/N/R)^{65}$-$(R/T)^{66}$-$(T/R)^{67}$-$(F/W/Y)^{68}$-R^{69}-$(T/A)^{70}$-$(E/T/Q/R)^{71}$-$(Q/E/T)^{72}$-. These results are highly similar to those based on the single crystal structure template, due to the structural resemblance of the two template sets.
  If the binary distance bin variable formulation was used instead, the amino acid preferences are still roughly the same: -$(F/A)^{63}$-$(R/F/W)^{64}$-$(T/R/N)^{65}$-$(R/T)^{66}$-$(T/R)^{67}$-$(W/F/Y)^{68}$-R^{69}-$(T/E/Q)^{70}$-$(E/T/Q)^{71}$-$(E/Q/N/T)^{72}$-.
- High resolution centroid-centroid forcefield
  Results from the weighed average sequence selection model are: -$(L/W)^{63}$-$(W/L/R/Y)^{64}$-$(R/E/D)^{65}$-$(R/S)^{66}$-$(K/H/Q)^{67}$-$(L/W)^{68}$-R^{69}-E^{70}-$(E/D/Q)^{71}$-$(H/Q/K)^{72}$-.

If the binary distance bin variable formulation was used instead for stage one, the fold specificity preferences become: -W^{63}-W^{64}-R^{65}-$(R/S)^{66}$-$(Q/S/H/T/K)^{67}$-W^{68}-R^{69}-$(E/D)^{70}$-$(E/D/Q)^{71}$-$(H/Q/K)^{72}$-. Interesting enough, despite the structural deviation between the two sets of MD templates, these results still show high degree of similarity to their counterparts based on the flexible templates from MD simulations with implicit solvent, except for position 72. They are also consistent with [62]'s suggestion for a super-potent peptide, which is 12-15 times more active than natural C3a.

A significant overlap between the solution space of the weighted average formulation and that of the binary distance bin variable formulation was observed through all runs. This implies the weighted average model can be used as an approximate method to substitute for the distance bin variable model on sequence selection problems with high complexity, as the former contains fewer variables and is thus more computationally efficient. However, for problems where the distance bin variable formulation can perform reasonably well, it is advisable to apply both models, since the distance bin variable model was observed to produce better fold specificity results than the weighted average model.

Table 11. Favored residues in the top 20 C3a sequences ranked by their fold specificities. Mutations versus the native sequence are highlighted in bold font. The topmost residue at each position corresponds to the one with the highest occurrence frequency in the 20 sequences.

Fold Specificity Results based on Template of [76]'s Crystal Structure									
Positions									
63	64	65	66	67	68	69	70	71	72
F	R	**T**	**R**	**T**	**W**	R	**T**	**E**	**E**
Y				**Q**	**Y**		**E**	**T**	**N**
W					**Q**			**R**	**Q**
									T

Table 12. Favored residues in the top 20 C3a sequences ranked by their fold specificities. Mutations versus the native sequence are highlighted in bold font. The topmost residue at each position corresponds to the one with the highest occurrence frequency in the 20 sequences.

Fold Specificity Results based on Flexible Templates from MD simulation with Generalized Born Implicit Solvent Model																				
High Resolution C^{α}-C^{α} Forcefield											**High Resolution Centroid-Centroid Forcefield**									
Weighted Average Forcefield Formulation											Weighted Average Forcefield Formulation									
Positions											Positions									
63	64	65	66	67	68	69	70	71	72		63	64	65	66	67	68	69	70	71	72
L	R	R	Q	H	A	R	A	S	H		L	R	R	Q	H	A	R	A	S	H
F	**F**	R	**T**	**R**	**F**	R	**E**	**T**	**E**		**W**	**W**	R	**R**	**N**	**W**	R	**D**	**N**	**Q**
L	**I**	**T**	**N**		**Y**			**Q**					**Q**		**Q**			**E**	**E**	**R**
W	**Y**							**N**							**S**				**Q**	**N**
															R					**S**
Binary Distance Bin Variable Formulation											Binary Distance Bin Variable Formulation									
Positions											Positions									
63	64	65	66	67	68	69	70	71	72		63	64	65	66	67	68	69	70	71	72
L	R	R	Q	H	A	R	A	S	H		L	R	R	Q	H	A	R	A	S	H
F	**F**	R	**T**	**R**	**F**	R	**E**	**T**	**E**		**W**	**W**	R	**R**	**Q**	**W**	R	**E**	**E**	**R**
L	**I**		**N**		**Y**			**Q**					**Q**	**N**	**R**			**D**	**N**	**Q**
	Y							**N**							**S**				**Q**	
															T					

5 Conclusions

In this paper, we reviewed our previous work on de novo protein design done by [1, 2] and presented our advances. The improvements are: (1) speeding up both the sequence selection stage and fold validation stage, and (2) incorporating higher degree of true protein backbone flexibility [88] by developing models that map sequences onto a flexible template with multiple structures. Extensive predictive results on compstatin, human β defensin-2, and C3a are discussed.

Acknowledgement. We gratefully acknowledge financial support from the National Science Foundation (CAF and DM), the National Institutes of Health (R01 GM52032: CAF; GM069736: CAF and DM), and the US Environmental Protection Agency (GAD R 832721-010: CAF). This work has not been reviewed by and does not represent the opinions of USEPA.

Table 13. Favored residues in the top 20 C3a sequences ranked by their fold specificities. Mutations versus the native sequence are highlighted in bold font. The topmost residue at each position corresponds to the one with the highest occurrence frequency in the 20 sequences.

Fold Specificity Results based on Flexible Templates from MD simulation with Explicit Water Molecules

High Resolution C^{α}-C^{α} Forcefield									
Weighted Average Forcefield Formulation									
Positions									
63	64	65	66	67	68	69	70	71	72
L	R	R	Q	H	A	R	A	S	H
F	R	**T**	**R**	**T**	**F**	R	**T**	**E**	**E**
		N	**T**	**R**	**Y**		A	**T**	**Q**
		R			**W**			**R**	**T**
								Q	
Binary Distance Bin Variable Formulation									
Positions									
63	64	65	66	67	68	69	70	71	72
L	R	R	Q	H	A	R	A	S	H
F	R	**T**	**R**	**T**	**W**	R	**T**	**E**	**E**
A	**F**	R	**T**	**R**	**F**		**E**	**T**	**Q**
	W	**N**			**Y**		**Q**	**Q**	**N**
									T

High Resolution Centroid-Centroid Forcefield									
Weighted Average Forcefield Formulation									
Positions									
63	64	65	66	67	68	69	70	71	72
L	R	R	Q	H	A	R	A	S	H
L	**W**	R	**R**	**K**	**L**	R	**E**	**E**	H
W	**L**	**D**	**S**	H	**W**			**D**	**Q**
	R	**E**		**Q**				**Q**	**K**
	Y								
Binary Distance Bin Variable Formulation									
Positions									
63	64	65	66	67	68	69	70	71	72
L	R	R	Q	H	A	R	A	S	H
W	**W**	R	**R**	**Q**	**W**	R	**E**	**E**	H
			S	**S**			**D**	**D**	**Q**
				H				**Q**	**K**
				K					
				T					

References

1. Klepeis, J.L., Floudas, C.A., Morikis, D., Tsokos, C.G., Argyropoulos, E., Spruce, L., Lambris, J.D.: Integrated structural, computational and experimental approach for lead optimization: deisgn of compstatin variants with improved activity. J. Am. Chem. Soc., **125**, 8422–8423 (2003)
2. Klepeis, J.L., Floudas, C.A., Morikis, D., Tsokos, C.G., Lambris, J.D.: Design of peptide analogs with improved activity using a novel de novo protein design approach. Ind. Eng. Chem. Res., **43**, 3817–3826 (2004)
3. Jin, W., Kambara, O., Sasakawa, H., Tamura, A., Takada, S.: De novo design of foldable proteins with smooth folding funnel: automated negative design and experimental verification. Structure, **11**, 581–590 (2003)
4. Drexler, K.E.: Molecular engineering: an approach to the development of general capabilities for molecular manipulation. Proc. Natl. Acad. Sci. USA, **78**, 5275–5278 (1981)
5. Pabo, C.: Molecular technology: designing proteins and peptides. Nature, **301**, 200 (1983)

6. Hardin, C., Pogorelov, T.V., Luthey-Schulten, Z.: Ab initio protein structure prediction. Curr. Opin. Struc. Biol., **12**, 176–181 (2002)
7. Moore, J.C., Arnold, F.H.: Directed evolution of a para-nitrobenzyl esterase for aqueous-organic solvents. Nat. Biotechnol., 14, 458–467 (1996)
8. Voigt, C.A., Mayo, S.L., Arnold, F.H., Wang, Z-G.: Computational method to reduce the search space for directed protein evolution. Proc. Natl. Acad. Sci. USA, 98, 3778–3783 (2001)
9. Skandalis, A., Encell, L.P., Loeb, L.A.: Creating novel enzymes by applied molecular evolution. Chem. Biol., 4, 889–898 (1997)
10. Ponder, J.W., Richards, F.M.: Tertiary templates for proteins. J. Mol. Biol., **193**, 775–791 (1987)
11. Desjarlais, J.R., Handel, T.M.: De novo design of the hydrophobic cores of proteins. Protein Sci., **4**, 2006–2018 (1995)
12. Koehl, P., Levitt, M.: De novo protein design I. In search of stability and specificity. J. Mol. Biol., **293**, 1161–1181 (1999)
13. Voigt, C.A., Gordon, D.B., Mayo, S.L.: Trading accuracy for speed: a quantitative comparison of search algorithms in protein sequence design. J. Mol. Biol., **299**, 789–803 (2000)
14. Desjarlais, J.R., Handel, T.M.: Side chain and backbone exibility in protein core design. J. Mol. Biol., **290**, 305–318 (1999)
15. Desmet, J., De Maeyer, M., Hazes, B., Lasters, I.: The dead-end elimination theorem and its use in side-chain positioning. Nature, **356**, 539–542 (1992)
16. Dahiyat, B.I., Mayo, S.L.: De novo protein design: fully automated sequence selection. Science, **278**, 82–87 (1997)
17. Tobi, D., Elber, R.: Distance-dependent pair potential for protein folding: results from linear optimization. Proteins: Structure, Function, and Bioinformatics, **41**, 40–46 (2000)
18. Tobi, D., Shafran, G., Linial, N., Elber, R.: On the design and analysis of protein folding potentials. Proteins: Structure, Function, and Bioinformatics, **40**, 71–85 (2000)
19. Loose, C., Klepeis, J.L., Floudas, C.A.: A new pairwise folding potential based on improved decoy generation and side chain packing. Proteins: Structure, Function, and Bioinformatics, **54**, 303–314 (2004)
20. CPLEX: Using the CPLEX Callable Library. ILOG, Inc. Mountain View, California (1997)
21. Sherali, H.D., Adams, W.P.: A Reformulation Linearization Technique for Solving Discrete and Continuous Nonconvex Problems. Kluwer Academic Publishing, Boston (1999)
22. Floudas, C.A.: Nonlinear and Mixed-Integer Optimization: Fundamentals and Applications. Oxford University Press, New York (1995)
23. Klepeis, J.L., Schafroth, H.D., Westerberg, K.M., Floudas, C.A.: Deterministic global optimization and ab initio approaches for the structure prediction of polypeptides, dynamics of protein folding and proteinprotein interaction. In: Friesner, R.A. (ed) Advances in Chemical Physics. Wiley, New York (2002)
24. Klepeis, J.L., Floudas, C.A., Morikis, D., Lambris, J.D.: Predicting peptide structures using NMR data and deterministic global optimization. J. Comput. Chem., **20**, 1354–1370 (1999)
25. Klepeis, J.L., Floudas, C.A.: Ab initio tertiary structure prediction of proteins. J. Global. Optim., **25**, 113–140 (2003)

26. Némethy, G., Gibson, K.D., Palmer, K.A., Yoon, C.N., Paterlini, G., Zagari, A., Rumsey, S., Scheraga, H.A.: Energy parameters in polypeptides. 10. J. Phys. Chem., **96**, 6472–6484 (1992)
27. Floudas, C. A.: Deterministic Global Optimization : Theory, Methods and Applications. Kluwer Academic Publishers, New York (2000)
28. Adjiman, C., Androulakis, I., Floudas, C.A.: A global optimization method, αBB, for general twice-differential constrained NPLs - I. Theoretical advances. Computers Chem. Engng., **22**, 1137–1158 (1998)
29. Adjiman, C., Androulakis, I., Floudas, C.A.: A global optimization method, αBB, for general twice-differentiable constrained NLPs - II. Implementation and computational results. Computers Chem. Engng., **22**, 1159–1179 (1998)
30. Adjiman, C., Androulakis, I., Floudas, C.A.: Global optimization of mixed-integer nonlinear problems. AIChE Journal, **46**, 1769–1797 (2000)
31. Sahu, A., Lambris, J. D.: Structure and biology of complement protein C3, a connecting link between innate and acquired immunity. Immunol. Rev., **180**, 35–48 (2001)
32. Sahu, A., Kay, B.K., Lambris, J.D.: Inhibition of human complement by a C3-binding peptide isolated from a phage displayed random peptide library. J. Immunol., **157**, 884–891 (1996)
33. Sahu, A., Soulika, A.M., Morikis, D., Spruce, L., Moore, W.T., Lambris, J.D.: Binding kinetics, structure activity relationship and biotransformation of the complement inhibitor compstatin. J. Immunol., **165**, 2491– 2499 (2000)
34. Morikis, D., Roy, M., Sahu, A., Torganis, A., Jennings, P.A., Tsokos, G.C., Lambris, J.D.: The structural basis of compstatin activity examined by structure-function-based design of peptide analogs and NMR. J. Biol. Chem., **277**, 14942–14953 (2002)
35. Soulika, A.M., Morikis, D., Sarias, M.R., Roy, M., Spruce, L., Sahu, A., Lambris, J.D.: Studies of structure-activity relations of complement inhibitor compstatin. J. Immunology, **170**, 1881–1890 (2003)
36. Soulika, A.M., Khan, M.M., Hattori, T., Bowen, F.W., Richardson, B.A., Hack, C.E., Sahu, A., Edmunds, L.H., Lambris, J.D.: Inhibition of heparin/ protamine complex-induced complement activation by comsptatin in baboons. Clin. Immunology, **96**, 212–221 (2000)
37. Nilsson, B., Larsson, R., Hong, J., Elgue, G., Ekdahl, K.N., Sahu, A., Lambris, J.D.: Compstatin inhibits complement and cellular activation in whole blood in two models of extracorporeal circulation. Blood, **92**, 1661–1667 (1998)
38. Fiane, A.E., Mollnes, T.E., Videm, V., Hovig, T., Hogasen, K., Mellbye, O.J., Spruce, L., Moore, W.T., Sahu, A., Lambris, J.D.: Compstatin, a peptide inhibitor of C3, prolongs survival of ex-vivo perfused pig xenografts. Xenotransplantation, **6**, 52–65 (1999)
39. Mollnes, T.E., Brekke, O.L., Fung, M., Fure, H., Christiansen, D., Bergseth, G., Videm, V., Lappegard, K.T., Kohl, J., Lambris, J.D.: Essential role of the C5a receptor in E coli-induced oxidative burst and phagocytosis revealed by a novel lepirudin-based human whole blood model of in ammation. Blood, **100**, 1869–1877 (2002)
40. Klegeris, A., Singh, E.A., McGeer, P.L.: Effects of c-reactive protein and pentosan polysulphate on human complement activation. Immunology, **106**, 381–388 (2002)

41. Sahu, A., Morikis, D., Lambris, J.D.: Compstatin, a peptide inhibitor of complement, exhibits species-specific binging to complement component c3. Mol. Immunology, **39**, 557–566 (2003)
42. Gordon, B.B., Hom, G.K., Mayo, S.L., Pierce, N.A.: Exact rotamer optimization for protein design. J. Comput. Chem., **24**, 232–243 (2003)
43. Pierce, N.L., Spriet, J.A., Desmet, J., Mayo, S.L.: Conformational splitting: a more powerful criterion for dead-end elimination. J. Comput. Chem., **21**, 999–1009 (2000)
44. Zou, J.M., Saven, J.G.: Statistical theory of combinatorial libraries of folding proteins: energetic discrimination of a target structure. J. Mol. Bio., **296**, 281–294 (2000)
45. Kuhlman, B., O'Neill, J.W., Kim, D.E., Zhang, K.Y.J., Baker, D.: Accurate computer-based design of a new backbone conformation in the second turn of protein 1. J. Mol. Bio., **315**, 471–477 (2002)
46. Kuhlman, B., Dantae, G., Ireton, G.C., Verani, G., Stoddard, B., Baker, D.: Design of a novel globular protein fold with atomic-level accuracy. Science, **302**, 1364–1368 (2003)
47. Kuhlman, B., Baker, D.: Native Protein Sequences Are Close to Optimal for Their Structures. Proc. Natl. Acad. Sci. USA, **97**, 10383–10388 (2000)
48. Dantas, G., Kuhlman, B., Callender, D., Wong, M., Baker, D.: A large scale test of computational protein design: folding and stability of nine completely redesigned globular proteins. J. Mol. Biol., **332**, 449–460 (2003)
49. Watters, A.L., Baker, D.: Searching for folded proteins in vitro and in silico. Eur. J. Biochem., **271**, 1615–1622 (2004)
50. Kuhlman, B., Baker, D.: Exploring folding free energy landscapes using computational protein design. Current Opinion in Structural Biology, **14**, 89–95 (2004)
51. Kortemme, T., Baker, D.: Computational design of protein-protein interactions. Current Opinion in Chemical Biology, **8**, 91–97 (2004)
52. Benson, D.E., Wisz, M.S., Hellinga, H.W.: Rational design of nascent metalloenzymes. Proc. Natl. Acad. Sci. USA, **97**, 6292–6297 (2000)
53. Goldstein, R.F.: Effcient rotamer elimination applied to protein sidechains and related spin glasses. Biophysics Journal, **66**, 1335–1340 (1994)
54. Looger, L.L., Dwyer, M.W., Smith, J.J., Hellinga, H.W.: Computational design of receptor and sensor proteins with novel functions. Nature, **423**, 185–190 (2003)
55. Richards, F.M., Hellinga, H.W.: Optimal sequence selection in proteins of known structure by simulated evolution. Proc. Natl. Acad. Sci. USA, **91**, 5803–5807 (1994)
56. Richards, F.M., Hellinga, H.W.: Construction of new ligand binding sites in proteins of known structure. I. Computer-aided modeling of sites with predefined geometry. J. Mol. Biol., **222**, 763–785 (1991)
57. Richards, F.M., Caradonna, J.P., Hellinga, H.W.: Construction of new ligand binding sites in proteins of known structure. II. Grafting of a buried transition metal binding site into Escherichia coli thioredoxin. J. Mol. Biol., **222**, 787–803 (1991)
58. Yang, W., Jones, L.M., Isley, L., Ye, Y., Lee, H-W., Wilkins, A., Liu, Z.R., Hellinga, H.W., Malchow, R., Ghazi, M., Yang, J.J.: Rational design of a calcium-binding protein. J. Am. Chem. Soc., **125**, 6165–6171 (2003)

59. Kraemer-Pecore, C.M., Wollacott, A.M., Desjarlais, J.R.: Computational protein design. Current Opinion in Chemical Biology, **5**, 690–695 (2001)
60. Kraemer-Pecore, C.M., Lecomte, J.T., Desjarlais, J.R.: A de novo redesign of the ww domain. Protein Science., **12**, 2194–2205 (2003)
61. Lim, W.A., Hodel, A., Sauer, R.T., Richards, F.M.: The crystal structure of a mutant protein with altered but improved hydrophobic core packing. Proc. Natl. Acad. Sci. USA, **91**, 423–427 (1994)
62. Ember, J.A., Johansen, N.L., Hugli, T.E.: Designing synthetic superagonists of c3a anaphylatoxin. Biochemistry, **30**, 3603–3612 (1991)
63. Dahiyat, B.I., Mayo, S.L.: Protein design automation. Protein Science, **5**, 895–903 (1996)
64. Su, A., Mayo, S.L.: Coupling backbone exibility and amino acid sequence selection in protein design. Protein Science, **6**, 1701–1707 (1997)
65. Malakauskas, S.M., Mayo, S.L.: Design, structure, and stability of a hyperthermophilic protein variant. Nat. Struct. Biol., **5**, 470–475 (1998)
66. Shimaoka, M., Shifman, J.M., Jing, H., Takagi, L., Mayo, S.L., Springer, T.A.: Computational design of an intergrin I domain stabilized in the open high affinity conformation. Nat. Struct. Biol., **7**, 674–678 (2000)
67. Mooers, B.H.M., Datta, D., Baase, W.A., Zollars, E.S., Mayo, S.L., Matthews, B.W.: Repacking the core of T4 lysozyme by automated design. J. Mol. Biol., **332**, 741–756 (2003)
68. Gillespie, B., Vu, D.M., Shah, P.S., Marshall, S.A., Dyer, R.B., Mayo, S.L., Plaxco, K.W.: NMR and temperature-jump measurements of de novo designed proteins demonstrate rapid folding in the absence of explicit selection for kinetics. J. Mol. Biol., **330**, 813–819 (2003)
69. Zhu, Y., Alonso, D.O., Maki, K., Huang, C.Y., Lahr, S.J., Daggett, V., Roder, H., DeGrado, W.F., Gai, F.: Ultrafast folding of alpha3D: a de novo designed three-helix bundle protein. Proc. Natl. Acad. Sci. USA, **100**, 15486–15491 (2003)
70. Kono, H., Saven, J.G.: Statistical theory for protein combinatorial libraries. Packing interactions, backbone exibility, and the sequence variability of a main-chain structure. J. Mol. Biol., **306**, 607–628 (2001)
71. Park, S., Yang, X., Saven, J.G.: Advances in computational protein design. Current Opinion in Structural Biology, **14**, 487–494 (2004)
72. Pokala, N., Handel, T.M.: Review: protein design-where we were, where we are, where we're going. Journal of Structural Biology, **134**, 269–281 (2001)
73. Dill, K.A.: Dominant forces in protein folding. Biochemisty, 29, 7133–7155 (1990)
74. Lee, C.: Predicting protein mutant energetics by self-consistent ensemble optimization. J. Mol. Biol., **236**, 918–939 (1994)
75. Klepeis, J.L., Floudas, C.A.: Free energy calculations for peptides via deterministic global optimization. J. Chem. Phys., **110**, 7491 (1999)
76. Huber, R., Scholze, H., Paques, E.P., Deisenhofer, J.: Crystal structure analysis and molecular model of human c3a anaphylatoxin. Hoppe-Seylers Z Physiol Chemie, **361**, 1389–1399 (1980)
77. Tuffery, P., Etchebest, C., Hazout, S., Lavery, R.: A new approach to the rapid determination of protein side chain conformations. J. Biomol. Struct. Dyn., **8**, 1267–1289 (1991)
78. Wilson, C., Mace, J.E., Agard, D.A.: Computational method for the design of enzymes with altered substrate specificity. J. Mol. Biol., **220**, 495–506 (1991)

79. Farinas, E., Regan, L. The de novo design of a rubredoxin-like Fe site. Protein Science, **7**, 1939–1946 (1998)
80. O. Prokopyev and H.X. Huang and P.M. Pardalos: Multi-quadratic Binary Programming. University of Florida, Research Report (2004)
81. Oral, M., Kettani, O.: A linearization procedure for quadratic and cubic mixed-integer problems. Operations Research, **40**, S109–S116 (1990)
82. Oral, M., Kettani, O.: Reformulating nonlinear combinatorial optimization problems for higher computational efficiency. European Journal of Operational Research, **58**, 236–249 (1992)
83. Pierce, N.A., Winfree, E.: Protein design is np-hard. Protein Engineering, **15**, 779–782 (2002)
84. Mallik, B., Katragadda, M., Spruce, L.A., Carafides, C., Tsokos, C.G., Morikis, D., Lambris, J.D.: Design and nmr characterization of active analogues of compstatin containing non-natural amino acids. Journal of Medicinal Chemistry, **48**, 274–286 (2005)
85. Fung, H.K., Rao, S., Floudas, C.A., Prokopyev, O., Pardalos, P.M., Rendl, F.: Computational comparison studies of quadratic assignment like formulations for the in silico sequence selection problem in de novo protein design. J. Comb. Optim., **10**, 41–60 (2005)
86. Saunders, C.T., Baker, D.: Recapitulation of protein family divergence using exible backbone protein design. J. Mol. Biol., **346**, 631–644 (2005)
87. Rajgaria, R., McAllister, S.R., Floudas, C.A.: Development of a novel high resolution calpha-calpha distance dependent force field using a high quality decoy set. Proteins: Structure, Function, and Bioinformatics, accepted for publication (2006)
88. Floudas, C.A.: Research challenges, opportunities and synergism in systems engineering and computational biology. AIChE Journal, **51**, 1872–1884 (2005)
89. Fung, H.K., Taylor, M.S., Floudas, C.A.: Novel formulation for the sequence selection problem in de novo protein design with exible templates. Optim. Methods & Software, in print (2006)
90. Guntert, P., Mumenthaler, C., Wuthrich, K.: Torsion angle dynamics for nmr structure calculation with the new program DYANA. J. Mol. Bio., **273**, 283–298 (1997)
91. Guntert, P.: Automated nmr structure calculation with CYANA. J. Mol. Bio., **278**, 353–378 (2004)
92. Ponder, J.: TINKER, software tools for molecular design. 1998. Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine. St. Louis, MO. (1998)
93. Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W., Kollman, P.A.: A 2nd generation force-field for the simulation Of proteins, nucleic-Acids, and organic-molecules. J. Am. Chem. Soc., **117**, 5179–5197 (1995)
94. Janssen, B.J.C., Huizinga, E.G., Raaijmakers, H.C.A., Roos, A., Daha, M.R., Nilsson-Ekdahl, K., Nilsson, B., Gros, P.: Structures of complement component C3 provide insights into the function and evolution of immunity. Nature, **437**, 505–511 (2005)
95. Dwyer, M.A., Looger, L.L., Hellinga, H.W.: Computational design of a biologically active enzyme. Science, **304**, 1967–1971 (2004)
96. Dwyer, M.A., Hellinga, H.W.: Periplasmic binding proteins: a versatile superfamily for protein engineering. Curr. Opin. Struc. Biol., **14**, 495–504 (2004)

97. Swift, J., Wehbi, W.A., Kelly, B.D., Stowell, X.F., Saven, J.G., Dmochowski, I.J.: Design of functional ferritin-like proteins with hydrophobic cavities. J. Am. Chem. Soc., **128**, 6611–6619 (2006)
98. Bunagan, M.R., Yang, X., Saven, J.G., Gai, F.: Ultrafast folding of a computationally designed trp-cage mutant: trp-cage. J. Phys. Chem. B., **110**, 3759–3763 (2006)
99. Cochran, F.V., Wu, S.P., Wang, W., Nanda, V., Saven, J.G., Therien, M.J., DeGrado, W.F.: Computational de novo design and characterization of a four-helix bundle protein that selectively binds a nonbiological cofactor. J. Am. Chem. Soc., **127**, 1346–1347 (2005)
100. Sood, V.D., Baker, D.: Recapitulation and design of protein binding peptide structures and sequences. J. Mol. Biol., **357**, 917–927 (2006)
101. Korkegian, A., Black, M.E., Baker, D., Stoddard, B.L.: Computational thermostabilization of an enzyme. Science, **308**, 857–860 (2005)
102. Lazar, G.A., Marshall, S.A., Plecs, J.J., Mayo, S.L., Desjarlais, J.R.: Designing proteins for therapeutic applications. Curr. Opin. Struc. Biol., **13**, 513–518 (2003)
103. Shukla, U.J., Marino, H., Huang, P., Mayo, S.L., Love, J.J.: A designed protein interface that blocks fibril formation. J. Am. Chem. Soc., **126**, 13914–13915 (2004)
104. Song, G., Lazar, G.A., Kortemme, T., Shimaoka, M., Desjarlais, J.R., Baker, D., Springer, T.A. Rational design of intercellular adhesion molecule-1 (ICAM-1) variants for antagonizing integrin lymphocyte function-associated antigen-1-dependent adhesion. J. Biol. Chem., **281**, 5042–5049 (2006)
105. Glover, F.: Improved linear integer programming formulations of nonlinear integer problems. Management Science, **22**, 455–460 (1975)
106. Floudas, C.A., Fung, H.K.: Mathematical modeling and optimization methods for de novo protein design. In: Rigoutsos, I., Stephanopoulos, G. (eds) Systems Biology II. Oxford University, New York, NY (2006)
107. Chan, D.C., Fass, D., Berger, J.M., Kim, P.S.: Core structure of gp41 from the HIV envelope glycoprotein. Cell, **89**, 263–273 (1997)
108. Malashkevich, V.N., Chan, D.C., Chutkowski, C.T., Kim, P.S.: Crystal structure of the simian immunodeficiency virus (SIV) gp41 core: conserved helical interactions underlie the broad inhibitory activity of gp41 peptides. Proc. Natl. Acad. Sci., **95**, 9134–9139 (1998)
109. Baritaki, S., Dittmar, M.T., Spandidos, D.A., Krambovitis, E.: In vitro inhibition of R5 HIV-1 infectivity by X4 V3-derived synthesis peptides. International Journal of Molecular Medicine, **16**, 333–336 (2005)
110. Bagnarelli, P., Fiorelli, L., Vecchi, M., Monachetti, A., Menzo, S., Clementi, M.: Analysis of the functional relationship between v3 loop and gp120 conext with regards to human immunodeficiency virus coreceptor usage using naturally selected sequences and different viral backbones. Virology, **307**, 328–340 (2003)
111. Galanakis, P.A., Spyroulias, G.A., Rizos, A., Samolis, P., Krambovitis, E. Conformational properties of HIV-1 gp120/v3 immunogenic domains. Current Medicinal Chemistry, **12**, 1551–1568 (2005)
112. Huang, C., Tang, M., Zhang, M., Majeed, S., Montabana, E., Stanfield, R.L., Dimitrov, D.S., Korber, B., Sodroski, J., Wilson, I.A., Wyatt, R., Kwong, P.D.: Structure of a v3-containing HIV-1 gp120 core. Science, **310**, 1025–1028 (2005)
113. Zolla-Pazner, S.: Identifying epitopes of HIV-1 that induce protective antibodies. Nature Reviews Immunology, **4**, 199–210 (2004)

114. Sia, S.K., Carr, P.A., Cochran, A.G., Malashkevich, V.N., Kim, P.S.: Short constrained peptides that inhibit HIV-1 entry. PNAS, **99**, 14664–14669 (2002)
115. Fung, H.K., Taylor, M.S., Floudas, C.A., Morikis, D., Lambris, J.D.: Redesigning complement 3a based on exible templates from both xray crystallography and molecular dynamics simulation. In preparation (2006)
116. Hoover, D.M., Rajashankar, K.R., Blumenthal, R., Puri, A., Oppenheim, J.J., Chertov, O., Lubkowski, J.: The structure of human β-defensin- 2 shows evidence of higher order oligomeration. J. Biol. Chem., **275**, 32911–32918 (2000)
117. García, J.R.C., Florian, J., Schulz, S., Krause, A., Rodríguez-Jiménez, F.J., Forssmann, U., Adermann, K., Kluver, E., Vogelmeier, C., Becker, D., Hedrich, R., Forssmann, W.G., Bals, R.: Identification of a novel, multifunctional β-defensin (human β-defensin 3) with specific antimicrobial activity. Cell and Tissue Research, **306**, 257–264 (2001)

An Improved Heuristic for Consistent Biclustering Problems

Artyom Nahapetyan[1], Stanislav Busygin[2], and Panos Pardalos[3]

[1] Center for Applied Optimization, University of Florida, artyom@ufl.edu
[2] Center for Applied Optimization, University of Florida, busygin@ufl.edu
[3] Center for Applied Optimization, University of Florida, pardalos@ufl.edu

Key words: Biclustering problems, Heatmap Builder software, Block Clustering, supervised biclustering, "Checkerboard" pattern, Human Gene Expression (HuGE) Index.

1 Introduction

Let matrix A represent a data set of m features and n samples. Each element of the matrix, a_{ij}, corresponds to the expression of the i-th feature in the j-th sample. Biclustering is a classification of the samples as well as features into k classes. In other words, we need to classify columns and rows of the matrix A. Doing so, let $S_1, S_2, \ldots, S_k$ and $F_1, F_2, \ldots, F_k$ denote the classes of the samples (columns) and features (rows), respectively. Formally biclustering can be defined as follows.

Definition 1. *A biclustering is a collection of pairs of sample and feature subsets* $\mathcal{B} = \{(S_1, F_1), (S_2, F_2), \ldots, (S_k, F_k)\}$ *such that*

$$S_1, S_2, \ldots, S_k \subseteq \{a^j\}_{j=1,\ldots,n},$$

$$\bigcup_{r=1}^{k} S_r = \{a^j\}_{j=1,\ldots,n},$$

$$S_\zeta \bigcap S_\xi = \emptyset, \zeta \neq \xi,$$

$$F_1, F_2, \ldots, F_k \subseteq \{a_i\}_{i=1,\ldots,m},$$

$$\bigcup_{r=1}^{k} F_r = \{a_i\}_{i=1,\ldots,m},$$

$$F_\zeta \bigcap F_\xi = \emptyset, \zeta \neq \xi,$$

where $\{a^j\}_{j=1,\ldots,n}$ *and* $\{a_i\}_{i=1,\ldots,m}$ *denote the set of columns and rows of the matrix A, respectively.*

By reordering the columns and rows of the matrix according to their classifications, the corresponding biclustering can be visualized using the Heatmap Builder software [HeatMap], where the color of a pixel is chosen according to the corresponding value of a_{ij}. Despite all varieties of classifications, our ultimate goal in a biclustering problem is to find a classification in which samples from the same class have similar values for the features that characterize the class. The visualization of a reasonable classification should reveal a block-diagonal or "checkerboard" pattern similar to one on Figure 1.

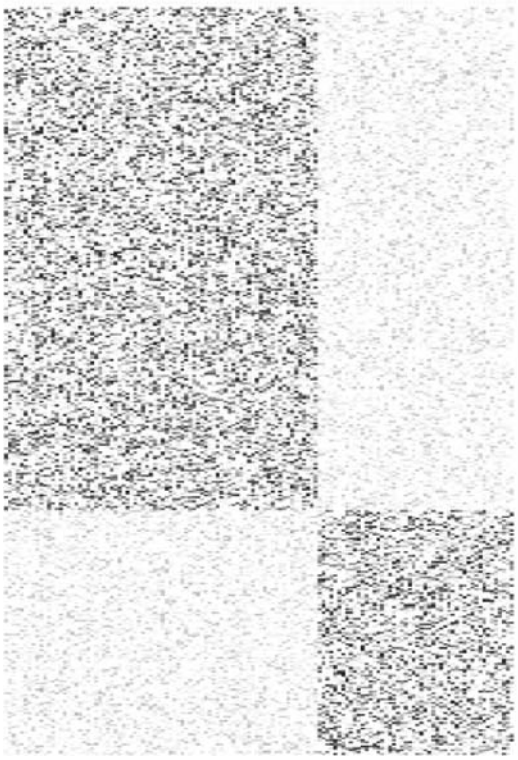


Fig. 1. An example of biclustering: "checkerboard" pattern.

One of the early algorithms to obtain an appropriate biclustering is proposed by Hartigan [H72], which is known as *Block Clustering*. Given a biclustering $\mathcal{B}$, the author employs the variability of the data in the block (S_r, F_r) to measure the quality of the classification. In the resulting problem a lower variability is preferable. However, to avoid a trivial solution with zero variability,

where each class consists of only one sample, it is required to fix the number of classes. A more sophisticated approach for biclustering was introduced by Cheng and Church [CC00], where the authors minimize the mean squared residual. They prove that the problem is NP-hard and propose a greedy algorithm to find an approximate solution to the problem. A simulated annealing technique to solve the problem is discussed by Bryan et al. [BCB05].

Dhillon [D01] discusses another biclustering method for text mining using a bipartite graph. In the graph the nodes represent features and samples, and each feature i is connected to a sample j with a link (i, j), which has a wight a_{ij}. The total weight of all links connecting features and samples from different classes is used to measure the quality of a biclustering. In particular, a lower value corresponds to a better biclustering. A similar method for microarray data is suggested by Kluger et al. [KBCG03].

Another method to tackle the problem is to treat the input data as a joint probability distribution between two discrete sets of random variables (see Dhillon et al. [DMM03]). The goal of the method is to find disjoint classes for both variables. A Bayesian biclustering technique based on the Gibbs sampling can be found in Sheng et al. [SMM03].

Recently, Busygin et al. [BPP05] have introduced a concept of *consistent biclustering*. Formally speaking, a biclustering $\mathcal{B}$ is consistent if in each sample (feature) from any set S_r (set F_r) the average expression of features (samples) that belong to the same class r is grater than the average expression of features (samples) from other classes. It has been shown that consistent biclustering implies cone separability of samples and features. The mathematical formulation of the problem belongs to the 0-1 fractional programming. To solve the supervised biclustering problem, the authors introduce additional variables to linearize the problem and propose an iterative heuristic procedure, where in each iteration it is required to solve a smaller size mixed integer problem. In this chapter we discuss an improved heuristic procedure, where in each iteration we solve a continuous linear problem. Numerical experiments on the same data confirm that our algorithm outperforms the previous result in the quality of the obtained solution as well as computational time.

In Section 1 we provide a brief discussion of the consistent biclustering. For details we refer to the paper by Busygin et al. [BPP05]. Section 3 introduces the application of the technique in the supervised biclustering problem. The heuristic algorithm and numerical experiments are described in Sections 4 and 5, respectively. Finally, Section 6 concludes the chapter.

2 Consistent Biclustering

Given a classification of the samples S_r, let $S = (s_{jr})_{n\times k}$ denote a 0-1 matrix where $s_{jr} = 1$ if the sample j is classified as a member of the class r, i.e., $a^j \in S_r$, and $s_{jr} = 0$ otherwise. Similarly, given a classification of the features F_r, let $F = (f_{ir})_{m\times k}$ denote a 0-1 matrix where $f_{ir} = 1$ if the feature i belong

to the class r, i.e., $a_i \in F_r$, and $f_{ir} = 0$ otherwise. Using those matrices construct corresponding *centroids* for the samples and features.

$$C_S = AS(S^TS)^{-1} = (c^S_{i\xi})_{m\times r} \tag{1}$$

$$C_F = A^TF(F^TF)^{-1} = (c^F_{j\xi})_{n\times r} \tag{2}$$

The elements of the matrices, $c^S_{i\xi}$ and $c^F_{j\xi}$, represent the average expression of the corresponding sample and feature in the class ξ, respectively. In particular,

$$c^S_{i\xi} = \frac{\sum_{j=1}^n a_{ij}s_{j\xi}}{\sum_{j=1}^n s_{j\xi}} = \frac{\sum_{j|a^j\in S_\xi} a_{ij}}{|S_\xi|},$$

and

$$c^F_{j\xi} = \frac{\sum_{i=1}^m a_{ij}f_{i\xi}}{\sum_{i=1}^m f_{i\xi}} = \frac{\sum_{i|a_i\in F_\xi} a_{ij}}{|F_\xi|}.$$

Consider the matrix C_S. Using the elements of the matrix, one can assign a feature to a class where it is most expressed. Doing so let assign the feature i to the class $\hat{r}$ if $c^S_{i\hat{r}} = \max_\xi\{c^S_{i\xi}\}$, i.e.,

$$a_i \in \hat{F}_{\hat{r}} \Longrightarrow c^S_{i\hat{r}} > c^S_{i\xi}, \qquad \forall\xi, \xi \neq \hat{r}. \tag{3}$$

It is noticed that the constructed classification of the features, $\hat{F}_r$, is not necessary to be the same as the classification F_r. Similarly, one can use the elements of the matrix C_F to classify the samples. In particular, assign the sample j to the class $\hat{r}$ if $c^F_{j\hat{r}} = \max_\xi\{c^F_{j\xi}\}$, i.e.,

$$a^j \in \hat{S}_{\hat{r}} \Longrightarrow c^F_{j\hat{r}} > c^F_{j\xi}, \qquad \forall\xi, \xi \neq \hat{r}. \tag{4}$$

As before, obtained classification $\hat{S}_r$ is not necessary to coincide with classification S_r.

Definition 2. *We refer to a biclustering $\mathcal{B}$ as a consistent biclustering if relations (3) and (4) hold for all elements of the corresponding classes, where matrices C_S and C_F are defined according to (1) and (2), respectively.*

Theorem 1. *Let $\mathcal{B}$ be a consistent biclustering. Then there exist convex cones $P_1, P_2, \ldots, P_k \subseteq \mathbb{R}^m$ such that only samples from S_r belong to the corresponding cone P_r, $r = 1, \ldots, k$. Similarly, there exist convex cones $Q_1, Q_2, \ldots, Q_k \subseteq \mathbb{R}^n$ such that only features from class F_r belong to the corresponding cone Q_r, $r = 1, \ldots, k$.*

Proof. For the proof, see [BPP05].

According to the definition, a biclustering is consistent if $F_r = \hat{F}_r$ and $S_r = \hat{S}_r$. Theorem 1 proves that a consistent biclustering implies the separability by cons. Despite the nice properties, for a given data set, it might be impossible to construct a consistent biclustering. The later is due to the fact that the

data set includes features and/or samples that are not evidently belong to any of the classes. However, one can delete some of the features and/or samples from the data set so that there is a consistent biclustering for the truncated data set. Our ultimate goal is to include into the truncated data as many features and samples as possible.

Another problem of our interest is to choose the most representative subset of samples and features. For instance, assume that there is a consistent biclustering for a given data set, and there is a feature, i, such that the difference between the two largest values of c^S_{ir} is negligible, i.e.,

$$\min_{\xi \neq \hat{r}} \{c^S_{i\hat{r}} - c^S_{i\xi}\} \leq \alpha,$$

where α is a small positive number. Although this particular feature is classified as a member of class $\hat{r}$, i.e., $a_i \in F_{\hat{r}}$, it is easy to violate the corresponding relation (3) by adding a slightly different sample to the data set. In other words, if α is a relatively small number, then it is not statistically evident that $a_i \in F_{\hat{r}}$, and the feature i cannot be used to classify the samples. The problem of choosing the most representative features and samples is important in the cases when performing feature tests and collecting a large number of samples are expensive and time consuming. Before we proceed to the formulation of the problem, let us define a notion of additive and multiplicative consistent biclusterings that are stronger than the consistent biclustering.

Instead of (3) and (4) consider the relations

$$a_i \in F_{\hat{r}} \Longrightarrow c^S_{i\hat{r}} > \alpha^S_i + c^S_{i\xi}, \qquad \forall \xi, \xi \neq \hat{r}, \tag{5}$$

and

$$a^j \in S_{\hat{r}} \Longrightarrow c^F_{j\hat{r}} > \alpha^F_j + c^F_{j\xi}, \qquad \forall \xi, \xi \neq \hat{r}, \tag{6}$$

respectively, where $\alpha^F_j > 0$ and $\alpha^S_i > 0$. Let α denote the vector of α^F_j and α^S_i.

Definition 3. *A biclustering $\mathcal{B}$ is called an additive consistent biclustering with parameter α or α-consistent biclustering if relations (5) and (6) hold for all elements of the corresponding classes, where matrices C_S and C_F are defined according to (1) and (2), respectively.*

Similarly, instead of (3) and (4) consider the relations

$$a_i \in F_{\hat{r}} \Longrightarrow c^S_{i\hat{r}} > \beta^S_i c^S_{i\xi}, \qquad \forall \xi, \xi \neq \hat{r}, \tag{7}$$

and

$$a^j \in S_{\hat{r}} \Longrightarrow c^F_{j\hat{r}} > \beta^F_j c^F_{j\xi}, \qquad \forall \xi, \xi \neq \hat{r}, \tag{8}$$

respectively, where $\beta^F_j > 1$ and $\beta^S_i > 1$. Let β denote the vector of β^F_j and β^S_i.

Definition 4. *A biclustering $\mathcal{B}$ is called a multiplicative consistent biclustering with parameter β or β-consistent biclustering if relations (7) and (8) hold for all elements of the corresponding classes, where matrices C_S and C_F are defined according to (1) and (2), respectively.*

It is easy to show that an α-consistent biclustering is a consistent biclustering for all values of $c^S_{i\xi}$ and $c^F_{j\xi}$. In the case of β-consistent biclustering, it is a consistent biclustering if $c^S_{i\xi} \geq 0$ and $c^F_{j\xi} \geq 0$. The latter usually holds in DNA microarray problems .

Using above definitions, we can formulate two problems of choosing the most representative subsets of features and samples. In the first one, we delete a least number of features and/or samples from a data set so that there exists an α-consistent biclustering for the truncated data set. In the second problem, one can choose to achieve a β-consistent biclustering by deleting a least number of features and/or samples. In those two problems, vectors α and β play a role of a threshold for choosing features and samples. However, large values of the vectors α and β can be very restrictive. As a result, some of the valuable features and samples might be excluded from the truncated data set. The optimal values of the parameters α and β should be tuned based on some experiments with the data.

3 Supervised Biclustering

In the real-life problems usually there is a set of data for which the classification is known. For instance, if some patients are diagnosed with *acute lymphoblastic leukemia (ALL)* or *acute myeloid leukemia (AML)* then their microarray data can be classified as ALL or AML. In the supervised biclustering we assume that there is a training data, i.e., a set of samples for which the classification is known, and it is accurate. Using the training data, one can classify the features as it is described in Section 1 and formulate consistent, α-consistent and β-consistent biclustering problems. Then solutions of the problems can be used to classify additional samples. The values of the vectors α and β can be adjusted to obtain a more compact set of representative features as well as reduce the number of misclassifications in the data.

Given a set of training data construct the matrix S and compute the values of $c^S_{i\xi}$ using the formula (1). Classify the features according to the following rule: the feature i belongs to the class $\hat{r}$, i.e., $a_i \in F_{\hat{r}}$, if $c^S_{i\hat{r}} > c^S_{i\xi}$, $\forall \xi \neq \hat{r}$. At last, construct the matrix F using the obtained classification. Let x_i denote a binary variable, which takes value one if the feature i is included in the computations and zero otherwise. Consider the following optimization problems.

CB:

$$\max_x \sum_{i=1}^{m} x_i \tag{9}$$

$$\frac{\sum_{i=1}^{m} a_{ij} f_{i\hat{r}} x_i}{\sum_{i=1}^{m} f_{i\hat{r}} x_i} > \frac{\sum_{i=1}^{m} a_{ij} f_{i\xi} x_i}{\sum_{i=1}^{m} f_{i\xi} x_i}, \quad \forall \hat{r}, \xi \in \{1, \ldots, k\}, \hat{r} \neq \xi, j \in S_{\hat{r}} \tag{10}$$

$$x_i \in \{0, 1\}, \quad \forall i \in \{1, \ldots, m\} \tag{11}$$

α-CB:

$$\max_x \sum_{i=1}^{m} x_i$$

$$\frac{\sum_{i=1}^{m} a_{ij} f_{i\hat{r}} x_i}{\sum_{i=1}^{m} f_{i\hat{r}} x_i} > \alpha_j + \frac{\sum_{i=1}^{m} a_{ij} f_{i\xi} x_i}{\sum_{i=1}^{m} f_{i\xi} x_i}, \quad \forall \hat{r}, \xi \in \{1, \ldots, k\}, \hat{r} \neq \xi, j \in S_{\hat{r}}$$

$$x_i \in \{0, 1\}, \quad \forall i \in \{1, \ldots, m\}$$

β-CB:

$$\max_x \sum_{i=1}^{m} x_i$$

$$\frac{\sum_{i=1}^{m} a_{ij} f_{i\hat{r}} x_i}{\sum_{i=1}^{m} f_{i\hat{r}} x_i} > \beta_j \frac{\sum_{i=1}^{m} a_{ij} f_{i\xi} x_i}{\sum_{i=1}^{m} f_{i\xi} x_i}, \quad \forall \hat{r}, \xi \in \{1, \ldots, k\}, \hat{r} \neq \xi, j \in S_{\hat{r}}$$

$$x_i \in \{0, 1\}, \quad \forall i \in \{1, \ldots, m\}$$

In the CB problem we are looking for the largest set of features, which can be used to construct a consistent biclustering. The α-CB and β-CB problems are similar to CB problem. The only difference is that the selected set of features have to allow constructing α-consistent and β-consistent biclusterings, respectively.

4 Heuristic Algorithm

All three above optimization problems belong to the fractional 0-1 programming, and it is difficult to find an optimal solution of the problems. In the paper by Busygin et al. [BPP05] the authors consider the β-CB problem and introduce a linearization of the problem. However, commercial mixed integer programming (MIP) solvers are not able to solve it due to the excessive number of variables and constraints. The authors introduce an iterative heuristic procedure, where in each iteration it is required to solve a linear 0-1 problem of a smaller size. In this section, we discuss a heuristic procedure to solve the problems, which iteratively solves continuous linear problems. Because the same algorithm can be applied to all three problems, in our discussion we focus only on the CB problem.

Observe that in the problems the expression $\sum_{i=1}^{m} f_{i\xi} x_i$ describes the cardinality of the set of features in the truncated data. In particular, if $x_i = 1$, $\forall i \in \{1, \ldots, m\}$ such that $f_{i\xi} = 1$, then it is equal to the cardinality of F_ξ. Given a vector x, let $F_\xi(x)$ denote the truncated set of features, i.e., $F_\xi(x) \subseteq F_\xi$ such that the features are included in the set $F_\xi(x)$ only if $x_i = 1$. If the optimal cardinality of the sets $F_\xi(x)$ are known, then they can be fixed at the optimal values, and the problem reduces to a linear one. In the heuristic procedure we employ this property and iteratively solve a series of linear programs by updating the cardinalities according to the current available solution.

In the first step of the algorithm (see Procedure 1), we assign $x_i^0 = 1$, $\forall i \in \{1, \ldots, m\}$, $F_\xi(x^0) = F_\xi$, $\forall \xi \in \{1, \ldots, k\}$, and $p = 0$. In the second step, we solve the following linear program, which can be obtained from the CB problem by fixing the cardinalities of the feature sets at the values $|F_\xi(x^p)|$ and relaxing the integrality of the variables x_i.

$$\max_x \sum_{i=1}^{m} x_i \tag{12}$$

$$\frac{\sum_{i=1}^{m} a_{ij} f_{i\hat{r}} x_i}{|F_{\hat{r}}(x_i^p)|} \geq \frac{\sum_{i=1}^{m} a_{ij} f_{i\xi} x_i}{|F_\xi(x_i^p)|}, \quad \forall \hat{r}, \xi \in \{1, \ldots, k\}, \hat{r} \neq \xi, j \in S_{\hat{r}} \tag{13}$$

$$x_i \in [0, 1], \quad \forall i \in \{1, \ldots, m\} \tag{14}$$

Let $p \leftarrow p+1$ and x^p denote the vector solution of the problem. According to the solution x^p, construct the sets $F_\xi(x^p)$, where the features are included in the set only if $x_i^p = 1$, i.e., $F_\xi(x^p) \subseteq F_\xi$ such that $x_i^p = 1$. If $\exists \xi \in \{1, \ldots, k\}$ such that $F_\xi(x^p) \neq F_\xi(x^{p-1})$ then go to Step 2 and solve the problem (12)-(14) with updated values of cardinalities. On the other hand, if $F_\xi(x^p) = F_\xi(x^{p-1})$, $\forall \xi \in \{1, \ldots, k\}$, then we have to check if $x_i^* = \lfloor x_i^p \rfloor$ is feasible to the constraint (13). If the feasibility is satisfied, then stop and return the value of the vector x^*. However, if the vector is not feasible, then we conclude that the variables x_i^p with fractional values cannot take value one, i.e., the corresponding features cannot be included in the set of the truncated features. Then we delete permanently those features from the data set and continue the process.

Observe that the solution x^* is feasible to the CB problem. In particular, x_i^* takes either value one or zero. Because by construction the sets $F_\xi(x^p)$ include only the features with $x_i^* = 1$, the feasibility to the inequality (13) implies the feasibility to the inequality (10). It is noticed that the strict inequality in (10) holds in practical problems because of the following reasons.

Recall that the objective (12) maximizes the number of features included in the truncated data set, and it is our benefit to have the values of the variables x_i as close as possible to one. However, because of the inequality (13), some variables take fractional values at optimality. As a result, some of the constraints (13) are tight at optimality. If $x^* = \lfloor x^p \rfloor$ is feasible to (13), then it is highly unlikely that the constraints remain tight.

5 Numerical Experiments

In the computational experiments we consider a well known data set, which consists of samples from patients diagnosed with *acute lymphoblastic leukemia (ALL)* or *acute myeloid leukemia (AML)* diseases (see [GST99]). This date set is used in the computations by Busygin et al. [BPP05] as well as other researchers (see, e.g., [BBNSY00], [BFY01], [WMCPPV01], [XK01]). Similar to the numerical experiments in [BPP05], we divide the data set into two groups, where the first group is used as a training data set and the second one, test data set, is used to verify the quality of the obtained classification. The training data set consists of 38 samples from which 27 are ALL and 11 are AML samples. The test data set consist of 20 ALL and 14 AML samples. Each sample in the sets consists of 7070 features.

Table 1. Computational results on ALL and AML samples: the CB and α-CB problems with different values of α.

	CB	10-CB	20-CB	30-CB	40-CB	50-CB	60-CB	70-CB	130-CB
Number or Features	7024	7021	7018	7014	7010	6959	6989	6960	4639
Number of Errors	2	2	2	2	1	1	1	1	1
CPU	1.66	1.91	2.08	2.21	2.84	90.52	32.79	24.35	6.91

Table 2. Computational results on ALL and AML samples: the CB and β-CB problems with different values of β.

	CB	1.05-CB	1.1-CB	1.2-CB	1.5-CB	2-CB	3-CB	5-CB	7-CB
Number or Features	7024	7017	7010	6937	6508	5905	5458	5173	5055
Number of Errors	2	2	1	1	1	1	1	2	3
CPU	1.66	1.68	1.7	37.55	28.45	17.67	6.39	6.44	4.73

Procedure 1 :

Step 1: Let $x_i^0 = 1$, $\forall i \in \{1, \ldots, m\}$, $F_\xi(x^0) = F_\xi$, $\forall \xi \in \{1, \ldots, k\}$, and $p = 0$.
Step 2: Solve the problem (12)-(14). Let $p \leftarrow p + 1$ and x^p denote the vector solution of the problem.
Step 3: Construct the set of features $F_\xi(x^p) \subseteq F_\xi$ such that $x_i^p = 1$.
Step 4: If $F_\xi(x^p) \neq F_\xi(x^{p-1})$ then go to Step 2.
Step 5: $x_i^* \leftarrow \lfloor x_i^p \rfloor$ If x_i^* is feasible to the constraint (13) then stop. Otherwise, delete permanently from further consideration all features with fractional values of x_i^p and go to the Step 2.

We run our heuristic algorithm to solve CB as well as α-CB and β-CB problems with different values for the parameters α and β. Although parameters α_j and β_j can take different values for different features, in our experiments we assume that there are all equal. In all cases, we obtain the "checkerboard" pattern similar to one on Figure 2.

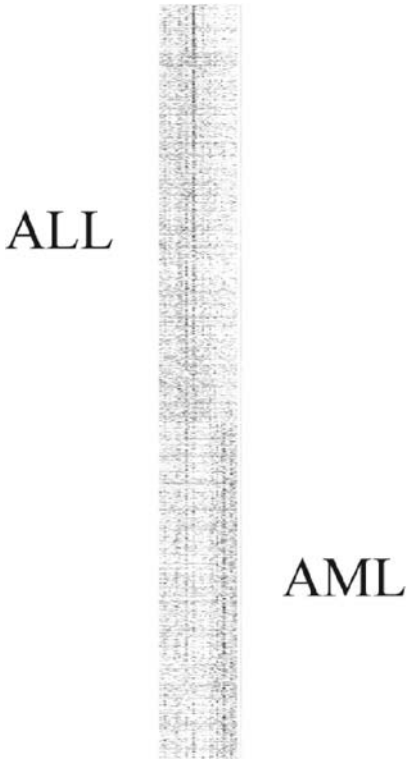


Fig. 2. "Checkerboard" pattern for ALL and AML samples.

Table 1 illustrates the results for the additive consistent biclustering with different values of the vector parameter α. The first row in the table represents the maximum number of features in the truncated data that allow constructing the corresponding biclustering. Using the obtained set of features, we classify the samples from the second group of data, and the second row in the table represents the number of misclassifications. Finally, the last row provides an information on the CPU time of the algorithm. As we can see from the table, a higher value of the parameter α better classifies the samples. In particular, for the values equal to 40 and higher there is only one error detected in the classification of the test data. In addition, observe that the number of selected features decreases with the increase of the parameter. These two observations

lead to a conclusion that a fewer but most representative features can be used to classify the data. The highest value of the parameter α for which we are able to obtain an α-consistent biclustering is 130. As for the CPU time, notice that it varies depending on the value of the parameter α but remains within reasonable limits.

A similar result can be obtained using β-consistent biclustering (see Table 2). In particular a higher value of β provides a better classification. However, the values of β grater than or equal to 5 are to restrictive and worsen the quality of the classification. It is noticed that in the paper by Busygin et al. [BPP05] the heuristic algorithm proposed by the authors converges after 15 minutes and is able to select 6681 features for the parameter $\beta_j = 1.1$, $\forall j \in \{1, \ldots, n\}$. Using the same parameter, our algorithm outperforms the previous result by selecting 7010 features within 1.68 seconds of CPU time.

Despite a small number of deleted features (in the case of the CB problem the number of deleted features is 46), the consistent biclustering is crucial to obtain a good classification for the features. In particular, if one classifies all features using the formula (3) and tests the classification using the second group of data, then usually the number of misclassifications is larger. In the case of AML and ALL samples, the number of misclassifications we obtain using this technique is 19. (Practically all ALL sample from the test set are classified as AML.)

In addition to the ALL and AML samples, we test our algorithm on Human Gene Expression (HuGE) Index. The samples are collected from healthy tissues of different parts of human body. The main purpose of the classification is to identify the features that are highly expressed in particular tissue. Table 3 illustrates the computational results of the CB and α-CB problems for different values of α. It is interesting to observe that in the most of the tissues, e.g., Blood, Brain, and Breast, the number of the selected features do not change for different values of α. On the other hand, some tissues, e.g., Ovary, are more "sensitive" to the changes of the parameter. Table 4 introduces the results for the multiplicative consistent biclustering. Although in these problems the set of "sensitive" tissues is larger than in the case of α-CB problems, some tissues, Cervix, Kidney, Placenta, Prostate, Spleen, and Stomach, preserve the same number of selected features. The last column in the table provides the data from the paper by Busygin et al. [BPP05] where the authors solve the multiplicative consistent biclustering problem with parameter $\beta_j = 1.1$, $\forall j \in \{1, \ldots, n\}$. Observe that for the same value of the parameter, our algorithm finds 162 more features.

6 Conclusion Remarks

In this chapter we have discussed the concept of consistent biclustering presented by Busygin et al. [BPP05]. For the supervised biclustering case, the additive and multiplicative variations of the problem are introduced to further

Table 3. Computational results on HuGE data set: CB and α-CB problems with different values of α.

		CB	α-CB			
Tissue type	#Sumples		$\alpha = 10$	$\alpha = 30$	$\alpha = 50$	$\alpha = 70$
Blood	1	472	472	472	472	472
Brain	11	615	615	615	615	615
Breast	2	903	903	903	903	903
Colon	1	367	366	363	360	355
Cervix	1	155	155	155	155	155
Endometrium	2	226	225	222	218	211
Esophagus	1	281	280	277	274	272
Kidney	6	159	159	159	159	159
Liver	6	440	440	440	440	440
Lung	6	102	102	102	102	102
Muscle	6	533	533	533	533	532
Myometrium	2	162	161	159	156	153
Ovary	2	257	255	251	246	240
Placenta	2	519	519	519	519	519
Prostate	4	281	281	281	281	281
Spleen	1	438	438	438	438	438
Stomach	1	447	447	447	447	447
Testes	1	522	521	520	518	515
Vulva	3	187	187	187	187	187
Total	59	7066	7059	7043	7023	6996

analyze the possibilities of choosing the most representative set of features. The heuristic algorithm presented in this chapter allows computing the set of truncated data. Unlike the algorithm presented in [BPP05], where it is required to solve a sequence of integer programs, our approach iteratively solves continuous linear problems. Computational results on the same data set conform that our heuristic algorithm outperforms the previous result in the quality of the solution as well as computational time. Although for the most of the values of the parameters α and β the heuristic algorithm converges to a solution, for some values it does not. However, in latter cases the algorithm converges after a small perturbation in the values of α and β.

Table 4. Computational results on HuGE data set: CB and β-CB problems with different values of β.

		CB	β-CB			[BPP05]
Tissue type	#Sumples		$\beta = 1.1$	$\beta = 1.5$	$\beta = 2$	$\beta = 1.1$
Blood	1	472	472	472	467	472
Brain	11	615	615	615	610	614
Breast	2	903	903	903	900	902
Colon	1	367	365	354	348	367
Cervix	1	155	155	155	155	107
Endometrium	2	226	224	212	190	225
Esophagus	1	281	278	269	259	289
Kidney	6	159	159	159	159	159
Liver	6	440	440	440	421	440
Lung	6	102	102	102	101	102
Muscle	6	533	533	533	515	532
Myometrium	2	162	160	153	142	163
Ovary	2	257	253	241	225	272
Placenta	2	519	519	519	519	514
Prostate	4	281	281	281	281	174
Spleen	1	438	438	438	438	417
Stomach	1	447	447	447	447	442
Testes	1	522	520	513	506	512
Vulva	3	187	187	187	182	186
Total	59	7066	7051	6993	6865	6889

References

[BBNSY00] Ben-Dor A., Bruhn L., Nachman I., Schummer M., Yakhini Z.: Tissue Classification with Gene Expression Profiles. Journal of Computational Bilology, **7**, 559–584 (2000)

[BFY01] Ben-Dor A., Ffiedman N., Yakhin Z.: Class Descovery in Gene Expression Data. Procidings of Fifth Annual International Conveference on Computational Malecular Biology, (2001)

[BCB05] Bryan K., Cunningham P., Bolshakova N.: Biclustering of Expression Data Using Simulated Annealing. Proceedings of the 18th IEEE Symposium on Computer-Based Medical Systems, 383–388 (2005)

[BPP05] Busygin S., Prokopyev O., Pardalos P.: Feature Selection for Consistent Biclustering via Fractional 0-1 Programming. Journal of Combinatorial Optimization, **10**, 7–21 (2005)

[CC00] Cheng Y., Church G. M.: Biclustering of Expression Data. Proceedings of the Eighth International Conference on Intelligent Systems for Molecular Biology, 93–103 (2000)

[D01] Dhillon I. S.: Co-Clustering Documents and Words Using Bipartite Spectral Graph Partitioning. Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), San Francisco, 26–29 (2001)

[DMM03] Dhillon I. S., Mallela S., Modha D. S.: Information-Theoretic Coclustering. Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 89–98 (2003)

[GST99] Golub T. R., Slonim D. K., Tamayo P., Huard C., Gaasenbeek M., Mesirov J. P., Coller H., Loh M. L., Downing J. R., Caligiuri M. A., Bloomfield C. D., Lander E. S.: Molecular Classification of Cancer: Class Discovery and Class Prodiction by Gene Expression Monitoring. Science, **286**, 531–537 (1999)

[H72] Hartigan J. A.: Direct Clustering of a Data Matrix. Journal of American Statistical Association, **67**, 123–129 (1972)

[HeatMap] HeatMap Builder Software, Quertermous Laboratory, Stanford University, http://quertermous.stanford.edu/heatmap.htm.

[KBCG03] Kluger Y., Basri R., Chang J. T., Gerstein M.: Spectral Biclustering of Microarray Data: Coclustering Genes and Conditions. Genome Research, 703–716 (2003)

[SMM03] Sheng Q., Moreau Y., De Moor B.: Biclustering Microarray Data by Gibbs Sampling. Bioinformatics, **19**, ii196–ii205 (2003)

[WMCPPV01] Weston J., Mukherjee S., Chapelle O., Pontil M., Poggio T., Vapnik V.: Feature Selection for SVMs. NIPS, (2001)

[XK01] Xing E.P., Karp R.M.: CLIFF: Clustering of High-Dimensional Microarray Data Via Iterative Feature Filtering Using Normilized Cuts. Bioinfomatics Discovery Note, **1**, 1–9 (2001)

The Steiner Tree Problem and Its Application to the Modelling of Biomolecular Structures

Rubem P. Mondaini

Federal University of Rio de Janeiro, COPPE, Centre of Technology
21.941-972 - P.O. Box 68511, Rio de Janeiro, RJ, Brazil
rpmondaini@gmail.com, mondaini@cos.ufrj.br

Key words: Steiner minimal tree, Fermat-Steiner problem, evenly spaced points, Chebyshev polynomials, Steiner Ratio function.

1 Introduction

There is now a paradigm in the study of an efficient modelling of biomolecular structure. The foundations of the physics and the elucidation of biological function of living organisms will be asserted on a clear geometrical language according the best ideas of D'arcy Thompson [1], Rashevsky [2], Schrödinger [3] and Anfinsen [4]. The present work aims to give a possible mathematical description of one of Nature's services of noticeable importance in the organization of life and its maintenance: the specific geometric form of macromolecular structure as provided by the mathematical problem of organization of Steiner Minimal Trees. The energy minimization process which lead to the formation of a biomacromolecule can be understood and modelled by the search process of organization of Steiner trees as the representatives of the possible molecular configurations corresponding to local minima of the free energy. Life maintenance and the survival of the living organism is guaranteed by the competence of staying away from the Global minimum structure and its associated Steiner Minimal Tree.

2 The Fermat-Steiner Problem. Motivations for Further Study

The Fermat-Steiner Problem in $\mathbb{R}^{\mathrm{d}}$ is the search for a point to be such that the sum of its distances (these are assumed to be Euclidean) to p other given points in $\mathbb{R}^{\mathrm{d}}$ is a minimum. This is not a sterile generalization of the old Fermat-Steiner problem in $\mathbb{R}^2$ for three given points [5] since it will be considered as the

basis of a evolutive process of molecular organization in terms of the stability analysis of globular clusters of points. These clusters are the prototypes for the modelling of early stages of biomacromolecules. This mathematical modelling will be done through the generalized Fermat Problem in $\mathbb{R}^3$. We now expose the guidelines of the Fermat Problem in $\mathbb{R}^{\mathrm{d}}$ and some elementary applications of the problem to the derivation of the coordinates of the Fermat-Steiner point of a triangle. We then consider p points in $\mathbb{R}^{\mathrm{d}}$ with coordinates given by $x_j^s s$, $1 \leq j \leq p$, $1 \leq s \leq d$. We look for a $(p+1)$-th point of coordinates x_{p+1}^s such that $\sum_{j=1}^{p} R_{p+1,j}$ = minimum where $R_{p+1,j}$ stands for the Euclidean distance between the $(p+1)$-th point and a generic j-th point, or

$$R_{p+1,j} = \sqrt{\sum_{s=1}^{d} \left(x_{p+1}^s - x_j^s\right)^2} \ . \tag{1}$$

This is an unconstrained optimization problem and the function $\sum_{j=1}^{p} R_{p+1,j}$ is strictly convex elsewhere as a sum of strictly convex functions $R_{p+1,j}$. The only exception is the case of p colinear points in which this function is convex only on the line joining them [6].

The optimization problem will be solved by

$$\frac{\partial}{\partial x_{p+1}^s} \sum_{j=1}^{p} R_{p+1,j} = 0 \ , \qquad 1 \leq s \leq d \ . \tag{2}$$

Let $\hat{\imath}_s$ be the unit vector of the s-direction. WE can write from eqs. (1), (2),

$$\sum_{j=1}^{p} \sum_{s=1}^{d} \left(\frac{x_{p+1}^s - x_j^s}{R_{p+1,j}} \right) \hat{\imath}_s = 0 = \sum_{j=1}^{p} \hat{r}_j \ , \tag{3}$$

where

$$\hat{r}_j = \frac{\mathbf{r}_{p+1} - \mathbf{r}_j}{R_{p+1,j}} \ . \tag{4}$$

The last equality of eq. (3) is the best characterization of the Fermat problem with Euclidean distance on $\mathbb{R}^{\mathrm{d}}$.

The first equality can be also written as

$$\sum_{s=1}^{d} \sum_{j=1}^{p} \cos \alpha_{js} \, \hat{\imath}_s = 0 \tag{5}$$

where $\cos \alpha_{js}$ is the direction co-sine of the vector $\mathbf{r}_{p+1} - \mathbf{r}_j$ with respect to the s-axis of coordinates.

The linear independence of the vectors $\hat{\imath}_s$ allow us to write

$$\sum_{j=1}^{p} \cos \alpha_{js} = 0 \ , \qquad 1 \leq s \leq d \ . \tag{6}$$

From eq. (1), we can also write

$$\sum_{s=1}^{d} \cos^2 \alpha_{js} = 1 \ , \qquad 1 \leq j \leq p \ . \tag{7}$$

Eqs. (6) and (7) are the fundamental constraints to be fulfilled by the direction co-sines in this formulation of the Fermat problem. The maximum number os free parameters is $dp-(d+p)$ to be chosen previously. We solve these equations below for the $p = 3$ problem. This problem has a unique solution as it follow from eq. (3) which can be written as

$$1 + \sum_{\substack{j=1 \\ j \neq k}}^{p} \cos \gamma_{jk} = 0 \ , \tag{8}$$

where

$$\cos \gamma_{jk} = \hat{r}_j \cdot \hat{r}_k \ .$$

There are p equations and $\binom{p}{2}$ unknowns. The unique solution for $p = 3$ is given by

$$\cos \gamma_{jk} = -\frac{1}{2}(1 - 3\delta_{jk}) \ , \tag{9}$$

where δ_{jk} is the Kronecker symbol.

There are also solutions of the form (9) for $p \geq 4$. These can be written as:

$$\cos \gamma_{jk} = -\frac{1}{p-1}(1 - p\delta_{jk}) \ . \tag{10}$$

In the following section, we shall prove that for points evenly spaced from a fixed origin, eq. (8) has a solution of the form (10) only for $p = 3, 4$. These points do correspond to the position of carbon and nitrogen atoms ($p = 3$) and α-carbon atoms ($p = 4$) in the structure of proteins.

We go back now to the solutions of eqs. (6), (7). We now propose as a solution for the direction co-sines $\cos \alpha_{js}$ the $p \times d$ matrix

$$\cos \alpha_{js} = \tag{11}$$

$$= \begin{pmatrix} \cos\theta_1 & \cos\theta_2 & \dots & \cos\theta_d \\ \cos\left(\theta_1 \pm \frac{2\pi}{p}\right) & \frac{\cos\theta_2}{\sin\theta_1}\sin\left(\theta_1 \pm \frac{2\pi}{p}\right) & \dots & \frac{\cos\theta_d}{\sin\theta_1}\sin\left(\theta_1 \pm \frac{2\pi}{p}\right) \\ \cos\left(\theta_1 \pm \frac{4\pi}{p}\right) & \frac{\cos\theta_2}{\sin\theta_1}\sin\left(\theta_1 \pm \frac{4\pi}{p}\right) & \dots & \frac{\cos\theta_d}{\sin\theta_1}\sin\left(\theta_1 \pm \frac{4\pi}{p}\right) \\ \vdots & \vdots & & \vdots \\ \cos\left(\theta_1 \pm \frac{2\pi(p-1)}{p}\right) & \frac{\cos\theta_2}{\sin\theta_1}\sin\left(\theta_1 \pm \frac{2\pi(p-1)}{p}\right) & \dots & \frac{\cos\theta_d}{\sin\theta_1}\sin\left(\theta_1 \pm \frac{2\pi(p-1)}{p}\right) \end{pmatrix}$$

This is a class of solutions with $d-1$ free parameters and $d-1 \leq pd-(p+d)$. We have used the identity

$$\sum_{k=0}^{p-1} e^{i\frac{2\pi k}{p}} \equiv 0$$

The case $d = 2$, $p = 3$ has a maximum of $6-5 = 1$ independent parameters, say $\alpha_{11} = \theta_1$. This solution can be written

$$\cos\alpha_{js} = \begin{pmatrix} \cos\theta_1 & \sin\theta_1 \\ \cos\left(\alpha \pm \frac{2\pi}{3}\right) & \sin\left(\alpha \pm \frac{2\pi}{3}\right) \\ \cos\left(\alpha \pm \frac{4\pi}{3}\right) & \sin\left(\alpha \pm \frac{4\pi}{3}\right) \end{pmatrix} \quad . \tag{12}$$

The case $d = 3$, $p = 3$ has $9 - 6 = 3$ parameters, say $\alpha_{11} = \theta_1$, $\alpha_{12} = \theta_2$, $\alpha_{13} = \theta_3$.

$$\cos\alpha_{js} = \begin{pmatrix} \cos\theta_1 & \cos\theta_2 & \cos\theta_3 \\ \cos\left(\theta_1 \pm \frac{2\pi}{3}\right) & \frac{\cos\theta_2}{\sin\theta_1}\sin\left(\theta_1 \pm \frac{2\pi}{3}\right) & \frac{\cos\theta_3}{\sin\theta_1}\sin\left(\theta_1 \pm \frac{2\pi}{3}\right) \\ \cos\left(\theta_1 \pm \frac{4\pi}{3}\right) & \frac{\cos\theta_2}{\sin\theta_1}\sin\left(\theta_1 \pm \frac{4\pi}{3}\right) & \frac{\cos\theta_3}{\sin\theta_1}\sin\left(\theta_1 \pm \frac{4\pi}{3}\right) \end{pmatrix} \quad . \tag{13}$$

We now restrict our analysis to the case of eq. (12). Let (x_4, y_4) and (a_j, b_j), $j = 1, 2, 3$ stand for the cartesian coordinates of the sought point and the vertices of the triangle, respectively. We fave from eqs. (3) and (5)

$$\begin{aligned} \cos\theta_1 &= \frac{x_4 - a_1}{R_{4,1}} & \sin\theta_1 &= \frac{y_4 - b_1}{R_{4,1}} \\ \cos\left(\theta_1 \pm \frac{2\pi}{3}\right) &= \frac{x_4 - a_2}{R_{4,1}} & \sin\left(\theta_1 \pm \frac{2\pi}{3}\right) &= \frac{y_4 - b_2}{R_{4,1}} \\ \cos\left(\theta_1 \pm \frac{4\pi}{3}\right) &= \frac{x_4 - a_3}{R_{4,1}} & \sin\left(\theta_1 \pm \frac{4\pi}{3}\right) &= \frac{y_4 - b_3}{R_{4,1}} \end{aligned} \tag{14}$$

We now proceed to the elimination of the free parameter θ_1 from eqs. (14). A straightforward but tedious manipulation will lead to

$$\frac{y_4 - b_1}{x_4 - a_1} = \frac{(a_2 - a_3)\sqrt{3} + b_1 + b_3 - 2b_1}{-(b_2 - b_3)\sqrt{3} + a_2 + a_3 - 2a_1} = \frac{a_2\sqrt{3} + b_2 - b_1 - x_4\sqrt{3}}{a_2 - b_2\sqrt{3} - a_1 + y_4\sqrt{3}} \quad . \tag{15}$$

The first and second equalities in eq. (15) correspond to the equations of two straight lines intersecting at the Fermat-Steiner point with coordinates (x_4, y_4) for a triangle in $d = 2$. These coordinates can be written

$$\begin{aligned} x_4 &= \frac{(A + C)[(BC - AD - b_1(B + D)\sqrt{3}) + a_1(B + D)^2\sqrt{3}]}{[(A + C)^2 + (B + D)^2]\sqrt{3}} \\ y_4 &= \frac{(B + D)[(BC - AD - a_1(A + C)\sqrt{3}) + b_1(A + C)^2\sqrt{3}]}{[(A + C)^2 + (B + D)^2]\sqrt{3}} \quad , \end{aligned} \tag{16}$$

where

$$A = a_2 - a_1 - b_2\sqrt{3} \ , \quad B = b_2 - b_1 + a_2\sqrt{3}$$
$$C = a_3 - a_1 - b_3\sqrt{3} \ , \quad D = b_3 - b_1 + a_3\sqrt{3} \ . \tag{17}$$

An interesting application of the formulae above will be given by deriving the equation of the geometrical locus of the Fermat-Steiner point of a triangle with two fixed vertices and the third variable along any continuous curve of the plain. Let us assume that the coordinates of the two fized vertices are given by $(-a, 0)$ and $(a, 0)$ and the variable vertex, (x, y). We have from eqs. (16):

$$x_4 = \frac{4a\sqrt{3}}{3} \frac{x(y + a\sqrt{3})}{x^2 + (y + a\sqrt{3})^2}$$
$$y_4 = \frac{a\sqrt{3}}{3} \frac{(y + a\sqrt{3})^2 - 3x^2}{x^2 + (y + a\sqrt{3})^2} \ . \tag{18}$$

From eqs. (2.14'), we can see that the geometrical locus is a circle with radius $2a\sqrt{3}/3$ and centre at $(0, -a\sqrt{3}/3)$, or

$$x_s^2 + \left(y_s + \frac{a\sqrt{3}}{3} \right)^2 = \frac{4a^2}{3} \ . \tag{19}$$

It should be stressed that there is not any assumed functional relation between the variables x, y. This means that we obtain the circle as a geometrical locus of the Fermat-Steiner point whatever the curve $y = y(x)$ is described by the variable vertex with coordinates (x, y). This property of the Fermat-Steiner point looks as the converse of a famous theorem of plane Euclidean Geometry:

Steiner Theorem: *All the geometrical constructions which can be done with a straightedge and a compass, can also be done with a straightedge only if a fixed circle and its centre are given previously.*

After the derivation of the geometrical locus above by using the methods of eqs. (6), (7), (15), we pass to the study of Steiner trees. These will be simply introduced by considering the possibility of additional points to the set of p fixed points of the present section in order to find a minimum spanning tree.

3 Consecutive Evenly Spaced Points

In this section we intend to present some solutions of the eqs. (3). We now propose to work with a sequence of points which are consecutive along a continuously differentiable curve and evenly spaced according the Euclidean

norm given by eq. (1). An Ansatz for the representation of its position vectors could be given by

$$\mathbf{r}_j = (r(\omega)\cos j\omega, r(\omega)\sin j\omega, jh(\omega))\,,\ 0 \le j \le p-1,\ \ 0 \le \omega \le 2\pi. \tag{20}$$

The position vectors of eq. (20) will satisfy the relations

$$\|\mathbf{r}_{j+1} - \mathbf{r}_j\| = \|\mathbf{r}_{j+2} - \mathbf{r}_{j+1}\|\ \ , \quad 0 \le j \le p-1\ \ , \tag{21}$$

where $\|\cdot\|$ stands for the Euclidean norm.

Our first set of solutions will be obtained by assuming that these points are the vertices of a regular p-gon with sidelength a. We then have the conditions

$$a^2 = 2r^2(1-\cos\omega) + h^2 \tag{22}$$

$$a^2 = 2r^2(1-\cos(p-1)\omega) + (p-1)^2h^2 \tag{23}$$

$$a^2\,\cos\frac{2\pi}{p} = r^2(2\cos\omega - 1 - \cos 2\omega) + h^2 \tag{24}$$

$$a^2\,\cos\frac{2\pi}{p} = r^2(\cos\omega - 1 + \cos(p-1)\omega - \cos(p-2)\omega) - (p-1)h^2\ \ . \tag{25}$$

After getting rid of the trivial case $p = 1$ and making some straightforward manipulations in the equations above, we have,

$$a^2 = \frac{2r^2(1-\cos\omega)^2}{1-\cos\frac{2\pi}{p}} \tag{26}$$

$$h^2 = \frac{2r^2(1-\cos\omega)}{1-\cos\frac{2\pi}{p}}\left(\cos\frac{2\pi}{p} - \cos\omega\right) \tag{27}$$

$$\cos(p-1)\omega - \cos\omega = p(p-2)\frac{h^2}{2r^2} \tag{28}$$

$$\cos(p-2)\omega - \cos 2\omega = p(p-4)\frac{h^2}{2r^2}\ \ . \tag{29}$$

It should be observed that the values $\omega = 2\pi/p$ and $\omega = 2\pi(p-1)/p$ are solutions of these equations for every $p \ge 2$. These solutions correspond to the inscription of p-gons of $p \ge 2$ sides in a closed plane curve, since we have in these cases $h = 0$. From eq. (27), the search for non-trivial solutions will be restricted to the range

$$\frac{2\pi}{p} < \omega < 2\pi\frac{(p-1)}{p}\ \ . \tag{30}$$

Since the left hand side of eqs. (28) and (29) are Chebyshev polynomials of degree $(p-1)$ and $(p-2)$ in the variable $\cos\omega$, respectively, we should have $p-1 \le 2$, $p-2 \le 2$. However, only for the value $p = 3$ the equations above become identities. We then have the result that the p-gons inscribed

in a curve given by eq. (20) are the equilateral triangles with coordinates of their second vertices in the range

$$\frac{2\pi}{3} < \omega < \frac{4\pi}{3} \ . \tag{31}$$

We now require the inscription of a regular tetrahedra which fourth vertex is the point

$$\mathbf{r}_3 = (r(\omega)\cos 3\omega, r(\omega)\sin 3\omega, 3h(\omega)) \ . \tag{32}$$

After connecting this point to the origin, we get the additional equation

$$a^2 = 2r^2(1-\cos 3\omega) + 9h^2 \ . \tag{33}$$

The solution of the system of equations (26), (27) with $p = 3$ and (33) for an arbitrary function $r(\omega)$ is

$$\omega = \omega_R = \pi \pm \arccos\left(\frac{2}{3}\right) \tag{34}$$

$$a(\omega_R) = \frac{10\sqrt{3}}{9} r(\omega_R) \tag{35}$$

$$h(\omega_R) = \frac{\sqrt{30}}{9} r(\omega_R) \ . \tag{36}$$

The two tetrahedra given by eq. (34) are mirror images. This solution is also valid for a sequence of regular tetrahedra glued together at common faces.

Our second set of solutions corresponds to work with configurations like those given by eqs. (10). All the unit vectors have been arranged with the same origin. The following conditions as derived from fig. (1) should be satisfied:

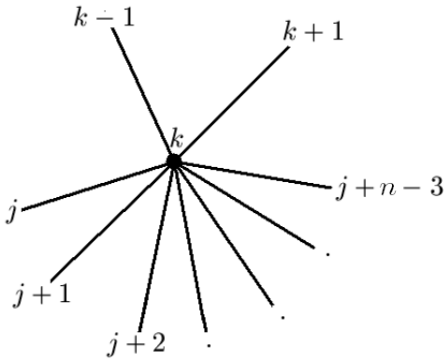


Fig. 1. A configuration of n unit vectors where all the $\binom{n}{2}$ angles are equal.

$$\frac{(\mathbf{r}_{j+l}-\mathbf{r}_k)}{\|\mathbf{r}_{j+l}-\mathbf{r}_k\|}\cdot\frac{(\mathbf{r}_{j+m}-\mathbf{r}_k)}{\|\mathbf{r}_{j+m}-\mathbf{r}_k\|}=-\frac{1}{n-1} \tag{37}$$

$$\frac{(\mathbf{r}_{j+l}-\mathbf{r}_k)}{\|\mathbf{r}_{j+l}-\mathbf{r}_k\|}\cdot\frac{(\mathbf{r}_{k\pm1}-\mathbf{r}_k)}{\|\mathbf{r}_{k\pm1}-\mathbf{r}_k\|}=-\frac{1}{n-1} \tag{38}$$

$$\frac{(\mathbf{r}_{k+1}-\mathbf{r}_k)}{\|\mathbf{r}_{k+1}-\mathbf{r}_k\|}\cdot\frac{(\mathbf{r}_{k-1}-\mathbf{r}_k)}{\|\mathbf{r}_{k-1}-\mathbf{r}_k\|}=-\frac{1}{n-1}\ , \tag{39}$$

where $l=0,1,2,\ldots,p-3$.

The set of $\binom{n}{2}$ equations (10) is given here by the set of eqs. (37)-(39). These are $\binom{n-2}{2}$, $2(n-2)$, 1, equations respectively on each subset and we have trivially

$$\binom{n}{2}\equiv\binom{n-2}{2}+2(n-2)+1\ . \tag{40}$$

After using an Ansatz like that given by eq. (20), eqs. (37)–(39) can be also written as

$$(n-1)^2[\,l,m]^2=[\,l\,][m] \tag{41}$$

$$(n-1)^2\{\,l\,\}_+^2=[\,l\,][\,\cdot\,] \tag{42}$$

$$(n-1)^2\{\,l\,\}_-^2=[\,l\,][\,\cdot\,] \tag{43}$$

$$(n-2)h^2=2r^2(1-\cos\omega)(1-(n-1)\cos\omega)\ , \tag{44}$$

where $l,m=0,1,2,\ldots,(n-3)$.

The symbols which appear in eqs. (41)-(44) can be written

$$\begin{aligned}[\,l,m]&=r^2[1+\cos(l-m)\omega-\cos(j+l-k)\omega-\cos(j+m-k)\omega]+\\&\quad+(j+l-k)(j+m-k)h^2\end{aligned} \tag{45}$$

$$\begin{aligned}\{\,l\,\}_+&=r^2(\cos(j+l-k-1)\omega-\cos(j+l-k)\omega+1-\cos\omega)+\\&\quad+(j+l-k)h^2\end{aligned} \tag{46}$$

$$\begin{aligned}\{\,l\,\}_-&=r^2(\cos(j+l-k+1)\omega-\cos(j+l-k)\omega+1-\cos\omega)-\\&\quad-(j+l-k)h^2\end{aligned} \tag{47}$$

$$[\,l\,]=r^2(1-2\cos(j+l-k)\omega)+(j+l-k)^2h^2 \tag{48}$$

$$[\,\cdot\,]=2r^2(1-\cos\omega)+h^2\ . \tag{49}$$

From eqs. (42) and (43), we get:

$$\{\,l\,\}_+=\{\,l\,\}_- \tag{50}$$

or

$$r^2\sin(j+l-k)\omega\sin\omega+(j+l-k)h^2=0\ . \tag{51}$$

We can now use eqs. (44) and (51) into eqs. (42) and

$$(n-1)^2\{\,l\,\}_+^2\{m\}_+^2 = [\,l,m]^2[\,\cdot\,]^2 \ . \tag{52}$$

We then have,

$$(n-3)^2A^2 = B^2 \tag{53}$$

and

$$(n-4)A^2 = 4B\left(A + \frac{n}{4}B\right) \ , \quad n \neq 2 \ , \tag{54}$$

where

$$A = (1-\cos(j+l-k)\omega)(1-\cos(j+m-k)\omega) \tag{55}$$

$$B = \frac{(1-\cos\omega)}{(1-(n-1)\cos\omega)}\sin(j+l-k)\omega \ \sin(j+m-k)\omega \ . \tag{56}$$

The feasible n-values are given by the equation

$$(n-3)(n^3-6n^2+8n+4) = 4(n-3)^2 \tag{57}$$

or

$$n = 3, 4 \ . \tag{58}$$

This result has a deep significance for the structure of biomacromolecules. For instance, the C_α atoms in the structure of an aminoacid are in the centre of a regular tetrahedron. The placement of the other atoms correspond to its vertices. This structure is a natural realization of the $n = 4$ case. The placement of carbon and the nitrogen of the carbonyl group are also a realization for the $n = 3$ case.

The third set of solutions corresponds to 3-dimensional version of the second set. Some information about this set was already presented into eqs. (32)-(36). We look for a relation like that

$$\mathbf{r}_{j+l} = A_1\mathbf{r}_{j+l-1} + A_2\mathbf{r}_{j+l-2} + \ldots + A_{l-1}\mathbf{r}_{j+1} + A_l\mathbf{r}_j \ . \tag{59}$$

This means that a l-th vector in a sequence will be given as a linear combination of the l previous vectors.

The third component of the vectors will satisfy

$$(j+l)h = A_1(j+l-1)h + A_2(j+l-2)h + \ldots + A_{l-1}(j+1)h + A_l jh \ . \tag{60}$$

The same identities for the two other coordinates can be written with an Argand representation,

$$rz_{j+l} = A_1 rz_{j+l-1} + A_2 rz_{j+l-2} + \ldots + A_{l-1} rz_{j+1} + A_l rz_j \tag{61}$$

where

$$z_j = (z)^j = e^{ij\omega} \ . \tag{62}$$

From eq. (61), we get:

$$z^l - \sum_{k=1}^{l} A_k z^{l-k} = 0 \ . \tag{63}$$

If we write eq. (60) for two sequences beginning at $j = j_1$ and $j = j_2$, we get

$$\sum_{k=1}^{l} A_k = 1 \ . \tag{64}$$

From eq. (64) we can see that eq. (63) has a $z = 1$ root:

$$(z-1)\left[z^{l-1} - \sum_{k=1}^{l-1}(A_1 + A_2 + \ldots + A_k - 1)z^{l-(k+1)}\right] = 0 \ . \tag{65}$$

From eqs. (64) and (60), we can write,

$$\sum_{k=1}^{l} k A_k = 0 \ . \tag{66}$$

From eqs. (66) and (64), we can see that eq. (63) has a second $z = 1$ root

$$(z-1)^2\left[z^{l-2} - \sum_{k=1}^{l-2}(kA_1 + (k-1)A_2 + \ldots + A_k - (k+1))z^{l-(k+2)}\right] = 0 \ . \tag{67}$$

A trivial solution will follows for the case $l = 4$. We have

$$z^2 + (2 - A_1)z + 3 - 2A_1 - A_2 = 0 \ . \tag{68}$$

For complex roots of unit modulus $|z| = 1$ like those of eq. (67),

$$A_1^2 + 4A_1 + 4A_2 - 8 < 0 \tag{69}$$

and

$$A_1 = 2(1 + \cos\omega) \ , \quad 0 \leq A_1 \leq 4 \ . \tag{70}$$

The two last equations lead to

$$2A_1 + A_2 = 0 \ . \tag{71}$$

We now form a system to be solved with eq. (71) and eqs. (64) and (65) for $l = 4$. We have,

$$A_2 = 2 - A_1 \ , \quad A_3 = A_1 \ , \quad A_4 = -1 \ . \tag{72}$$

As a result, we can write for the linear combination satisfied by the sequence of position vectors given by eq. (20) for $l = 4$:

$$\mathbf{r}_{j+4} = A_1\mathbf{r}_{j+3} + 2(1 - A_1)\mathbf{r}_{j+2} + A_1\mathbf{r}_{j+1} - \mathbf{r}_j \ . \tag{73}$$

The existence of the last relation is enough for thinking about generic tetrahedra as the elementary "cells" to be constructed with edges of equal length and vertices given by eq. (20). For n vertices, there are $n_T = n - 3$ tetrahedra.

There are two values corresponding to regular tetrahedra. They are:

$$\omega = \omega_R = \pi \pm \arccos\left(\frac{2}{3}\right) \ . \tag{74}$$

These values will correspond to two sequences of regular tetrahedra, which are joined together at common faces along each sequence. Actually, they correspond to the same geometric structure. This structure is also chiral for $n \geq 6$. It is known in the literature by the name of 3-sausage [7]. There are non-chiral configurations with a 2-fold axis of symmetry ($n = 3$) and a 3-fold axis ($n = 4, 5$). The structures with values $\pm\omega$ and $\omega \neq \omega_R$ are sequences of non-regular tetrahedra. They are chiral themselves and chiral to each other for $n \geq 3$.

4 Steiner Points and Steiner Trees

Let M be a metric manifold and A a finite set of points on it. We consider the subsets of A such that each pair of points on them could be connected by an edge of minimal length. These edges will be geodesic arcs of the manifold M. A tree is a collection of points and their connecting edges. From this definition we can see that loops should be discarded. A spanning tree (SP) of a subset is a tree which connects all the points of this subset. Among all possible spanning trees of a set A with length $l_{\mathrm{SP}}(s, A)$, there is at least one which overall length is a minimum. This will be called the Minimum Spanning Tree of a set A, MST(A). Its length will be given by

$$l_{\mathrm{MST}}(A) = \min_{(s-\mathrm{trees})} l_{\mathrm{SP}}(s, A) \ . \tag{75}$$

If in the search of a minimum spanning tree, we allow for the introduction of additional points of the manifold M on each set A, we can get spanning trees of smaller overall length. A Steiner tree is defined with the additional requirements of three tangent lines to three geodesic edges meeting at an angle of 120° on each additional (Steiner) point. Among all the Steiner trees t of a set A which length is $l_{\mathrm{ST}}(t, A)$, there is one which overall length is minimum. This is the Steiner Minimal tree of the set A, SMT(A). Its length is given by

$$l_{\mathrm{SMT}}(A) = \min_{(t-\mathrm{trees})} l_{\mathrm{ST}}(t, A) \ . \tag{76}$$

The Minimum Spanning tree MST(A) is considered to be the worst approximation to the Steiner Minimal tree, SMT(A) for each set $A \subset M$. This

approximation is also called the "worst cut". A usual measure for this approximation is the Steiner Ratio of the set $A \subset M$. It is given by

$$\rho(A) = \frac{l_{\mathrm{SMT}}(A)}{l_{\mathrm{MST}}(A)} \ . \tag{77}$$

The infimum of all values $\rho(A)$ for all sets A is the Steiner Ratio ρ_M of the manifold M or

$$\rho_M = \inf_{A \subset M} \rho(A) \ . \tag{78}$$

In order to implement these ideas we take a generic curve and points evenly spaced along it [8]. These points will be also evenly spaced according the Euclidean distance introduced in $\mathbb{R}^3$ for curves $\mathbf{r}(\phi) = (x(\phi), y(\phi), z(\phi))$ such that

$$x'^2 + y'^2 + z'^2 = a^2 \tag{79}$$

where a is a constant and $(') = \frac{d}{d\phi}$.

For points $\mathbf{r}_j = (x_j(\phi), y_j(\phi), z_j(\phi)) = (x(\phi_j), y(\phi_j), z(\phi_j))$, eq. (79) is enough to satisfy the requirement of evenly spaced consecutive points or

$$\|\mathbf{r}_{j+2} - \mathbf{r}_{j+1}\| = \|\mathbf{r}_{j+1} - \mathbf{r}_j\| \tag{80}$$

where $\|\cdot\|$ stands for the Euclidean norm.

We shall form a full Steiner tree [7] with n external vertices and $(n-2)$ Steiner nodes which position vectors satisfy eq. (80) and are given by the Ansatz:

$$\mathbf{r}_j(\phi) = (r(\phi)\cos(j\phi), r(\phi)\sin(j\phi), jh(\phi)) \ , \quad 0 \le j \le n-1 \tag{81}$$

$$\mathbf{r}_{S_k}(\phi) = (r_S(\phi)\cos(k\phi), r_S(\phi)\sin(k\phi), kh_S(\phi)) \ , \quad 1 \le k \le n-2 \ . \tag{82}$$

We now assume a fishbone or path-topology [7] for the Steiner tree. This means that the point $\mathbf{r}_{S_1}$ is connected to the points $\mathbf{r}_0$ and $\mathbf{r}_1$; the point $\mathbf{r}_{S_{n-1}}$ is connected to $\mathbf{r}_{n-2}$ and $\mathbf{r}_{n-1}$. The points $\mathbf{r}_{S_k}$ are connected consecutively. The vertice of the intermediate couples $\mathbf{r}_j$, $\mathbf{r}_{S_k}$ with $j = k$ are also connected. The assumption of a path-topology will be commented below. We are also to provide a scheme in order to favour the modelling of protein structures.

From the requirement of meeting edges at Steiner points with angles of $2\pi/3$, we have,

$$h_S \equiv h \tag{83}$$

$$h_S^2 = r_S^2 A_1 (A_1 + 1) \tag{84}$$

where

$$A_1 = 1 - 2\cos\phi \ . \tag{85}$$

Furthermore, the requirement for full Steiner trees with $(n-2)$ vertices should be guaranteed by a property to be satisfied by the associated spanning

tree: the smallest angle θ_1 between two consecutive edges whose vertices $\mathbf{r}_j$ are given by eq. (81), should be lesser than $2\pi/3$. We can write

$$\cos\theta_1 = -1 + \frac{(A_1+1)^2}{2[F^2+A+1]} > -\frac{1}{2}\,,\quad F \equiv \frac{h}{r}\,. \tag{86}$$

The last equation can be also written in the form

$$\cos\theta_1 = \max\left(-\frac{1}{2}, -1 + \frac{(A_1+1)^2}{2[F^2+A+1]}\right)\quad . \tag{87}$$

This modelling will depends on the function $F(\phi) \equiv h(\phi)/r(\phi)$. However, we can circumvent this dependence with a new modelling which operates with two variables [8], $(\phi, F(\phi) \equiv h(\phi)/r(\phi))$.

After connecting all the consecutive points which position vectors are given by eq. (81), we can write for a candidate of spanning tree:

$$l_{S_{P_1}} = (n-1)\sqrt{F^2+A_1+1}\ . \tag{88}$$

Since we have adopted a path-topology for the Steiner tree, it will be straightforward to write for its Euclidean length after using eqs. (81)–(84) as

$$l_{S_{T_1}} = (n-2)r + [(n-3)A_1 - 1]\frac{F}{\sqrt{A_1(A_1+1)}} + \\ +2r\sqrt{1 + (A_1+1)\frac{F}{\sqrt{A_1(A_1+1)}} + (A_1^2+A_1+1)\frac{F^2}{A_1(A_1+1)}} \tag{89}$$

For a large set of points we will have, instead eqs. (88) and (89),

$$l_{S_{P_1}} = rn\sqrt{F^2+A_1+1} \tag{90}$$

$$l_{S_{T_1}} = rn\left(1 + F\sqrt{\frac{A_1}{A_1+1}}\right)\quad . \tag{91}$$

We notice that the ratio $l_{S_{T_1}}/l_{S_{P_1}}$ does not correspond to a candidate for a Steiner Ratio Function. It will be just the convex envelope of the function which will be derived by following the prescription given into eqs. (75)–(78) as will be shown in the next section. However, this convex envelope function will be useful in order to define the domain of the ϕ-variable from the lower and upper values of the Steiner Ratio.

Let us call ρ_e the function of the convex envelope for $n \gg 1$. We have from eqs. (90) and (91),

$$\rho_e = \frac{l_{S_{T_1}}}{l_{S_{P_1}}} = \frac{1 + F\sqrt{\frac{A_1}{A_1+1}}}{\sqrt{F^2+A_1+1}}\ . \tag{92}$$

The extreme of this function will be obtained by solving the equation

$$0 = \frac{d\rho_e}{d\phi}$$
$$= -\frac{(F - \sqrt{A_1(A_1+1)})}{\sqrt{(F^2+A_1+1)^3}} \left\{ \frac{dF}{d\phi} - \sin\phi \frac{[F(F+\sqrt{A_1(A_1+1)}) + A_1 + 1]}{\sqrt{A_1(A_1+1)^3}} \right\} \quad (93)$$

The first factor in eq. (93 leads to $F = \sqrt{A_1(A_1+1)}$ which means $\rho_e = 1$. The second factor gives after integration

$$F(\phi) = \sqrt{A_1+1}\tan\left(\arctan\sqrt{A_1} + \alpha\right) \quad (94)$$

where α is a constant.

From eq. (92) and for the ω-values which satisfy eq. (94), we have,

$$\rho_e = \cos\alpha \ . \quad (95)$$

If b and B are the lower and upper bounds to be imposed on the Steiner Ratio, they can be also imposed on the convex envelope function or

$$b \leq \rho_e \leq B \ . \quad (96)$$

The largest region of ρ_e value corresponds to $b = 0.5$ (Moore's bound) and $B = 1$. We then have,

$$\frac{\pi}{3} \geq \alpha \geq 0 \ . \quad (97)$$

If we choose to work with a two variable modelling, the conditions for an extremum can be written.

$$0 = \frac{\partial\rho_e}{\partial\phi}$$
$$= \frac{\sin\phi}{\sqrt{A_1(A_1+1)^3(F^2+A_1+1)^3}} \left[F(F^2 - A_1^2 + 1) - \sqrt{A_1(A_1+1)^3} \right] \quad (98)$$

$$0 = \frac{\partial\rho_e}{\partial F} = -\frac{\left(F - \sqrt{A_1(A_1+1)}\right)}{\sqrt{(F^2+A_1+1)^3}} \quad (99)$$

The extrema of the surface $\rho(F,\phi)$ will be on the curve $F = \sqrt{A_1(A_1+1)}$. In order to analyze if these extrema will correspond to maxima or minima, we consider the curve,

$$F = \sqrt{A_1(A_1+1)} + \eta \quad (100)$$

where η is a real number.

The values (F,ϕ) which satisfy eq. (100) with $\eta \neq 0$ are in the hypergraph ($\eta > 0$) or the subgraph ($\eta < 0$) of the functions $F = \sqrt{A_1(A_1+1)}$. We now assume that $\rho(F,\phi) > 1$ for these values and we get from eqs. (92) and (100) that $\eta^2 < 0$. This absurd will lead us to the conclusion that $\rho(F,\phi) < 1$ for all those values (F,ϕ) and $F = \sqrt{A_1(A_1+1)}$ is a curve or relative maxima.

5 Generic Sequences of Points and the Construction of the Steiner Ratio Function

This section starts from the assumption that for a large number of points, every sequence can be considered as formed by subsequences of evenly spaced points. The method to be introduced for constructing these subsequences has very useful applications in the modelling of protein structure where the vertices of subsequences correspond to the placement of atoms in the backbone of α-helices.

In order to realize all the subsequent calculations, we extend our representation of coordinates of the position vectors given into eqs. (81) and (82) to represent surfaces. We can now write for the position vectors of evenly spaced consecutive points:

$$\mathbf{r}_j(\phi, h) = (r(\phi, h)\cos(j\phi), r(\phi, h)\sin(j\phi), jh),\ 0 \le j \le n-1 \tag{101}$$

$$\mathbf{r}_{S_k}(\phi, h_S) = (r_S(\phi, h)\cos(k\phi), r_S(\phi, h_S)\sin(k\phi), kh_S),\ 1 \le k \le n-2 \tag{102}$$

The requirement of meeting edges at each Steiner point with angles $2\pi/3$ leads to analogous equations of eqs. (83) and (84).

We are now ready for introducing the sub-sequences of evenly spaced but non-consecutive points [9]. These subsequences will be represented $(P_j)_{m,l_{Pmax}}$ and $(S_k)_{m,l_{Smax}}$ and can be written as

$$\mathbf{r}_j,\ \mathbf{r}_{j+m}, \mathbf{r}_{j+2m}, \ldots, \mathbf{r}_{j+l_P m}, \ldots, \mathbf{r}_{j+l_{Pmax} m}\,,\quad 0 \le j \le m-1 \tag{103}$$

$$\mathbf{S}_k, \mathbf{S}_{k+m}, \mathbf{S}_{k+2m}, \ldots, \mathbf{S}_{k+l_S m}, \ldots, \mathbf{S}_{k+l_{Smax} m}\,,\quad 1 \le k \le m\ , \tag{104}$$

respectively.

There are m subsequences on each equation above, as can be easily seen. The value $(m-1)$ is the number of skipped points necessary to form the subsequence. The values l_{Pmax} and l_{Smax} can be given from the restriction imposed on the generic integer index of the position vectors of eqs. (103) and (104). We have,

$$j + l_P m \le n-1\ ,\quad 0 \le l_P \le l_{Pmax}\ , \tag{105}$$

$$k + l_S m \le n-2\ ,\quad 0 \le l_S \le l_{Smax}\ . \tag{106}$$

It follows from the last equations that

$$l_{Pmax} = \left[\frac{n-j-1}{m}\right]\ ;\quad 0 \le j \le m-1\ , \tag{107}$$

$$l_{Smax} = \left[\frac{n-k-2}{m}\right]\ ;\quad 1 \le k \le m\ . \tag{108}$$

The square brackets in the last equations $[x]$ stand for the greatest integer value $\le x$.

The number of points on the subsequences given by eqs. (103) and (104) is $(l_{Pmax}+1)$ and $(l_{Smax}+1)$, respectively. We define now sequences of n and $(n-2)$ points instead those given by eqs. (101) and (102), respectively. They are given by

$$\mathbb{P}_m = \bigcup_{j=0}^{m-1} (P_j)_{m,l_{Pmax}} \quad , \tag{109}$$

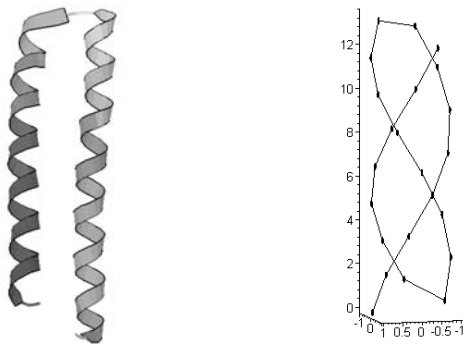
$$\mathbb{S}_m = \bigcup_{k=1}^{m} (S_k)_{m,l_{Smax}} \quad . \tag{110}$$

This is inspired by the linking of α-helices and β-sheets by turns in the tertiary structure of a protein. The original sequences of eqs. (101) and (96) are trivially included in the scheme presented here and are given by $\mathbb{P}_1 = (P_0)_{1,n-1}$ and $\mathbb{S}_1 = (S_1)_{1,n-3}$, respectively. An elementary check can be made in order to show that the composed sequences have the same number of points of the original sequences. We have,

$$\sum_{j=0}^{m-1} (l_{Pmax}+1) = m + \sum_{j=0}^{m-1} \left[\frac{n-j-1}{m}\right] = m+n-m = n\,, \tag{111}$$

$$\sum_{k=1}^{m} (l_{Smax}+1) = m + \sum_{k=1}^{m} \left[\frac{n-k-2}{m}\right] = m+n-m-2 = n-2 \tag{112}$$

In fig. (2) we can see the comparison of a tertiary structure of a protein and the sequence $\mathbb{P}_3$ with $n = 23$ points which is formed by the union of the subsequences $(P_0)_{3,7}$, $(P_1)_{3,7}$ and $(P_2)_{3,6}$.



(a) A tertiary structure of a protein

(b) The $\mathbb{P}_3$ sequence

Fig. 2. The turns which link the α-helices are modelled here by segments connecting consecutive points at the ends of two subsequences.

A path-topology will be also adopted for the Steiner trees organized with the subsequences given by eqs. (103) and (104). The first point S_k will be connected to the two first points $\mathbf{r}_j$ and $\mathbf{r}_{j+m}$, $\forall k$, j. The last point $\mathbf{S}_{k+l_{Smax}m}$ will be connected to the two last points $\mathbf{r}_{j+(l_{Pmax}-1)m}$ and $\mathbf{r}_{j+l_{Pmax}m}$, $\forall k$, j. The points $\mathbf{S}_{k+l_Sm}$ are connected consecutively $\forall k$, $1 \leq l_S \leq m$. All the intermediate couples of points such as $\mathbf{r}_{j+l_Pm}$, $\mathbf{S}_{k+l_Sm}$ with $j = k$ and $2 \leq l_P = l_S \leq m-2$, $\forall j$. All the connection topologies for the resulting trees formed from the new sequences $\mathbb{P}_m$ and $\mathbb{S}_m$ can be obtained by connecting conveniently the subsequences. The Euclidean length of the resulting trees does not depend on the special combination chosen for the subsequences.

The coordinates of points of the subsequences can be written analogously to eqs. (103) and (104), we get,

$$\mathbf{r}_{j+l_Pm} = (r(\phi,h)\cos(j+l_Pm)\phi, r(\phi,h)\sin(j+l_Pm)\phi, jh)\,, \quad 0 \leq j \leq m-1\,, \tag{113}$$

$$\mathbf{S}_{k+l_Sm} = (r_{S_m}(\phi,h_{S_m})\cos(k+l_Sm)\phi, r_{S_m}(\phi,h_{S_m})\sin(k+l_Sm)\phi, kh_{S_m})\,, \quad 1 \leq k \leq m\,. \tag{114}$$

The requirement of meeting edges at angles of $2\pi/3$ on each Steiner point results,

$$h_{S_m} = h\ , \tag{115}$$

$$m^2 h_{S_m}^2 = r_{S_m}^2 A_m(A_m+1)\ , \tag{116}$$

where

$$A_m = 1 - 2\cos(m\phi)\ . \tag{117}$$

The restriction to full Steiner trees with $(n-2)$ Steiner points gives for the angle θ_m between consecutive edges of the spanning tree of a subsequence

$$-\frac{1}{2} < \cos\theta_m = -1 + \frac{(A_m+1)^2}{2\left[m^2F^2 + A_m + 1\right]}\ . \tag{118}$$

This can be also written as

$$\cos\theta_m = \max\left(-\frac{1}{2}, -1 + \frac{(A_m+1)^2}{2\left[m^2F^2 + A_m + 1\right]}\right)\ . \tag{119}$$

The Euclidean length of the spanning tree $\mathbb{P}_m$ corresponding to the union of the subsequences $(P_j)_{m,l_{Pmax}}$, $0 \leq j \leq m-1$ is given by

$$l_{SP_m} = r_m\left[(n-m)\sqrt{m^2F^2 + A_m + 1} + (m-1)\sqrt{F^2 + A_1 + 1}\right]\ . \tag{120}$$

We notice the contribution of the turns between α-helices to our modelling. They correspond to the last term into eq. (120). The value $(m-1)$ is the number of skipped points necessary to form this union sequence. This is also the number of turns necessary to connect the subsequences.

The Euclidean length of the Steiner Tree organized with the sequences $\mathbb{P}_m$ and $\mathbb{S}_m$ according the path-topology adopted for each sub-tree formed by the subsequences $(P_j)_{m,l_{Pmax}}$ and $(S_k)_{m,l_{Smax}}$ can be written

$$l_{ST_m} = r_m \left[n - 2 + [(n-m-2)A_m - m]\, m \frac{F}{\sqrt{A_m(A_m+1)}} + \right.$$
$$\left. + 2\sqrt{1 + (A_m+1)\frac{F}{\sqrt{A_m(A_m+1)}} + (A_m^2 + A_m + 1)\frac{F^2}{A_m(A_m+1)}} \right] \quad (121)$$

The present sort of modelling has many interesting features. The application of these methods to biomolecular structure can provide the unveiling of patterns and mechanisms which are used by Nature to stabilize the biomolecules.

By following the usual prescription for a Steiner Ratio, as given by eqs. (4.3), (78), we have

$$\rho(\phi, h) = \frac{\min_{(m)}\{l_{ST_m}\}}{\min_{(m)}\{l_{SP_m}\}} \,, \quad (122)$$

where $\{l_{ST_m}\}$, $\{l_{SP_m}\}$ represent the set of Euclidean distances corresponding to $\mathbb{S}_m$ and $\mathbb{P}_m$ respectively for all values of m. They are given by eqs. (120) and (121). The "$\min_{(m)}$" process should be understood in terms of a new function formed in a piecewise way from the functions which correspond to chosen m-values. Actually, the modelling is only effective for $m = 1, 2, 3$, since there is a maximum of three surfaces meeting at a point.

This formula for the Steiner Ratio can be also simplified by taking into consideration another restriction imposed by eq. (119): In order to form a Steiner tree, the only feasible value of m is $m = 1$, since this corresponds to the largest region of feasible ϕ-values according the lower and upper bounds of $\rho(\phi, h)$. We then get, instead eq. (122),

$$\rho(\phi, h) = \frac{l_{ST_1}}{\min_{(m)}\{l_{SP_m}\}} \,. \quad (123)$$

6 An Unconstrained Optimization Problem. The Euclidean Steiner Ratio for Very Large Number of Points

We write eq. (123) in the limit of $n \gg 1$. We have from eqs. (120) and (121),

$$\rho(\phi, h) = \frac{r + h\sqrt{\frac{A_1}{A_1+1}}}{\min_{(m)}\left\{\sqrt{m^2h^2 + r^2(A_m+1)}\right\}} \,. \quad (124)$$

The lower (b) and upper (B) bounds to be imposed on the convex envelope function (eq. (92)) of the function given into eq. (124) can be shown to satisfy

$$\frac{1-b^2}{b^2} \geq \frac{\left(h_S - r\sqrt{A_1(A_1+1)}\right)^2}{\left(h_S\sqrt{A_1} + r\sqrt{A_1+1}\right)^2} \geq \frac{1-B^2}{B^2} \ . \tag{125}$$

From eqs. (125) and (84), we can also write,

$$r \geq r_S \geq \frac{\left(\sqrt{A_1} - \frac{\sqrt{1-b^2}}{b}\right) r}{\sqrt{A_1}\left(1 + \frac{\sqrt{A_1(1-b^2)}}{b}\right)} \ . \tag{126}$$

Since the first inequality should be always satisfied from eq. (86), it remains to require that

$$\sqrt{A} - \frac{\sqrt{1-b^2}}{b} \geq 0 \ , \tag{127}$$

or

$$\cos\omega \leq 1 - \frac{1}{2b^2} \leq 1 - \frac{1}{2B^2} \ . \tag{128}$$

Table 1. A range of ϕ-values to include Du's greatest lower bound [10]. It also includes Graham-Hwang's [11].

b	$\cos\phi$
1/2	-1
0.522232	$-5/6$
0.547722	$-4/6$
$\sqrt{3}/3$	$-3/6$
0.612372	$-2/6$
0.654653	$-1/6$

In the range of ϕ-values of table (1), we can write eq. (124) in the following form:

$$\rho(\phi, h) = \max_{(m)} \rho_m(\phi, h) = \max_{(m)} \left\{ \frac{1 + F\sqrt{\frac{A_1}{A_1+1}}}{\sqrt{m^2F^2 + A_m + 1}} \right\} \ . \tag{129}$$

This unconstrained optimization problem needs a non-derivative method in order to be solved [8, 12]. Fortunately, it can also be solved for the special function of eq. (129) by noticing that the $\rho_m(\phi, h)$, $m = 1, 2, 3$, intersect at a point of coordinates

$$\phi_R = \pi - \arccos\left(\frac{2}{3}\right) \ , \quad h_R = \frac{\sqrt{30}}{9}$$

$$\rho(\phi_R, h_R) = \frac{1}{10}\left(3\sqrt{3} + \sqrt{7}\right) = 0.78419037337\ldots \tag{130}$$

We now form the compact domain

$$\left(H_{\rho_1} - \bigcup_{m \geq 2} H_{\rho_m} \right) \bigcap \left\{ (\phi, h) \,\middle|\, \pi - \arccos\left(\frac{1}{6}\right) \leq \phi \leq \pi + \arccos\left(\frac{1}{6}\right) \right\}, \tag{131}$$

where H_{ρ_1}, H_{ρ_m} are the subgraphs of the functions $\rho_1 = 1$, $\rho_m = 1$, $m \geq 2$. $(A - B)$ is the set which points belong to A but not to B. The functions $\rho_m(\phi, h)$, $\forall m$ are continuously decreasing from the boundaries of this compact towards its interior. By the Weierstrass theorem, there will be a Global Minimum in this domain [13]. The uniqueness will follows since the surfaces $\rho_m(\phi, h)$, $m = 1, 2, 3$, intersect only at the point given by eq. (130).

7 Concluding Remarks

The use of Steiner Trees for modelling the biomolecular structures is a research topic which gives several insights for solving many Optimization problems since there is always an analogous problem which must have been solved by Nature. This is one of best examples of the two way road of Mathematical Biology: to understand biological phenomena with powerful mathematical methods, as well as to develop more mathematics with the insights provided by Nature. This is in the best tradition of true interdisciplinary science. The research on Steiner Trees has provided a deep introduction to the study of a full geometrical foundation of intramolecular interactions and can unveil the foundations of a unified treatment only dreamed by men like Schrödinger, Rashevsky and D'arcy Thompson. There are many special problems to be studied, from amide planes to the discovery of special surfaces inside the biomolecular structure. From the discussion of the problem with a finite number of points to its application to the modelling of proteins and the dynamics of protein folding. We think that it is worthwhile to continue the work in this research field. Some new results on the topics mentioned above will be published elsewhere.

References

1. Thompson, D.W.: On Growth and Form. Dover, New York (1992)
2. Rashevsky, N.: Mathematical Biophysics: Physico-Mathematical Foundations of Biology. Dover, New York (1960)
3. Schrödinger, E.: What is Life? Cambridge U.P., Cambridge (1992)
4. Anfinsen, C.B.: Principles that govern the Folding of Protein Chains (Nobel Lecture). Science, **181**, 223–230 (1973)
5. Courant, R., Robbins, H.: What is Mathematics? Oxford U.P., Oxford (1978)
6. Kuhn, H.W.: Steiner's Problem Revisited. In: Dantzig, G.B., Eaves, B.C. (eds.) Studies in Mathematics, vol. 10, Studies in Optimization. The Mathematical Association of America (1974)

7. Du, D.-Z., Smith, W.D.: Disproofs of Generalized Gilbert-Pollak Conjecture on the Steiner Ratio in Three or more Dimensions. Journ. Comb. Theory **A74**, 115–130 (1996)
8. Mondaini, R.P.: An Analytical Method for derivation of the Steiner Ratio of 3D Euclidean Steiner Trees. Journ. Glob. Optim., *in press* (2007)
9. Mondaini, R.P.: Steiner Trees as Intramolecular Networks of the Biomacromolecular Structures. In: BIOMAT 2005, International Symposium on Mathematical and Computational Biology. World Scientific Co. Pte. Ltd., 327–342 (2006)
10. Du, D.-Z.: On Steiner Ratio Conjectures. Annals of Operations Research **33(6)**, 437–449 (1991)
11. Graham, R.L., Hwang, F.K.: Remarks on Steiner Minimal Trees I. Bull. Inst. Acad. Sinica **4**, 177–182 (1976)
12. Walsh, G.R.: Methods of Optimization. Wiley (1975)
13. Bazaraa, M.S., Sherali, H.D., Shetty, C.M.: Nonlinear Programming: Theory and Algorithms. Wiley, 2nd ed. (1993)

Phenotypic Switching and Mutation in the Presence of a Biocide: No Replication of Phenotypic Variant

Brenda Tapia-Santos[1] and Jorge X. Velasco-Hernández[2]

[1] Facultad de Matemáticas, Universidad Veracruzana, Xalapa, Ver., A.P. 270, C.P. 91090 México, bretasa@gmail.com
[2] Programa de Matemáticas Aplicadas y Computación, Instituto Mexicano del Petróleo (IMP), México, D.F. 07730, México, velascoj@imp.mx

Summary. A model for three competing bacterial strains that incorporates mutation and/or phenotypic switching is studied. We consider three different strains: wild, mutated and phenotypic bacteria generated by an inhibitor introduced in the environment. Our model considers that all new phenotypic bacteria are sensitive to the inhibitor and there is no phenotypic replication. Two steady state regimes are identified, finding that the strain surviving is the one arising from mutation of the wild strain. The model may also show three steady state regimes with the persistence of the three bacteria in the system.

Key words: Phenotypic switch, mutation, competition, inhibitor, chemostat.

1 Introduction

A simple example of competition occurs in the chemostat [8–10, 13, 17], where we can study the mechanisms of population interaction under simplified, controllable laboratory conditions. These conditions, for example the application of an inhibitor, are imposed by the experimentalist. For this reason, the chemostat is an important laboratory piece for applied sciences. In medicine, for example, it is used to study a common problem: one strain of bacteria to be affected by an antibiotic while another is resistant, in this problem, one wishes to determinate whether the resistant strain will outcompete the nonresistant strain [16]. If it can, the antibiotic will not be effective in treating the disease. The original model about this problem is that of Lenski and Hattingh [11]; but there are more authors, as Hsu and Waltman [5] and Hsu, Li and Waltman [4], who have worked in models where the inhibitor is introduced to the system in an external form.

An alternative problem happens when the inhibitor is produced for one competitor at some cost to its own growth [3, 6].

In both of two problems the inhibitor can be lethal or not lethal, depending on the interference of the inhibitor with the reproduction of the organism. Works cited before ([3–6, 11]) consider models with two independent strains, one of them sensitive to an inhibitor and the other resistant to it.

However, the effect of the inhibitor generate resistance and mutate (like the resistance presented by *E. coli* to quinolonas [15,20]) or may induce a phenotypic switch (for example the resistance of *E. faecium* to glycopeptides [21] or the resistance to ampicillin [1, 12]). Scientific evidence exists about both resistance mechanisms; Miller [15] has shown that during exposure of *E. coli* to β-lactam antibiotics (penicillin, cefuroxan, etc), the bacteria presents a defense mechanism that temporally inhibits cell division, inducing resistance through mutation, limiting the bactericidal effects of these drugs. Balaban *et. al.* [1] provide evidence for the persistence of single cells of *E. coli* exposed to penicillin, they shown that the fraction of cells surviving have not genetically acquired antibiotic resistance, they regrow a new population that is sensitive to the antibiotic [1].

In this paper we address the following problem. In a constant environment, where resources are provided at a constant rate and where we have introduced an inhibitor, are there survival of bacteria when both of the resistance mechanisms (mutation and phenotypic switching) are present?, and, if that is the case, what kind of resistant strain will survive?. These questions we address using a competition model in a chemostat between three populations, that is we consider the competition for one nutrient of the planktonic cells u, v_0 and v_1. We consider an environment where a concentration of biocide (inhibitor) p is introduced and we will observe that, in the general case, it is possible to have the coexistence of the three species.

The paper is organized as follows. Section 2 deals with our model construction and a first study of the existence and stability of steady states. Section 3 includes the second part of the study of the model. In Section 4 we present our numerical results, the conclusions are given in Section 5 and some of the proofs are in Appendix A.

2 Mutation and Phenotypic Change

We consider a competition model in a chemostat, given by system (2), where u indicates a wild type bacteria sensitive to a biocide present in the environment and whose concentration we denote by p. Wild bacteria mutate or undergo phenotypic switching to avoid the action of the biocide. The first mechanism is an adaptation process to the new environment, and it does not depend on the biocide since it is a random phenomena. Here it is modeled with the term μu where μ is the mutation rate. The other mechanism is a consequence of the biocide presence, and it is represented by the term $\alpha(1-\phi(p))$, where α is the

phenotypic switching rate and the term $1-\phi(p)$ indicates us a smaller switching if we have a small quantity of biocide in the environment; v_0 denotes the mutated strain concentration and v_1 indicates the concentration of phenotypic variant. The two latter strains are resistant to the biocide. The competition between the bacteria is resource competition for one nutrient whose concentration is denoted by S. We denote the set of values that our variables can take (phase space) by Ω:

$$\Omega = \{(S, u, v_0, v_1, p) \mid S \geq 0, u \geq 0, v_0 \geq 0, v_1 \geq 0, p \geq 0\} \tag{1}$$

As in [16] we consider that nutrient is added to the chemostat at a constant rate S^0D and is eliminated either by washout at a rate SD, or from consumption by bacteria. Strain u is considered sensitive to the biocide so its rate of consumption is affected by it. This effect is represented by the function $\phi(p)$, the degree of inhibition that the biocide exerts upon strain u, $\phi(p) = e^{-\lambda p}$. Increase of strain u is a function of both u and v_1 strains since the latter is the same strain u in a reproductive sense (phenotypic variant) that is, all new phenotypic bacteria are sensitive to the biocide. The loss term includes the natural death rate D, the mutation rate (occurring at rate μu) and the phenotypic variation that occurs at a rate $\alpha(1-\phi(p))u$.

The strain v_0 is recruited either by consumption of nutrient or by new mutated bacteria at a rate μu. In this case, nutrient intake is different from u and v_1 due to different genetical requirements for each bacteria. The bacteria dies at a rate D.

Increase in numbers by v_1 is due only to phenotypic switching, nutrient intake does not affect its growth since all offspring are of type u. Bacteria of this type dies at a rate D.

Finally, as in [16], p indicates the concentration of biocide added to the chemostat at a constant rate p^0D that is lost due to washout and due to the absorption of biocide by the bacteria v_0 and v_1.

So, considering these assumptions we have constructed the model:

$$S' = (S^{(0)} - S)D - \frac{Sm_1}{a_1 + S}(u\phi(p) + v_1) - \frac{Sm_2 v_0}{a_2 + S} \tag{2}$$

$$u' = \frac{Sm_1}{a_1 + S}(u\phi(p) + v_1) - (\mu + D + \alpha(1 - \phi(p)))\, u \tag{3}$$

$$v_0' = v_0 \left(\frac{Sm_2}{a_2 + S} - D \right) + \mu u \tag{4}$$

$$v_1' = \alpha(1 - \phi(p))u - Dv_1 \tag{5}$$

$$p' = (p^{(0)} - p)D - \frac{hp}{k + p}(v_0 + v_1) \tag{6}$$

Notice that the growth of v_1 appears in the u-equation, by our hypothesis all new phenotypic bacteria are sensitive to the biocide. In the same way, there is not phenotypic replication.

2.1 Identical Consumption Functions

Obviously a_i and m_i, $i = 1, 2$ are the Michaelis-Menten constant and maximal growth rate for the wild and phenotypic bacteria and mutated bacteria, respectively; k and h are the constants involved in the absorption of biocide by the resistant bacteria (v_0 and v_1). To a first approximation and to reduce the number of the parameters, we assume for a moment that $m_1 = m_2 = m$ and $a_1 = a_2 = a$. Small deviations from these values will not affect the results. Later in the paper we will weaken this restriction. v_0 and v_1 are the same as u, the only difference being that, is on the effects induced by the biocide onto them. To scale the variables first, scale the units of concentration of S, u, v_1 and v_0 by the input concentration $S^{(0)}$. This scaling replaces the parameter a by $a/S^{(0)}$. Then we scale time by the dilution rate (with units 1/time), i.e. $\bar{t} = Dt$. This reduces the input/washout rate to unity and replaces m by m/D, μ by μ/D and α by α/D. Finally we scale p by $p^{(0)}$, which has the effect of scaling $p^{(0)}$ to unity and replaces h by $S^{(0)}h/Dp^{(0)}$ and k by $k/p^{(0)}$. In system (2), $\phi(p) = e^{-\lambda p}$, so this now is written as $\phi(p) = e^{-\lambda p^{(0)}(p/p^{(0)})}$. The new variable is $p/p^{(0)}$ and the new parameter is $\lambda p^{(0)}$ (see [2–5]). If one makes these changes in system (2), we obtain

$$S' = 1 - S - \frac{Sm}{a+S}(u\phi(p) + v_1 + v_0) \tag{7}$$

$$u' = \frac{Sm}{a+S}(u\phi(p) + v_1) - (\mu + 1 + \alpha(1 - \phi(p)))\, u \tag{8}$$

$$v_0' = v_0\left(\frac{Sm}{a+S} - 1\right) + \mu u \tag{9}$$

$$v_1' = \alpha(1 - \phi(p))u - v_1 \tag{10}$$

$$p' = 1 - p - \frac{hp}{k+p}(v_0 + v_1) \tag{11}$$

where $\phi(p) = e^{-\lambda p}$ in the new variables.

Adding the first four equations yields

$$S' + u' + v_0' + v_1' \leq 1 - S - u - v_0 - v_1$$

or, using a comparison theorem,

$$S(t) + u(t) + v_0(t) + v_1(t) \leq 1 + ce^{-t} \tag{12}$$

Thus, all four concentrations are bounded since each element of the sum is positive. Moreover, the coordinates of any omega limit point must satisfy $S + u + v_0 + v_1 \leq 1$. Since $p(t)$ satisfies

$$p'(t) \leq 1 - p(t)$$

then

$$\limsup_{t\to\infty} \quad p(t) \leq 1$$

As a consequence, the right hand side of (7) is bounded, so when one can show that the limit as t tends to infinity of a variable exists, then the limit of the time derivative is zero.

We turn now to the equilibrium or rest points of the system (7). As in [2–5], we define λ_0 as the usual chemostat parameter reflecting the break-even concentration for v_0 given by the solution of the equation

$$\frac{\lambda_0 m}{a+\lambda_0} = 1$$

Then the feasible equilibrium points are $E_0 = (1, 0, 0, 0, 1)$ which always exists and the equilibria $E_1 = (\lambda_0, 0, 1-\lambda_0, 0, p^*)$ where p^* is the positive root of $(1-z)(k+z) = hz(1-\lambda_0)$. E_1 exists when $\lambda_0 < 1$. By calculating the variational matrix of the system (7) and replacing E_0 and E_1 into it, we obtain

Lemma 1 - The equilibrium E_0 is locally stable if λ_0>1 and

$$\frac{m}{(m-1)\lambda_0 + 1} < \frac{\mu + 1 + \alpha_1}{e^{-\lambda} + \alpha_1}$$

- The equilibrium E_1 is locally stable whenever $\lambda_0 < 1$ (i.e. whenever it exists).

The proof is a straight-forward computation, which we present in Appendix A.

We note that one of the conditions to do E_0 locally asymptotically stable does not hold when the non-trivial equilibrium point E_1 exists; thus if we have the existence of both equilibria E_0 and E_1, the point E_0 is unstable; otherwise if E_1 does not exist, then E_0 is globally stable.

Theorem 1 There is no non-trivial equilibrium point with all coordinates different to zero for the system (7).

Proof. When we assume $u \neq 0, v_0 \neq 0$ and $v_1 \neq 0$ we are studying a full problem of competition among bacteria in a chemostat. Considering $\phi(p) \neq 1$, from the zerocline of (10) we get $u = \frac{v_1}{\alpha(1-\phi(p))}$, from (9) we obtain $\frac{Sm}{a+S} = \frac{v_0-\mu u}{v_0}$. Replacing these two expressions into (8) and after some algebra we obtain the expression for $\phi(p)$:

$$\mu v_1 \phi(p) + \mu\alpha(1-\phi(p))(v_0+v_1) + v_0\alpha(1-\phi(p))^2 = 0 \qquad (13)$$

Since $1-\phi(p) > 0$ the expression given in (13) does not make sense if v_0 and v_1 are both positive. If $v_1 = 0$ we need $v_0 = 0$ (and viceversa) to satisfy (13). Therefore, for $\phi(p) \neq 1$ we do not have a non-trivial equilibrium point with all bacteria types present (see figure (1)).

When $\phi(p) = 1$ there is not phenotypic switching and the deal of the model, the study of two resistance mechanisms, is lost (in fact, we can see from the v_1 equation that $v_1 \to 0$ when $t \to \infty$).

2.2 Global Stability Analysis of the Equilibrium Points

If $E_1 = (\lambda_0, 0, 1-\lambda_0, 0, p^*)$ does not exist, all the solutions of the system (7) tend to E_0, that is, E_0 is globally asymptotically stable and we can prove this with the Lyapunov function $V(S, u, v_0, v_1) = 1 - (S + u + v_0 + v_1)$ (we can see that $\dot{V} = -1 + (S + u + v_0 + v_1) \leq 0$ due to the relation (12)). Biologically this means that the bacteria is lost from the system.

When $E_1 = (\lambda_0, 0, 1-\lambda_0, 0, p^*)$ exists, we get

Theorem 2 If $0 < \lambda_0 < 1$ where $\frac{m\lambda_0}{a+\lambda_0} = 1$ then the equilibrium point $E_1 = (\lambda_0, 0, 1-\lambda_0, 0, p^*)$ is globally asymptotically stable.

Proof. Consider the Lyapunov function

$$V(S, u, v_0, v_1, p) = \int_{\lambda_0}^{S} \frac{\eta - \lambda_0}{\eta} d\eta + c_1(u + v_1) + c_2 \int_{1-\lambda_0}^{v_0} \frac{\eta - (1-\lambda_0)}{\eta} d\eta$$

where $c_1, c_2 > 0$ are to be determined. We can see that $V(E_1) = 0$ and $V(S, u, v_0, v_1, p) \in C^1(\Re_+^5, \Re)$. Moreover, $V(S, u, v_0, v_1, p) > 0$ for all $(S, u, v_0, v_1, p) \in \Delta \setminus E_1$ where

$$\begin{aligned} \Delta &= \Omega - \{(S, 0, 0, 0, p)\} \\ &= \{(S, u, v_0, v_1, p) | S \geq 0, u \geq 0, v_0 \geq 0, v_1 \geq 0, p \geq 0\} - \{(S, 0, 0, 0, p)\} \end{aligned}$$

in fact, for the first integral:

- *For $S > \lambda_0$ we have $\frac{\eta - \lambda_0}{\eta} > 0$, then the integral is positive.*
- *For $S < \lambda_0$ we have $\frac{\eta - \lambda_0}{\eta} < 0$, then the integral is positive.*

We can apply the same procedure to the second integral and since $c_1, c_2 > 0$ we have $V(S, u, v_0, v_1, p) > 0$. The derivative of V is given by.

$$\begin{aligned} \dot{V} &= \frac{S - \lambda_0}{S} \dot{S} + c_1(\dot{u} + \dot{v_1}) + c_2 \left(\frac{v_0 - (1-\lambda_0)}{v_0} \right) \dot{v_0} \\ &= A(S) + B(S, p)u + C(S)v_1 + D(S)v_0 - c_2 \frac{\mu(1-\lambda_0)u}{v_0} \qquad (14) \\ &< A(S) + B(S, p)u + C(S)v_1 + D(S)v_0 \end{aligned}$$

where

$$A(S) = \frac{(S-\lambda_0)(1-S)}{S} - c_2(1-\lambda_0)\left(\frac{Sm}{a+S} - 1\right)$$

$$B(S,p) = \frac{-(S-\lambda_0)m\phi(p)}{a+S} + c_1\left(\frac{Sm}{a+S}\phi(p) - (\mu+1)\right) + c_2\mu$$

$$C(S) = \frac{-(S-\lambda_0)m}{a+S} + c_1\left(\frac{Sm}{a+S} - 1\right)$$

$$D(S) = \frac{-(S-\lambda_0)m}{a+S} + c_2\left(\frac{Sm}{a+S} - 1\right)$$

From $A(S)$ we define

$$T(S) = \frac{(S-\lambda_0)(1-S)}{S(1-\lambda_0)\left(\frac{Sm}{a+S} - 1\right)}$$

Then $T(S) > 0$ for all $S \in (0,1) \setminus \lambda_0$, $T(1) = 0$, and we have

$$\lim_{S\to 0^+} T(S) = +\infty$$

$$\lim_{S\to\lambda_0} T(S) = \lim_{S\to\lambda_0} \frac{\frac{\lambda_0 - S^2}{S^2}}{(1-\lambda_0)\left(\frac{am}{(a+S)^2}\right)} = \frac{m\lambda_0}{a}$$

So, defining $c_2 = \frac{m\lambda_0}{a}$ we obtain $A(S) < 0$. Moreover, when we replace this value in $D(S)$ get $D(S) = 0$.

Since the terms $B(S,p)$ y $C(S)$ depend on c_1 y c_2 we need to consider the cases $c_1 = c_2$ and $c_1 \neq c_2$. For the case $c_1 = c_2$, we obtain that $C(S)$ is equal to $D(S)$, so $C(S) = 0$; also,

$$B(S,p) = \frac{-(S-\lambda_0)m\phi(p)}{a+S} + \frac{m\lambda_0}{a}\left(\frac{Sm}{a+S}\phi(p) - 1\right) = \frac{m\lambda_0}{a}(\phi(p) - 1) \leq 0$$

Therefore we get $\dot{V} \leq 0$ for all $(S, u, v_0, v_1, p) \in \Delta$. In the case $c_1 \neq c_2$, from $C(S)$ we define

$$T_1(S) = \frac{(S-\lambda_0)\frac{m}{a+S}}{\frac{Sm}{a+S} - 1}$$

We apply to $T_1(S)$ the same procedure from $T(S)$ and we get $c_1 = \frac{m\lambda_0}{a}$, that is, $c_1 = c_2$, then $\dot{V} \leq 0$.

We now seek the maximum invariant region in the set $\{(S, u, v_0, v_1, p) \mid \dot{V} = 0\}$. Since $A(S), B(S,p) \leq 0$, $D(S) = C(S) = 0$ and $\dot{V}$ is given by (14); it must be the case when $A(S) = 0$, that is,

$$(S-\lambda_0)^2\left(S+\frac{a}{\lambda_0}\right)=0$$

Then $S=\lambda_0$. *Moreover,* $B(\lambda_0,p)<0$ *so,* $u=0$. *The values* $S=\lambda_0$, $u=0$ *forces* $v_1=0$ *and in consequence* $v_0=1-\lambda_2$ *and* $p=p^*$, *where* p^* *is the positive root of* $(1-z)(g+z)=hz(1-\lambda_0)$. *The only invariant set in this region is the rest point* E_1. *Then, by LaSalle invariance principle [7] we get,* E_1 *is globally asymptotically stable.*

3 Different Consumption Functions

In this section we study model (2) considering the difference that the inhibitor has on each bacteria. For this we assume that the consumption functions in (2) are different, that is $m_1\neq m_2$ and $a_1\neq a_2$.

We scale all variables of model (2) as before (section 2.1) except m_1, m_2, a_1 and a_2 since in this case these are replaced by m_1/D, m_2/D, $a_1/S^{(0)}$ and $a_2/S^{(0)}$ respectively. Then, making the changes we obtain

$$S'=1-S-\frac{Sm_1}{a_1+S}(u\phi(p)+v_1)-\frac{Sm_2}{a_2+S}v_0 \tag{15}$$

$$u'=\frac{Sm_1}{a_1+S}(u\phi(p)+v_1)-(\mu+1+\alpha(1-\phi(p)))u \tag{16}$$

$$v_0'=v_0\left(\frac{Sm_2}{a_2+S}-1\right)+\mu u \tag{17}$$

$$v_1'=\alpha(1-\phi(p))u-v_1 \tag{18}$$

$$p'=1-p-\frac{hp}{k+p}(v_0+v_1) \tag{19}$$

where $\phi(p)=e^{-\lambda p}$ in the new variables. As before, if we add the first four equations yields

$$S'+u'+v_0'+v_1'\leq 1-S-u-v_0-v_1$$

or, using a comparison theorem,

$$S(t)+u(t)+v_0(t)+v_1(t)\leq 1+ce^{-t}$$

Thus, all four concentrations are bounded since each element of the sum is positive. Moreover, the coordinates of any omega limit point must satisfy $S+u+v_0+v_1\leq 1$.
Since $p(t)$ satisfies

$$p'(t)\leq 1-p(t)$$

then

$$\limsup_{t\to\infty} \quad p(t) \leq 1$$

As in section 2.1, when the wild bacteria u is not present or when either of the resistant bacteria v_0 or v_1 are not present, the system solutions tend to the only one equilibria state $E_0 = (1, 0, 0, 0, 1)$. A non trivial equilibrium, if it exist, can be either only the mutated bacteria present or alternatively, all kinds of bacteria present. When only the mutated bacteria is present in the chemostat, we have that $u = 0$ and $v_1 = 0$ and we obtain the equilibrium point $E_1 = (\lambda_2, 0, 1 - \lambda_2, 0, p_*)$, where λ_2 is defined as the solution of the equation

$$\frac{\lambda_2 m_2}{a_2 + \lambda_2} = 1$$

and p_* is the positive root of $(1 - z)(k + z) = hz(1 - \lambda_2)$. E_1 exists when $0 < \lambda_2 < 1$. By calculating the variational matrix of the system (15) and replacing E_0 and E_1 into it, we obtain

Lemma 2 - The equilibrium E_0 is locally stable if $\lambda_2 > 1$ and $\frac{m_1}{a_1+1} < \frac{\mu+1+\alpha_1}{e^{-\lambda}+\alpha_1}$ where $\alpha_1 = \alpha(1 - e^{-\lambda})$

- The equilibrium E_1 is locally stable if $\frac{\lambda_2 m_1}{a_1+\lambda_2} < \frac{\mu+1+\alpha(1-\phi(p_*))}{\phi(p_*)+\alpha(1-\phi(p_*))}$

The proof is a straight-forward computation, Appendix A. We will show that the first condition in the stability of E_0 is not true due the existence of the non trivial equilibrium point E_1, so if we have the existence of both equilibrium points E_0 and E_1; the point E_0 is unstable, on the other hand, if E_1 does not exist biologically, the equilibrium point E_0 can be locally stable if the second condition above is true.

Now, for the existence of a non trivial equilibrium point we have the next result.

Theorem 3 A non trivial equilibrium point E_c where $u \neq 0, v_0 \neq 0$ and $v_1 \neq 0$ exists for the system (15) if $m_2 < \mu + 1 < \frac{m_1}{a_1+1}$, E_1 does not exist (that is, $\lambda_2 > 1$) and E_0 is unstable. That is, just if

$$\frac{m_1}{a_1 + 1} > \max\left\{\mu + 1, \frac{\mu + 1 + \alpha_1}{e^{-\lambda} + \alpha_1}\right\} \tag{20}$$

$$m_2 < \min\left\{a_2 + 1, \mu + 1\right\}$$

On the other hand, this equilibrium point has biological sense if and only if $G(\phi) < \frac{m_1}{a_1+1}$ where $G(\phi) = \frac{1+\mu+\alpha(1-\phi)}{\phi+\alpha(1-\phi)}$ and $\phi = e^{-\lambda p}$.

Proof. From the zerocline of (19) we have an expression for v_1+v_0 in function of the variable p, we denote this by G_1 and it is given by

$$G_1 = \frac{(1 - p)(k + p)}{hp}$$

From the zerocline of (18), and due to $1-\phi(p)\neq 0$*, we get a expression for* u*. Replacing this expression in the rest of the zeroclines and doing some algebra we obtain another expression for* v_1+v_0 *in function of the variable* p *(remember that* $\phi(p)=e^{-\lambda p}$*), namely* G_2

$$G_2=\frac{(1-S)\left[\mu+\alpha\left(1-\phi(p)\right)\left(1-f_2(S)\right)\right]}{G\left(\phi(p)\right)\left[1-f_2(S)\right]\left[\phi(p)+\alpha\left(1-\phi(p)\right)\right]+f_2(S)\mu}$$

where

$$S=\frac{a_1[\mu+1+\alpha\left(1-\phi(p)\right)]}{m_1[\phi(p)+\alpha\left(1-\phi(p)\right)]-[\mu+1+\alpha\left(1-\phi(p)\right)]}$$
$$G\left(\phi(p)\right)=\frac{1+\mu+\alpha\left(1-\phi(p)\right)}{\phi+\alpha\left(1-\phi(p)\right)}$$
$$f_2(S)=\frac{m_2S}{a_2+S}$$

A condition to guarantee a non trivial equilibrium point is the existence of a positive intersection between these two expressions, which depend only on the variable p*, in the interval* $(0,1)$ *(the extremes of this are not solutions for the* p*-equation when* v_0 *and* v_1 *are different from zero). Following this idea, we need that both of this expressions have sense (are positive) for* $p\in(0,1)$*.*
The expression given by G_1 *has biological sense due to* $1>p>0$*; on the other hand,* G_2*-expression has biological sense if and only if*

$$G\left(\phi(p)\right)<min\Big\{m_1,\frac{m_1}{a_1+1},\frac{a_2m_1}{a_2+a_1(m_2-1)}\Big\}=\frac{m_1}{a_1+1} \tag{21}$$

where the last equality is because $\lambda_2>1$

Now, defining $F(p)=G_1-G_2$ *we can see that, if* $m_2<\mu+1<\frac{m_1}{a_1+1}$*,* E_1 *does not exist (that is,* $\lambda_2>1$*) and* E_0 *is unstable,* F *has a root in* $(0,1)$*.*

So if the non-trivial equilibrium point exists, it can not coexist with E_1, the case when only the mutated bacteria is present. The dynamics of the general system (15), is observed in the figure (2).

3.1 Global Stability

When all the non trivial equilibrium points do not exist, all the solutions tend to E_0, that is, the equilibrium point E_0 is globally stable (this assertion can be proved taking the Lyapunov function $V(S,u,v_1,v_0,p)=1-(S+u+v_1+v_0)$ in Ω) and this means that all bacteria die (see section 2.2).

Lemma 3 If $\frac{m_2}{a_2+1}>1$ and $\frac{m_1}{a_1+1}>\frac{\mu+\alpha_1+1}{\phi(1)+\alpha_1}$ where $\alpha_1=\alpha(1-\phi(1))$, then the dimension of the stable manifold of E_0 for the system (15) is 3 with the S-axis and the p-axis as two of its eigenvectors.

Proof. The plane $u = v_0 = v_1 = 0$ is invariant and it is in fact an orbit of our equations (15). On this orbit $p', S' > 0$ if $S, p < 1$ and $p', S' < 0$ otherwise. Now, the other three eigenvalues come from the submatrix (see proof of lemma 1 in Appendix A)

$$A = \begin{pmatrix} \frac{m_1}{a_1+1}\phi(1) - (\mu + 1 + \alpha(1-\phi(1))) & 0 & \frac{m_1}{a_1+1} \\ \mu & \frac{m_2}{a_2+1} - 1 & 0 \\ \alpha(1-\phi(1)) & 0 & -1 \end{pmatrix}$$

The characteristic polynomial associated with the submatrix A is given by:

$$\left(\frac{m_2}{a_2+1} - 1 - x\right) \cdot \left[x^2 + \left(\mu + 2 + \alpha_1 - \frac{m_1\phi(1)}{a_1+1}\right)x + \left(\mu + 1 + \alpha_1 - \frac{m_1(\phi(1)+\alpha_1)}{a_1+1}\right)\right] = 0 \quad (22)$$

where $\alpha_1 = \alpha(1-\phi(1))$. Note that the polynomial between brackets in (22) has a positive discriminant, so, its roots are real. By the hypothesis: $\frac{m_1}{a_1+1} > \frac{\mu+\alpha_1+1}{\phi(1)+\alpha_1}$, then we have that the independent term in the quadratic polynomial is negative and therefore this polynomial has two real roots one being negative and the other positive. And, we have too that the first term in (22) gives us the positive root $\frac{m_2}{a_2+1} - 1$. The conclusion follows.

If E_0 and E_1 exist we have that, by theorem 3, is not possible that the non-trivial equilibrium point E_c exists.

Theorem 4 If $0 < \lambda_2 < 1$, and either of the next two conditions hold

(i) $\hat{\lambda}_1 > 1$
(ii) $\lambda_2 < \hat{\lambda}_1 < 1$

where $\frac{m_2\lambda_2}{a_2+\lambda_2} = 1$ and $\frac{m_1\hat{\lambda}_1}{a_1+\hat{\lambda}_1} = 1$ and, in addition, either of the next two conditions hold

$$m_1 < m_2 \quad , \quad a_2 \leq a_1 \quad (23)$$

$$\frac{m_2}{a_2} < \frac{m_1}{a_1} \quad (24)$$

then the equilibrium point $E_1 = (\lambda_2, 0, 1-\lambda_2, 0, p_*)$ is globally asymptotically stable for the system (15).

Proof. Consider the Lyapunov function

$$V(S, u, v_0, v_1, p) = \int_{\lambda_2}^{S} \frac{\eta - \lambda_2}{\eta} d\eta + c_1(u + v_1) + c_2 \int_{1-\lambda_2}^{v_0} \frac{\eta - (1-\lambda_2)}{\eta} d\eta$$

where $c_1, c_2 > 0$ are to be determined. We can see that $V(E_1) = 0$ and $V(S, u, v_0, v_1, p) \in C^1(\Re^5_+, \Re)$. Moreover, $V(S, u, v_0, v_1, p) > 0$ for all $(S, u, v_0, v_1, p) \in \Delta \setminus E_1$ where

$$\begin{aligned}\Delta &= \Omega - \{(S, 0, 0, 0, p)\} \\ &= \{(S, u, v_0, v_1, p) | S \geq 0, u \geq 0, v_0 \geq 0, v_1 \geq 0, p \geq 0\} - \{(S, 0, 0, 0, p)\}\end{aligned}$$

in fact, for the first integral:

- *For $S > \lambda_2$ we have $\frac{\eta - \lambda_2}{\eta} > 0$, then the integral is positive.*
- *For $S < \lambda_2$ we have $\frac{\eta - \lambda_2}{\eta} < 0$, then the integral is positive.*

We can apply the same procedure to the second integral and since $c_1, c_2 > 0$ we have $V(S, u, v_0, v_1, p) > 0$. The derivative of V is given by.

$$\begin{aligned}\dot{V} &= \frac{S - \lambda_2}{S}\dot{S} + c_1(\dot{u} + \dot{v_1}) + c_2\left(\frac{v_0 - (1 - \lambda_2)}{v_0}\right)\dot{v_0} \\ &= A(S) + B(S, p)u + C(S)v_1 + D(S)v_0 - c_2\frac{\mu(1 - \lambda_2)u}{v_0} \qquad (25) \\ &< A(S) + B(S, p)u + C(S)v_1 + D(S)v_0\end{aligned}$$

where

$$A(S) = \frac{(S - \lambda_2)(1 - S)}{S} - c_2(1 - \lambda_2)\left(\frac{Sm_2}{a_2 + S} - 1\right)$$

$$B(S, p) = \frac{m_1\phi(p)}{a_1 + S}[S(c_1 - 1) + \lambda_2] - c_1 - \mu(c_1 - c_2)$$

$$C(S) = \frac{-(S - \lambda_2)m_1}{a_1 + S} + c_1\left(\frac{Sm_1}{a_1 + S} - 1\right)$$

$$D(S) = \frac{-(S - \lambda_2)m_2}{a_2 + S} + c_2\left(\frac{Sm_2}{a_2 + S} - 1\right)$$

From $A(S)$ we define

$$T(S) = \frac{(S - \lambda_2)(1 - S)}{S(1 - \lambda_2)\left(\frac{Sm_2}{a_2 + S} - 1\right)}$$

Then $T(S) > 0$ for all $S \in (0, 1) \setminus \lambda_2$, $T(1) = 0$, and we have

$$\lim_{S \to 0^+} T(S) = +\infty \qquad \lim_{S \to \lambda_2} T(S) = \lim_{S \to \lambda_2} \frac{\frac{\lambda_2 - S^2}{S^2}}{(1 - \lambda_2)\left(\frac{a_2 m_2}{(a_2 + S)^2}\right)} = \frac{m_2\lambda_2}{a_2}$$

So, defining $c_2 = \frac{m_2\lambda_2}{a_2}$ we obtain $A(S) < 0$. Moreover, when we replace this value in $D(S)$ get $D(S) = 0$.

Since the terms $B(S,p)$ y $C(S)$ depend on c_1 y c_2 we need to consider the cases $c_1 = c_2$ and $c_1 \neq c_2$. For the case $c_1 = c_2$, we obtain

$$C(S) = \frac{\lambda_2}{a_2(a_1+S)}[(m_1-m_2)S + (a_2m_1 - a_1m_2)]$$

We must to observe that the term between brackets is a line in the variable S, we denote it by $L(S)$. By hypothesis (23) we have that the slope of $L(S)$ (m_1-m_2) is negative and $L(0) = a_2m_1 - a_1m_2 < 0$, that is, the complete line $L(S)$ is negative for all $S \in (0,1)$. Now, for $B(S,p)$ we get

$$B(S,p) = \frac{m_1\phi(p)}{a_1+S}[S(\frac{m_2\lambda_2}{a_2} - 1) + \lambda_2] - \frac{m_2\lambda_2}{a_2}$$

in the last expression the line between brackets, denoted by $l(S)$ is positive due to $l(0) > 0$ and its slope is positive, so we have

$$B(S,p) = \frac{m_1\phi(p)}{a_1+S}l(S) - \frac{m_2\lambda_2}{a_2} \leq \frac{m_1}{a_1+S}l(S) - \frac{m_2\lambda_2}{a_2} = C(S)$$

and since $C(S) < 0$ we obtain that $B(S,p) < 0$ and in consequence we have $\dot{V} \leq 0$ for all $(S,u,v_0,v_1,p) \in \Delta$.

For the case $c_1 \neq c_2$, from $C(S)$ we define

$$T_1(S) = \frac{(S-\lambda_2)\frac{m_1}{a_1+S}}{\frac{Sm_1}{a_1+S} - 1}$$

Then

$$T_1'(S) = \frac{-m_1(m_1-1)\left(\hat{\lambda}_1 - \lambda_2\right)}{\left(S(m_1-1) - a_1\right)^2}$$

that is, independently of the condition (i) or (ii), $T_1(S)$ is a decreasing function of S; also satisfy $T_1(0) > 0$, $T_1(\lambda_2) = 0$ and

$$\lim_{S\to\hat{\lambda}_1^+} T_1(S) = +\infty \qquad \lim_{S\to\hat{\lambda}_1^-} T_1(S) = -\infty$$

Then, for case (i), we choose $c_1 = T_1(0) > 0$. Since $T_1(S)$ is decreasing $T_1(0) \geq T_1(S)$ for all $S \in (0,1)$ and, due to the denominator of $T_1(S)$ is negative in $(0,1)$ ($\hat{\lambda}_1 > 1$ in this case), we get $C(S) < 0$ for all $S \in (0,1)$.

In the case (ii), due to the properties for $T_1(S)$, exists $c_1 > 0$ such that

$$\max_{(0,\hat{\lambda}_1)} T_1(S) \leq c_1 \leq \min_{(\hat{\lambda}_1,1)} T_1(S)$$

If c_1 does not exist, then there exists $\beta > 0$ such that the equation $T_1(S) = \beta$ has at least two distinct roots η_1, η_2 satistying $0 < \eta_1 < \lambda_2 < \hat{\lambda}_1 < \eta_2 < 1$

(roots are not in $[\lambda_2, \hat{\lambda}_1)$ *since* $T_1(S) < 0$*); however this is a contradiction.*

Analyzing the critical values for $T_1(S)$ *we obtain*

$$T_1(0) = \frac{\lambda_2 m_1}{a_1} \leq c_1 \leq (1-\lambda_2)\frac{m_1\hat{\lambda}_1}{a_1(1-\hat{\lambda}_1)} = T_1(1) \tag{26}$$

so, with c_1 *satisfying (26) we have* $C(S) < 0$ *for case* (ii)*.*

Due to hypothesis (24), the relation (26) and taking $c_2 = \frac{m_2\lambda_2}{a_2}$ *we have* $c_2 < c_1$ *and therefore*

$$B(S,\phi) \leq \frac{m_1\phi(p)}{a_1+S}R(S) - c_1 \leq \frac{m_1}{a_1+S}R(S) - c_1 = C(S)$$

Where $R(S) = S(c_1 - 1) + \lambda_2$*, is positive due to* $R(0) > 0$ *and its slope* $c_1 - 1$ *is positive (by hypothesis (24)). So, due to* $C(S) \leq 0$ *we get* $B(S,\phi) \leq 0$ *for all* $S \in (0,1)$*.*

We have shown $\dot{V} \leq 0$ *for all* $(S,u,v_0,v_1,p) \in \Delta$*. We now seek the maximum invariant region in the set* $\{(S,u,v_0,v_1,p) \mid \dot{V} = 0\}$*. Since* $A(S), B(S,p)$*,* $C(S) \leq 0$*,* $D(S) = 0$ *and* $\dot{V}$ *is given by (25); it must be the case when* $A(S) = 0$*, that is,*

$$(S-\lambda_2)^2\left(S + \frac{a_2}{\lambda_2}\right) = 0$$

Then $S = \lambda_2$*. Moreover,* $B(\lambda_2, p) < 0$ *and* $C(\lambda_2) < 0$ *so,* $u = v_1 = 0$*. The values* $S = \lambda_2$*,* $u = v_1 = 0$ *forces* $v_0 = 1 - \lambda_2$ *and in consequence* $p = p_*$*, where* p_* *is the positive root of* $(1-z)(g+z) = hz(1-\lambda_2)$*. The only invariant set in this region is the rest point* E_1*. Then, by LaSalle invariance principle [7] we get,* E_1 *is globally asymptotically stable.*

The graphical behavior for the global stability of the equilibrium point $E_1 = (\lambda_2, 0, 1-\lambda_2, 0, p_*)$, is shown in figure (3).

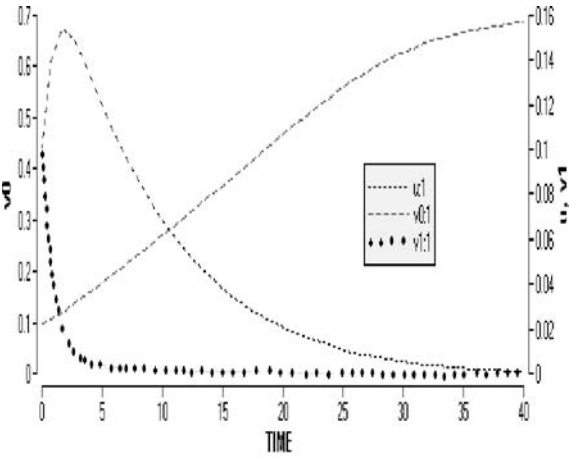
4 Numerical Results

The numerical simulations of the systems (7) and (15) are shown considering two graphics, phase plane and trajectory in time. The initial conditions we consider are $C1(0.8\,, 0.02\,, 0.01\,, 0.01\,, 0.5)$, $C2(0.2\,, 0.01\,, 0.8\,, 0.01\,, 0.5)$, $C3(0.8\,, 0.02\,, 0.01\,, 0.3\,, 0.5)$ and $C4(0.2\,, 0.01\,, 0.01\,, 1\,, 0.5)$.

Figure (1) shows the global behavior of the solutions of (7) illustrating the global stability of E_1 (Theorem 2). Figure (2) shows the local behavior of the solutions of (15), illustrating the case of the coexistence of all the possible

equilibrium points when the function of consumption are different for each bacteria (Theorem 3).

Figure (3) shows the global stability of the equilibrium point given by $E_1 = (\lambda_2, 0, 1-\lambda_2, 0, p_*)$ of (15). Since the global stability of the equilibrium point E_1 requires to be unique in the whole region of interest, the graphics ((3)) only shows the points E_1 (circle) and the trivial equilibrium point $E_0 = (1, 0, 0, 0, 1)$ (asterisk).



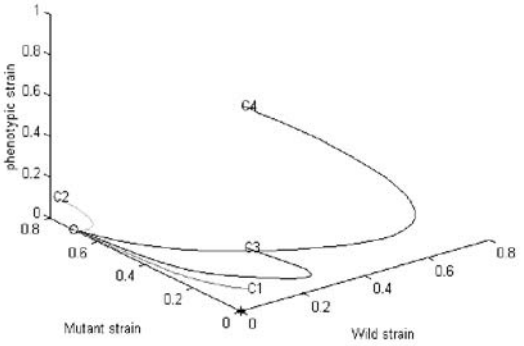


Fig. 1. The parameters are $m = 1.1$, $a = 0.03$, $\alpha = 0.2$, $\mu = 0.07$, $\lambda = 0.1$, $h = 0.001$ and $k = 0.05$, that is, we are considering the same consumption function of nutrient for the three bacteria. The equilibrium points $E_0 = (1, 0, 0, 0, 1)$ (asterisk) and $E_1 = (0.3, 0, 0.7, 0, 0.99)$ (circle) coexist and we can see the global stability of E_1. This means that only the mutated bacteria, v_0, survive.

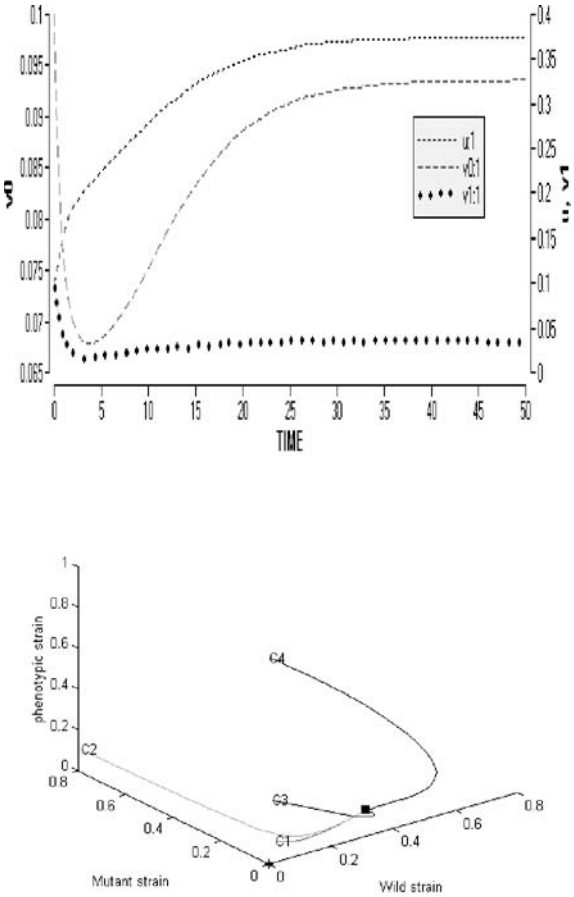


Fig. 2. Considering different consumption function of the nutrient and taking the parameters $a_1 = 0.1$, $a_2 = 0.2$, $m_1 = 1.4$, $m_2 = 1.01$, $\alpha = 1$, $\mu = 0.07$, $\lambda = 0.1$, $h = 0.001$ and $k = 0.05$. The equilibrium points $E_0 = (1, 0, 0, 0, 1)$ (asterisk) and E_c (square) coexist (we must remember that E_1 must not exists in order to guarantee the existence of E_c). The analytical expression for E_c is very hard to find. We can see that, due to the existence of E_c, all the three bacteria survive in the environment although the v_1 or phenotypic strain decreases from the original concentration.

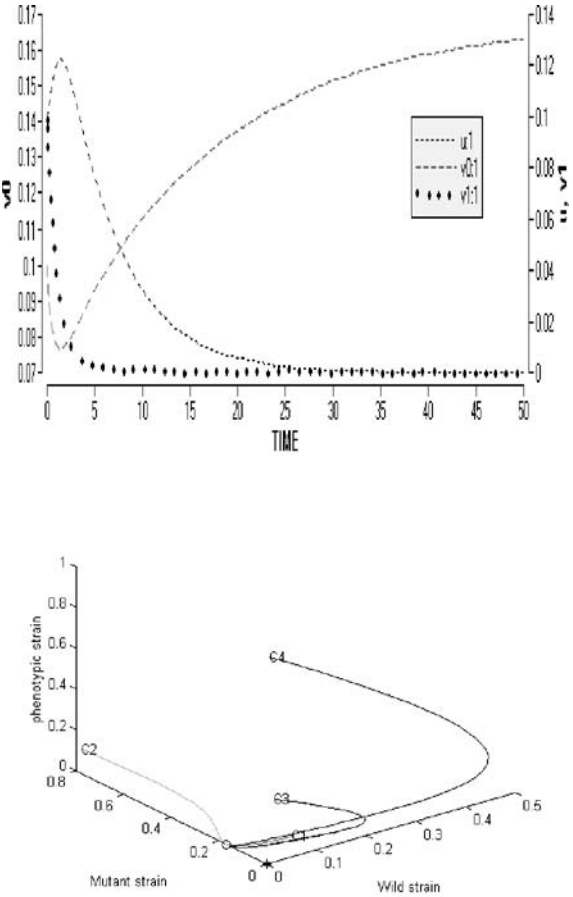


Fig. 3. The parameters are $m_1 = 1.1$, $m_2 = 1.6$, $a_1 = 0.1$, $a_2 = 0.5$, $\mu = 0.07$, $\lambda = 0.1$, $\alpha = 0.2$, $h = 0.001$ and $k = 0.05$. We have the global stability of the equilibrium point $E_1 = (0.83, 0, 0.166, 0, 0.99)$ (circle) that is, only the bacteria v_0 survive.

5 Discussion and Conclusion

In contrast to other models ([2–5]), we consider three strains in competition for a nutrient and in presence of one inhibitor. The resistance from mutation does not depend on the inhibitor concentration since this is a random phenomena. However, the resistance due to phenotypic switching depends on the inhibitor concentration. There is scientific evidence about the behavior of bacteria for each one of this resistance mechanisms ([1, 15]); for this reason, our goal was

to determine if there is survival of bacteria when both resistance mechanisms appear, as it happens when these appear separately. In our model, we have assumed that there is not phenotypic replication and that all new phenotypic bacteria is sensitive to the inhibitor. We have shown that

1. If we assume that the mutant and the phenotypic strain appear from the wild strain (so, they have the same consumption function, (7)) then the surviving bacteria is the mutant, as occurs in Miller's results [15].
2. If we assume that genetic differences in bacterial strain induce different consumption functions, then both resistant types may coexist together with the wild type: there is no competitive exclusion. This result depends on the values of the parameters of the model.
3. With an appropriate selection of parameters, we can get competitive exclusion and, as before, the mutant strain is the one that survives in the environment.

Appendix A: Proofs

Proof (Proof of Lemma 1). The variational matrices at E_0 and E_1 for the system (7) are, respectively

$$A_0 = \begin{pmatrix} -1 & -\frac{m}{a+1}\phi(1) & -\frac{m}{a+1} & -\frac{m}{a+1} & 0 \\ 0 & A_{0_{22}} & 0 & \frac{m}{a+1} & 0 \\ 0 & \mu & \frac{m}{a+1}-1 & 0 & 0 \\ 0 & \alpha(1-\phi(1)) & 0 & -1 & 0 \\ 0 & 0 & -\frac{h}{k+1} & -\frac{h}{k+1} & -1 \end{pmatrix}$$

$$A_1 = \begin{pmatrix} -A_{1_{11}} & -\phi(p^*) & -1 & -1 & 0 \\ 0 & A_{1_{22}} & 0 & 1 & 0 \\ A_{1_{31}} & \mu & 0 & 0 & 0 \\ 0 & \alpha(1-\phi(p^*)) & 0 & -1 & 0 \\ 0 & 0 & -\frac{hp^*}{k+p^*} & -\frac{hp^*}{k+p^*} & -A_{1_{55}} \end{pmatrix}$$

Where

$$\begin{aligned} A_{0_{22}} &= \frac{m}{a+1}\phi(1) - (\mu + 1 + \alpha\,(1-\phi(1))) \\ A_{1_{11}} &= 1 + \frac{am}{(a+\lambda_0)^2}(1-\lambda_0) \\ A_{1_{22}} &= \phi(p^*) - (\mu + 1 + \alpha\,(1-\phi(p^*))) \\ A_{1_{31}} &= \frac{am}{(a+\lambda_0)^2}(1-\lambda_0) \\ A_{1_{55}} &= 1 + \frac{kh}{(k+p)^2}(1-\lambda_0) \end{aligned}$$

the corresponding characteristic polynomials in x for A_0 is

$$\Big(1+x\Big)^2 \ \Big(\frac{m}{a+1}-1-x\Big)P(x)=0,$$

where

$$P(x)=x^2-\Big[\frac{me^{-\lambda}}{a+1}-(\mu+2+\alpha_1)\Big]x-\Big[\frac{m}{a+1}(e^{-\lambda}+\alpha_1)-(\mu+1+\alpha_1)\Big]$$

and $\alpha_1=\alpha(1-e^{-\lambda})$. Then, E_0 is locally asymptotically stable if and only if $\frac{m}{a+1}-1<0$, and, simultaneously, the roots of $P(x)$ have negative real part. These conditions can be summarized as

$$\lambda_0>1 \qquad \text{and} \qquad \frac{m}{(m-1)\lambda_0+1}<\frac{\mu+1+\alpha_1}{e^{-\lambda}+\alpha_1}$$

the characteristic polynomial associated to A_1 in x is

$$\left(1+\frac{kh(1-\lambda_0)}{(k+p^*)^2}+x\right)\ (x+1)\ \left(x+\frac{am}{(a+\lambda_0)^2}(1-\lambda_0)\right)Q(x)=0$$

where

$$Q(x)=x^2+\Big[2-\phi(p^*)+\mu+\alpha(1-\phi(p^*))\Big]x+\Big[1-\phi(p^*)+\mu\Big]$$

From the Routh-Hurwitz criteria we have that the roots of the polynomial $Q(x)$ have negative real part. Moreover, the roots are real. So, E_1 is a locally stable node.

Proof (Proof of Lemma 2). The Jacobian matrix of the system (15) is, for a general point $E=(S,u,v_0,v_1,p)$ in Ω, is given by:

$$A=\begin{pmatrix} -A_{11}-\frac{Sm_1}{a_1+S}\phi(p) & -\frac{Sm_2}{a_2+S} & -\frac{Sm_1}{a_1+S} & -\frac{Sm_1}{a_1+S}u\phi'(p) \\ A_{21} & A_{22} & 0 & \frac{Sm_1}{a_1+S} & \left(\frac{Sm_1}{a_1+S}+\alpha\right)u\phi'(p) \\ A_{31} & \mu & \frac{Sm_2}{a_2+S}-1 & 0 & 0 \\ 0 & \alpha(1-\phi(p)) & 0 & -1 & -\alpha u\phi'(p) \\ 0 & 0 & \frac{-hp}{k+p} & \frac{-hp}{k+p} & -A_{55} \end{pmatrix}$$

Where

$$A_{11} = 1+\frac{a_1 m_1}{(a_1+S)^2}\left(u\phi(p)+v_1\right)+\frac{a_2 m_2}{(a_2+S)^2}v_0$$
$$A_{21} = \frac{a_1 m_1}{(a_1+S)^2}\left(u\phi(p)+v_1\right)$$
$$A_{22} = \frac{Sm_1}{a_1+S}\phi(p)-(\mu+1+\alpha\left(1-\phi(p)\right))$$
$$A_{31} = \frac{a_2 m_2}{(a_2+S)^2}v_0$$
$$A_{55} = 1+\frac{kh}{(k+p)^2}(v_0+v_1)$$

Replacing E_0 we obtain the characteristic polynomial in x:

$$\Big(1+x\Big)^2\ \Big(\frac{m_2}{a_2+1}-1-x\Big)P(x)=0$$

where

$$P(x)=x^2-\Big[\frac{m_1 e^{-\lambda}}{a_1+1}-(\mu+2+\alpha_1)\Big]x-\Big[\frac{m_1}{a_1+1}(e^{-\lambda}+\alpha_1)-(\mu+1+\alpha_1)\Big]$$

and $\alpha_1=\alpha(1-e^{-\lambda})$. So, E_0 is locally asymptotically stable if and only if

$$\begin{aligned}&\frac{m_2}{a_2+1}<1\\ &\frac{m_1}{a_1+1}<min\Big\{\frac{\mu+2+\alpha_1}{e^{-\lambda}},\frac{\mu+1+\alpha_1}{e^{-\lambda}+\alpha_1}\Big\}=\frac{\mu+1+\alpha_1}{e^{-\lambda}+\alpha_1}\end{aligned}\tag{27}$$

On the other hand, replacing E_1 into the Jacobian matrix, we obtain the characteristic polynomial:

$$\left(1+\frac{kh(1-\lambda_2)}{(k+p_*)^2}+x\right)\ (x+1)\ \left(x+\frac{a_2 m_2}{(a_2+\lambda_2)^2}(1-\lambda_2)\right)Q(x)=0$$

where $Q(x)=x^2+Ax+C$

$$A = 2+\mu+\alpha\left(1-\phi(p_*)\right)-\frac{\lambda_2 m_1}{a_1+\lambda_2}\phi(p_*)$$
$$B = \mu+1+\alpha(1-\phi(p_*))-\frac{\lambda_2 m_1}{a_1+\lambda_2}\left(\alpha(1-\phi(p_*))+\phi(p_*)\right)$$

So, E_1 will be locally stable if the roots of $Q(x)$ have negative real part, that is

$$\frac{\lambda_2 m_1}{a_1+\lambda_2}<min\Big\{\frac{\mu+2+\alpha(1-\phi(p_*))}{\phi(p_*)},\frac{\mu+1+\alpha(1-\phi(p_*))}{\phi(p_*)+\alpha(1-\phi(p_*))}\Big\}$$
$$=\frac{\mu+1+\alpha(1-\phi(p_*))}{\phi(p_*)+\alpha(1-\phi(p_*))}$$

Acknowledgement. BTS acknowledges support of a Doctoral CONACyT grant. JXVH acknowledges partial financial support from an International Fellowship of the Santa Fe Institute.

References

1. Balaban, N.Q., Merrin, J, *et. al.*: Bacterial Persistence as a Phenotypic Switch. Science **305**, 1622–1625 (2004)
2. Braselton, J.P., Waltman, P.: A competition model with dynamically allocated inhibitor production. Mathematical Biosciences **173**, 55–84 (2001)
3. Hsu, S.B., Waltman, P.: A survey of mathematical models of competition with an inhibitor. Mathematical Biosciences **187**, 53–91 (2004)
4. Hsu, S.B., Li, Y.S., Waltman, P.: Competition in the presence of a lethal external inhibitor. Mathematical Biosciences **167**, 177–199 (2000)
5. Hsu, S.B., Waltman, P.: Analysis of a model of two competitors in a chemostat with an external inhibitor. Journal of Applied Mathematics (SIAM) **52**, 528–540 (1992)
6. Hsu, S.B., Waltman, P.: Competition in the chemostat when one competitor produces a toxin. Japaniese Journal of Industrial Applied Mathematics **15**, 471–490 (1998)
7. Khalil, H.: Nonlinear Systems. Prentice-Hall (1996)
8. Jones, D.A., Le, D., Kojouharov, H.V., Smith, H.L.: The Freter model: A simple model of biofilm formation. Journal of Mathematical Biology **47** 2, 137–152 (2003)
9. Leenheer, P., Li, B., Smith, H.L.: Competition in the chemostat: Some remarks. Canadian Applied Mathematics Quarterly **11** 3, 229–248 (2003)
10. Leenheer, P., Smith, H.L.: Feedback control for the chemostat. Journal of Mathematical Biology **46**, 48–70 (2003)
11. Lenski, R.E., Hattingh, S.: Coexistence of two competitors one resource and one inhibitor: a chemostat model based on bacteria and antibiotics. Journal of Theoretical Biology **122**, 83–93 (1986)
12. Levin, B.R.: Noninherited Resistance to Antibiotics. Science **305**, 1578–1579 (2004)
13. Li, B., Smith, H.L.: How Many Species Can Two Essential Resources Support?. Journal of Applied Mathematics (SIAM) **62**, 336–366 (2001)
14. Markus, L.: Asymptotically autonomous differential systems. In: Lefschetz, S. (ed.) Contributions to the theory of Nonlinear Oscillations III. Annals of Mathematics Studies **36** Princeton University Press, Princeton, NJ, 17–29 (1956)
15. Miller, C., *et. al.*: SOS Response Induction by β-Lactams and bacterial defense against antibiotic Lethality. Science **305**, 1629–1631 (2004)
16. Smith, H. L., Waltman, P.: The theory of the chemostat. Dynamics of microbial competition. Cambdrige University Press (1995)
17. Stemmons, E.D., Smith, H.L.: Competition in a chemostat with wall attachment. Journal of Applied Mathematics (SIAM) **61** 2, 567–595 (2000)
18. Thieme, H.R.: Convergence results and a Poincaé-Bendixon trichotomy for asymtotically autonomous differential equations. Journal of Mathematical Biology **30**, 755–763 (1992)

19. Thieme, H.R.: Mathematics in Population Biology. Princeton University Press, Princeton, NJ (2003)
20. El ciclo del azufre, *http://www.monografias.com/trabajos4/azufre/azufre.shtml*.
21. Mecanismos de resistencia, *www.virtual.unal.edu.co/cursos/odontologia/2005205*.

From Spatial Pattern in the Distribution and Abundance of Species to a Unified Theory of Ecology: The Role of Maximum Entropy Methods

John Harte

Energy and Resources Group, 310 Barrows Hall, University of California, Berkeley, CA 94720 USA jharte@berkeley.edu

Key words: Distribution and abundance of species, spatial pattern in ecology, Random Placement model, species-area relationships, maximum entropy framework, unified theory of Ecology.

1 Introduction

A central focus of ecological research is to understand the distribution and abundance of species (DAS) across a range of spatial and temporal scales. Such knowledge is critical to our ability to design sustainable ecological reserves, to predict extinction rates under habitat loss or climate change, to estimate species diversity from incomplete census data, and to decipher the fundamental forces shaping ecosystems [16, 25, 44]. Moreover, it is increasingly becoming evident that the spatial structure of populations influences the potential for coexistence of species and the likelihood that cooperative phenomena will emerge in the face of selfish or competitive behaviors [33, 36].

Success in this endeavor would ideally consist of a quantitative description and a theoretical understanding of empirical patterns, scaling rules, and the mechanisms that generate those patterns and rules. It is instructive to locate current research on DAS on a possibly analogous historical path: that of accumulating insight into celestial dynamics and gravity. The following only slightly oversimplifies that latter history:

Accumulation of data	→ Distillation of patterns	→ Deduction of mechanism
from observation	from data	from pattern
(Brahe)	(Kepler)	(Newton)

Current research on DAS is spread out across all three of these steps. Within the past couple of decades, a tremendous effort has been made to accumulate data on DAS, and thus to accomplish for ecology what Brahe did for planetary orbits. Particularly of note here is the pioneering work of Hubbell [27], who had the foresight to initiate tropical-forest census plots that now provide a wealth of spatially-explicit vegetation data that are shared with ecologists. Other spatially-explicit plant censuses are now also available (see, for example, [13, 17]), as are less detailed, but nevertheless valuable, animal census data from, for example, the North American annual Christmas bird counts and the United Kingdom's grid-based censuses.

Here I will review the types of DAS patterns that ecologists look for in the data and a selection of statistical models that have been proposed to describe these patterns; in addition, I present some ideas about how one of these models in particular, which appears to best describe patterns, might help guide us on the path toward award a future unified theory of ecology.

In Section II, I describe a suite of metrics used by ecologists to describe different facets of DAS across spatial scales. In Section III, I review some old and new statistical models developed to characterize the actual functional forms of these metrics, focusing on the scaling properties of spatial distributions. Then, in Section IV, I compare the central predictions of these models and in Section V, I compare predictions to observations. In Section VI, I explain why the maximum entropy inference framework, which was the foundation for one of the models introduced in Section III, is particularly well-suited for ecology. Finally, in Section VII, I propose an agenda for developing a theory of ecology that would include, but extend beyond, an explanation of patterns in DAS; it is based on the closely related foundations of the maximum likelihood inference method or MaxEnt as developed and expounded by Jaynes [30–32] and the principle of maximum entropy production (MEP) as derived by Dewar [11].

2 Scaling Metrics in Spatial Ecology

Throughout this chapter, I consider the distribution and abundance of species within some arbitrary area, A_0. This area might be a square kilometer plot located within Amazonia and within which trees or insects or birds are censused, or a 100 m^2 plot located within a meadow and within which forbs and grasses are censused, or a square meter patch of soil within which microorganisms are censused. By the term 'scaling metrics' I refer to measures that describe phenomena in or across cells of specified area within A_0; the area of those cells will be referred to here as the scale of analysis. One such metric is the probability distribution $P_i^{(n_0)}(n)$. Here the superscript is a species label indicating the total abundance of the species in A_0, the largest scale under consideration. The subscript i is a scale label, indicating the cell area $A_i = A_0/2^i$. The variable n is the abundance of the species in an arbitrary A_i cell, so P

is the distribution of abundances across the cells of scale i. Implicit in the way we label this distribution is the assumption made throughout this article that the shape of the distribution function depends only on i and n_0. This is a strong assumption because it implies that the influence of species traits, such as the rooting architecture of the individuals within a tree species or the food requirements of individuals within a bird species, exert their influence on spatial pattern only insofar as those traits influence n_0.

$P_i^{(n_0)}(n)$ is an example of a species-level metric, in that it describes the distribution of individuals within a species. In contrast, community-level metrics describe the spatial distribution of species within a community or describe the distribution of abundances of species across a community of species. Table 1 provides a partial list of species- and community-level metrics that are used in ecology to describe scaling properties of DAS. Metrics can also be distinguished depending on whether they describe the distribution within cells of area A_i (a one-point function) or whether they describe the joint distribution across more than 1 cell (a multi-point function) and thus contain information about correlations across space. $P_i^{(n_0)}(n)$ is an example of a one-point function. In the Table, $\chi^{(n_0)}(A_i, D)$ and $\chi(A_i, D)$ are examples of two-point functions at species-level and community-level, respectively. Finally, community-level metrics can be distinguished by whether or not they depend on the assumption that species are distributed independently of each other across space. In the Table, only the community-level distributions $\chi(A_i, D)$ and $F_i(N)$ are sensitive to that assumption [22].

$P_i^{(n_0)}(n)$ is a particularly fundamental metric because from knowledge of its functional form, the functional forms of many other useful metrics can be derived. As shown elsewhere [38], the aggregation index $\Omega_i^{(n_0)}$ can be expressed as

$$\Omega_i^{(n_0)} = \frac{2}{1 - P_i^{(n_0)}(0)} - \frac{1}{1 - P_{i-1}^{(n_0)}(0)} \ . \tag{1}$$

To relate the range-area relationship (RAR) and the species-area relationship (SAR) to $P_i^{(n_0)}(n)$, we note that $1 - P_i^{(n_0)}(0)$ is the probability that species n_0 is present in a cell of area A_i. Hence, from its definition in Table 1, the RAR can be expressed [21] as

$$R_i^{(n_0)} = \left(1 - P_i^{(n_0)}(0)\right) A_0 \ . \tag{2}$$

The SAR is the expected number of species at scale i, which is then given by:

$$S_i = \sum_{\text{species}} \left(1 - P_i^{(n_0)}(0)\right) \ . \tag{3}$$

Consider, next, the endemics-area relationship (EAR). A species is endemic to a cell if all of its individuals are located just in that cell, and thus

Table 1. Spatial Metrics

Species-Level Metrics		
$P_i^{(n_0)}(n)$	Species-level spatial abundance distribution	Probability that n individuals of a species with total abundance n_0 in A_0 are found in cell of area A_i
$R_i^{(n_0)}$	Range-area relationship (RAR)	Box-counting measure of occupancy for a species with n_0 individuals in A_0: relates range size of a species (= number of occupied cells × cell area) to cell area A_i
$\chi^{(n_0)}(A_i, D)$	Species-level commonality function	Probability a species with total abundance n_0 in A_0 is found in two A_i cells separated by a distance D
$\Omega_i^{(n_0)}$	Measure of aggregation	Density of conspecifics at distance $A_i^{1/2}$ from average individual, relative to random distribution for a species with n_0 individuals in A_0
Community-Level Spatial Metrics		
S_i	Species-area relationship (SAR)	Relates number of all species occurrences in a census cell to area A_i of that cell
E_i	Endemics-area relationship (EAR)	Relates number of species unique to a cell within A_0 to cell A_i; $E_0 = S_0$ by definition
$\chi(A_i, D)$	Community-level commonality function	Fraction of all species in common to two cells of area A_i separated by a distance D
$\Phi_i(n)$	Species-abundance distribution	Fraction of all possible species occurrences with n individuals in cells of area A_i
$F_i(N)$	Community-level spatial-abundance distribution	Probability N individuals from all species are in cell A_i

$$E_i = \sum_{\text{species}} P_i^{(n_0)}(n_0) \ . \tag{4}$$

We note that Eq. 4 yields $E_0 = \sum_{\text{species}} 1 = S_0$, as it must because all the species are endemic to A_0 by our definition of endemicity.

The scale-dependent species abundance distribution in the Table is given by

$$\Phi_i^{(n_0)} = \frac{1}{S_0} \sum_{\text{species}} P_i^{(n_0)}(n) \ . \tag{5}$$

And, under the assumption that species are distributed independently of each other, the community-level abundance distribution can be related to the $P_i^{(n_0)}(n)$ by

$$F_i(N) = \sum_{n(j)} \prod_{\text{species}} P_i^{(n_0)}(n(j))\, \delta(N, \sum_{n(j)} n(j)) \ . \tag{6}$$

Here $n(j)$ is an abundance variable for the j^{th} species and the equation simply states that the probability of a total of N individuals in an A_i cell is given by the sum over the joint probabilities for all the various combinations of individual species abundances that add up to the value N. $\prod_{\text{species}}$ denotes a product over all species and $\sum_{n(j)}$ denotes the sum over all combinations of abundances of species.

The two-point species-level metric, $\chi^{(n_0)}(A_i, D)$, contains information about the simultaneous presence of a species in two cells separated by a distance D. This two-point metric cannot, in general, be derived solely from knowledge of the $P_i^{(n_0)}(n)$, although as I discuss briefly in Section III, for some models of spatial pattern such a derivation is possible. The community level commonality metric, $\chi(A_i, D)$, can be derived from the $\chi^{(n_0)}(A_i, D)$ provided the species are assumed to be distributed in space independently from each other:

$$\chi(A_i, D) = \frac{\sum_{\text{species}} \chi^{(n_0)}(A_i, D)}{\sum_{\text{species}} \left(1 - P_i^{(n_0)}(0)\right)} \ . \tag{7}$$

Although the Table only gives 1- and 2-point functions, more complex metrics can be defined; at scale i, the most detailed spatial information is contained in the 2^i-point metric,which is the probability of any particular assignment of the n_0 individuals across all 2^i cells [5, 22].

With these definitions of spatial scaling metrics, and some relationships among them presented, I turn now to a discussion of 4 particular models that have been advanced to predict the actual functional form of these metrics in ecology.

3 Models of Spatial Pattern in Ecology

Spatially explicit models in ecology can be statistical, and thus based on assumed mathematical properties of distributions, or dynamical, and thus based on biologically-based assumptions about processes such as birth, death, migration and dispersal, competition, etc. In some cases, see MaxEnt below, the distinction between these two categories is difficult to maintain (for another example, see the "neutral theory of ecology", [28]). Many specific models

developed to explain one, or a subset, of the metrics in Table 1 have been proposed and I will not attempt to review them all. Instead, I review here four models that predict the forms of the metrics of spatial pattern in Table 1. These models span a very wide range of assumptions about spatial structure, and one of the models, HEAP, is actually a framework within which a continuum of spatial models can be embedded. Another, MaxEnt is more of a theoretical framework than it is a model, and as we will see in the final Section, its potential applicability to ecology may extend far beyond the prediction of spatial structure.

The one exception to the ability of these four models to predict the metrics in Table 1 is the species abundance distribution $\Phi_i(n)$ at scale $i = 0$. This distribution is just the set of abundances $\{n_0\}$ found in the largest area under consideration, A_0, and in all but one of the models it will be assumed as a given. I will show, however, that in the fourth model, MaxEnt, the possibility exists that this distribution can be derived for any particular taxonomic group such as plants from simply knowing for that group the total number of species and of individuals, and the total amount of living biomass, in A_0.

3.1 The Random Placement Model (RPM)

Imagine distributing the n_0 individuals of a species onto a landscape whose entire area A_0 is gridded into 2^i cells of area A_i, by letting the probability that any particular individual lies in any particular A_i cell be 2^{-i}. For obvious reasons this is called the "random placement model" [3], and in it the function $P_i^{(n_0)}(n)$ is given by the binomial distribution.

$$P_i^{(n_0)}(n) = \frac{n_0!}{n!(n_0-n)!}(2^{-i})^n(1-2^{-i})^{n_0-n} \ . \tag{8}$$

Because of the simplicity of its assumptions, the RPM [3] is often taken as a null model in ecology. That is, by looking for deviations from its predictions, the RPM is used to determine whether "anything interesting is going on". As I will argue later, there may be more useful, yet equally basic, null models.

The shape of the distribution in Eq. 8 depends on whether n_0 is greater or less than 2^i. In the former case, the function is hump-shaped, with a maximum at $n \sim n_0/2^i$, while in the latter case the function is monotonically decreasing in n.

Using Eq. 3, the species area relationship (SAR) in this model is given by

$$S_i = \sum_{\eta} \Phi_0(\eta)[1-(1-2^{-i})^{\eta}] \tag{9}$$

where the function $\Phi_0(\eta)$ is the probability density describing the distribution of abundances in A_0 and the variable η in this expression is just a dummy variable for the total number of individuals in a species. Tests of Eq. 9 require

inserting the empirical species abundance distribution $\Phi_0(\eta)$ into the equation because the model does not predict that distribution.

Because the assumptions of the RPM preclude spatial correlations, the model predicts that the species-level and community-level metrics, χ, in Table 1 are independent of distance D between cells.

3.2 A Fractal Model

To define this model, I introduce two probability parameters. At the species level, each species is assigned a probability $\alpha_i^{(n_0)}$ that describes the probability that if that species is present in A_0 and also in a grid cell of area A_{i-1}, then it is present in a pre-specified one of the two A_i cells that comprise the A_{i-1} cell [21]. Defining

$$\lambda_i^{(n_0)} \equiv \prod_{j=1}^{i} \alpha_j^{(n_0)} \tag{10}$$

it is easy to show that $\lambda_i = 1 - P_i^{(n_0)}(0)$, or in other words, λ_i is the probability that the species is present in an arbitrarily chosen A_i cell [21], [22]. Note that from Eq. 2, the Range-Area Relationship for each species is uniquely specified in terms of the $\lambda-$parameters.

If the α's are independent of scale, i, so that each $\alpha_i^{(n_0)} = \alpha^{(n_0)}$ then the spatial distributions of each species are self-similar. The spatial distribution of each species then has a fractal dimension that is a function of the α-value for that species. Now the $\lambda_i^{(n_0)} = (\alpha^{(n_0)})^i$ and the box-counting measure of occupancy, $R_i^{(n_0)}$, is related to cell area by

$$R_i^{(n_0)} = A_0(\alpha^{(n_0)})^i \ . \tag{11}$$

Recalling that $A_i = A_0/2^i$, Eq. 11 is mathematically equivalent to the power-law form $R_i^{(n_0)} \sim A_i^y$, with $y = -\log_2(\alpha^{(n_0)})$. The fractal dimension of the spatial distribution, D, is related to y by $D = 2(1-y)$.

Using Eq. 3, the SAR can be expressed as

$$S_i = \sum_{\text{species}} (\alpha^{(n_0)})^i \ . \tag{12}$$

At the community level, I introduce another probability parameter, a_i, defined to be the ratio of the number of species found at scale $i-1$ to that at scale i. In terms of this parameter, we can write

$$S_i = S_0 \prod_{j=1}^{i} a_j \ . \tag{13}$$

If the a_i are scale-independent, then Eq. 13 reads, in analogy with Eq. 11:

$$S_i = S_0(a)^i \ . \tag{14}$$

Recalling that $A_i = A_0/2^i$, this is mathematically equivalent to the power-law form of the SAR, $S_i \sim A_i^z$, with $z = -\log_2(a)$.

For the case of scale-independence at both species level (α's are scale-independent) and community level (a's are scale-independent), we can equate the expressions for S_i in Eqs. 12 and 14, and thereby arrive [21] at an important rigorously true "impossibility theorem" relating the probability parameters at the species level and the community level:

$$a^i = \langle (\alpha^{(n_0)})^i \rangle_{\text{species}} \tag{15}$$

where $\langle \, . \, \rangle_{\text{species}}$ denotes an average over all species. I refer to Eq. 15 as an impossibility theorem because it cannot in general be satisfied for all $i-$values or even over an adjacent pair of $i-$values, $(i, i+1)$. The only case in which self similarity can hold at both community level and at species level is if the α's are all equal to each other and at the same time equal to a.

If the α's are scale invariant, then we arrive at a power-law Range-Area Relationship for each species (Eq. 11), whereas if a is scale invariant then we arrive at a power-law SAR (Eq. 14). But both metrics cannot be power-law except in the case in which all the powers are equal. In Section V, I discuss which, if either, of these options "nature chooses".

The species level spatial abundance distributions $P_i^{(n_0)}(n)$ can be derived in the Fractal Model subject to a boundary condition at some smallest scale, $i = m$, at which there is assumed to be either 0 or 1 individual within each cell. Leaving out the species label (n_0) for simplicity, I define a probability distribution, $Q_i(n)$ for $n \geq 1$, that is conditional on the species being present in A_i by

$$Q_i(n) = \frac{P_i(n)}{1 - P_i(0)} \ . \tag{16}$$

By scaling up from that smallest scale, $i = m$, the following recursion relationship can be derived [18] for the conditional spatial probability distribution

$$Q_{i-1}(n) = 2(1-\alpha)Q_i(n) + (2\alpha - 1)\sum_{\eta=1}^{n-1} Q_i(n-\eta)Q_i(\eta) \tag{17}$$

subject to the boundary condition $Q_m(n) = \delta_{n,1}$. It has been shown [18] that the solution to this equation has the property that the average abundance of the species over *occupied cells* of area $A_i = 2^{m-i}A_m$ is $(2\alpha)^{m-i}$. Because there are n_0 individuals in A_0, consistency is achieved if

$$m \log(2\alpha) = \log(n_0) \ . \tag{18}$$

The solution to Eq. 17, at any fixed scale, i, and for $n \geq 1$, is unimodal, with a rising power-law dependence of the $Q_i(n)$ on n for small n, and a

faster-than-exponential fall off at large n. For $n = 0$, self similarity results in the simple expression $P_i^{(n_0)}(0) = 1 - \alpha^i$.

In the publication in which Eq. 17 was first derived [18], we replaced the species-level parameter, α, by the community-level parameter, a, and claimed that the community $Q_0(n)$ could be interpreted as a species-abundance distribution $\Phi_0(n)$. Implicit in that argument, though not recognized or stated in the publication, was the assumption that the α's for all species are identical and equal to the community parameter, a. The assumption that species have identical α's, however, does not hold for any actual ecosystem. Therefore the solutions to Eq. 17 must be interpreted as a prediction of the shape and scale dependence of the conditional species-level spatial abundance distributions under the Fractal Model.

The commonality metrics in Table 1 are also constrained in the self similarity model. If self similarity holds at the community level (and therefore, by the impossibility theorem, not at the level all the species), then we have shown using a scaling argument [19] that the community-level commonality metric behaves as $\chi(A_i, D) = b(A_i/D^2)^z$, where z is the SAR exponent and b is a constant. For a species with a scale-invariant α, the species level commonality function, $\chi^{(n_0)}(A_i, D)$, can be derived using a similar scaling argument, yielding $\chi^{(n_0)}(A_i, D) = b'(A_i/D^2)^y$, where $y = -\log_2(\alpha)$.

3.3 The HEAP Model

If a collection of distinguishable objects are randomly assigned to the two halves of a box, the binomial distribution is obtained. Now consider what happens if the objects are assumed to be indistinguishable. At the first division of a box, the counting rule becomes HEAP: the hypothesis of equal allocation probabilities. In particular, for n_0 objects, the n_0+1 options: $(0, n_0)$, $(1, n_0-1)$, $(2, n_0-2)$, $\ldots$, $(n_0, 0)$ are all equally likely [22] and so

$$P_1^{(n_0)}(n) = \frac{1}{n_0 + 1} \ . \tag{19}$$

The HEAP model assumes that at each successively smaller scale division, the same rule applies. Thus, if at the first division, n out of the n_0 objects happen to be assigned to the left side of the box, then the probability that there will be n' objects in the upper left quadrant (the upper half of the left half of the box) is just $1/(n+1)$. As shown in [22], the HEAP allocation rule leads directly to the following recursion relation for the species-level spatial abundance distribution $P_i^{(n_0)}(n)$:

$$P_i^{(n_0)}(n) = \sum_{q=n}^{n_0} \frac{P_{i-1}^{(n_0)}(q)}{(q+1)} \ . \tag{20}$$

From the solutions to Eq. 20, all the other metrics in Table 1, with the exception of the commonality distributions, can be easily calculated using Eqs.

1 – 6. Elsewhere [23], I have shown that in the HEAP model, the species-level commonality metrics, $\chi^{(n_0)}(A_i, D)$ can be derived from two coupled exact recursion relationships. Moreover, if the species are assumed to be distributed independently of one another, then the community-level commonality metric follows from Eq. 7.

There is an alternative formulation of the HEAP model that may provide a better basis for developing a mechanistic understanding of it. In particular, Eq. 20 and the HEAP statistical rule from which that recursion relation is derived are mathematically equivalent [22, 23] to the following "assembly" or "colonization" rule. Consider the individuals of a single species being allocated to the two halves of a cell in which the species is known to exist. Assume the individuals are allocated, one individual at a time until all the individuals have been allocated using the following sequential colonization rule:

$$\beta(\text{left}\,|l, r) = \frac{l+1}{l+r+2}\ . \tag{21}$$

Here β is the probability that if there are l and r individuals in the left and right halves of the cell, then the next individual is allocated to the left half. Symmetrically, individuals are allocated to the right with probability $\beta(\text{right}\,|l, r) = (r+1)/(l+r+2)$. This assembly rule can be applied successively to generate spatial distributions at finer and finer scale. Thus, if n_0 individuals are known to exist in an A_0, then Eq. 21 generates a division of those individuals into the two A_1 cells that comprise A_0. Continuing in this way, the individuals in each of the two A_2 cells comprising each A_1 cell can be assembled. This process can be iterated down to finer spatial scales.

Note that the assembly rule given by Eq. 21 promotes aggregation; at any stage in the assembly process, the half that has more individuals has a greater than 50% probability of attracting the next individual. For example, if $l = 5$ and $r = 4$, then the probability that the 10^{th} individual is added to the left half is $(5+1)/(5+4+2) = 6/11 > 1/2$. In contrast, in the RPM, the probability that a new individual is added to the left half is always exactly $1/2$.

Eq. 21 can be seen as a special case of the more general allocation rule:

$$\beta(\text{left}\,|l, r) = \frac{l\theta+1}{l\theta+r\theta+2}\ . \tag{22}$$

When $\theta = 0$, the rule is equivalent to the random-placement model [3], while $\theta = \infty$ yields a maximally-aggregated distribution in which, at every scale i, all individuals of each species are located in just one of the 2^i A_i cells. Within a range of negative values of θ, Eq. 8 yields a distribution of individuals that is more uniform than random. The value $\theta = 1$ corresponds to HEAP. Not all spatial distributions are generated by repeated application of Eq. 22; e.g., the negative binomial distribution and the Fractal Model distribution do not appear to correspond to any value of θ. Further generalizations of Eq. 22 have been discussed by Conlisk et al. [4], who also consider the case in

which the bisection scheme that defines scale in Eqs. 20 – 22 is replaced by a more general $K-$section scheme, in which at each scale iteration, a given cell is divided into K equal-area cells. This more general class of models now includes as a special case ($K = 2^i$, $\theta = 1$) a form of the negative binomial distribution [5] that is conditional on the total abundance being exactly n_0.

3.4 The MaxEnt Method

Edwin Jaynes [30–32], building on the work of Claude Shannon [45], derived a remarkable result about inference within a maximum likelihood framework. Suppose you seek the least biased estimate of the functional form of the probability distribution $p(n)$, subject to a set of M constraints that you accept from prior knowledge and that can be expressed in the form of M equations:

$$\sum_n f_k(n)p(n) = \langle f_k \rangle \ , \tag{23}$$

where n is summed over all of its possible values, and the index k runs from 1 to M. Then, given the information from these constraints, the best inference as to the shape of $p(n)$ is the function that maximizes the "information entropy":

$$I = -\sum_n p(n) \ln(p(n)) \tag{24}$$

subject to those constraints. Maximization is carried out using the method of Lagrange multipliers. This procedure yields:

$$p(n) = \frac{\exp\left(-\sum_k \lambda_k f_k(n)\right)}{Z(\lambda_1, \lambda_2, \ldots, \lambda_M)} \ , \tag{25}$$

where Z, the partition function, is given by:

$$Z(\lambda_1, \lambda_2, \ldots, \lambda_M) = \sum_n \exp\left(-\sum_k \lambda_k f_k(n)\right) \ , \tag{26}$$

and the λ_k are given by the solutions to:

$$\frac{\partial \ln(Z)}{\partial \lambda_k} = -\langle f_k \rangle \ . \tag{27}$$

If it should turn out that the $p(n)$ obtained from maximizing Eq. 24 subject to imposed constraints fails to provide reliable predictions for a data set, then that indicates that some of the assumed constraints did not really hold and/or that additional constraints do hold but were not included in the M constraint equations.

The MaxEnt prodedure can be readily applied to spatial ecology. Consider the species-level spatial abundance distribution, $P_i^{(n_0)}(n)$. A trivial constraint

on P is that $\sum_n P(n) = 1$, but this, alone, simply leads to the prediction of an n-independent $P(n)$. Another constraint arises because we know the total number of individuals, n_0, at the largest scale A_0. In particular, at scale i we know that the average number of individuals in an A_i cell is $n_0/2^i$. Hence a non-trivial constraint equation is:

$$\sum_n n P_i^{(n_0)}(n) = \frac{n_0}{2^i} \ . \tag{28}$$

Here the sum extends from $n = 0$ to $n = n_0$, the range of possible values of n. Letting $x = \exp(-\lambda)$, where λ is the single Lagrange multiplier for this single-constraint problem, the normalized MaxEnt solution is:

$$P_i^{(n_0)}(n) = \frac{x^n}{Z(x)} \ , \tag{29}$$

where Z is equal to:

$$Z = \sum_n x^n \tag{30}$$

and x is, from Eq. 29, given by the solution to:

$$\sum_n n x^n = \left(\frac{n_0}{2^i}\right) Z \ . \tag{31}$$

Consider, first, the case $i = 1$, corresponding to the allocation of objects between two halves of a box. The unique real-valued, non-negative solution to Eqs. 27, 28 is $x = 1$, or $\lambda = 0$, and thus from Eqs. 29, 30,

$$P_1^{(n_0)}(n) = \frac{1}{\sum_n 1} = \frac{1}{n_0 + 1} \ . \tag{32}$$

This is identical to the result in Eq. 19, and thus the HEAP and MaxEnt models give identical predictions for $P_1^{(n_0)}(n)$, the allocation of individuals between two halves of a plot. At finer spatial scales, however, the two models diverge.

If the MaxEnt procedure is applied naively, with a single constraint, to the prediction of the species abundance distribution $\Phi_0(n)$, the constraint equation would read

$$\sum_n n \, \Phi_0(n) = \frac{N_0}{S_0} \tag{33}$$

where N_0 is the total number of individuals, summed over all species, in A_0, and S_0 is the total number of species in A_0. The sum in Eq. 33 ranges from $n = 1$ to $n = N_0 - S_0 + 1$, which is the maximum possible number of individuals there can be in any one species.

In MaxEnt, with a single constraint on the mean value of n (Eq. 33), the resulting $p(n)$ is always an exponentially decreasing function of n. In reality, observed species abundance distributions in ecosystems are nearly always either

unimodal or, if monotonically decreasing do so with approximate $1/n$ behavior rather than exponentially. Thus, some other constraint must be invoked. Elsewhere we have shown that more realistic species abundance distributions result from introducing a second constraint equation [24]. One such possible additional constraint derives from an approximate empirical relation [10, 14] between the abundance of a species and the body mass of individuals within the species:

$$n(m) \sim \beta m^{-3/4} \ , \tag{34}$$

where β is a constant. From Eq. 31, we have:

$$\sum_n n\, m(n)\, \Phi_0(n) = \beta^{4/3} \sum_n n^{-1/3} \Phi_0(n) = \frac{M_0}{S_0} \ . \tag{35}$$

Here, M_0 is the total mass of all individuals in all species within A_0. With this additional constraint, species-abundance distributions that are unimodal result from realistic values of the ratios on the right hand sides of Eqs. 33 and 35. These distributions have the realistic property that the predicted $\Phi_0(n)$ are approximately lognormal but skewed toward an excess of species with $n < n_{\text{modal}}$.

Although it has been argued that Eq. 34 is a pervasive pattern in ecosystems [14], there is consider able scatter around that central tendency and available evidence may also be consistent with other mass-abundance relationships. I present next a possible way to derive both the mass-abundance relationship, and simultaneously the species-abundance relationship at scale A_0, using the MaxEnt approach.

Consider a probability distribution $R_0(n, m)$ defined over the species in A_0. Here n labels species abundance and m labels a typical mass value for a species. R_0 is the probability that a species picked from the species pool in A_0 has abundance n and the average mass of the individuals of that species is m. From R_0, the species-abundance distribution is obtained by summing over m:

$$\sum_m R_0(n, m) = \Phi_0(n) \ . \tag{36}$$

One constraint equation for R_0 is the normalization condition $\sum_{n,m} R_0(n, m) = 1$. In addition, as in Eq. 35,

$$\sum_{n,\, m} R_0(n, m)\, n\, m = \frac{M_0}{S_0} \ . \tag{37}$$

And, as in Eq. 33,

$$\sum_{n,\, m} R_0(n, m)\, n = \frac{N_0}{S_0} \ . \tag{38}$$

The central tendency of the mass-abundance relationship can be derived from the distribution, R, by solving for the "ridge line" of that distribution.

The solution to this two-variable MaxEnt problem remains to be explored and tested against data.

Finally, the two-point and possible higher-point distributions can be derived within the MaxEnt framework. Because the MaxEnt derivation of the $P_i^{(n_0)}(n)$ is not recursive, but rather results in an independent calculation at each scale, consistency relationships exist across adjacent scales. It can be shown that the imposition of consistency requirements across scales results in predictions of at least some of the $N-$point distributions. Some examples are shown in Figure 1 for low abundances, in which case the calculations are simple. As n_0 and/or i increases, the number of consistency constraints becomes exceeded by the number of possible $N-$point diagrams and so redundancy classes of diagrams emerge, within which probabilities are not distinguishable without additional constraints.

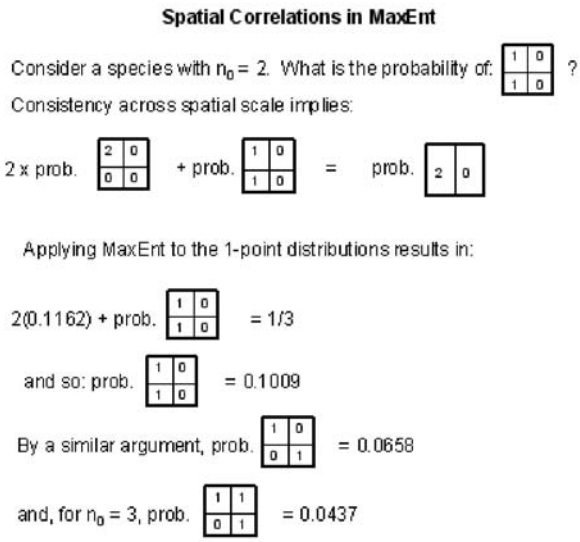


Fig. 1. Illustration of how imposing consistency across scales under MaxEnt permits calculation of $N-$point distributions from 1-point distributions. The pictorial equation that results from consistency across scales states that there are three distinct ways to allocate both individuals to the left side of the box: (both in the upper left; both in the lower left; one in the upper left and one in the lower left). MaxEnt predicts that the probability of both on the left is 1/3, and that the probability of both in a quadrant is 0.1162. Hence the joint probability of one in the upper left and one in the lower left is $1/3 - 2(0.1162) = 0.1009$.

4 Intercomparisons of the Four Models

Because $P_i^{(n_0)}(n)$ is so fundamental in spatial ecology, I focus here on how the models differ in their predictions of this distribution. First I present the prediction of each model for the relationship between the variance, σ_i^2, of the distribution and the mean, $\langle n_i \rangle = n_0/2^i$:

Random Placement Model: $$\sigma_i^2 = (1 - 2^{-i})\langle n_i \rangle \tag{39}$$

Fractal Model: $$\sigma_i^2(n_0, i) = \left(\alpha^{-1-i} - 1\right) \langle n_i \rangle^2 - \left(\alpha^{-1} - 1\right) \langle n_i \rangle \tag{40}$$

HEAP: $$\sigma_i^2 = \left[\left(\frac{4}{3}\right)^i - 1\right] \langle n_i \rangle^2 + \left[\left(1 - \frac{2}{3}\right)^i\right] \langle n_i \rangle \tag{41}$$

MaxEnt: $$\sigma_i^2 \to \langle n_i \rangle^2 + \langle n_i \rangle \ . \tag{42}$$

Eq. 39 follows directly from Eq. 8. In the Fractal Model, an expression for the variance of n across cells in which it is present can be obtained by writing a recursion relation for that variance using Eq. 17 [1]. Denoting this conditional variance by σ'^2 and the conditional mean of n_i by $\langle n_i' \rangle$, the result is $\sigma'^2 = [\langle n_i' \rangle^2 - \langle n_i' \rangle](\alpha^{-1} - 1)$. From this expression, the unconditional variance can then be derived using $P_i^{(n_0)}(0) = 1 - \alpha^i$, yielding Eq. 40. Eq. 41 is derived [22] from an analytical solution, expressed as a finite sum, to Eq. 20. In Eq. 42 the $\to$ indicates that the variance approaches the rhs as i increases. A numerical analysis of Eqs. 29 – 31 indicates that for $n_0 > 2$ and $i > 3$, the parameter x is well approximated by $x \approx n_0/(n_0 + 2^i)$ and from that expression Eq. 42 can be derived. A closed-form expression for the variance of n under MaxEnt in the general case has not been derived.

As seen from these expressions, the models differ considerably in their variance to mean ratio (VMR). The Random Placement Model generally results in the lowest VMR, meaning that occupation numbers in the cells tend to tightly cluster around the mean. Thus in this model there is the least probability of large aggregates within small cells and also the fewest number of cells with no individuals. The HEAP and MaxEnt distributions have the same form, a linear combination of a term that is constant and a term that is proportional to $\langle n \rangle$, but differ at small scales (large i) where the VMRs predicted by HEAP exceed those from MaxEnt because of the $(4/3)^i$ coefficient. In that sense, MaxEnt is intermediary between random and HEAP, predicting fewer cells with significantly sub-mean or supra-mean occupancy than HEAP but more such cells than in the random model.

Values of α have to fall between 1/2 and 1, equaling 1/2 only if $n_0 = 1$, and so the coefficient multiplying $\langle n_i \rangle^2$ on the rhs of Eq. 40 is generally greater than 1 and decreases with increasing n_0. At fixed i and increasing n_0, the VMR under self similarity is considerably smaller than that under MaxEnt or HEAP because α must approach 1 as n_0 increases.

The limiting value of σ_i^2 under MaxEnt, given in Eq. 42, is identical to the variance of the unconditional (on n_0) negative binomial distribution; but the MaxEnt model is conditional on total abundance, and so the MaxEnt result for $P(n)$ and the negative binomial distribution are fundamentally different.

The actual shape of the $P_i^{(n_0)}(n)$ in each model depends on the scale parameter, i, and on the total abundance of the species, n_0. In addition, in the self similarity model, the shape also depends on the parameter, m, characterizing the scale at which there is at most 1 individual per cell, or equivalently, on the parameter α, which is related to m and n_0 by Eq. 18.

Figure 2 compares the species-level spatial abundance distributions for the case in which $n_0 = 4$, $i = 2$. The additional specification of $m = 4$ and thus $\alpha = 2^{-1/2}$, is made for the self similarity model. Note that HEAP and MaxEnt are nearly indistinguishable for this case, but differ considerably from the predicted distributions in the random and the self similar models. As shown in the next Section, in the limit in which $2^i \gg n_0$, all the models converge toward the random placement model.

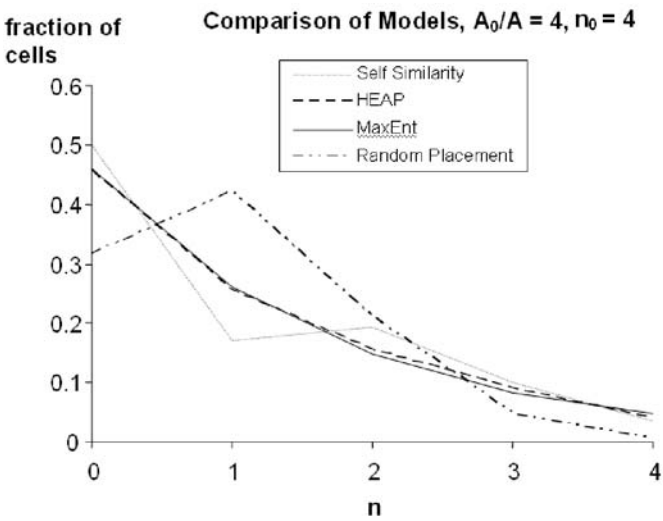


Fig. 2. Comparison of species-level spatial abundance distributions, $P_2^{(4)}(n)$ for the four models.

5 Tests of Models: an Overview

In a series of papers, we have characterized in detail the successes and failures of these models to capture spatial patterns [5,17,21,22]. I will not repeat here

such a systematic set of comparisons but rather highlight the essential results that emerge from these previous comparisons.

5.1 Summary of Tests of the Fractal Model

Because the Fractal Model, requires specification of the parameter, m, for each species, we have thoroughly evaluated that model separately from the others [17] using data from a serpentine grassland in California, where on a 64 m^2 plot every individual plant was identified to species and located within a grid of cell size 1/4 m^2. By assuming that the community-level probability parameter a is scale-independent, the model is consistent with an observed power-law species-area relationship for the serpentine site [17], but the model cannot predict the value of a and thus of the SAR exponent z.

Moreover, many observed species-area relationships do not follow a power-law form and thus the model is readily falsified at the community level in many ecosystems. At species-level, studies of plant species distributions both at the serpentine site [17] and in the UK [35] suggest that some, but by no means all, plant species do exhibit an approximately fractal spatial distribution and that the fractal dimension varies with the overall abundance of the species in a manner consistent with the model. Using spatially explicit census data on breeding birds in the UK and on ferns in Canada, the model has also successfully predicted the relationship (Eqs. 11, 18) between box counting range and abundance of species [21] under the ad hoc assumption that the $m-$values of all species are identical and given by $2^m = \sum_{\text{species}} n_0 = N_0$.

My summary assessment of the Fractal Model is that the predicted power-law behaviors are sometimes, but not generally, observed in the spatial structure of ecosystems at either species- or community-level. The model has at least as many as failures as successes.

5.2 Summary of Tests of RPM, HEAP, and MaxEnt

To test the other 3 models, we have made use primarily of plant census data from three sites. One is the serpentine grassland mentioned above. The second is a 50 ha wet tropical forest plot in Panama [6–8, 27] where every individual tree with diameter-at-breast-height greater than 1 cm has been identified and point located, and the third is a 9.68 ha dry tropical forest plot in Costa Rica [13], where the same type census as in the 50 ha plot was carried out. At each of these sites, censusing has been carried out in more than one year, so that the temporal dynamics of pattern can be established. The total number of individuals and species in each site is given in Table 2.

5.3 Species-Level Spatial Abundance Distributions

Generally, the Random Placement Model poorly describes observed spatial scaling patterns for vegetation, except in the limit $2^i \gg n_0$ where the models

Table 2. Description of three sites from which spatially-explicit census data are used here to test spatial theories.

Site	64 m^2 serpentine grassland plot in CA (2005 census)	50 ha wet tropical forest plot at BCI, Panama (19xx census)	9.68 ha dry tropical forest plot at San Emilio, CR (19xx census)
# plant species in entire plot	28	305	138
# individual plants in entire plot	61,000	235,308	12,851

are difficult to distinguish. The essential failure of Eq. 8 is that it under-predicts the amount of clumping of individual plants within each species; its hump-shape contrasts with the generally-observed monotonically decreasing shape of the $P_i^{(n_0)}(n)$, whatever the relationship between n_0 and 2^i. Stated differently, the Random Placement Model under-predicts the number of vacant cells and over-predicts the probability that a cell will have the average number of individuals.

Figures 3a, 3b show the spatial abundance data for a single species, with $n_0 = 3144$, at the BCI plot. Figure 3a shows a case in which $2^i < n_0$ and provides an example of the typical failure of the Random Placement Model for plants. It also shows MaxEnt outperforming HEAP and, again, in a way that is typical; for the high-abundance species, HEAP tends to over-predict the number of cells with 0 individuals, or equivalently over-predicts the VMR relative to MaxEnt. Figure 3b shows data for the same species but at a finer spatial scale so that $2^i \gg n_0$. Here, the 3 models all converge and are in good agreement with observation.

5.4 Species-area Relationships

Testing the ability of each of the models to predict the $P(n)$ for all the species at all our censused sites has not yet been carried out, but a test of the ability of the models to predict the species-area relationship at each site partially accomplishes that goal because by Eq. 3, all of the species' spatial distributions, evaluated at $n = 0$, enter into the prediction.

A comparison of the observed species-area relationship at the serpentine site with the Random Placement Model prediction for $P_i^{(n_0)}(n)$ (Eq. 9) shows that the model prediction deviates considerably from observation. Generally Indeed, for the serpentine site, the empirical species-area relationship fits a power law model very well ($S = cA^z$, with $z = 0.21xx$, $R^2 = 0.999$), whereas the Random Placement Model predicts considerable curvature (negative second derivative) on log-log axes. Such power law behavior is not always observed – for example, at the BCI and San Emilio sites – but even in those

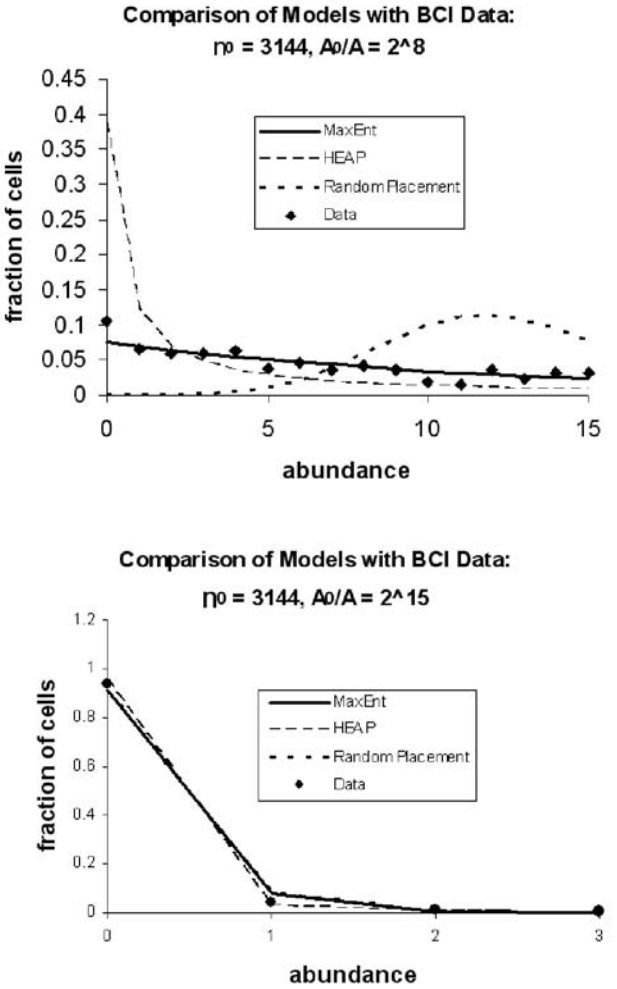


Fig. 3. Test of MaxEnt, HEAP and the Random Placement Model predictions for $P_i^{(3144)}(n)$ for (a) $i = 8$ ($A_0/A = 256$) and (b) $i = 15$ ($A_0/A = 32768$) within the 50 ha BCI plot.

cases, the Random Placement Model poorly matches observation. Figure 4a compares the BCI SAR against both MaxEnt and HEAP; clearly MaxEnt outperforms HEAP. A similar result holds for the San Emilio data (Figure 4b). Using the empirical species-abundance distribution for the serpentine site, species-area relationship is predicted by MaxEnt to be of power-law form, although the slope is predicted to be ~ 0.17 instead of ~ 0.21.

5.5 Generalized HEAP Models

The HEAP model is a special case, $\theta = 1$, of the more general class of spatial models defined by Eq. 22. Because such a wide variety of distributions are all characterized by the value of a single parameter, θ, likelihood testing is readily carried out without concern for degrees-of-freedom adjustment. In Harte et al. [22] we conducted such likelihood tests for each species at our three census sites and found that the likelihood-maximizing value of θ tends to decrease from 1 (the HEAP value) or larger, toward 0 (the random placement value), as species abundance in A_0 increases.

In Harte et al. [22] we noted that this trend toward a smaller best-fitting $\theta-$value as n_0 increases suggests the existence of a negative effect of crowding on the tendency of individuals within a species to aggregate. Such density-dependent regulation of spatial distributions is not included in the HEAP model, but any mechanistic model that is proposed to explain HEAP at low population density should incorporate such an effect of crowding as density increases. Although the MaxEnt model does not exactly correspond to a particular $\theta-$value in the family of generalized HEAP models at any scale (except $A = A_0/2$, where it corresponds to $\theta = 1$), it does predict a crowding trend that is in the right direction: the MaxEnt predictions for $P_i^{(n_0)}(n)$ approach the random model distribution faster than do the $\theta = 1$ HEAP model predictions as abundance increases.

This is illustrated in Figure 5, which shows data on the distribution of ant colonies in an open field in Colorado (Reithel, pers. Comm..). The scale of analysis here is $A = A_0/256$, and $n_0 = 97$. Thus n_0 is less than 2^i but not much less, and so the HEAP distribution ($\theta = 1$) is approaching, but still distinct from, RPM. As seen in the figure, MaxEnt and RPM well describe the data but HEAP deviates considerably.

6 Advantages of the Maximum Entropy Framework in Ecology

Although HEAP and MaxEnt yield identical predictions at the scale $A = A_0/2$, and in many ways have more in common than either does with the Random Placement Model or the Fractal Model, there are some fundamental differences. A drawback of the HEAP model is that distributions at finer

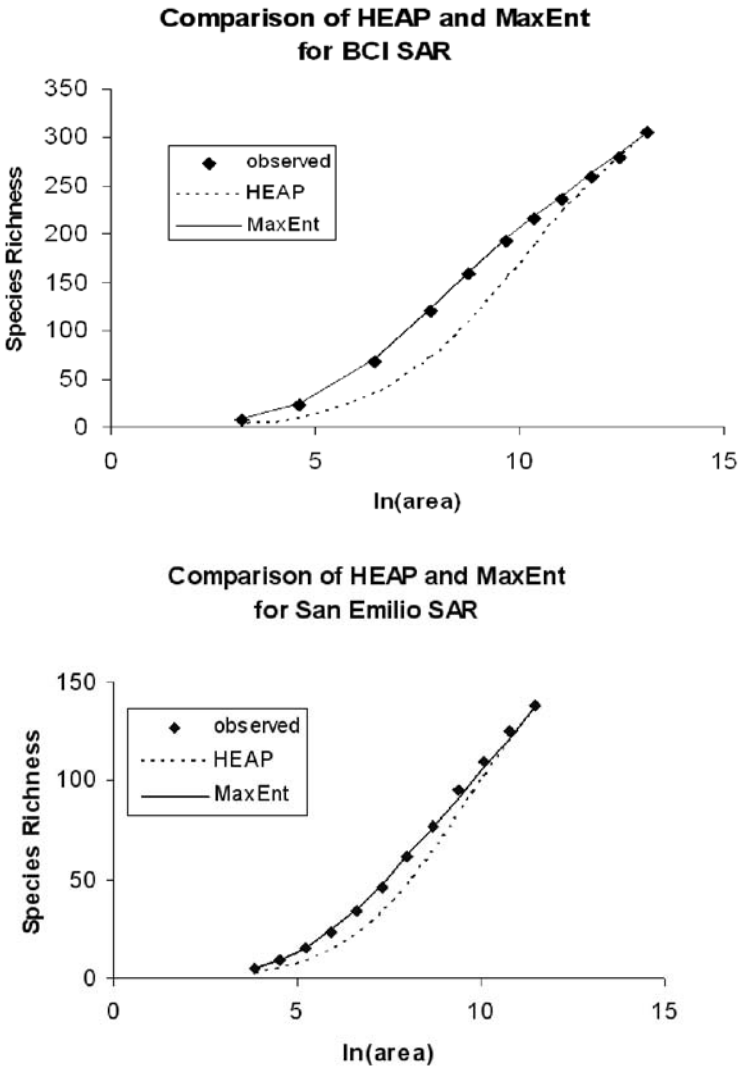


Fig. 4. Test of MaxEnt and HEAP for the Species-Area Relationship at (a) the 50 ha BCI plot, and (b) the 9.68 ha San Emilio plot.

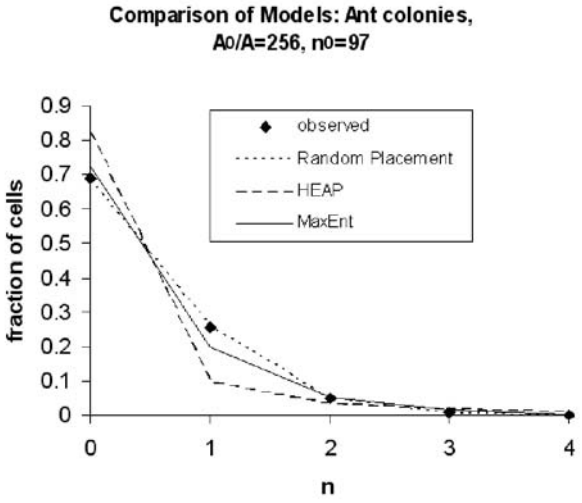


Fig. 5. The function $P_i^{(n0)}(n)$ for ant colonies in an open field in Colorado. Here $n_0 = 97$ and the scale of analysis is $i = 8$. Data from J. Reithel, pers. comm.

spatial scales are determined recursively from distributions at coarser scales, with scale change defined by successive bisections of A_0. It has been pointed out [20, 37] that this bisection procedure (or any analogous $n-$section procedure) can introduce unrealistic features into landscapes created by such models. For example, nearest-cell correlations depend on whether the bisection line between the two cells is a primary bisection (which creates two cells of area $A_0/2$) or a higher order bisection (which, for example, might create two cells of area $A_0/4$ out of a cell of area $A_0/2$). While it has been shown that these features are avoided if appropriate "user rules" are followed [39], it would be more satisfying to not have to invoke such user rules. Because the MaxEnt approach is not recursive in nature, it avoids the necessity of user rules to interpret the structure of space within the model framework.

Another drawback of HEAP is that the mechanistic origin of this statistical model is not identified; it is a purely statistical model. Although our applications of the MaxEnt principle give no explicit role to traditional mechanisms in ecology, such as birth, death, dispersal and competition, they can incorporate systemic knowledge in the form of constraints. The ecological constraint equations are analogous to the conservation laws used to derive Gibbsian statistical mechanics from MaxEnt [30]; in that sense only, the MaxEnt predictions of patterns in DAS are mechanistic in origin.

Our first application of MaxEnt to the derivation of the species-abundance distribution (Eqs. 33 – 35) relied upon the constraint of an emprical size-abundance relationship, which has been criticized because there is consider-

able scatter in the data around the central tendency explained by the power-law model. But in MaxEnt, the constraint Eq. 18 does not require that all species fall exactly on a straight line on a log(abundance) versus log(body size) graph; rather the constraint equations are statements about expectation values. The MaxEnt approach is thus advantageous for ecology because constraint equations that stem from models that only capture central tendencies can be used to make inferences, even when there is considerable scatter around the central tendencies. Our second application of MaxEnt to the derivation of the species abundance distribution and the mass-abundance distribution (Eqs. 36 – 38) has the added advantage that the distribution of the scatter around the central tendency is predicted as well.

7 Toward a Unified Theory of Ecology

The search for a unified theory of ecology has not, historically, been a preoccupation of ecologists. One reason is the widely recognized evolutionary and ecological importance of, and the aesthetic delight derived from, the existence of the enormous complexity of ecosystems. This complexity includes the remarkable variety of differences among species and even among populations and individuals within species, and has led to skepticism that a widely-applicable elegant theory with few or no adjustable parameters could exist. Referring back to the historical timeline characterizing progress in developing a theory of gravity, some might argue that in ecology the progression will never get further than the Keplerian stage, and may not even get that far; there will be no Newton.

How would we recognize a unified theory of ecology if we had one? I suggest three requisite traits: a unified theory should be comprehensive, lean, and falsifiable.

A *comprehensive* theory would predict the central tendencies, though certainly not all the variability around the central tendencies, of the spatial, temporal, energetic, material, and informational aspects of ecosystems and biodiversity. We emphasize the term "central tendencies". No theory of ecology is going to capture all the details. Just as Boyle's law, $PV = nRT$, does not accurately predict the behavior of gases at very high pressure or temperature (in other words, it does not provide a description of all the behaviors of gases under all conditions), so a comprehensive theory of ecology will often fail to explain the details of many ecologists field observations.

A *lean* theory needs to be as simple as possible, but no simpler. It will be parsimonious in the sense that it explains a lot with a little; the ratio of the number of distinctly different phenomena it predicts to the number of assumptions and adjustable parameters must be $\gg 1$. Antoine de Saint Exupery expressed it well in "Wind, Sand and Stars": *"In anything at all, perfection is finally attained not when there is no longer anything to add but when there is no longer anything to take away".*

A *falsifiable* theory sticks it neck out. Theory that is not falsifiable is not science. Falsifiability requires explicitly-stated assumptions and an absence of excessive adjustable parameters, and so leanness is a necessary condition for falsifiability. Because a unified theory will only predict central tendencies, it will, strictly speaking, be frequently falsified. That is, in numerous confrontations with data, observations will inevitably deviate from predictions. These "failures" should not necessarily be viewed as a reason to discard the theory, but rather can be used as a source of insight into why the theory works to the extent it does. Moreover, from its failures, many important insights into the importance of neglected mechanisms will emerge and that will help advance understanding. The difficulty, of course, is knowing when to discard an entire theory because of failures of its predictions.

The issue of pattern versus process frequently confounds discussions in theoretical ecology. Mechanisms at one level of description may appear as phenomenological at another. Newtonian gravity provided a mechanism that explained the patterns Kepler detected in planetary orbits. But Newton did not have a mechanism for how mass produces gravity. Einsteinian gravity does have a mechanism (mass curves space and that affects the orbits of objects) ... but then how does mass curve space? The demand for mechanism can lead to an infinite regress. We suggest that in ecology there are instances where it may be most useful to close the circle with self-consistency constraints that avoid the question of what is cause and what is effect. Of course clusters of interacting variables, with undecipherable causation amongst them, may sometimes be usefully viewed as causal units with respect to other clusters.

Some components of a unified theory may already exist. The Metabolic Theory of Ecology, initially proposed by West, Brown and Enquist [46] and further developed by numerous authors such as Banavar et al. [2] and Ettienne et al. [15], provides an elegant way to begin to understand from some fundamental assumptions many patterns in the energetics of ecosystems. In the pure form of this theory, the only organismal trait that matters is body size. The Neutral Theory of Ecology [28] demonstrates that the neglect of differences in birth and death rates among individuals within and across species still leads to reasonable predictions about community structure. And the MaxEnt framework may provide an even more fundamental and assumption-free component of a future unified theory.

The promise of the MaxEnt approach can be summarized as follows. Suppose we are given an area A_0, and within that area a total number of species, S_0, a total number of individuals, N_0, and a total biomass, M_0. From the ratios M_0/N_0 and M_0/S_0, MaxEnt then predicts the species abundance distribution at scale A_0 (Eqs. 34, 35). From the resulting distribution of the $\{n_0\}$ the spatial distributions $P_i^{(n_0)}(n)$ can be worked out at any scale from Eqs. 29 – 31, and then all the other one-point metrics in Table 1 follow. Consistency requirements on the $P_i^{(n_0)}(n)$ across spatial scales then place constraints on the $N-$point metrics such as commonality, as exemplified in Fig. 1. A large

number of predictions about ecological patterns in a prescribed but arbitrary area A_0 would flow from the input of just three numbers: S_0 , N_0, and M_0.

Elegant and powerful as this agenda would be if it works, it may be possible to go further along the path toward a unified theory of ecology by further exploiting MaxEnt. Additional progress might stem from the recent work of Dewar [11], who showed that the Jaynes MaxEnt framework allows a rigorous derivation of another entropy-related principle: Maximum Entropy Production (MEP).

The principle of MEP asserts that dissipative processes tend to maximize the amount of entropy production; Dewar's proof is somewhat analogous to the proof in classical equilibrium thermodynamics that the macrostate with maximum entropy is associated with the most microstates and thus is most probable. For non-equilibrium dissipative systems, Dewar showed that the transition to the MEP state is most probable because there are more paths associated with that transition than with any other. The result derives from maximization of "path entropy" $S = -\sum_{\Gamma} p_{\Gamma} \log p_{\Gamma}$ where p_{Γ} is the probability of the path Γ through the microstate phase space. Application of MEP to the dissipative atmospheric and oceanic fluid motion that redistributes heat on Earth was shown (well before Dewar's proof) to lead to a simple and remarkably accurate calculation, with no adjustable parameters, of the latitudinal gradient of Earth's surface temperature [40–42]. This is an accomplishment yet to be matched by simulation of the Navier Stokes equations.

At three levels of complexity in ecology, MEP might be applicable in ecology. At the level of the single organism, there is considerable interest in the relationship between the basal metabolism rate of organisms (or respiration rate in plants) and the mass of the organism. For adult life stages, a power law relationship of the form metabolism $\sim$ mass$^{3/4}$ has been observed across a wide range of the masses, in species ranging from whales to shrews, and from phytoplankton to redwood trees. At the same time, there is evidence that in younger, growing, individuals, a linear relationship may be more prevalent. The 3/4−power behavior has been derived from several different sets of assumptions about the properties of the nutrient-transport system within an individual organism [2,15,46]. I suggest that MEP might provide a framework for estimating how food energy is divided between growth and metabolism, and thus could explain the dependence of the metabolism-mass relationship on growth stage.

At the whole ecosystem level, a collection of organisms spanning many species is usefully characterized by a food web, describing the flows of energy from one species to another due to trophic interactions (A eats B). Many regularities in the topology of food webs have been documented across a variety of habitat types (For a recent review, see [12]). Because predation is an entropy-creating act, it should be possible to determine the web topology that maximizes entropy production and thus derive predictions for web regularities.

A third potential application is at the regional to global spatial scale, at which energy and material exchanges between the land surface and the atmosphere are in part influenced by the properties of terrestrial vegetation. Preliminary work examining the consequences of MEP to this system has been initiated by Kleidon [34], but many questions remain. Further work along these lines could help us understand, at the largest spatial scales, the self-consistent solution "chosen by nature" to the coupled system in which terrestrial vegetation influences carbon cycling, surface albedo of vegetation, and transpiration rates, all of which influence the climate system that in turn influences the characteristics of terrestrial vegetation.

If MEP can illuminate food webs, organismal energetics, and the coupled climate-vegetation system, and MaxEnt allows us to deduce scaling laws and abundance distributions from pre-specified values of S_0, N_0, and M_0, a major advance would be made in ecological theory. Even more ambitiously, however, it is possible that from MEP, the partitioning of available energy and other resources into distinct species (S_0) and individuals (N_0), as well as the total biomass (M_0), can be deduced, thus closing the circle and providing a means of deducing all of the above from first principles. Figure 6 shows schematically the outlines of a possible future comprehensive and unified theory of ecology.

To conclude, I wish to emphasize one fundamental simplification that underlies all of the models and ideas discussed here: the assumption of the principle of constituent equivalence (PCE). By PCE I mean the a-priori indistinguishability (after possible re-scaling in some cases) of the living constituents of the theory. Indistinguishability can apply to the individuals within a species or, more strongly, to the individuals across species. In all four spatial models discussed in Section III, PCE is assumed to describe the spatial preferences of individuals within species; species differences arise simply because of different n_0 values across species. More generally, PCE can refer to the a-priori equivalence of per-capita demographic parameters of all individuals of all species, as in the neutral theory of Hubbell [28], who refers to the assumption as both "neutrality" and "the hypothesis of ecological equivalence" [29]. Or it can refer to the mass-scaled energy requirements of all individuals within broad taxonomic groups, as in the Metabolic Theory.

PCE is a profoundly simplifying assumption and one that understandably attracts considerable criticism. It is likely that if all individual organisms were as alike as are all electrons, few people would be attracted to ecology as a science, and certainly not as a hobby. Moreover, evolution could not operate and biology would then be a minor subfield of physics. Yet at the same time, some of the most promising theoretical developments in ecology stem from exactly this assumption. In the coming years, attempts to resolve this seeming paradox will provide an exciting challenge to ecologists searching for a unified understanding of Darwin's "entangled bank":

"It is interesting to contemplate an entangled bank, clothed with many plants of many kinds, with birds singing on the bushes, with various insects flitting about, and with worms crawling through the damp earth, and to re-

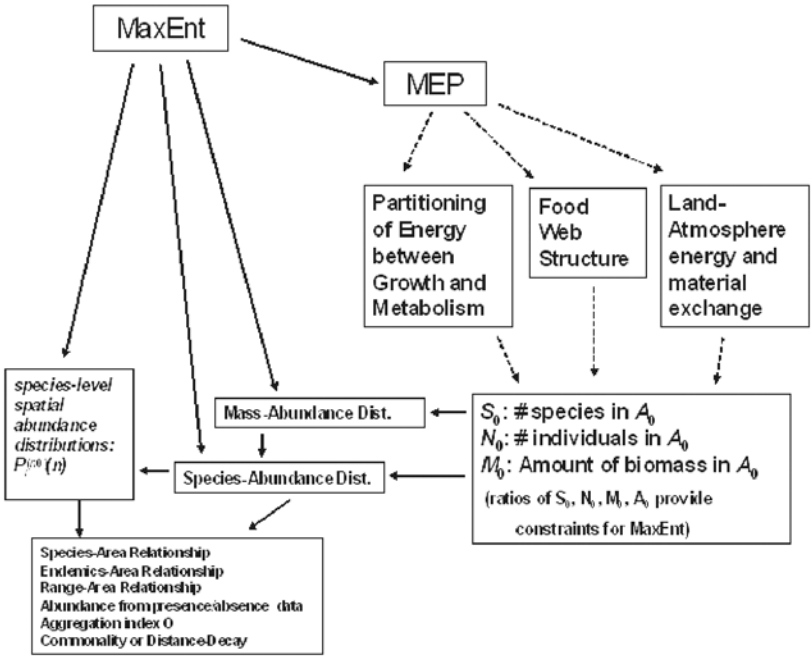


Fig. 6. Possible roadmap of a unified theory of ecology; solid arrows correspond to completed calculations ([24] and also Sections III and VII of this Chapter), while the dashed arrows will be the subject of future exploration.

flect that these elaborately constructed forms, so different in each other, and dependent on each other in so complex a manner, have all been produced by laws acting around us."

Charles Darwin

References

1. Banavar, J., Green, J., Harte, J., Maritan, A.: Finite Size Scaling in Ecology. Physical Review Letters, **83**(20), 4212–4214 (1999)
2. Banavar, J., Maritan, A., Rinaldo, A.: Size and Form in Efficient Transportation Networks. Nature, **399**, 130–132 (1999)
3. Coleman, B.: On random placement and species-area relations. Journal of Mathematical Biosciences, **54**: 191–215 (1981)
4. Conlisk, E., Conlisk, J., Harte, J.: The Impossibility of Estimating a Negative Binomial Clustering Parameter from Presence-Absence Data. American Naturalist, In Press (2006)
5. Conlisk, E., Bloxham, M., Conlisk, J., Enquist, B., Harte, J.: A Class of Models of Spatial Distributions. Ecological Monographs, In Press (2006)

6. Condit, R., Hubbell, S. P., LaFrankie, J. V. Sukumar, R. Manokaran, N., Foster, R., Ashton, P.S.: Species-area and species-individual relationships for tropical trees: a comparison of three 50-ha plots. Journal of Ecology, **84**, 549–562 (1996)
7. Condit, R., Hubbell, S.P., Foster, R.B.: Changes in tree species abundance in a Neotropical forest: impact of climate change. Journal of Tropical Ecology, **12**, 231–256 (1996)
8. Condit, R.: Tropical Forest Census Plots. Springer-Verlag and R. G. Landes Company, Berlin, Germany, and Georgetown, Texas (1998)
9. Condit, R., Ashton, P. S., Baker, P., Bunyavejchewin, S. Gunatilleke, S., Gunatilleke, N., Hubbell, S. P., Foster, R. B., Itoh, A., LaFrankie, J. V., Seng, L. H., Losos, E., Manokaran, N., Sukumar, R., Yamakura, R.: Spatial patterns in the distribution of tropical tree species. Science, **288**, 1414–1418 (2000)
10. Damuth, J.: Population Density and Body Size in Mammals. Nature, **290**, 699–700 (1981)
11. Dewar, R.: Information Theory Explanation of the Fluctuation Theorem, Maximum Entropy Production and Self-Organized Criticality in Non-Equilibrium Stationary States. Journal of Physics A, Math. Gen., **36**, 631–641 (2003)
12. Dunne, J.: The Network Structure of Food Webs. In: Pascual, M. and Dunne, J. (eds) Linking Structure to Dynamics in Food Webs. Oxford University press, Oxford, UK, 27–86 (2006)
13. Enquist, B. J., West, G. B., Charnov, E. L., Brown, J. H.: Allometric scaling of production and life history variation in vascular plants. Nature, **401**, 907–911 (1999)
14. Enquist, B., Niklas, K.: Invariant Scaling Relations across Tree-Dominated Communities. Nature, **410**, 655–660 (2001)
15. Ettienne, R., Apol, M., Olff, H.: Demystifying the West, Brown, & Enquist Model of the Allometry of Metabolism. Functional Ecology, **20**, 394–399 (2006)
16. Gaston, K., Blackburn, T.: Pattern and Process in Macroecology. Blackwell Scientific, Oxford (2000)
17. Green, J., Harte, J., Ostling, A.: Species Richness, Endemism, and Abundance Patterns: Tests of Two Fractal Models in a Serpentine Grassland. Ecology Letters, **6**, 919–928 (2003)
18. Harte, J., McCarthy, S., Taylor, K., Kinzig, A., Fischer, M.: Estimating Species-Area Relationships from Plot to Landscape Scale using Species Spatial-Turnover data. Oikos, **86**, 45–54 (1999)
19. Harte, J., Kinzig, A., Green, J.: Self-Similarity in the Distribution and Abundance of Species. Science, **284**, 334–336 (1999)
20. Harte, J.: Scaling and Self-Similarity in Species Distributions: Implications for Endemism, Spatial Turnover, Abundance, and Range. In: Proceedings of the 1997 Workshop on Scaling in Biology. Oxford University Press, Oxford (2000)
21. Harte, J., Blackburn, T., Ostling, A.: Self Similarity and the Relationship between Abundance and Range Size. American Naturalist, **157**, 374–386 (2001)
22. Harte, J., Conlisk, E., Ostling, A., Green J., Smith, A.: A theory of Spatial Structure in Ecological Communities at Multiple Spatial Scales. Ecological Monographs, **75**(2), 179–197 (2005)
23. Harte, J.: Toward a Mechanistic Basis for a Unified Theory of Spatial Structure in Ecological Communities at Multiple Spatial Scales. In: Storch, D. and Marquet, P. (eds) Proceedings of the Scaling Biodiversity Workshop. Prague, 2004, SFI/CTS (2006)

24. Harte, J., Zillio, T.: Biodiversity Scaling Patterns are Governed by Maximum Entropy. Science, in review (2006)
25. He, F., Legendre, P.: Species diversity patterns derived from species area models. Ecology, **83**, 1185–1198 (2002)
26. He, F., Gaston, K.: Occupancy, spatial variance, and the abundance of species. The American Naturalist, **162**, 366–375 (2003)
27. Hubbell, S.P., Foster, R.B.: Diversity of Canopy Trees in a Neotropical Forest and Implications for Conservation. In: Sutton, S.L., Whitmore, T.C., and Chadwick, A.C. (eds) Tropical Rain Forest: Ecology and Management. Blackwell Scientific Publications, Oxford, 25–41 (1983)
28. Hubbell, S.: The Unified Neutral Theory of Biodiversity and Biogeography. Monographs in Population Biology **32**. Princeton University Press, Princeton NJ (2001)
29. Hubbell, S.: Neutral Theory and the Evolution of Ecological Equivalence. Ecology, **87**(6), 1387–1398 (2006)
30. Jaynes, E.: Information theory and Statistical Mechanics. Physical Review, **106**, 620–630 (1957)
31. Jaynes, E.: Information Theory and Statistical Mechanics. In: Ford, K. (ed) Brandeis Summer Institute 1962, Statistical Physics. Benjamin, New York, NY, 181–218 (1963)
32. Jaynes, E.: Where do We Stand on Maximum Entropy. In: Levine, R. and Tribus, M. (eds) The Maximum Entropy Principle. MIT Press, Cambridge, MA pp. 15-118 (1979)
33. Kinzig, A., Harte, J.: Selection of Microorganisms in a Spatially Heterogeneous Environment: Implications for Plant Access to Nitrogen. Journal of Ecology, **86**, 841–853 (1998)
34. Kleidon, A.: Beyond Gaia: Thermodynamics of Life and Earth System Functioning. Climatic Change (2004)
35. Kunin, W.: Extrapolating Species Abundance across Spatial Scales. Science, **281**, 1513–1515 (1998)
36. MacLean, R., Gudelj, I.: Resource Competition and Social Conflict in Experimental Populations of Yeast. Nature, **441**, 498–501 (2006)
37. Maddux, R.: Self Similarity and the Species Area Relationship. American Naturalist, **163**, 616–626 (2004)
38. Ostling, A., Harte, J., Green, J.: Self Similarity and Clustering in the Spatial Distribution of Species. Science, **290**: 671 (Technical Comment: www.sciencemag.org/cgi/content/full/290/5492/671a) (2000)
39. Ostling, A., Harte, J., Green, J., Kinzig, A.: Self Similarity, the Power Law Form of the Species-Area Relationship, and a Probability Rule: A Reply to Maddux. American Naturalist, **163**, 627–633 (2004)
40. Paltridge, G.: Global Dynamics and Climate: A System of Minimum Entropy Exchange. Quarterly Journal of the Royal Meteorological Society, **101**, 475–484 (1975)
41. Paltridge, G.: The Steady State Format of Global Climate. Quarterly Journal of the Royal Meteorological Society, **104**, 927–945 (1978)
42. Paltridge, G.: Climate and Thermodynamic Systems of Maximum Dissipation. Nature, **279**, 630–631 (1979)
43. Plotkin, J., Potts, M., Leslie, N. Manokaran, N., LaFrankie, J., Ashton, P.: Species-area curves, spatial aggregation, and habitat specialization in tropical forests. Journal of Theoretical Biology, **207**, 81–99 (2000)

44. Rosenzweig, M.: Species Diversity in Space and Time. Cambridge University Press, Cambridge, UK (1995)
45. Shannon, C.: A Mathematical Theory of Communication. Bell Systems Technology Journal, **27**, 379–423; 623–656 (1948)
46. West, G., Brown, J., Enquist, B.: A General Model for the Origin of Allometric Scaling Laws in Biology. Nature, **413**, 628–631 (1997)

Protein Structure and Its Folding Rate

Alexei V. Finkelstein, Dmitry N. Ivankov, Sergiy O. Garbuzynskiy, and Oxana V. Galzitskaya

Laboratory of Protein Physics, Institute of Protein Research, Russian Academy of Sciences, 4 Institutskaya str., Pushchino, Moscow Region, 142290, Russian Federation alexey@finkelstein.ru

Summary. In the first part of this paper we overview protein structures, their spontaneous formation ("folding") and thermodynamic and kinetic aspects of this phenomenon. It is stressed that universal features of folding are observed near the point of thermodynamic equilibrium between the native and denatured states of the protein. Here the "two-state" ("denatured state" $\leftrightarrow$ "native state") transition proceeds without accumulation of metastable intermediates, and only the transition state, i.e., the most unstable state in the folding pathway, is outlined by its essential influence on the folding/unfolding kinetics. In the second part of the paper, a theory of protein folding rates and related phenomena is presented. First, it is shown that the protein size determines the range of protein's folding rates in the vicinity of the point of thermodynamic equilibrium between the native and denatured states of the protein. Then we present methods for calculating folding and unfolding rates of globular proteins from their sizes, stabilities and either 3D structures or amino acid sequences. And, at last, we show that the same theory outlines the location of the protein folding nucleus (i.e., the structured part of transition state) in a reasonable concordance with experimental data.

Key words: Protein folding nucleus, polypeptide backbone, Anfinsen's experiments, Levinthal paradox, folding pathways, "all-or-none" transition.

1 Introduction

Protein physics is grounded on three fundamental experimental facts: (i) many proteins have well defined three-dimensional structures [1,3]; (ii) many protein chains are capable of self-organization, i.e., they form their native structures spontaneously in appropriate environment [4, 5], and (iii) the native state of many proteins is separated from the unfolded state of the chain by an "all-or-none" phase transition [6]. The latter ensures robustness of protein action: as a bulb, the protein either has a correct structure and works correctly, or does not work at all.

The aim of this paper is to overview modern understanding of physical principles of arrangement and self-organization of protein structures.

2 Protein Structure

Protein is a heteropolymer (Figure 1) with regular backbone and unique (for each protein) sequence of amino acid residues, having side groups of 20 kinds. In an "operating" protein its chain is folded in a strictly specified ("*native*") structure, which exists under normal biological conditions but decays under action of various denaturants, such as temperature, acid, some chemicals like urea, etc. Some (∼10% of) protein chains, however, have no fixed structure by themselves, but obtain it by interacting with other molecules [7].

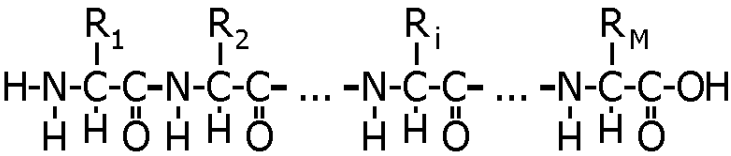


Fig. 1. Chemical structure of protein chain: regular polypeptide backbone $(\text{NH-CH-CO})_M$ with various side groups (R_1, R_2, ... R_M), whose sequence in the chain is unique for each protein (as established by Sanger in 1950s). NH groups of the backbone can form hydrogen bonds with CO groups of the other amino acid residues.

In the late 1950s, Perutz and Kendrew solved the first structures of protein crystals and demonstrated highly intricate protein spatial structures [1, 2]. Later, the identity (to small fluctuations) of structures of various proteins in a crystal and in solution was demonstrated by NMR spectroscopy [3].

Before considering protein physics, it is not out of place to remind to a reader that proteins "live" under various environmental conditions leaving an obvious mark on their structures [8].

Roughly, according to the "environmental conditions" and general structure, proteins can be divided into three large groups (Figure 2).

1. Fibrous proteins form vast aggregates; their structure is usually highly hydrogen-bonded, highly regular and maintained mainly by interactions between different polypeptide chains.
2. Membrane proteins, although partly project into water, mostly "live" in water-lacking membranes. Their intramembrane portions are highly regular (like fibrous proteins) and highly hydrogen-bonded, but restricted in size by the membrane thickness ($30 - 40$ $\mathring{A}$).
3. Water-soluble (living in water) globular proteins are less regular (especially the small ones). Their structure is maintained by interactions of the

chain with itself and sometimes with various other molecules (co-factors). Typical globule is formed by a chain of 50 – 200 amino acid residues; typical size of the globule is $30 - 50\ \mathring{A}$.

The reason for different construction of these proteins is that proteins within membranes or aggregates have poor access to water, and the less water is around, the more valuable the hydrogen bonds are. Such bonds (between regularly positioned N-H and O=C groups ob protein backbone) reinforce the regular secondary structures of protein chain, and thus, the less water is around, the more regular stable protein structures ought to be.

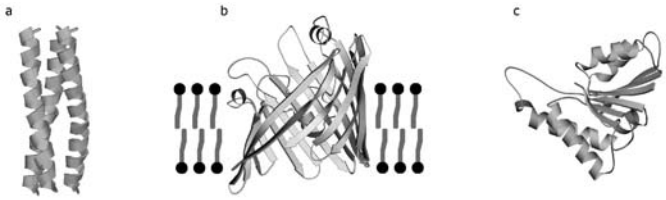


Fig. 2. Typical shapes of proteins. Secondary structures are outlined schematically: regular α-helices are presented as helical ribbons, regular β-strands (forming β-sheets) as arrows and irregular chain regions as threads. Side chains are not shown for the sake of simplicity. (a) An aggregate typical of fibrous proteins (a fragment); (b) membrane protein (lipids, with polar heads and non-polar tails, are shown only schematically); (c) water-soluble globular protein. Adapted from [8].

The above classification is certainly extremely rough. Some proteins may comprise a fibrous "tail" and a globular "head" (like myosin, for example), and so on.

To date we know millions protein sequences and tens of thousands three-dimensional (3D) protein structures. Most (∼95%) of known 3D structures belong to water-soluble globular proteins: they are easily isolated and then studied by X-ray in crystals and by NMR in solution. As to membrane and fibrous proteins, their solved 3D structures are relatively few; they and comprise ∼5% and ∼1% of known protein 3D structures, respectively [9].

For similar experimental reasons, protein folding is also better studied for water-soluble globular proteins. That's why, when speaking about "protein structure" and "protein structure formation" one often actually means regularities shown for water-soluble globular proteins only.

Following this tradition, we will also concentrate on water-soluble globular proteins.

Moreover, we will concentrate mostly on relatively small "single-domain" proteins that form one compact protein globule (Figures 2c, 3). A "single-domain" structure is typical of small water-soluble globular proteins. Large proteins usually consist of two-three or even more domains [8–10].

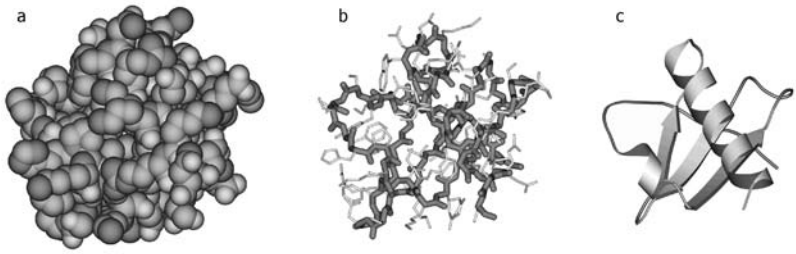


Fig. 3. The structure of a small globular protein shown as a space-filling atomic model (a), as a wire model with the chemically regular backbone outlined by dark line and various side chains by light lines (b), and as a ribbon diagram that outlines the backbone fold and the secondary structures (c). The secondary structures are presented by two α-helices (the turns within each helix being fasten together by backbone-backbone hydrogen bonds) and three extended β-strands (shown as arrows), which, being fasten together by backbone-backbone hydrogen bonds between them, form a single β-sheet. This and the other drawings of protein structures are done using their Protein Data Bank [11] coordinates by the MOLSCRIPT [12] program. The drawing is adapted from [8].

The regular hydrogen-bonded secondary structures (α-helices and β-sheets) form the framework inside of the globule, while the elements with "free" (i.e., saturated with water) hydrogen bond donors and acceptors (irregular loops and tails, edges of β-sheets, ends of α-helices) remain at the surface and contact to water.

Although 3D structures and amino acid sequences of all proteins (membrane, fibrous and globular) are unique, statistical analysis shows that sequences of globular proteins look almost "random" [13], – in the sense that they contain neither blocks of non-polar and polar amino acids, typical of membrane proteins, nor repeats, typical of fibrous proteins (Figure 4).

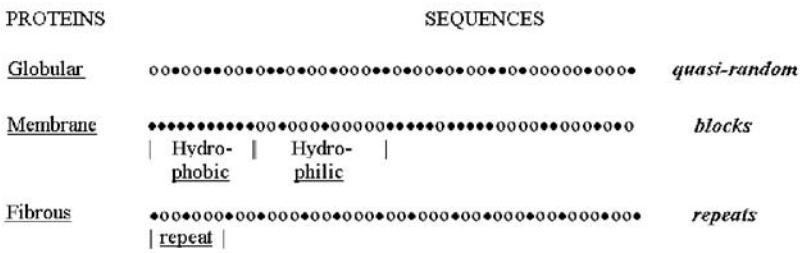


Fig. 4. Typical alteration of hydrophobic (•) and polar (○) residues in sequences of water-soluble globular proteins, of membrane proteins and of fibrous proteins.

3 Phase Transitions in Protein Molecules

3.1 Reversible Denaturation of Protein Structures

Figures 2, 3 show "solid" protein structures. However, depending on ambient conditions, the most stable state of a protein molecule may be not solid but molten or even extremely swollen, "unfolded": then the protein "denatures" and loses its native, "working" 3D structure.

Usually protein denaturation is studied *in vitro* (in a test tube); then it is caused by an abnormal, "non-physiological" temperature or by a excess of a denaturant (like H^+, OH^- or urea). However, decay of the "solid" protein structure (and its subsequent refolding) can occur also in a living cell, e.g., during trans-membrane transport of proteins.

Denaturation and renaturation of the water-soluble globular proteins is best studied, and we will speak about them only.

It is well established that denaturation of small proteins is a cooperative transition with a simultaneous abrupt change of various characteristics of the molecule. The narrow transition regions suggest that the transition embraces many amino acid residues.

Moreover, denaturation of a single-domain protein occurs as an "all-or-none" transition [6, 14]. The latter means that the transition embraces the domain as a whole, and that only the initial (native) and the final (denatured) states amount to visible quantities, while "semi-denatured" states are unstable and practically absent . (Though, of course, they do exist to a very small quantity, since a native molecule cannot come to its denatured state without passing the intermediate forms, and their presence has a crucial effect on kinetics of the transition, which we will discuss in the next chapter.)

The "all-or-none" transition is a microscopic analog of the first-order phase transitions in macroscopic systems (e.g., crystal melting). However, unlike the true phase transitions, the "all-or-none" transitions in proteins have a non-zero temperature width, since this transition embraces a microscopic system. It should be specified that the "all-or-none" denaturation actually concerns small proteins and to separate domains of large proteins, while denaturation of a large protein is a sum of denaturations of its domains [15, 16].

To prove that melting is an "all-or-none" transition, one has to compare (1) "effective latent heat" of transition calculated from its width (i.e., the amount of heat consumed by one independent melting unit) with (2) "calorimetric heat" of this transition, i.e., the amount of heat consumed by one melting protein molecule [6].

Denaturation of a single-domain protein is usually reversible. This is known since Anfinsen's experiments of 1960's [4]: a protein can renature (when the ambient conditions come back to "physiological" ones) if it is not too large and has not been subjected to substantial chemical modifications after the *in vivo* folding (and if the protein solution is sufficiently diluted to avoid aggregation). In this case, a "mild" (without chemical decay) destruction of

the protein's native structure (by temperature, denaturant, etc.) is reversible, and the native structure spontaneously restores after environmental conditions have become normal.

The reversibility of protein denaturation is very important: it shows that the entire information necessary to build up the protein 3D structure is contained in its amino acid sequence and that the protein structure itself (to be more exact, the structure of a not too modified and not too large protein) is thermodynamically stable. This allows one to use thermodynamics to study and describe de- and renaturation transitions.

Heat denaturation of proteins is usually accompanied by a large heat effect, ∼1 kcal per mole of amino acid residues; however, the native state of a protein is more stable than its unfolded state by not more than a few kcal/mol even under physiological conditions, where the native state is the most stable.

3.2 How Do Denatured Proteins Look Like?

Numerous *thermodynamic* experiments have shown that there are no cooperative transitions within the denatured state of a protein molecule. Therefore, it was initially assumed that the denatured protein is always a very loose random coil (as it is in a "very good" solvent like concentrated solution of urea [17]).

However, *structural* studies of denatured proteins reported on some large-scale rearrangements within the denatured state, that is, on some stable "intermediates" between the completely unfolded coil and the native state of proteins (Figure 5).

Apart from biochemical activity, only the ordering of side chains *always* abruptly changes during denaturation; this is observed by, e.g., NMR. On the other hand, the ordering of protein main chain (i.e., its secondary structure) and the density of protein molecule can be virtually preserved in some cases and strongly changed in the others, depending on denaturation conditions [19–21].

Protein chain looks like a random coil in a "very good" solvent like concentrated solution of urea or GuHCl; then its hydrodynamic volume is proportional to the chain length in the power 3/2 or close to 3/2. However, in poor solvents (e.g., water) the volume of denatured globule is often only a little larger than the volume of the native protein. Thus, the above studies have revealed a universal, or rather, nearly universal intermediate of protein unfolding and folding, which is now known as the "*molten globule*" [21] (Figure 6).

On the face of it, the molten globule properties are contradictory. Its secondary structure is usually developed nearly as that of the native protein. On the other hand, the molted globule (like the completely unfolded protein) has nearly no ordering of side chains, which is so typical of the native protein. However, some portion of the side chain native-like contacts evidently remains in the molten globule; NMR shows that this concerns aromatic rather than to

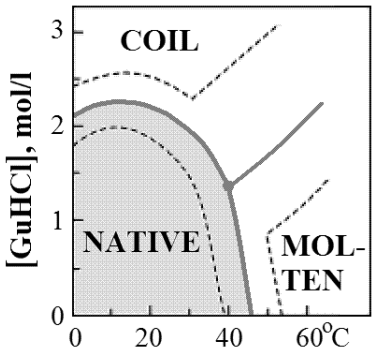


Fig. 5. Phase diagram of conformational states (at pH 1.7) of a single-domain protein lysozyme at various temperatures in solution of guanidine hydrochloride (GuHCl) denaturant of various concentrations: the solid NATIVE state, the completely unfolded COIL, and a more compact temperature-denatured state (MOLTEN). The solid line corresponds to the mid-transition, the dashed lines outline the transition zones (from the proportion $\approx 9:1$ in favor of one state to $\approx 1:9$ in favor of another). One can see that the COIL – MOLTEN transition is much wider than the others. Adapted from [18].

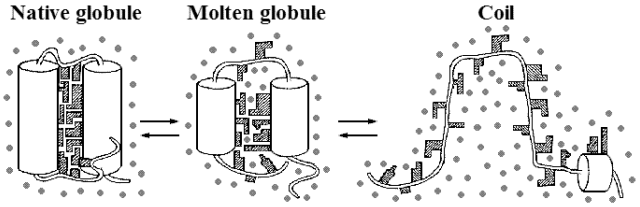


Fig. 6. Schematic model of the molten protein globule [21–24] in comparison to the native protein and the coil. For simplicity, the protein is shown to consist of only two helices and irregular loops. The backbone is covered with numerous side chains (dashed). Reinforced by H-bonds, the secondary structures are stable until the globule is "dissolved" by a solvent. Usually, waters are unable to do this without a strong denaturant. In the molten globule, side chains lose their close packing, but acquire the free movements, i.e., they lose energy but gain entropy. The waters (•) come into pores of the molten globule (that appear when the close packing is lost), but, until the denaturant is not too strong, cause no further decay of the globule; a stronger denaturant converts the globule into the coil.

aliphatic side chains. The molten globule is a little less compact than the native protein (as the hydrodynamic volume measurements report), and its core is virtually as compact as that of the native protein (as the "middle-angle" X-ray scattering reports); on the other hand, a rather high rate of hydrogen exchange shows that at least separate solvent molecules easily penetrate into the globule [21, 22]. To some extent, these contradictions may be conciliated by assumption that the molten globule has a relatively dense, native-like core and more loose loops. However, the assumption that the molten globule has the completely native core and the completely unfolded loops is in a clear contradiction with experiment, and specifically with the absence of the ordered environment of aromatic side chains in the molten globule observed by the near UV CD and by the NMR spectroscopy.

For very many (though not for all) proteins, the "molten globule" arises at a moderate denaturing impact upon the native protein, and decays (turns into a random coil) only under the impact of a concentrated denaturant. The molten globule-like state often occurs after temperature denaturation ("melting"), and this melting has been always observed to be the "all-or-none" transition. The molten globule usually does not undergo cooperative melting with increasing temperature (see Figure 5), but its unfolding caused by a strong denaturant looks like a cooperative S-shaped transition.

However, some proteins (especially small ones) unfold directly into a coil without the mediating molten globule state; and many other proteins are converted into the molten globule by some denaturing agents (e.g., by temperature or by acid), while other agents (e.g., urea) directly convert them into a coil. In all these cases, though, the decay of the native state is the "all-or-none", i.e., the first order phase transition [21].

3.3 Why Denaturation of a Globular Protein Is the First-order Phase Transition?

To understand protein denaturation, one has to explain why there exist two equally stable phases of the protein chain (which does not take place in a usual polymers [25, 26]), and why these two phases are separated by a free energy barrier (which is necessary for the "all-or-none" transition) [23, 24]. That is, one has to explain why the protein globule cannot decay by gradual swelling, like usual polymer globules do.

In doing so, one has to take into account the main peculiarities of proteins (those which differ them from usual polymers): that each protein has one chain fold distinguished by its peculiar stability; that flexible side groups are attached to a more rigid backbone of the protein chain; and that the native protein is packed as tightly as a molecular crystal.

Side groups of protein chains are capable of rotational isomerization, i.e., jumps from one allowed conformation to another. Each jump requires some vacant volume around the jumping side chain; but the native protein fold is distinguished by a tight packing (contributing to a peculiar stability of this

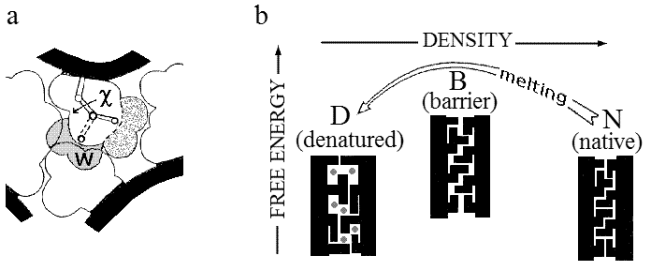


Fig. 7. (a) A sketch of the side chain packing. Only a small peace of the globule is shown. The dashed region W corresponds to an alternative rotamer of the side chain (χ being its rotational angle); this rotamer is forbidden by close packing of the native globule. Appearance of this new rotamer W requires additional vacant volume of at least 30 $\mathring{A}^3$ (i.e., the volume of a methyl group), or $\approx 1/5$ of the average amino acid volume. Nearly the same volume is required for H_2O penetration into the core. (b) Origin of the free-energy barrier between the native and any denatured state of the protein. The "barrier" state B arises at a small expansion of the native, closely packed state N. The pores formed in the state B *already* cause a great increase in the van der Waals energy, but *yet* allow neither side chain liberation nor penetration of the solvent (ffl) inside the protein; both these effects require a further swelling of the globule, which leads to the denatured state D. Adapted from [23].

fold) that excludes these jumps. Besides, the side chains are bound to the rigid backbone, which is especially rigid inside the globule where the chain forms α- and β-structures that are necessary to involve the backbone into the dense globule (Figures 3, 7). These structures are stable at least until the solvent penetrates into the globule (that requires approximately the same free volume as the jumps of side groups).

One can consider two scenarios of protein denaturation: a uniform and a non-uniform swelling of the native globule. The latter is accompanied with creation a boundary between the more and less expanded parts of the globule, and the increased free energy of this boundary (i.e., the surface tension) creates a free energy barrier separating the native and the denatured phases. Thus, the "non-uniform" scenario is incompatible with a gradual swelling; as we will see later, this scenario is of crucial importance for kinetics of protein unfolding and folding.

A uniform swelling can be gradual, in principle. However, a uniform expansion also cannot avoid a necessity to overcome a free energy barrier. When a globule expands, each of the rigid secondary structures has to move as a whole (at least at the beginning of the globule's uniform expansion), with the entire forest of side chains attached. Therefore, uniform expansion of the closely packed globule through movement of α- and β-structures creates approximately equal free spaces near each side chain, and these spaces are either insufficient for isomerization of each of the side chains (when the globule's expansion is too small), or sufficient for isomerization of many of them at once.

This means that liberation of the side chains (as well as solvent penetration) does not occur gradually but only at once, when the globule's expansion crosses some threshold, the "barrier" (Figure 7). These two events can make a less dense state of the protein chain as stable as its native state, but only after the density barrier has been passed.

Thus, a large ("post-barrier") expansion liberates the rotational isomerization and leads (at sufficiently high temperature) to decreasing free energy.

On the contrary, a small ("pre-barrier") expansion of the native globule *always* increases its free energy: it *already* increases the globule's energy, but *does not* increase its entropy, because the small expansion does not *yet* liberate the rotational isomerization of the side chains.

As a result, a small increase in protein's energy is not accompanied by increase in its entropy. This leads to an "energy gap" between the native fold and its "misfolded" competitors (see Figure 16 below). Just because of this gap protein denaturation does not occur gradually: it occurs as a jump over the free energy barrier, in accordance with the "all-or-none" principle. It has been shown that the "all-or-none" folding phase transition requires a selected amino acid sequence that only provides a large energy gap between the native and all the other, structurally dissimilar conformations [27–29].

In the other words, the protein tolerates, without a change, modification of ambient conditions up to a certain limit, and then melts altogether, like a solid body. This resistance and hardness of protein, in turn, provides reliability of its biological functioning, and therefore must be maintained by biological evolution [8].

3.4 Protein Folding in Vivo

In a living cell, protein is synthesized by a ribosome that makes a protein chain (whose sequence is encoded by mRNA) residue by residue, from its N- to C-end, and not quite uniformly: there are temporary rests of the synthesis at the "rare" codons (they correspond to tRNAs which are rare in the cell, and these codons are rare in the cell's mRNAs, too). It is assumed that the pauses may correspond to the boundaries of structural domains that can help a quiet maturation of the domain structures. The biosynthesis takes about a minute and yielding of a "ready" folded protein lasts as long: the experiment does not see any difference [30, 31].

Some enzymes, like prolyl-peptide- or disulfide-isomerases accelerate *in vivo* folding. They catalyze slow, if unaided, *trans* $\leftrightarrow$ *cis* conversions of prolines and formation (and decay) of S-S bonds.

Protein chain folds under the protection of special proteins, chaperons. These are the cell's trouble-shooters that fight the aggregation, since, in a cell, folding takes place in a highly crowded molecular environment. There is no reason to assume, though, that anything other than the amino acid sequence determines protein conformation in the cell [32, 33].

It looks as though the biosynthetic machinery (ribosomes + chaperons + ...), besides of chemical synthesis of the protein chain, serve only as a kind of incubator, which does not determine the protein structure (at least if the protein is not very large and does not consist of many domains) but rather provides "hothouse" conditions for its maturation, - just like a usual incubator helps a nestling to develop but does not determine what will be developed, a chicken or a duckling.

Unfortunately, it is difficult to follow the *in vivo* folding of a nascent protein chain against the background of the huge ribosome. It is known, though, that the first synthesized domains of multi-domain proteins are able to fold before the biosynthesis of the whole chain is completed [30,33]; but there is virtually no date of this kind for single-domain proteins.

Therefore, most of experiments on protein folding are done *in vitro*.

As above, we will consider mostly the single-domain water-soluble globular proteins which are studied much better than the others.

3.5 Protein Folding in Vitro

In about 1960, a remarkable discovery was done: it was shown that a globular protein is capable of spontaneous folding *in vitro* [4]. If protein chain has not been heavily chemically modified after the initial (*in vivo*) folding, then the protein gently (without chain damaging) unfolded by temperature, denaturant, etc., spontaneously "renatures", i.e., restores its activity and structure after solvent "normalization". True, the effective renaturation requires a careful selection of experimental conditions; otherwise, aggregation (including famous amyloid formation [34,35]) can prevent the protein from folding.

Furthermore, it was demonstrated [5] that the protein chain synthesized chemically, without any cell or ribosome, and placed in the proper ambient conditions, folds into a biologically active protein.

The phenomenon of spontaneous folding of protein native structures allows us to detach, at least to a first approximation, the study of protein folding physics from the study of protein biosynthesis.

Protein folding *in vitro* is the most simple (and therefore, the most interesting for a physicist) case of pure *self*-organization: here nothing "biological" (but for the sequence!) helps protein chain to fold.

The Levinthal Paradox

The ability of proteins (and RNA) to fold spontaneously immediately raised a fundamental problem that has come to be known as the Levinthal paradox [36]. It reads as follows.

On the one hand, the same native state is achieved by various folding processes: *in vivo* on the ribosome, *in vivo* after translocation through the membrane, *in vitro* after denaturation with various agents ... The existence of the spontaneous and correct folding of chemically synthesized protein chains

suggests that the native state is thermodynamically the most stable state under the "biological" conditions.

On the other hand, a chain has zillions of possible conformations (at least 2^{100} for a 100-residue chain, since at least two conformations are possible for each residue), and the protein can "feel" the right stable structure only if it is achieved exactly, since even a 1? deviation can strongly increase the chain energy in the closely packed globule. Thus, the chain needs at least $\sim 2^{100}$ picoseconds, or $\sim 10^{10}$ years to sample all possible conformations in its search for the most stable fold.

Then, a question arises: how can the chain find its most stable structure within a "biological" time (minutes) at all?

The paradox is that, on the one hand, the achievement of the same (native) state by a variety of processes is (in physics) a clear-cut evidence of its stability. On the other hand, Levinthal's estimate shows that the protein simply does not have enough time to prove that the native structure is the most stable among all possible structures!

In order to solve this paradox, Levinthal suggested existence of specific folding pathways, and hypnotized that the native fold is simply an end of the protein-specific pathway rather than the most stable fold of its chain. Should this pathway be narrow, only a small part of the conformational space would be sampled, and the paradox would be avoided. In other words, Levinthal suggested that the native protein structure is under kinetic rather than under thermodynamic control, i.e., that it corresponds not to the global but rather to the easily accessible free energy minimum.

Folding Pathways and Folding Intermediates

The question as to whether the protein structure is under kinetic or thermodynamic control is not a purely speculative question. It is raised again and again when one faces practical problems of protein physics and engineering. For example: when trying to predict protein structure from its sequence, what have we to look for? The most stable or the most rapidly folding structure? When designing a *de novo* protein, what have we to do? To maximize stability of the desired fold or to create a rapid pathway to this fold?

A discussion on protein folding mechanisms started immediately after discovery of the spontaneous folding. It seems that the first proposed hypothesis was that by Phillips who suggested that the folding nucleus is formed by the N-end of the nascent protein chain, and the remaining part of the chain wraps around it [37]. This appealing hypothesis is present in some works up to now. However, it has been refuted experimentally, as far as single-domain proteins are concerned. The elegant works by Goldenberg & Creighton have shown that the N-terminus has no special role in the *in vitro* folding: it is possible to glue the ends of the chain of a small protein with a peptide bond, and it folds into the correct 3D structure, nevertheless [38]. Moreover, it is possible to cut this circular chain so that to make a new N-end at the former middle of the

chain; and it folds, nevertheless, to the former native structure. Nowadays, protein engineering routinely produces circularly permuted proteins.

In an effort to solve the folding problem, Ptitsyn proposed a model of stepwise protein folding [39] (Figure 8). Later given the name "framework model", this hypothesis stimulated investigation of folding intermediates. It postulated a stepwise involvement of different interactions in the protein structure formation, stressed the importance of rapidly folded α-helices and β-hairpins at the initial folding steps, gluing of these helices and hairpins into a native-like globule, and crystallization of the final structure within this globule at the last step of folding.

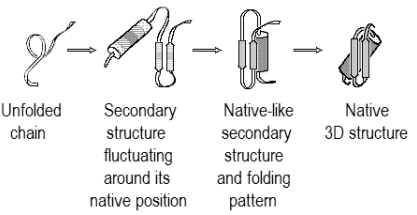


Fig. 8. Framework model of stepwise folding [39]. The secondary structures are shown as cylinders (α-helices) and arrows (β-strands). Both predicted intermediates have been already observed; the first is now known as the "pre-molten globule" and the other as the "molten globule" [21].

The cornerstone of this concept was then hypothetical and now well known folding intermediate, the "molten globule", which was later discovered and studied first as the equilibrium state of a "weakly denatured" protein [40, 41] and then as a kinetic folding intermediate [42].

Experiment shows that different properties of the native protein have two quite different rates of restoring. Nearly-native volume and secondary structure restore within a second, while side-chain order, tertiary structure and biochemical activity take minutes to restore. This is an evidence for accumulation of some "intermediate" state of the protein molecule (as shown, the molten globule [21]) at the beginning of the folding process.

The molten globule is an early folding intermediate in the *in vitro* folding of many proteins at physiological conditions [21, 43, 44]. It takes a few milliseconds to form, while the complete restoration of native properties of a 100 – 300 residue chain can take seconds for some proteins and hours for others. Thus, the rate-limiting folding step concerns formation of the native "solid" protein from the molten globule rather than formation of the molten globule from the coil.

The molten globule is not the only intermediate observed in protein folding. The "pre-molten" globule (that also fits the "framework model") was observed [45] to precede the molten globule formation. As a kinetic intermediate it was discovered with the use ultra-fast (sub-millisecond) measuring tech-

niques [46, 47]. In addition, the proteins with disulfide bonds allow trapping various intermediates indicative of the order of S-S bond formation, etc. [17].

The "kinetic control" hypothesis initiated very intensive studies of folding intermediates. Actually, it was clear almost from the very beginning that the metastable intermediates are not obligatory for folding (since the protein can fold also near the point of equilibrium between the native and denatured states, where the transition is of the "all-or-none" type, which excludes any metastable intermediates). The idea was, though, that the intermediates, if trapped, would help to trace the folding pathway, like intermediates in a complicated (bio)chemical reaction trace its pathway. This was, as it is now called, "chemical logic". However, this logic worked only in part when it came to the protein folding. The intermediates (like molten globules) were found for many proteins, but the main question as to how the protein chain can rapidly find its native structure among zillions of alternatives remained unanswered.

"Two-state" and "Multi-state" Protein Folding

A progress in the understanding of protein folding [48, 49] has been achieved just by investigation of those proteins, which fold without "unnecessary complications" (previously widely used to trace the folding pathway): without accumulation of any intermediates at the folding pathways, without *cis-trans* proline isomerization, and without S-S bond formation. The folding (and the unfolding) kinetics looks very simple in this case: all the properties of the native (or denatured) protein are restored synchronically, following the single-exponential kinetics [50]. For some proteins, this simplicity is observed in a wide range of conditions, including the denaturant-free water ("biological zone" in Figure 9), the zone of the reversible thermodynamic transition between two phases (the native and the denatured state) and the unfolding zone; these proteins obtained a name of "two-state proteins". For the other, "multi-state" proteins, the two-state folding occurs only in the transition zone, if any, while the unfolding demonstrates a "two-state" manner (Figure 9). Usually, the complicated folding demonstrates three phases, and the corresponding proteins obtained a name of "three-state proteins" [48–51]. Thus, the most universal features of folding (and unfolding) can be observed just in and around the transition zone, while the moving off this zone towards the "biological" conditions reveals individualities of various proteins (which are the "unnecessary complications", when we try to understand the basics of protein folding).

The above statement looks, in a sense, paradoxical. Indeed, what can we get from investigation of folding (or unfolding) in the transition zone, where we cannot accumulate any transition intermediates? The answer is: just here we can most readily, though indirectly observe the folding transition state, whose stability (or, more exactly, instability) determines the folding (and unfolding) rate [48–54]. The transition state corresponds to the free energy maximum on the folding/unfolding pathway, - or, better to say, to the free energy saddle

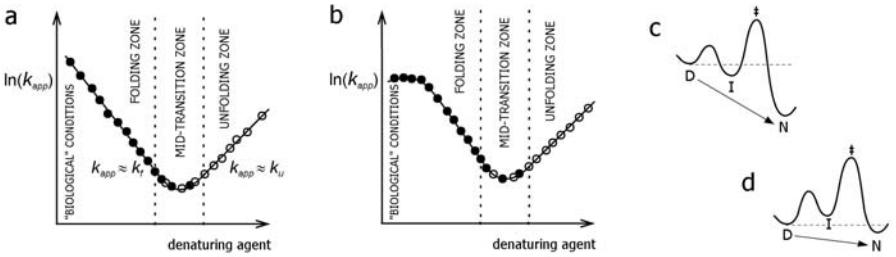


Fig. 9. Typical appearance of a "chevron plot" presenting an apparent rate of the folding/unfolding process (k_{app}) vs. the denaturant concentration (or the temperature). The closed circles correspond to folding occurs when the protein comes from the denaturing to renaturing medium. The open circles correspond to unfolding that occurs when the protein is transferred from the native to denaturing conditions. Note that circles of both kinds overlap at mid-transition. (a) Typical plot for a protein having the two-state folding throughout the whole range of experimental conditions. For the two-state folding, different characteristics of the protein change with equal rate k_{app}, and $k_{app} = k_f + k_u$, where k_f is the folding rate and k_u is the unfolding rate: thus, $k_{app} \approx k_f$ in the folding zone (where $k_f \gg k_u$), $k_{app} \approx k_u$ in the unfolding zone (where $k_f \ll k_u$) and$k_f \approx k_u \approx k_{app}/2$ at the mid-transition [51]. (b) Typical plot for the rate-limiting step of folding and unfolding of a "multi-state" folding protein: such a protein has the two-state folding only close to the mid-transition (i.e., the point of thermodynamic equilibrium between the native and denatured states), but not at the "biological" conditions where a multi-state folding occurs (and some state(s) arise(s) and/or decay(s) much faster than the complete transition occurs). (c, d): Free energy changes along the pathway of the "multi-state" protein folding at "biological" conditions (c) and close to the mid-transition (d); N is the native state, D the denatured, I the metastable (molten globule-like) folding intermediate and ‡ the unstable transition state.

point on the network of these pathways. The folded part of the transition state is called "folding nucleus", and the way of folding via formation of a nucleus (which usually consists of amino acid residues remote in protein chain [55,56]) obtained a name of "nucleation-condensation" mechanism of folding.

Folding Nucleus

"Folding nucleus" plays a key role in protein folding: its instability determines the folding and unfolding rate-limiting steps. It should be stressed that the folding nucleus is not the molten globule, although some of their characteristic may be similar [54]: the nucleus corresponds to the free energy maximum, while the molten globule corresponds to the free energy minimum [21]. It has been shown that the nucleus looks like some part of 3D structure of the native protein [51,54]. So far, there is only one, very difficult experimental method to identify the folding nuclei in proteins: to find residues whose mutations affect the folding rate by changing the transition state stability as strongly as that of the native protein [51, 54] (Figures 10, 11).

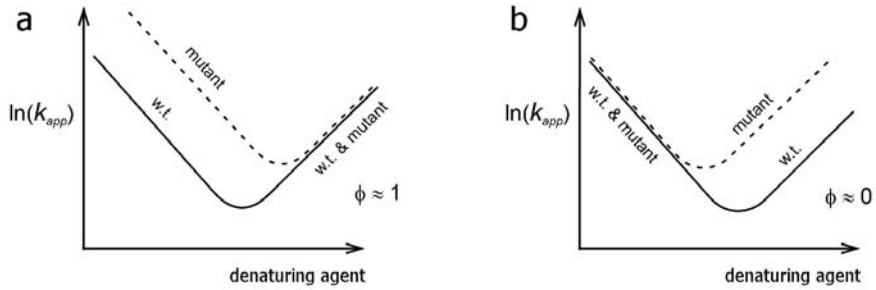


Fig. 10. Folding nucleus identification using site-directed mutations (a scheme). (a) Mutation of a residue, having its native environment and conformation (i.e., its native interactions) already in the folding nucleus changes the mutant's folding rate rather than its unfolding rate. (b) Mutation of a residue, which remains denatured in the transition state has the opposite effect.

The participation of a residue in the folding nucleus is expressed by the residue's ϕ value. ϕ is defined as $\Delta \ln k_f / \Delta \ln K$, where k_f is the folding rate constant, $K = k_f/k_u$ is the folding-unfolding equilibrium constant, and Δ means the mutation-induced shift of the corresponding value. According to the model of a native-like folding nucleus [51, 54], $\phi = 1$ means that the residue has its native conformation and environment already in the transition state (i.e., that this residue is in the folding nucleus), while $\phi = 0$ means that the residue remains unfolded in the transition state. The values $\phi \approx 0.5$ are ambiguous: either the residue is at the surface of the nucleus, or it is in one of the alternative nuclei, belonging to different folding pathways. It is noteworthy that the values $\phi < 0$ and $\phi > 1$ (which would be inconsistent with the model of a native-like folding nucleus) are extremely rare and never concerns a residue with a reliable measured $\Delta \ln K$.

It has been shown that proteins with different sequences but similar 3D structures often have similar folding nuclei [57–60]. However, there are many exceptions [61]. It has been shown also that circular permutation, changing the protein topology, sometimes changes [62, 63] and sometimes does not change [64] the transition state.

4 Theory of Protein Folding

All experimental data we discussed, though exciting, cannot answer the main question as to how a protein manages to find its native, apparently the most stable structure among zillions of others within those minutes or seconds that are assigned for its folding.

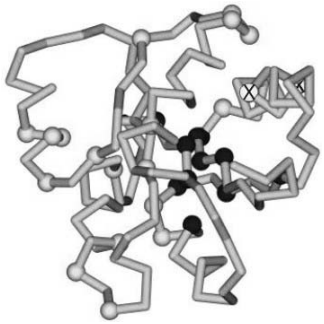


Fig. 11. Experimentally outlined folding nucleus for CheY protein [57]. The experimentally studied residues are shown as beads against the background of the native chain fold. The residues forming the folding nucleus are shown in black. Usually, the nucleus is shifted to the surface and does not coincide (though partially overlaps) with the protein's hydrophobic core. The gray beads indicate the residues that are not involved in the nucleus. The crossed beads show two residues that are difficult for interpretation since their $\Delta \ln K$ values are close to zero.

4.1 Solution of the Levinthal Paradox

The difficulty of this "Levinthal problem" is that it cannot be solved in direct experiment. Indeed, suppose that the protein has some structure that is more stable than the native one but folds very slowly. How can we find it if the protein does not do so itself? Shall we wait for $\sim 10^{10}$ years?

However, is there a real contradiction between "the most stable" and the "rapidly folding" structure? Maybe, the stable structure *automatically* forms a focus for the "rapid" folding pathways, and therefore it is *automatically* capable of fast folding?

Before considering *kinetic* aspects of protein folding, let us recall some basic facts concerning protein *thermodynamics* (as above, we will consider single-domain globular proteins only). This will help us to understand what chains and what folding conditions we have to consider. The facts are as follows:

- Protein folding and unfolding are usually reversible "all-or-none" transitions: only the native and denatured states of the chain are present (close to the denaturation point) in a visible quantity, while the other states are virtually absent. An "all-or-none" folding phase transition requires the amino acid sequence that provides a large energy gap between the native and the other folds.
- The denatured state, at least that of small proteins unfolded by a strong denaturant, is often the random coil.
- Even under physiological conditions the native state of a protein is only by a few kcal/mol more stable than its unfolded state (and these states have equal stability at mid-transition, naturally).

Thus, to solve the "Levinthal paradox" and to show that the most stable chain fold can be found within a reasonable time, we could, to a first approximation, consider only the rate of the "all-or-none" transition between the coil and the most stable structure. And we may consider only the case when the most stable fold is close to thermodynamic equilibrium with the coil, all other forms of the chain being unstable close to the "all-or-none" transition midpoint. Here the analysis is the simplest: it must not consider accumulating intermediates. True, the maximal folding rate is achieved when the native fold is much more stable than the coil (Figure 9), and then the observable intermediates often arise. But let us consider the situation when the folding is not the fastest but the simplest ...

Since the "all-or-none" transition requires a large energy gap between the most stable structure and the other ones, *we will assume that the considered amino acid sequence provides such a gap.*

We are going to show you that the "gap condition" provides a rapid folding pathway to the global energy minimum, to estimate the rate of folding, and to prove that the most stable structure of a normal size domain can fold within seconds or minutes [8, 65].

To prove that the most stable chain structure is capable of rapid folding, it is sufficient to prove that at least one rapid folding pathway leads to this structure. Additional pathways can only accelerate the folding since the rates of parallel reactions are additive. (One can imagine water leaking from a full to an empty pool through cracks in the wall between them: when the cracks cannot absorb all the water, each additional crack accelerates filling of the empty pool. And, by definition of the "all-or-none" transition, all semi- and mis-folded forms together are too unstable to absorb a significant fraction of the folding chains and trap them.)

A rapid pathway must include not too many steps, and, first of all, it must not require overcoming of a too high free energy barrier.

An $L-$residue chain can attain its lowest-energy fold in L steps, each adding one fixed residue to the growing structure. *If* the free energy went downhill along the entire pathway, a 100-residue chain would fold in ~ 100 – 1000 ns, since the growth of a structure (e.g., an α-helix) by one residue (τ) is known to take a few nanoseconds [66]. Protein folding takes much more than 1 μs only because of the free energy barrier, since most of the folding time is spent on climbing up this barrier and falling back, rather than on moving along the folding pathway.

According to the conventional transition state or Kramers theories [67,68], characteristic time ($\equiv 1/k$) of the process is estimated as

$$\text{TIME} \sim \tau \times \exp(+\Delta F^{\ddagger}/RT) \; . \tag{1}$$

Here T is the temperature, R the gas constant, τ the time of one step ($\sim$ ns), and $\Delta F^{\ddagger}$ the free energy of transition state relatively to the initial one, i.e., the barrier height.

Our main question is: how high is the free energy barrier $\Delta F^{\ddagger}$ on the pathway leading from unfolded to the lowest-energy structure?

If the fold-stabilizing contacts start to arise only when the chain comes very close to its final structure (that is, if the chain has to lose almost all its entropy *before* the energy starts to decrease), the initial free energy increase would form a very high free energy barrier (proportional to the *total* chain entropy lost). The Levinthal paradox claiming that the lowest-energy fold cannot be found within any reasonable time, since this involves exhaustive sampling of all chain conformations, originates exactly from this "golf course" picture of energy landscape (loss of the entire entropy *before* the energy gain).

However, this paradox can be avoided if there is a folding pathway where the entropy decrease is immediately or nearly immediately covered by the energy decrease (as in the usual first order phase transitions).

Let us consider a *sequential* (Figure 12) folding pathway. At each step of this process, one residue leaves the coil and takes its final position in the lowest-energy 3D structure. This pathway looks a bit artificial, but it is exactly the pathway of unfolding of the lowest-energy structure, went in the opposite direction. The detailed balance law [69] reads that direct and reverse reactions must follow the same pathway under the same conditions (and we already agreed to consider the mid-point of the folding-unfolding equilibrium). The advantage of considered this folding scenario is that it obviously exists (though the others are also not excluded); second, it allows us to consider only those residue-residue contacts which exist in the native protein [70].

Thus, we can replace a difficult analysis of folding by a simpler analysis of unfolding, and consider the folding nucleus instead of the nucleus of unfolding: these two nuclei coincide at the thermodynamic mid-transition conditions!

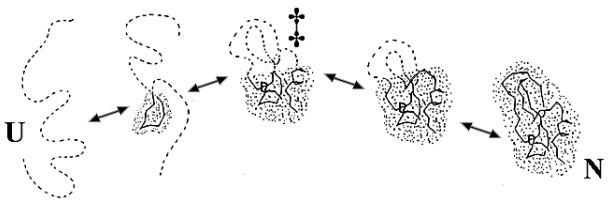


Fig. 12. Sequential folding (and unfolding) pathway [65]. U is the unfolded state, N the native state, ‡ the transition state. The folded part (dotted) is native-like. The bold line shows the backbone fixed in this part; the fixed side chains are not shown for the sake of simplicity (the volume that they occupy is dotted). The dashed line shows the unfolded chain.

Since, in the equilibrium point, the free energies of the native and the unfolded phases are equal, the additional free energy $\Delta F^{\ddagger}$ of the nucleus is due only to the boundary between the native and the unfolded phases. The largest boundary which will be met at the optimal phase transition pathway,

i.e., when the boundary between phases moves along the longest axes of the globule (Figure 12), includes not more than $\approx L^{2/3}$ out of L residues of the chain. This corresponds to a folding nucleus embracing about a half of the globule. The energy of this boundary can be estimated as $\approx 1/3\,\varepsilon L^{2/3}$ [65], where ε is the protein denaturation energy per residue, and 1/3 is a fraction of residue's contacts that are lost at a 2-dimensional surface in the 3-dimensional space. The ε value has been experimentally estimated as ≈ 1 kcal/mol, or $\approx 1.5\ RT$ at the room temperature [14]. Besides, depending on the protein topology and the boundary position, the surface of the nucleus may be or may be not covered by the unfolded closed loops, whose Flory entropy creates an additional surface tension that adds to the conventional surface energy of the boundary. At the very maximum (when all loops have equal length), this entropic term is

$$T\Delta S^{\ddagger,\,\mathrm{Flory}} = -L^{\,2/3} \cdot \frac{1}{6} \cdot \frac{5}{2} \cdot \ln(3L^{1/3})\ . \tag{2}$$

Here $L^{2/3}$ is the number of surface amino acid residues, the multiplier 1/6 reflects that only 1 out of 6 possible directions of the surface residue is consistent with beginning of a loop, the multiplier 5/2 is used instead of the Flory multiplier 3/2 because each loop avoids the space occupied by the globule, and $3L^{1/3}$ is the average loop length, when $L/2$ non-globular residues are divided into $(1/6)L^{2/3}$ equal-length loops. However, in a more typical case of random division of $L/2$ residues into $(1/6)L^{2/3}$ loops, the term $T\Delta S^{\ddagger,\,\mathrm{Flory}}$ does not exceed $RT\,.\,L^{2/3}$ [65] (which is, though, very close numerically to the estimate (3.2) when $L < 10^3$).

Thus, the free energy barrier for folding (and unfolding) in the mid-transition can be estimated as

$$\Delta F^{\ddagger} \approx (1 \pm 0.5) RT \cdot L^{2/3}\ . \tag{3}$$

$\Delta F^{\ddagger}$ is $\approx 1.5 L^{2/3} RT$, when the boundary is densely covered by loops, and $\approx 0.5 L^{2/3} RT$, when the boundary is free of them [65]. Since a characteristic time of rearrangement of one residue τ is $\approx 10\,\mathrm{ns}$ [66], it takes $\sim 10\,\mathrm{ns} \times \exp(1.5 L^{2/3})$ to overcome the free energy barrier of nucleation in the first case, and only $\sim 10\,\mathrm{ns} \times \exp(0.5 L^{\,2/3})$ in the second. This range is exactly consistent with the observed times of protein folding near the mid-transition (Figure 13). It has been also estimated that knotting of a 100-residue chain can increase the estimate of the folding time no more than twice, and of a 400-residue chain by an order of magnitude at most [71].

The reason for the obtained "non-Levinthal" estimate of achievement of the lowest-energy structure,

$$\mathrm{TIME} \sim exp\,[(1 \pm 0.5) L^{2/3}] \times 10\,ns\ , \tag{4}$$

is that the entropy decrease is almost immediately compensated for by the energy gain along the sequential folding pathway [73], and the free energy barrier occurs owing to the surface effects only.

It is noteworthy that the sequential folding pathway does not require any rearrangement of the dense globular part (which could take a lot of time): all rearrangements occur in the swollen (coil) phase and therefore are rapid.

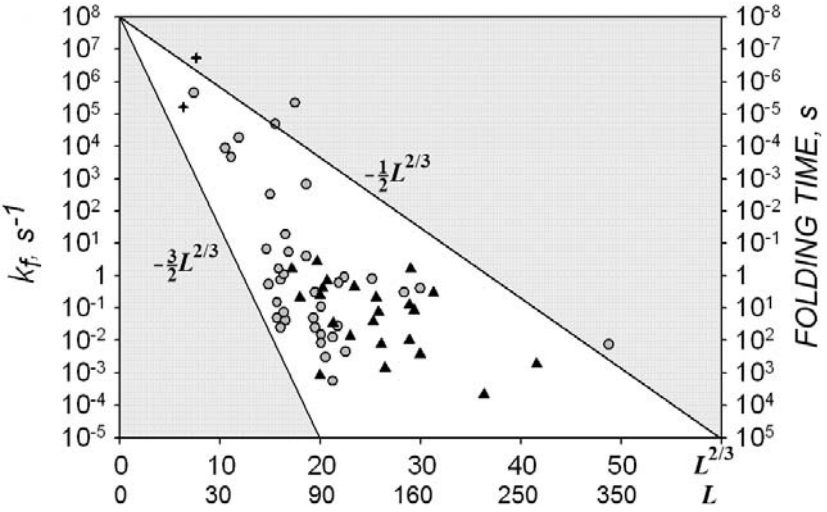


Fig. 13. Observed folding rate (and time) at the point of equilibrium between the unfolded and the native states vs. $L^{2/3}$ (L being the number of residues in the chain). The circles and triangles correspond to two-state and multi-state folding proteins, respectively. All these proteins do not contain S-S bonds or large ligands. Symbols + correspond to the α-helical and β-hairpin peptides. The theoretically predicted region of folding rates (white) is between the lines corresponding to $t = 10\,\mathrm{ns} \times \exp(0.5L^{2/3})$ and $t = 10\,\mathrm{ns} \times \exp(1.5L^{2/3})$. The folding time vs. $L^{2/3}$ correlation is 0.65. Adapted from [72].

Two notes in conclusion of this chapter:

- it is noteworthy that numerous computer simulation of folding never encounter any kinetically inaccessible structure of 3D protein models [74].
- it is remarkable that a recent mathematical estimate [75] of the maximal time necessary to find the most stable structure of a chain molecule in a $d-$dimensional space accords to equation $\ln(\mathrm{TIME}) \sim L^{(1-1/d)} \cdot \ln(L)$, in accordance with the above given estimates (2) – (4), obtained from simple physical considerations.

4.2 Theories of Protein Folding Rates

The above described theory not only "demystifies" (as it was written in [72]) "the protein folding problem cast in terms of the Levinthal paradox", but also opens a way to creation of more detailed theories of protein folding rates.

These theories have to take into account such factors as non-uniform distribution of strongly and weakly interacting amino acid residues within the globule, the folding pattern ("topology") formed by the native fold of the protein chain and folding under strongly non-equilibrium conditions.

The non-uniform distribution of weak and strong interactions between the nucleus and the remaining part of the protein contributes only the term of less than $\sim L^{1/2}RT$ in $\Delta F^{\ddagger}$ [65]. This effect is of secondary importance when the folding takes place close to the mid-transition where $\Delta F^{\ddagger} \sim L^{2/3}RT$. However, the non-uniformity effect seems to be important when the native state is much more stable than the denatured one [74, 76–78]. Besides, it is of a crucial importance for a theory of those, relatively small changes in the folding rates, which are caused by mutations of protein's residues.

When protein folds under the non-equilibrium conditions, the transition state free energy becomes smaller than that in the case of equilibrium conditions (see Figure 16 below). It was shown that this free energy $\Delta F^{\ddagger}$ scales with the chain length L as $\Delta F^{\ddagger} \sim L^{1/2}RT$ when the native state is moderately more stable than the denatured one [76,78], and as $\Delta F^{\ddagger} \sim (4 \div 6) \ln (L) \, . \, RT$ when the native state is much more stable than the denatured one [77,78].

A recent review of Shakhnovich [74] gives an excellent overview of effects that occur at non-equilibrium folding.

The influence of the protein chain topology upon the folding time has been estimated using a "contact order" parameter (CO) [79, 80]. The CO is equal to the average chain separation of contacting (in the native fold) residues, divided by the chain length. The CO is low for a structure rich of local contact, and high for a structure rich in contacts between remote chain regions. Specifically, CO is low for α-helices and high for β-structures; this may explain the observed 50-fold difference specifically in folding rates of α-helices and β-hairpins [81, 82].

A high CO value reflects the existence of many long closed loops in the fold of a complicated topology; thus, CO is roughly proportional to the "1 ± 0.5" multiplier in Equations (3), (4) [83]. When CO is taken into account in addition to the folding rate on the chain length dependence sown in Figure 13, the correlation of theory with experiment rises by additional 10% and reaches 0.74 [83]. The correlation is nearly the same for two- and multi-state proteins, in a contrast to results obtained for "pure" CO [79] , which are good for two-state proteins only.

Interestingly, that is even higher correlation (0.82) is observed between the length of an α-helix-free part of protein chain and the observed protein folding rate in denaturant-free water [84]; in this case, again, the correlation is nearly the same for two- and multi-state proteins (Figure 14). It is noteworthy that

this highly accurate prediction does not need any knowledge of the protein's tertiary structure: it can be obtained from amino acid sequence alone, since one can predict [85] α-helices form the sequence.

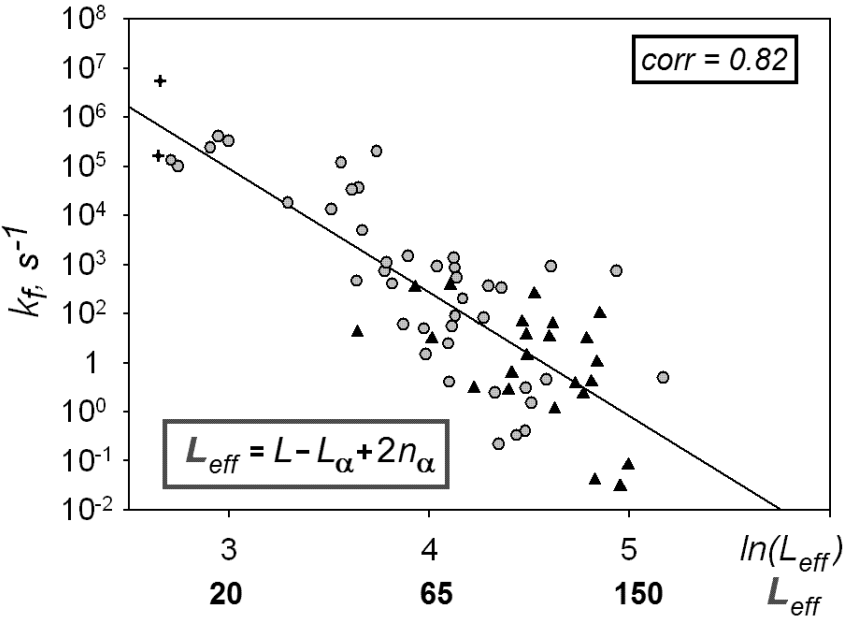


Fig. 14. Correlation between observed protein folding rate in denaturant-free water and the number of amino acid residues in the protein chain less the number of α-helical hydrogen bonds in this chain. Adapted from [84].

A more detailed and physically strict scheme to estimate of protein folding rates [86] can be obtained from analysis of the networks of folding pathways, or rather, the networks of pathways of unfolding of native protein structures (Figure 15). In addition to predicting protein folding/unfolding rates at various conditions, a theory based on this scheme is also able to find approximate localization of the protein folding nuclei [87–89].

The above described theoretical results show, in accordance with experiment, that the most stable protein fold must be found out within minutes. They explain also why the very large protein should consist of the separately folding domains ("foldons"): otherwise, the chains of more than ~ 400 residues would fold too slowly (within days) even when the protein chain topology is simple.

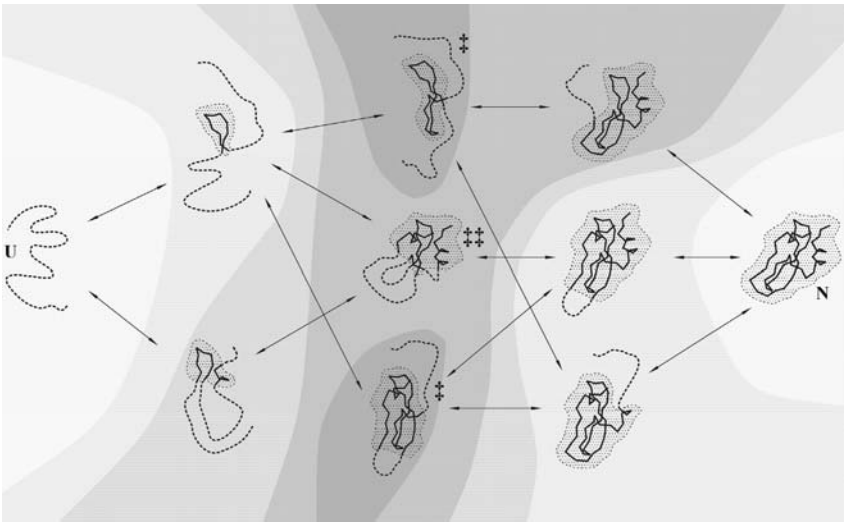


Fig. 15. Free energy landscape of protein chain; the darker the color, the higher the free energy. Arrows show possible transitions between semi-folded microstates (a "transition" being adding one chain link to the growing globule or removing of one chain link from the globule). These transitions are parts of various folding-unfolding pathways, one of which shown in Figure 12. Transition states are marked with ‡. The main transition state, i.e., the one of the lowest free energy, is marked with ‡‡.

4.3 Protein Folding Under Strongly Non-equilibrium Conditions

So far, we only considered the folding rate close to the mid-transition, where only one (the "native") fold competes with the coil, and all other globular forms, even taken together, are unstable.

If we come to an environment that stabilizes the native protein fold, it stabilizes the other globular and semi-folded structures of the chain as well. At first, this stabilization only increases the folding rate (see the rise of the left chevron limb from the mid-transition in Figures 9, 10), since the folding nucleus also is stabilized, and the competing (with the native fold) misfolded structures are still unstable (relative to the initial unfolded state as well) due to the energy gap between them and the native fold (Figure 16a).

However, this acceleration proceeds up to a certain limit only: the maximal folding rate is achieved when the "misfolded" states become as stable as the initial unfolded state, i.e., when the globular folding intermediates become stable (see the plateau at the top of the left chevron limb in Figure 16b). The further increase in stability of the folded states leads to a rapid misfolding followed by a relatively slow conversion into the native state (Figure 16b). This is a rather complicated process; it includes differences in coming to the native state by the protein main chain and its side chains [74].

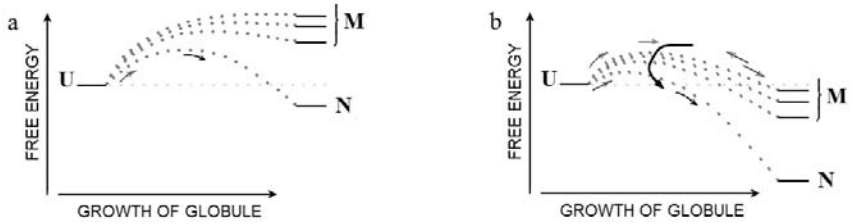


Fig. 16. Folding under conditions when (a) the most stable fold N is only a little more stable than the unfolded chain U, and (b) when N is much more stable than U, and some mis- or semi-folded structures (shown by numerous free energy levels M) are also more stable than U. Dotted lines schematically show the free energy changes along the folding pathways leading to different structures; maxima of these lines correspond to the transition states on these pathways. Adapted from [65].

5 Concluding Remarks

Actually, protein folding resembles crystallization (where a kind of the Levinthal paradox also exists, since the decrease in the number of configurations during crystallization is really huge) [90]. When crystallization [67] occurs close to the freezing temperature, a perfect large monocrystal (the lowest-energy structure) arises, although extremely slowly. As the temperature decreases a little, the monocrystal grows faster; and the further temperature decrease leads to a rapid formation of many chips rather than of a perfect large monocrystal.

The above given scheme of entropy-by-energy compensation along the folding pathway and the conclusion that it can solve the Levinthal paradox are applicable to formation of the native protein structure not only from the coil but also from the molten globule or from another intermediate. However, for these scenarios, all estimates would be much more cumbersome, while these processes (a variety of which are discussed in [21, 36, 37, 39, 47, 49, 73–80, 82]) do not show (on experiment) drastic advantages in the folding rate. Therefore, we will not go here beyond the simplest case of the coil-to-native globule transition.

In the very conclusion, it should be noted that a hierarchic scheme of protein folding [36, 37, 39, 91], as well as many simplified "protein folding funnel" models [92, 93] do not solve the Levinthal paradox since they cannot provide a *simultaneous* [94, 95] explanation for all the above discussed major features observed for protein folding: (i) spontaneous, non-assisted folding of a unique native structure within non-astronomical time, (ii) the fact that the same native structure can be achieved at very different conditions (including the thermodynamic mid-transition) and with very different folding rates, and (iii) co-existence, to a visible quantity, of only the native and unfolded molecules during folding of single-domain proteins near the thermodynamic mid-transition between the native and denatured states.

On the contrary, a nucleation-condensation mechanism can account for all these major features simultaneously and thus resolves the Levinthal paradox and opens a way to a theory of protein folding nuclei and protein folding/unfolding rates.

Acknowledgement. We are grateful to E.A. Davydova for assistance in preparation of the paper.

This work is supported by a grant from the "Physical and Chemical Biology" Program of the Russian Academy of Sciences, by grants 04-04-49682, 05-04-48750-a of the Russian Foundation for Basic Research, by an INTAS grant and by an International Research Scholar's Award from the Howard Hughes Medical Institute

References

1. Kendrew, J.C., Bodo, G., Dintzis, H.M., Parrish, H., Wyckoff, H., Phillips, D.C.: Nature, **181**, 662 (1958)
2. Perutz, M.F., Rossmann, M.G., Cullis, A.F., Muirhead, G., Will, G., North, T.: Nature, **185**, 416 (1960)
3. Wüthrich, K.: NMR of Proteins and Nucleic Acids. John Wiley & Sons, New York (1986)
4. Anfinsen, C.B., Haber, E., Sela, M., White, F.H.: Proc. Natl. Acad. Sci. USA, **47**, 1309 (1961)
5. Gutte, B., Merrifield, R.B.: J. Amer. Chem. Soc., **91**, 501 (1969)
6. Privalov, P.L., Khechinashvili, N.N.: J. Mol. Biol., **86**, 665 (1974)
7. Uversky, V.N., Gillespie, J.R., Fink, A.L.: Proteins, **41**, 415 (2000)
8. Finkelstein, A.V., Ptitsyn, O.B.: Protein Physics. A Course of Lectures. Academic Press, An Imprint of Elsevier Science, Amsterdam – Boston – London – N.Y. – Oxford – Paris – San Diego – San Francisco – Singapore – Sydney – Tokyo (2002)
9. Andreeva, A., Howorth, D., Brenner, S.E., Hubbard, T.J.P., Chothia, C., Murzin, A.G.: Nucl. Acid Res., **32**, D226 (2004)
10. Branden, C., Tooze, J.: Introduction to Protein Structure. Garland Publ. Inc., N.Y. (1999)
11. Bernstein, F.C., Koetzle, T.F., Meyer, Jr., E.F., Brice, D., Kennard, O., Shimanouchi, T., Tasumi, T.: J. Mol. Biol, **112**, 535 (1977)
12. Kraulis, P.J.: J. Appl. Cryst, **24**, 946 (1991)
13. Poroikov, V.V., Esipova, N.G., Tumanyan, V.G.: Biofizika (Moscow), **21**, 397 (1976)
14. Privalov, P.L.: Adv. Protein Chem., **33**, 167 (1979)
15. Nojima, H., Ikai, A., Oshima, T., Noda, H.: J. Mol. Biol., **116**, 429 (1977)
16. Privalov, P.L.: Adv. Protein Chem., **35**, 1 (1982)
17. Creighton, T.E.: Proteins. W.H. Freeman & Co., New York, 2nd ed. (1991)
18. Tanford, C.: Adv. Prot. Chem., **23**, 121 (1968)
19. Kuwajima, K., Sugai, S.: Biophys. Chem., **8**, 247 (1978)
20. Dolgikh, D.A., Abaturov, L.V., Bolotina, I.A., Brazhnikov, E.V., Bychkova, V.E., Bushuev, V.N., Gilmansin, R.I., Lebedev, Yu.O., Semisotnov, G.V., Tiktopulo, E.I., Ptitsyn, O.B.: Eur. Biophys. J., **13**, 109 (1985)

21. Ptitsyn, O.B.: Adv. Protein Chem., **47**, 83 (1995)
22. Dobson, C.M.: Curr. Biol., **4**, 636 (1994)
23. Shakhnovich, E.I., Finkelstein, A.V.: Biopolymers, **28**, 1667 (1989)
24. Finkelstein, A.V., Shakhnovich, E.I.: Biopolymers **28**, 1681 (1989)
25. Lifshitz, I.M., Grosberg, A.Yu., Khokhlov, A.R.: Rev. Mod. Phys., **50**, 683 (1979)
26. Kayaman, N., Guerel, E.E., Baysal, B.M., Karasz, F.: Macromolecules, **32**, 8399 (1999)
27. Shakhnovich, E.I., Gutin, A.M.: Nature, **346**, 346 (1990)
28. Goldstein, R.A., Luthey-Schulten, Z.A., Wolynes, P.G.: Proc. Natl. Acad. Sci. USA, **89**, 4918 (1992); *ibid.*, 9029
29. Finkelstein, A.V., Gutin, A.M., Badretdinov, A.Ya.: In: Biswas, B.B., Roy, S. (eds) Proteins: Structure, Function and Protein Engineering. Subcellular Biochemistry, Plenum Press, **24**, 1 (1995)
30. Stryer, L.: Biochemistry, W.H. Freeman & Co., New York, 4th ed. (1995)
31. Kolb, V.A., Makeev, E.V., Spirin, A.S.: EMBO J., **13**, 3631 (1994)
32. Ellis, R.J., Hartl, F.U.: Curr. Opin. Struct. Biol., **9**, 102 (1999)
33. Hardesty, B., Tsalkova, T., Kramer, G.: Curr. Opin. Struct. Biol., **9**, 111 (1994)
34. Fändrich, M., Forge, V., Buder, K., Kittler, M., Dobson, C. M., Diekmann, S.: Proc. Natl. Acad. Sci. USA, **100**, 15463 (2003)
35. Lührs, T., Ritter, C., Adrian, M., Riek-Loher, D., Bohrmann, B., Döbeli, H., Schubert, D., Riek, R.: Proc. Natl. Acad. Sci. USA, **102**, 17342 (2005)
36. Levinthal, C.: J. Chim. Phys. Chim. Biol., **65**, 44 (1968)
37. Phillips, D.C.: Sci. Am., **215**, 78 (1966)
38. Goldenberg, D.P., Creighton, T.E.: J. Mol. Biol., **165**, 407 (1983); *ibid.*, **179**, 527 (1984)
39. Ptitsyn, O.B.: Doklady AN SSSR (Moscow), **210**, 1213 (1973)
40. Kuwajima, K., Sugai, S.: Biophys. Chem., **8**, 247 (1978)
41. Dolgikh, D.A., Gilmanshin, R.I., Brazhnikov, E.V., Bychkova, V.E., Semisotnov, G.V., Venyaminov, S.Yu., Ptitsyn, O.B.: FEBS Letters, **136**, 311 (1981)
42. Dolgikh, D.A., Kolomiets, A.P., Bolotina, I.A., Ptitsyn, O.B.: FEBS Letters, **164**, 88 (1984)
43. Dyson, J.H., Wright, P.E.: Annu. Rev. Phys. Chem., **47**, 369 (1996)
44. Jackson, S.E.: Fold. Des., **3**, R81 (1998)
45. Uversky, V.N., Semisotnov, G.V., Pain, R.H., Ptitsyn, O.B.: FEBS Letters, **314**, 89 (1992)
46. Jones, C.M., Henry, E.R., Hu, Y., Chan, C-K, Luck, S.D., Bhuyan, A., Roder, H., Hofrichter, J., Eaton, W.A.: Proc. Natl. Acad. Sci. USA, **90**, 11860 (1993)
47. Shastry, M.C.R., Roder, H.: Nature Struct. Biol., **5**, 385 (1998)
48. Fersht, A.R.: Curr. Opin. Struct. Biol., **5**, 79 (1995)
49. Dobson, C.M., Karplus, M.: Curr. Opin. Struct. Biol., **9**, 92 (1999)
50. Kragelund, B.B., Robinson, C.V., Knudsen, J., Dobson, C.M., Poulsen, F.M.: Biochemistry **34**, 7217 (1995)
51. Matouscheck, A., Kellis, Jr., J.T., Serrano, L., Bycroft M., Fersht, A.R.: Nature, **346**, 440 (1990)
52. Segawa, S.-I., Sugihara, M.: Biopolymers, **23**, 2473 (1984)
53. Fersht, R.: Curr. Opin. Struct. Biol., **7**, 3 (1997)
54. Matouscheck, A., Kellis, Jr., J.T., Serrano, L., Fersht, A.R.: Nature, **340**, 122 (1989)

55. Itzhaki, L.S., Otzen, D.E., Fersht, A.R.: J. Mol. Biol., **254**, 260 (1995)
56. Abkevich, V.I., Gutin, A.M., Shakhnovich, E.I.: Biochemistry, **33**, 10026 (1994)
57. López-Hernándes, E., Serrano, L.: Fold. Des., **1**, 43 (1996)
58. Martinez, J.C., Serrano, L.: Nat. Struct. Biol., **6**, 1010 (1999)
59. Riddle, D.S., Grantcharova, V.P., Santiago, J.V., Alm, E., Ruczinski, I., Baker, D.: Nat. Struct. Biol., **6**, 1016 (1999)
60. Perl, D., Welker, C., Schindler, T., Schroder, K., Marahiel, M.A., Jaenicke, R., Schmid, F.X.: Nat. Struc. Biol., **5**, 229 (1998)
61. Steensma, E., van Mierlo, C.P.M.: J. Mol. Biol., **282**, 653 (1998)
62. Viguera, A.R., Serrano, L., Wilmanns, M.: Nat. Struct. Biol., **3**, 874 (1999)
63. Lindberg, M.O., Tangrot, J., Otzen, D.E., Dolgikh, D.A., Finkelstein, A.V., Oliveberg, M.: J. Mol. Biol., **314**, 891 (2001)
64. Otzen, D.E., Fersht, A.R.: Biochemistry, **37**, 8139 (1998)
65. Finkelstein, A.V., Badretdinov, A.Ya.: Fold. Des., **2**, 115 (1997)
66. Zana, R.: Biopolymers, **14**, 2425 (1975)
67. Ubbelohde, A.R.: Melting of Crystal Structure. Clarendon Press, Oxford (1965)
68. Moore, J.W., Pearson, R.G.: Kinetics and Mechanism. J. Wiley, New York (1981)
69. Lifshiz, E.M., Pitaevskii, L.P.: Physical Kinetics. Pergamon, London (1981)
70. G, N.: Int. J. Pept. Prot. Res., **7**, 313 (1975)
71. Finkelstein, A.V., Badretdinov, A.Ya.: Fold. Des., **3**, 67 (1998)
72. Finkelstein, A.V., Ivankov, D.N., Galzitskaya, O.V.: Uspekhi Biol. Khim. (Moscow), **45**, 3 (2005)
73. G, N.: Ann. Rev. Biophys. Bioeng., **12**, 183 (1983)
74. Shakhnovich, E.: Chem Rev., **106**, 1559 (2006)
75. Fu, B., Wang, W.: Proc. ICALP 2004, Lecture Notes in Computer Science **3142**, 630 (2004)
76. Thirumalai, D.: J. Phys. (Orsay, Fr.), **5**, 1457 (1995)
77. Gutin, A.M., Abkevich, V.I., Shakhnovich, E.I.: Phys. Rev. Let., **77**, 5433 (1996)
78. Wolynes, P.G.: Proc. Natl. Acad. Sci. USA, **94**, 6170 (1997)
79. Plaxco, K.V., Simons, K.T., Baker, D.: J. Mol. Biol., **277**, 985 (1998)
80. Fersht, A.R.: Proc. Natl. Acad. Sci. USA, **97**, 1525 (2000)
81. Finkelstein, A.V., Galzitskaya, O.V.: Physics of Life Reviews, **1**, 23 (2004)
82. Eaton, W.A., Muñoz, V., Hagen, S.J., Jas, G.S., Lapidus, L.J., Henry, E.R., Hofrichter, J.: Annu. Rev. Biophys. Biomol. Struct., **29**, 327 (2000)
83. Ivankov, D.N., Garbuzynsky, S.O., Alm, E., Plaxco, K.V., Baker, D., Finkelstein, A.V.: Protein Sci., **12**, 2057 (2003)
84. Ivankov, D.N., Finkelstein, A.V.: Proc. Natl. Acad. Sci. USA, **101**, 8942 (2004)
85. Jones, D.T.: J. Mol. Biol., **292**, 195 (1999)
86. Ivankov, D.N., Finkelstein, A.V.: Biochemistry, **40**, 9957 (2001)
87. Galzitskaya, O.V., Finkelstein, A.V.: Proc. Natl Acad. Sci. USA, **96**, 11299 (1999)
88. Garbuzynskiy, S.O., Finkelstein, A.V., Galzitskaya, O.V.: J. Mol. Biol., **336**, 509 (2004)
89. Garbuzynskiy, S.O., Finkelstein, A.V., Galzitskaya, O.V.: Mol. Biol. (Moscow), **39**, 1032 (2005)

90. Finkelstein, A.V.: Slow relaxations and nonequilibrium dynamics in condenced matter. In: Barrat, J.-L., Feigelman, M., Kurchan, J., Dalibard, J. (eds) UFJ NATO ASI. Les Houches, Session **77** 2002. EDP Sciences, Les Ulis – Paris – Cambridge and Springer-Verlag, Berlin – Heidelberg – New York – Hong Kong – London – Milan – Paris – Tokyo, 649 (2003)
91. Baldwin, R.L., Rose, G.D.: Trends Biochem. Sci., **24**, 26 (1999); *ibid.*, 77
92. Chan, H.S., Dill, K.A.: Proteins, **30**, 2 (1998)
93. Bicout, D.J., Szabo, A.: Prot. Sci., **9**, 452 (2000)
94. Bogatyreva, N.S., Finkelstein, A.V.: Prot. Eng., **14**, 521 (2001)
95. Finkelstein, A.V., J. Biomol. Struct. Dyn., **20**, 311 (2002)

Index