

Preface

This is not a biophysics book.

The readers of this book will indeed find a number of topics in it which commonly classify as biophysics; an example is the discussion of the electrostatic properties of biomolecules. But the book's ambition is different from being yet another introduction into biophysics. I therefore like to explain my motivation for the selection of topics I made.

The title of the book establishes a link between computational biology on the one side, and statistical mechanics on the other. Computational biology is the name of a new discipline. It is probably fair to say that it is, in fact, an emerging one, since a new name is not sufficient to establish a new discipline. Quantitative methods have been applied to biological problems for many years; this has led to a vast number of different subdisciplines of established fields of science: there is a mathematical biology, a biomathematics, a biostatistics, a bioinformatics, a biophysics, a theoretical biology, a quantitative biology... this list is certainly not exhaustive.

All these subdisciplines emerged out of the existing fields and developed by an act of transfer: the use of a method or a mathematical approach within a new context, its application to new problems. One may expect that at some point in the future all these different subdisciplines may merge into a common discipline. For the time being and for the lack of a definite name, let us call this discipline computational biology.

This book wants to contribute a particular element to this field; the use of statistical mechanics methods for the modelling of the properties of biological systems. Statistical physics is the scientific discipline which was developed in order to understand the properties of matter composed of many particles. Traditionally, it has been applied to non-living matter: gases, liquids, and solids. Meanwhile, it is increasingly applied to what is nowadays called 'soft matter' which encompasses complex objects like colloids, membranes and biomolecules, hence objects which do not clearly fall into any one of the classic categories.

Statistical mechanics methods are therefore indeed essential in their application to biophysical problems, since they are needed to understand the static and dynamic properties of biomolecules, complex molecular machines and even whole cell behaviour.

But there is a second aspect for which these methods can prove important, and this relates to the information content of the biological systems. Biology

is built on recognition processes: DNA strands have to recognize each other, proteins have to identify DNA binding sites etc. In bioinformatics, these recognition problems are commonly modelled as pattern recognition problems: this mapping is the basis of the enormous success of the field of modern genomics.

Moving beyond genomics, however, to the biology of whole systems, it becomes increasingly clear that an understanding of the physical properties of biological systems becomes more and more important for processes involving biological information content: DNA is not merely a string spelled out in a four-letter alphabet, but it is also an elastic string. There is biological information contained in its *structure* and its *dynamics*. Biology employs physical mechanisms to organize its information processes. This becomes particularly evident, as we begin to understand, in the properties of chromatin, the DNA-protein complex making up the chromosomes in the nuclei of eukaryotic cells (a topic which is discussed in the book). Or, in the particle interaction networks upon which the cellular machinery relies.

This book is placed at just this interface: between biological recognition on the one hand, and the statistical physics methods that can be employed to understand the underlying mechanisms.

The first part of the book gives a concise introduction into the main concepts of statistical mechanics, equilibrium and nonequilibrium. The exposition is introductory in the choice of topics addressed, but still mathematically challenging. Whenever possible, I have tried to illustrate the methods with simple examples.

The second part of the book is devoted to biomolecules, to DNA, RNA, proteins, and chromatin - i.e., the progression of topics follows more or less what is commonly known as the central dogma of molecular biology. In this part, mostly equilibrium statistical mechanics is needed. The concern here is to understand and model the processes of base-pair recognition and (supra-) molecular structure formation.

The third part of the book is devoted to biological networks. Here, both equilibrium and nonequilibrium concepts introduced in the first part are used. The presentation covers several of the nonequilibrium statistical physics approaches described in Chapter 2 of Part I, and illustrates them on biologically motivated and relevant model systems.

Throughout the book, *Exercises* and *Tasks* are scattered, most of them in the first part. They are intended to motivate the readers to participate actively in the topics of the book. The distinction between *Exercises* and *Tasks* is the following: *Exercises* should be done in order to verify that the concept

that was introduced has been understood. *Tasks* are more ambitious and usually require either a more involved calculation or an additional idea to obtain the answer.

A final technical note: the book is complemented by a detailed index. Keywords listed are marked in *italics* throughout the text.

In the course of shaping the idea of this book and writing it, I profited from discussions with many colleagues. I am especially thankful to Arndt Benecke, Dennis Bray, Martin Brinkmann, Luca Cardelli, Enrico Carlon, Avi Halperin, Andreas Hildebrandt, Martin Howard, Oliver Kohlbacher, Hans Meinhardt, Thomas Lengauer, Hans-Peter Lenhof, Annick Lesne, Ralf Metzler, Johan Paulsson, Andrew Phillips, Wilson Poon, Helmut Schiessel, Bernard Vandenbunder, Jean-Marc Victor, Pieter Rein ten Wolde, Edouard Yeramian.

Finally, I particularly like to thank Andreas Hildebrandt and Martin Howard for their detailed and helpful remarks on an early version of the manuscript.

I gratefully acknowledge the hospitality of the Institut d'Électronique, de Microélectronique et de Nanotechnologie (IEMN) in Villeneuve d'Ascq and the Institut des Hautes Études in Bures-sur-Yvette, where parts of this book were written.

1. Molecular Clocks and Time

1.1. The Clock

1.2. The Clock

1.3. The Clock

2. DNA

2.1. The Clock

2.2. The Clock

2.3. The Clock

3. RNA

3.1. The Clock

3.2. The Clock

3.3. The Clock

4. Proteins

4.1. The Clock

4.2. The Clock

4.3. The Clock

4.4. The Clock

4.5. The Clock