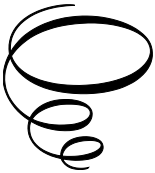
An Introductory Course on Molecular Biology

An Introductory Course on Molecular Biology

By

Ramón Serrano

Cambridge Scholars Publishing



An Introductory Course on Molecular Biology

By Ramón Serrano

This book first published 2023

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN (10): 1-5275-1009-3
ISBN (13): 978-1-5275-1009-8

For my wife Mariche, light of my shadows, and for our sons and daughters, sons-in-law and daughters-in-law and our 11 grandchildren. Without them this book would have been completed much quicker but everything would be meaningless.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I want to thank the many students of the grade on Biotechnology at the Polytechnic University of Valencia (UPV), Spain, who attended my courses on Molecular Biology and Genetic Engineering, because by answering my questions and asking me questions they helped to clarify a lot of concepts.

I want to recognize the work of the architect Rafael Ligorit for the layout of the book. Thanks also, to Hilario Teruel, consul of Burkina Faso, who always encouraged me to finish this book, my son Ramón, a professional translator, and proof reader Laurence Fenton, for a full correction of the manuscript. Last, but not least, thanks to the computer scientist Ramón Nogales who helped a lot with the computer work. Organizations such as Wikipedia, ResearchGate and others that offer figures and illustrations free of copyright on the internet have been very useful, especially given the enormous amounts of money demanded by some editorials for reproduction of figures and illustrations.

My gratitude to Prof. Alberto Sols (Research Institute on Enzymology, Autonomous University of Madrid, Spain), Richard Reeves (University of Louisiana, New Orleans, USA), Efraim Racker (Cornell University, Ithaca, New York, USA) and Gerald R. Fink (Whitehead Institute, Massachusetts Institute of Technology, Boston, USA) for being my teachers in scientific research and for dealing good-naturedly with my difficult character. sometimes I learned with them most of what I know and I will always be indebted to them.

Finally, thanks to the Polytechnic University of Valencia, Spain, and the Research Institute on Molecular and Cellular Plant Biology (IBMCP) for providing me the opportunity to teach and enabling me to continue research after retirement as Emeritus Professor.

INTRODUCTION

This book is intended to introduce students and scholars to Molecular Biology, the modern approach to living organisms that has revolutionized biology and generated novel biotechnological applications. The best text book in the field of Molecular Biology is that of Berg, Tymoczkko, Gatto and Stryer (2015). However, that book is a mixture of Biochemistry and Molecular Biology, which in my opinion are not the same thing.

I want to pinpoint in this introduction two great paragraphs from Steven Pinker (2018, page 378):

Any curriculum will be pedagogically ineffective if it consists of a lecturer yammering in front of a blackboard, or a textbook that students highlight with a yellow marker. People understand concepts only when they are forced to think them through. A second impediment to effective teaching is that pupils don't spontaneously transfer what they learned from a concrete example to another in the same abstract category.

Science is not a game with an arbitrary rulebook; it's the application of reason to explaining the universe and to ascertaining whether its explanations are true.

Finally, we should consider the thinking of the great Russian science fiction author Ivan Efremov (1952, 2):

The school must teach the latest knowledge. A lot of time is wasted teaching past things.

Another consideration is that the best way to think about a subject is to discuss it frequently with others, as Socrates and Plato did. For example, an isolated person like the one in the figure below gets bored while the group in the figure below has a good time.

Finally, Albert Einstein said once that: "*You do not really understand something unless you can explain it to your grandmother.*"

I would also like to make a digression about scientific literature. The Thomson Reuters impact factor of a journal in 2022 is calculated as follows:

A = citations in 2022 in all journals to articles published in journal X during 2020 and 2021.
B = number of articles published in journal X during 2020 and 2021.

Then the 2022 impact factor of journal X is A/B.

This method has some problems: (a) the citation period of only 2 years discriminates against papers that are slow to be accepted; and (b) it does not correct for the size of one scientific field (for example, biomedicine is much larger than plant sciences). A fashionable field has more journals that could provide citations (A above), but one particular journal might publish a similar number of articles in all fields (B above)

Other classifications of journals are: (a) journals with no publication fee (with the exception of color figures) but where institutions must pay to have access to the digital form of the journal or to receive the printed form in the library; and (b) journals with expensive publication fees (from €2,000 up to €10,000). These are the Open Access journals. In the latter case, some potential authors of articles may be excluded based on financial but not scientific criteria.

Randy W. Schekman, the 2013 Nobel Prize winner in Physiology and Medicine from the University of California at Berkeley, has said: *"Pressure to publish in luxury journals encouraged researchers to cut corners and pursue trendy fields of science instead of doing more important work. The problem was exacerbated by editors who were not active scientists but professionals who stimulated studies that were likely to make a splash*".

Randy W. Schekman (left) and Jorge E. Hirsch (right).

In order to evaluate the scientific production of scientists, Jorge E. Hirsch, presently professor of Physics at the University of California at San Diego, USA, proposed the "*h index*" (Hirsch 2005, 16569-16572). This is defined as the number of articles with a number of citations ≥ *h*. A scientist has an "*h index*" of X if X of his N articles have at least X citations each, and the other N-X papers have < X citations. It is calculated by ordering publications by decreasing number of citations. See an example below.

How to calculate the "h index" of an author:

Example:	citations		citations	
article 1	6		7	
article 2	4		6	
article 3	3	*h* index = 3	5	
article 4	2		4	h index = 4
article 5	2		2	
article 6	1		1	
article 7	1		1	

The advantage of this index is that the impact factor of the journals does not matter because journals with a high impact factor may have poor articles with few citations. The important thing is the number of citations received by the articles of one scientist, independently of the impact factor of the journals where the articles are published.

References

Berg, Jeremy M., Tymoczko, John L., Gatto, Gregory J. and Stryer, Lubert. 2015. *Biochemistry*. New York: W.H. Freeman and Company

Efremov, Ivan. 1952.*The Andromeda Nebula.* Translation from Russian by Maria K. Molodaya Gvardiya, Moscow, Russia, Foreign Language Publishing House

Hirsch, Jorge E. 2005. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences (USA)*, vol. 102, pages 16569-16572

Pinker, Steven. 2018. *Enlightenment Now. The case for Reason, Science, Humanism and Progress*. Great Britain, Allen Lane Penguin Books

CHAPTER 1

WHAT IS MOLECULAR BIOLOGY?

1.1 What is not and what is Molecular Biology

Molecular Biology is not the study of the molecules of living organisms. Biochemistry, also called Biological Chemistry, fits this definition because it studies biological molecules and their chemical reactions in cells (metabolism). The adjective "Molecular" before Biology does not refer to all molecules of living organisms but only to the two crucial ones: nucleic acids and proteins (structural and catalytic, as well as motors, membrane transporters and regulators). Molecular Biology is not the classical study of genes. This corresponds to Genetics, which deals with genes in relation to heredity, loci and genetic variation.

Molecular Biology is an approach to biological phenomena that is based on the atomic structures and genetic modifications of the mechanisms and physiological functions of the two crucial molecules for life: nucleic acids (DNA and RNA) and proteins. It is a reductionist approach derived from Biochemistry and from Genetics and it has developed a specific tool called Genetic Engineering (also known as Recombinant DNA Technology) to isolate genes, modify them *in vitro* and reintroduce them into organisms to investigate the physiological functions of genes and encoded proteins. It has also developed physical methods to determine the atomic structure and mechanism of these macromolecules. In X-Ray Diffraction, ordered molecules in fibers or crystals give a three-dimensional atomic picture of nucleic acids and proteins. More recently, in Cryo-Electron Microscopy (cryo-EM), many pictures of single, unordered molecules at liquid helium temperature are combined to provide similar atomic pictures of macromolecules because at such low temperatures the molecules are static, and without need of crystallization. Both Genetic Engineering and structural methods require a specific informatics approach: Bioinformatics, which is crucial to generate and utilize the big data provided by the above methods.

Some time ago, the physicist Ernest Rutherford (1871-1937, Nobel Prize in Chemistry 1908, Figure 1.1) said the following: “All science is either Physics or stamp collecting”. He was right because Physiology is fantasy, just correlations without cause-effect connections, Biochemistry may be an artefact because it consists of just “in vitro” studies and Genetics is just phenotypes and loci and does not reach the molecular level (genes and proteins). However, Molecular Biology has raised Biology to the level of an exact science because Genetic Engineering can prove hypotheses about the functions of genes and proteins, and structural studies of nucleic acids and proteins can demonstrate their mechanisms.

Figure 1.1. Ernest Rutherford. Courtesy of gettyimages.com.

Genetic Engineering can be used to generate mutants of specific genes with either gain-of-function (over-expression) or loss-of-function (knock-out) and then check the phenotypes of the organisms. If a function of an organism is worsened by the loss of function of a gene, this is a proof of the role of the gene and its corresponding protein in this function. If a function of an organism is improved by the gain of function of a gene, this is a demonstration that this gene and its corresponding protein not only participate in this function but that they are rate-limiting.

Molecular Biology explains “what is life”, the question and title of a famous 1944 book of Erwin Schrödinger, by testing hypotheses thanks to Genetic Engineering and uncovering biological mechanisms at the atomic level thanks to X-Ray Crystallography and Cryo-Electron Microscopy. These methodologies provide biotechnological tools such as transgenic organisms and rational drug design.

1.2 Gene nano-programs and protein nano-machineries explain the hallmarks of life

The mystery of life has always intrigued scientists. With regards to the chemical activities of organisms, the Swedish chemist Jöns Jacob Berzelius proposed in 1815 that organic compounds (those made by organisms) could only be produced by a God-given "vital force". But a German disciple of Berzelius, Friedrich Wöhler, synthesized urea in the laboratory in 1828, without the use of animals, and so dispelled the need of a mysterious vital force.

However, the chemical activities of organisms (metabolism) are relatively simple functions of life and cells are much more complex than just a bag of enzymes (the biological catalysts).

The six hallmarks of life, as adapted from SCIENCEFACTS, https://www.scifacts.net/biology/what-is-life/, are:

1. *Carbon-based and cellular organization*
   Living things are made from organic compounds (carbon-based); they are composed of cells and are either unicellular or multi-cellular.
2. *Reproduction and heredity*
   Asexual or sexual reproduction, the later involving the joining of gametes or sex cells. Daughter cells and children organisms inherit features from mother cells and parent organisms, respectively.
3. *Metabolism*
   The chemical reactions of living things; nutrients are taken and utilized as sources of energy (catabolism) and as sources of pillars to build up cell components (anabolism).
4. *Growth and development*
   Cells enlarge and divide; development of multi-cellular organisms occurs by growth and differentiation of cells.
5. *Response to stimuli and homeostasis*
   Cells and organisms make changes in response to external stimuli to keep their internal environment within a narrow range.

6. Adaptation and evolution

Variation of individuals is important for adaptation of species to certain environments; selection of adaptation traits is the basis for evolution.

Living things have all these features, while other beings, such as viruses, demonstrate only a few of these characteristics, specifically numbers 2 and 6 above, and are not living organisms. Because all living things have these common features (the hallmarks of life) and utilize the same basic molecules, nucleic acids and proteins as key molecules and carbohydrates and lipids as secondary ones, it has been proposed that all present organisms derived from a single ancestral cell called LUCA (Last Universal Common Ancestor), who lived about 4 billion years ago, see https://phys.org/news/2018-12-luca-universal-common-ancestor.html).

The emerging picture of Molecular Biology is a scientific understanding of the complexity of life with no need to invoke mysterious forces in organisms ("vitalism" theories). Molecular Biology was born in the middle of the 20th century through the realization that life consisted of molecular phenomena that transcended classical biochemistry and genetics. Genes are complex nano-programs of DNA that dictate all cellular activities and determine inheritance and evolution (hallmarks 2 and 6). Sophisticated nano-machineries made from proteins (and sometimes including RNA) replicate and express genes, convert chemical energy into mechanical and osmotic energy, move molecules across biological membranes, and regulate all cellular activities. See Figure 1.2 for a graphic description of these statements.

1.3 Molecular Biology is a basic science but its methods have enormous applications

Molecular Biology has raised Biology from a descriptive level into the category of an "exact science" because hypotheses can be tested by genetic modifications of organisms and the mechanisms of nucleic acids programs and protein nano-machineries can be understood at the atomic level through either X-ray Crystallography or Cryo-Electron Microscopy.

Genetic Engineering is the basic tool that proves the physiological function of genes and proteins by generating gain-of-function and loss-of-function mutants. Basically, modifications of isolated genes are made "in vitro" to generate gain-of-function (higher expression or higher protein

activity) or loss-of-function (knock out/null mutants or mutants with partial loss of function) and then these mutated genes are introduced into organisms by replacing the wild type copy. Phenotypes are investigated to ascertain what functions of the organism are modified. A final development has been the study of thousands of genes at the same time with special microarrays containing thousands of nano-samples in small pieces of glass. This process has been named "Genomics", and the word "omics" has been generalized to massive studies with proteins (Proteomics).

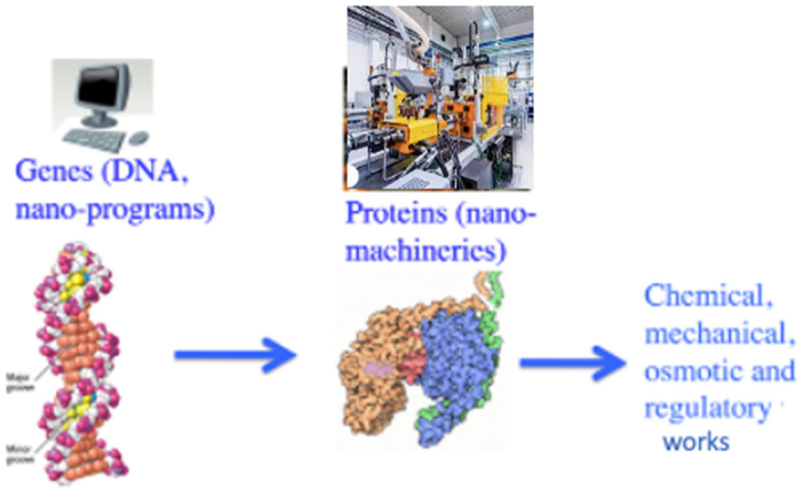


Figure 1.2. Molecular Biology is an approach to understanding life by considering proteins and nucleic acids as the "hardware" (nano-machineries) and "software" (nano-programs), respectively, of living cells. These crucial molecules are studied at the atomic level and modified to test hypotheses.

An example is shown in Table 1.1. The physiological function during intracellular acid stress of the two prolyl isomerases in the model plant *Arabidopsis thaliana* was investigated using two kinds of transgenic plants: (a) single (*rof1* or *rof2)* and double (*rof1 rof2*) mutants with loss-of-function (knock out) by insertion of T-DNA (the part of the Ti plasmid of *Agrobacterium tumefaciens* transferred to plant cell genomes by conjugation); and (b) gain-of-function by over-expression of *ROF2* (OE *ROF2*) through transformation with a chimeric gene containing the strong 2x35S viral promoter and the coding regions of *ROF2*. Intracellular acid stress was produced by the addition of acetic acid, which diffuses into

cells in the protonated form and dissociates protons inside. The experiment demonstrates that these two prolyl isomerases are redundant (single mutants have little phenotype but the double mutant is clearly more sensitive to the acid than the wild type) and required for maximum tolerance to intracellular acidification. On the other hand, the improved tolerance to intracellular acidification resulting from over-expression of *ROF2* suggests that this prolyl isomerase gene is limiting for tolerance to intracellular acid stress. Over-expression of a gene corresponding to a not-limiting step would not improve stress tolerance.

Table 1.1. Arabidopsis ROF1 and ROF2 prolyl isomerases modulate germination in media containing weak organic acids. Percentage of germination and seedling establishment of Arabidopsis wild-type, the *rof1* mutant, *rof2* mutant, *rof1 rof2* double mutant, and a line over-expressing ROF2 (OE *ROF2*) four days after planting in normal MS medium (pH 5.5) and in this medium supplemented with 3.5 mM acetic acid. Data from the author.

Percentage of small germinated plants with green cotyledons					
	Wild type	*rof1*	*rof2*	*rof1, rof2*	OE *ROF2*
Control	99	98	98	98	99
Acetic acid	43	36	28	18	75

These methodologies, developed to test the physiological function of genes, have also led to transgenic organisms (GMOs or Genetically Modified Organisms) with applications in agriculture (useful transgenic crops), animal breeding (useful transgenic animals) and medicine (useful transgenic microorganisms producing drugs, hormones or antibodies). One example of a genetically modified organism widely utilized is the insect-resistant potato, which expresses insecticide proteins from *Bacillus thuringiensis* called Bt and is encoded by *Cry* genes (see Figure 1.3). Other examples are the fast-growing salmon (which over-expresses growth hormone) and the insulin-producing bacteria (which express the human insulin gene).

An irrational rejection of transgenic crops has been developed in Europe by environmental (Green) parties and non-scientific populist organizations such as Greenpeace. This rejection has no scientific basis and it is hoped it will go away with time because GMOs are the future of agriculture and cattle raising. In America and China there is no such rejection. When the

Spanish discoverers of America brought tomatoes and potatoes from the new world to Europe (16th century) they were immediately eaten in Spain but it took two centuries for the French and Germans to eat these wonderful healthy foods. Non-Mediterranean Europeans are very conservative and easily confused by environmental activists. The European Union, for example, had for a period a Scientific Advisor, the English scientist Prof. Anne Glover, but as she made a document in favor of transgenic organisms she was inmediately fired. Clearly, the European Union operates by the non-scientific ideology of Greenpeace and has no respect for scientists.

Figure 1.3. Transgenic potato plants that express the Bt protein and are resistant to the potato beetle (*Leptino* sp.) (top of picture) and control plants (bottom of picture). Courtesy of Professor Francisco Garcia-Olmedo (School of Agricultural Engineers, Madrid, Spain).

Atomic structures require the biochemical purification of proteins from either cell tissues or from transgenic microorganisms or convenient organisms expressing the protein of interest. X-ray diffraction of crystals or cryo-Electron Microscopy (cryo-EM) of single particles provide the required structures. Protein crystals provide a static array of molecules for X-ray diffraction, but by using single protein molecules at the liquid helium temperature (-269 ºC) during Cryo-EM, molecules are static and, as indicated above, the three-dimensional structure can be solved without the tedious crystallization.

Atomic structures allow us to understand the mechanisms of the protein nano-machineries involved in chemical, mechanical, osmotic and regulatory works in cells. An example of a complex protein machinery is the mitochondrial Fo.F1-ATPase or ATP synthase of bacteria, mitochondria

and chloroplasts, which is composed of 31 protein subunits and has had its atomic structure solved by modern Cryo-Electron Microscopy (see Figure 1.4). Mitochondria carry out oxidation of reduced substrates with oxygen and the respiratory complexes couple the chemical energy of red-ox reactions to pumping protons out of mitochondria. This is the chemiosmotic theory of Peter Mitchell, who received the Nobel Prize in Chemistry in 1978 and buried the previous model based on chemical phosphorylated intermediates as occurs in glycolysis. The electrochemical proton gradient (pH difference and electrical potential) drives ATP synthesis by forcing mitochondrial Fo.F1-ATPase to run in reverse, working as an ATP synthase and coupling proton uptake by mitochondria to ATP synthesis.

The enzyme was discovered by Efraim Racker and is made of two parts: Fo - with the “o” coming from conferring sensitivity to the drug oligomycin, binding to the trans-membrane part - and F1 - factor 1 of the oxidative phosphorylation with ATPase activity protruding from the membrane. Both parts contain static subunits and rotary ones. The mechanism was solved by the model of Paul Boyer and the enzyme structure of John E. Walker and both shared the Nobel Prize in Chemistry in 1997. This protein nano-machinery works as a rotary motor: ATP hydrolysis (ATP + H_2O -> ADP + Pi + 30 kJ/mol) forces rotation of the rotor part with respect to the stator part (counter clock-wise from the top of F1), resulting in the pumping of protons out of the mitochondria. When the proton gradient generated by the mitochondrial respiratory chain is big enough (> 40 kJ/mol), protons move into mitochondria though the ATPase, and force rotation of the rotor part clockwise from the top of F1, i.e. in the opposite direction to before, and the enzyme is converted into ATP synthase. Rotation of the rotor part modifies the three catalytic subunits (β) sequentially from active site open (O), to active site binding ADP + Pi (L), to active site binding preferentially ATP + H_2O (T). Alternatively, ATP hydrolysis when no respiration is working results in a T > L> O sequence. See Figure 8.19.

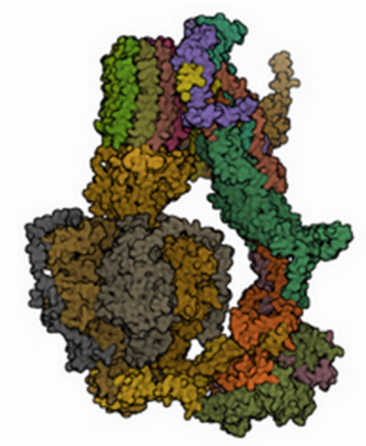
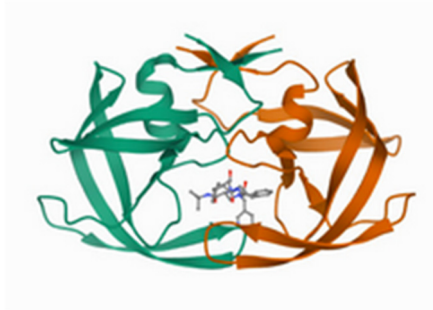


Figure 1.4. Structure of mitochondrial Fo.F1 ATPase. Courtesy of Protein Data Bank.

In addition to providing insight into mechanisms, the atomic structures of crucial proteins in human diseases, such as viral proteins and oncogenic proteins, paved the way for rational drug design. Major examples are inhibitors of the essential protease of human immunodeficiency virus such as Crixivan (Figure 1.5 A) and inhibitors of tyrosine kinases such as Gleevec (imatinib mesylate), which inhibits the oncogenic Bcr-Abl fusion protein active in chronic myelogenous leukemia (Figure 1.6 A). Both cases illustrate the point that determination of the atomic structure of targets is essential for rational drug design. Organic molecules are then designed to block the active site of the protein. In the first case the inhibitor resembles the peptide substrate of HIV protease (Figure 1.5B). In the second case, Gleevec binds to the catalytic cleft of the kinase without resembling the substrates of the enzyme (Figure 1.6 B).

A



B

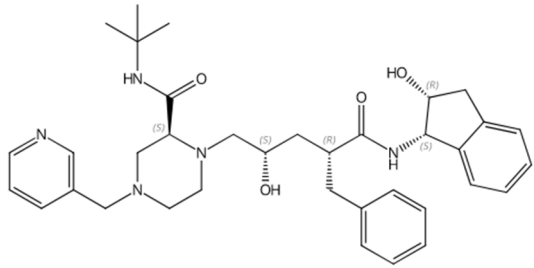
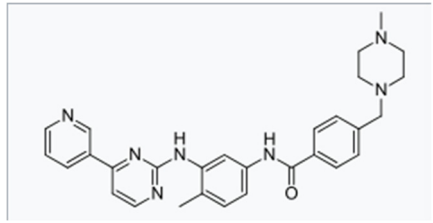


Figure 1.5. The design of indinavir as an inhibitor of the essential protease of Human Immunodeficiency Virus (HIV). A: determination of the atomic structure of HIV protease was essential for the rational design of indinavir (Crixivan, bound to the enzyme); B: this drug resembles a peptide substrate of the protease. Courtesy of Protein Data Bank (A) and SciFinder (B).

A



B

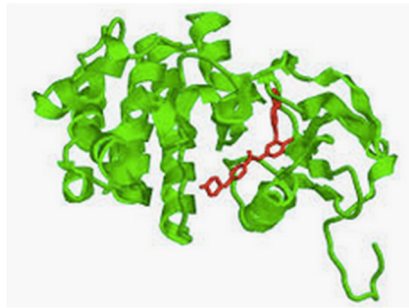


Figure 1.6. Gleevec was designed to bind the catalytic cleft of the Bcr-Abl protein tyrosine kinase. A: structure of Gleevec; B: binding of the inhibitor into the catalytic cleft. A courtesy of es.m.wikipedia.or and B courtesy of The Medical Dictionary.

The three steps of rational drug design are:

1. Identifying the protein responsible for a disease.
2. Solving the atomic structure of this protein.
3. Designing an organic molecule able to block the active site of the protein.

Another methodology of Molecular Biology is Bioinformatics, which is the analysis of the sequences and structures of genes and proteins. This has become especially prevalent with the “big data” that resulted from the “omic” approaches (global studies of all the genes and proteins of an organism – called genomics and proteomics, respectively). Algorithms for “big data” analysis have enormous importance in health and marketing

studies. An example is the genotyping of human beings and microorganisms, plants and animals. The sequencing of different genomes from the same species has shown differences between individuals belonging to geographical and racial groups. This has become a very profitable business for an American company named "23andMe" that utilizes saliva samples from customers to isolate DNA and analyze SNPs (Single Nucleotide Polymorphisms - changes of one nucleotide at a position of the genome in different individuals) to estimate ancestry and health risks. More than 600,000 oligonucleotides are printed onto small chips by the company Illumina (San Diego, California, USA) to be hybridized with the DNA samples of the organism by automated techniques and the resulting map is correlated with phenotypes (https://www.23andme.com/?mdc2=true).

1.4. An example of a Molecular Biology approach: growth control by proton transport

Physiological and biochemical studies have shown that cell plasma membranes have transport systems for two monovalent cations: protons and sodium. Two types of transporters exist: primary active transporters and secondary transporters. The first couple light or chemical energy (red-ox reactions or ATP hydrolysis) to pump H^+ or Na^+ out of cells, creating the electrochemical gradient of these cations. Secondary transporters couple the uptake of nutrients or the efflux of toxic molecules to the energy released by the downhill movements of H^+ or Na^+ into cells. Animal cells and some prokaryotes have a primary transport of sodium and a secondary one of protons, while fungi, plants and most prokaryotes have a primary transport of protons and a secondary one of sodium. In addition to nutrient uptake, intracellular pH and sodium concentrations seem to have a more specific regulatory role in cell growth (see Chapters 8, 9 and 10). A Molecular Biology approach to this important aspect of Biology has the following steps:

1. Isolation by biochemical methods of the plasma membrane H^+-ATPase (Pma1) from the model organism *Saccharomyces cerevisiae* (baker's yeast) and demonstration *in vitro* (upon reconstitution in lipid membranes) that the enzyme couples ATP hydrolysis to pump protons into vesicles. See Table 10.1.
2. Isolation by Genetic Engineering methods of the *PMA1* gene encoding the yeast proton pump.

3. Generation by reverse genetics of mutant yeast cells with different levels of proton pump activity.
4. Correlation between growth rate and Pma1 activity in the above series of mutants (Table 1.2).
5. Expression in mouse fibroblasts of the yeast *PMA1* gene, resulting in fast growth and tumorigenic transformation (Table 1.3).
6. The conclusion is that proton transport regulates growth not only in yeast, where H^+ transport is primary, but also in mouse fibroblasts, where H^+ transport is secondary to Na^+ transport. Cancer is a disease of unregulated growth control where proton efflux and intracellular pH are greater than normal and the above results open novel therapeutic strategies centered on proton transport.

Table 1.2. Correlation between activity of yeast plasma membrane H^+-ATPase in a series of yeast mutants generated by genetic engineering and intracellular pH and growth rate in media at pH 4.0. Data from the author.

Relative ATPase Activity (%)	Growth rate (h^{-1})	Intracellular pH
30	0.05	5.4
50	0.07	5.6
60	0.10	5.9
100	0.15	6.1

Table 1.3. Correlation between H^+-ATPase activity, tumorigenic transformation and intracellular pH of mouse fibroblasts expressing different mutants of yeast plasma membrane H^+-ATPase. Data from the author.

(%) ATPase activity of the introduced gene	(%) Tumorigenic transformation	Cell pH
No gene introduced	<5	7.1
100 (wild type)	50	7.3
10	<5	7.1
20	10	-
70	30	7.2
300	70	7.4

1.5. Hallmarks of Molecular Biology

Molecular Biology can be differentiated from Biochemistry and from Genetics, the two most related disciplines. We may list the most important distinguishing features (hallmarks) as the conclusion of this chapter:

(1) The basic concept of Molecular Biology is that biological phenomena can be understood through knowledge of the genes involved (software) and their encoded proteins (hardware), the two central molecules of life. This generalization reflects the molecular unity underlying the biological diversity of living organisms (Chapter 1).

(2) In addition to the classical tools of Biochemistry and Genetics, the specific tools of Molecular Biology are Genetic Engineering (gene manipulation), physical methods (X-ray crystallography and Cryo-Electron Microscopy) applied to the atomic structure of genes and proteins and, finally, Bioinformatics (to manipulate normal and "Big" data). These tools have generated novel applications in Medicine and Agriculture (Chapters 1-4).

(3) The function of genes and encoded proteins can be demonstrated by genetic modification. The mechanisms of protein machineries can be investigated at the atomic level through structural studies (Chapters 3 and 4).

(4) The genomic revolution is transforming Molecular Biology and its applications. Biological phenomena are tackled by studying thousands of genes at the same time (Chapters 2 and 3).

(5) Gene expression and protein synthesis are precise and regulated by complex machineries of many protein subunits and RNAs, explaining inheritance and the production of the proteins encoded by genes (Chapter 8).

(6) Protein nano-machineries execute chemical, mechanical, osmotic and regulatory works in cells (Chapters 5-6 and 8-9), explaining all the mysteries of life.