Molecular Biology and Genomics

分子生物学与基因组学

Cornel Muelhardt

Translated by
E . W . Beese , M . D .
Munich , Germany

前　　言

　　真是令人难以置信,在短短的五年内我们就准备要发行本书的第四版了。在第四版中,我们再次扩充了内容;然而即使专题的覆盖面一直在增加,本书的主旨仍然保持基本不变。例如,本次将微测序法、焦磷酸测序与 RNA 干扰包括了进来。同时还着手编写了与年轻科研工作者职业生涯规划相关的短篇——毕竟一个人不能仅依靠科研的神圣而生存。另外还循例做了大量小的修订。尽管自发行以来已做了大量的工作,仍留有很大的改进空间。你可以做任何可能之事,但最终都处在一种不确定性的掌控之中!本书难免会有一些印刷错误,希望读者不吝指正。

<div align="right">(刘斌　译)</div>

第一版前言

真的想读一下前言吗?

读过 Hubert Rehms 的《实验者系列:蛋白质生物化学与蛋白质组学》的前言与结束语并细细品味后,再读一下 Siegfried Bars 在"如何在德国做科研:研究员与将成为研究员者之马基雅弗利准则"中的语句。即使新手不倾向于相信这些,然而事实确是如此。换句话说,最成功的研究者,其出众之处在于对包括出版在内的各种场合都能应付自如,广交益友,善于与朋友沟通,而非什么开创性的成就。我们的目标是把从学生成就为教授的时间尽可能地缩短。然而,一旦成为教授,研究工作就交与他人完成,这正如一个成功的军师不是依靠自己的战刀,而是通过向野心勃勃的指挥官出谋划策,取得理论上的胜利。在此身份下,一个人仅仅是间接与科研相联系。如果在读过本书后你没有放弃,那么你可能已经显示了献身于分子生物学研究所需要的"沉着"。换句话说,为了避免被实验室司空见惯又永无休止的挫折折磨得发疯,冷静的头脑是必需的。

"不要气馁⋯⋯!其他人即使没有成功也仍在奋力地工作;刚开始时得不到结果很正常。不要放弃!—— 要么就去选择另一个更合适的职业。"

(刘斌　译)

致　　谢

　　以前我一直认为,作者在书的开头对朋友、亲属的溢于言表的致谢是作者义务无趣的表现。今天,我终于意识到致谢其实是历经磨练、发自内心的真情流露。感谢 Marc Bedoucha, Igor Bendik, Hans-Georg Breitinger, Dorothee Foernzler, Christophe Grundschober, Andreas Humeny, Frank Kirchhoff, Nina Meier, Nicoletta Milani 和 Ruthild Weber,是他们的贡献促成本书的成功。把最温馨、最特别的感谢送给 Cord-Michael Becker,多年来她一直支持着我。

　　最后,也感谢 Spektrum Verlag 和本书的策划者 Ulrich Moltmann,他们提出编写这样一本书的设想,我才得以实现这一夙愿;感谢 Jutta Hoffmann,是 Jutta 不时而友好的敦促,使我没有半途而废;感谢致力于本书每一版初稿修订的 Bettina Saglio。如果上帝真的存在的话,那就请保佑他们能与更多严谨的作者合作。

　　众所周知,没有任何事物是完美无缺的。这本书也是一样,因此我非常乐意接受任何建设性的批评。

<div align="right">(刘斌　译)</div>

目　　录

（刘斌　译）

Contents

Foreword

It is hard to believe that we are now already producing the fourth edition within such a brief time period of only five years. In this edition we have once again increased its dimensions; here, however, the basics have remained more or less unchanged, although the number of "specialties" covered continues to increase. This time, for instance, minisequencing, pyrosequencing, and RNA-interference have been included. In addition, the time has now come to write a short introduction concerning the planning of careers for younger researchers—after all, one does not live from the nobility of research alone. Numerous, minute corrections have had to be performed as well, since nothing is perfect. This book as well, in spite of the enormous amount of work which has already been carried out in its production, nevertheless continues to demonstrate room for improvement. One can do whatever is possible, but entropy always seems to win in the end! However, there are enough typographical errors still to be found here as well and we are therefore thankful for every piece of advice which we might receive!

Foreword to the First Edition

Do you actually want to read a foreword?

Read the preface and concluding comments in Hubert Rehms' "Experimentator: Proteinbiochemie" ("Protein Biochemistry" in the "Experimentator" series from Spektrum Verlag) and, finally, after getting a taste for them, read those statements in Siegfried Bars' "Forschen auf Deutsch: Der Machiavelli für Forscher—und solche, die es werden wollen" ("How to Do Research in German: The Machiavellian Guideline for Researchers—and for Those Who Would Like to Become Such" from Verlag Harri Deutsch). Even if the beginner is not inclined to believe so, the wind actually blows as has been described here. The most successful researcher, namely, is not to be distinguished by any pioneering accomplishments, but rather through his/her clever actions in scenes involving publications and in his/her large number of friends. The goal is to keep the phase as short as possible between one's time as a student and that during which one is authorized to function as a professor. Once one has finally reached this goal, however, the research is then left to the others; in the same way in which a successful military strategist no longer reaches for his/her own saber, but instead only carries out his/her advances theoretically through the actions of ambitious military leaders. At this stage, one merely has an indirect relationship to research. Should you not have given up after completing this reading material, you may possibly demonstrate the necessary composure to dedicate yourself to research in molecular biology. This composure, namely, will prove to be quite essential in order to avoid being driven crazy from the endless series of failures which can be expected in the daily life which is to be experienced in the laboratory.

"Don't let yourself be discouraged […]! The others are also working themselves to a frazzle without any success; it is quite normal that no results are to be observed in the beginning. Just keep it up!—Or take up a reasonable occupation."

With this in regard …
Cornel Mülhardt

Acknowledgments

1 What Is Molecular Biology?

Gib nach dem löblichen
von vorn Schöpfung anzufangen!
Zu raschem Wirken sei bereit!
Da regst du dich nach ewigen Normen
durch tausend, abertausend Formen,
und bis zum Menschen hast du Zeit.

Yield to the noble inspiration
To try each process of creation,
And don't be scared if things move fast;
Thus growing by eternal norms,
You'll pass through many a thousand forms,
Emerging as a man at last.

–Goethe*, Faust*[1]

Pursuits in the field of molecular biology involve genetic engineering and techniques such as cloning, and they may be interpreted as being quite adventuresome and almost divine. As a molecular biologist, declaring how you spend your days may garner boundless admiration from some people and disapproval from others. Consequently, you must consider carefully with whom you are speaking before describing your activities. You should not mention how many problems and how much frustration you must endure daily, because one group will be disillusioned, and the other will inevitably ask why you continue to work in this field.

I confess that I like the debates about whether we should clone humans, although the discussion often is not engaged at a scientific level. One individual makes a stupid suggestion, and one half of the media and the nation explain why the cloning of humans may not be justified by any means. It seems much ado about nothing, and I cannot comprehend what advantage it would be for me to clone myself. Why should I wish to invest $75,000 for a small crybaby, whose only common ground with me is that it looks like me as I did 30 years ago, if I could instead come to a similar result in a classic manner for the price of a bouquet of flowers for my wife and a television-free evening?

This example reveals the problems that are associated with molecular biology. Incited by spectacular reports in the media, everybody has his or her own, usually quite extreme, ideas on the benefits or disasters related to this topic. In reality, the picture is deeply distressing. Molecular biology, also known as the *molecular world*, generally deals with minute quantities of chiefly clear, colorless solutions. There are no signs of ecstatic scholars who have gone wild in the cinematic setting of flickering, steaming, gaudy-colored liquids. The molecular world deals with molecules, evidence for the existence of which has been attempted in many textbooks, although there is generally little more than a fluorescing spot to be seen on the agarose gel. Every procedure seems to take 3 days, and no Nobel Prize is associated with any of them.

---

[1] This and the following quotations are from Goethe's *Faust* (3rd ed. dtv-Gesamtausgabe, München 1966) and from Goethe's *Faust* (John Shawcross, translator. Allen Wingate Publishers, 1959).

Molecular biology is primarily a process of voodoo—sometimes everything works, but usually nothing works. Very unusual parameters seem to play a role in the result of an experiment, which should represent the subject of the research itself—the last, great taboo of modern science. Based on the empirical data, I have come to the realization that the results of an experiment can be calculated as the quotient of air pressure and the remaining number of scribbled note pages in the drawer raised to the power of grandmother's dog. Until I have been able to verify this experimentally, however, I will continue to limit myself to the classic explanations based on mathematics, which have admittedly enabled this profession to progress quite far.

1.1 The Substrate of Molecular Biology: The Molecular World for Beginners

War es ein Gott, der diese Zeichen schrieb,
die mir das innre Toben stillen,
das arme Herz mit Freude füllen
und mit geheimnisvollem Trieb
die Kräfte der Natur rings um mich her enthüllen?

Was it a God, who traced this sign,
Which calms my raging breast anew,
Brings joy to this poor heart of mine,
And by some impulse secret and divine
Unveils great Nature's labors to my view!

Molecular biology is by no means the biology of molecules, which is better characterized by biochemistry. Molecular biology is the study of life as reflected in DNA. It is a small, closely knit world that differentiates itself from all other related fields, including zoology, botany, and protein biochemistry. Only a few individuals in the field of molecular biology are biologists, and those who are biologists would be most happy to deny it. If you are not a biologist, be assured that you now find yourself in the best of company and that the small amount of material needed to master this work can quickly be understood.

Molecular biology deals with nucleic acids, which come in two forms: deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). The chemical differences between the two substances are minimal. They are both polymers that are made of four building blocks each, deoxynucleotides in DNA and nucleotides in RNA.

The nucleotides of RNA are composed of a base constituent (adenine, cytosine, guanine, or uracil), a glucose component (ribose), and a phosphoryl residue, whereby two nucleotides are connected to each other through phosphoglucose bonds. In this way, one nucleotide can be connected to the next, forming a long chain known as a *polynucleotide*, which is designated as RNA. DNA is formed in almost the same manner. The deoxynucleotides of DNA are composed of a base constituent (adenine, cytosine, guanine, or thymine), 2′-deoxyribose (instead of ribose), and a phosphoryl residue.

More details can be found in textbooks of biochemistry, cell biology, or genetics. I limit my comments on these topics to a few aspects that are significant for laboratory practice.

Deoxynucleotide is a tongue twister, which is avoided whenever possible. Researchers instead speak of *nucleotides* when using these substances in the laboratory, although they generally refer to the deoxy variant.

5′-end

thymidine

ribose

cytidine

2′-deoxyribose

guanosine

adenosine

3′-end
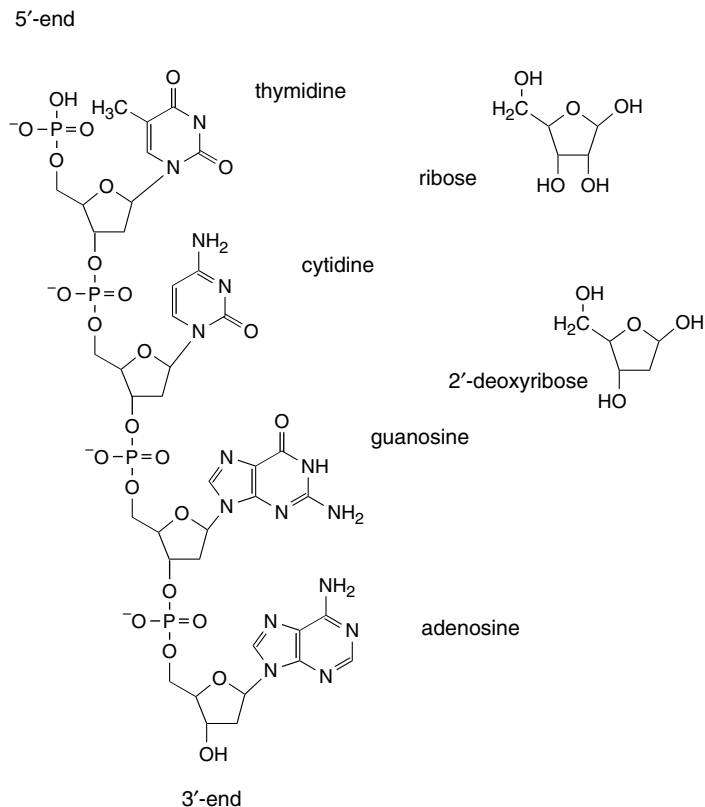
Figure 1-1. The four nucleotides and the way in which a DNA helix typically is portrayed.

The names of the nucleotides are based on the names of their bases: **adenosine** (A), **cytidine** (C), **guanosine** (G), **thymidine** (T), or **uridine** (U). Many confuse these bases with the nucleotides (Figure 1-1), although this rarely makes a difference in practice. The abbreviations A, C, G, T, and U have been used to immortalize all the DNA and RNA sequences known to the world.

The phosphoryl group of the nucleotides is connected to the 5′ carbon atom of the sugar component. The synthesis of nucleic acid begins with a nucleotide at whose 3′-OH group (i.e., the hydroxyl group at the 3′ carbon atom) a phosphodiester bond connects the phosphate group with the next nucleotide. Another nucleotide can be connected to the 3′-OH group of the nucleotide and so on. Synthesis proceeds in the 5′ to 3′ direction, because a 5′ end and a 3′ end exist at all times, and a new nucleotide is always added to the 3′ end. All known sequences are read from the 5′ end to the 3′ end, and attempts to alter this convention can only create chaos.

If no 3′-OH group exists, it is impossible to connect a new nucleotide. In nature, this situation is rarely observed, although it plays a large role in the sequencing performed in the laboratory. In addition to the four 2′-deoxynucleotides, 2′,3′-dideoxynucleotides are used in DNA synthesis. They can be connected to the 3′ end of a polynucleotide in the normal manner, but because a 3′-OH group is lacking, the DNA cannot be lengthened any further. This principle is essential for understanding sequencing, a subject that will be discussed in more detail later.

Each nucleotide or polynucleotide with a phosphoryl group at its 5′ end can be bonded (ligated) to the 3′ end of another polynucleotide with the aid of a DNA ligase. In this way, even

larger DNA fragments can be bonded to one another. Without the phosphoryl group, nothing works, and elimination of a phosphoryl residue with a phosphatase can be used to help inhibit such ligations.

Two nucleotides bonded to one another are known as a *dinucleotide*, and three together are known as a *trinucleotide*. If more nucleotides are bonded together in a group, the structure is called an *oligonucleotide*, and if very many are bonded together, the entire structure is known as a *polynucleotide*. The boundary between oligonucleotides and polynucleotides is not defined precisely, but oligonucleotides usually are considered to contain fewer than 100 nucleotides.

Whether the structure is a mononucleotide, dinucleotide, oligonucleotide, or polynucleotide, the substance always represents a single molecule, because a covalent bond is formed between each of the initial molecules involved. The length of the chain formed in this way plays no role, because even a nucleic acid that is 3 million nucleotides long is a single molecule.

Polynucleotides have a remarkable and important characteristic: Their bases pair specifically with other bases. Cytosine always pairs with guanine, and adenine pairs with thymine (in DNA) or with uracil (in RNA); no other combinations are functional. The more bases that pair with one another, the more stable this combination becomes, because hydrogen bonds are formed between the bases of each pair, and the power of these bonds is cumulative. A sequence that pairs perfectly with another sequence is said to be *complementary*. The nucleic acid sequences of two complementary nucleic acids are completely different and pair in a mirror image form. The following example clarifies this arrangement:

<div align="center">

5′-AGCTAAGACTTGTTC-3′

3′-TCGATTCTGAACAAG-5′

</div>

The orientation of both sequences proceeds in an opposite direction (reading from the 5′ to 3′ end), because the two chains must spatially be aligned in opposition for the pairing to function properly. In the normal 5′ to 3′ method of writing, the complementary sequences are read as follows:

<div align="center">

5′-AGCTAAGACTTGTTC-3′

</div>

and

<div align="center">

5′-GAACAAGTCTTAGCT-3′

</div>

The sequences appear to be so different that you must look very closely to see that they are complementary to one another, and you can see the value of the application of the 5′ to 3′ reading convention. A beginner can very quickly write a primer strand sequence of

<div align="center">

3′-TCGATTCTGAACAAG-5′

</div>

as

<div align="center">

5′-TCGATTCTGAACAAG-3′

</div>

so that the sequence, despite its normal appearance, looks completely different from that of the original sequence.

DNA typically is made up of two complementary chains, and RNA usually is made of a single chain. When pairs occur in RNA, they are found within the same chain and play a more significant role in the three-dimensional structure of the RNA, whereas double-stranded DNA must be considered a linear molecule.

Have you noticed the error? Double-stranded DNA does not consist of a single molecule; it is made up of a double-stranded structure with a chain and an opposing chain. The two chains can

be separated from one another at any time if sufficient energy is supplied; this process is called *denaturation*. Ten seconds at 95°C (203°F) is sufficient for performing complete separation of both strands. If they join together again, a process known as *hybridization*, they are said to *anneal*. The entire procedure is almost arbitrarily reversible.

DNA is frequently called the *molecule of the life* by the yellow press. Why does such a very long, but quite boring molecule have such a lengthy name? Nature has optimally made use of the characteristics of nucleic acids to create a confusing diversity of life, and this is carried out in the following manner. Because the bases pair, a complementary strand can be synthesized from a single strand of DNA; another strand can be synthesized from this second strand, which is identical in structure to that of the initial strand. In this way, DNA can be duplicated as often as desired while preserving the sequence of its bases. The process, which is carried out by special enzymes called **DNA polymerases**, is known as **replication**.

Because RNA and DNA are similar chemically, a complementary RNA molecule can be synthesized from a DNA strand. This process, which is carried out with the aid of **RNA polymerases**, is known as **transcription**.

With an extremely complicated apparatus involving very many molecules (ribosomes), the sequencing information available in RNA can be used for the synthesis of a protein, a process known as **translation**. Proteins are long polymers, whose synthesis proceeds in much the same way as that of nucleic acids—an amino acid exists to which a second amino acid is added, then a third, and so on, until a polypeptide is finally formed, which is commonly known as a *protein*. Twenty different amino acids are known to exist, but with only four bases, how can the information for a protein be hidden in RNA? The trick is that three bases (codons) together contain the information to specify one amino acid. This code is interpreted in the same manner in almost all living cells on this planet: AAA signifies lysine, CAA represents glutamine, and so forth. Together, there are 64 triplet codons (4^3) that code for 20 different amino acids and 3 stop codons; the system is somewhat redundant, and the genetic code is therefore degenerate. For example, AAA and AAG each code for lysine. If you look through all of the 64 triplet codes, you can see that the first two bases are frequently decisive for the particular amino acid that is to be installed, whereas the third base often proves to be insignificant. This arrangement is very useful for molecular biologists, because they can cause mutations in DNA sequences without altering their coding characteristics.

Not every DNA sequence codes for a protein; in human chromosomes, more than 90% of the available sequences have no recognizable significance. The debate continues about whether these stretches should be considered as junk DNA or as elements with some unknown function, but let us instead concern ourselves with the miserable remnants.

The remnants are the **genes**, transcription units on the long DNA molecule that are made up of regulatory and transcribed regions (Figure 1-2). Their functions are indicated by their names. The regulatory region controls whether the other regions are transcribed or not. The transcribed area can be divided into two subregions, the **exons** and **introns**. This division reflects a peculiarity of transcription in higher organisms (in this case, everything that is not a bacterium): **splicing**. Directly after its synthesis, the transcribed RNA is immediately processed by a complex apparatus, the spliceosome; this process excises parts of the RNA sequence and reattaches the remnants, a complicated procedure that ultimately results in a final RNA with a somewhat shorter length than before. The sequences excised during the splicing are introns on the DNA level, and those that remain are exons. The remaining RNA can again be divided into a region that serves as an information source for the synthesis of a protein, also as a coding area or an **open reading frame** (ORF), and one region each at the 5′ and 3′ termini of the RNA, which are known as **untranslated regions** (UTR) and whose function is still being investigated. The 5′ end of the RNA has a methylated G nucleotide (**5′ cap**), and a **poly A tail** (sequence of 100 to 250 adenosines) is located at the 3′ end. The A residues are not encoded on the DNA but are added by the action of poly A polymerase using ATP as a substrate.
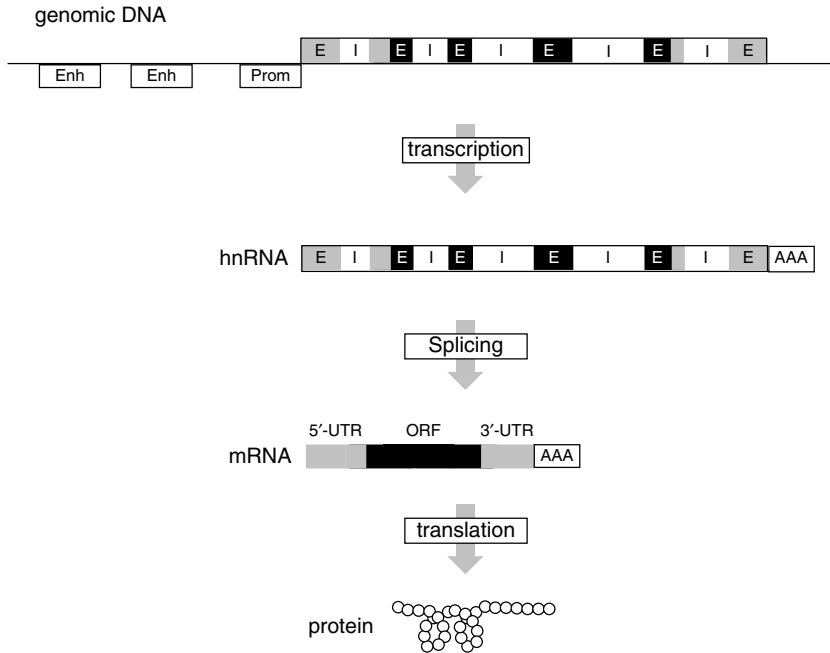
Figure 1-2. The eukaryotic gene and its designations. The gene contains a regulatory region with enhancers (Enh) and promoters (Prom), and a transcribed area where RNA is translated. The heterogeneous nuclear RNA (hnRNA) is a one-to-one copy of the DNA with a poly A tail (AAA) that contains exons (E) and introns (I). The introns are excised while still in the cell nucleus through the action of a spliceosome complex to eventually form messenger RNA (mRNA). Only the open reading frames (ORF) of the mRNA are translated into a protein, whereas the untranslated regions (UTR) at the 5′ and the 3′ termini are in part responsible for the stability of the RNA, in part for their localization, and in part for nothing at all.

The structure signals the existence of mRNA[2] and protects it from attack by exonucleases, thereby increasing the half-life of the transcript, which codes for a protein.

All the regions described are arbitrary. It is currently impossible to discern whether a DNA sequence originates from the exon of a gene or from one of its many junk sequences. Many clever people are working to rectify this situation, and their eventual success will save us all a lot of work.

In molecular biology, experimenters also work with a handful of proteins, including polymerases, restriction enzymes, kinases, and phosphatases, in processing nucleic acids, although the functions of these proteins are not fully understood, which is not absolutely necessary as long as they do function. If this is not the case, the experimenter must usually only order new ones, and this part of the work therefore needs no further explanation.

---

[2] Messenger RNA (mRNA) is one of the three primary classes of RNA. Although it represents only approximately 2% of the RNA in a cell, it does play an important role because mRNA contains the sequencing information for the synthesis of proteins. Ribosomal RNA (rRNA) represents the most important component of the ribosomes, and transfer RNA (tRNA) occurs as small RNA segments to which the amino acids are coupled before they are installed in the protein sequence. Normal cells contain several other types of RNA, which are rarely mentioned, with the exception of small interfering RNA (siRNA) and microRNA (miRNA), which target corresponding mRNAs and thereby participate in post-transcriptional gene silencing.

1.2 What Is Required for This Work?

For a job in a laboratory with the authorization to carry out experiments in genetic engineering, you need three pipettes with which you can transfer volumes between almost 0 and $1000\,\mu L$; everything else seems to be a luxury. Most teams have high demands and little equipment. Even a place to write is frequently not available, or it must be shared with somebody else. The experimenter should be happy with what she or he gets and should try to make the best of it.

Laboratory requirements include a large quantity of bottles for buffers, graduated cylinders, glass pipettes, centrifuges, chemicals, refrigerators, freezers, deep-freezers, polymerase chain reaction (PCR) machines, and scales. Anyone who is not yet a professional in this field should refrain from trying to set up a laboratory, at least if he or she does not want to lose a year of life.

What should the workplace look like? Everybody is happy to see a well-ordered workplace—tidiness is half the battle, as the old adage goes. An alternate approach says that a chaotic workplace, where no staff member is to be found, at least looks as if work were being performed, whereas a spotless, totally empty surface cannot convince a boss that anyone has much enthusiasm for the work.

The people in a laboratory can be differentiated into anarchists and nitpickers. Find out to which group you belong. Because both are determined genetically and can be influenced only to a small degree, do not waste your time trying to change your colleagues. Learn to live with the situation, although you do not have to be satisfied with everything. Some colleagues, for instance, tend to extensively use premixed solutions from others because they are too lazy to prepare their own. Experiments are compromised when the experimenter is faced with an empty bottle of buffer. A popular remedy is to change the labels on the bottles according to the motto "when the label says Tris, there is guaranteed to be no Tris inside"; this is effective as long as the colleague is not aware of the system of re-labeling employed. This method is at best an emergency brake for extremely incomprehensive contemporaries, and it does not help to improve relationships with other colleagues. Worse, experimenters can forget their own systems after the next vacation.

How should a molecular biologist work? Many are in the habit, with great conviction, of upholding the principle of creative chaos. This may work well until the experimenter finds five racks in the refrigerator filled with test tubes that have accumulated over the past 3 weeks, and each of them is numbered consecutively from 1 to 24. You should get used to following a certain system in your work from the start so that you can understand what you have done, especially during some particularly "fervent" phases.

You should get used to a reasonably **clean mode of working**. Nucleic acids are quite stable molecules that fear only one thing: nucleases. Unfortunately, a little bit of nuclease is to be found almost everywhere, beginning with our own fingers. A slovenly worker will soon find a plasmid-DNA preparation that is made up of many small DNA pieces that are no longer useful. Consequently, standard practice is to autoclave everything with which the DNA might have come in contact: pipette tips, reaction vessels, glass bottles, and solutions. Some of this labor can be avoided if you learn to work in a clean manner (i.e., free from nucleases). The plastic reaction vessels, for instance, are practically always free of nucleases on delivery, even if it is not so indicated on the packaging, and they do not necessarily need to be autoclaved if you are capable of tipping them out of the bag without touching them, which may take some practice. The acid test is working with bacterial media. Anyone who is capable of using the same bottle of LB medium for an entire week without any bacteria or yeast growing profusely inside has proved that he or she has learned the tricks of the trade. However, even the most extensive autoclaving is useless if you subsequently rummage in the glass containing the sterile Eppendorf tubes using your dirty little fingers!

Experimenters should become accustomed to using a **laboratory log book**. This is required to maintain an overview, and it is a regulation. Never wait any longer than a week to make your entries because you can quickly forget many of the details. The best thing to do is complete this work at the end of a long, successful workday, perhaps while enjoying a cup of coffee. At the same time, you

can consider what you intend to do the next day. I am well aware that this suggestion is about as realistic as the wish that all people will be good tomorrow.

1.3 Safety in the Laboratory

A molecular biologist is confronted with a whole series of **safety regulations** in the laboratory. Many do not have much to do with biology, and the same guidelines are valid in a chemical laboratory, such as those that primarily deal with the question of safe working methods and with the prevention of accidents. For instance, eating, drinking, and smoking are forbidden in the laboratory, and safety codes demand that laboratory coats; firm, closed, surefooted shoes (i.e., no sandals or no high heels); and a pair of goggles be worn at all times. Normal eyeglasses were allowed in the past, but they are no longer considered adequate. Employers are obligated to make safety goggles with corrective lenses available to their employees, but it is no surprise that such demands from people who spend their days pipetting miniscule amounts of saline and protein solutions from one plastic vessel into another are chiefly ignored. Nevertheless, everyone is urged to conform to these rules.

Experimenters encountering dangerous substances are obliged to wear **protective gloves** for their own protection and for the protection of their fellow workers. For instance, it is not enough to put on gloves when working with ethidium bromide; you must also take them off when you are no longer working with the substance to avoid leaving a wafer-thin film of poison behind on door handles, telephone receivers, and water faucets.

Gloves are essential, but life with them is not easy. The most unsuitable gloves in the laboratory are vinyl gloves because they are difficult to put on, fit badly, and are easily punctured. Latex gloves, which can be found in different sizes and thicknesses, are substantially better. They fit like a second skin, but they are unfortunately allergenic, especially the powdered varieties, and over the course of months and years, most people develop skin problems. If you ultimately develop a severe allergy to latex, even the powder on the gloves of your neighbor may cause you to have problems. The newest design for gloves is the nitrile glove, which is not as elastic as latex and is a little more expensive, but it is not associated with any allergenic risk.

A second group of safety regulations involves **interaction with radioactive substances**. Handling these substances is forbidden unless the particular institution is authorized to handle them. Even with institutional authorization, the experimenter must seek further permission and instruction. Authorities responsible for radiation protection will explain in detail which radionuclides and what quantities are allowed and in which rooms and under the observance of which safety regulations you may work.

Biologic safety is further regulated by the **genetic engineering safety regulations**. These regulations establish that the responsibility lies fully with the project manager, including the administrative aspects and the instruction and supervision of employees. Project managers must complete a scientific or medical university education and must have at least 3 years of experience in the field of genetic engineering. They must have completed a program of continuing education dealing with the topic of genetic engineering, safety precautions, and legal regulations.

Safety regulations for genetic engineering are naturally interesting for the young researcher. They assign organisms that have been altered through genetic engineering to specific **risk groups**. Group 1 generally includes all microorganisms and cell cultures representing no risk, those in group 2 demonstrate a slight risk, those in group 3 pose a moderate risk, and those in group 4 have a high risk for causing harm if they are released into the environment. These designations have been established by scientific and governmental organization. In Germany, classification has been directed by the Zentrale Kommission für die biologische Sicherheit (ZKBS; the Central Committee for Biological Safety), which is equivalent to the American Biological Safety Association (ABSA). *Escherichia coli*

K-12, the forbearer of most strains used in the laboratory, is classified in risk group 1, in the same way as, for instance, the laboratory mouse, *Drosophila melanogaster*, or the CHO, PCL2, or HeLa cell lines. In the higher-risk groups are primarily viruses such as the adenovirus (group 2), hepatitis B virus (group 2), hepatitis C virus (group 3), and human immunodeficiency virus (group 3) and cell lines that can release these viruses.

The risk groups correlate with **stages of safety**, which designate in which type of rooms the respective experiments may be carried out. A laboratory for biosafety level 1 must have four walls, a ceiling, windows, and at least one door, whereby the room can be kept closed. It cannot be a storeroom (because "the work should be performed in sufficiently large rooms and/or regions"), and it must be designated as being a working area for genetic engineering. The list of these regulations is considerably longer. The laboratory, for instance, must have a sink, at which investigators can wash their hands at the end of the workday, after cleaning the room and fighting all the "bugs." Other regulations duplicate previously established ordinances. For instance, the experimenter must wear a laboratory coat and may not perform any oral pipetting. Although eating, drinking, smoking, and applying makeup are forbidden, the storage of food, personal items, and cosmetics is allowed as long as these things do not come into contact with "organisms that have been changed by way of genetic engineering."

Biosafety level 2 represents a more earnest situation in laboratories. There is a warning sign signifying *biohazards*, a disinfectant must stand on the sink, and the doors and windows must be kept closed. No aerosols may be released, work surfaces must be disinfected, and everything that comes into contact with genetically modified organisms (GMOs) must be autoclaved and disinfected. The responsible authorities must also provide protective clothing. Such laboratories may be entered only by those who are explicitly authorized.

In level 3 biosafety laboratories, the windows do not need to be kept closed, because the ordinance anticipates that they cannot be opened. Investigators can enter such a laboratory only by way of a sluice, in which suitable protective clothing is donned. An autoclave must be available, the laboratory must be maintained at a negative pressure, and an individual may not work there alone.

A level 4 biosafety laboratory is more or less a shelter into which those who are weary of life must enter the region disguised as astronauts. The two-page list of official safety precautions is twice as long as that for the other safety stages, and it almost relieves experimenters of the desire to work. There are no level 4 laboratories existing in Germany, but there are four in the United States, with another three under construction. Still, they are rare elsewhere, so a broader description here is superfluous.

Disposal of biologic wastes and sewage is regulated through the safety regulations for genetic engineering. The wastes from biosafety level 1 facilities may be eliminated without any preliminary treatment, as long as no increase in or danger of infection is expected. Beginning with biosafety level 2, such waste must be inactivated or autoclaved. You should make it a habit to autoclave cultures, plates with cell cultures, pipette tips that have been contaminated by DNA, bacteria, and similar by-products of your productive actions before disposing of them to prevent being accused by somebody of having worked sloppily. This approach also corresponds with the rules of **Good Laboratory Practice** (GLP) established by the Organization for Economic Cooperation and Development (OECD). These guidelines also regulate other areas of laboratory life fully (e.g., the method by which experiments are documented), and they are used for well-managed laboratory work in Germany. Anyone who is interested in the GLP can find information in the Internet (http://www.oecd.org/ehs/glp.htm; http://www.glpguru.com; http://www.uic.edu/~magyar/Lab_Help/Ighome.html).

You can learn more about genetic engineering and safety through Internet sources. For example, the Robert-Koch-Institute, Germany's highest authority in this field, publishes useful information on the Internet, such as a list of organisms and vectors, general opinions concerning topics related to genetic engineering, and other information rated by the ZKBS (http://www.rki.de, under the keyword genetic engineering). You also can consult the local commissions for biologic safety.

Literature

Ashbrook PC, Renfrew MM. (1990) Safe Laboratories: Principles and Practices for Design and Remodelling. Chelsea, Michigan: Lewis Publishers.

Rayburn SR. (1990) The Foundations of Laboratory Safety. New York: Springer-Verlag.

Block S. (1991) Disinfection, Sterilization and Preservation. Philadelphia: Lea & Febiger.

Flemming DO, Hunt DL (eds). (1995) Laboratory Safety. Principles and Practices. Washington, DC, American Society for Microbiology.

2 Fundamental Methods

Heute, seh ich, will mir nichts gelingen.

Today, I see, I'm not in clover.

In this chapter, some of the methods described are required almost daily in the practice of molecular biology. For this reason, records must be kept in a more detailed fashion, so that a mixture can be remade in the event of an emergency.

The material in this chapter allows you to see if you have what it takes to be a molecular biologist, because it is not the larger concepts that lead to failure in practice; it is instead the smaller things in the daily routine. There are more than enough ingenious techniques to be found in the literature, but what do you do if the DNA does not precipitate as expected? A researcher with more than one method available will ultimately be the winner.

2.1 Differences in Nucleic Acids

As described in Chapter 1, all nucleic acids are principally constructed in the same manner. Nevertheless, they do differ in some points according to their specific origin, which is of considerable significance for the practice of molecular biology. A DNA chain made up of 500 base pairs (bp) is substantially easier to handle than one 500,000 bp long, and genomic DNA is obtained in a totally different manner from plasmid DNA.

Genomic DNA constitutes the genome, the complete gene complement of an organism. In higher organisms, the genomic DNA is found within the cell nucleus. It is composed of a duplicate set of chromosomes, the numbers of which vary according to the particular organism.

Each chromosome is an individual DNA double strand, which is millions of nucleotides long and must therefore be maintained by many proteins that are bonded to the DNA. For the purification of genomic DNA, the trick is to dispose of these proteins without causing the DNA to crumble too extensively.

The total number of base pairs in a genome varies in different organisms. In humans, it is about 3.2 billion; in the fruit fly, *Drosophila*, it is approximately 180 million base pairs. The size of the genome is not proportional to the size of the organism. The number of base pairs in different mammals, for instance, is more or less equal, whereas some plants can reach considerably higher numbers. Most of the genome is made up of *junk* DNA. The share of important sequences is represented by a maximum of 10% of the DNA, and the remainder appears to be nonsense. Much of it consists of *parasitic* DNA, repetitive sequences and transposons (so-called jumping DNA elements) that increase in our genome over time. Ll elements, the most frequent transposons in the mouse, are estimated to make up 5% to 10% of the murine genome. A first analysis of the human genome showed that the different repetitive elements that we carry make up 40% to 50% of the entire amount of DNA (Li et al., 2001)—that would be 1.6 billion bp that most probably will turn out to be useless (Table 2-1).

Bacteria also contain genomic DNA, although it is constructed differently. It takes the form of an individual, ring-shaped chromosome, which is not localized in the nucleus of the cell but is instead