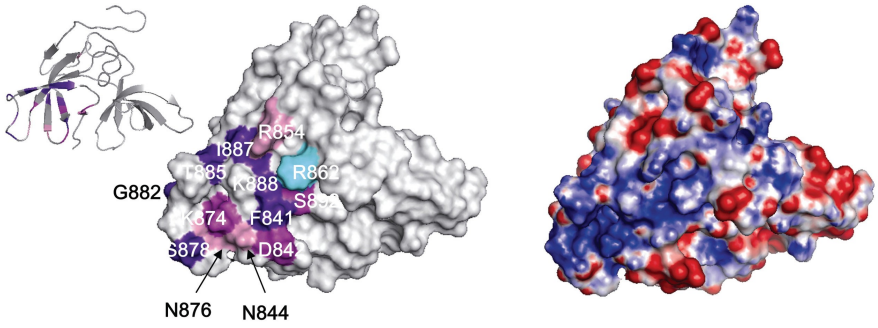
Chromatin and Genetic Differences - When is DNA read?

Karin Lammers* and Jory Sonneveld†
Faculties of Science and Humanities
University of Leiden and University of Utrecht

July 25, 2009



Supervisors: Dr. Ria van der Lecq
Dr. ir. John van Noort

*Email: Karin.Lammers@las.uu.nl
†Email: Jory.Sonneveld@las.uu.nl

Front page image[1]

Left: Biological view of chemicals on chromatin. Right: Physical view of charge distribution in chromatin.

[1]Figure taken from http://www-dsv.cea.fr/en/instituts/institut-de-biologie-et-de-technologies-de-saclay-ibitec-s/informations-scientifiques/actualites-scientifiques/rad9-de-saccharomyces-cerevisiae-contient-un-domaine-en-tandem-tudor-qui-reconnait-l-adn.

Acknowledgements

We would like to express our gratitude to Dr. Ria van der Lecq for her support and her critical comments on the interdisciplinary part of our thesis. We would also like to thank our disciplinary advisor Dr. John van Noort for his inspiration and his enthusiastic, clear explanations.

Abstract

The two disciplines of Physics and Biology have been "cooperating" for a long time in the field of biophysics, but usually use quite different approaches and have difficulties integrating the two views. In this article, various biological and physical models of chromatin (the part of the cell nucleus consisting of DNA and certain proteins), in particular the processes of how and when DNA is read, will be discussed. Biology tends to look at the processes observed by experiments, while physicists try to create models of possible (not always naturally occurring) processes. The problem each discipline tries to solve is how a particular gene on DNA is read; this determines what a cell does and is an important aspect of gene regulation. Of each discipline, we will discuss several experiments or models of chromatin; thereafter we will attempt to integrate the observations from biology with the models from physics. This way we will try to get a more comprehensive answer to the question of how and when DNA is read in cells.

Contents

1 Introduction

People differ from each other not only because they have very different DNA; most differences are a result of certain genes that are "on" or "off" at a particular moment. A gene is a portion of DNA, and if one wants to know whether a gene is "switched on" one actually asks when the DNA will be read.

In this article, a biologist and a physicist will try to find a way to obtain a more integrated solution to the possible ways DNA is read in all our cells. Many biological explanations and physical models exist for this, and at a first glance they do not seem compatible. They seem to put an emphasis on completely different factors in gene expression. We will try to integrate the two perspectives on this problem and in that way attempt to get a more comprehensive picture of the problem.

1.1 What is chromatin?

All cells of human beings contain the same genome consisting of 35,000 genes. We are looking at chromatin, a complex mixture of DNA, and proteins. DNA is wrapped into tight packages by proteins, also known as histones, inside every human cell [16]. In her article "Genetisch Ongelijk", Martine Segers explains that "beads" of histone make it possible for two meters of DNA to be folded into a cell by winding it around "beads" without getting tangled up. Histones are considered the "packing material" of DNA, which makes it possible for the DNA to fit into small packages in each of our cells. See figure 1 of how DNA is folded around proteins into a structure of chromatin.

Little "histone tails" are sticking out of the package. On top of these tails, chemicals can be attached. For this, four chemical groups are most commonly used: methyl, acetyl, the phosphor group, and the protein ubiquitin. These chemical groups can literally block the way for other proteins that "read" the DNA, or it can modify the histone so that the proteins cannot reach the DNA. DNA is tightly packed so that it cannot easily be read, or less tightly packed in order to make it easier to read. A lot of chemicals can be attached, so many types of histone modifications are possible for one bead [16].

A chemical group can be attached to the DNA itself; this does not change the DNA-code but only affects the readability of the DNA. On top of the DNA-code, all the proteins surrounding and influencing the DNA are called the epigenetic code [16].

Research on the subject of chromatin can give important information on various illnesses. When epigenetic encoding is defect in certain cells, many malfunctions can result in the human body; one of these is cancer. There are medicines that stop a certain form of leukemia by stopping certain enzymes from taking acetyl groups off histones [16]. The result is that cancer cells

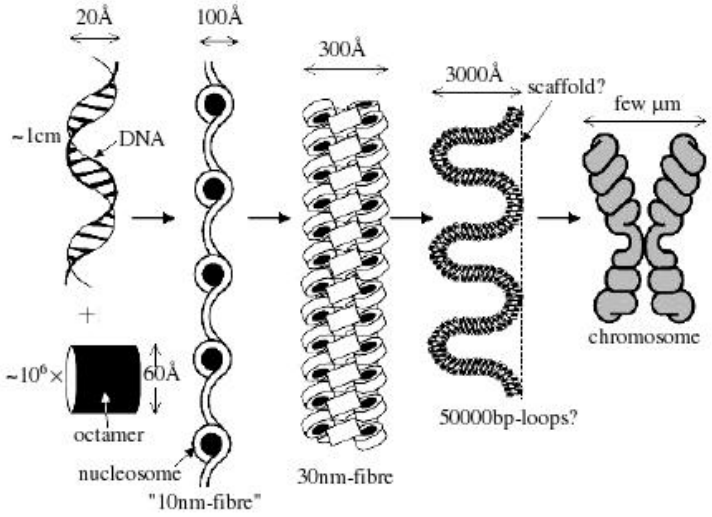


Figure 1: The steps of folding DNA into chromatin [10].

specialize to become harmless cells. In order to understand more defects and ways to cure these, it is important to know the exact effect of epigenetic encoding.

1.2 The field of biophysics

The field of biophysics is defined as *"the branch of biology that applies the methods of physics to the study of biological structures and processes"*[1] by Dictionary.com and as *"physics as applied to biological problems"*[2] by Wordnet of Princeton University. This can then be seen as the physicist's view on biological processes. Biologists and physicists have their own way of solving a problem, and will barely make use of the other possible techniques and methods of the other discipline. Two different explanations of a phenomenon, however, could together give a more reliable answer than one.

One way physicists attempt to answer the questions about chromatin is by treating the DNA cluster as a rubber rod with a certain charge distribution [10]. The DNA will be read only if the cluster unrolls. The higher the elasticity of the DNA, the easier it is to bend and thus unfold. Also, the more the particles of chromatin attract each other, the harder it is to unfold the structure. Furthermore, physicists see cells as a complex system with many moving particles. Although the average energy stays constant, there are many small collisions between the DNA and, for instance, some of the many water molecules in a cell. This is what causes the instability of the compactly folded and therefore seemingly stable DNA structure. For a physicist, then, whether DNA is likely to be read depends on the elasticity

of the structure, the distribution of attracting and repelling charges, and small changes in energy due to collisions of particles.

Biologists try to solve the same problem in a different way. They look at the chemical substances around the DNA because, according to them, this is what determines whether the DNA will be accessible for the enzymes that read the DNA. When genes in a cell are "switched on", a cell produces certain proteins. Biologists look at the structures and shape of the entire chromatin, the way it is modified and its effect on protein production within a cell to determine which genes are "switched on". The physicists' simplified model of DNA and the biologists' view of it as a complex structure are, at a first glance, not quite compatible.

1.3 Interdisciplinary Approach

How this epigenetic structure regulates that certain genes are switched "on" and others "off" is a problem that seems to be easily solved by biologists alone. However, DNA and its packing material consist of more than chemicals and acid. Physicist think that charge distribution and elasticity of strands of DNA are a determining factor in whether DNA can be read [10]. When we look at the epigenetic structure of chromatin from the biological view, knowledge from physics is needed to interpret the experiment. The models from physics give the basic assumptions that make the biological process possible. Physics creates models of the instability of chromatin; without this instability there would not be any modifications or regulation processes for biology to observe. DNA is so tightly packed that it seems impossible for enzymes that read the DNA to reach it. However, because chromatin is so flexible, collisions with water molecules (which are abundant in cells) are enough to cause instabilities.

As an example of why both physics and biology are important in studying chromatin, we can take a look at research done in the neurosciences. For people suffering from depression, medicines have been developed in the field of biology by studying the effect of chemicals on the brain. Drugs have been developed to suppress depressive thoughts. Recently, physics has been used to develop a new method to stimulate or deactivate certain brain regions. Transcranial magnetic stimulation (TMS) uses electromagnetic induction to get electric current (neurons) to flow in the brain [7]. At the moment, research is being done on how this can be used on people suffering from depression. These are two ways of solving the same problem; by trying to solve problems in different ways, one can learn of factors (electric flow of neurons in the brain) that were never considered before; this way scientists get a more comprehensive picture of the problem.

Both biologists and physicists have their own ways of trying to find out what makes DNA more easily read, but these methods are nowhere near similar. Schiessel says in his physical review of chromatin that both of these

disciplines can give an interesting perspective.

> *Undoubtedly chromatin lies at the heart of many essential biological processes, ranging from gene expression to cell division. Most of these processes are controlled by a huge number of specific proteins. Their investigation clearly belongs to the realm of biologists and is beyond the scope of a physicist.* [10]

However, this same physical review elaborates on the physics of chromatin: "*New physical methods like micromanipulation experiments allow us to gain access to certain physical properties of chromatin*" and "*the interest of physicists in single nucleosomes was mainly sparked by the above mentioned fact that the core particle carries a large negative net charge*" [10]. We hope to resolve the differences between biological and physical views on chromatin and give a more complete picture of the process of DNA being read.

2 The Biological Perspective on Chromatin

DNA is the foundation of all our living cells (see appendix A for a list of definitions of terms). It determines if a cell becomes a muscle cell or a brain cell, but it also determines certain characteristics of our bodies. The complete picture of how our genes make up our characteristics, or, in other words, how our genes determine our phenotype, is still uncertain. From this research area there is upcoming proof that not only our DNA determines our phenotype, but that there is a whole range of other influences that also play a part. Certain proteins, also known as histones, may play a great part. Histones are the packing material of our DNA; DNA is wrapped around the histones as shown in figure 2. This is how the DNA is "folded" into small packages that are present in all our cells. Scientists think that the interaction between the DNA and histones, and in particular when they change shape, plays a great part in the accessibility of DNA. This is important, because when DNA is more easily accessed, the proteins that read and translate DNA can do a better job, or even work at all [16].

In most cases, DNA gives instructions to make proteins. These proteins have a whole range of functions within a cell. The instructions given are on heritable material called chromosomes. Chromosomes are structures located in the nucleus; they are consist of DNA and proteins called histones. DNA is a double-stranded, helix-shaped structure composed out of four elements: the nucleotides A, T, C and G. The arrangement of the nucleotides is "the code" for the amino acids, and amino acids combine to form proteins. The human genome has 23 chromosome pairs, each made of a long strand of DNA-molecule (the double-stranded and helix-shaped DNA structure) and protein (histone) that packs the DNA so that it forms a small, compact structure. This whole structure, or the packed DNA, is called chromatin.

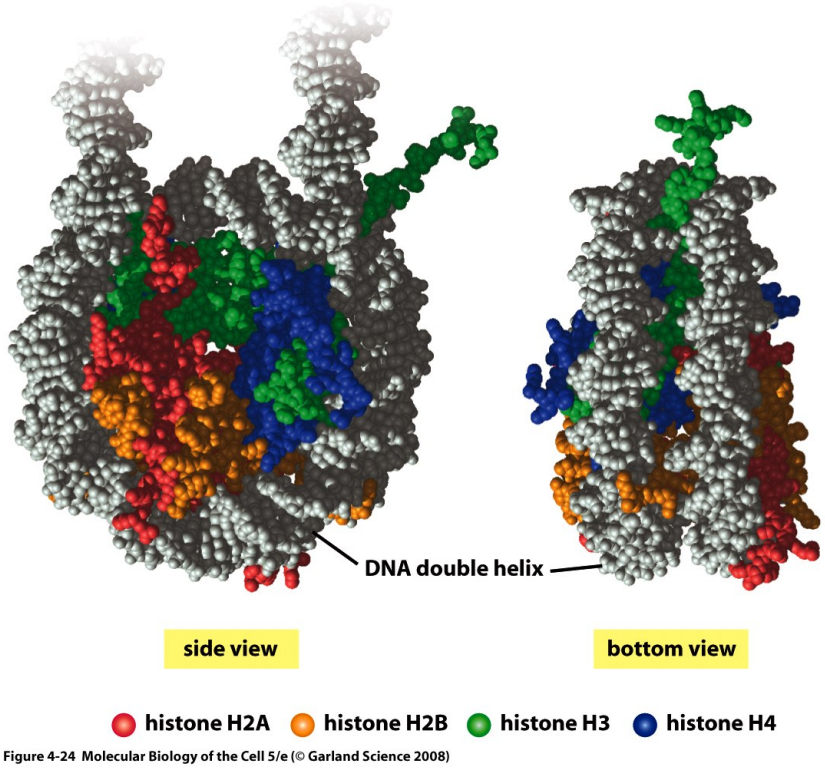


Figure 2: DNA "folded" into small packages by histones [3].

Almost every cell in the human body contains two variants of every chromosome, one from the mother and one from the father (except for the sex chromosomes) [3].

Genes are located on the DNA-strand and are pieces of DNA-strand that contain "the code" for proteins (or the RNA-molecule). Cells do not constantly read the entire DNA. This would cause a constant protein production, which is too much for the cells to cope with. A cell just reads certain genes that produce proteins for the use in that particular cell (or a cell nearby). There then has to be a regulation that holds back certain genes and "reveals" other genes for protein production. Histones and the DNA are regulating this process; which genes are read depends on the shape of the histone-DNA complex [3].

2.1 DNA and RNA polymerase

DNA does not "code" directly for proteins but uses RNA. If a cell needs a protein, a particular piece of DNA is read. Of this particular piece, the gene, a single-stranded copy is made, also known as RNA. Before DNA can be copied, a cell has to make certain preparations. First the double-stranded structure has to be taken apart, so that the piece that needs to be copied is "visible"; this process is also known as transcription. The enzymes (proteins that catalyze a process) that cause the transcription are called RNA-polymerase. This enzyme "walks" on the DNA-strand and "reads" it as shown in figure 3. DNA-polymerase is the enzyme that copies DNA into another double-stranded DNA-strand (not into a single-stranded RNA-strand). Both have the ability to combine nucleotides (the single elements that DNA and RNA are made of) with the template, which is the model of the piece of DNA that needs to be copied. Transcription factors support the whole process by, for example, putting the RNA-polymerase at the beginning spot on the DNA, or by splitting the double DNA-strand [3].

Because DNA is wrapped together with the histones, polymerase needs extra support to be able to read the DNA. This is another example of support by transcription factors; these particular ones are called transcription activators. Transcription activators bind to the DNA and attract RNA-polymerase. Then another protein complex mediates between the activator and the polymerase. Finally, there is a group of proteins that changes the shape of the chromatin (and its histones) so that the access to the DNA is free [3].

The DNA and RNA-polymerase have to get through the chromatin to reach the DNA. On their way they change the composition and shape of the histones. The histones can for example change their position a bit so that the polymerase and the transcription factors can do their work. It could be that RNA-polymerase by itself changes the shape of the histones, but it is also possible it receives help for this job [19].

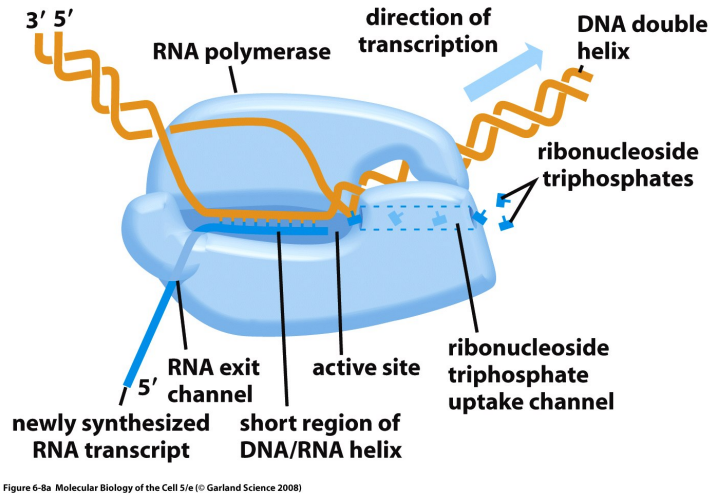


Figure 3: RNA polymerase "walks" on the DNA-strand and "reads" it. [3].

When RNA is produced, it can be converted into a protein. The nucleotides, the lego-pieces of the DNA and RNA, are not the same as the amino acids, the lego-pieces of a protein. There has to be another "code" to enable DNA and RNA to be read and produce proteins. This happens through "codons". A codon is a set of three nucleotides that can be translated to a particular amino acid. Because protein production is not the subject of this article, it will not be discussed in further detail [3].

2.2 Histones and chromatin

The DNA and histones together form a nucleosome. A nucleosome is a histone complex together with a piece of DNA-strand. The DNA-strand is wrapped around a "coin-shaped" complex of histones, as seen in figure 4. When all the nucleosomes are put together, they form a shape that looks like "beads on a string". The "string" in this structure is the DNA, and the "beads" are the histone complexes. This histone complex, the yellow discs in the picture, consists of eight parts.

The DNA-strand is wrapped tightly around this complex, but the structure is dynamic; there can be loops in some places. The histone complex has tails attached to it that are made up of amino acids. These tails reach far out of the DNA-histone complex of the nucleosome. The tails can be modified; this will also change the shape of the DNA-histone complex. When this happens, the DNA comes off the histones, and unrolls a bit; this makes the DNA more reachable for proteins.

In addition, a cell has certain energy-dependent remodelling complexes.

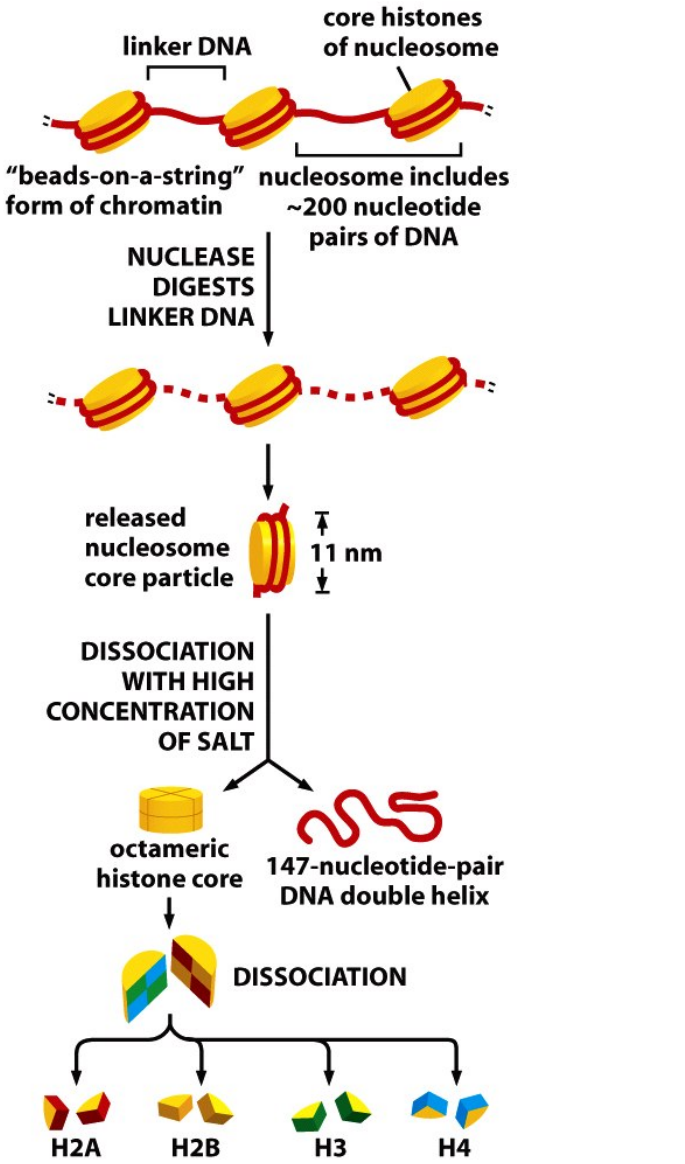


Figure 4: The coin-shaped complex of histones and the "beads-on-a-string" complex [3].

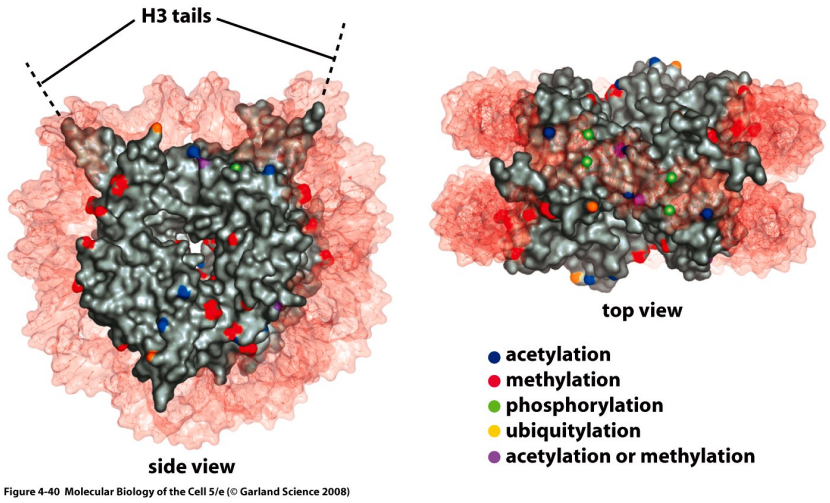


Figure 5: Different molecules that can bind and modify the histone [3].

When genes are turned "on" or "off" (more or less readable), these complexes are working locally on the structure of the chromatin. The proteins that are attached to the DNA can also have an influence on the readability of the DNA. These proteins can attract or reject the histone, which makes the DNA-histone interaction stronger or weaker. In a cell the nucleosomes are close together; this way the chromatin is very small, and the nucleosomes have interactions with each other, i.e. they can attract or reject other nucleosomes [3]. These are all ways to look at the interaction between DNA and its surrounding proteins, and all have their influences on the structure of chromatin and the ability to reach and read the DNA.

DNA-methylation plays a part in gene regulation. It will turn genes "off" for a long period of time by attaching a methyl group to a nucleotide "C". This way the gene cannot be read because the methyl group sticks out; it is literally blocking the RNA-polymerase. DNA-methylation probably evolved to turn "off" viral genes. Viruses can put their DNA in our DNA; in this way it is positive if a cell can "shut down" some particular genes by methylation [3].

Recent studies suggest that histones play a big part in gene regulation, a much larger part than was accepted before. Histones can make portions of the DNA "visible", but they can also "hide" particular parts; in this way, genes are turned "on" or "off". The way histones regulate the gene is still being investigated, but researchers think that local remodelling of the histone can make parts of the DNA more or less reachable [16].

2.3 Histone modifications

The histone structure of a nucleosome is very dynamic. Molecules can bind or come off several parts of the histone, which will cause a local modification. A lot of different molecules can cause this modification, as shown in figure 5. There has been a lot of research on acetylation of the histone, or the binding of an acetyl group, a molecule with "COCH3", to a histone. When this happens, the histone's shape changes so that the DNA will be more reachable. As a consequence, there is a structural modification that makes it possible for transcription factors to reach the DNA. In addition to this, acetylation of the histone will make the chromatin less stable. Because the entire structure of the DNA and all the proteins that surround it is less stable, the transcription factors will have easier access [3].

A lot of transcription activators have histone acetyltransferase, and transcription repressors have histone acetylase. This means that there is a possibility that acetylation of the histone will make the DNA more reachable for polymerase, and deacetylation exactly the opposite. In other words, acetylation can cause genes to be turned "on", and deacetylation can cause them to be turned "off". Histones play their part in gene regulation. Acetylation makes it possible for the transcription factors to reach through the nucleosome [19].

Histones have tails that can reach far out of the DNA-helix. These tails are depicted in figure 6. Research suggests that these tails play a role in the stability of the nucleosome because they mediate between the different nucleosomes. The tails are very dynamic, and because they influence the stability of the nucleosome they can also affect the ability to reach the DNA. Because of this, the interaction between the DNA and the tails is very important for transcription. Acetylation of the tails can probably cause the DNA to bind less firmly to the histone; maybe it could cause destabilization for the entire chromatin structure. However, the overall effect is that acetylated tails have less interaction with the rest of the chromatin, or the DNA and histones [19].

2.4 Influences on transcription

The nucleosome and the complete chromatin structure control the transcription or the reading of DNA. The way DNA is wrapped can promote or obstruct the transcription. Both of the structures are dynamic and can change their shape. This remodeling can have different causes, some of which were described above. Most of these causes are currently under investigation. An example that we have not discussed is the binding of phosphor, which will result in less attraction between the DNA and the tails of histones, or less binding between the histone and the rest of the chromatin. This could cause destabilization of the chromatin, which results in higher accessibility

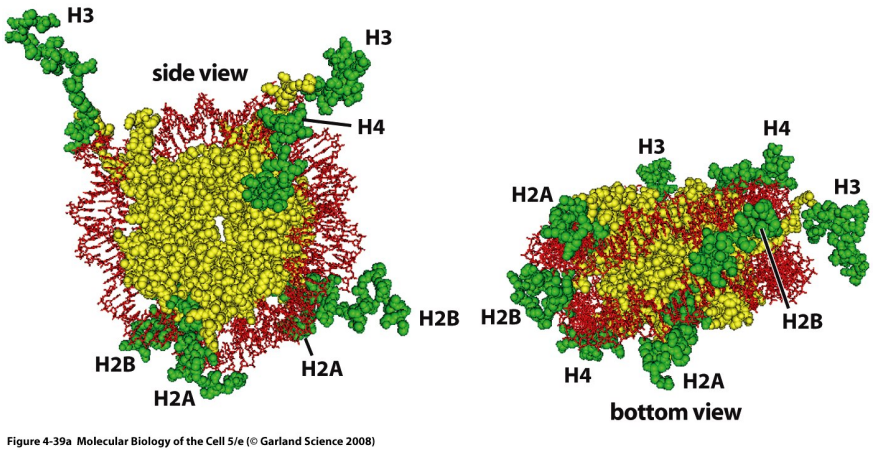


Figure 6: Nucleosome with part of the histone tails; side and bottom view [3].

of DNA [19].

It is possible that the histone composition plays a major part in the existence of cancer cells. For instance, when a histone blocks access to the tumor suppressor gene, a gene that stops cell duplication and brings a natural stop to cancer cells, the gene cannot be read and the cancer can spread. A shortage of methyl groups in the diet (methyl groups are in folic acid and vitamin B12) could increase chances of overweight, diabetes, or cancer in the offspring (this has only been tested with mice). These effects disappear when the offspring has a proper diet with all the needed methyl groups. Perhaps this problem is also caused by histone/DNA interaction, because methyl groups and histones have an important interaction [16].

2.5 Histone Acetylation

For a better understanding on the subject of histone modulation we focus on one example: histone acetylation. A nucleosome is a histone complex associated with a double DNA-strand. The DNA-strand is wrapped around a complex of histones. When one of these histones is acetylated, meaning that an acetyl group is bound to the histone, it can change shape. Deacetylation, when an acetyl group comes off the histone, can have a similar effect. The changing shape that the (de)acetylation will cause makes the DNA more or less reachable. Previous research suggested that the DNA will be more reachable when the histones are acetylated, and less reachable when deacetylated [19]. This (de)acetylation process takes place at the lysine residues of the histone. Lysine is an amino acid that is very common in proteins and is the part of the histone that becomes acetylated or deacetylated [14].

One of the big questions in this kind of research is whether this type of histone modification is important in specific gene regulation. In other words, does histone (de)acetylation turn any specific genes "on" or "off"? This suggests a kind of regulating factor of the histone modification on gene expression. This could mean that certain acetylation patterns may be responsible for specific biological processes. For an answer to these questions researchers looked at eleven lysines in the four histones of the core (the whole histone complex consists of eight histones, of which only four make up the core). By looking at the genetic and biochemical structures of individual acetylation sites, scientists could determine their roles in gene expression. Histone acetyltransferase and deacetylase as coregulator (coregulators are parts that make transcription, DNA replication, and DNA repair possible) in transcription are also a regulatory factor of histone modification [14]. Acetyltransferase is an enzyme that catalyzes the acetyl transfer, and deacetylase catalyzes the process of detaching acetyl groups [3]. In yeast, various acetyltransferases and deacetylases show distinct specificities towards the acetylation sites [14]. The mechanisms by which (de)acetylation affects gene expression are thought to involve two major pathways. In the first, histone acetylation may alter the folding properties of the chromatin fiber and thereby change the accessibility of the DNA. Another way is that the modulated lysines provide specific binding surfaces for repressors or activators of gene expression; this way a single lysine can regulate the binding of proteins to the chromatin [14].

Scientists suggest that there might be a certain "code" of histone modification that relates to gene expression. There is no evidence for a real "code", but researchers did find certain patterns consisting of combinations of hyper- and hypoacetylation of lysines. In other words, not only acetylation and deacetylation are of functional influence, but also the amount of acetylation plays a part. First, it was found that not only hyperacetylation (a lot of acetyl groups attached to lysines of the histone) but also hypoacetylation could cause gene expression. These lysines may provide regulatory factors that coordinate certain groups of genes. Combinations of hyperacetylation and hypoacetylation patterns influence different gene groups. The hyper- and hypoacetylated lysines may provide binding surfaces for certain regulatory factors that coordinate the activity of the different gene groups.

The groups of expressed genes share certain characteristics. They mediate similar biological processes, have some structural similarities, and share a preference for the binding of specific transcription factors [14]. Because these histone modulations are regulating specific gene groups, there may be possibilities for coordinating this regulation - possibly in the form of certain drugs.

2.6 Summary

DNA is "folded" into tight packages by being wrapped around a histone complex. One histone complex together with a part of DNA-strand wrapped around it is called a nucleosome. Chromatin is the entire structure of all the nucleosomes packed together.

A gene is a piece of DNA-strand that can be transcribed into a protein or into a piece of RNA. These genes only come to expression when the gene is read. To be able to read the DNA, certain proteins, like RNA polymerase, must have access to the gene. Histones can have influence on the readability of the DNA.

Modifying histones is one way for proteins to have better access to DNA. Because of this, histone modification is a gene-regulating factor. One way to modify histones is by attaching or detaching certain molecules, like acetyl groups. A consequence of this modification can be a slight change of shape that gives proteins access to read a particular gene.

3 The Physical Perspective on Chromatin

Through changing salt concentrations, temperature, externally applied tension, and the charge of chromatin, various dynamical and mechanical properties of chromatin could be studied in the field of physics [10]. There are various models (i.e. theories) of chromatin, which can be classified into the two categories of models of single nucleosomes (i.e. DNA wrapped around one histone complex) and models of the complete chromatin structures. The entire chromatin structure is also called the "beads-on-a-string" complex; examples of models of this structure are solenoid or zig-zag models. In the models of single nucleosomes, slight changes in energy can cause unwrapping of DNA from the nucleosome, which makes it easier for DNA to be read. Sliding (instead of unwrapping) of DNA is another way of making it more "accessible" to be read. The effect of acetylation of histone "tails" on the mechanical stability of nucleosomes has also been studied in the field of physics.

3.1 Relevant Concepts

Before discussing the most prominent models of chromatin in the discipline of physics, some concepts will be explained in this section.

The nucleosome In physics, single nucleosomes can be seen as a spool of 11 nm wide and 6 nm high (see figure 7) with a large negative electric charge. This is a result of the fact that DNA, wrapped around histones, is more negatively charged than the positively charged histones themselves [11]. More precisely, the "tails" of the histones are positively charged.

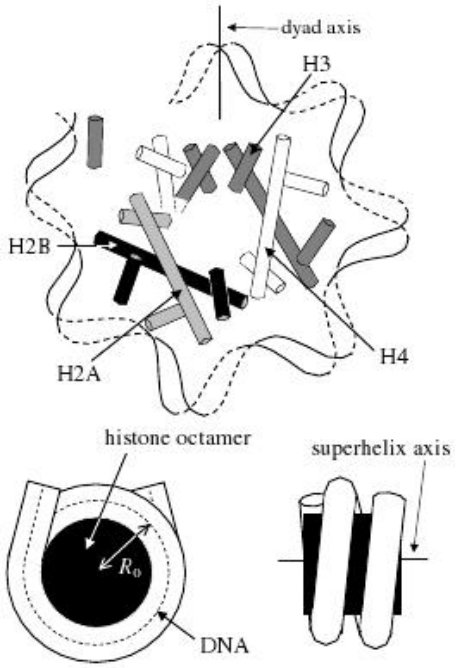


Figure 7: A simplified model of the nucleosome, where the octamer is replaced by a cylinder and the DNA by a worm-like chain [10].

At the center of the histone complex we find an "octamer". This consists of two copies of each of the four core histone proteins (H2A, H2B, H3 and H4; see figure 7) [6]. The "tails" sticking out of the octamer core particle are histone tails. These are rich in the positively charged amino-acid lysine [10], and thus cationic (a cation is a positively charged ion).

The ''Beads-on-a-String" Complex DNA is wrapped around different histones. There are around a 146 base pairs of DNA per nucleosome. The DNA going from one histone complex to the next, or DNA between nucleosomes, is called the linker DNA and is usually about 50 base pairs long. The linker histone H1 binds the DNA to the histone core at the entry and exit points of the DNA. In several studies it was found that when the linker histone H1 is present, the complete structure of chromatin, also called a fiber, is dense and no unwrapping, sliding or transcription through the octamer takes place [17].

The complete complex of chromatin can be seen as "beads-on-a-string", where the histone octamers are the beads, and the linker DNA (the DNA between the nucleosomes) the "string".

Salt concentrations Changing salt concentrations in chromatin is also known as changing ionic conditions (salt, indeed, consists of ions). At high salt concentrations, more "screening" takes place: this means that the electrostatic attraction or repulsion between particles is eliminated or "screened out" for larger distances [10]. Attraction or repulsion are then only active at short distances, or, put differently, they are only short-range.

Thermal Fluctuations DNA can unwrap and nucleosomes can reposition so DNA is easier to read as a result of thermal fluctuations. In a cell, many water molecules bounce off each other and off the DNA. Every second, an amount of collisions of the order of 10^{12} takes place in a cell (see appendix B). A change in amplitude of thermal fluctuations (changes in temperature) occurs when changes in energy take place. This can be seen from the temperature-dependence of the average kinetic energy, or energy of motion, of a free-moving molecule at a certain temperature ($\frac{1}{2}mv^2 = k_BT$, where k_B is Boltzmann's constant and T is the temperature in degrees Kelvin) [4]. The total energy of all molecules together stays the same; however, the energy of individual molecules can fluctuate, resulting in our "thermal fluctuations". These collisions are enough to cause a small part of the DNA to come off a nucleosome. It was recently found that collisions of molecules with chromatin can stretch the flexible chromatin up to one third on top of its original length [13].

Bending Energy DNA has a bending energy which can be measured in experiments by applying external tension [10]. This energy depends on temperature and the persistence length of DNA. Persistence length is a mechanical property describing the stiffness of a polymer. It can be compared with the way a garden hose behaves; a short part seems to be less flexible (and thus straight) than a longer part. 150 base pairs of DNA can be regarded as straight and is the persistence length of DNA. DNA is in contact with the histone at different places. Contact of DNA with the surface of the histone might change bending properties and thus the bending energy [10]. DNA does not bend uniformly around the octamer, which makes it also harder to calculate the bending energy of DNA.

Adsorption Energy Another type of energy to describe a nucleosome is the adsorption energy of DNA. Adsorption is the binding of molecules or particles to a surface. The adsorption energy is, in this case, the energy that is needed to bind the DNA to the surface of the histone (about 6 k_BT per binding site [10]). When DNA binds to the nucleosome, there is an increase in energy of the structure equal to the adsorption energy. The probability of DNA binding to the surface is proportional to the exponent of the adsorption energy (see equation 1, where e is the natural number with a value of about

2.78, k_B is Boltzmann's constant of 1.38×10^{-23} Joules/Kelvin, and T is the temperature in degrees Kelvin) [12].

$$P_{\text{adsorption}} \sim \mathrm{e}^{\frac{-E_{\text{adsorption}}}{k_B T}} \tag{1}$$

One can see from equation 1 that when the adsorption energy decreases, the probability that DNA will adsorb onto the histone protein increases. In the case of chromatin, adsorption energy is largely a result of electrostatic energy.

Electrostatic Energy Potential energy is energy that is stored in a physical system. This energy has "potential" to be converted into, for example, kinetic energy. Electrostatic energy is the potential energy of a charge in an electric field of static charges [9]. The word static here indicates that the electric field does not depend on time. When the charge is moved, the potential energy increases or decreases, depending on whether the charge "wanted" to move that direction already because of attraction or repulsion (in this case energy is released) or force has to be exerted on the particle in order to move it (in this case, work is done *on* the particle and the electric potential energy increases). The electrostatic energy between different nucleosomes would be similar to the positive or negative work done to bring all of them together to their positions from infinity (see equation 2, where W is the work done, q_i and q_j are the charges, and r_{ij} is the distance between them) [9].

$$W = \frac{1}{2} \sum_{i=1}^{n} q_i \Big(\sum_{j=1, j \neq i}^{n} \frac{1}{4\pi\epsilon_0} \frac{q_j}{r_{ij}} \Big) \tag{2}$$

An important finding is that DNA bending energy cancels, for a large part, the adsorption energy of the nucleosome [10]. From these properties, the adsorption energy of DNA can be calculated, as well as the electrostatic interaction energy between different nucleosomes. When the adsorption energy decreases, the DNA unwraps and is more easily read [10]. It, however, still costs energy (new energy will be stored) to bend DNA. We will see this in several models of the single nucleosome.

3.2 Models for the single nucleosome

There are several models to describe the system of the single nucleosome. First, a model where the adsorption energy of DNA and the electrostatic interaction are taken as the main factor for the wrapping structure is discussed. After that follows a description of results of research on changing ionic conditions and its effect on unwrapping of DNA. Subsequently, several theoretical models of nucleosome repositioning which are caused primarily

by thermal fluctuations are discussed. Finally, acetylation of histone tails, the best-studied modification of histone proteins in biology, is discussed from a physical perspective.

3.2.1 Overcharged Chain-Sphere Complex

One model to describe the packing of DNA around a single octamer is the overcharged chain-sphere complex. In this model, the octamer core is treated as a perfect sphere, and the DNA wrapped around it is the chain in the sphere-chain complex. Histone octamers interact with the oppositely negatively charged DNA [10]. DNA is an example of highly negatively charged macro-ions. Counterions (which are in this case ions charged oppositely to the DNA) are released when DNA is adsorbed onto the surface of the histone. This release of counterions (as a result of the system seeking higher entropy [4]) leads to an overcharged sphere-chain complex of chromatin. In other words, the nucleosome has a net charge. As explained above, the electrostatic contribution of the nucleosomes to the energy has to compete with energy cost of deforming the DNA to bring DNA and histone into contact.

Helmut Schiessel discusses several of these simplified model systems of the nucleosome in his article about chromatin, for highly charged and weakly charged cases, and for single spheres and multi-sphere complexes. In each of these models, the optimal wrapping length (length of DNA wrapped around one nucleosome) for a state of lowest free energy was examined. For a so-called iso-electric wrapping length, the system is neutral and has the lowest free energy. As mentioned before, systems are more likely to occupy states with the lowest energy, where this probability is proportional to the Boltzmann factor e^{-E/k_BT}.

The model of the overcharged chain-sphere complex is not a very reliable one. The models of iso-electric wrapping length and larger wrapping lengths do not occur in nature [10]. Besides that, many of the studied cases had used a screening length not found under physiological (i.e. consistent with the normal functioning of an organism) conditions. A weak screening was assumed and the bending of DNA was not taken into account. Also, the so-called lamellar arrangement of the wrapped chain assumed in these models is not a good model for wrapped DNA and neither are the perfect spheres for the proteins [10].

3.2.2 Unwrapping Transition Due to Changing Ionic Conditions

The point at which DNA starts to unwrap from the histone core is called the unwrapping transition. Research has been done on the effect of salt concentrations on the way DNA is wrapped around the octamer. The nucleosomal complex is only stable at intermediate salt concentrations [10]. Ionic strength refers to the concentration of ions in a substance. Taking

into account the bending energy of the chain, the electrostatic attraction between single DNA molecules and the sphere (the octamer core), and the monomer-monomer repulsion, it is found that only a small fraction of the chain of DNA is wrapped for a vanishing ionic strength (i.e. low salt concentrations) [11], [20]. As salt is added (i.e. increasing ionic strength), the nucleosome loses rotational symmetry, which is gained again as a consequence of more and more wrapping after much more salt is added, or, in other words, at very high ionic strengths. The chain stays nearly completely wrapped up to very high salt concentrations; at higher ionic strength (meaning higher salt concentrations), the chain unwraps because of screening of chain-sphere attraction. Here, the ions of the salt are "in the way" of the electrostatic interaction; the electric field of charged particles at larger distances is not felt anymore. Under physiological conditions, high ionic strengths are found, and the interaction between the DNA-chain and the octamer is short-range. The histone core particle is thus unstable at very high and very low salt concentrations, which makes it more susceptible to unwrapping under physiological conditions where salt concentrations are high. These experimental results were, however, based on short chain length; chains of arbitrary length were not included.

3.2.3 Nucleosome Repositioning

Nucleosome repositioning is a modification of the nucleosome within a short region of the genome so that DNA-binding proteins can gain access to their attachment sites [6]. Nucleosomes move up so that a part of DNA is exposed, then called naked DNA. In other words, if one looks at chromatin as the beads-on-a-string complex, a part of string between two beads is longer after repositioning. Under certain conditions (such as higher temperatures: repositioning is more likely to occur at 37 ℃ than at 4 ℃) nucleosome repositioning occurs spontaneously through thermal fluctuations [10]. This can be seen in experiments as well as from calculations in theoretical models. In experiments, it was found that nucleosomes can reposition through sliding and rolling of DNA without breaking contact of DNA with the octamer [10]. It was also found that nucleosome mobility is increased when histone tails (or the linker histone H1 [17]) are absent in chromatin.

In repositioning, not all binding sites (where DNA is in contact with the octamer) break contact at once [10]. Loop diffusion is an example of a mechanism that requires breakage of only a few contact points. There are two kinds of loops with different energy and repositioning dynamics: small loops (or bulges) and big loops. Another model of nucleosome repositioning is corkscrew motion, or twist diffusion, of small defects requiring breakage of only one or two contact points at a time.

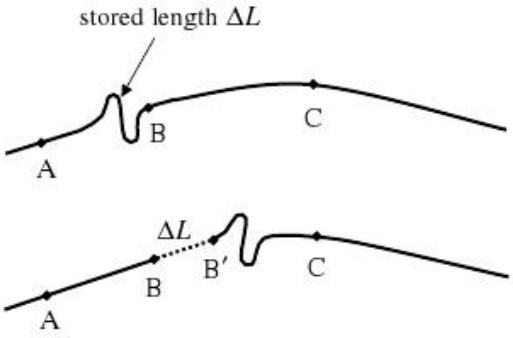


Figure 8: When a loop of length ΔL passes through a piece of DNA-chain at point B, that piece of DNA-chain is moved a distance ΔL to point B′ [10].

Bulge or Small loop diffusion Repositioning of nucleosomes without breaking contact with the DNA chain might be possible through small moving intranucleosomal loops. These loops cause DNA molecules to move by a distance of the length of this loop in way similar to how a caterpillar moves (see figure 8) [10]. Thermal fluctuations lead to the unwrapping of pieces of DNA chain at the ends of the adsorbed DNA chain. Some length of linker DNA can be pulled in before the piece is readsorbed, leaving an intranucleosomal loop. The rate of diffusion of loops depends on the stiffness of the chain (i.e. its bending energy), the adsorption energy per unit length, and the size of the protein (more precisely, the radius of curvature of the DNA centerline) [10].

Large loop repositioning What happens with small loops, also happens with larger loops in structures where loops are more separated and longer "caterpillar" movements are made [10]. By calculating the energy of such a loop and minimizing it, one can find the ground state (state with the least energy) loop for a given "loose" length of chain. Structures in nature are most likely to be found in a state of lower energy (see equation 1) [4]. Loops with smaller energy have simple uncrossed structures [10]. The optimal shape changes to a crossed type when the excess ("loose") length of the loop is of the order of 500 Ångström, or 50 nanometer (a nanometer, abbreviated nm, is 1×10^{-9} m), which is called the critical point. As a comparison, the nucleosome has a diameter d of about 11 nm and a circumference of $\pi d \approx 3.14 \times 11\text{nm} = 34.54$ nm. One can expect the energy to be minimum and loops to be uncrossed for lengths shorter than the persistence length and crossed when the excess length exceeds the persistence length of the DNA (remember that the persistence length is the length for which the

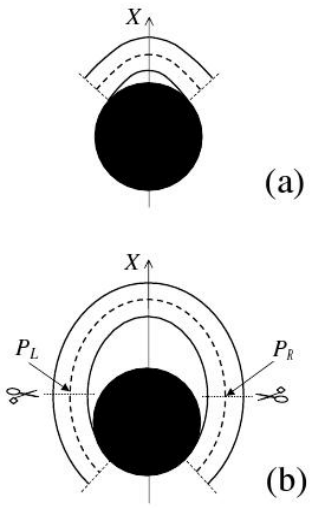


Figure 9: Loops with a length lower than the critical excess length increase energy when length is added (a). The energy of loops larger than the critical excess length decreases when length is added (b) [10].

DNA is more uncontrollable and starts to fold similarly to a garden hose). Then the maximum energy would occur at the critical point; however, the excess length at the critical point is considerably less than the persistence length of DNA. This is because of the shape of the histone around which the DNA is wrapped. A loop with a length below the critical point has a radius of curvature less than the radius of the histone, so that more length will decrease the "gap" and increase the energy (see figure 9 (a)). Loops with a length above the critical point have a much larger radius of curvature and will decrease energy when more length is added, because for these lengths the adsorption energy starts dominating over the bending energy (see figure 9 (b)) [10].

For shorter segments of DNA the repositioning takes place on a scale of hours, whereas for intermediate segments this scale is of the order of seconds to minutes [10]. If the starting position of the loop is at the end, where the DNA enters the histone core particle, large jumps are preferred, and the nucleosome is likely to jump to the other end, where the DNA leaves the histone core. Small jumps take place as well, but less frequently. If the starting position is in the middle, small loops are required initially and the relaxation process of the initial position of this loop is slower than when loops start at the entrance or exit point of DNA. Gel electrophoresis is one type of experiment in which repositioning of nucleosomes can be followed and in which can be seen how this evolves with time.

Twist diffusion A one base pair twist defect is one missing or one extra base pair. If this defect, formed through temperature fluctuations, forms at one end, and then "migrates" to the other end, the entire nucleosome has moved up like a corkscrew by one base pair and turned around its axis by an angle of 36° [10]. The DNA chain can be seen as a chain of particles connected by harmonic springs in a potential that varies periodically in space. We have elastic energy of the DNA, which can be calculated using the harmonic springs, and the adsorption energy of the contact points and is here replaced by a spatially periodic potential. Minimization of these energies shows that defects are likely to move along the chain [10].

Discussion of Models of Nucleosome Repositioning On short fragments of DNA, slow bulge diffusion (timescale of hours) or fast twist diffusion (timescale of seconds) can take place. Under physiological conditions, there is not enough free DNA for large loop repositioning to take place [10]. Repositioning depends on thermal fluctuations and adsorption strength (which is affected by, among other things, the ionic conditions). Twist diffusion slows down to a timescale of hours when the DNA induces strong rotational positioning of the nucleosome. For homogenous DNA, the corkscrew motion is predicted to be the faster one. For long fragments, large loop repositioning is possible and predicted to be much faster than bulge diffusion. There is, however, not much experimental evidence on large loop repositioning [10].

In experiments with two specific nucleosome families (SWI/SNF and ISWI) it seemed likely that one of them (the former) uses large loop repositioning [10]. In the experiments with the first family (SWI/SNF), disruption of many contacts (points where DNA is adsorbed onto the octamer) and transfer of an octamer to another DNA was observed, whereas in experiments with the other family (ISWI, or imitation SWI) there seemed not to be any interruption of nucleosome-DNA contacts even though repositioning did take place.

RNA polymerase is also known to cause nucleosome repositioning. It transcribes DNA through nucleosomes, moving the nucleosomes in the opposite direction of transcription without disrupting the structure. The backwards direction of repositioning could be explained by a model that assumes RNA polymerase crosses the nucleosome in a loop. This loop would have a different shape than the one discussed before. The corkscrew motion of a loop with the shape of this model causes RNA polymerase to get stuck, so it will not be seen in *in vivo* experiments [10].

3.2.4 Histone Tails

The influence of histone tail modifications on the readability of DNA has been mentioned before in section 2. Brower-Toland *et al* studied the acetylation of histone tails and the resulting mechanical stability of nucleosomes.

They found that acetylation causes the positive charge of the tails to decrease [5]. One then expects the binding strength of tails to negatively charged DNA to decrease.

Schiessel says in his article that transcriptionally active regions show acetylation of core histone tails [10]. He suggests that this might overlap with the more general principle of charge neutralization. Brower-Toland *et al* indeed found that acetylation of histone tails decreases overall affinity of histones to DNA in nucleosomes [5]. They also found that an increase in acetylation of histone tails decreases nucleosome stability.

Brower-Toland *et al* argue that two domains of each histone protein play important roles in the nucleosome. The first is the globular domain, which is important for the foundational architecture of the nucleosome. The domain of "tails" has a more dynamic and complex role; the tails are less structured than the globular domains and are thought to interact with DNA in a way that depends on the charge [5].

Brower-Toland *et al* used three "signatures" to judge the stability of the nucleosome: the amount of outer turn DNA, the amount of inner turn DNA, and the peak force [5]. The amount of outer turn DNA is detected at low force and determines how tightly the outer turn DNA is wrapped around the histone core. The amount of inner turn DNA gives a measure of the locations of the off-dyad strong interactions. Off dyad strong interactions are interactions located about 36 base pairs away from the dyad axis of symmetry [2]. The peak force is a measure of these strong interactions, and is indicative of the bond strength. The amount of inner turn DNA and the size of the peak force are detected at high force.

Upon removing the histone tails, outer turn DNA was reduced by 60 percent, the inner turn DNA was minimally affected, and the peak force was significantly reduced, which shows that histone tail removal decreases overall affinity of histones to DNA in nucleosomes [5]. This is similar to a decrease in the adsorption energy of DNA. Brower-Toland *et al* found that results of acetylation of histone tails resembled the results of histone tail removal, and that this modification also decreases overall affinity of histones to DNA in nucleosomes.

Removal of histone tails completely cancels the electrostatic interactions due to the high positive charge of histone tails and the negative charge of DNA. Histone tail acetylation eliminates only interactions at specific lysine residues. Brower-Toland *et al* speculate that the drastic changes seen after several acetylations are due to "critical residues" which directly affect nucleosome stability [5].

Acetylation decreases the resistance of the nucleosome to mechanical folding, and decreases the average amount of DNA stably bound to the histone protein, Brower-Toland *et al* argue. Some transcription regulatory

[2] A dyad is symmetry about a two-fold axis [15].

factors bind to specifically acetylated histone tails; the modification of histone tails which leads to charge neutralization thus exposes histone tails for transcription [5].

Brower-Toland *et al* say that much of the work on chromatin dynamics has used a single nucleosome and usually studied them under conditions not (completely) matching ideal physiological conditions. This was the case of several of the models on the single nucleosome discussed before [10], as well as some parts of the experiments discussed above of histone acetylation.

3.3 Models for the entire chromatin structure

There is still much controversy about the higher-order structures that are often called and depicted by "beads-on-a-string" [10]. Two models of these structures exist: solenoid and crossed-linker models (see figure 10). In the solenoid model of the complex of chromatin, the chain of nucleosomes forms a helical structure where the linker DNA is bent in between the nucleosomes. In zig-zag or crossed-linker models, straight DNA linkers connect nucleosomes in a zig-zag pattern.

From micrographs, a zig-zag pattern could be seen at lower salt concentrations [10]. This structure became more compact when the ionic strength was raised towards physiological value. Computer simulations and analytical approaches supported crossed-linker models. The solenoid model, often depicted in biology textbooks, could not be excluded for physiological conditions under which it was not possible to identify the geometry of the structure a few years ago [10]. Now, however, single chromatin fibres can be stretched through micromanipulation techniques so that different structures can be seen. With single molecule force spectroscopy ("pulling" on ends of chromatin structures), DNA was found to be folded in a helical structure [13]. With this method, the higher-order chromatin structure could be directly studied under physiological conditions. Another way to study the structure is through using the magnetic susceptibility of a base pair of DNA that can help discern between the two structures as well as the nucleosome axes, which are oriented differently towards the fibre axis in each structure [10].

3.3.1 Two-Angle Model

One example of a crossed-linker model is the two-angle model. The value of the exit angle (when it "exits" the nucleosome) θ of the DNA with respect to the entry angle (where DNA "enters" the nucleosome) is determined at the single-nucleosome level where DNA is wrapped a non-integer number of turns around the nucleosome. The other angle in the two-angle model is the dihedral or rotational angle ϕ between the axes of neighbouring histone octamers (see figure 11). A dihedral angle is the angle between two planes,

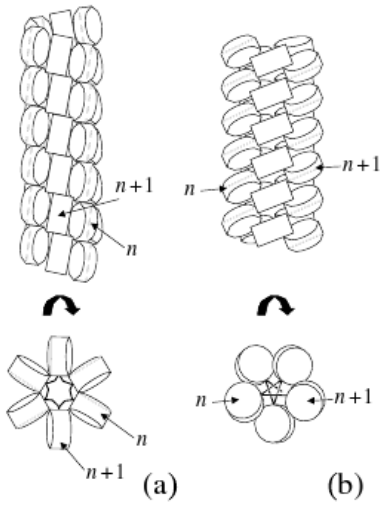


Figure 10: The crossed linker (a) and the zig-zag (b) models. In reality, the structures are more irregular. [10].

which are in figure 11 the planes formed between the outer and middle "rods". There is a preferred value of ϕ. The two-angle model uses the linker length B and the angles θ and ϕ as parameters.

With this model, several mechanical properties of the chromatin structure can be calculated, such as elasticity. It was seen that for small entry-exit angles (θ), the structure has a minimal potential energy at a "swollen state" [10], i.e. the chromatin fiber is swollen and thus more accessible. For larger entry-exit angles, the structure has a minimal potential energy at a "condensed state". Changing θ results in a structural transition between the two. In the condensed state, an external stretching force can cause the system to

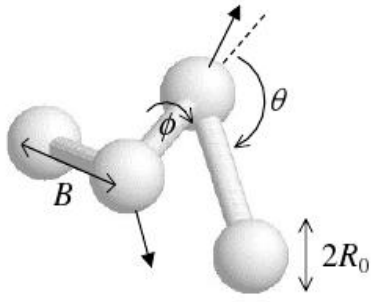


Figure 11: The two angles of the two-angle model. The diameter of the nucleosome is $2R_0$ and the linker length is B [10].

change to a swollen state.

A weakness of the two-angle model is that not all theoretical analyses and calculations of the entire chromatin structure matched the experimental results and physiological conditions [10]. Additionally, Kruithof *et al* found more recently that, unlike what is claimed in Schiessel's article, linker histones do not affect the length or stiffness of the DNA-nucleosome "string" and actually stabilize its folding [13].

3.4 Summary

In physics, there is a large separation between the studies of single nucleosomes and higher order structures of chromatin called the 30 nanometer fiber. The results of higher order structure research are much more controversial than those of the single nucleosome [13], [10]. With single-molecule force spectroscopy it was recently found that (under physiological conditions) chromatin has a helical structure, unlike the other dominating model of a zig-zag structure used in, for example, the two-angle model. Even though the evidence was strong, it was an *in vitro* experiment, which is no proof for *in vivo* situations [13].

For single nucleosomes it was found that nucleosome repositioning, which facilitates reading of DNA, takes place through loop diffusion or twist diffusion. This is, in turn, caused by thermal fluctuations, which are a result of the many water molecules bouncing around and transferring energy in a cell. The energy needed to bend DNA and the energy needed to bind DNA to the surface of a histone protein play a large role in the way the chromatin is folded. The more compactly DNA is folded, the more difficult it will be to read the DNA.

The acetylation of histone tails is also important in the folding of DNA. Electrostatic interaction between the positively charged histone tails and the negatively charged DNA causes tight folding of the structure. Acetylation causes charge neutralization, which results in a more open structure where DNA is easier to read.

4 Conflicts Between Disciplinary Views

The disciplines of physics and biology have some methods, assumptions, terminology, and things that are thought to be important in common. However, the differences in these are more abundant than their similarities.

4.1 Assumptions

To start with, both physicists and biologists, at least in this area of research, assume there is an objective reality, or only one true way the processes around chromatin work. Both of the sciences take the technical aspects and

all the existing information of the research for granted. They think that what you see through a microscope is reality and that DNA transcribes in the way that biology suggests. Evidence comes from experimental data and observations are considered as the truth. Both disciplines accept these data, but use their own terminology to describe this as a result of their different perspectives.

Both disciplines postulate that chromatin is acting within the natural laws of physics; charges repel or attract each other and thermodynamics can be used to describe thermal fluctuations, calculate probabilities, and give a statistical description of a complex structure. In physics, the fundamental assumption of thermodynamics is used, which says that "a closed system is equally likely to be in any of the quantum states accessible to it; all accessible quantum states are assumed to be equally probable" [12].

4.2 Methods and Epistemology

Both physics and biology use roughly the same epistemology. Both sciences are empirical, rational, and experimental. Their methods, however, somewhat differ. Physics makes use of the same experimental methods as biology; however, the theoretical methods in physics are different from some of those used in biology.

In physics, there are various theoretical methods to describe the process of folding and unfolding of DNA. Computer simulations are often used to predict the behavior of chromatin, but also to study in more detail than possible in experiments. Theoretical calculations can be made in various ways: a variational approach, a geometric approach, or a perturbative approach. Perturbation theory is a systematic procedure for obtaining approximate solutions to the perturbed problem by building on the known exact solutions to the unperturbed case [8]. The variational approach is a self-consistent field theory using an analogy to quantum theory [10]. In the geometric approach, the nucleosome was reduced to a point-charge, whereafter it was increased to a perfect sphere with a perfectly smeared out charge to study the electrostatic effects of chromatin [10]. The experimental results from biology are used in theoretical physics. X-ray scattering on core particles and gel-electrophoresis are examples of experimental methods of which the results are used to verify the results of theoretical models.

4.3 Terminology

Both disciplines are not very strict when it comes to terminology. In a physicist's article on chromatin, terms as "the fiber", "monomers", and "nucleosome core particle" or "nucleosome" frequently occurred, without specifying whether histone, chromatin as a whole, or DNA was meant [10]. With "the fiber" or "30-nm fiber" (indicating the size of the structure) physicists mean

the higher-order structure of multiple nucleosomes. The word "monomer" indicated single DNA molecules, and "nucleosome core particle" denoted the "core" of the nucleosome: the octamer core. Nucleosome was used properly to describe the histone with DNA wrapped around it, but also improperly to discuss DNA wrapped around "the nucleosome". From this use of terminology, it can be inferred that it is less important in physics whether linker DNA was meant or a charged rod, as long as the effects could be studied. It seems that biology would be more organized when it comes to terminology. But also in this discipline the terms are frequently used incorrectly or merely not so strictly. Often one can only tell from the context whether the subject is DNA itself, the histone complex, the histone core, the whole nucleosome or even the entire chromosome.

One very important difference, in the subject at focus, is the charge in physics as compared to the chemical groups in biology. Many of these chemical groups that attach to nucleosomes are charged; this is the same "charge" physicists talk about. Although biologists mention the word "charge", they do not pay much attention to this, while for a physicist this may be a subject of more importance. This leads us to a very important difference between the two disciplines of biology and physics: the focus of their research.

4.4 Emphasis and Common Ground

Different things are important to each discipline, and usually completely unimportant to the other. In physics, many models have idealized shapes and circumstances. The histone is, for instance, reduced to a perfect cylinder. This is not a naturally occurring phenomenon, and causes inconsistencies of some predictions based on these models and experimental values [10]. Several physicists create models that are not applicable to natural processes. An example is the number of times DNA is wound around a histone; Schiessel describes this in his model of a highly charged single sphere complex. It might help physicists to understand why the cases of more turns do not occur in nature, but to a biologist it seems to be wasted time to study unnatural phenomena.

Biology and physics have different ways of looking at chromatin. Not only their perspectives differ - for example, the chemical charges as opposed to the chemical groups that are attached to chromatin - but also their scopes of research. For example, the elasticity or the interactions between the atoms is mainly studied in physics, while the shape of DNA and modification of chromatin is important in biology.

When Schiessel discusses the two leading theories of the shape of the higher-order structures of multiple nucleosomes, one can see another difference between the theories of physics and the views of biology. One of the two models is the solenoid model, in which the chain of nucleosomes forms a helical structure and in which the linker DNA is bent to connect neigh-

boring nucleosomes. The other is the crossed-linker or zig-zag model, in which a three-dimensional zig-zag pattern is formed and the linker DNA is straight. Schiessel argues that experiments with two different microscopes (electron cryomicroscopy and atomic force microscopy) show a zig-zag motif that becomes more compact when the ionic strength is raised towards the physiological value. Unfortunately, neither method can help find the structure *at* physiological levels, so the solenoid strucutre cannot be excluded. In contrast, Schiessel writes, "[The solenoid model] is in fact the structure that is depicted in most standard textbooks on cell biology". In other words, in physics the models are seen as highly controversial, and in biology the solenoid model is taken as the true model. Even the more recent article of Kruithof *et al* who used magnetic tweezers and found that the nucleosomes actually form a helical structure does not pose this as a proof that this is the "true" structure; it is merely "strong evidence" [13].

It is important for biologists to take the physical approach into account, but it is not enough at this point to explain the entire phenomenon. An integration of the perspectives of biology and physics looks at biological processes and will try to create physical models that fit. This means that, in creating common ground and regarding the problem at hand, physicists are limited by what exists in nature and biologists try a more mathematical approach than they are used to.

5 An Interdisciplinary Understanding of Chromatin

Physics and biology are already using each other's results despite their little cooperation. Even physicists, who so often claim their discipline is the only true science, need experimental results to compare with their theoretical models. Nevertheless, the results of physics and biology are mostly non-overlapping and on different aspects of biological processes. The results of the two disciplines, specifically in this situation, mostly complement one another.

5.1 Using Results of Other Discipline

The two disciplines of biology and physics are in fact already using each other's results. Physical models or laws can make suitable predictions for biologists. This means that biology can use some models and laws of physics to better understand biological processes. For example, laws of motion and thermodynamics describing thermal fluctuations show that chromatin is an instable structure that will easily be modified.

On the subject of chromatin, the two disciplines need each other. Physics cannot create any models without biological data, and biology cannot understand the whole process without the background of the physical models. There is not much overlap between the two; physicists suggest models to

interpret the data or the underlying mechanisms that make the process possible. On this aspect the two disciplines are not in the same research area. But on some processes they are; an example are the charges (physics) or the chemicals itself (biology) that affect the readability of DNA. The focus here is, of course, on the same charged chemicals. An acetyl group is not only a chemical group that can modify the histone by changing its shape, but also a chemical group with associated charges that interact with the charges of the histone and attached DNA. An example of two kinds of medicine affecting the brain to suppress depression was given in the introduction; perhaps there is a way to integrate these two perspectives so that scientists can come up with a better medicine that takes both the physical and biological/chemical aspects of the depression in account. For chromatin, the same might be true. It has charges and elastic properties, as well as chemical attachments and associated modifications. The natural processes of chromatin are affected by all these properties. Research should take all these aspects into account.

5.2 Complements

The two different views do not contradict but complement one another. Physics is studying the very same chromatin as biology, but merely at a different level. Physics looks at the processes as a result of small disturbances (thermal fluctuations) and electric charges, and biology looks at the substances attached to the chromatin and the influences it has on gene transcription. Both physics and biology study acetylation, but again, physics looks at its consequences for DNA adsorption and biology at the consequences on the level of gene transcription.

In biology, research is done on how various chemical groups placed differently on nucleosomes affect transcription. This could possibly be connected with loop diffusion. A chemical group could cause a loop, and the placement of this chemical group could have an effect on the number of loops made. The placement of the chemical group and loop diffusion are then just two different stages of one and the same process that results in easier-to-read DNA. Loops can exist because of the natural thermal fluctuations, which are only likely to be explained in physics. One could say that this biological process has a physical foundation that makes the process possible.

Of course there is an important issue when we talk about combining the perspectives on chromatin of biology and physics. The question is if all science, including biology, can ultimately be represented in terms of physics. This is not the same as looking for the fundamental issues underlying the biological processes, such as the thermal fluctuations that make modification of the chromatin possible. This a more mathematical approach of biology than is usually used in biological experiments [18].

For physicists, too, there is a difference, because there are limits to nature that physicists do not always take into account. Models can give an

explanation of how a DNA-strand can be wrapped around the histone in a way that does not waste a lot of energy; this does not mean, however, that this is the way it is observed in nature. Nature is restricted by the course of evolution; certain structures may be possible and successful, but do not exist in nature.

6 Conclusion

The models discussed and integrated in this article show how chemical groups, electrostatic charge differences, thermal fluctuations, and other factors affect gene expression. The differing DNA itself is not the only way characteristics of cells are determined; genes can be switched "on" or "off", regulating their expression. What biology and physics try to explain, is why genes in one cell are read, and not in the next. In this paper explanations for this phenomenon from two different perspectives, physics and biology, are given. These different perspectives have their focus on different subjects of the same phenomenon. These different subjects are, in most cases, different stages of the same process. An exception are the effects of acetylation of positively charged tails, which are intensively studied by biologists, and are studied by some physicists as well. In this case physicists look at the electrostatic interactions and their consequences on the folding of DNA, while biologists look at the effects of this folding: whether DNA is approachable for reading. Putting physical and biological explanations together seems to give a more complete picture of the process of how DNA is read. It would be good for both disciplines to take this into account and take more effort to understand each other's work.

Appendices

A List of Terms

Below follows a list of technical terms in used this paper. For most definitions, Henderson's Dictionary of Biology was used [15]. Any differences in biological and physical interpretations are elaborated in section 4.

Glossary

acetyl group	A molecule with COCH3.
acetylase	An enzyme that helps the detachment of acetyl groups.
acetylation	The binding of an acetyl group.
acetyltransferase	An enzyme that catalyzes the transfer of an acetyl group.
adsorption	Adhesion of molecules to the surface of a solid or liquid.
amino acid	One of the elements, the lego-pieces, of proteins.
cation	A positively charged ion.
chromatin	The structure of packed DNA with its associated histone complex that makes up the chromosome.
chromosome	Heritable structure that is located in the nucleus of a cell. A chromosome is composed of double-stranded DNA and the associated histones.
counterion	The ion that accompanies an ionic species in order to maintain electric neutrality.
DNA	Deoxyribonucleic acid, a very large linear molecule which acts as the store of genetic information in all cells.
dyad	Symmetry about a twofold axis.
electric potential	Potential energy per unit of charge that is associated with a static (time-invariant) electric field.
entropy	Thermodynamic term describing the disorder or randomness of a system.

enzyme	Protein that catalyzes a process.
gene	Pieces of DNA or RNA-strand that contain the code for proteins.
histone	Proteins that "fold" DNA into tight packages.
histone octamer	Two copies of each of the four core histone proteins (H2A, H2B, H3 and H4) at the center of the histone complex.
ion	Atom or molecule that has acquired an electric charge by gaining or losing one or more electrons.
ionic strength	Measure of the concentration of ions in a solution.
lysine	An amino acid that is very common in proteins; it is the part of the histone that becomes acetylated or deacetylated.
monomer	A molecule that is the repeating unit of a polymer (here, a DNA molecule is often meant).
nucleosome	One histone complex together with a piece of DNA-strand; this makes up for a "coin shape".
nucleotide	Basic chemical subunit of DNA and RNA; one of the four elements, the Lego-pieces, of DNA.
octamer	See histone octamer.
persistence length	Mechanical property describing the stiffness of a polymer.
phenotype	All the physical characteristics of an organism.
physiological	Consistent with the normal functioning of an organism.
polymer	A large molecule (macromolecule) composed of repeating structural units.
polymerase	An enzyme that "walks" on the DNA or RNA-strand and "reads" it.

potential	Potential energy that is associated with a certain field. See also electric potential.
RNA	Ribonucleic acid; large linear molecule made up of a single chain of ribonucleotide subunits. RNA is single-stranded and important in protein production.
screening	(electric field screening) Damping of electric field caused by mobile charge carriers. Charges of other particles can be screened out by for example high salt concentrations; in this case the electric field of the screened out particles is not felt.

B Amount of collisions in a cell per second [3]

The energy of one molecule in linear motion is equal to $k_B T$ [4], with $k_B = 1.38 \times 10^{-23}$ Joules/Kelvin Boltzmann's constant and T the temperature in Kelvin. Room temperature is around 300 Kelvin (more precisely, $273.1 + 20 = 293.1$ Kelvin). We can set this equal to the kinetic energy of a particle with mass m and velocity v; assuming there is no potential:

$$\frac{1}{2}mv^2 = k_B T \tag{3}$$

From here, we can calculate the velocity of a water molecule at room temperature:

$$v = \sqrt{\frac{2k_B T}{m}} = \sqrt{\frac{2 \times 1.38 \times 10^{-23}\mathrm{J/K} \times 300\mathrm{K}}{18 \times 1.67 \times 10^{-27}\mathrm{kg}}} \approx 525\,\mathrm{m/s} \tag{4}$$

meters/second where the mass is taken to be that of 18 protons. Protons and neutrons have similar masses of 1.67×10^{-27} kg; the mass of electrons (9.1×10^{-31} kg) can be ignored. A molecule of water, or H_2O, is made of 2 hydrogen molecules (most commonly 1 proton, no neutrons) and 1 oxygen molecule (most commonly 8 protons, 8 neutrons).

Now the average time for a collision can be calculated using that

$$v = \frac{dx}{dt} \approx \frac{\Delta x}{\Delta t} \tag{5}$$

[3] For this we have made use of http://www.uic.edu/classes/phys/phys450/MARKO/N003.html, 05/05/2009.

and if we take Δx to be the size of a water molecule (0.2 nanometers) we have

$$t = \frac{\Delta x}{v} = \frac{2 \times 10^{-10}\text{m}}{525\,\text{m/s}} \approx 3.81 \times 10^{-13}\,\text{seconds} \tag{6}$$

which results in

$$\frac{1\,\text{collision}}{3.81 \times 10^{-13}\text{s}} = 2.625 \times 10^{12}\,\text{collisions/second.} \tag{7}$$

References

[1] Dictionary.com unabridged (v 1.1), Mar 2009. http://dictionary.reference.com/browse/biophysics.

[2] Wordnet 3.0, Mar 2009. http://dictionary.reference.com/browse/biophysics.

[3] Bruce Alberts, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, and Peter Walter. *Molecular Biology of The Cell*. Garland Science, fifth edition, 2008.

[4] Stephen J. Blundell and Katherine M. Blundell. *Concepts in Thermal Physics*. Oxford University Press, 2006.

[5] Brent Brower-Toland, David A. Wacker, Robert M. Fulbright, John T. Lis, W. Lee Kraus, and Michelle D. Wang. Specific Contributions of Histone Tails and Their Acetylation to the Mechanical Stability of Nucleosomes. *Journal of Molecular Biology*, 346:135–146, 2005.

[6] T.A. Brown. *Genomes 3*. Garland Science Publishing, third edition, 2007.

[7] A.A. Gershon, P.N. Dannon, and L. Grunhaus. Transcranial Magnetic Stimulation in the Treatment of Depression. *American Journal of Psychiatry*, 160(5):835–845, 2003.

[8] David J. Griffiths. *Introduction to Quantum Mechanics*. Pearson Prentice Hall, 2005.

[9] David J. Griffiths. *Introduction to Electrodynamics*. Pearson Education, Inc, San Francisco, CA, USA, 2008.

[10] Helmut Schiessel. The Physics of Chromatin. *Journal of Physics: Condensed Matter*, pages 699–774, May 2003.

[11] S.N. Khrapunov, A.I. Dragan, A.V. Sivolob, and A.M. Zagariya. Mechanisms of stabilizing nucleosome structure. Study of dissociation of histone octamer from DNA. *Biochimica et Biophysica Acta*, 1351(1, 2):213–222, 1997.

[12] Charles Kittel and Herbert Kroemer. *Thermal Physics*. W.H. Freeman and Company, second edition, 1980.

[13] Maarten Kruithof, Fan-Tso Chien, Andrew Routh, Colin Logie, Daniela Rhodes, and John van Noort. Single-molecule force spectroscopy reveals a highly compliant helical folding for the 30nm chromatin fiber. *Nature Structural and Molecular Biology, Advance Online Publication*, April 2009.

[14] Siavash K. Kurdistani, Saeed Tavazoie, and Michael Grunstein. Mapping Global Histone Acetylation Patterns to Gene Expression. *Cell*, 117:721–733, 2004.

[15] Eleanor Lawrence, editor. *Henderson's Dictionary of Biology*. Prentice Hall, 13 edition, 2005.

[16] Martine Segers. Genetisch Ongelijk. *Natuurwetenschap en Techniek*, 72(4):20–27, April 2004.

[17] S. Pennings, G. Meersseman, and E.M. Bradbury. Linker histones H1 and H5 prevent the mobility of positioned nucleosomes. *Proceedings of the National Academy of Sciences of the United States of America*, 91(22):10275–10279, 1994.

[18] N. Rashevsky. Physics, biology, and sociology: A reappraisal. *Bulletin of mathematical biology*, June 1966.

[19] Alan P. Wolffe and Dmitry Guschin. Review: Chromatin Structural Features and Targets That Regulate Transcription. *Journal of Structural Biology*, 129:102–122, 2000.

[20] Thomas D. Yager, Cynthia T. McMurray, and K.E. Van Holde. Salt-induced release of DNA from nucleosome core particles. *Biochemistry*, 28(5):2271–2281, 1989.