Die Zelle

Eine chemische Fabrik im Nanomaßstab

Peter Schuster

Institut für Theoretische Chemie, Universität Wien, Österreich
und
Österreichische Akademie der Wissenschaften

Bundesgymnasium/Bundesrealgymnasium

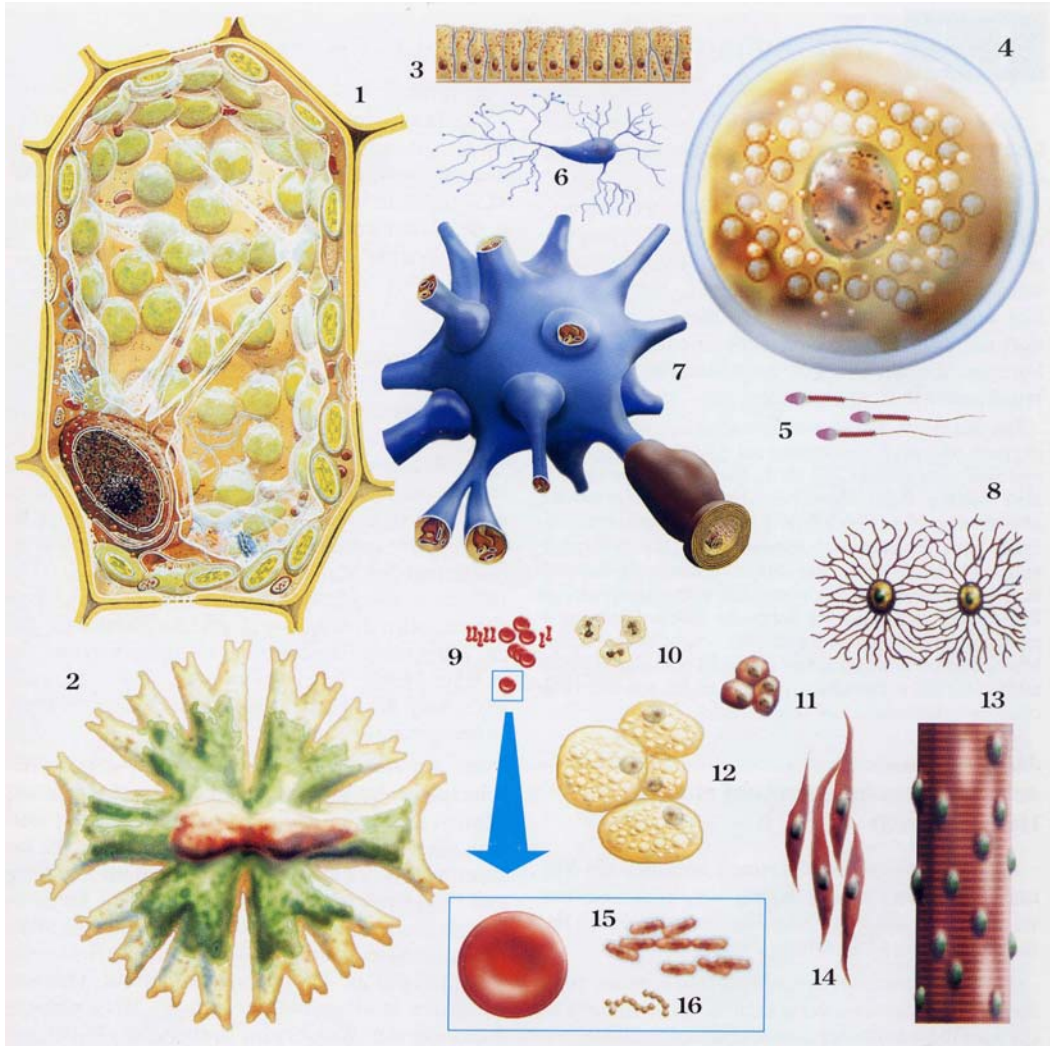
Wr. Neustadt, 10.12.2008

Web-Page for further information:

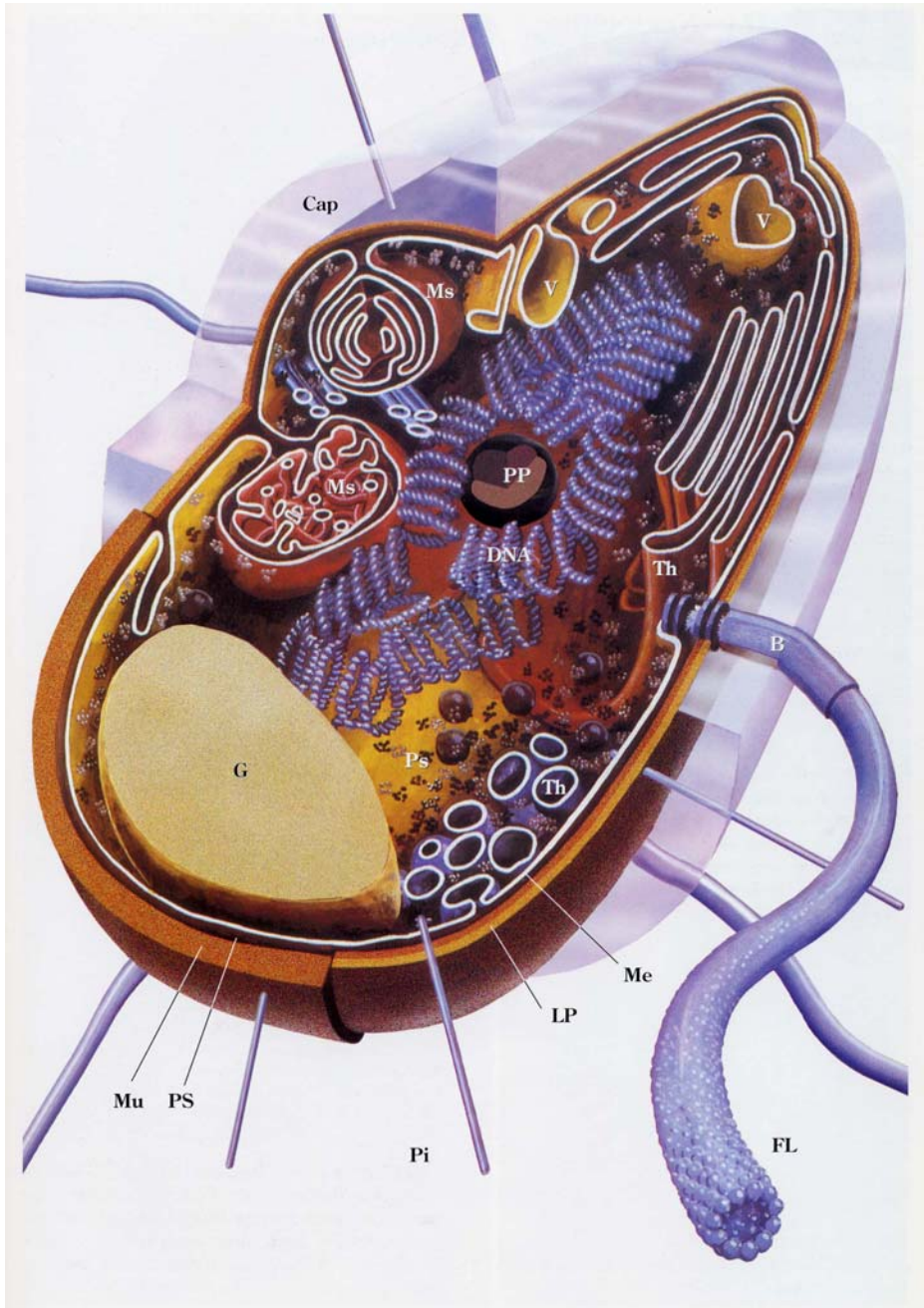
http://www.tbi.univie.ac.at/~pks

Photographs are taken from

Joachim Ude und Michael Koch,
Die Zelle. Atlas der Ultrastruktur, 2.Auflage
Gustav Fischer Verlag Jena Stuttgart, 1994



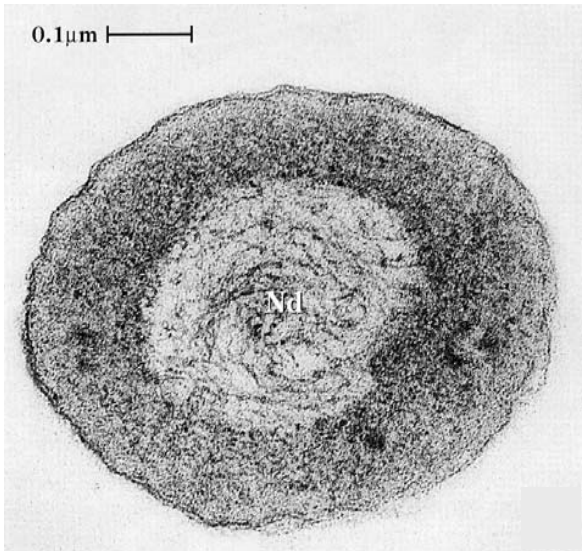
Beispiele für die Größen- und Formenvielfalt von Zellen. In der Natur kommen Größenunterschiede von 1: 500 000 vor. Während manche Mikrokokken nur eine Länge von 0,15 - 0,2 µm aufweisen, gibt es einzelne Pflanzenzellen von einer Länge bis zu 30 cm. Die Axone von Nervenzellen aus dem Rückenmark des Menschen können bis zu einem Meter lang werden. 1 - Pflanzenzelle; 2 - die Grünalge Micrasterias crux melitensis; 3 - menschliche Epithelzellen; 4 - Eizelle des Menschen; 5 - Samenzellen; 6 - Pyramidenzelle der Großhirnrinde; 7 - große motorische Ganglienzelle des Rückenmarks; 8 - Knochenzellen; 9 - Erythrozyten; 10 - Granulozyten; 11 - Leberzellen; 12 - Fettzellen; 13 - Abschnitt einer quergestreiften Muskelfaser; 14 - glatte Muskelzellen; 15 - Stäbchenbakterien (stark vergrößert); 16 - Mikrokokken.

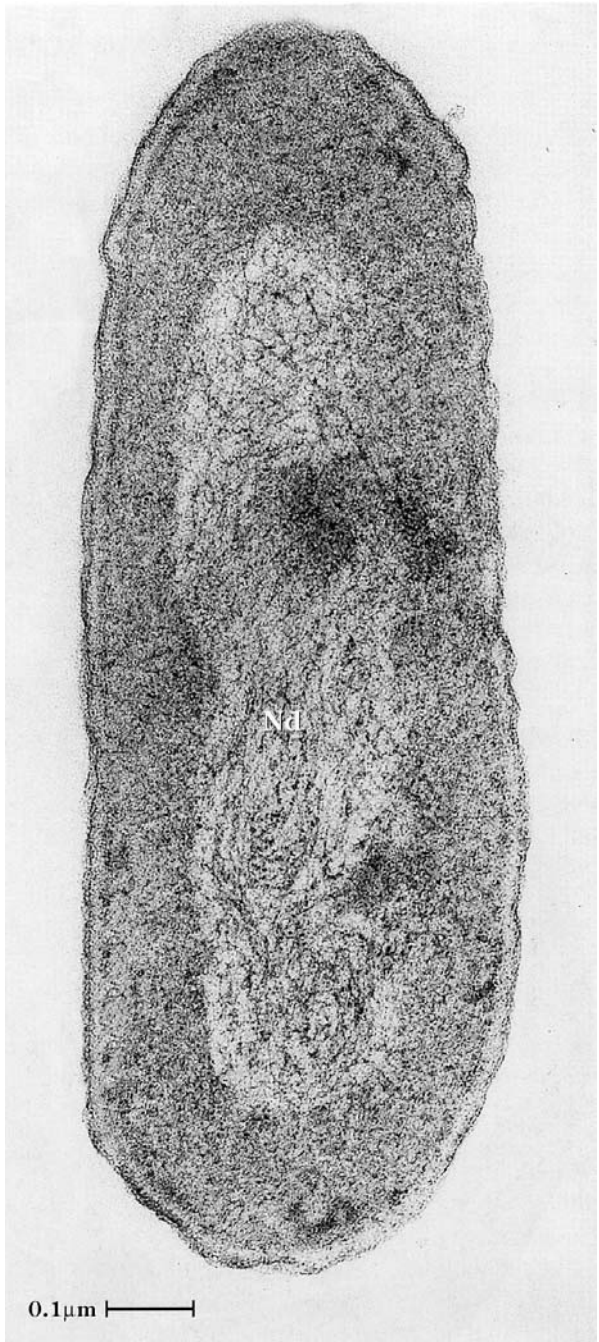


Die prokaryotische Zelle

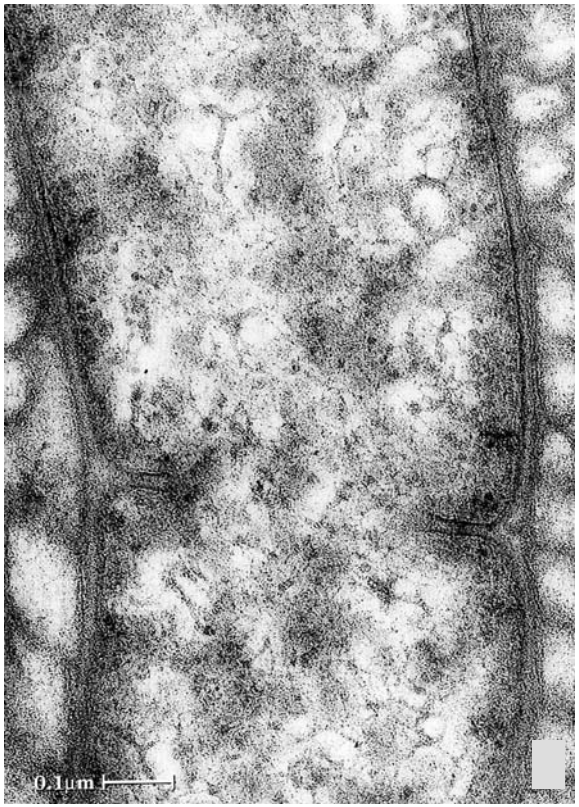
Raster-Elektronenmikroskop (***Scanning electron microscope***, SEM)

Transmissions-Elektronenmikroskop (TEM)

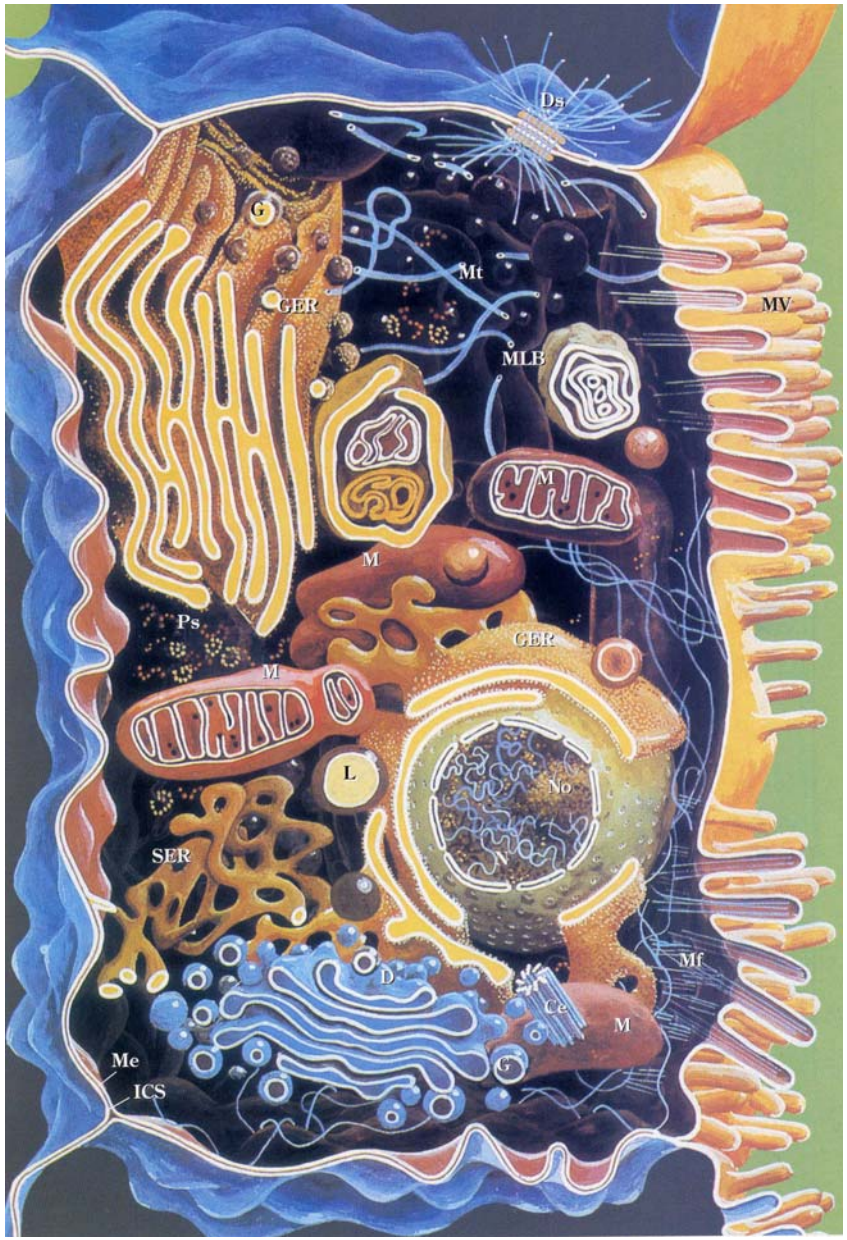


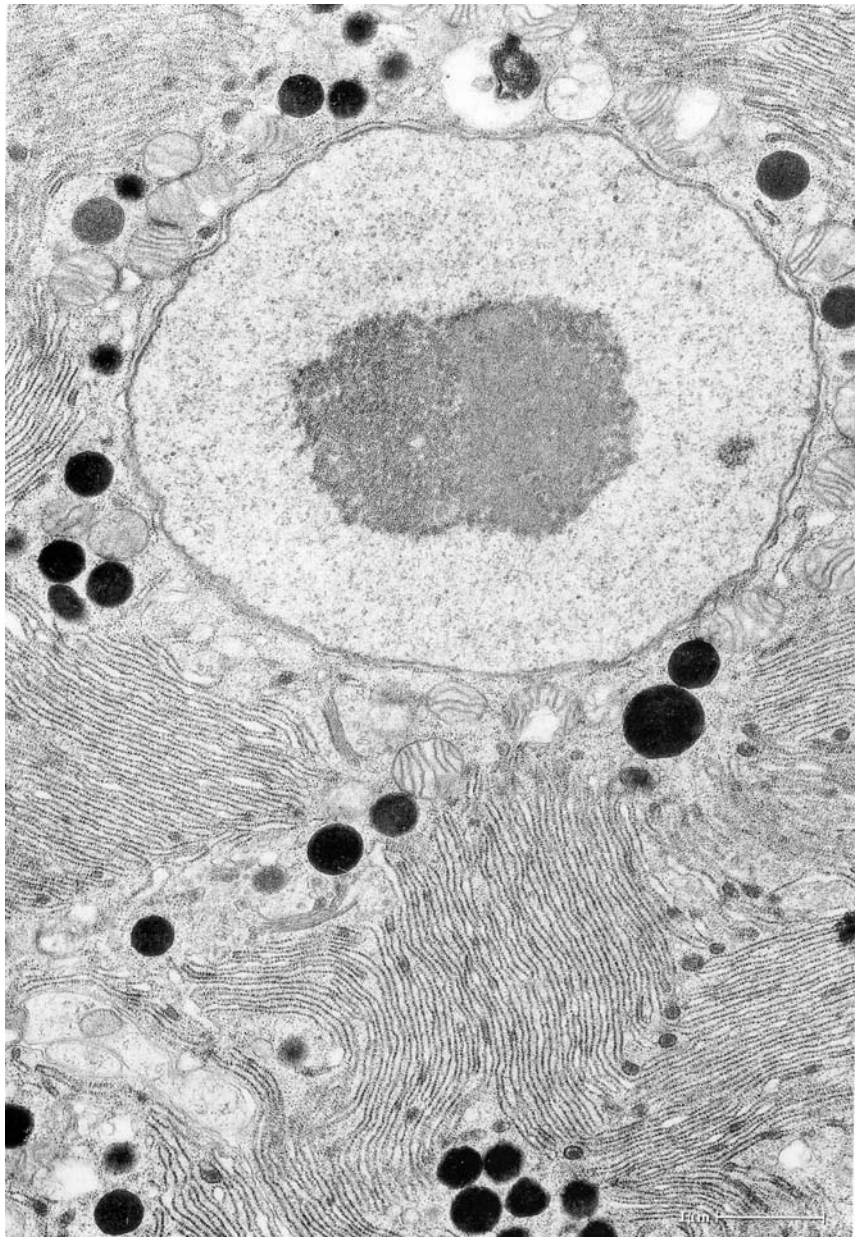


Elektronenmikroskopische Schnitte durch
Escherichia coli Bakterien

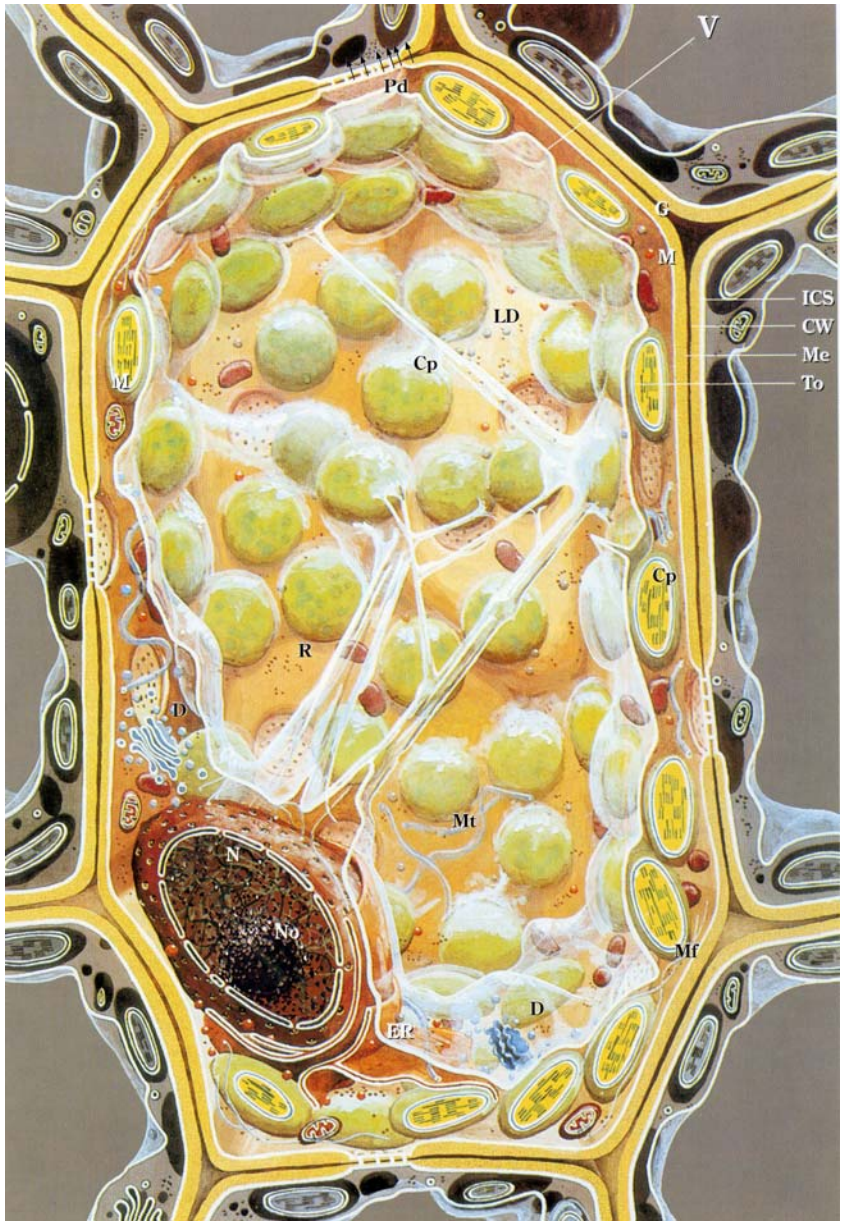


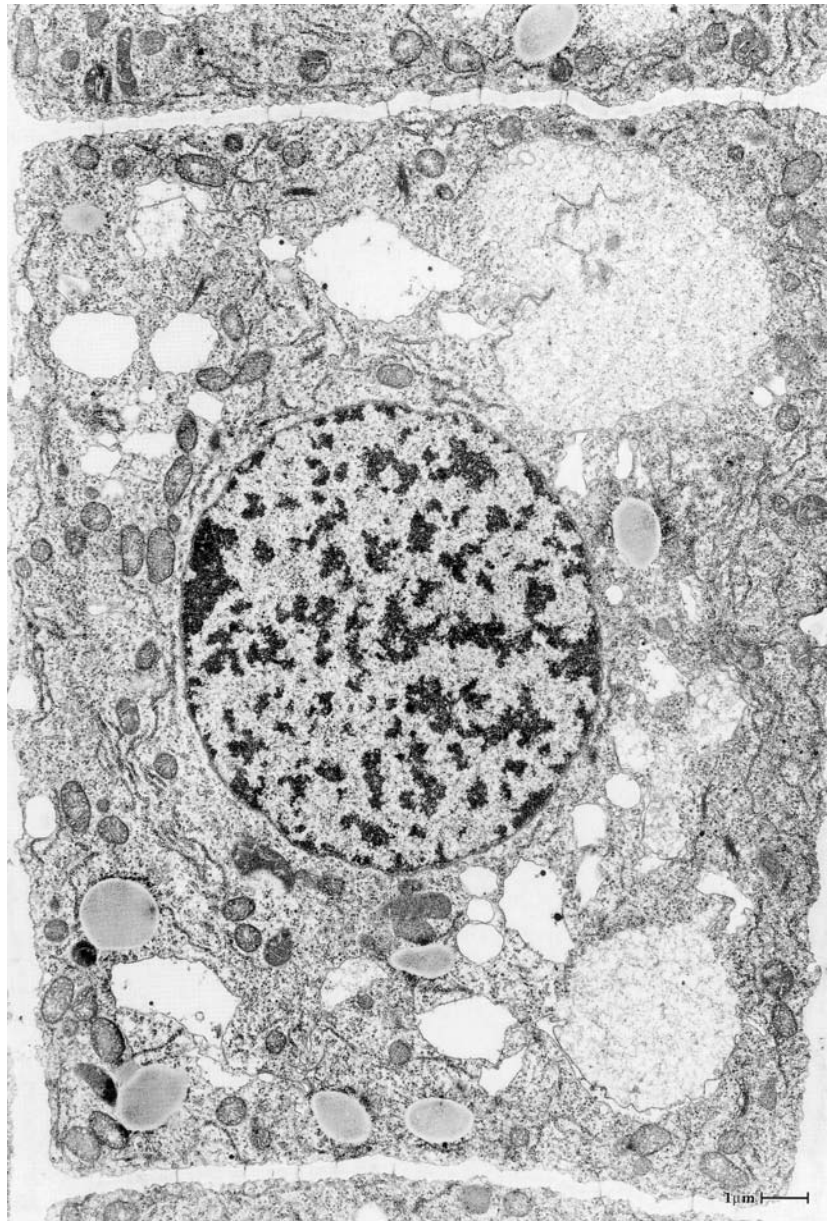
Zellteilung in ***Corynebacterium periplanetae***



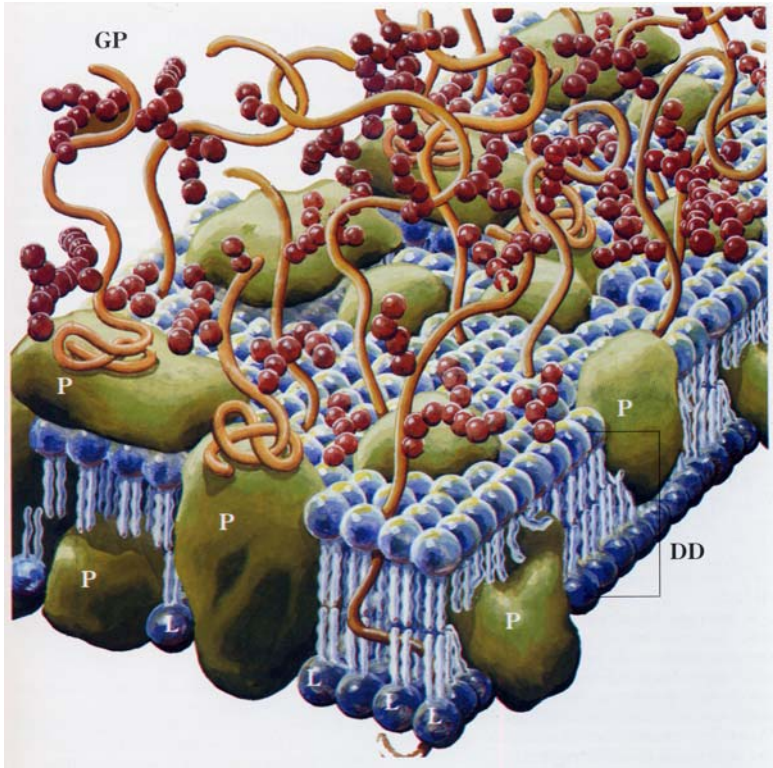


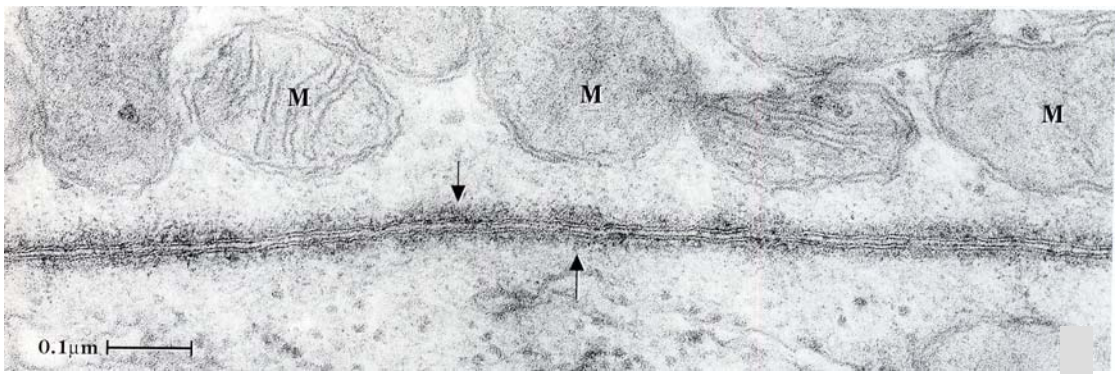
Die tierische eukaryotische Zelle

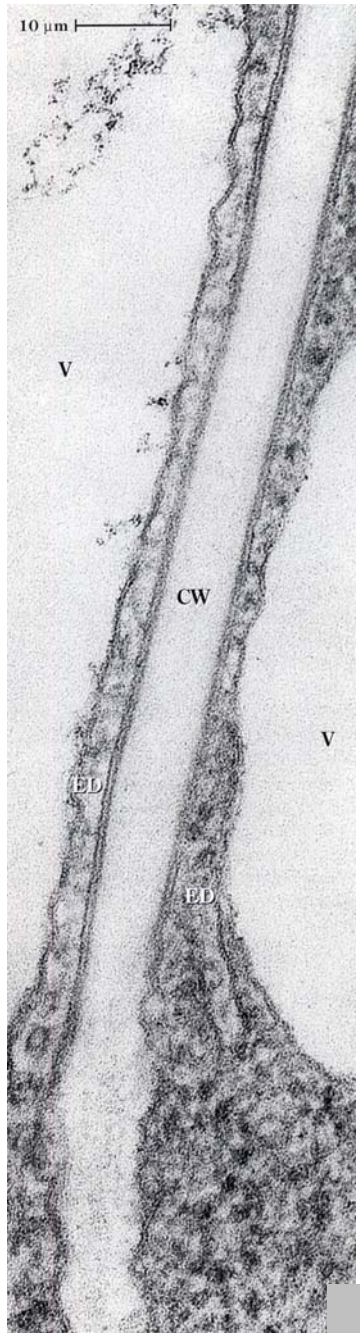




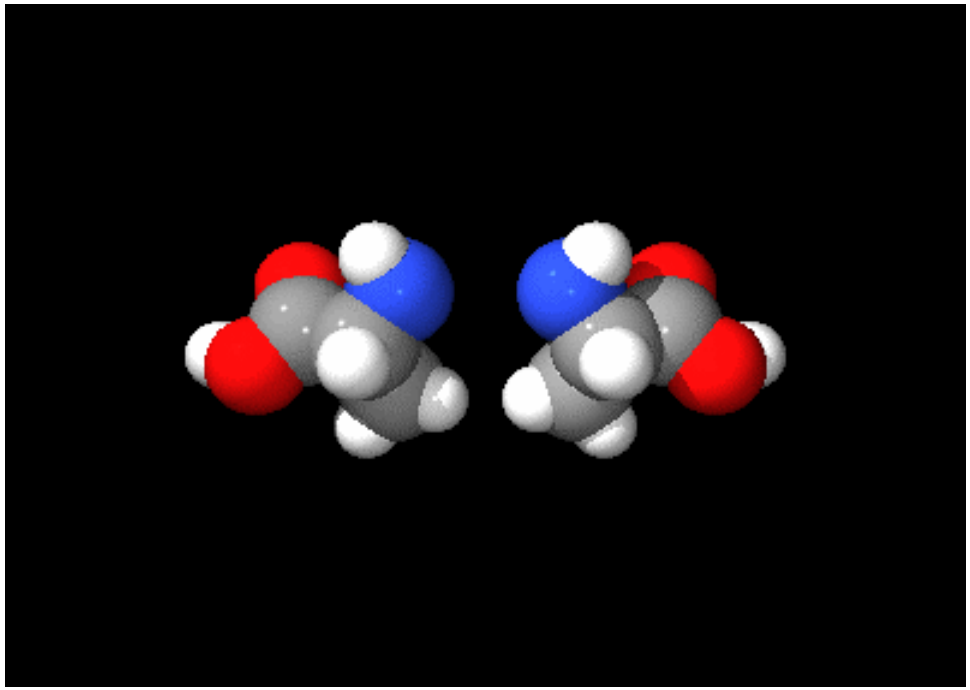
Die pflanzliche eukaryotische Zelle

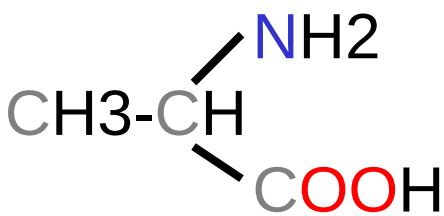




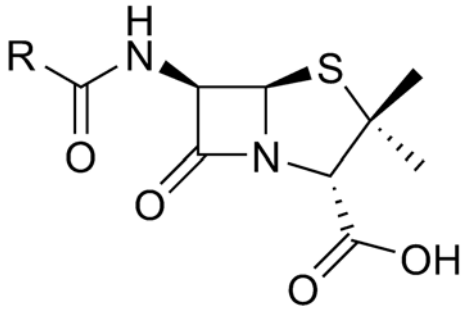


Zellmembran und Zellwand



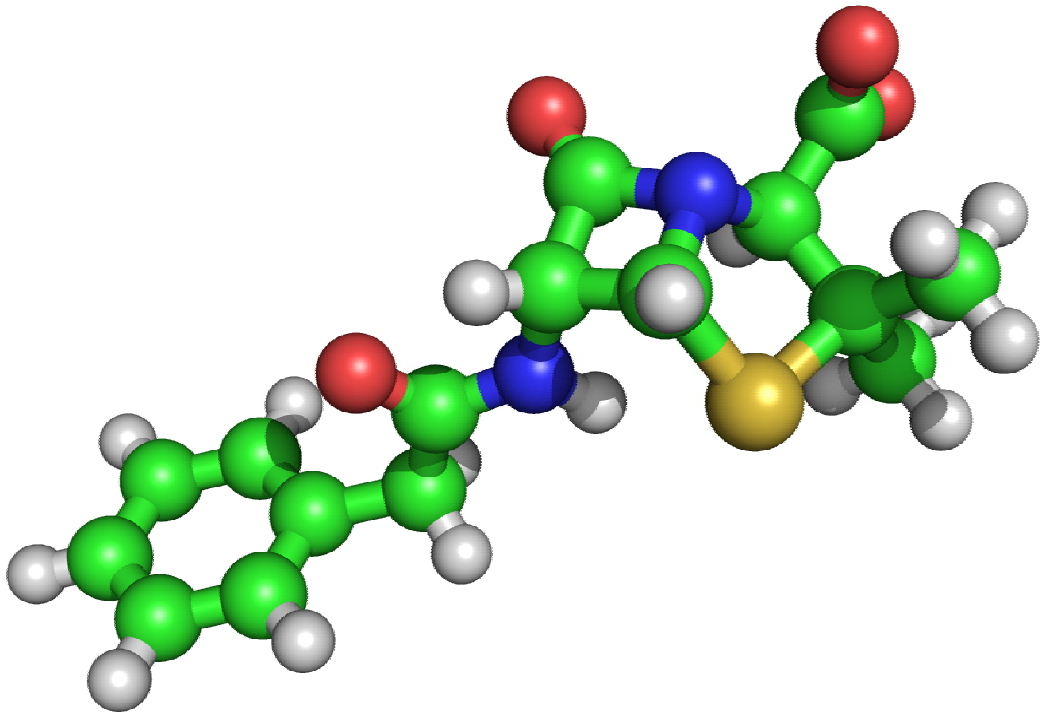


Chiralität bei Aminosäuren: D,L-Alanin

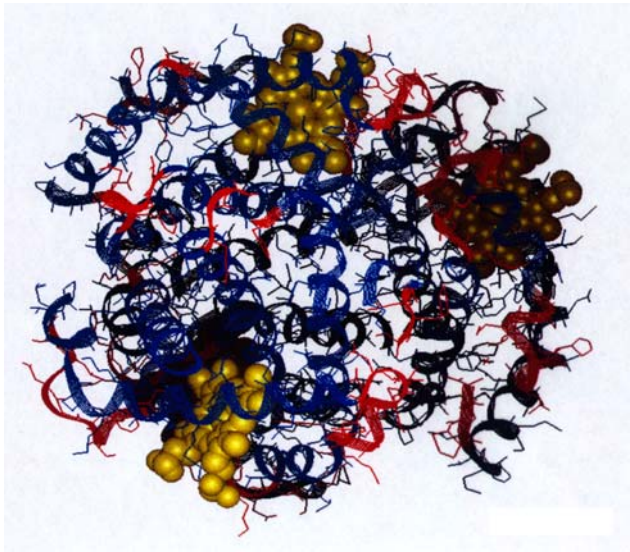


R = - CH_2-C_6H_5

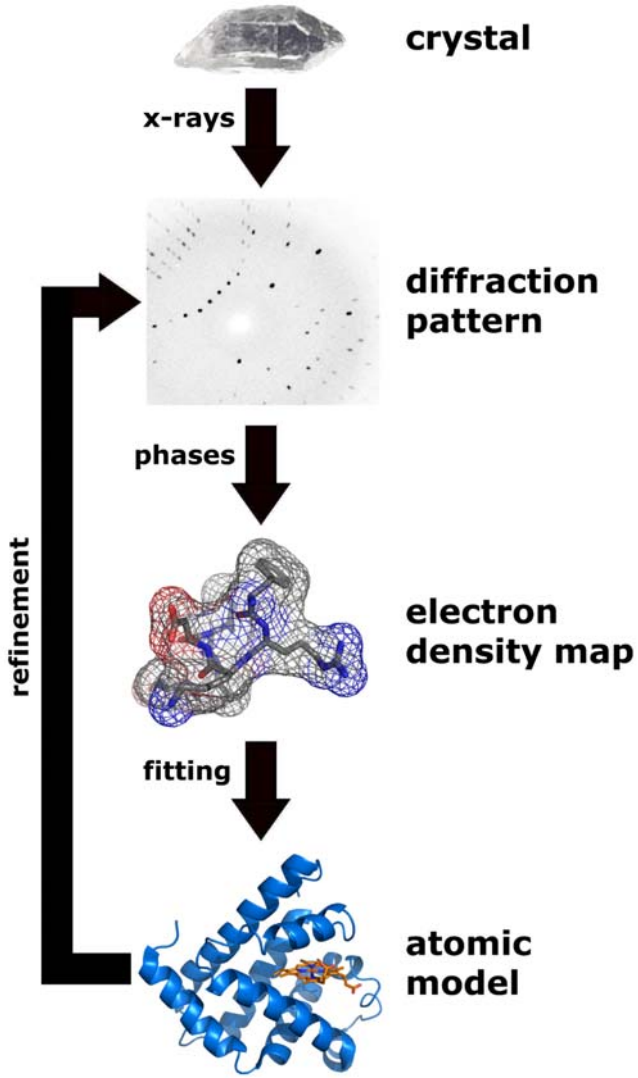
Penicillin G: 41 Atome
Molgewicht 334.4 Da



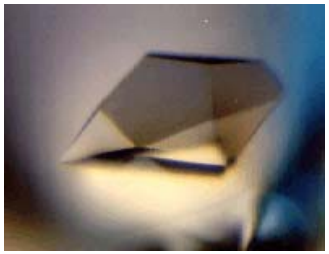
Molekulare Struktur von Penicillin



Hämoglobin: ≈ 10 000 Atome
Molgewicht: 64 000 Da

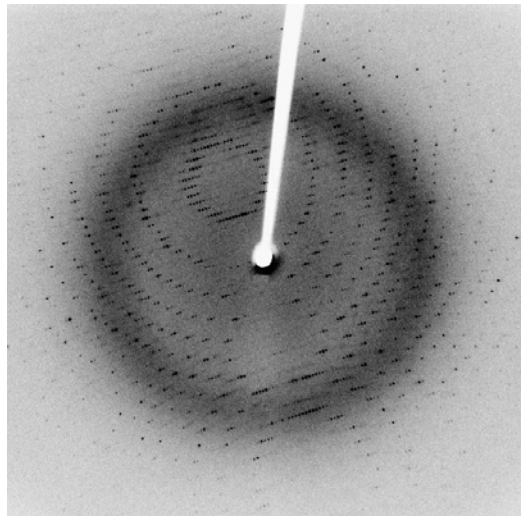


Röntgenkristallographie

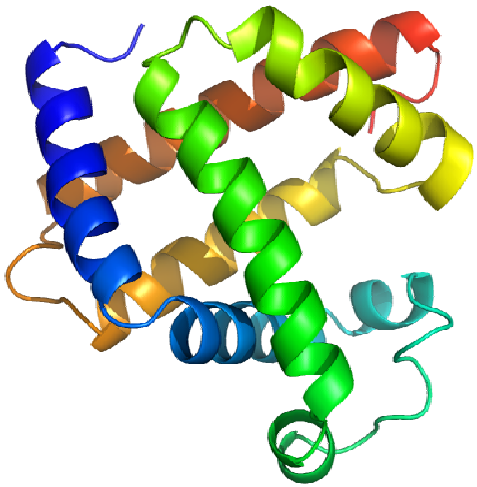


Proteinkristall

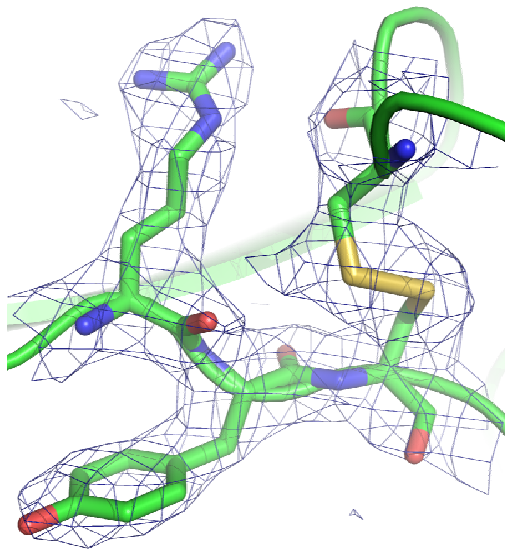
Röntgendiffraktometer



Diffraktionspattern

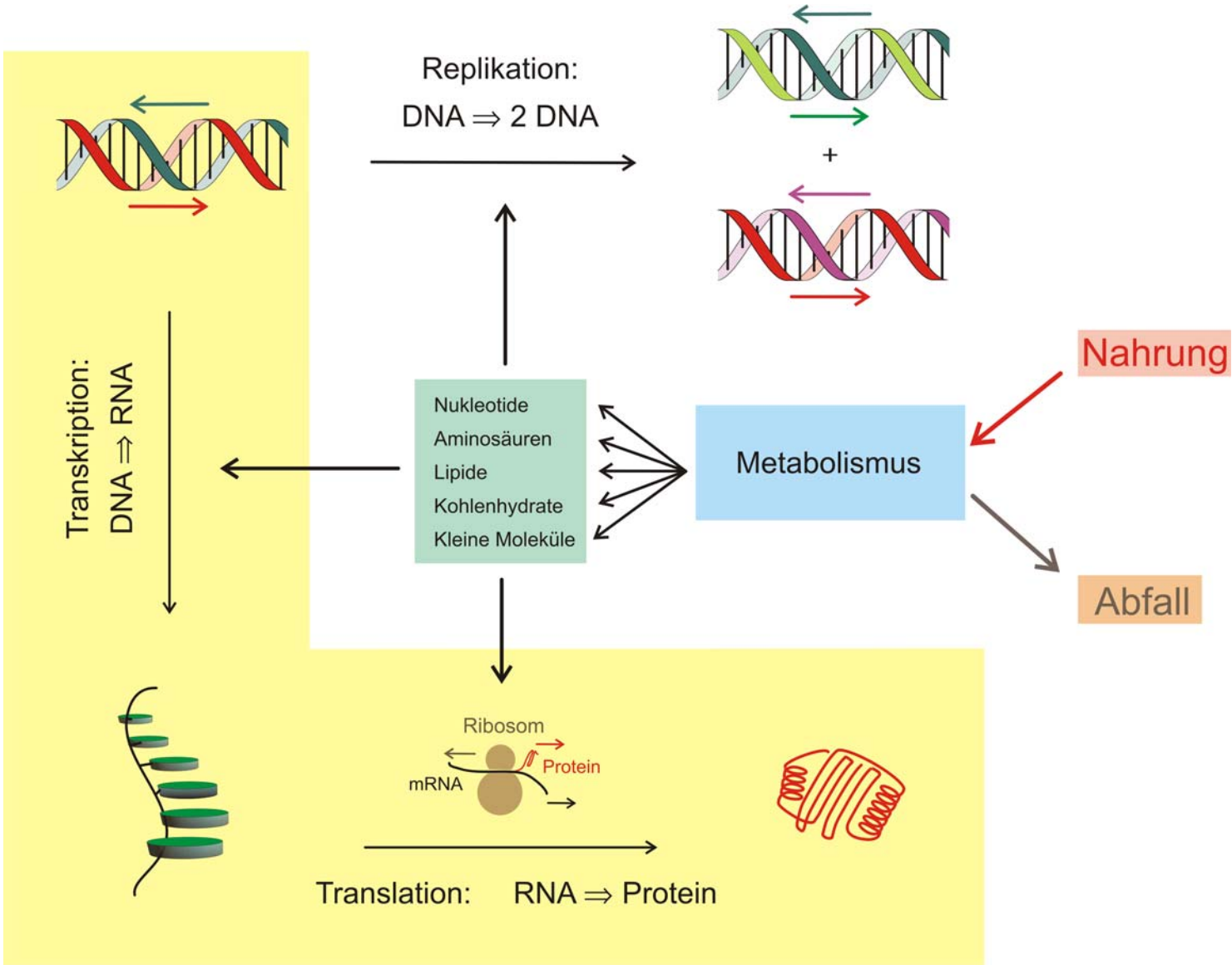


Molekulare Struktur



Elektronendichtemodellierung

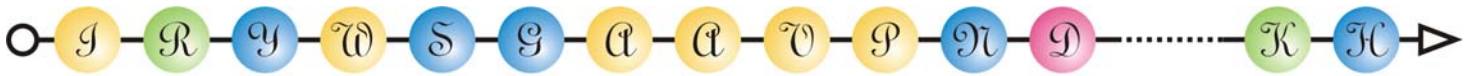
Proteinkristallographie



Verarbeitung der biologischen Information in der Zelle

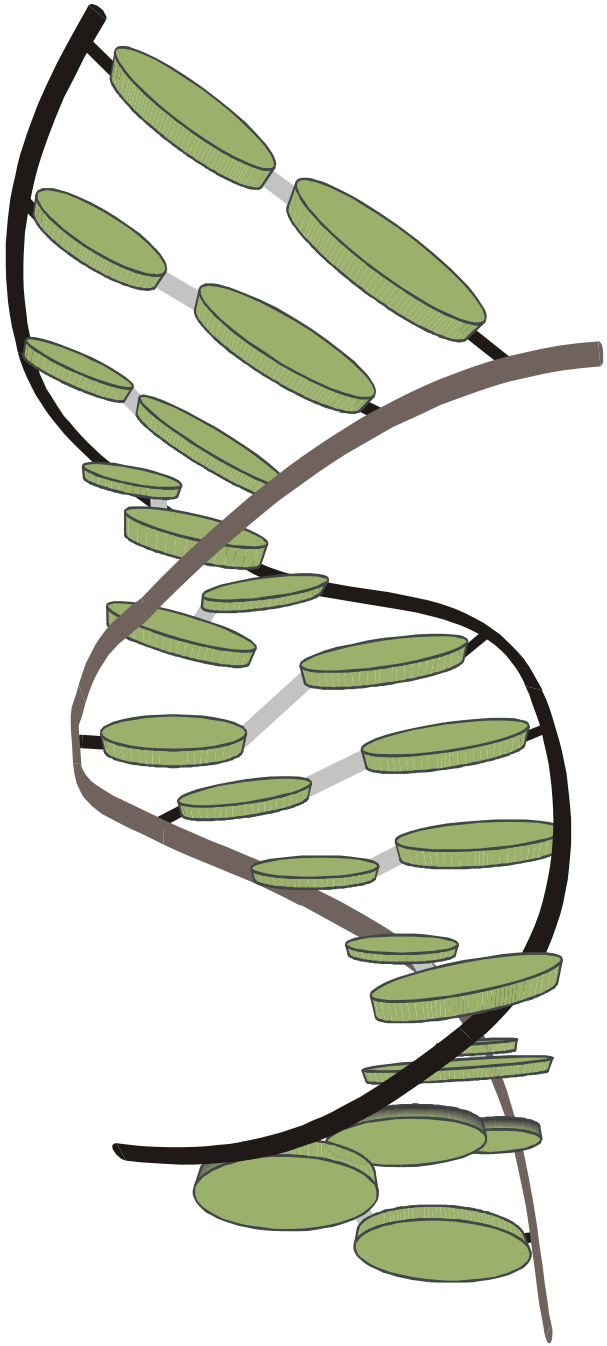
A ≡ Adenine
G ≡ Guanine
T ≡ Thymine
C ≡ Cytosine

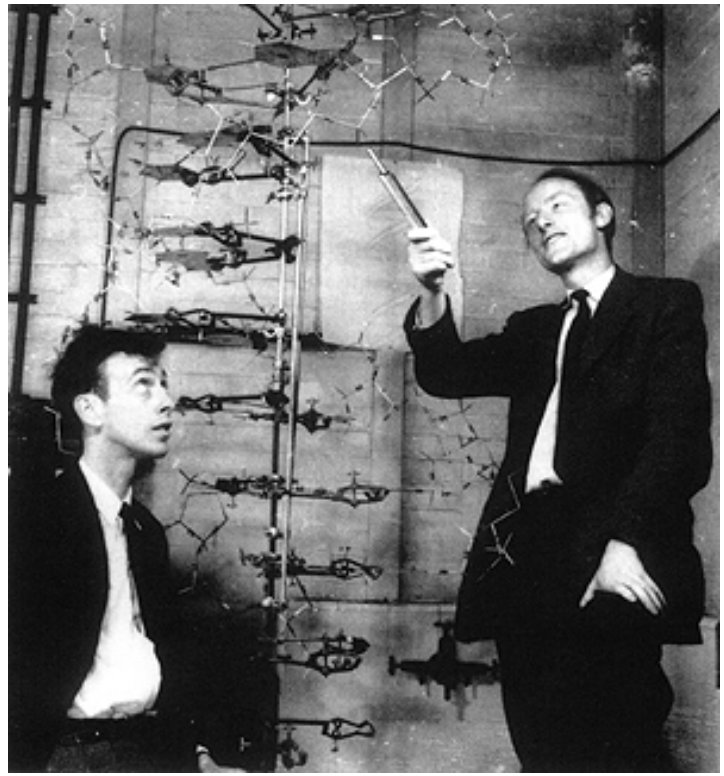
Deoxyribonucleic acid - DNA



𝔄 ≡ **ala**nine
𝔊 ≡ **gly**cine
𝔐 ≡ **met**hionine
𝔖 ≡ **ser**ine
ℭ ≡ **cys**teine
ℌ ≡ **his**tidine
ℜ ≡ **as**paragi**n**e
𝔗 ≡ **thr**eonine
𝔇 ≡ **asp**artic acid
ℑ ≡ **iso**l**eu**cine
𝔓 ≡ **pro**line
𝔙 ≡ **val**ine
𝔈 ≡ **glu**tamic acid
𝔎 ≡ **ly**sine
𝔔 ≡ **gl**utami**n**e
𝔚 ≡ **tryp**tophane
𝔉 ≡ **phe**nyl alanine
𝔏 ≡ **leu**cine
ℜ ≡ **arg**inine
𝔜 ≡ **tyr**osine

Protein

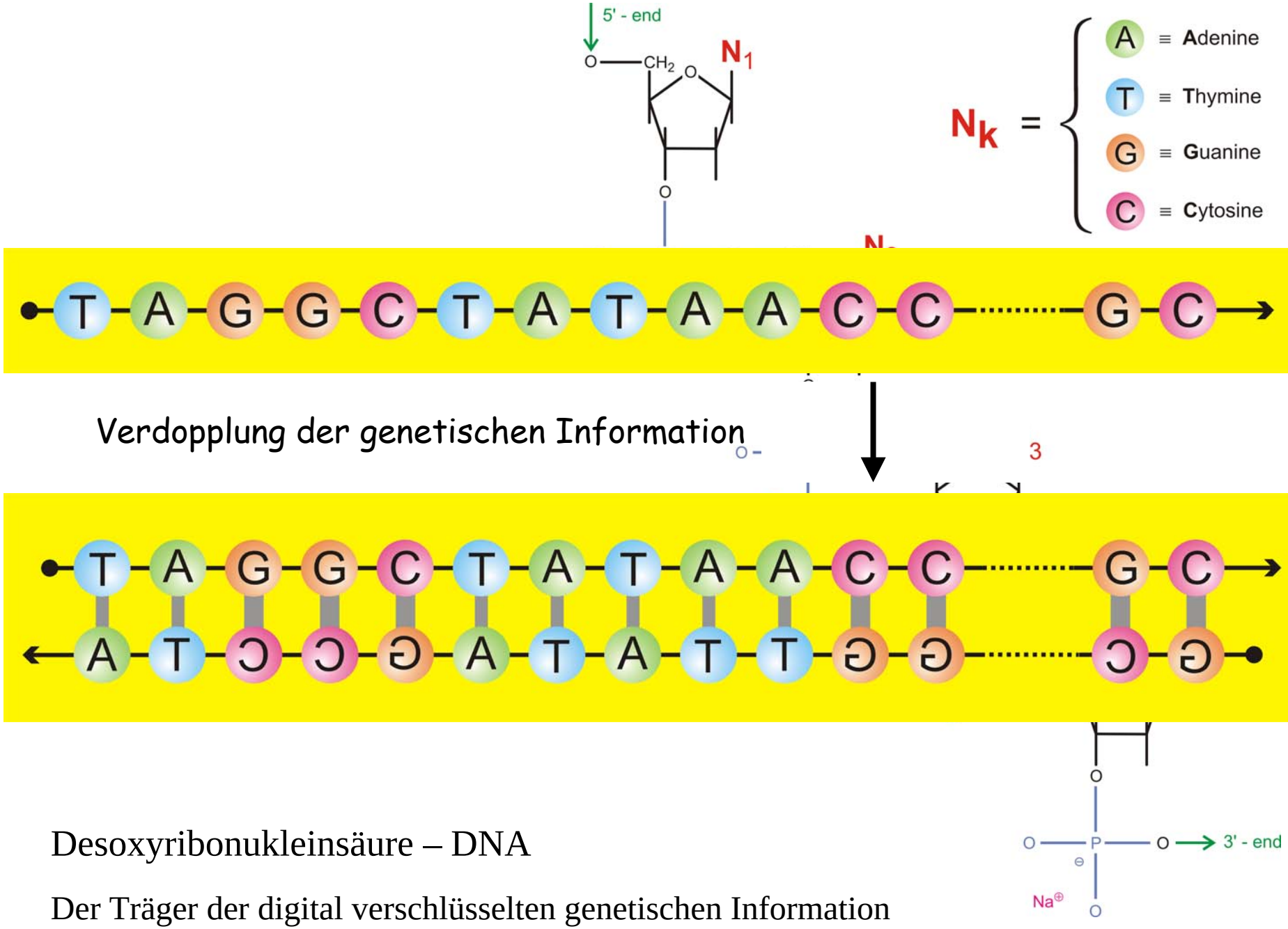




James D. Watson, 1928- , and Francis Crick, 1916- ,
Nobel Prize 1962

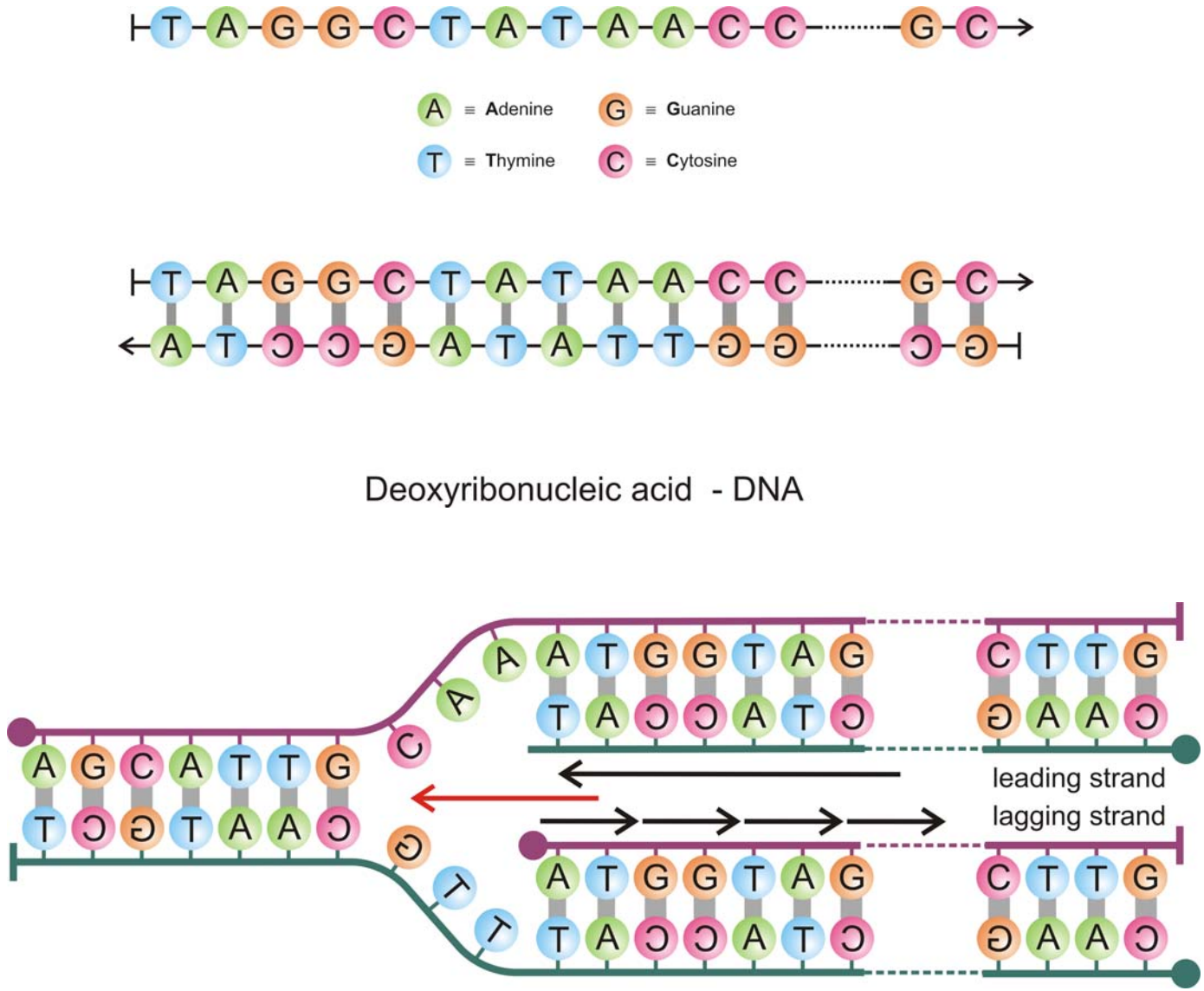
1953 – 2003 fifty years double helix

The three-dimensional structure of a short double helical stack of B-DNA

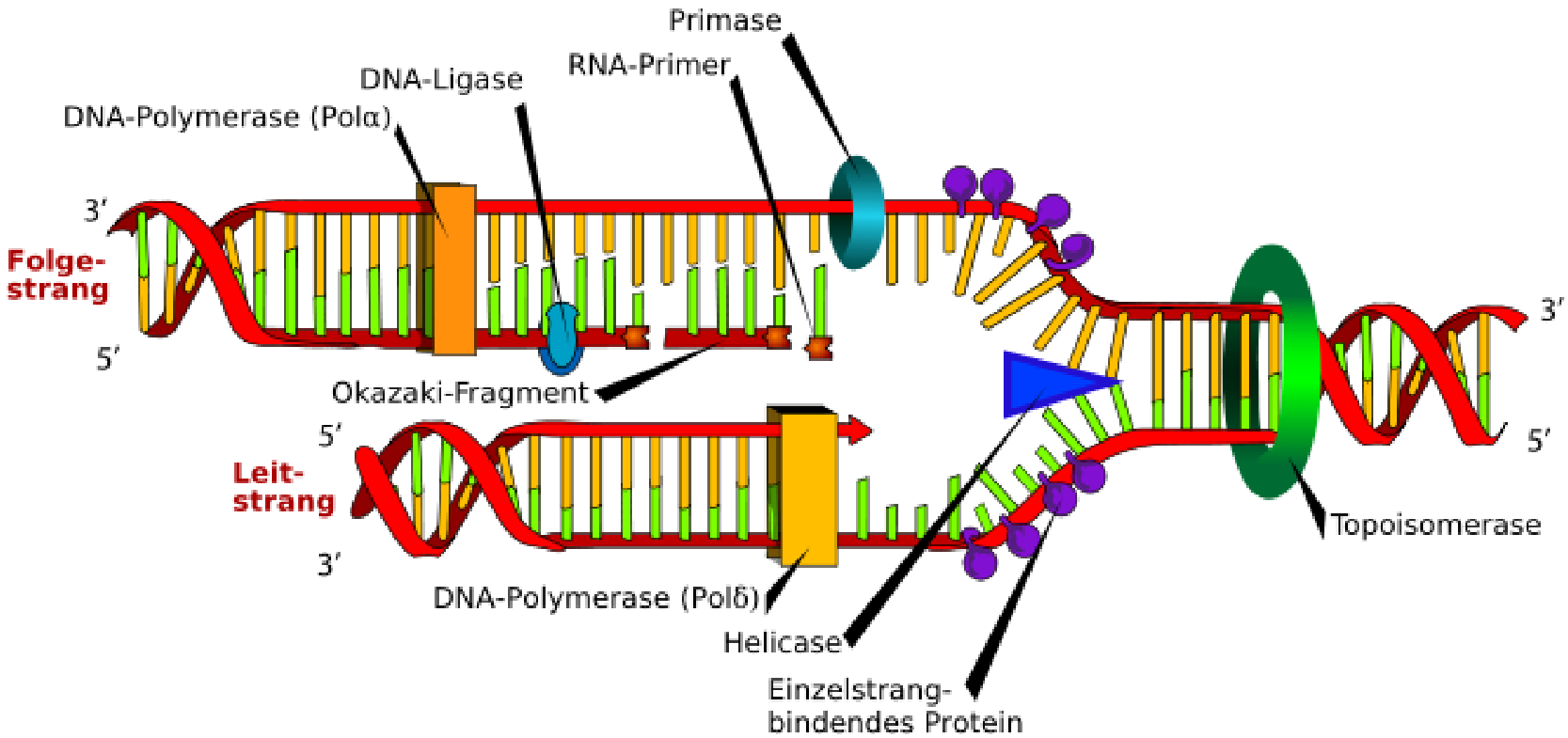


Desoxyribonukleinsäure – DNA

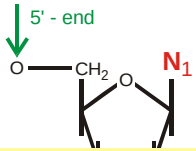
Der Träger der digital verschlüsselten genetischen Information



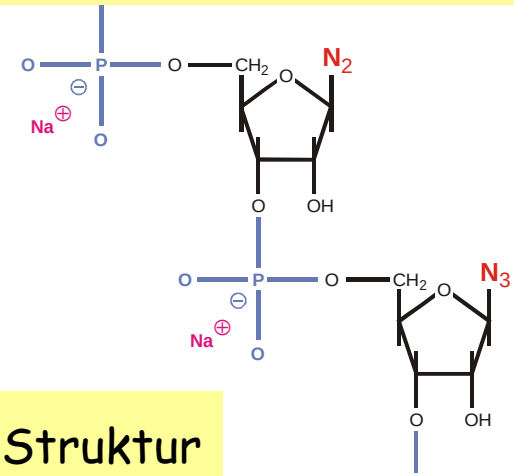
Der Mechanismus der DNA Replication

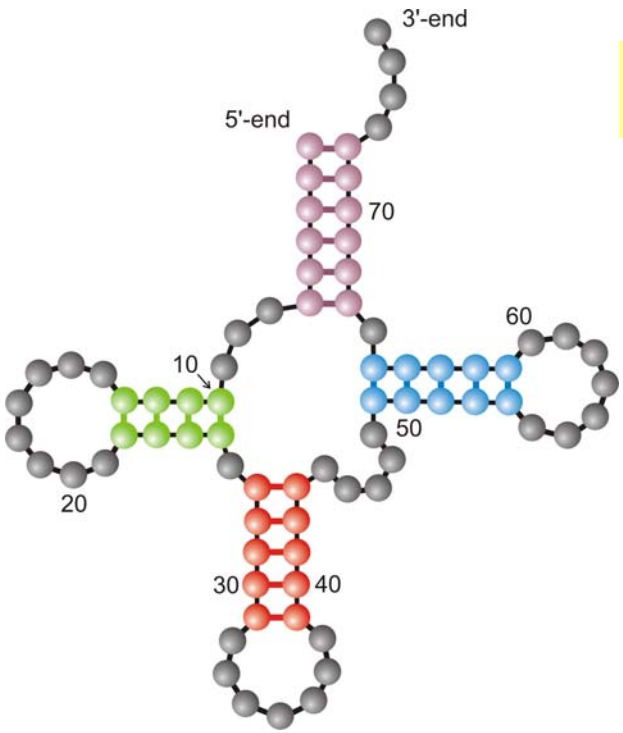


Replikation eines DNA Doppelstranges

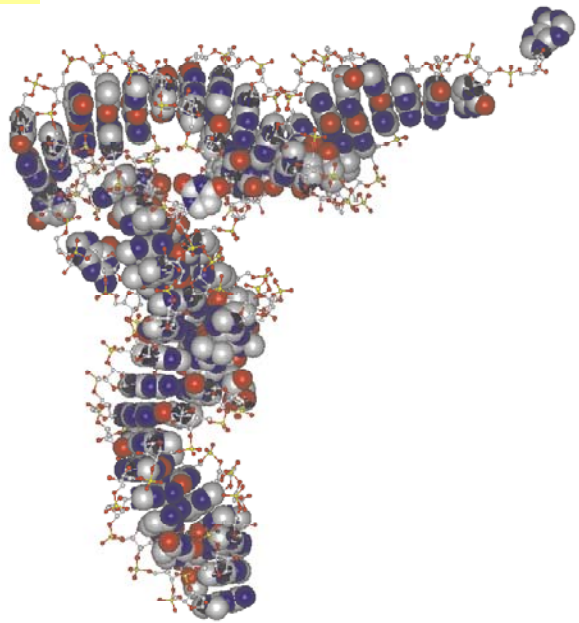


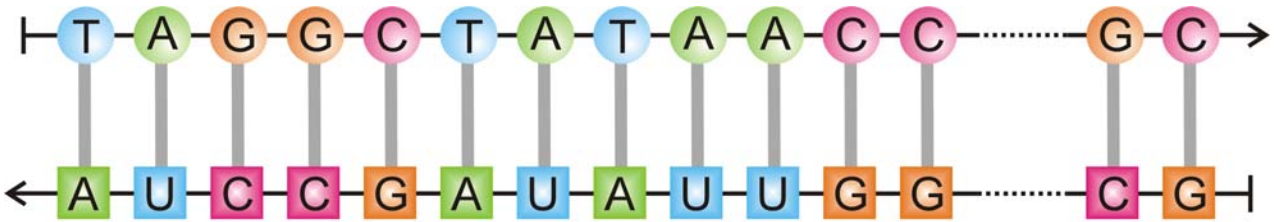
5'-end GCGGAUUUAGCUCAGUUGGGAGAGCGCCAGACUGAAGAUCUGGAGGUCCUGUGUUCGAUCCACAGAAUUCGCACCA 3'-end



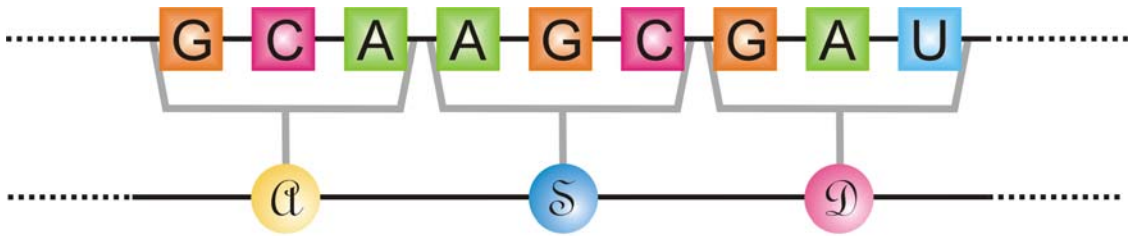


Definition der RNA Struktur

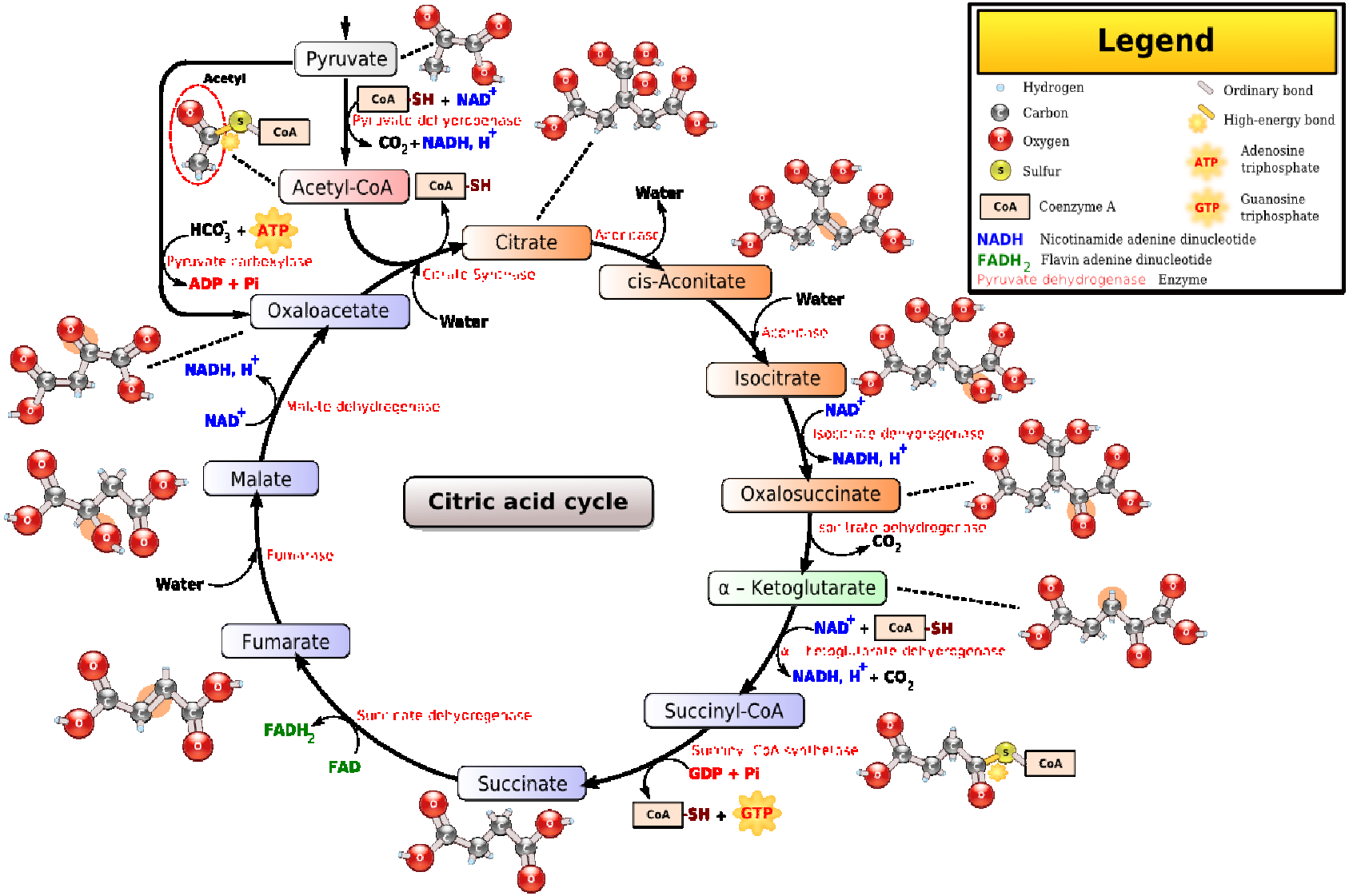




Transcription - DNA → RNA

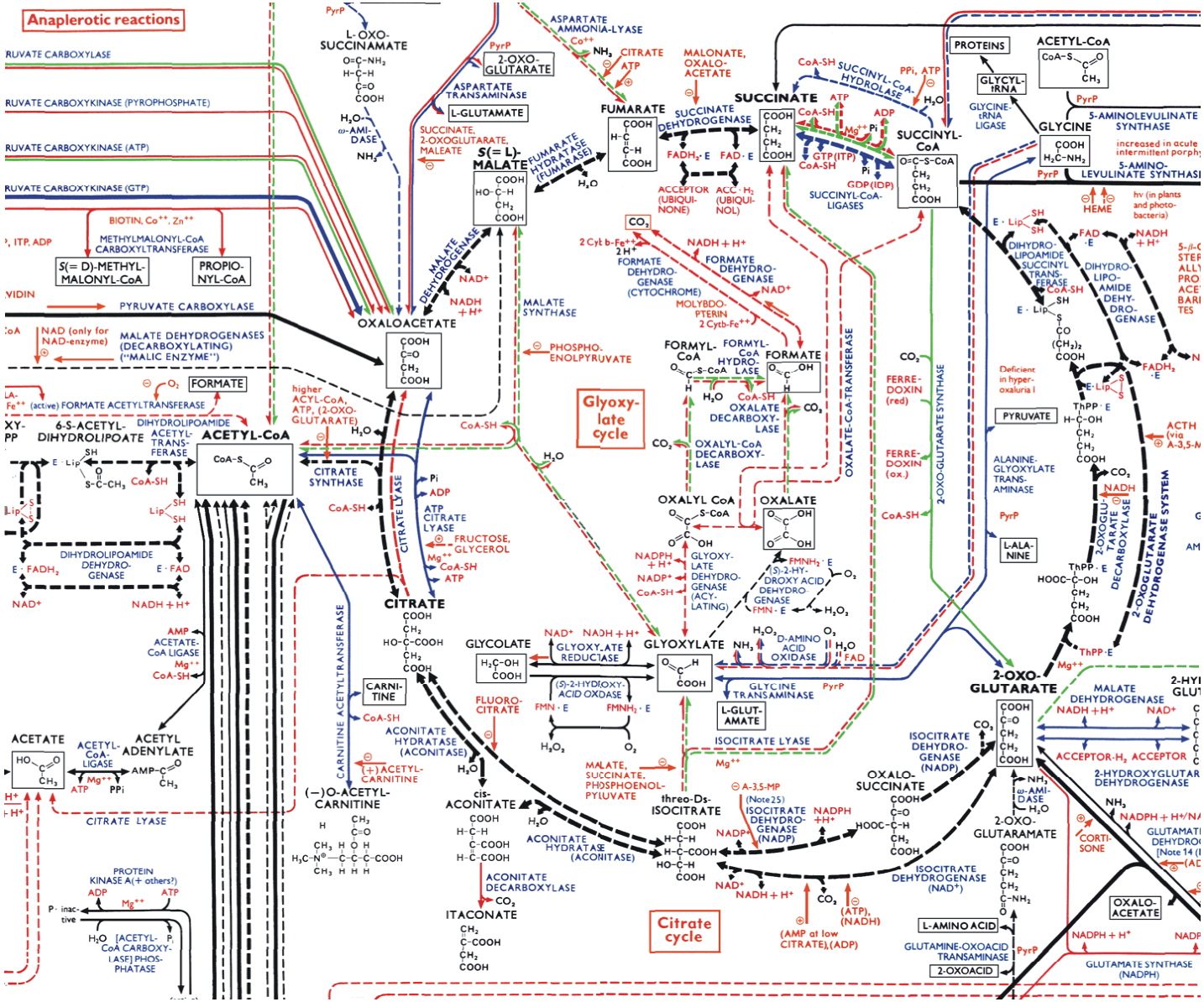


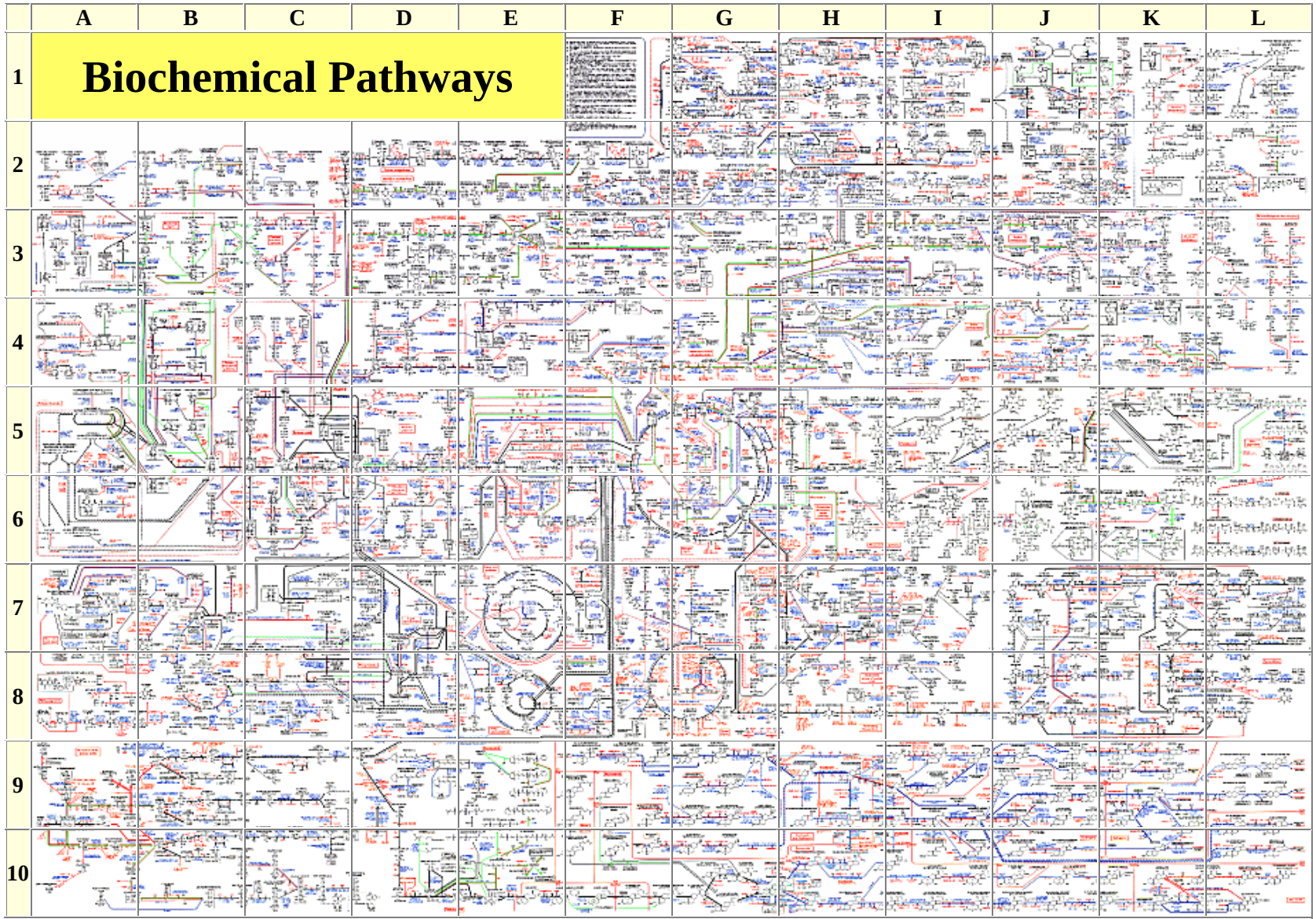
Translation - RNA → Protein



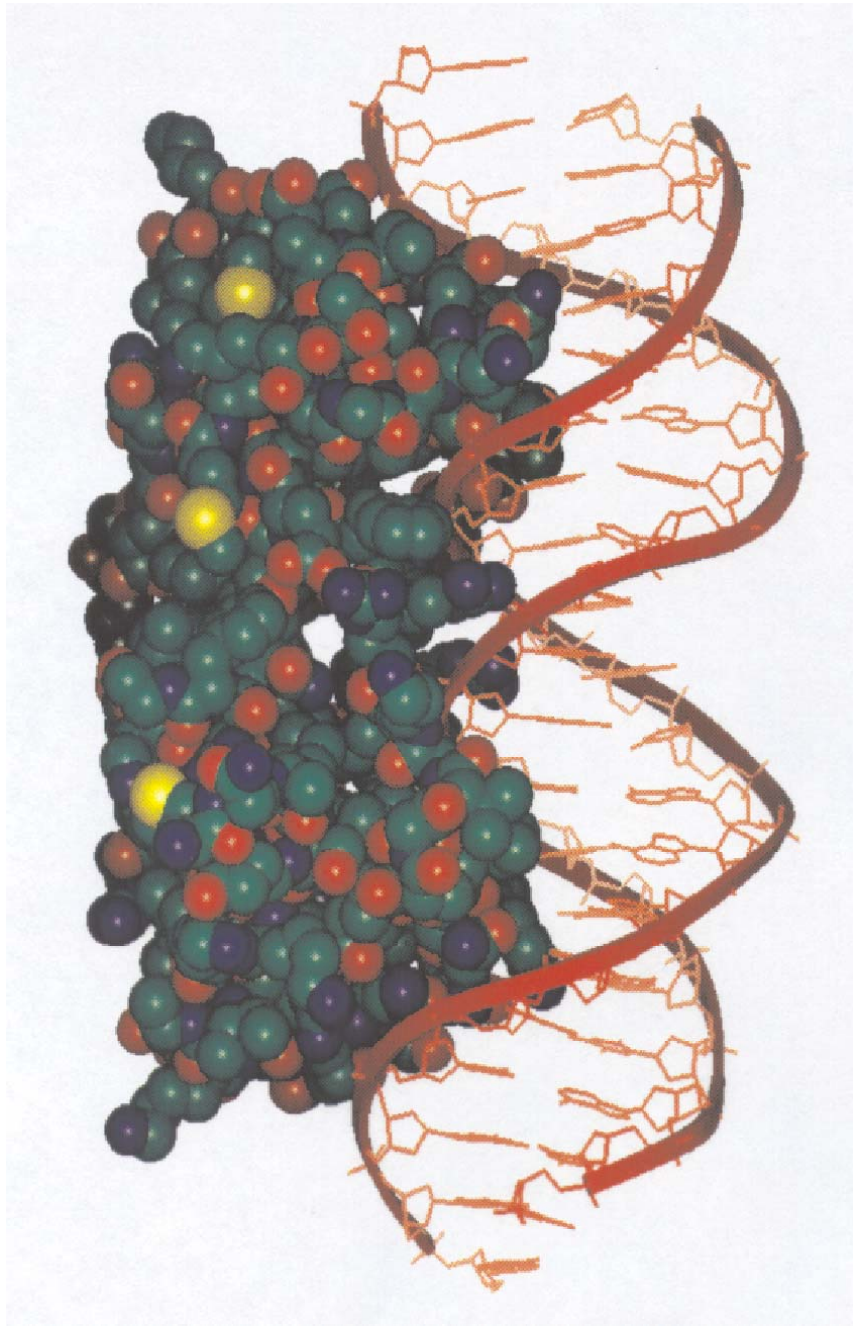
Der Zitronensäure- oder Krebszyklus als Energiequelle der Zelle in den Mitochondrien

The citric acid or Krebs cycle (enlarged from previous slide).

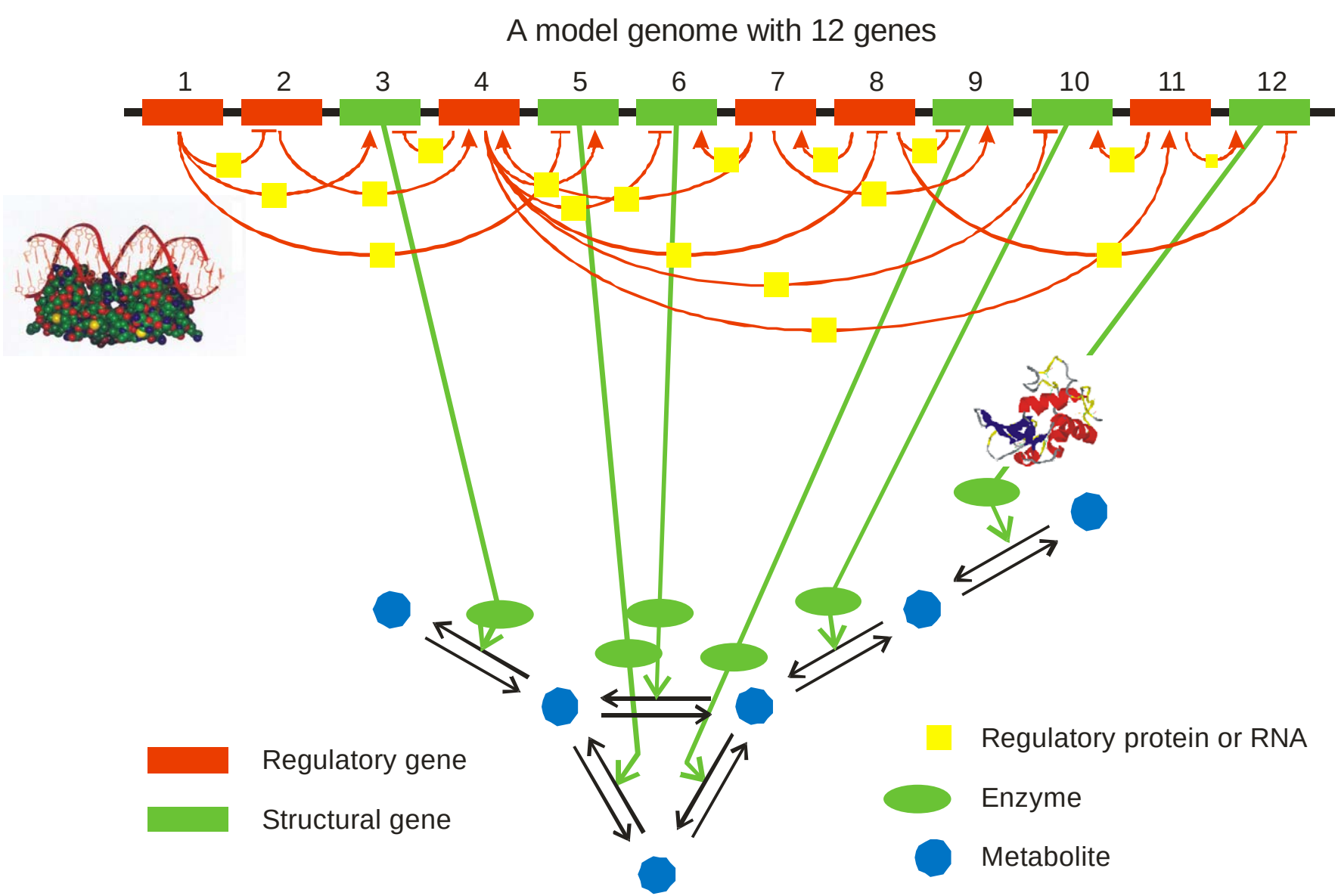




The reaction network of cellular metabolism published by Boehringer-Mannheim.



Die molekulare Struktur des Komplexes aus dem Regulationsprotein **cro-repressor** und der spezifischen Bindungsstelle an der λ-Phagen **B-DNA**

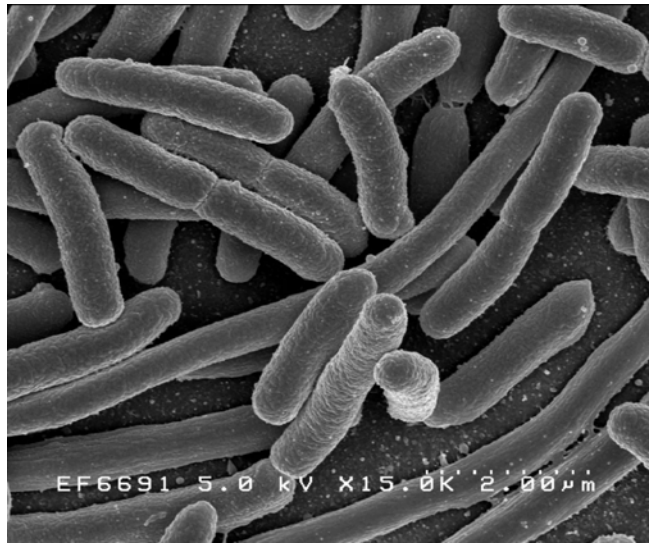


Sketch of a genetic and metabolic network

E. coli: Genomlänge 4×10^6 Nucleotides

Zahl der Zelltypen 1

Zahl der Gene 4 460



Mensch: Genomlänge 3×10^9 Nucleotides

Zahl der Zelltypes 200

Zahl der Gene ≈ 30 000

Komplexität in der Biologie

Die Schwierigkeit einer Definition des Begriffs "Gen".

Helen Pearson,
Nature **441**: 399-401, 2006

WHAT IS A GENE?

The idea of genes as beads on a DNA string is fast fading. Protein-coding sequences have no clear beginning or end and RNA is a key part of the information package, reports **Helen Pearson**.

'Gene' is not a typical four-letter word. It is not offensive. It is never bleeped out of TV shows. And where the meaning of most four-letter words is all too clear, that of gene is not. The more expert scientists become in molecular genetics, the less easy it is to be sure about what, if anything, a gene actually is.

Rick Young, a geneticist at the Whitehead Institute in Cambridge, Massachusetts, says that when he first started teaching as a young professor two decades ago, it took him about two hours to teach fresh-faced undergraduates what a gene was and the nuts and bolts of how it worked. Today, he and his colleagues need three months of lectures to convey the concept of the gene, and that's not because the students are any less bright. "It takes a whole semester to teach this stuff to talented graduates," Young says. "It used to be we could give a one-off definition and now it's much more complicated."

In classical genetics, a gene was an abstract concept — a unit of inheritance that ferried a characteristic from parent to child. As biochemistry came into its own, those characteristics were associated with enzymes or proteins, one for each gene. And with the advent of molecular biology, genes became real, physical things — sequences of DNA which when converted into strands of so-called messenger RNA could be used as the basis for building their associated protein piece by piece. The great coiled DNA molecules of the chromosomes were seen as long strings on which gene sequences sat like discrete beads.

This picture is still the working model for many scientists. But those at the forefront of genetic research see it as increasingly old-fashioned — a crude approximation that, at best, hides fascinating new complexities and, at worst, blinds its users to useful new paths of enquiry.

Information, it seems, is parceled out along chromosomes in a much more complex way than was originally supposed. RNA molecules are not just passive conduits through which the gene's message flows into the world but active regulators of cellular processes. In some cases, RNA may even pass information across generations — normally the sole preserve of DNA.

An eye-opening study last year raised the possibility that plants sometimes rewrite their DNA on the basis of RNA messages inherited from generations past[1]. A study on page 469 of this issue suggests that a comparable phenomenon might occur in mice, and by implication in other mammals[2]. If this type of phenomenon is indeed widespread, it "would have huge implications," says evolutionary geneticist Laurence Hurst at the University of Bath, UK.

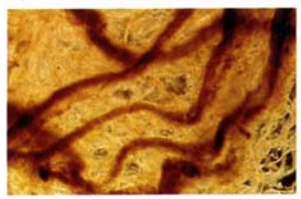
"All of that information seriously challenges our conventional definition of a gene," says molecular biologist Bing Ren at the University of California, San Diego. And the information challenge is about to get even tougher. Later this year, a glut of data will be released from the international Encyclopedia of DNA Elements (ENCODE) project. The pilot phase of ENCODE involves scrutinizing roughly 1% of the human genome in unprecedented detail; the aim is to find all the sequences that serve a useful purpose and explain what that purpose is. "When we started the ENCODE project I had a different view of what a gene was," says contributing researcher Roderic Guigo at the Center for Genomic Regulation in Barcelona. "The degree of complexity we've seen was not anticipated."

> "We've come to the realization that the genome is full of overlapping transcripts."
> — Phillip Kapranov

Under fire

The first of the complexities to challenge molecular biology's paradigm of a single DNA sequence encoding a single protein was alternative splicing, discovered in viruses in 1977 (see 'Hard to track', overleaf). Most of the DNA sequences describing proteins in humans have a modular arrangement in which exons, which carry the instructions for making proteins, are interspersed with non-coding introns. In alternative splicing, the cell snips out introns and sews together the exons in various different orders, creating messages that can code for different proteins. Over the years geneticists have also documented overlapping genes, genes within genes and countless other weird arrangements (see 'Muddling over genes', overleaf).

Alternative splicing, however, did not in itself require a drastic reappraisal of the notion of a gene; it just showed that some DNA sequences could describe more than one protein. Today's assault on the gene concept is more far reaching, fuelled largely by studies that show the previously unimagined scope of RNA.



Spools of DNA (above) still harbour surprises, with one protein-coding gene often overlapping the next.

P. PLAILLY/SPL

The one gene, one protein idea is coming under particular assault from researchers who are comprehensively extracting and analysing the RNA messages, or transcripts, manufactured by genomes, including the human and mouse genome. Researchers led by Thomas Gingeras at the company Affymetrix in Santa Clara, California, for example, recently studied all the transcripts from ten chromosomes across eight human cell lines and worked out precisely where on the chromosomes each of the transcripts came from[3].

The picture these studies paint is one of mind-boggling complexity. Instead of discrete genes dutifully mass-producing identical RNA transcripts, a teeming mass of transcription converts many segments of the genome into multiple RNA ribbons of differing lengths. These ribbons can be generated from both strands of DNA, rather than from just one as was conventionally thought. Some of these transcripts come from regions of DNA previously identified as holding protein-coding genes. But many do not. "It's somewhat revolutionary," says Gingeras's colleague Phillip Kapranov. "We've come to the realization that the genome is full of overlapping transcripts."

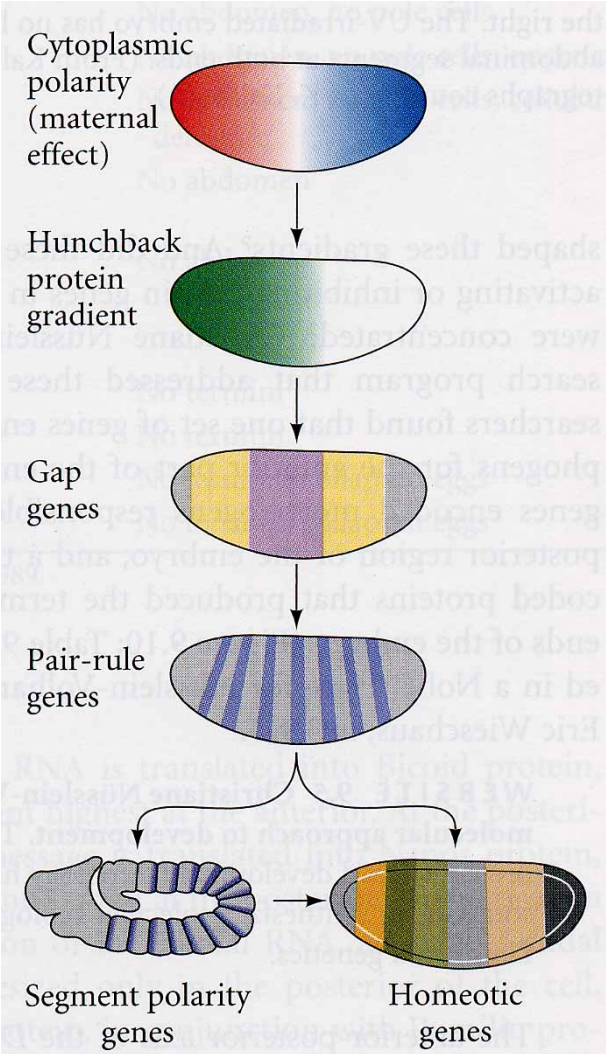
Other studies, one by Guigo's team[4], and one by geneticist Rotem Sorek[5], now at Tel Aviv University, Israel, and his colleagues, have hinted at the reasons behind the mass of transcription. The two teams investigated occasional reports that transcription can start at a DNA sequence associated with one protein and run straight through into the gene for a completely different protein, producing a fused transcript. By delving into databases of human RNA transcripts, Guigo's team estimate that 4–5% of the DNA in regions conventionally recognized as genes is transcribed in this way. Producing fused transcripts could be one way for a cell to generate a greater variety of proteins from a limited number of exons, the researchers say.

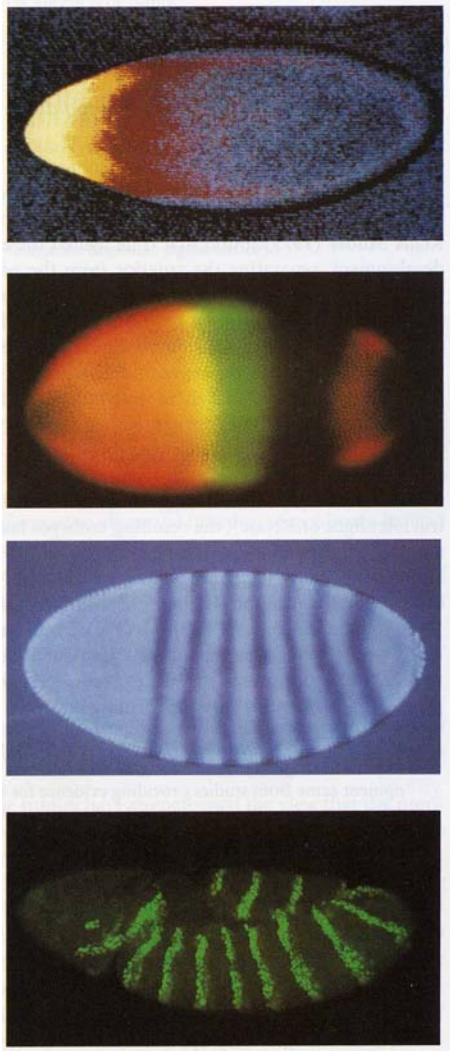
Many scientists are now starting to think that the descriptions of proteins encoded in DNA know no borders — that each sequence reaches into the next and beyond. This idea will be one of the central points to emerge from the ENCODE project when its results are published later this year.

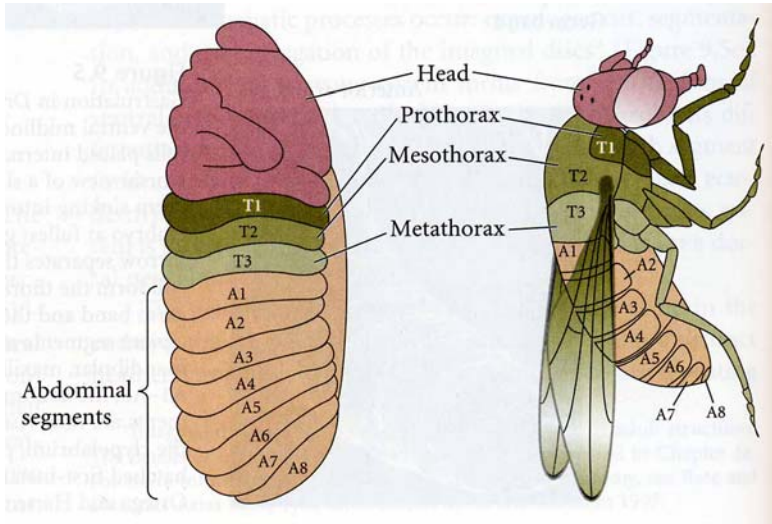
Kapranov and others say that they have documented many examples of transcripts in which protein-coding exons from one part of the genome combine with exons from another

ENCODE stands for
ENCyclopedia Of DNA Elements.

ENCODE Project Consortium.
Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project.
Nature **447**:799-816,2007







Cascades, A ⇒ B ⇒ C ⇒ ... , and networks of genetic control

Turing pattern resulting from reaction-diffusion equation ?

Intercelluar communication creating positional information

Development of the fruit fly *drosophila melanogaster*: Genetics, experiment, and imago

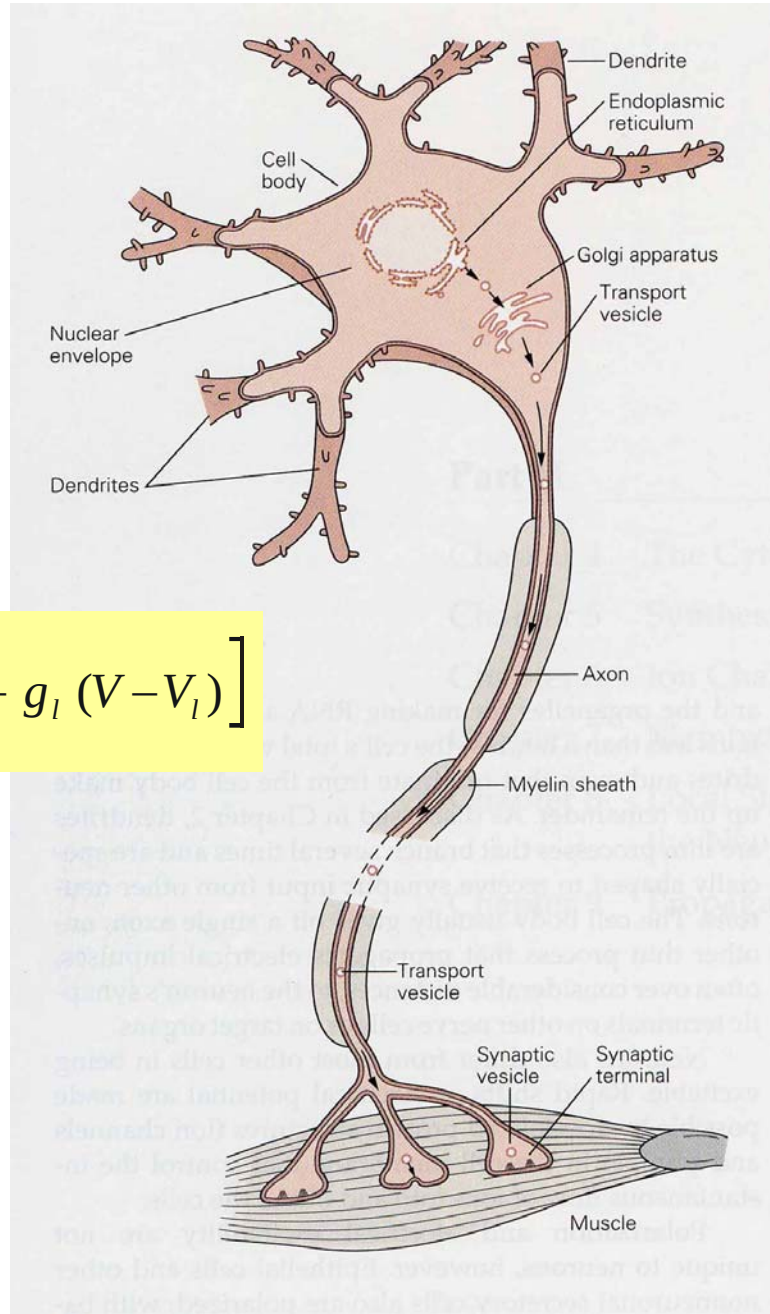
$$\frac{dV}{dt} = \frac{1}{C_M}\left[I - g_{Na}\, m^3\, h\,(V - V_{Na}) - g_K\, n^4\,(V - V_K) - g_l\,(V - V_l)\right]$$

$$\frac{dm}{dt} = \alpha_m\,(1-m) - \beta_m\, m$$

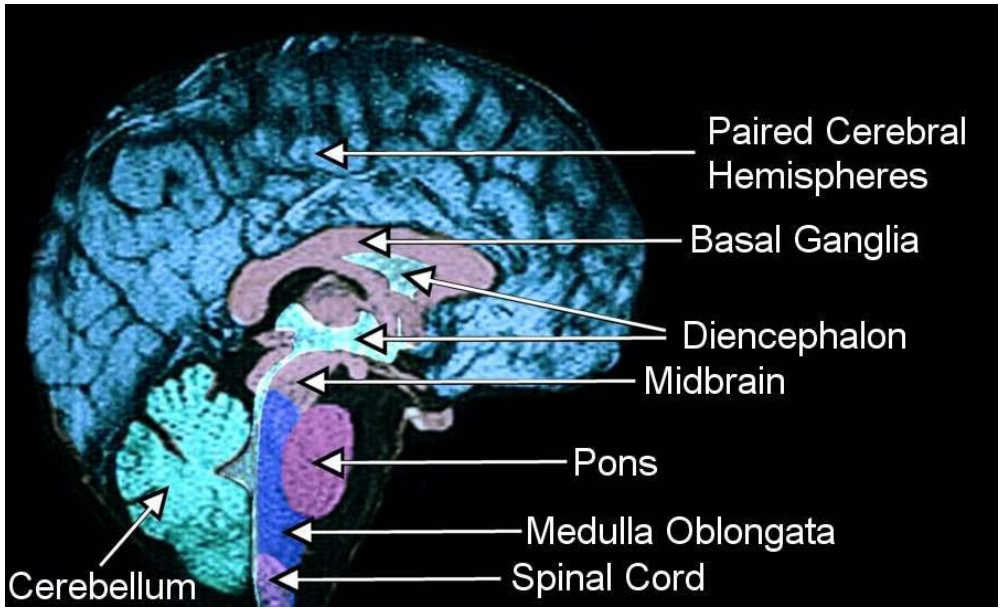
$$\frac{dh}{dt} = \alpha_h\,(1-h) - \beta_h\, h$$

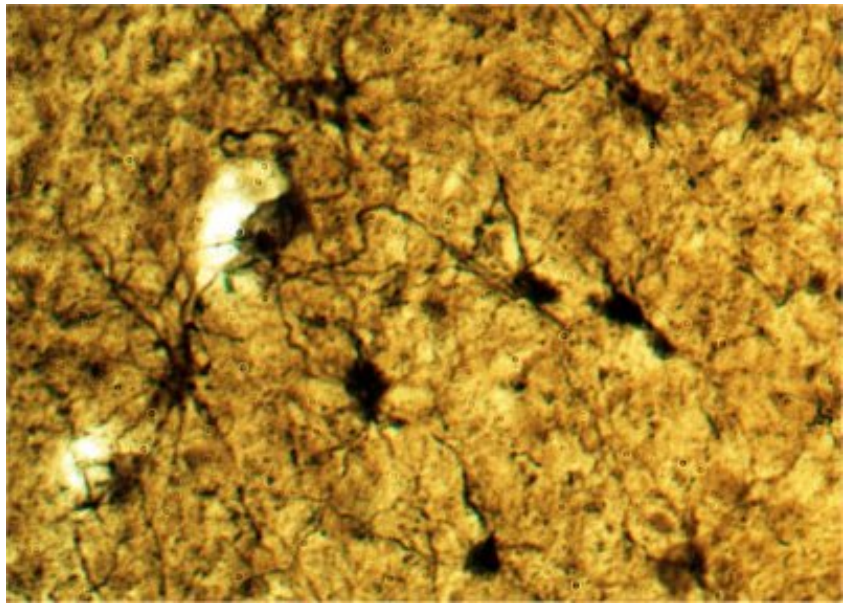
$$\frac{dn}{dt} = \alpha_n\,(1-n) - \beta_n\, n$$

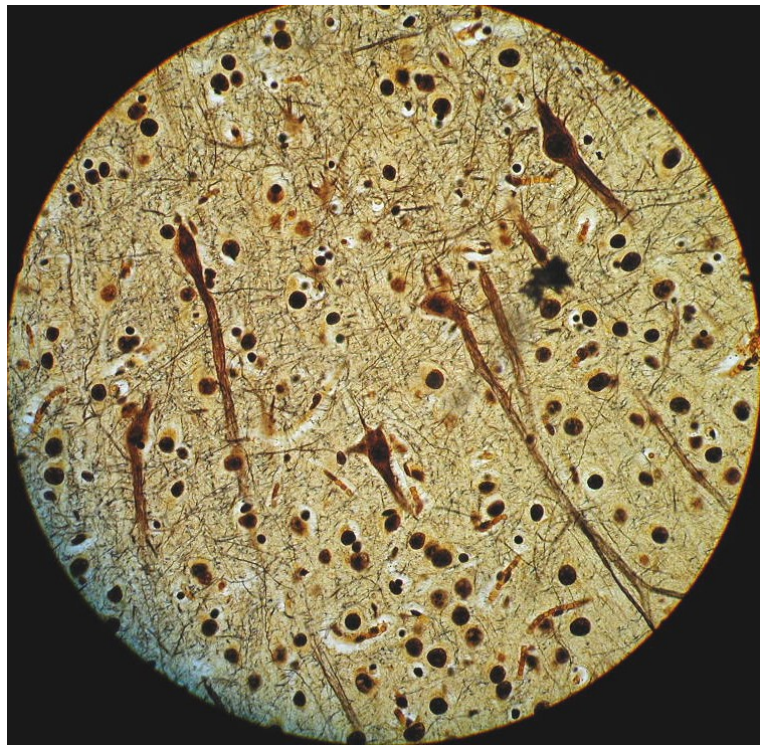
Hogdkin-Huxley OD equations



A single neuron signaling to a muscle fiber







The human brain

10^{11} neurons connected by ≈ 10^{13} to 10^{14} synapses

Web-Page for further information:

http://www.tbi.univie.ac.at/~pks