第1回

生体関連化学B

Chemical Biology

筑波大学数理物質系
横川　雅俊

生体関連化学A (大石先生)

講義内容(Contents)

応用理工学類(大学1～4年生)では(In undergraduate course):

1年生　生物学(自由科目)
1st year　Biological Science

2年生　生物工学概論(専門選択科目)
2nd year　Introduction of Biological Engineering

3年生　生命科学(専門選択科目)
3rd year　Life Science

3年生　生物工学(専門選択必修科目)
3rd year　Biological Engineering

本講義では、これらの講義内容を基礎として遺伝子工学の技術と応用について解説する。
This class　provides an overview of genetic engineering　based on the above-mentioned lectures.

生体関連化学A (大石先生)

詳しい講義内容（Detailed Contents）

必要な資料等はこちらで準備します。Purchase of a textbook is not necessary, because I will give you materials（lecture data）.

内容：Contents

- 復習：核酸と遺伝子の基礎　Review: Basic of Nucleic Acid and Gene
- 遺伝子工学の概要　Introduction of Genetic Engineering
- ベクター　Vector
- クローニング　Cloning
- ベクターによる生産　Production by Vector
- 遺伝子診断　Gene Diagnosis
- 遺伝子治療　Gene Therapy
- DNAナノテクノロジー　DNA nanotechnology

成績（Grading）

主に、出席とレポート（学期末の）により評価する。
The grade will be based on the attendance and a report.

欠席（Absence）

学会等で欠席する場合は、指導教員の捺印がある欠席届け（自作）を提出。
If you are absence due to the conferences, you should submit a statement of the reasons with the signature of supervisor.

本講義の目的

Humanized mice

バイオサイエンス、バイオテクノロジーを学ぶ上で必要な分子生物学・生化学・分析化学・生命情報学の基礎を学ぶ。又、生物と化学・物理の境界領域における研究のトピックスを紹介すると共に、これらの異なる領域の融合がもたらす新しいバイオテクノロジーの、医学、工学、理学、農学などにおける広大な応用範囲について議論する。

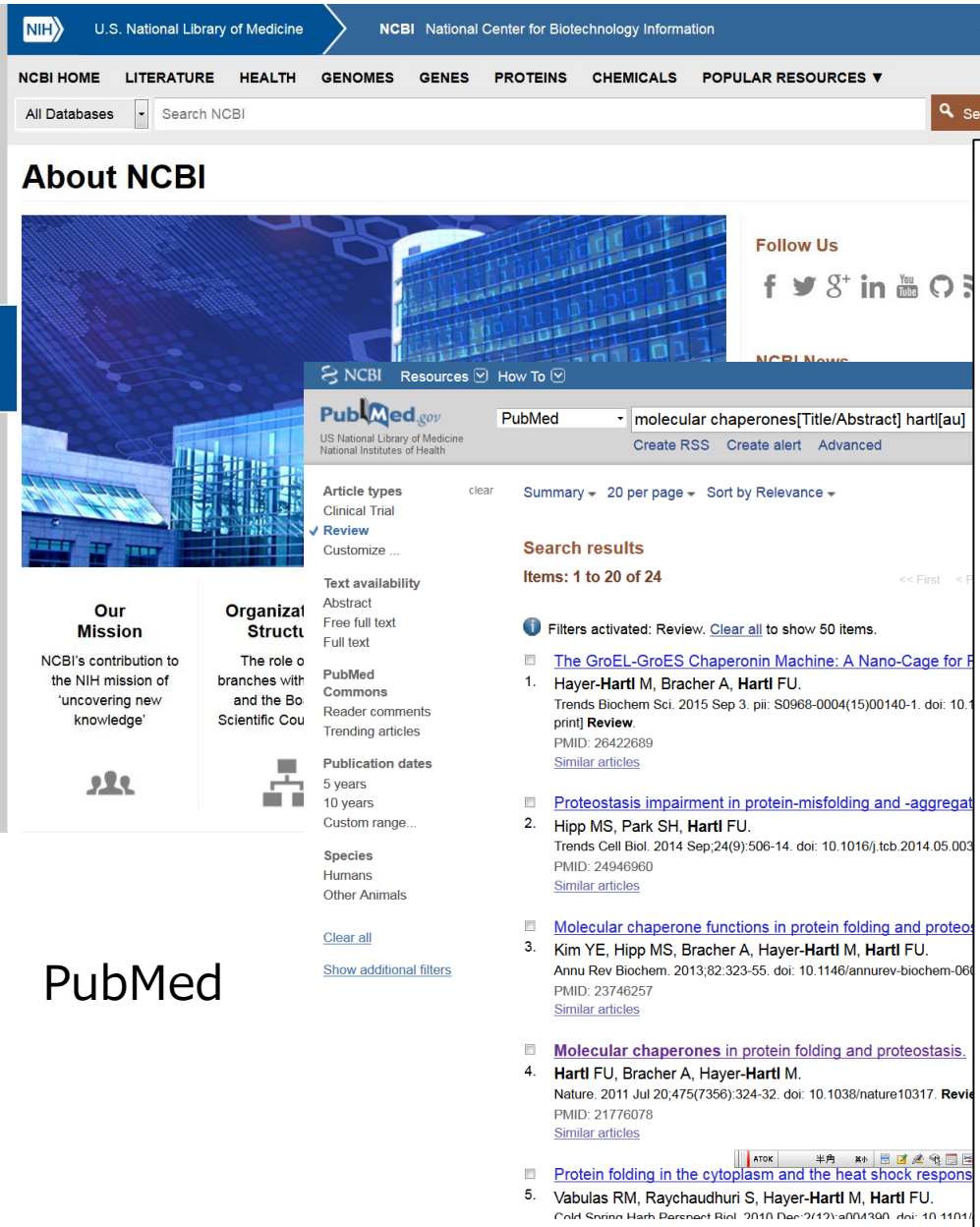
そもそも、バイオテクノロジーってなに！？

Biotechnology is a broad discipline in which biological processes, organisms, cells or cellular components are exploited to develop new technologies. New tools and products developed by biotechnologists are useful in research, agriculture, industry and the clinic. (nature.com)

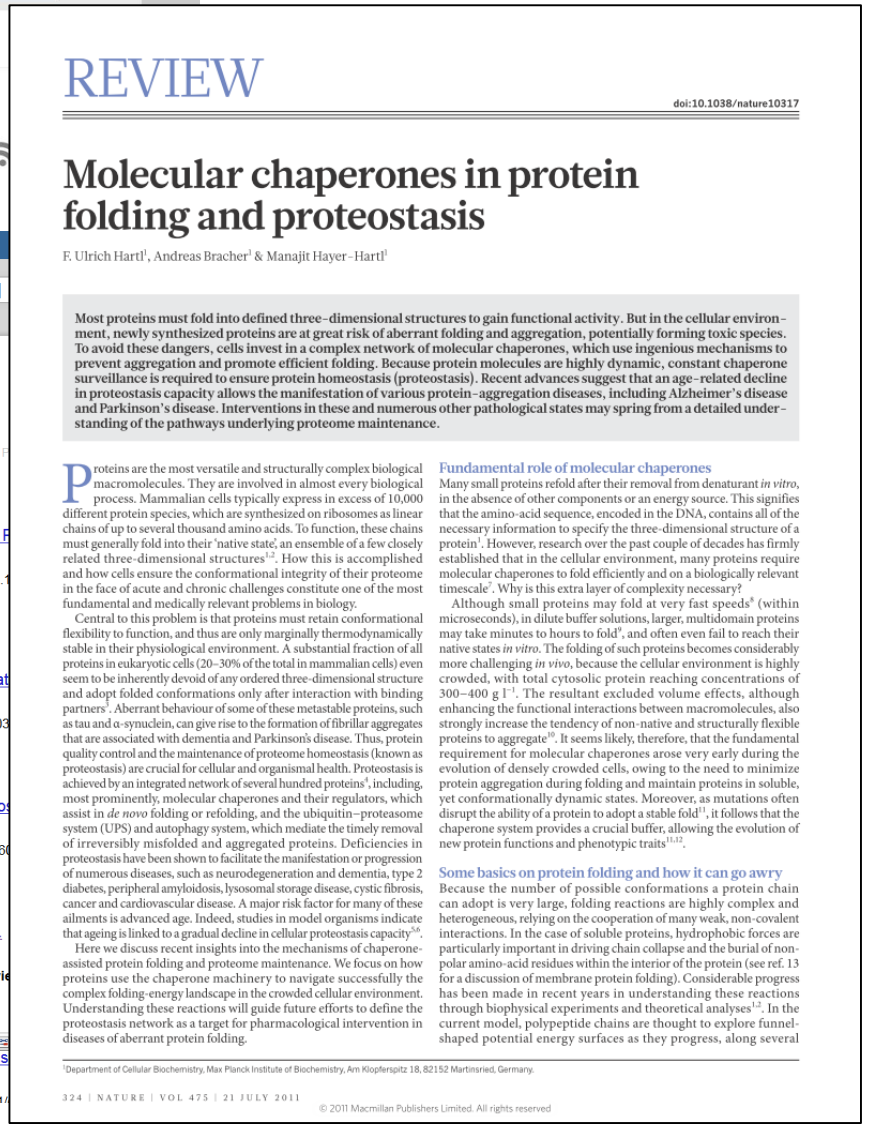
Body of knowledge related to the use of organisms, cells or cell-derived constituents for the purpose of developing products which are technically, scientifically and clinically useful. Alteration of biologic function at the molecular level (i.e., GENETIC ENGINEERING) is a central focus; laboratory methods used include TRANSFECTION and CLONING technologies, sequence and structure analysis algorithms, computer databases, and gene and protein structure function analysis and prediction. (NCBI, MeSH Database)

National Center for Biotechnology Information

米国立バイオテクノロジー情報センターHP



PubMed



REVIEW

doi:10.1038/nature10317

Molecular chaperones in protein folding and proteostasis

F. Ulrich Hartl[1], Andreas Bracher[1] & Manajit Hayer-Hartl[1]

Most proteins must fold into defined three-dimensional structures to gain functional activity. But in the cellular environment, newly synthesized proteins are at great risk of aberrant folding and aggregation, potentially forming toxic species. To avoid these dangers, cells invest in a complex network of molecular chaperones, which use ingenious mechanisms to prevent aggregation and promote efficient folding. Because protein molecules are highly dynamic, constant chaperone surveillance is required to ensure protein homeostasis (proteostasis). Recent advances suggest that an age-related decline in proteostasis capacity allows the manifestation of various protein-aggregation diseases, including Alzheimer's disease and Parkinson's disease. Interventions in these and numerous other pathological states may spring from a detailed understanding of the pathways underlying proteome maintenance.

Proteins are the most versatile and structurally complex biological macromolecules. They are involved in almost every biological process. Mammalian cells typically express in excess of 10,000 different protein species, which are synthesized on ribosomes as linear chains of up to several thousand amino acids. To function, these chains must generally fold into their 'native state', an ensemble of a few closely related three-dimensional structures[1,2]. How this is accomplished and how cells ensure the conformational integrity of their proteome in the face of acute and chronic challenges constitute one of the most fundamental and medically relevant problems in biology.

Central to this problem is that proteins must retain conformational flexibility to function, and thus are only marginally thermodynamically stable in their physiological environment. A substantial fraction of all proteins in eukaryotic cells (20–30% of the total in mammalian cells) even seem to be inherently devoid of any ordered three-dimensional structure and adopt folded conformations only after interaction with binding partners[3]. Aberrant behaviour of some of these metastable proteins, such as tau and α-synuclein, can give rise to the formation of fibrillar aggregates that are associated with dementia and Parkinson's disease. Thus, protein quality control and the maintenance of proteome homeostasis (known as proteostasis) are crucial for cellular and organismal health. Proteostasis is achieved by an integrated network of several hundred proteins[4], including, most prominently, molecular chaperones and their regulators, which assist in *de novo* folding or refolding, and the ubiquitin–proteasome system (UPS) and autophagy system, which mediate the timely removal of irreversibly misfolded and aggregated proteins. Deficiencies in proteostasis have been shown to facilitate the manifestation or progression of numerous diseases, such as neurodegeneration and dementia, type 2 diabetes, peripheral amyloidosis, lysosomal storage disease, cystic fibrosis, cancer and cardiovascular disease. A major risk factor for many of these ailments is advanced age. Indeed, studies in model organisms indicate that ageing is linked to a gradual decline in cellular proteostasis capacity[5,6].

Here we discuss recent insights into the mechanisms of chaperone-assisted protein folding and proteome maintenance. We focus on how proteins use the chaperone machinery to navigate successfully the complex folding-energy landscape in the crowded cellular environment. Understanding these reactions will guide future efforts to define the proteostasis network as a target for pharmacological intervention in diseases of aberrant protein folding.

Fundamental role of molecular chaperones

Many small proteins refold after their removal from denaturant *in vitro*, in the absence of other components or an energy source. This signifies that the amino-acid sequence, encoded in the DNA, contains all of the necessary information to specify the three-dimensional structure of a protein[1]. However, research over the past couple of decades has firmly established that in the cellular environment, many proteins require molecular chaperones to fold efficiently and on a biologically relevant timescale[7]. Why is this extra layer of complexity necessary?

Although small proteins may fold at very fast speeds[8] (within microseconds), in dilute buffer solutions, larger, multidomain proteins may take minutes to hours to fold[9], and often even fail to reach their native states *in vitro*. The folding of such proteins becomes considerably more challenging *in vivo*, because the cellular environment is highly crowded, with total cytosolic protein reaching concentrations of 300–400 g l^{-1}. The resultant excluded volume effects, although enhancing the functional interactions between macromolecules, also strongly increase the tendency of non-native and structurally flexible proteins to aggregate[10]. It seems likely, therefore, that the fundamental requirement for molecular chaperones arose very early during the evolution of densely crowded cells, owing to the need to minimize protein aggregation during folding and maintain proteins in soluble, yet conformationally dynamic states. Moreover, as mutations often disrupt the ability of a protein to adopt a stable fold[11], it follows that the chaperone system provides a crucial buffer, allowing the evolution of new protein functions and phenotypic traits[11,12].

Some basics on protein folding and how it can go awry

Because the number of possible conformations a protein chain can adopt is very large, folding reactions are highly complex and heterogeneous, relying on the cooperation of many weak, non-covalent interactions. In the case of soluble proteins, hydrophobic forces are particularly important in driving chain collapse and the burial of non-polar amino-acid residues within the interior of the protein (see ref. 13 for a discussion of membrane protein folding). Considerable progress has been made in recent years in understanding these reactions through biophysical experiments and theoretical analyses[1,2]. In the current model, polypeptide chains are thought to explore funnel-shaped potential energy surfaces as they progress, along several

[1]Department of Cellular Biochemistry, Max Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany.

そもそも、バイオテクノロジーってなに！？

生物学の知見を元に、実社会に有用な利用法をもたらす技術

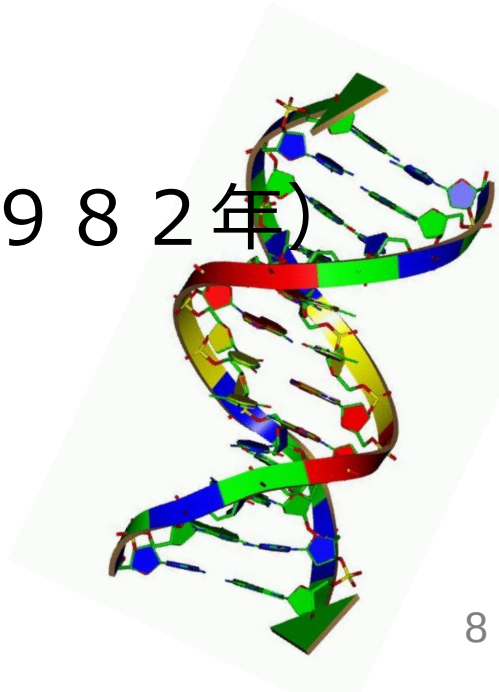
(nature.com)

古くは「醸造・発酵技術」、近年では、再生医療や創薬に応用される「細胞工学」、生物の遺伝情報をつかさどるDNAの組み替え技術を研究する「遺伝子工学」等を包括する学問領域

遺伝子工学

- ヒトゲノム解析計画（２００３年完了）
  - **癌などの疾病に関わる遺伝子の解明**
  - **エピジェネティクス**
    （分化能・リプログラミング・癌化などにかかわる）

- 遺伝子組換え技術
  - 遺伝子組換え医薬：インスリン（１９８２年）
  - 遺伝子組換え作物（１９８０年代）
  - 基礎生物学・医学研究



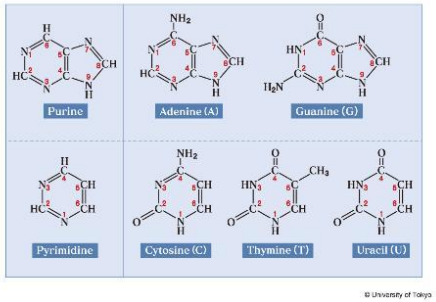
遺伝・遺伝子・DNAとは・・・

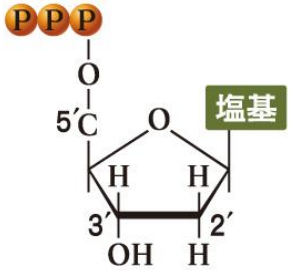
- 遺伝とは、形質が親から子へと伝達されること
  - 形質とは、遺伝子の発現を受け、個体に観察形態的・機能的単位・・・毛髪や肌の色など
  - 表現型とは、個々の観察できる形質のタイプ
- 遺伝子とは、生物の設計図
  - 遺伝する形質を発現させる遺伝物質（DNA情報）の事。染色体上の定まった場所にあり、遺伝の単位。
- 遺伝子は情報であり、ＤＮＡはそれを記録する記憶媒体

DNA二重螺旋構造の発見（1953年）

1869　白血球から酸性物質DNAの単離　（ミーシャー）

1920s DNAが4種類の塩基と糖、リン酸から構成される



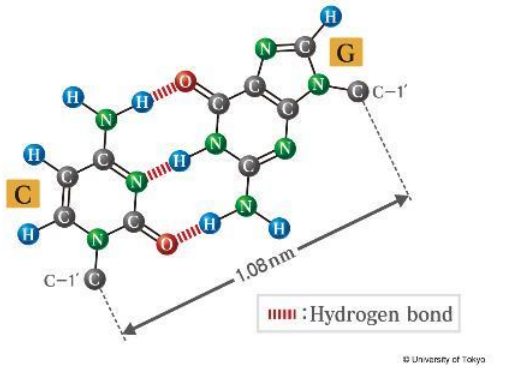


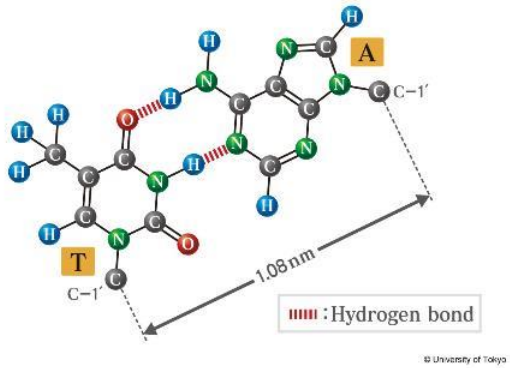
1944　遺伝子を支配するのはDNA　（アベリー）
形質転換物質にRNA分解酵素やタンパク質分解酵素処理を施しても形質転換能力を失わないが，DNA分解酵素で処理すると失活する。すなわち，形質転換物質はDNAである

それまでは，タンパク質が形質転換物質であるという説が有力であった

DNA二重螺旋構造の発見（1953年）

1950　どの生物のDNA中のアデニン（グアニン）の量とチミン（シトシン）の量が同じ<シャルガフ則>

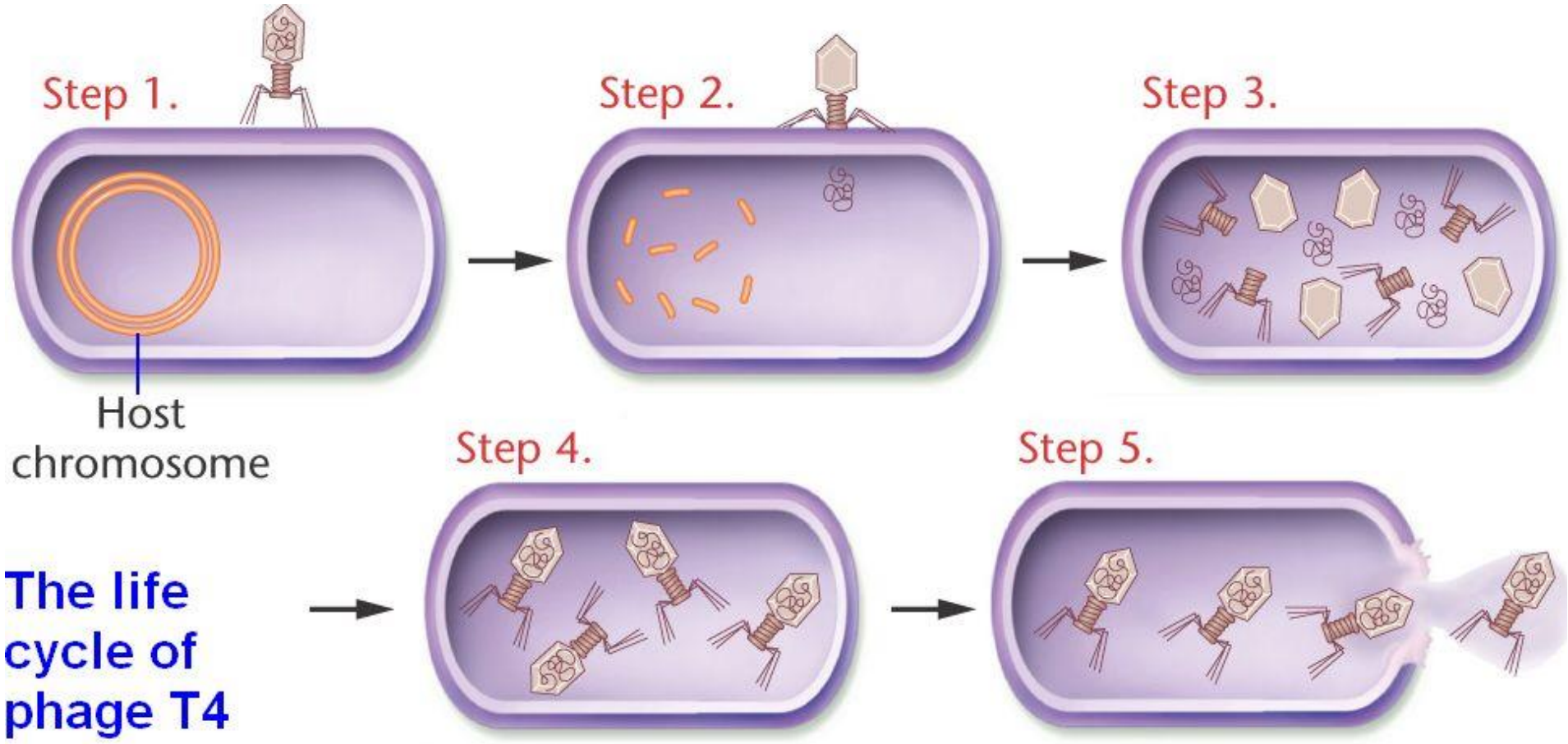




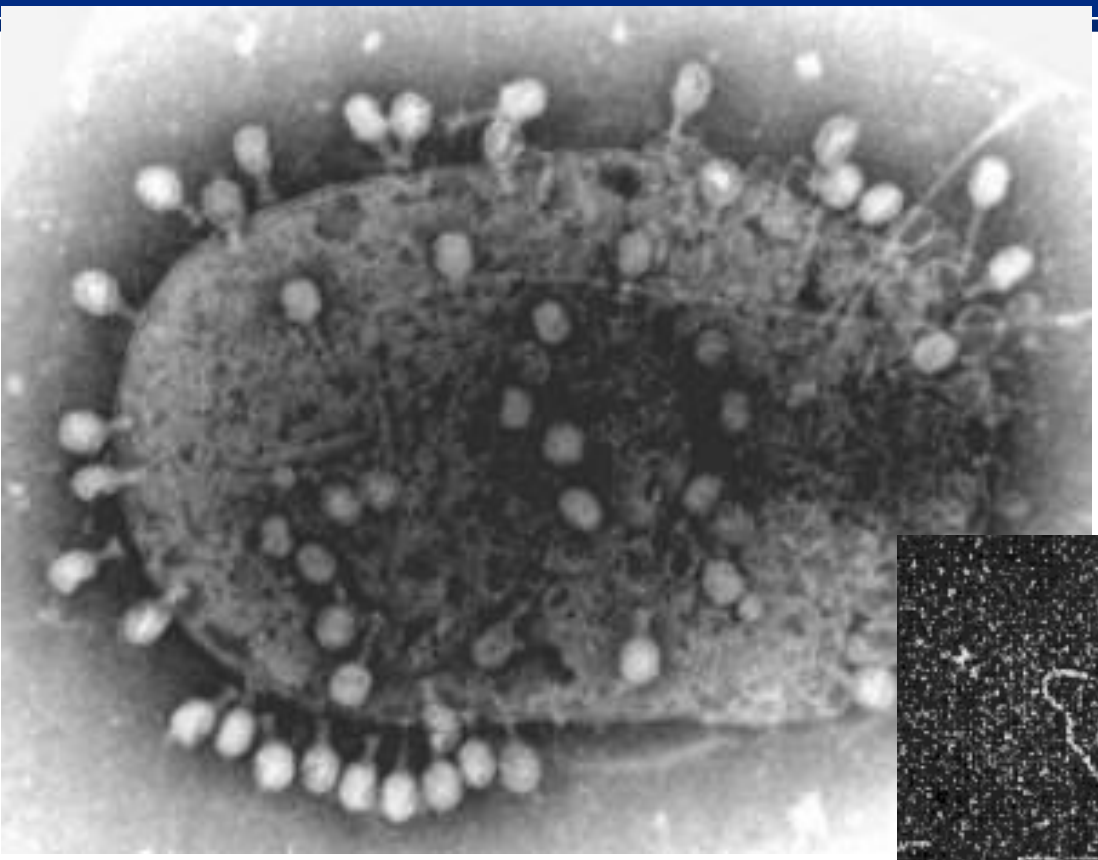
A ＋ G ＝ T ＋ C

1952　遺伝子の本体はタンパク質ではなくDNA
ファージDNAを^{32}Pで標識，タンパク質を^{35}Sで標識し大腸菌に感染させたところ，^{32}Pのみが大腸菌に体内に入り，増殖したファージが得られた。

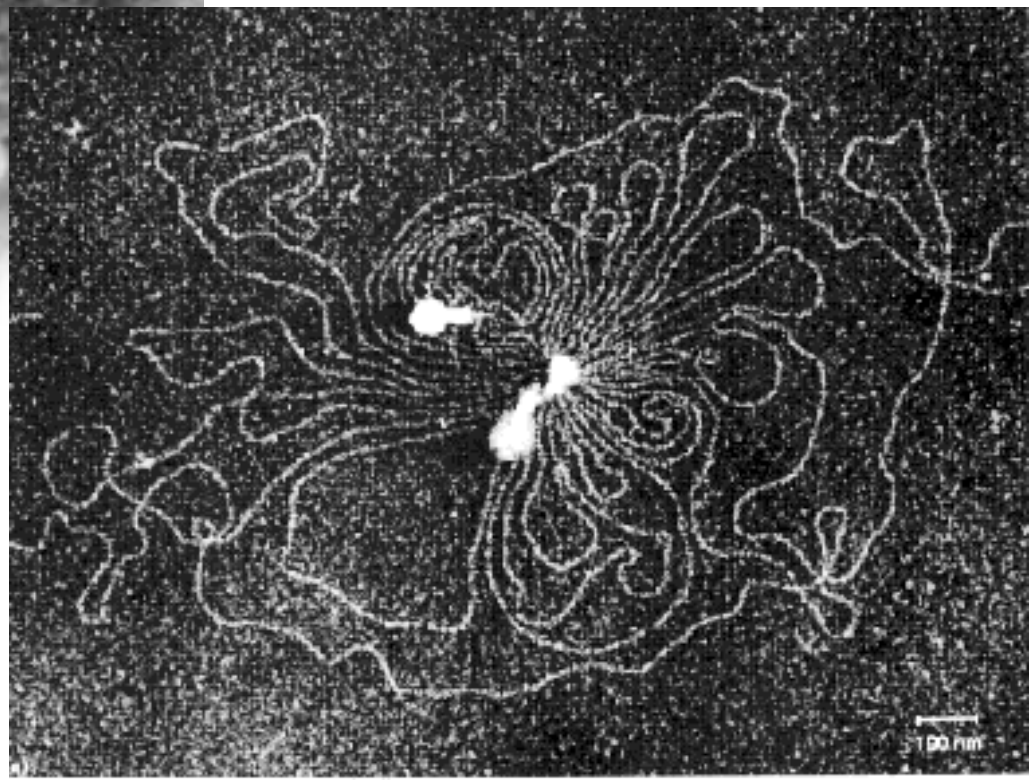
ウイルスの増殖


Step 1.
Step 2.
Step 3.
Host chromosome
Step 4.
Step 5.
The life cycle of phage T4

遺伝子組換え技術



Phage T4 virions poised on the surface of an *E. coli* cell. (Courtesy of Cornell Integrated Microscopy Center.)



A classic dark field image of the DNA molecule released by rupture of a single T4 capsid

DNA二重螺旋構造の発見（1953年）

1895　レントゲンがX線を発見

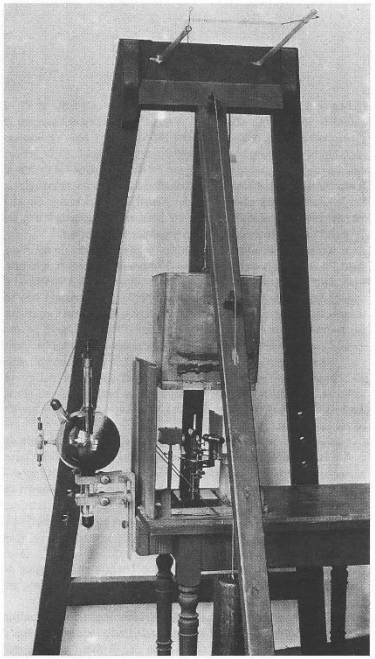
1912　ラウエがX線回折現象を発見

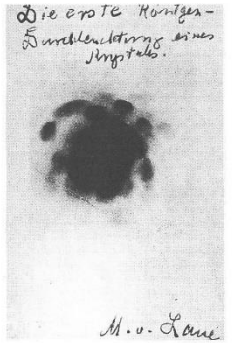
(1914年　ノーベル物理学賞)

1913　ブラッグ親子がブラッグの法則を発表。 X線回折による構造解析が可能に

1940s タンパク質分子はα型と言われるらせん状の構造をしている（アストベリー）

1951　タンパク質の2次構造・αヘリックスの発見（ポーリング）
DNA三重螺旋構造モデル

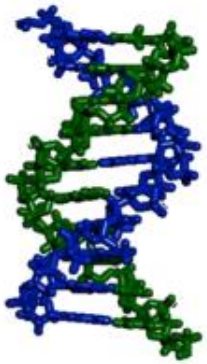




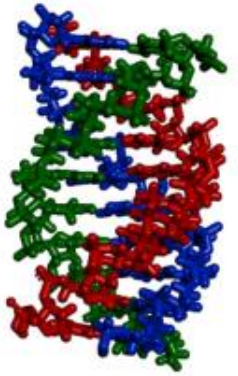
最初のラウエ・ダイアグラム
自分たちの発見に感動しながら，ラウエは同僚たちに最初のラウエ・ダイアグラムの写真を送っている。

左写真は
ヴァルター・フリードリッヒとパウル・クニッピングの実験装置
1912年4月，この実験装置でX線の干渉が発見された。この装置の現物が，現在ミュンヘンのドイツ博物館にある。

アーミン・ヘルマン 著
「アインシュタイン時代－物理学が世界史になる－」
地人書館(1993年刊) より引用

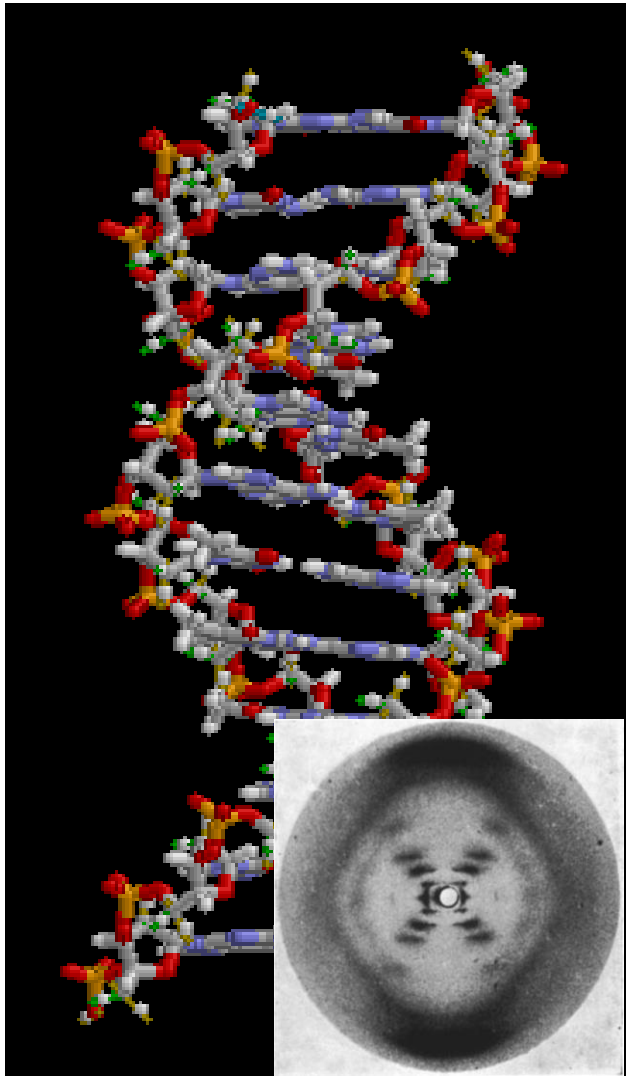


DNA二重鎖



DNA三重鎖

DNA二重螺旋構造の発見（1953年）



1869　白血球から酸性物質DNAの単離

1920s DNAが4種類の塩基と糖、リン酸から構成される

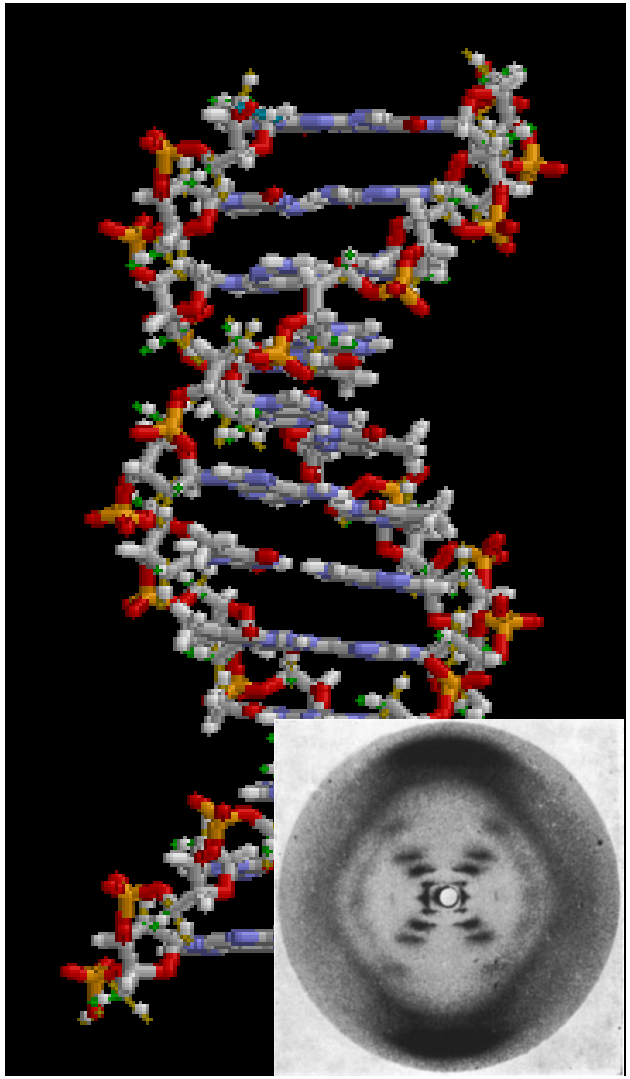
1944　形質転換物質がDNA

1950　どの生物のDNA中のアデニン（グアニン）の量とチミン（シトシン）の量が同じ

1952　遺伝子の本体はタンパク質ではなくDNA

1953　フランクリンによるDNAのX線回折写真の報告

DNA二重螺旋構造の発見（1953年）



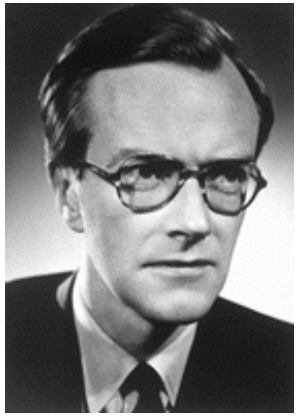
R. Franklin, *Nature*, **171**, 740 (1953)

J. Watson

(1928-)
動物学・分子生物学

F. Crick

(1916-2004)
物理学→生物学



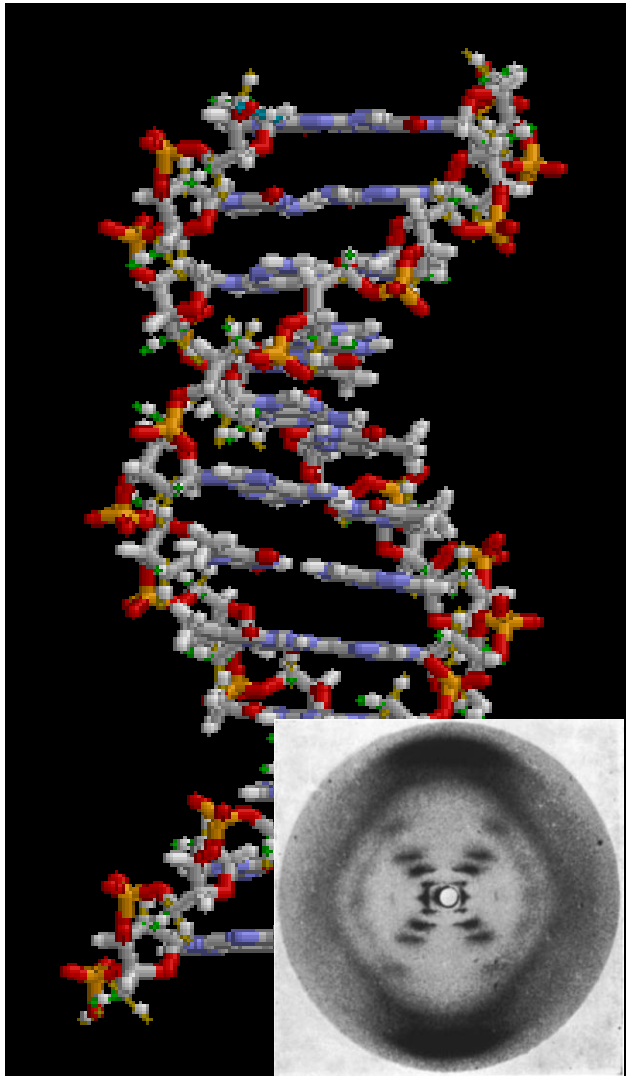
M. Wilkins

(1916-2004)
物理学→生物物理

R. Franklin

(1920-1958)
物理化学・結晶学

DNA二重螺旋構造の発見（1953年）



J. Watson

結晶化失敗

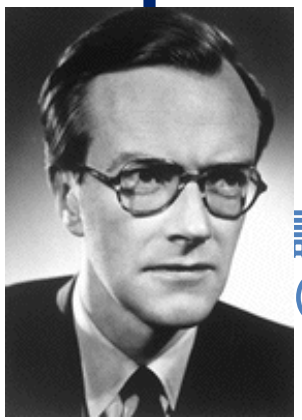
共同研究者

Wilkinsに内緒で
データを分析

F. Crick

結晶化失敗

意見を聞くために
写真を提示



M. Wilkins

結晶化失敗

競争相手

講師・実験助手
（仲が良くなかった）

R. Franklin

DNA提供

R. Franklin, *Nature*, **171**, 740 (1953)

複数のDNA構造を結晶化

DNA二重螺旋構造の発見（1953年）

equipment, and to Dr. G. E. R. Deacon and the captain and officers of R.R.S. *Discovery II* for their part in making the observations.

[1] Young, F. B., Gerrard, H., and Jevons, W., *Phil. Mag.*, **40**, 149 (1920).
[2] Longuet-Higgins, M. S., *Mon. Not. Roy. Astro. Soc., Geophys. Supp.*, **5**, 285 (1949).
[3] Von Arx, W. S., Woods Hole Papers in Phys. Oceanog. Meteor., **11** (3) (1950).
[4] Ekman, V. W., *Arkiv. Mat. Astron. Fysik. (Stockholm)*, **2** (11) (1905).

MOLECULAR STRUCTURE OF NUCLEIC ACIDS

A Structure for Deoxyribose Nucleic Acid

We wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

A structure for nucleic acid has already been proposed by Pauling and Corey[1]. They kindly made their manuscript available to us in advance of publication. Their model consists of three intertwined chains, with the phosphates near the fibre axis, and the bases on the outside. In our opinion, this structure is unsatisfactory for two reasons: (1) We believe that the material which gives the X-ray diagrams is the salt, not the free acid. Without the acidic hydrogen atoms it is not clear what forces would hold the structure together, especially as the negatively charged phosphates near the axis will repel each other. (2) Some of the van der Waals distances appear to be too small.

Another three-chain structure has also been suggested by Fraser (in the press). In his model the phosphates are on the outside and the bases on the inside, linked together by hydrogen bonds. This structure as described is rather ill-defined, and for this reason we shall not comment on it.

This figure is purely diagrammatic. The two ribbons symbolize the two phosphate—sugar chains, and the horizontal rods the pairs of bases holding the chains together. The vertical line marks the fibre axis

We wish to put forward a radically different structure for the salt of deoxyribose nucleic acid. This structure has two helical chains each coiled round the same axis (see diagram). We have made the usual chemical assumptions, namely, that each chain consists of phosphate diester groups joining β-D-deoxyribofuranose residues with 3′,5′ linkages. The two chains (but not their bases) are related by a dyad perpendicular to the fibre axis. Both chains follow right-handed helices, but owing to the dyad the sequences of the atoms in the two chains run in opposite directions. Each chain loosely resembles Furberg's[2] model No. 1; that is, the bases are on the inside of the helix and the phosphates on the outside. The configuration of the sugar and the atoms near it is close to Furberg's 'standard configuration', the sugar being roughly perpendicular to the attached base. There is a residue on each chain every 3·4 A. in the *z*-direction. We have assumed an angle of 36° between adjacent residues in the same chain, so that the structure repeats after 10 residues on each chain, that is, after 34 A. The distance of a phosphorus atom from the fibre axis is 10 A. As the phosphates are on the outside, cations have easy access to them.

The structure is an open one, and its water content is rather high. At lower water contents we would expect the bases to tilt so that the structure could become more compact.

The novel feature of the structure is the manner in which the two chains are held together by the purine and pyrimidine bases. The planes of the bases are perpendicular to the fibre axis. They are joined together in pairs, a single base from one chain being hydrogen-bonded to a single base from the other chain, so that the two lie side by side with identical *z*-co-ordinates. One of the pair must be a purine and the other a pyrimidine for bonding to occur. The hydrogen bonds are made as follows: purine position 1 to pyrimidine position 1; purine position 6 to pyrimidine position 6.

If it is assumed that the bases only occur in the structure in the most plausible tautomeric forms (that is, with the keto rather than the enol configurations) it is found that only specific pairs of bases can bond together. These pairs are: adenine (purine) with thymine (pyrimidine), and guanine (purine) with cytosine (pyrimidine).

In other words, if an adenine forms one member of a pair, on either chain, then on these assumptions the other member must be thymine; similarly for guanine and cytosine. The sequence of bases on a single chain does not appear to be restricted in any way. However, if only specific pairs of bases can be formed, it follows that if the sequence of bases on one chain is given, then the sequence on the other chain is automatically determined.

It has been found experimentally[3,4] that the ratio of the amounts of adenine to thymine, and the ratio of guanine to cytosine, are always very close to unity for deoxyribose nucleic acid.

It is probably impossible to build this structure with a ribose sugar in place of the deoxyribose, as the extra oxygen atom would make too close a van der Waals contact.

The previously published X-ray data[5,6] on deoxyribose nucleic acid are insufficient for a rigorous test of our structure. So far as we can tell, it is roughly compatible with the experimental data, but it must be regarded as unproved until it has been checked against more exact results. Some of these are given in the following communications. We were not aware of the details of the results presented there when we devised our structure, which rests mainly though not entirely on published experimental data and stereochemical arguments.

It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

Full details of the structure, including the conditions assumed in building it, together with a set of co-ordinates for the atoms, will be published elsewhere.

We are much indebted to Dr. Jerry Donohue for constant advice and criticism, especially on interatomic distances. We have also been stimulated by a knowledge of the general nature of the unpublished experimental results and ideas of Dr. M. H. F. Wilkins, Dr. R. E. Franklin and their co-workers at

King's College, London. One of us (J. D. W.) has been aided by a fellowship from the National Foundation for Infantile Paralysis.

J. D. Watson
F. H. C. Crick
Medical Research Council Unit for the Study of the Molecular Structure of Biological Systems,
Cavendish Laboratory, Cambridge.
April 2.

[1] Pauling, L., and Corey, R. B., *Nature*, **171**, 346 (1953); *Proc. U.S. Nat. Acad. Sci.*, **39**, 84 (1953).
[2] Furberg, S., *Acta Chem. Scand.*, **6**, 634 (1952).
[3] Chargaff, E., for references see Zamenhof, S., Brawerman, G., and Chargaff, E., *Biochim. et Biophys. Acta*, **9**, 402 (1952).
[4] Wyatt, G. R., *J. Gen. Physiol.*, **36**, 201 (1952).
[5] Astbury, W. T., Symp. Soc. Exp. Biol. 1, Nucleic Acid, 66 (Camb. Univ. Press, 1947).
[6] Wilkins, M. H. F., and Randall, J. T., *Biochim. et Biophys. Acta*, **10**, 192 (1953).

Molecular Structure of Deoxypentose Nucleic Acids

While the biological properties of deoxypentose nucleic acid suggest a molecular structure containing great complexity, X-ray diffraction studies described here (cf. Astbury[1]) show the basic molecular configuration has great simplicity. The purpose of this communication is to describe, in a preliminary way, some of the experimental evidence for the polynucleotide chain configuration being helical, and existing in this form when in the natural state. A fuller account of the work will be published shortly.

The structure of deoxypentose nucleic acid is the same in all species (although the nitrogen base ratios alter considerably) in nucleoprotein, extracted or in cells, and in purified nucleate. The same linear group of polynucleotide chains may pack together parallel in different ways to give crystalline[1–3], semi-crystalline or paracrystalline material. In all cases the X-ray diffraction photograph consists of two regions, one determined largely by the regular spacing of nucleotides along the chain, and the other by the longer spacings of the chain configuration. The sequence of different nitrogen bases along the chain is not made visible.

Oriented paracrystalline deoxypentose nucleic acid ('structure *B*' in the following communication by Franklin and Gosling) gives a fibre diagram as shown in Fig. 1 (cf. ref. 4). Astbury suggested that the strong 3·4-A. reflexion corresponded to the internucleotide repeat along the fibre axis. The ~ 34 A. layer lines, however, are not due to a repeat of a polynucleotide composition, but to the chain configuration repeat, which causes strong diffraction as the nucleotide chains have higher density than the interstitial water. The absence of reflexions on or near the meridian immediately suggests a helical structure with axis parallel to fibre length.

Diffraction by Helices

It may be shown[5] (also Stokes, unpublished) that the intensity distribution in the diffraction pattern of a series of points equally spaced along a helix is given by the squares of Bessel functions. A uniform continuous helix gives a series of layer lines of spacing corresponding to the helix pitch, the intensity distribution along the *n*th layer line being proportional to the square of J_n, the *n*th order Bessel function. A straight line may be drawn approximately through

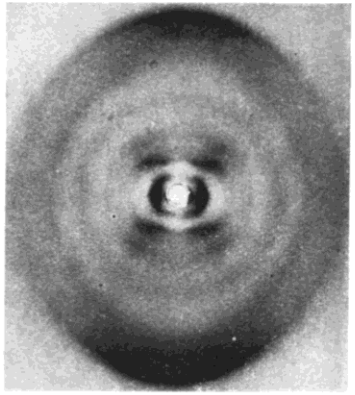


Fig. 1. Fibre diagram of deoxypentose nucleic acid from *B. coli*. Fibre axis vertical

the innermost maxima of each Bessel function and the origin. The angle this line makes with the equator is roughly equal to the angle between an element of the helix and the helix axis. If a unit repeats *n* times along the helix there will be a meridional reflexion (J_n^2) on the *n*th layer line. The helical configuration produces side-bands on this fundamental frequency, the effect[5] being to reproduce the intensity distribution about the origin around the new origin, on the *n*th layer line, corresponding to *C* in Fig. 2.

We will now briefly analyse in physical terms some of the effects of the shape and size of the repeat unit or nucleotide on the diffraction pattern. First, if the nucleotide consists of a unit having circular symmetry about an axis parallel to the helix axis, the whole diffraction pattern is modified by the form factor of the nucleotide. Second, if the nucleotide consists of a series of points on a radius at right-angles to the helix axis, the phases of radiation scattered by the helices of different diameter passing through each point are the same. Summation of the corresponding Bessel functions gives reinforcement for the inner-

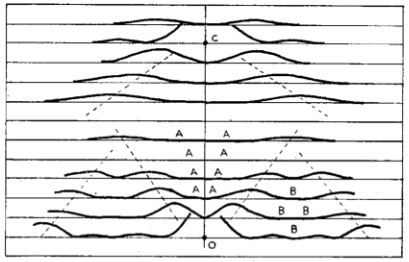


Fig. 2. Diffraction pattern of system of helices corresponding to structure of deoxypentose nucleic acid. The squares of Bessel functions are plotted about 0 on the equator and on the first, second, third and fifth layer lines for half of the nucleotide mass at 20 A. diameter and remainder distributed along a radius, the mass at a given radius being proportional to the radius. About *C* on the tenth layer line similar functions are plotted for an outer

DNA二重螺旋構造の発見（1953年）

We wish to thank Prof. J. T. Randall for encouragement; Profs. E. Chargaff, R. Signer, J. A. V. Butler and Drs. J. D. Watson, J. D. Smith, L. Hamilton, J. C. White and G. R. Wyatt for supplying material without which this work would have been impossible; also Drs. J. D. Watson and Mr. F. H. C. Crick for stimulation, and our colleagues R. E. Franklin, R. G. Gosling, G. L. Brown and W. E. Seeds for discussion. One of us (H. R. W.) wishes to acknowledge the award of a University of Wales Fellowship.

M. H. F. Wilkins
Medical Research Council Biophysics Research Unit,
A. R. Stokes
H. R. Wilson
Wheatstone Physics Laboratory, King's College, London.
April 2.

[1] Astbury, W. T., Symp. Soc. Exp. Biol., 1, Nucleic Acid (Cambridge Univ. Press, 1947).
[2] Riley, D. P., and Oster, G., *Biochim. et Biophys. Acta*, **7**, 526 (1951).
[3] Wilkins, M. H. F., Gosling, R. G., and Seeds, W. E., *Nature*, **167**, 759 (1951).
[4] Astbury, W. T., and Bell, F. O., Cold Spring Harb. Symp. Quant. Biol., **6**, 109 (1938).
[5] Cochran, W., Crick, F. H. C., and Vand, V., *Acta Cryst.*, **5**, 581 (1952).
[6] Wilkins, M. H. F., and Randall, J. T., *Biochim. et Biophys. Acta*, **10**, 192 (1953).

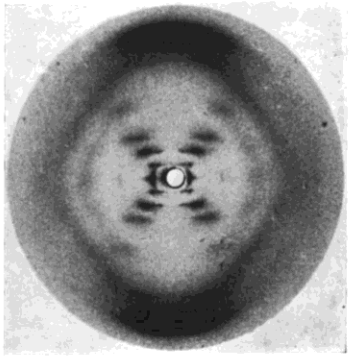
Molecular Configuration in Sodium Thymonucleate

Sodium thymonucleate fibres give two distinct types of X-ray diagram. The first corresponds to a crystalline form, structure *A*, obtained at about 75 per cent relative humidity; a study of this is described in detail elsewhere[1]. At higher humidities a different structure, structure *B*, showing a lower degree of order, appears and persists over a wide range of ambient humidity. The change from *A* to *B* is reversible. The water content of structure *B* fibres which undergo this reversible change may vary from 40–50 per cent to several hundred per cent of the dry weight. Moreover, some fibres never show structure *A*, and in these structure *B* can be obtained with an even lower water content.

The X-ray diagram of structure *B* (see photograph) shows in striking manner the features characteristic of helical structures, first worked out in this laboratory by Stokes (unpublished) and by Crick, Cochran and Vand[2]. Stokes and Wilkins were the first to propose such structures for nucleic acid as a result of direct studies of nucleic acid fibres, although a helical structure had been previously suggested by Furberg (thesis, London, 1949) on the basis of X-ray studies of nucleosides and nucleotides.

While the X-ray evidence cannot, at present, be taken as direct proof that the structure is helical, other considerations discussed below make the existence of a helical structure highly probable.

Structure *B* is derived from the crystalline structure *A* when the sodium thymonucleate fibres take up quantities of water in excess of about 40 per cent of their weight. The change is accompanied by an increase of about 30 per cent in the length of the fibre, and by a substantial re-arrangement of the molecule. It therefore seems reasonable to suppose that in structure *B* the structural units of sodium thymonucleate (molecules on groups of molecules) are relatively free from the influence of neighbouring



Sodium deoxyribose nucleate from calf thymus. Structure *B*

molecules, each unit being shielded by a sheath of water. Each unit is then free to take up its least-energy configuration independently of its neighbours and, in view of the nature of the long-chain molecules involved, it is highly likely that the general form will be helical[3]. If we adopt the hypothesis of a helical structure, it is immediately possible, from the X-ray diagram of structure *B*, to make certain deductions as to the nature and dimensions of the helix.

The innermost maxima on the first, second, third and fifth layer lines lie approximately on straight lines radiating from the origin. For a smooth single-strand helix the structure factor on the nth layer line is given by:

$$F_n = J_n(2\pi rR) \exp i\, n(\psi + \tfrac{1}{2}\pi),$$

where $J_n(u)$ is the nth-order Bessel function of u, r is the radius of the helix, and R and ψ are the radial and azimuthal co-ordinates in reciprocal space[2]; this expression leads to an approximately linear array of intensity maxima of the type observed, corresponding to the first maxima in the functions J_1, J_2, J_3, etc.

If, instead of a smooth helix, we consider a series of residues equally spaced along the helix, the transform in the general case treated by Crick, Cochran and Vand is more complicated. But if there is a whole number, m, of residues per turn, the form of the transform is as for a smooth helix with the addition, only, of the same pattern repeated with its origin at heights mc^*, $2mc^*$ . . . etc. (c is the fibre-axis period).

In the present case the fibre-axis period is 34 A. and the very strong reflexion at 3·4 A. lies on the tenth layer line. Moreover, lines of maxima radiating from the 3·4-A. reflexion as from the origin are visible on the fifth and lower layer lines, having a J_5 maximum coincident with that of the origin series on the fifth layer line. (The strong outer streaks which apparently radiate from the 3·4-A. maximum are not, however, so easily explained.) This suggests strongly that there are exactly 10 residues per turn of the helix. If this is so, then from a measurement of R_n the position of the first maximum on the nth layer line (for n 5≮), the radius of the helix, can be obtained. In the present instance, measurements of R_1, R_2, R_3 and R_5 all lead to values of r of about 10 A.

Since this linear array of maxima is one of the strongest features of the X-ray diagram, we must conclude that a crystallographically important part of the molecule lies on a helix of this diameter. This can only be the phosphate groups or phosphorus atoms.

If ten phosphorus atoms lie on one turn of a helix of radius 10 A., the distance between neighbouring phosphorus atoms in a molecule is 7·1 A. This corresponds to the P . . . P distance in a fully extended molecule, and therefore provides a further indication that the phosphates lie on the outside of the structural unit.

Thus, our conclusions differ from those of Pauling and Corey[4], who proposed for the nucleic acids a helical structure in which the phosphate groups form a dense core.

We must now consider briefly the equatorial reflexions. For a single helix the series of equatorial maxima should correspond to the maxima in $J_0(2\pi rR)$. The maxima on our photograph do not, however, fit this function for the value of r deduced above. There is a very strong reflexion at about 24 A. and then only a faint sharp reflexion at 9·0 A. and two diffuse bands around 5·5 A. and 4·0 A. This lack of agreement is, however, to be expected, for we know that the helix so far considered can only be the most important member of a series of coaxial helices of different radii; the non-phosphate parts of the molecule will lie on inner co-axial helices, and it can be shown that, whereas these will not appreciably influence the innermost maxima on the layer lines, they may have the effect of destroying or shifting both the equatorial maxima and the outer maxima on other layer lines.

Thus, if the structure is helical, we find that the phosphate groups or phosphorus atoms lie on a helix of diameter about 20 A., and the sugar and base groups must accordingly be turned inwards towards the helical axis.

Considerations of density show, however, that a cylindrical repeat unit of height 34 A. and diameter 20 A. must contain many more than ten nucleotides.

Since structure *B* often exists in fibres with low water content, it seems that the density of the helical unit cannot differ greatly from that of dry sodium thymonucleate, 1·63 gm./cm.^{3} [1,5], the water in fibres of high water-content being situated outside the structural unit. On this basis we find that a cylinder of radius 10 A. and height 34 A. would contain thirty-two nucleotides. However, there might possibly be some slight inter-penetration of the cylindrical units in the dry state making their effective radius rather less. It is therefore difficult to decide, on the basis of density measurements alone, whether one repeating unit contains ten nucleotides on each of two or on each of three co-axial molecules. (If the effective radius were 8 A. the cylinder would contain twenty nucleotides.) Two other arguments, however, make it highly probable that there are only two co-axial molecules.

First, a study of the Patterson function of structure *A*, using superposition methods, has indicated[6] that there are only two chains passing through a primitive unit cell in this structure. Since the $A \rightleftharpoons B$ transformation is readily reversible, it seems very unlikely that the molecules would be grouped in threes in structure *B*. Secondly, from measurements on the X-ray diagram of structure *B* it can readily be shown that, whether the number of chains per unit is two or three, the chains are not equally spaced along the fibre axis. For example, three equally spaced chains would mean that the nth layer line depended on J_{3n}, and would lead to a helix of diameter about 60 A. This is many times larger than the primitive unit cell in structure *A*, and absurdly large in relation to the dimensions of nucleotides. Three unequally spaced chains, on the other hand, would be crystallographically non-equivalent, and this, again, seems unlikely. It therefore seems probable that there are only two co-axial molecules and that these are unequally spaced along the fibre axis.

Thus, while we do not attempt to offer a complete interpretation of the fibre-diagram of structure *B*, we may state the following conclusions. The structure is probably helical. The phosphate groups lie on the outside of the structural unit, on a helix of diameter about 20 A. The structural unit probably consists of two co-axial molecules which are not equally spaced along the fibre axis, their mutual displacement being such as to account for the variation of observed intensities of the innermost maxima on the layer lines; if one molecule is displaced from the other by about three-eighths of the fibre-axis period, this would account for the absence of the fourth layer line maxima and the weakness of the sixth. Thus our general ideas are not inconsistent with the model proposed by Watson and Crick in the preceding communication.

The conclusion that the phosphate groups lie on the outside of the structural unit has been reached previously by quite other reasoning[1]. Two principal lines of argument were invoked. The first derives from the work of Gulland and his collaborators[7], who showed that even in aqueous solution the —CO and —NH_2 groups of the bases are inaccessible and cannot be titrated, whereas the phosphate groups are fully accessible. The second is based on our own observations[1] on the way in which the structural units in structures *A* and *B* are progressively separated by an excess of water, the process being a continuous one which leads to the formation first of a gel and ultimately to a solution. The hygroscopic part of the molecule may be presumed to lie in the phosphate groups ($(C_2H_5O)_2PO_2Na$ and $(C_3H_7O)_2PO_2Na$ are highly hygroscopic[8]), and the simplest explanation of the above process is that these groups lie on the outside of the structural units. Moreover, the ready availability of the phosphate groups for interaction with proteins can most easily be explained in this way.

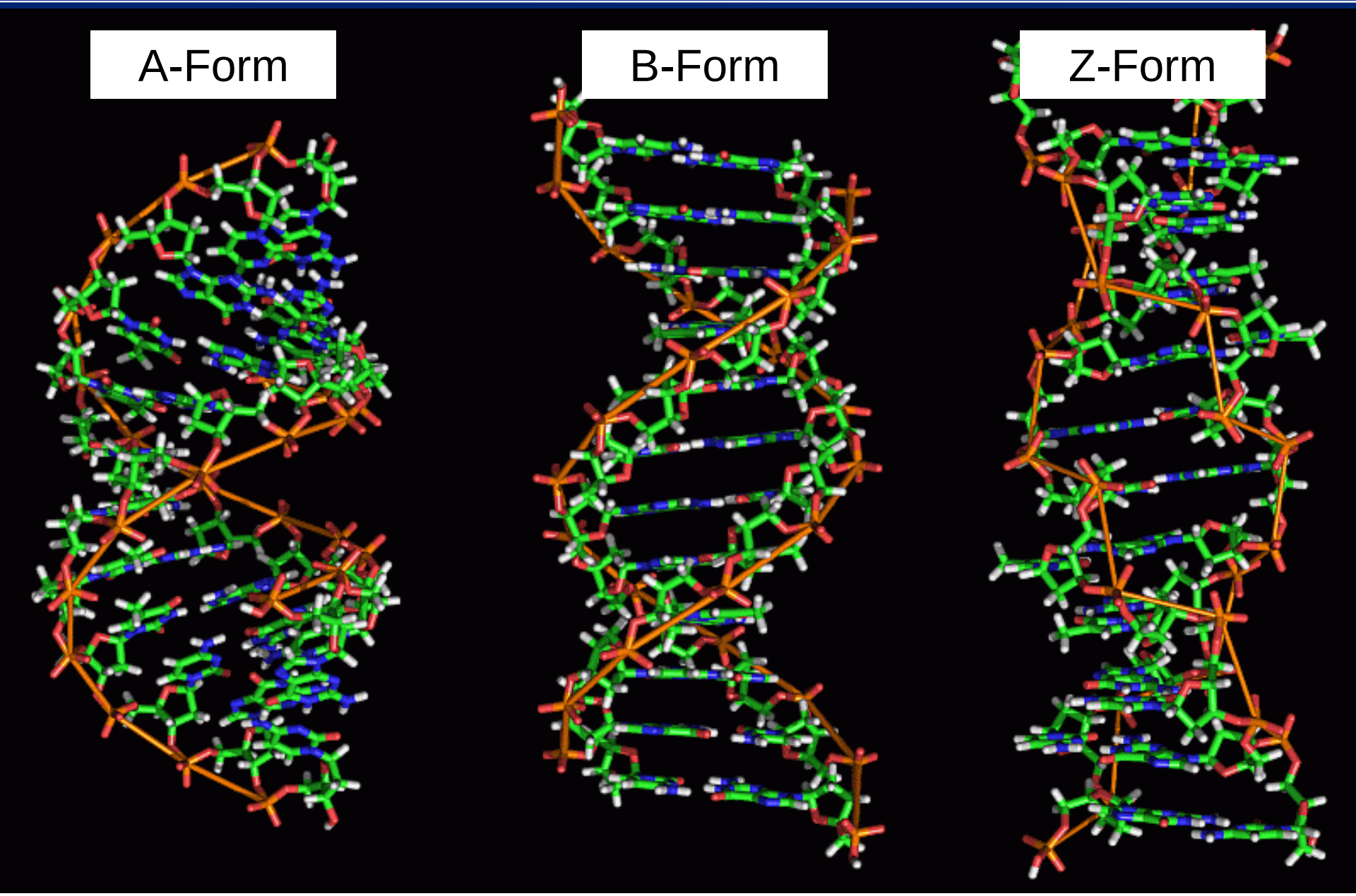
We are grateful to Prof. J. T. Randall for his interest and to Drs. F. H. C. Crick, A. R. Stokes and M. H. F. Wilkins for discussion. One of us (R. E. F.) acknowledges the award of a Turner and Newall Fellowship.

Rosalind E. Franklin*
R. G. Gosling
Wheatstone Physics Laboratory, King's College, London.
April 2.

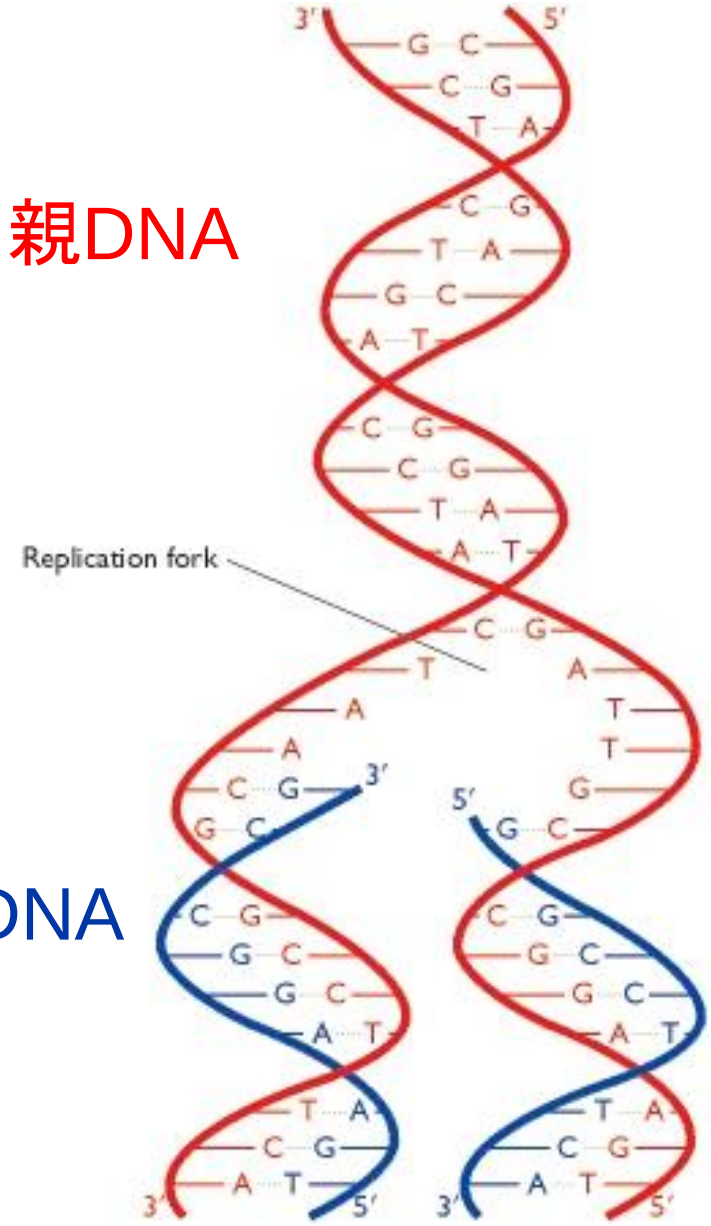
* Now at Birkbeck College Research Laboratories, 21 Torrington Square, London, W.C.1.

[1] Franklin, R. E., and Gosling, R. G. (in the press).
[2] Cochran, W., Crick, F. H. C., and Vand, V., *Acta Cryst.*, **5**, 501 (1952).
[3] Pauling, L., Corey, R. B., and Branson, H. R., *Proc. U.S. Nat. Acad. Sci.*, **37**, 205 (1951).
[4] Pauling, L., and Corey, R. B., *Proc. U.S. Nat. Acad. Sci.*, **39**, 84 (1953).
[5] Astbury, W. T., Cold Spring Harbor Symp. on Quant. Biol., **12**, 56 (1947).
[6] Franklin, R. E., and Gosling, R. G. (to be published).
[7] Gulland, J. M., and Jordan, D. O., Cold Spring Harbor Symp. on Quant. Biol., **12**, 5 (1947).
[8] Drushel, W. A., and Felty, A. R., *Chem. Zent.*, **89**, 1016 (1918).

様々なDNAの結晶構造



DNA複製　（DNAの半保存的自己複製）



DNA複製

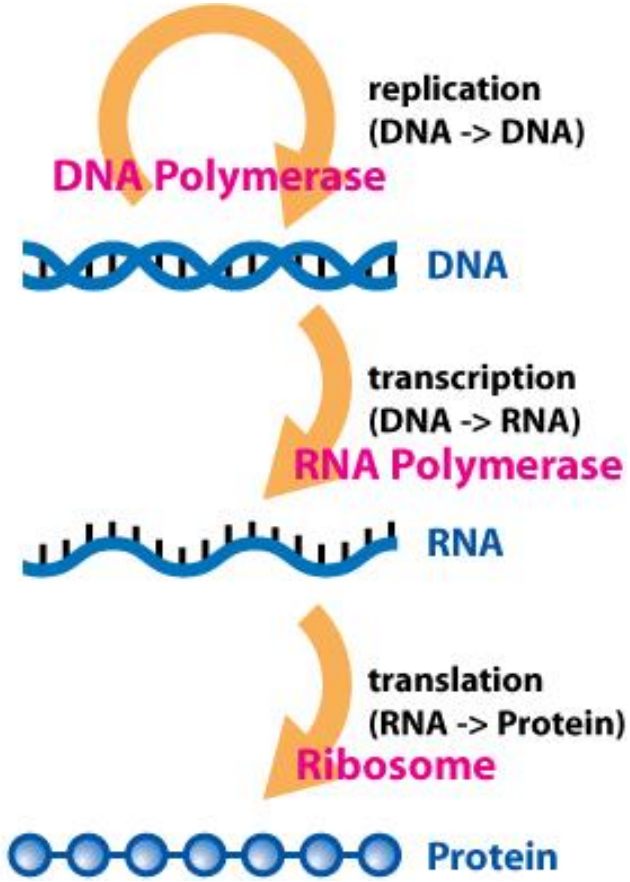
1. 塩基対の水素結合が壊れ、2つの鎖がよじれを直し分かれ、鋳型となる
2. 細胞中の遊離ヌクレオチドが鋳型となる鎖の塩基の配列と特定な関係で並ぶ
3. 並んだヌクレオチドの隣同士が結合される

娘DNA

1958年に、メセルソンによって証明

セントラルドグマ（Central dogma）

- 分子生物学の中心原理　（一部修正が加えられた）

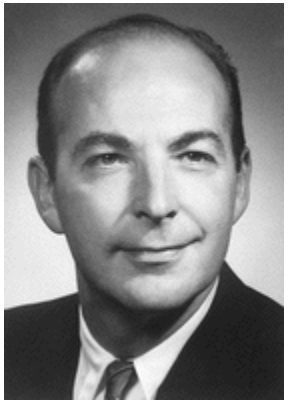


– DNA複製
  - DNAとは、「生命の設計図」？

– 転写
  - DNAの情報をコピーする (RNA)

– 翻訳
  - mRNAが持つタンパク質の設計図をもとに、タンパク質を合成する

DNA生合成機序の発見



A. Kornberg（コーンバーグ，生化学者，1918-2007）

DNAの生合成のメカニズムを解明
（1959年にノーベル生理学・医学賞）

仮説　ポリヌクレオチドの3′末端の水酸基と、デオキシリボヌクレオチドのリン酸部位が結合し、DNA鎖は一塩基分ずつ伸張する

実験　4種のデオキシリボヌクレオチドと大腸菌の抽出物を混合

結果　0.005%の収量－＞DNA合成酵素の発見