Molecular Biology/The Coding Problem
Unification of Genetics and Biochemistry

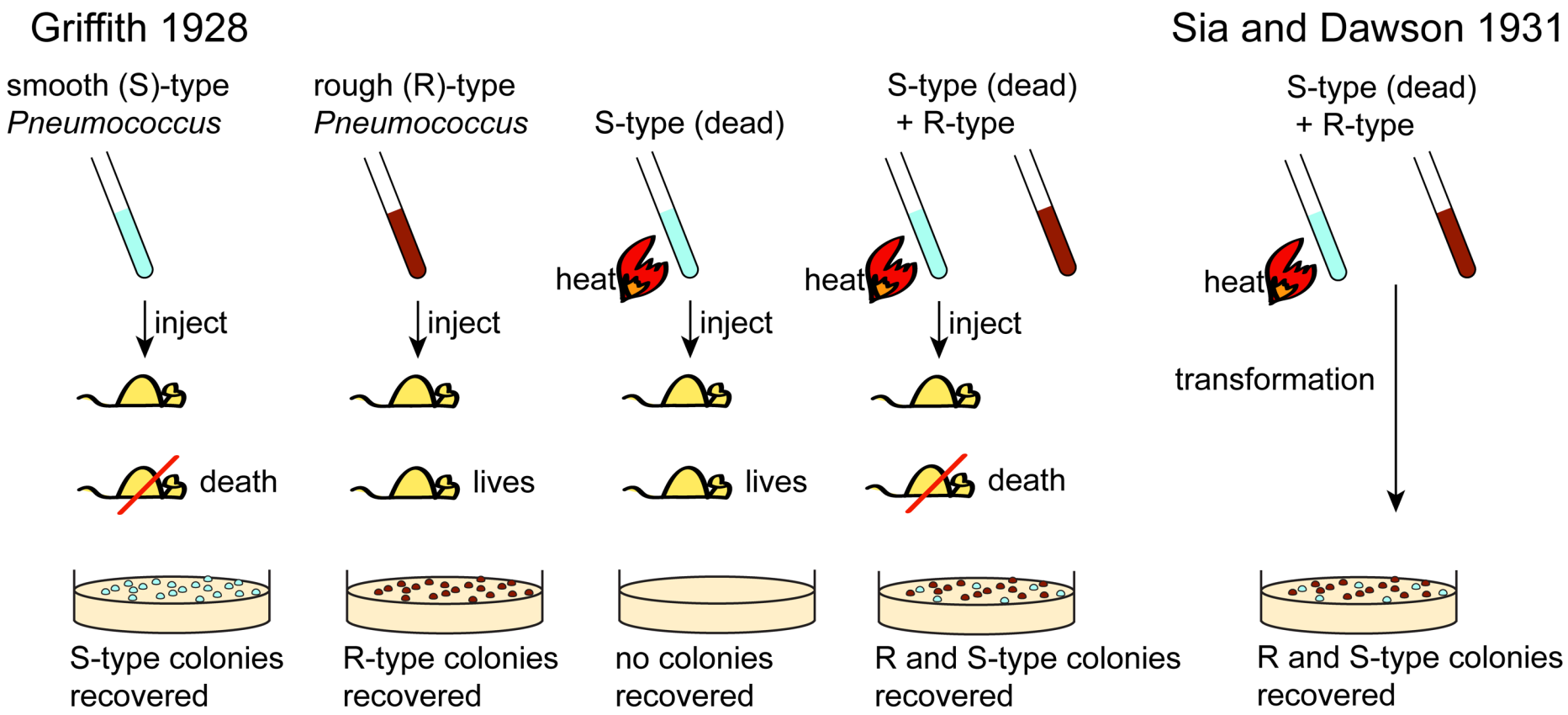
What is the physical basis of heredity? What are chromosomes made of?

- 1869-1871 Friedrich Miescher studying nuclei of pus cells isolates DNA describes “nuclein”

- 1889 Richard Altmann shows the substance can be split into protein and nucleic acid

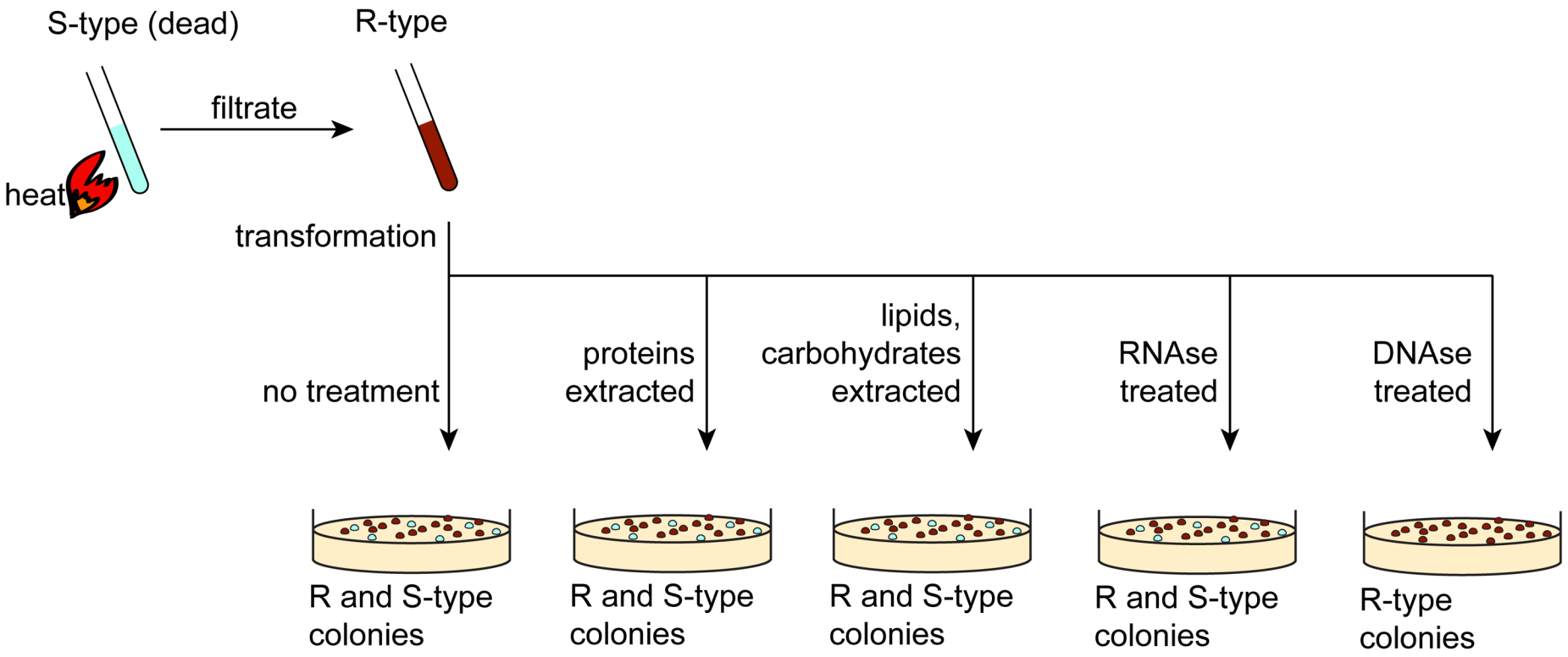
- 1895 Edmund Beecher Wilson writes:
  - Now chromatin is known to be closely similar to, if not identical with, a substance known as nuclein – which analysis shows to be a tolerably definite compound composed of nucleic acid (a complex organic compound rich in phosphorus) and albumin [protein]. And thus we reach the remarkable conclusion that inheritance may, perhaps, be effected by the physical transmission of a particular chemical compound from parent to offspring.

Bacterial transformation

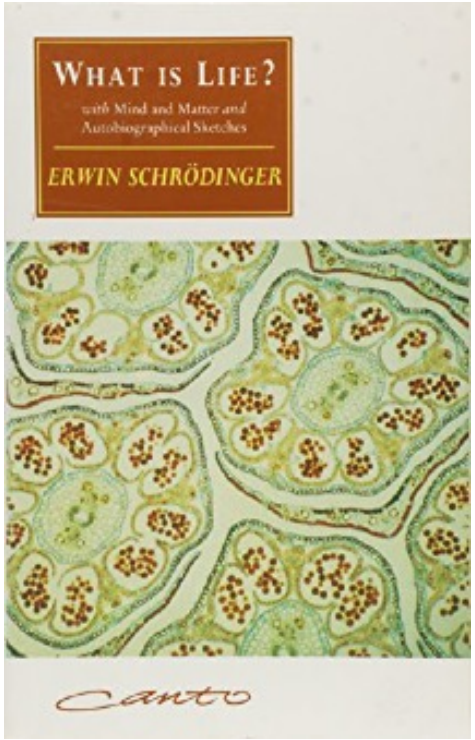


DNA is the heredity molecule

Avery et al. 1944



"What Is Life?"
The physical Aspect of the Living Cell



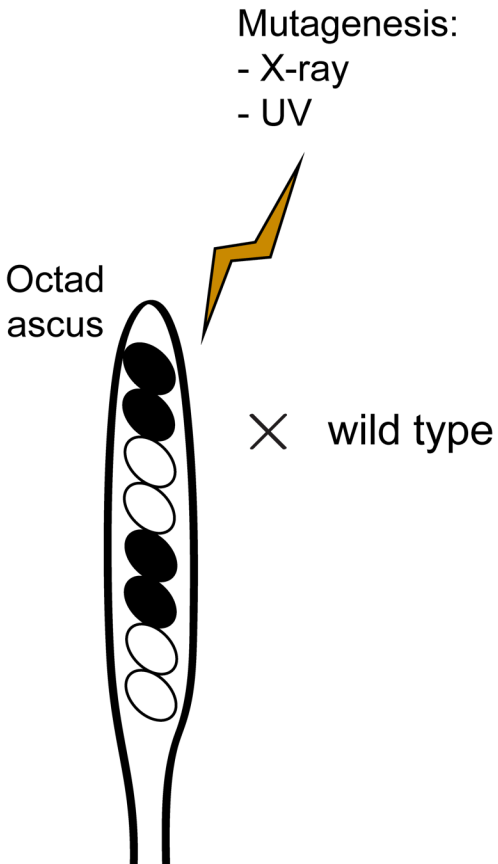
- Published in 1944- how can events that take place within an organism be accounted for by physics and chemistry?

- Organisms evade the 2nd law of thermodynamics and exhibit “negative entropy”
  - Must have a way to capture and store information

- Introduced the idea of an ”aperiodic crystal” that could contain genetic information in its configuration of covalent bonds
  - “Code script”

- Refers to Muller’s X-ray data (via Delbruck) to estimate the size of a gene

One Gene One Enzyme



Substance required for growth	Culture number	Treatment	References
Thiamin	1,090*	X-ray	Beadle and Tatum, 1941; Tatum and Bell, 1945
	9,185	X-ray	Tatum and Bell, 1945
	17,084	X-ray	Tatum and Bell, 1945
	50,005	Ultraviolet	Houlahan and Mitchell, unpublished
	56,501	X-ray	Tatum and Bell, 1945
Thiazole	18,558	X-ray	Tatum and Bell, 1945
Riboflavin	51,602	Ultraviolet	Mitchell and Houlahan, 1945
Pyridoxin	299*	X-ray	Beadle and Tatum, 1941; Stokes *et al.*, 1943
Nicotinic acid	3,416	X-ray	Beadle and Coonradt, 1944; Bonner and Beadle, unpublished
	4,540	X-ray	Beadle and Coonradt, 1944; Bonner and Beadle, unpublished
	39,401	Ultraviolet	Bonner and Beadle, unpublished
Pantothenic acid	5,531	X-ray	Tatum, 1944
	34,556R	Ultraviolet	This paper
p-Aminobenzoic acid	1,633	X-ray	Tatum and Beadle, 1942; Thompson *et al.*, 1943
Inositol	37,101	Ultraviolet	Beadle, 1944
	37,401R	Ultraviolet	Beadle, 1944
	46,316R	Ultraviolet	Beadle, 1944
	46,802R	Ultraviolet	Beadle, 1944
	64,001R	Ultraviolet	Beadle, 1944
Choline	34,486	Ultraviolet	Horowitz and Beadle, 1943; Horowitz et al., 1945
	34,542R	Ultraviolet	This paper
	37,903R	Ultraviolet	This paper
	47,904	Ultraviolet	Horowitz *et al.*, 1945
	66,210R	Ultraviolet	This paper
Ornithine	21,502	X-ray	Srb and Horowitz, 1944
	27,947	X-ray	Srb and Horowitz, 1944
	29,997	Ultraviolet	Srb and Horowitz, 1944
	34,105	Ultraviolet	Srb and Horowitz, 1944
Citrulline	30,330	Ultraviolet	Srb and Horowitz, 1944
	33,442	Ultraviolet	Srb and Horowitz, 1944
Arginine	36,703	Ultraviolet	Srb and Horowitz, 1944
Isoleucine, valine	16,117	X-ray	Bonner *et al.*, 1943
Leucine	33,757	Ultraviolet	Regnery, 1944; Ryan and Brand, 1944
Lysine	4,545	X-ray	Doermann, 1944
Methionine	4,894	X-ray	Horowitz *et al.*, 1945; Buss, 1944
Proline	21,863	X-ray	Horowitz *et al.*, 1945; Bonner, unpublished
Anthranilic acid	40,008	Ultraviolet	Tatum *et al.*, 1944; Tatum and Bonner, 1944
Indole	10,575	X-ray	Tatum *et al.*, 1944; Tatum and Bonner, 1944
Valine	33,050	Ultraviolet	Horowitz *et al.*, 1945; this paper
Adenine	3,254	X-ray	Pierce and Loring, in press
Cytidylic acid	H263[a]	X-ray	Loring and Pierce, 1944
	1,298	X-ray	Loring and Pierce, 1944
	45,203R	Ultraviolet	Loring and Pierce, 1944
Nitrite or other reduced nitrogen	14,789	X-ray	Horowitz *et al.*, 1945

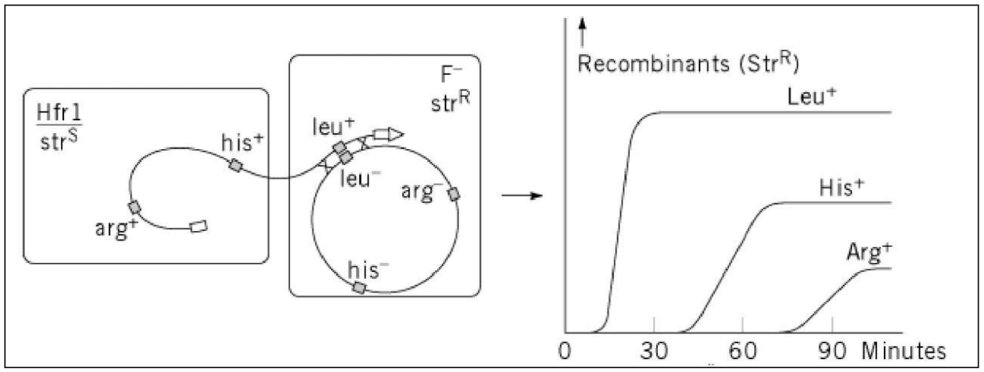
Conjugation Mapping

TABLE 1. TYPES ISOLATED FROM SINGLE AND MIXED CULTURES. MUTANTS USED ARE INDICATED ON THE LETTERED LINES.

	From single and mixed	From mixed only	From single and mixed
A	$B^-M^-P^+T^+$ $B^-M^+P^+T^+$*	and $B^+M^+P^+T^+$ *	$B^+M^+P^+T^-$ $B^+M^+P^-T^+$
B	$B^-M^-P^+T^+R$ **	and $B^+M^+P^+T^+R$ * $B^+M^+P^+T^+$ *	$B^+M^+P^-T^-$ **
C	$B^-M^-P^+T^+$ **	and $B^+M^+P^+T^+R$ * $B^+M^+P^+T^+$ *	$B^+M^+P^-T^-R$ **
D	$B^-M^-P^+T^+R$ **	and $B^+M^+P^+T^+R$ *	$B^+M^+P^-T^-R$ **
E	$B^-\Phi^-C^-P^+T^+$ $B^-\Phi^+C^-P^+T^+$ $B^+\Phi^-C^-P^+T^+$	and $B^+\Phi^+C^+P^+T^+$ * $B^-\Phi^+C^+P^+T^+$ $B^-\Phi^-C^+P^-T^+$	$B^+\Phi^+C^+P^-T^+$ $B^+\Phi^+C^+P^+T^-$

* Prototroph.
** See A for biochemical variations.
The letters refer to requirements for essential metabolites as follows:
B = biotin; M = methionine; Φ = phenylalanine; P = proline; C = cystine; T = threonine
R = Resistance to virus T1.

Lederburg and Tatum, 1946



Conjugation mapping by interrupted mating
Adapted from Wollman, Jacob, Hayes, 1956
(image source what-when-how.com)

TABLE 2

Relative proportions of various nutritional cell types in a mixed culture of $B-\phi-C-T+L+B_1+V_1^s$ and $B+\phi+C+T-L-B_1-V_1^r$

TYPE	NUMBER OF THIS TYPE ISOLATED*	NUMBER OF PROTOTROPHS	RATIO OF THIS TYPE TO PROTOTROPHS	REMARKS
$B-\phi-C-T+L+B_1+V_1^s$	(Parental type. Present in large excess)			
$B+\phi+C+T-L-B_1-V_1^r$	(Parental type. Present in large excess)			
$B+\phi+C+T+L+B_1+$	86		1.00	Prototrophs
$B+\phi+C+T+L+B_1-$	36	37	0.97	Thiamineless
$B+\phi+C+T-L+B_1+$	2	31	0.06	Threonineless
$B+\phi+C+T+L-B_1+$	4	55	0.07	Leucineless
$B-\phi+C+T+L+B_1+$	5	56	0.09	Biotinless
$B+\phi-C+T+L+B_1+$	1	52	0.02	Phenylalanineless
$B+\phi+C-T+L+B_1+$	1	19	0.05	Cystineless
$B+\phi+C+T+L-B_1-$	3	16	0.19	Possible single-reversion type
$B-\phi-C+T+L+B_1+$	2	41	0.05	Possible single-reversion type
$B-\phi+C+T+L+B_1-$†	3	28	0.11	
$B-\phi+C+T-L+B_1+$†	(Isolated in a similar experiment)			
$B-\phi+C+T+L-B_1+$†	(Isolated in a similar experiment)			

* These figures do not include results of tests of virus resistance. Of 49 prototrophs tested, 20 (41%) were resistant. Seven out of 20 thiamineless (35%) were resistant.
† It should be noted that these types represent double-requirement recombination types.

Tatum and Lederburg, 1947

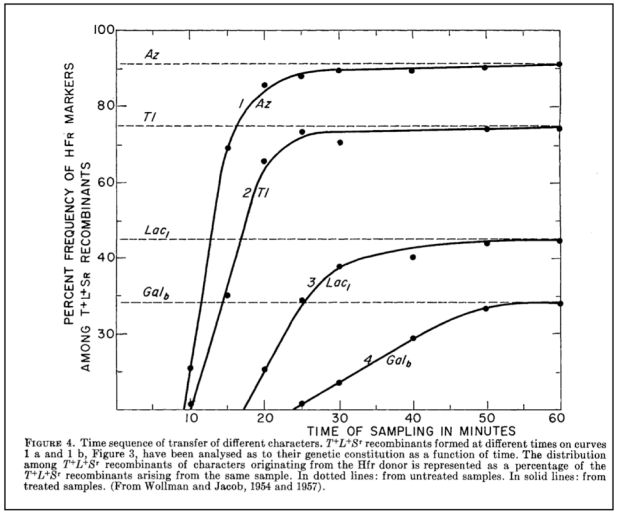


FIGURE 4. Time sequence of transfer of different characters. $T^+L^+S^r$ recombinants formed at different times on curves 1 a and 1 b, Figure 3, have been analysed as to their genetic constitution as a function of time. The distribution among $T^+L^+S^r$ recombinants of characters originating from the Hfr donor is represented as a percentage of the $T^+L^+S^r$ recombinants arising from the same sample. In dotted lines: from untreated samples. In solid lines: from treated samples. (From Wollman and Jacob, 1954 and 1957).

Wollman and Jacob 1954, 1957

How are genes regulated? The lac operon

Two E. coli strains:
- *Str*^{S}, *Hfr* donor cell
- *Str*^{R}, *F*$^{-}$ recipient cell

Structural genes:

lacZ (encodes β-galactosidase)
mutants: Z^{+},Z^{-}

lacY (encodes β-galactoside permease)
mutants: Y^{+},Y^{-}

lacA (encodes β-galactoside transacetylase)
mutants: A^{+},A^{-}

Regulator genes:

lacI (encodes lactose repressor)
mutants: I^{+} ,I^{-} ,I^{-D} , I^{S}

lacO (operator sequence)
mutants: O^{+} , O^{c}

Experiments:

	β-gal activity		
	w/ lactose	w/o lactose	inducibility
$Z^{+}, I^{+},O^{+} \longrightarrow Z^{-}, I^{-},O^{+}$	+	-	ind
$Z^{-}, I^{-},O^{+} \longrightarrow Z^{+}, I^{+},O^{+}$	+	-	ind
$Z^{-}, I^{s},O^{-} \longrightarrow Z^{+}, I^{+},O^{+}$	-	-	unind
$Z^{+}, I^{+},O^{c} \longrightarrow Z^{-}, I^{+},O^{+}$	+	+	const

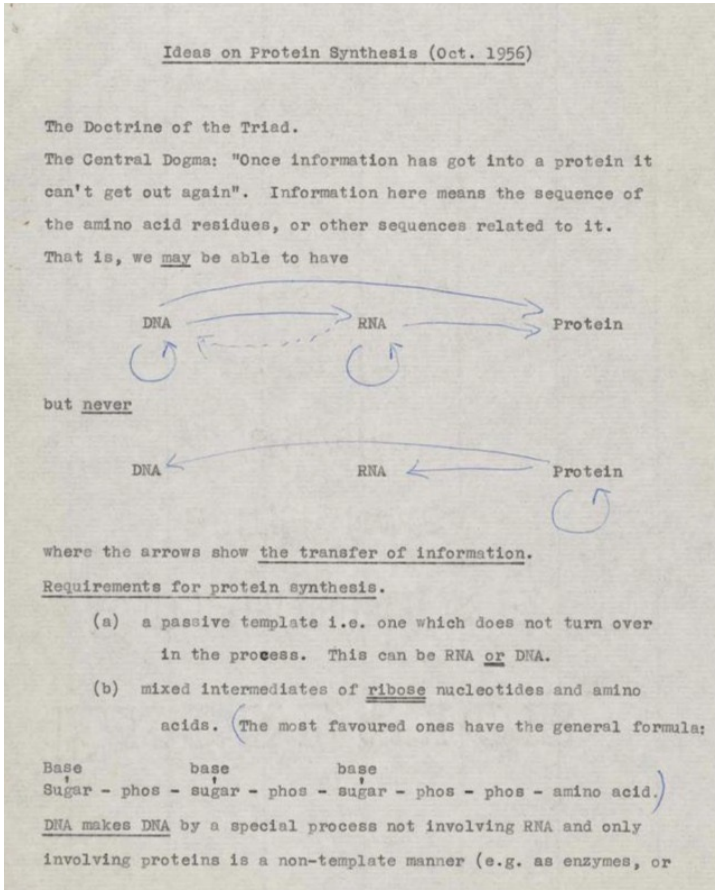
$Z^{+},Y^{-}\ I^{+},O^{c} \longrightarrow Z^{-}, Y^{+},I^{+}O^{+}$

Key findings on the way to coding

- Hershey and Chase 1952
  - infected bacteria with labeled phage
    - labeled sulfur did not enter the cell
    - labeled phosphorus did enter the cell

- Watson/Crick/Franklin/Wilkins 1953
  - Double Helix structure of DNA
    - Chargaff's rules (A = T and G = C)

- Crick develops "central dogma" ~1956



Ideas on Protein Synthesis (Oct. 1956)

The Doctrine of the Triad.

The Central Dogma: "Once information has got into a protein it can't get out again". Information here means the sequence of the amino acid residues, or other sequences related to it. That is, we may be able to have

DNA → RNA → Protein

but never

DNA ← RNA ← Protein

where the arrows show the transfer of information.

Requirements for protein synthesis.

(a) a passive template i.e. one which does not turn over in the process. This can be RNA or DNA.

(b) mixed intermediates of ribose nucleotides and amino acids. (The most favoured ones have the general formula:

Base base base
Sugar - phos - sugar - phos - sugar - phos - phos - amino acid.)

DNA makes DNA by a special process not involving RNA and only involving proteins is a non-template manner (e.g. as enzymes, or

Cell-free translation synthesis system

THE DEPENDENCE OF CELL- FREE PROTEIN SYNTHESIS IN E. COLI UPON NATURALLY OCCURRING OR SYNTHETIC POLYRIBONUCLEOTIDES

By Marshall W. Nirenberg and J. Heinrich Matthaei*

NATIONAL INSTITUTES OF HEALTH, BETHESDA, MARYLAND

Communicated by Joseph E. Smadel, August 3, 1961

What's in the buffer?

- 100 mM Tris pH 7.8
- 10 mM Magnesium acetate
- 50 mM KCl
- 6 mM BME
- 1mM ATP
- 5 mM PEP
- Pyruvate kinase
- 19 L-amino acids (- valine)
- 0.03 each GTP, CTP, UTP
- C^{14} L-valine

- Various extract fractions

Effects of ‘soluble’ and ‘ribosomal’ RNA

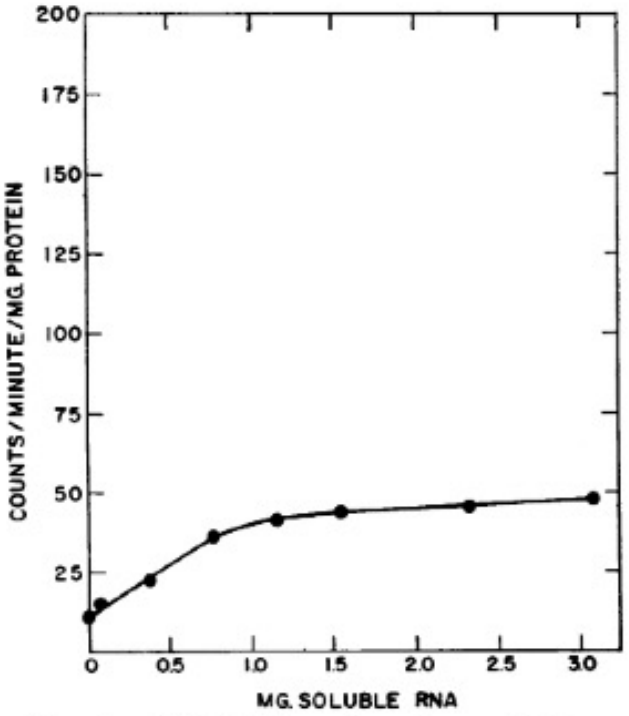


FIG 1.—Stimulation of amino acid incorporation into protein by *E. coli* soluble RNA. Composition of reaction mixtures is specified in Table 1. Samples were incubated at 35° for 20 min. Reaction mixtures contained 4.4 mg. of Incubated-S-30 protein.

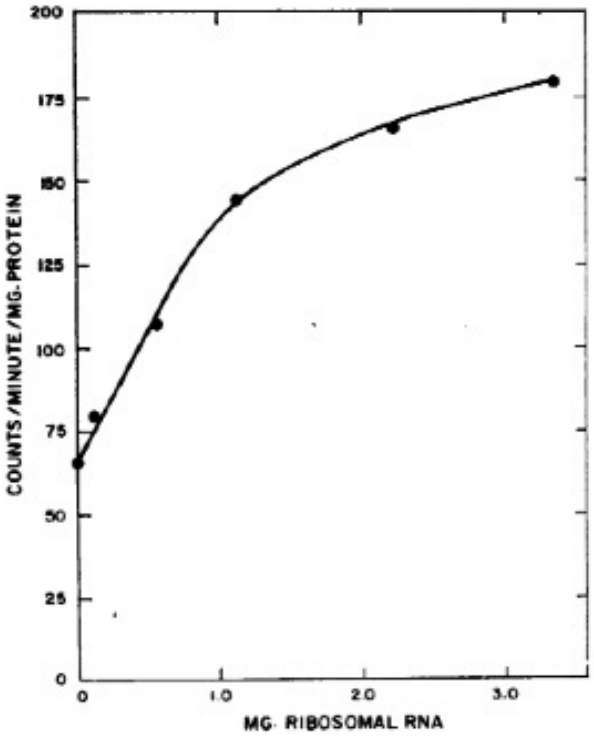


FIG. 2.—Stimulation of amino acid incorporation into protein by *E. coli* ribosomal RNA in the presence of soluble RNA. Composition of reaction mixtures is specified in Table 1. Samples were incubated at 35° for 20 min. Reaction mixtures contained 4.4 mg of Incubated-S-30 protein and 1.0 mg *E. coli* soluble RNA.

CODING!!!
PolyU encodes poly-Phenylalanine

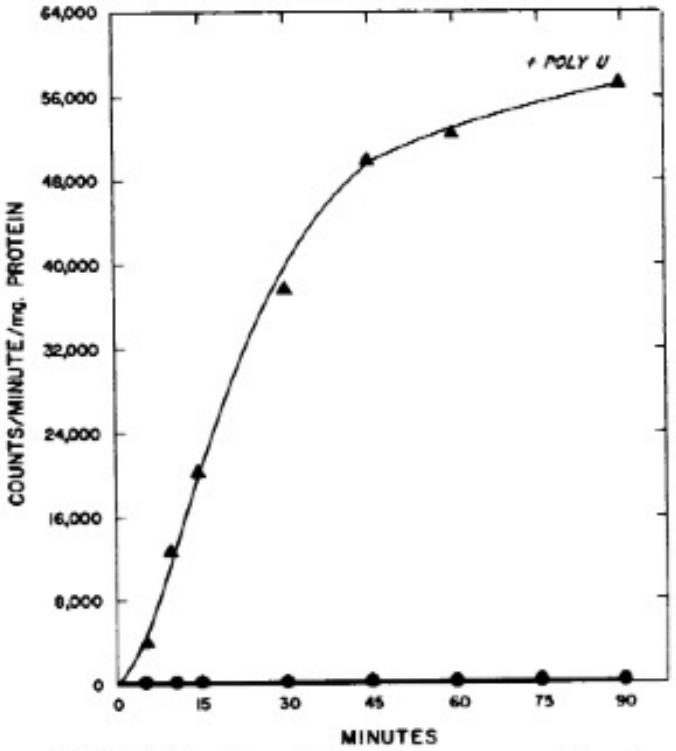


FIG. 6.—Stimulation of U-C^{14}-L-phenylalanine incorporation by polyuridylic acid. ● without polyuridylic acid; ▲ 10 μg polyuridylic acid added. The components of the reaction mixtures and the incubation conditions are given in Table 1. 0.024 μmole U-C^{14}-L-phenylalanine (~500,000 counts/min) and 2.3 mg Incubated-S-30 protein were added/ml of reaction mixture.

PolyU is really specific for incorporation of phenylalanine

TABLE 8

SPECIFICITY OF AMINO ACID INCORPORATION STIMULATED BY POLYURIDYLIC ACID

Experiment no.	C^{14}-amino acids present	Additions	Counts/min/mg protein
1	Phenylalanine	Deproteinized at zero time	25
		None	68
		+ 10 μg polyuridylic acid	38,300
2	Glycine, alanine, serine, aspartic acid, glutamic acid	Deproteinized at zero time	17
		None	20
		+ 10 μg polyuridylic acid	33
3	Leucine, isoleucine, threonine, methionine, arginine, histidine, lysine, tyrosine, tryptophan, proline, valine	Deproteinized at zero time	73
		None	276
		+ 10 μg polyuridylic acid	899
4	S^{35}-cysteine	Deproteinized at zero time	6
		None	95
		+ 10 μg polyuridylic acid	113

Components of the reaction mixtures are presented in Table 1. The unlabeled amino acid mixture was omitted. 0.015 μM of each labeled amino acid was used. The specific activities of the labeled amino acids are present in the *Methods and Materials* section. 2.3 mg of protein of preincubated S-30 enzyme fraction were added to each reaction mixture. All samples were incubated at 35° for 30 min.

Deciphering the Triplet Code

- Benzer (1955) Fine structure of a gene
  - rII mutants of T4 phage

- Crick, Brenner et al., 1961 lay out evidence for a non-overlapping triplet

- Khorana works out how to chemically synthesize oligonucleotides

- Nirenberg and Leder
  - Used short synthetic triplet RNAs
  - By a ribosome binding assay, they could determine which aa-charged tRNAs associated with each triplet