

BRK-013 · Deciphering the genetic code: 1954 to 1966

BRK · Conceptual Breaks

Juan Segurola · All Rights Reserved

Knowing that a sequence can carry hereditary information does not tell us how that sequence specifies a protein. Four kinds of nucleotide and twenty common amino acids pose a mapping problem. Counting possible arrangements can constrain an answer, but it cannot determine which arrangement is assigned to which amino acid or how a cell performs the assignment.

Between 1954 and 1966, that problem became experimentally tractable. Proposals about the code's structure were confronted with mutations, cell-free synthesis, short RNA sequences and chemically prepared templates. The change was from plausible patterns to experimentally discriminated assignments, with several methods supplying different parts of the answer.

1. A combinatorial possibility is not yet a code

Gamow's 1954 proposal linked DNA structure to possible arrangements specifying amino acids. Its historical importance lies partly in making the coding question explicit. It should not be presented as the experimentally established mechanism of translation. A geometrical scheme can be ingenious and still fail as a description of what cells do. [1]

The elementary counting argument is useful but limited. Four symbols give four one-symbol units, sixteen ordered pairs and sixty-four ordered triplets. If twenty distinct amino-acid assignments must be represented by fixed-length units, pairs are insuffi-

cient and triplets are sufficient in number. That arithmetic does not prove that nature uses triplets.

It also does not decide whether the units overlap, whether there are separators, where reading begins or whether different units can specify the same amino acid. Those are additional structural questions. A code is not deciphered by noticing that one numerical capacity exceeds another.

2. Mutations test the organization of reading

Crick, Barnett, Brenner and Watts-Tobin's 1961 work used mutations and combinations of mutations to infer general properties of the code. The recovery of function with certain combinations of additions and deletions supported reading in successive groups of three. This was genetic evidence about organization, not a complete dictionary of individual triplets. [2]

An explanatory sequence makes the logic visible. Suppose a reader groups symbols as ABC DEF GHI JKL. Insert one symbol near the beginning and every later grouping changes. Remove a symbol farther along, and the original grouping can be restored beyond that point, although the intervening segment remains altered. Three added symbols can also preserve the later grouping while changing the local content.

This is an illustration of reading-frame logic, not one of the actual sequences or phenotypic measurements from the experiment. Real interpretation required attention to which mutant combinations restored function and what alternative mechanisms could explain the pattern. Recovery of function is an observable consequence from which sequence organization is inferred.

3. A synthetic template gives a chemical assignment

Nirenberg and Matthaei's 1961 experiments supplied synthetic polyribonucleotides to a cell-free protein-synthesizing system. Polyuridylic acid, containing repeated uracil nucleotides, directed production of a product rich in phenylalanine. Together with the emerging triplet interpretation, this provided an assignment of UUU to phenylalanine. [3]

The cell-free preparation mattered because it separated a specific biochemical operation from the full complexity of a living cell. Investigators could change the added RNA and measure incorporation into a product. A template became an experimental input rather than only a hypothetical message inferred from inheritance.

But one repeated symbol leaves many questions untouched. It does not establish all assignments, distinguish every possible starting position or explain the whole apparatus. The result is powerful because its scope is specific: under the tested conditions, this synthetic RNA changed which amino acid was incorporated into a polypeptide.

4. More informative templates require more informative chemistry

Mixed synthetic polymers could constrain assignments through their composition, but a known mixture of bases is not the same as a known order of bases. Different triplets can contain the same counts of each nucleotide. A statistical association between composition and amino-acid incorporation may therefore leave ordering ambiguities.

Khorana's program used polynucleotides with defined repeating sequences to make the input more informative. His 1968 Nobel lecture describes this work retrospectively. The relevant advance

was control over sequence patterns, allowing different candidate assignments to be separated by what the templates produced. [4]

For example, an indefinitely repeated two-symbol pattern can be read in alternating triplets. A repeated three-symbol pattern permits three possible starting frames. These constructed cases show why the outputs must be interpreted together with the template's periodicity and the possible initiation positions. A chemically orderly input can still generate more than one candidate reading.

Short-triplet binding assays supplied another kind of constraint: a triplet could be tested for its effect on the association of an aminoacyl-transfer RNA with ribosomes. Nirenberg's Nobel lecture describes how these assays complemented protein-synthesis experiments. Binding and completed synthesis are different observables, so agreement between them is more informative than merely repeating one measurement. [5]

5. Convergence without one decisive experiment

By the mid-1960s, the assignment problem was being resolved through the combination of these methods. Genetic alterations constrained how the message was grouped; synthetic templates connected RNA patterns to products; binding experiments tested individual triplet associations. None alone supplied every part of the mechanism. [4,5]

This division of experimental labor also explains why the history should not be organized as a race won by one person. Preparing materials, maintaining active biochemical systems, designing sequence patterns and interpreting mutant combinations were different contributions to a common inferential problem. The evidence chain is more informative than a list of awards.

The outcome did not show that sequence operates independently of machinery. Transfer RNAs, enzymes, ribosomes and initiation and termination processes make the mapping physically effective. A printed table is a description of regular assignments, not a causal agent that translates itself.

6. What deciphering made possible

The conceptual break was the replacement of an unrestricted coding speculation by experimentally constrained relations between sequence and synthesis. A proposed assignment could now be challenged by changing a template, examining a mutation or comparing distinct assays. The code became something one could investigate one relation at a time.

That achievement also created new questions. How are assignments implemented with sufficient fidelity? How are starts and stops recognized? Why does the mapping have the organization it does? Establishing a dictionary did not answer all those questions, but it changed them from demands on an unknown mapping into investigations of a mapping with experimentally established structure.

Related reading

BRK Article 015, “Immunity becomes selection among biological variants,” examines another shift in which a proposed biological relation acquired tests through the organization of cells and molecules.

References

- [1] Gamow, G. “Possible Relation between Deoxyribonucleic Acid and Protein Structures.” *Nature* 173, 318 (1954). <https://doi.org/10.1038/173318a0>

- [2] Crick, F. H. C., Barnett, L., Brenner, S., and Watts-Tobin, R. J. "General Nature of the Genetic Code for Proteins." *Nature* 192, 1227–1232 (1961). <https://doi.org/10.1038/1921227a0>
- [3] Nirenberg, M. W., and Matthaei, J. H. "The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides." *Proceedings of the National Academy of Sciences* 47, 1588–1602 (1961). <https://doi.org/10.1073/pnas.47.10.1588>
- [4] Khorana, H. G. "Nucleic Acid Synthesis in the Study of the Genetic Code." Nobel Lecture (1968). <https://www.nobelprize.org/prizes/medicine/1968/khorana/lecture/>
- [5] Nirenberg, M. W. "The Genetic Code." Nobel Lecture (1968). <https://www.nobelprize.org/prizes/medicine/1968/nirenberg/lecture/>