Information That Records and Alters Its Own Conditions

a mechanistic account of the origin of the genetic code

A mechanistic and computational account of the origin of biological information

Dustin Lane

Independent Researcher

lanedustin91@gmail.com

Preprint · 27 June 2026

Information That Records and Alters Its Own Conditions: a mechanistic account of the origin of the genetic code

A mechanistic and computational account of the origin of biological information

Dustin Lane · Independent Researcher · lanedustin91@gmail.com

Mechanism specification, grounded evidence, and a reproducible computational model

Abstract. This document argues that the genetic code is not a static cipher — a fixed lookup table mapping symbols to meanings — but a *condition-instructed writing* process: its symbol→meaning mapping is set by physical conditions, and the act of writing a mark is a metabolic conversion that consumes the condition and records it, so every mark is at once a memory of a past condition and a change to the system that made it. Underneath the account runs a structural corollary: the genetic material and the cell's core metabolic currency are the same nucleotides — the pool that records heredity is the pool that runs metabolism — so heredity and metabolism are one chemical system, within which the RNA→DNA conversion is a flux branch point. On this view the mapping has two physically instantiated layers — condition→symbol and symbol→referent, realised in condition-responsive RNA and implemented by molecules rather than by a living cell; the record it writes is heritable and proofread; and the mapping is not assigned but emerges from the chemistry, contingent on the molecular inventory rather than fixed in advance. The argument (1) specifies the mechanism, (2) grounds every step in measured biology, and (3) provides a reproducible model of the contingency such a code must exhibit — stating honestly that the model bounds the mapping layer's capacity, while the memory that defines the system is realised in the written sequence itself and measured in Section 3.

Keywords: origin of the genetic code; RNA world; nucleotide metabolism; ribozymes; condition-dependent gene regulation; molecular evolution

1. The question, restated

The origin-of-information problem asks how a coded correspondence between two physically distinct chemical domains — nucleic acids and the molecules they specify — could arise without a designer assigning it. It is natural to picture the genetic code as a *cipher*: a fixed table mapping symbols to meanings, of the kind engineers build. Posed that way the problem is genuinely hard, because an arbitrary table has astronomically many alternatives and no physical reason to select one; a mapping that carries information-theoretic capacity but no symbol→referent correspondence, or one that a single physical curve forces rather than freely assigns, does not by itself resolve it. We argue that the genetic code is not that kind of object, and that picturing it so is part of what makes its origin look intractable.

We make three claims, each meant to be checked against chemistry and data.

Claim 1 (the object). The genetic code is not a cipher of the engineered kind. Its mapping is *set by* the physico-chemical conditions, and the act of marking is a metabolic event that changes the system — so the code is a condition-coupled, self-modifying, memory-writing process, not a static lookup. And the memory it writes is not passive: because the same condition that drives a write also sets what is written, the record biases the cell toward the recurrence of that condition — so the system writes, into heritable sequence, an improved capacity to respond to its environment (evidenced at the sequence level in §3.5). The feature that makes it look hard to originate (apparent arbitrariness of a frozen table) is an artifact of modelling it as the wrong kind of object.

Claim 2 (the mechanism). That object is reachable from prebiotic chemistry, because condition-directed catalysis — environment setting structure, structure setting which reaction happens — was present in the RNA world, and the very same metal-and-metabolite-dependent chemistry runs the extant code. The unity is concrete at the level of the molecules: the nucleotides are at once the genetic alphabet and the backbone of the cell's energy, redox and group-transfer currency, so the pool that records heredity also runs metabolism, and the reductase step that converts RNA precursors to DNA is a branch point in that shared flux (Lundin 2015). Chemistry and biology are one continuous system on this axis, not two.

Claim 3 (the quantification). The contingency that distinguishes an emergent code from an authored one is decidable and measurable by degeneracy, which we compute. We are explicit that this computation captures only the mapping layer; the memory layer — the part that makes the code biological — is specified mechanistically and grounded in data rather than simulated.

Formally. The thesis compresses to a single statement: `written_sequence = f(conditions, machinery and its fidelity, template binding-context)`, and the biological *function* of what is written is a readout of that same writing context — so the genotype–phenotype map is itself condition-mediated. A technological code is the degenerate special case in which the conditions and the context drop out and only a fixed table remains; the genetic code is the general case, in which they do not.

2. Two kinds of code

This contrast between two kinds of code organises the mechanism developed in §3.

Property	Technological code (ASCII, a cipher)	Chemical–biological code (the genetic code)
The table	Exogenous: a convention imposed from outside the physics.	Endogenous: implemented *by* the chemistry; there is no table separate from the molecules and their conditions.
Response to environment	Invariant: temperature, pH, ionic state do not change what a symbol means.	Condition-set: the mapping shifts with conditions, because the enzymes that write/erase are metabolite sensors (§3.2).
Effect of writing	Memoryless / open-loop: writing or reading leaves the machinery and the table untouched; same input, same output, forever.	Memory-writing / closed-loop: writing consumes the condition and deposits a persistent, inherited record; the past condition is stored and the system is changed by storing it (§3.3).
Role of the medium	Substrate-independent: the same code runs on silicon, paper, tape.	Substrate-dependent: the medium (sequence, its adopted structure, the metabolite pools) *is* part of the meaning (§3.4).

A technological code — filled, fixed encode/decode tables and a static mapping — is the left column. The object an origin account must explain is the right column — and that object is better drawn as a branching *constraint graph* (a flow chart) than a table: its outcomes are *enforced* by conditions satisfied in sequence — coincidences met, thresholds crossed — not stipulated row by row, so the filled table is a projection read off the graph rather than the mechanism itself (developed in §4). The remainder of this document develops that object, mechanism first and evidence throughout.

3. The system

All biological claims below are paraphrased from peer-reviewed sources retrieved and verified via PubMed; DOIs are in Section 8. Labels: **[Established]** = drawn from the peer-reviewed literature; **[Interpretation]** = our reading built on it; **[Established + Interpretation]** = an established basis carried into an interpretive synthesis.

3.1 The foundational writing — condition-dependent addition of nucleotides (di-, tri-, and higher-order)

The system has an origin, and the most basic writing event — the one that builds the genetic polymer itself, more foundational than the genetic code or any regulatory mark placed upon it — is the addition of one nucleotide to the next. Whether a bond forms, and how far the chain grows from a dimer to a trimer to a higher-order oligomer, is set by the conditions, not by any imposed table. We anchor the account on the one progression that is not in dispute: **prebiotic → RNA → DNA**.

The monomers form under prebiotic conditions. **[Established]** On the anoxic early Earth, geochemistry already runs a metabolism: native Fe, Co and Ni metals catalyse H_2 + CO_2 to pyruvate and acetyl-CoA, replacing the function of over 120 enzymes with no RNA, nitrogen or light required (Mrnjavac 2024), and at alkaline hydrothermal vents pH gradients across Fe(Ni)S barriers drive CO_2 fixation to α-ketoglutarate (Camprubi/Lane 2017). The monomers of all three macromolecule classes then form together from a single cyanosulfidic feedstock (Patel/Sutherland 2015). The building blocks are available; the question is how a *sequence* gets written from them.

The writing is condition-dependent nucleotide addition. **[Established + Interpretation]** Activated ribonucleotides add to one another on mineral surfaces — montmorillonite clay, and the iron–sulfur chimneys of alkaline vents — and the elemental composition, pH, the mineral, and the activation chemistry decide the outcome: unmodified monomers reach only dimers, while imidazole-activated monomers extend to higher oligomers (Burcar 2015), with montmorillonite-catalysed synthesis tuned over pH and salinity (Joshi 2015). The same chemistry that joins di- to tri- to higher-order oligomers is governed throughout by the conditions, not by a stored rule.

Add a template, and the same write becomes a copy — the prebiotic special case of the master relation. **[Established + Interpretation]** The untemplated addition above gains a template, and the writing now copies it: the master relation `written_sequence = f(conditions, machinery and its fidelity, template binding-context)` appears here in its simplest form — the machinery term at its prebiotic minimum (no enzyme), the template supplying the binding-context — so what is written is set by template and conditions alone. An activated monomer first binds the template, then forms a 5′–5′ imidazolium-bridged *dinucleotide* intermediate, and template-directed primer extension proceeds with the catalytic metal ion Mg^{2+} deciding the outcome — over 90% of the template-bound intermediate is extended rather than hydrolysed, an enzyme-like specificity supplied by the template itself (Walton/Szostak 2019). This template-directed copying is the prebiotic form of replication, by which RNA is propagated as hereditary information; which sequences are accessible, and at what fidelity, is set by the reaction conditions (Duzdevich/Szostak 2020). At its limit this copying writes not merely a longer strand but a functional one: an evolved RNA polymerase ribozyme synthesises an enzymatically active hammerhead ribozyme from an RNA template — RNA writing working RNA (Wochner/Holliger 2011). The template says which base; the conditions say whether, and how faithfully, it is written.

The chemistry of addition stays condition-dependent when proteins take over. **[Established]** Protein polymerases do not change the chemistry in kind: two divalent metal ions catalyse each nucleotidyl-transfer step (Yang 2006), and the identity of the metal sets the fidelity of the addition (Sirover & Loeb 1977). The

mechanism by which a nucleotide is written into the chain is metal-dependent — condition-dependent — from mineral-surface oligomerization to the extant replisome.

"Higher-order" is also folded structure, and it too is condition-set. [Established + Interpretation] As sequence accumulates it adopts higher-order structures whose formation depends on the conditions — G-quadruplexes stabilised by K^+ (Monsen 2022) and i-motifs by acidic pH (Zeraati 2018) — and these structures gate access to the template and bias what is written next. The polymer's own condition-set folding becomes part of its writing context.

The machinery that runs the addition is primordial. [Established + Interpretation] On the Trifonov consensus chronology the nine Miller-experiment amino acids (Gly, Ala, Asp, Val, Pro, Ser, Glu, Leu, Thr) are recruited earliest (Trifonov 2000; Trifonov 2004); our prior computations find the conserved catalytic cores built from those earliest amino acids (the two-aspartate-plus-metal centre of the polymerases), with eight of ten writer catalytic engines at LUCA grade — present in bacteria, archaea and eukaryotes (exemplified by MutS across all domains, Eisen 1998, and the 2-oxoglutarate/Fe(II) oxygenase fold conserved prokaryote→human, Chowdhury 2014) — for these oxygenases it is the *fold* that is this old; their O_2-dependent catalysis is a later overlay, not a claim that oxygen-sensing operated at origin. The two-metal addition chemistry, and the engines that run it, are as old as cellular life. This verdict rests on two independent instruments — residue-level chronology and cross-superkingdom domain phylogeny — agreeing that the writing chemistry is ancient and the specificity derived; bulk amino-acid composition, by contrast, shows nothing, because modern ortholog gain has erased the ancestral signal, leaving only the conserved catalytic core to carry it. (Honest caveat: universal presence is consistent with deep ancestry but does not by itself exclude ancient horizontal transfer.)

prebiotic→RNA→DNA. [Established + Interpretation] Metal identity already sets what an RNA can do — Fe(II), abundant in the anoxic ocean, confers a catalysis (single-electron transfer) that Mg(II) does not (Hsiao 2013) — and DNA comes last, written as a covalent *extension* of RNA: ribonucleotide reductase makes DNA's building blocks from RNA's (2′-OH → H; Nordlund & Reichard 2006), nascent DNA extending from an RNA primer. DNA is the chemically stabilised younger copy, not a separate code. This conversion is a branch point in metabolic flux: the ribonucleotides consumed are the same pool that builds RNA and carries the cell's energy and redox currency, and ribonucleotide reductase is the single de-novo gate out of it — a chemically demanding radical reaction judged unlikely to have been a ribozyme, hence a protein innovation at the RNA→DNA boundary (Lundin 2015) — and it runs one way, to a U-DNA world after the reductase and then a T-DNA world after thymidylate synthase (Forterre 2005).

The anchor in one view:

Stage	The writing event	Grounding (refs in §8)
Prebiotic (mineral floor)	Monomers form; activated nucleotides add on mineral surfaces — di-, tri-, higher-order — with length and yield set by pH, metals, mineral and activation.	Native Fe/Co/Ni metabolism (Mrnjavac 2024); vent CO_2 fixation (Camprubi/Lane 2017); one cyanosulfidic feedstock (Patel/Sutherland 2015); oligomerization on minerals (Burcar 2015; Joshi 2015).
RNA world	Template-directed addition copies a sequence: a dinucleotide intermediate, Mg^{2+}-gated, template-specified — up to a working ribozyme copied from an RNA template; RNA propagated as hereditary information.	Template-directed primer extension (Walton/Szostak 2019); copying fidelity set by conditions (Duzdevich/Szostak 2020); RNA writes active RNA (Wochner/Holliger 2011); Fe(II) sets RNA catalysis (Hsiao 2013).
RNA→DNA seam	DNA written as a covalent extension of RNA; its building blocks made from RNA's.	RNR converts RNA→DNA building blocks (Nordlund & Reichard 2006).

Extant	Protein polymerases add by two-metal catalysis with metal-set fidelity; condition-set higher-order structure gates further writing.	Two-metal nucleotidyl transfer (Yang 2006); metal-set fidelity (Sirover & Loeb 1977); G4/i-motif (Monsen 2022; Zeraati 2018).

The through-line: condition-dependent writing at every stage, no moment of assembly. [Interpretation] Read down the anchor, the through-line is exact: writing is the condition-dependent addition of nucleotides at every stage — from mineral-surface oligomerization, through template-directed copying, across the seam to DNA, into the polymerase. One naturally looks for the moment a static cipher self-assembles; on this progression there is no such moment, because writing was a condition-dependent chemical addition from the first dinucleotide.

Why the medium went double-stranded — an information upgrade, not only durability. [Established + Interpretation] Calling DNA the stabilised copy understates the second strand. The textbook reason for the duplex is fidelity and template-directed repair; there is a second reason, about *writing*. At a single G/C-rich address the two antiparallel strands can present *two different* structures — a G-quadruplex on the G-strand and an i-motif on the C-strand — and these are not one signal read twice but an antagonistic pair: stabilising the i-motif versus the G-quadruplex at the same proto-oncogene promoter drives expression in opposite directions (Saha 2020), and placing an oxidised guanine on the coding versus the template strand of the same element flips transcription up versus down (Fleming 2017). The two structures answer *different* conditions — the G-quadruplex licensed by K^+ and guanine oxidation (Siebenmorgen & Zacharias 2017; Fleming 2017), the i-motif by pH (Zeraati 2018) — so one locus reads two environmental variables on its two strands and can output either way depending on which the conditions favour: a two-input switch made of the duplex itself. A single strand cannot do this. RNA folds the G-quadruplex readily but disfavours the i-motif — the RNA i-motif is roughly half as stable per $C{\cdot}C^+$ pair (Snoussi 2001) — and, being single-stranded, it has no complement on which to present the second channel at all. [Interpretation] So the complementary strand is two things at once: a second, orthogonal condition-channel that single-stranded RNA lacks, and — through semi-conservative replication — the engine that partitions the two strands, and with them the two channels, into two daughters (§3.3). That is a functional driver for the RNA → DNA and single → double transitions with nothing to do with chemical stability: the second strand upgrades the environmental-encoding system from one channel in one cell to two orthogonal channels distributable between two cells.

3.2 The same principle, now an epigenetic write — the mapping *is* metabolism

A distinct, downstream layer — epigenetic writing — obeys the same conditional grammar without touching the genetic text. [Established] The writing of §3.1 is *genetic*: it builds and copies the base sequence itself. A separate layer writes reversible marks *onto* the finished polymer — methyl and acetyl groups on histones and on DNA bases — without altering that sequence: a 5-methylcytosine still pairs and is read as a cytosine. This is the *epigenetic* layer, and we keep it firmly distinct from genetic writing; the only parallel we draw is that its mapping is set the same way — by conditions, now the cell's metabolite pools, with no separable table. The enzymes that deposit and remove these marks cannot fire without specific metabolites: DNA and histone methyltransferases write from S-adenosylmethionine; the demethylating dioxygenases (the TET family and the JmjC histone demethylases) require 2-oxoglutarate, molecular O_2 and Fe(II) — the same 2OG/Fe(II) chemistry as the ancient oxygenases of §3.1; histone acetyltransferases require acetyl-CoA, the same acetyl-CoA the prebiotic metal chemistry made. Three measured anchors make the dependence concrete:

- Acetyl-CoA supply gates the write: histone acetyltransferases consume acetyl-CoA, and its availability tunes the mark — shown for the glucose → citrate → ATP-citrate-lyase route, where ACLY-derived acetyl-CoA is required for the acetylation increases seen during differentiation and glucose availability itself tunes acetylation (Wellen 2009). That route is one major source of the acetyl-CoA pool, not the only one. [Established]

- Oxygen is read directly by the eraser to control fate: the H3K27 demethylase KDM6A senses O_2 directly (its paralog KDM6B does not), and KDM6A loss — like hypoxia — blocks H3K27 demethylation and blocks cellular differentiation; restoring methylation homeostasis reverses the block, so "oxygen directly affects chromatin regulators to control cell fate" (Chakraborty 2019). **[Established]**
- A metabolite can jam the erasers genome-wide: the oncometabolite 2-hydroxyglutarate is a competitive inhibitor of the α-ketoglutarate-dependent dioxygenases, including TET and the histone demethylases, so an IDH mutation shifts methylation across the genome (Xu 2011). And the writer's deployment is itself condition-gated by a metabolic kinase: AMPK phosphorylates TET2 to stabilize it, high glucose impedes that phosphorylation, and the anti-diabetic drug metformin restores it (Wu 2018). **[Established]**

The marks are the code's addressing logic, not a modulatory layer. **[Interpretation]** These are not a "modulatory layer" on a fixed code; they are the code's addressing logic — the same conditional grammar as §3.1, applied now to marks rather than to the sequence. Change the conditions and you change which mark is written, the precise property a technological code is engineered to lack. And the addressing is not one condition to one mark: a single metabolic state is read at once by many of these enzymes — the TET dioxygenases, the JmjC histone demethylases, and the RNA N^6-methyladenosine erasers FTO and ALKBH5 (Jia 2011; Zheng 2013) all sharing one 2OG/O_2/Fe(II) chemistry, and the guanine-oxidation repair arm coupled to the same O_2/ redox state — each with a different substrate affinity, so that KDM6A reads oxygen at a threshold its paralog KDM6B does not (Chakraborty 2019). One condition therefore writes not a single symbol but a graded pattern across many marks, fixed by the relative thresholds of the readers — the positive reason the mapping cannot be reduced to a separable lookup table, only to the kinetics of the enzymes that read it.

One eraser chemistry spans all three layers, and it reaches past the marks to the translation program. **[Established + Interpretation]** The erasers just named are not three mechanisms but one: FTO and ALKBH5 are members of the AlkB family of Fe(II)/2-oxoglutarate dioxygenases (Fedeles 2015), the same Fe(II)/2-oxoglutarate superfamily and the same ferryl Fe(IV)=O chemistry — consuming 2-oxoglutarate to succinate and CO_2, requiring O_2 (Krebs 2007) — that erases 5-methylcytosine on DNA (the TET enzymes) and methyl-lysine on histones (the JmjC demethylases). Across DNA, histone and mRNA the economy is therefore single: written from S-adenosylmethionine, erased by the 2-oxoglutarate clock. The unity is provable by a single lesion to it — the oncometabolite R-2-hydroxyglutarate jams all three layers at once, and on the mRNA layer its inhibition of FTO raises global N^6-methyladenosine, destabilises the *MYC* and *CEBPA* growth transcripts, and arrests the cell cycle (Su 2018) — the same metabolite that retains the chromatin marks above (Xu 2011). And the clock does more than choose which mark is written; it sets how the cell translates what it has written. Under stress the 5′ untranslated regions of newly made transcripts retain N^6-methyladenosine — the FTO eraser held off — and a single retained mark there licenses cap-independent translation initiation, selecting which messages become protein when the 7-methylguanylate-cap route is down (Zhou/Qian 2015). So one metabolic state, read through one dioxygenase chemistry, fixes the mark on all three molecular layers and the translation mode that reads them: the condition → write line is carried to its endpoint, condition → which-protein-is-made — the self-altering half of the thesis extended from the genome onto the proteome.

Why metabolism and the code are one system, not two wired together **[Established + Interpretation]****.** The coupling is physical, and it runs through the nucleotide — as it has since the primordial writers were first built on cofactor-binding folds (§3.1). The deepest form of the coupling is that the genetic monomers and the metabolic currency are one pool: the same ribonucleotides that polymerise into RNA are the cell's energy and redox tokens — ATP is at once an RNA monomer and the universal energy currency, and the redox cofactor NAD is found as a covalent 5′ cap on a subset of bacterial RNAs (Cahová 2015) — so assembling the polymer spends the metabolic pool and degrading it returns it, and RNA synthesis and turnover are themselves metabolic flux. The substrates consumed to write are nucleotide cofactors: acetyl-CoA donates the acetyl

group, S-adenosylmethionine the methyl, NAD^+ the ADP-ribose (PARP), FAD the redox equivalents for demethylation. Each is built on an adenosine core — the adenosyl group of S-adenosylmethionine, the AMP halves of NAD^+ and FAD, the 3′-phospho-ADP of coenzyme A — so the reagent that licenses every mark carries a nucleotide at its centre. S-adenosylmethionine makes the point sharpest: it is an adenosyl group transferred from ATP onto the amino acid methionine, so the universal methyl donor is itself a nucleotide–amino-acid conjugate — the two ancestral alphabets, nucleotide and amino acid, fused in the very molecule that writes the methyl mark. The same nucleotide chemistry is both the cell's metabolic currency and the alphabet that reads, writes, and marks the genome, so there is no wire to cut. And that nucleotide core is a recognition handle, not the reactive group — the methyl rides on the sulfonium, the acetyl on the thioester, the hydride on nicotinamide, the redox on flavin — so the adenosine does no chemistry and serves only as a grip. Nor is this confined to the marking cofactors: the energy (ATP, GTP), redox (NAD^+, FAD) and group-transfer (coenzyme A, nucleotide-sugars) currencies are all nucleotide-based, so the genome's reagents come from the same nucleotide economy of which the genome is the most durable ledger. That a catalytically inert nucleotide sits at the centre of every one of these is the signature of a maker that bound nucleotides: the cofactors are most parsimoniously read as molecular fossils of an early metabolism run by nucleic-acid enzymes before coded protein synthesis, in which the cofactor's nucleotide moiety was the part the ribozyme recognised (White 1976). And the prediction this fossil reading makes is observed directly in the extant polymerase: the redox cofactors NAD^+, NADH and desphospho-coenzyme A are written into RNA not by a dedicated capping enzyme but as non-canonical initiating nucleotides by RNA polymerase itself — in both bacteria and eukaryotic RNAP II, with promoter sequence at and upstream of the start site setting the efficiency — so the cofactor's nucleotide moiety is still incorporated into the chain as a nucleotide, exactly as a nucleotide-binding maker would have placed it (Bird 2016); and the cap so written is itself a condition-set fate symbol, for unlike the 7-methylguanylate cap it does not support translation but routes the transcript to decay (Jiao 2017). On this reading the coupling between metabolism and the genome is not only physical but ancestral — the frozen remainder of the same RNA world §3.1 reconstructs for the writing chemistry itself.

3.3 The template is the memory — writing consumes the condition and leaves a record

The same templated addition that writes the polymer (§3.1) is also what gives it a memory, because the writing is a copy.

(i) The template is the memory. [Established + Interpretation] In template-directed addition the existing strand carries the information and the new strand is written against it; this copying is precisely how RNA is propagated as hereditary information in the prebiotic replication model (Duzdevich/Szostak 2020). The memory is not a separate mechanism added later — it is intrinsic to templated nucleotide addition: read the template, write the complement. The parental strand is the stored record; the new strand is the re-writable copy. And the copy must be faithful to be a memory: in an RNA-only replicating system a low-fidelity polymerase ribozyme diverges toward random sequence and loses function within a few generations, while a high-fidelity variant maintains the heritable information and even evolves improved variants (Papastavrou/Joyce 2024) — fidelity is what makes the template a memory and not noise.

(ii) The write spends the condition. [Established] Each addition consumes an activated monomer, and in the later marking layer each deposition consumes its metabolite — the dioxygenase erasers consume 2-oxoglutarate and release succinate, and the accumulating α-ketoglutarate antagonist 2-HG competitively inhibits those same dioxygenases (Xu 2011). The pool that licensed the write is depleted by it. [Interpretation] The code therefore feeds back on the very conditions it reads — a closed loop, unlike the open loop of a decoder whose dictionary is untouched by use.

And the spending regenerates the substrate, not only the pool. [Established + Interpretation] The erasers do not merely remove a mark; they manufacture the blank for the next write. Active demethylation runs through

TET oxidation of 5-methylcytosine to 5-carboxylcytosine and its excision by thymine-DNA glycosylase, which leaves an abasic site — a position with no base of its own, and therefore no Watson–Crick instruction for what to restore (He 2011). What is written there is a separate, conditional decision, not a readout of the erased mark: filled faithfully against the intact complementary strand, or — when the blank is met without a template — written by the route, as REV1 inserts deoxycytidine opposite an abasic site, a base chosen by the enzyme rather than dictated by a template (Nelson 1996). So the same conditional machinery that reads a mark also produces a non-instructive blank and then writes it — the write→erase→rewrite loop named in the title, closed within one chemistry, and the maintenance-level instance of the same "the machinery is the template" principle that §3.5 shows in the reverse transcriptases.

(iii) The extant maintenance system is a late elaboration of the same template logic. [Established + Interpretation] The strand-age principle of templated copying reappears, much later, in the maintenance of the methylation mark: after replication the mark is hemimethylated (on the parental strand only), UHRF1 reads that one-strand state through its SRA domain — flipping the 5-methylcytosine out of the duplex — and recruits DNMT1 to re-write the mark onto the daughter strand (Sharif 2007; Avvakumov 2008; Bostick 2007). It is the same parental-template-as-memory, new-strand-as-copy logic of §3.1, now applied to a methyl mark rather than a base. And the record is heritable across divisions: multigenerational single-cell tracking resolves the inheritance of endogenous DNA lesions and the emergence of sister-cell asymmetry across up to four generations (Panagopoulos 2025). And because the medium is double-stranded, that record is not merely stored but *partitionable*: semi-conservative replication hands the two strands — and with them the two condition-channels of §3.1 — to the two daughters, so one condition-instructed write can be inherited asymmetrically across the same multi-division interval these measurements resolve.

Storing the condition changes the system — the record is response-improving. [Interpretation] The past condition becomes a stored memory — the template — and storing it has changed the system: a strand written, a pool spent, the next read biased. That bias is the point: a mark written under a condition can pre-position the cell for its return, so the record is response-improving rather than a mere log. The property is demonstrable and ancient: in the bacterial adaptive response to alkylation, the Ada protein repairs a methyl lesion by transferring the methyl onto itself, and that self-methylation — the record of the damage — turns Ada into a transcriptional activator of the repair regulon, so a cell that has met the condition is better equipped to survive its recurrence, conserved across diverged bacteria (Nakabeppu & Sekiguchi 1986; Sedgwick & Vaughan 1991). It is heritable capacity-to-respond written by the environment into the sequence (§3.5) — the property a static-table model cannot represent, present from the first templated addition.

3.4 The two layers, made physical — condition → structure → function

The two layers — condition→symbol and symbol→referent — are not only modelled (§4); they have a concrete, non-living, condition-set instantiation in RNA itself, in which both are directly measurable.

The encode layer: condition → structure. [Established + Interpretation] Physical conditions funnel an RNA into a defined fold. Temperature opens or closes an RNA-thermometer hairpin in a zipper-like equilibrium (Kortmann & Narberhaus 2012); K^+ folds a guanine tract into a G-quadruplex (Monsen 2022); acidic pH assembles a cytosine tract into an i-motif (Zeraati 2018); a divalent metal selects a catalytic fold (Hsiao 2013); and a small-molecule ligand reshapes an aptamer or ribozyme on binding (Winkler 2004). RNA folding is a function of the conditions — a real, enumerable map from {pH, ionic strength, temperature, ligand} to adopted structure.

The decode layer: structure → binding and catalysis. [Established + Interpretation] The adopted fold then sets what the RNA binds and what reaction it runs. The thermometer's open form admits the ribosome and turns translation on (Kortmann & Narberhaus 2012); the metabolite-bound *glmS* fold becomes a self-cleaving ribozyme, so the structure *is* the catalyst and the metabolite is its switch (Winkler 2004); the correctly folded

polymerase ribozyme copies an RNA template into a working ribozyme (Wochner/Holliger 2011); the Fe(II)-loaded fold runs electron-transfer chemistry that the Mg(II) fold does not (Hsiao 2013). Structure → function is equally a real, enumerable map.

The genetic code's own second layer is the same kind of object — chemistry-set, not arbitrarily assigned. **[Established + Interpretation]** The instantiation above is folding→catalysis; the code's *literal* symbol→referent map — codon→amino acid — is itself physically grounded. Across hundreds of independent RNA binding sites, cognate coding triplets cluster around their own amino acids' RNA binding sites at vanishing probability ($P \approx 5\times10^{-45}$ for codons, 2×10^{-46} for anticodons), evidence for a stereochemical era in which a majority of amino acids entered the code through direct chemical interaction rather than assignment (Yarus 2009). And the assignment chemistry runs without protein: de-novo flexizyme ribozymes charge amino acids onto tRNAs (Ohuchi 2007), amino-acid-specific RNA aptamers bind free amino acids enantioselectively (Geiger 1996), and the peptide bond that joins them is formed by a ribozyme — the ribosome's peptidyl-transferase centre, with no protein within 18 Å of its active site (Polacek 2001). **[Interpretation]** So the code's actual two layers — codon→amino acid, and the peptide chemistry that reads it — are realisable in non-living RNA and have a chemical, not arbitrary, basis. The symbol→referent step — historically the harder one to account for — is itself physically grounded, by the folding→catalysis analogue and by the genetic code's own codon→amino-acid map.

The composed map is non-living and set entirely by the chemistry. **[Interpretation]** Composed, the two maps are a physical encode/decode table: conditions in, structure as the intermediate symbol, binding-and-catalysis out — with no living organism and no designer-chosen dictionary, the mapping set entirely by the chemistry. It also has the property a genetic system needs and a salt crystal or a thermostat does not: it integrates *coordinated* inputs — the *glmS* ribozyme reads a metabolite (glucosamine-6-phosphate) that is itself the building block of another macromolecular pathway, so the structure adopted, and therefore the catalysis run, is set jointly by the ionic and ligand environment. This is the same two-layer logic the model in §4 abstracts, here carried out in measured RNA chemistry rather than simulation: encode and decode performed by molecules, with no living organism and no designer-chosen dictionary. (The same comparison explains why the heritable medium is double-stranded and not single: a duplex can present a G-quadruplex on one strand and an i-motif on its complement — two condition-channels with opposite outputs that a single RNA strand cannot offer; §3.1.)

The same two layers, filled in measured metabolic variables — the chromatin-mark layer. **[Established + Interpretation]** The RNA instantiation above is one physical encode/decode table; the chromatin-mark layer of §3.2 is a second, in which the conditions are metabolite pools and the readout is a written mark. Each row's condition→reader→mark dependency is measured (cited); reading the rows together as one filled grid is our interpretation.

Condition (the encode input)	Reader state — encode	Written mark / outcome — decode	Measured anchor
O_2 high · αKG high · acetyl-CoA high	KDM6A and the TET / JmjC dioxygenases run (O_2 + αKG present); HAT is supplied	H3K27 demethylated, 5mC oxidised, H3K27ac written → differentiation-permissive	Chakraborty 2019; Wellen 2009
O_2 low (hypoxia) · αKG high	KDM6A stalls (its O_2 threshold unmet); the paralog KDM6B, not O_2-gated at that threshold, is unaffected	H3K27me retained at KDM6A-dependent sites → differentiation blocked	Chakraborty 2019
O_2 high · 2-HG high (IDH-mutant)		marks retained genome-wide (hyper-methylation)	Xu 2011

	the αKG-dependent dioxygenases (TET + histone demethylases) are competitively inhibited		
acetyl-CoA low	HAT under-supplied	acetyl mark not written	Wellen 2009

One physical variable is read at several thresholds — O_2 by KDM6A but not by its paralog KDM6B (Chakraborty 2019) — so a single condition writes a graded pattern across marks, not one symbol: the positive reason the mapping is the kinetics of the readers, not a separable lookup. This is the encode→decode table of §3.4 in measured metabolic variables — the physical counterpart of the model's §4 table, grounded in extant enzymology rather than abstracted to vary the inventory.

3.5 The same writing logic, observed in living cells — structure routes the machinery, fidelity is biased on purpose, before any nucleus

The formula *written_sequence = f(conditions, machinery and its fidelity, template binding-context)* is not only prebiotic chemistry (§3.1) or extant marking (§3.2); all three of its terms are visible, intact, in living bacteria — which, having no nucleus, are the ancestral case. The same condition→structure→function logic that §3.4 shows in RNA catalysis runs here on the writing of DNA itself.

The binding context — the fold — decides which machinery writes. **[Established + Interpretation]** At repeat loci that fold into non-B structure, the structure, not enzyme abundance, sets the writing capacity. The mismatch-repair endonuclease preferentially binds a G-quadruplex over its own cognate duplex but does not process it there — the fold dictating where repair engages, a structural element in the regulation of programmed antigenic variation (Savitskaya 2023); and the expansion writing is replication-independent, catalysed by human cell extracts with no S-phase, through hairpin intermediates and stimulated by MutSβ (Stevens 2013). Which structure forms is then set by the conditions: K^+ selects the G-quadruplex (Siebenmorgen & Zacharias 2017), acidic pH switches the same locus to the i-motif (Zeraati 2018), and an oxidised guanine read by OGG1 toggles the structural address on or off (Fleming 2017). The fold is the address, and the conditions write it.

The fidelity term is condition-set and locus-addressed — biased on purpose, not by failure. **[Established + Interpretation]** "Fidelity" in the formula is not always "as faithful as possible." Bacterial reverse-transcriptase elements write DNA directly from an RNA template as a condition-triggered response — the retron writes a defined RNA/DNA product and deploys it in anti-phage defence (Millman 2020) — and diversity-generating retroelements implement programmed, adenine-specific biased writing: the reverse transcriptase's own misincorporation, restricted to adenine positions and aimed at a template-defined target, harnessed as managed diversification rather than error (Naorem 2017). And in the limit the machinery *is* the template: the antiphage reverse transcriptase Drt3b synthesises a defined poly(AC) DNA strand with no nucleic-acid template at all, its own conserved active-site residues enforcing the base alternation (Deng 2026) — protein sequence specifying DNA sequence — while a related defence reverse-transcribes a noncoding RNA into an entire *de novo* gene, switched to its expressible double-stranded form only on infection (Tang 2024). The machinery — its bias, and in the limit the very sequence it lays down — is itself a written, addressed, condition-deployed variable.

The adaptor that reads the code also biases a write. **[Established + Interpretation]** A transfer RNA decodes a codon into an amino acid — and is also used to begin the writing of DNA. Reverse transcription does not start de novo: a specific cellular tRNA anneals to a defined primer-binding site, and the reverse transcriptase extends the tRNA's own 3′ end into DNA, so the tRNA selects and primes where the write begins — a highly specific, cognate self-tRNA, selectively packaged for the purpose (characterised for tRNA-Lys3 in HIV-1; Abbink & Berkhout 2008). It biases a write in its translational role too: a tRNA's modification state sets which codons are read, and so which messages are translated — measured in the writers' own codon content below. The same adaptor biases writing at both levels, nucleotide and protein.

No compartment is ancestral; the nucleus is a consequence of writing. [Established + Interpretation] These reverse-transcriptase defence systems are a modern overlay — bacterial anti-phage immunity — but they write with the same ancient chemistry that ran before them: the two-divalent-metal nucleotidyl-transfer of the replicative polymerases, in enzymes descended from mobile group II introns — the progenitors of the spliceosome, the LINE retrotransposons and telomerase — whose proliferation within an early genome is thought to have necessitated the evolution of the nuclear envelope itself (Lambowitz & Belfort 2015). DNA writing predates the nucleus; the compartment is downstream of the writing, not a precondition for it.

The licensing conditions are the ancestral ones. [Established + Interpretation] The variables that set the structural address are the variables of the first cells: phylogenomic reconstruction of the inorganic requirements of the universal cellular components places the first cells in pools of high K^+/Na^+ ratio with Zn^{2+}, Mn^{2+} and phosphate, at anoxic inland geothermal fields rather than in the sea (Mulkidjanian 2012) — high K^+ to license the G-quadruplex address, the divalent metals to run the two-metal engine. The conditions that instruct fate-relevant biased writing in a cell today are the conditions of the origin.

The same write, measured in one genome — codon content as heritable response-capacity. [Established + Interpretation] The claim that the write records a condition *and improves the future response to it* can be made empirical at the sequence level, on a real genome. In *E. coli* K-12 MG1655 (RefSeq GCF_000005845.2; 4,226 coding sequences), the stress-induced DNA-writers — the error-prone polymerase *dinB*, the SOS polymerase-V subunits *umuC*/*umuD*, and the stress class-Ib ribonucleotide reductase *nrdF* — retain precisely the modification-dependent synonymous codons that translation-optimised genes purge. Against the ribosomal-protein reference set (54 genes; Mann–Whitney one-sided, Benjamini–Hochberg across five amino-acid axes) the writers are enriched for the MiaA-dependent Leu-UUR codons (subset median 0.28 vs 0.04; BH-q = 0.003) and the MnmA-thiolated Gln-CAA codon (0.46 vs 0.11; BH-q = 0.034); the other three axes are null. These are the mechanistically correct codons: the wobble-uridine modifications decode exactly these NNA/NNG codons (Moukadiri 2009), and MiaA modification is required to translate the master stress regulator RpoS, whose own message is Leu-UUR-enriched (Thompson & Gottesman 2014) — reproduced here as the positive control (*rpoS* Leu-UUR 0.295 vs *rpoD* 0.056) — while the same modified-tRNA→codon-biased-translation coupling tunes the DNA-damage writers in yeast (Begley 2007). The honest negative matters and matches this document's voice: against the whole-genome average the writers show no enrichment — their subset medians sit at the genome-wide cognate fraction (Leu-UUR 0.28 vs the genome-wide 0.257) — because the constitutive class-Ia reductase *nrdAB* is itself codon-optimised; the signal appears only against the correctly specified optimised reference. [Interpretation] The reading is the formula's: a gene carrying the stress-modified codons is one that translates better when that stress recurs, so the codon content of the writers is a stable, heritable write of the capacity to respond to the condition that deploys them — biological function as a readout of the writing context, shown in measured sequence rather than in the model.

All three terms observed in compartment-free cells, not inferred. [Interpretation] Read together, the three terms of the formula are not inferred from a model but observed in compartment-free cells: the conditions license a structure, the structure routes the machinery, and the machinery's deliberately biased fidelity writes the change — condition-instructed, structure-addressed, replication-independent, and older than the nucleus. This is shown in measured genomic sequence — in the writing of the genetic polymer itself.

3.6 Metabolism wrote the first writing — the inputs came before the mapping, and both came before proteins

At the origin, the conditions that instructed the writing were metabolism itself. [Established + Interpretation] Nucleotides are not only the genetic alphabet but the cofactor backbone of metabolism — NAD, FAD, coenzyme A, S-adenosylmethionine, the cyclic messengers — so the abundance of an activated nucleotide is a physical readout of metabolic state, and a catalyst that incorporates the abundant

species writes that state into sequence. The pieces are already in this paper. A metabolite switches a ribozyme: glucosamine-6-phosphate turns the glmS fold into a self-cleaving catalyst, the structure the enzyme and the metabolite its switch (Winkler 2004). A nucleotide instructs which nucleotide is written: ribonucleotide reductase's specificity site reads a bound deoxynucleotide and sets which one it makes (Nordlund & Reichard 2006), and when NAD is abundant the polymerase writes it directly as a 5′ cap (§3.2). The bond-making fold still runs this way: in their signalling role the cGAS/DncV-like nucleotidyltransferases synthesise diverse cyclic di- and tri-nucleotides de novo, an active-site lid selecting purine or pyrimidine (Whiteley/Kranzusch 2019) — an activated transferase, and a nucleotide product whose composition is set by the fold. The write was instructed by the metabolic environment before there was anything else to instruct it.

Two distinctions keep the claim honest: there are two codes, and at the origin there were no proteins. **[Interpretation]** Metabolism wrote the first, compositional layer — which nucleotides, in what modified and cyclic forms, gating which catalytic events; the sequential codon–amino-acid mapping is a second layer, and it came not from abundance but from direct chemical affinity, cognate triplets clustering at their own amino acids' RNA binding sites at vanishing probability (Yarus 2009). These are complementary layers of one origin, and both sit inside the formula — metabolism its conditions, stereochemistry its template-context, condition-gated catalysis its machinery. Those catalysts were not proteins, which postdate the code; the allosteric protein writers above are a later re-implementation of the logic. At the origin the condition-gated writers were ribozymes — metabolite-switched and amino-acid-charging (Ohuchi 2007) — and metal catalysis on the prebiotic floor of §3.1. That metabolism authored the first writing is the strong, still-unclosed claim of a metabolism-first origin, owed the hard step from compositional to heritable sequence; the mechanism that makes it conceivable — metabolite-gated, structure-selected writing — is established, and on the account given here it never went away.

4. The quantitative model, and its honest scope

To give the contingency claim a number, we built and ran a reproducible model (Python 3.12 / NumPy / SciPy; the source and outputs are provided in the accompanying reproducibility package). Thirty-two condition-states map through a symbol channel to thirty-two referents (a two-layer system). Per-condition accessibility is the product of two independent random surfaces drawn fresh for each molecular inventory (a bind×write AND-gate); a Hill function of sharpness n converts accessibility into write reliability (the analog→digital, proofreading-like step — ultrasensitivity in the sense of Goldbeter & Koshland 1984); redundancy R lets several symbols share a referent. **Fitness is the mutual information between condition and referent, referencing no target map** — selection rewards distinguishability, never a particular code.

Real outputs:

- A minimal channel (one symbol per meaning, shallow switching) lifts the condition–referent mutual information only modestly above a random baseline and plateaus well short of the channel ceiling in all 200 inventories — a structural limit set by the shallow mapping, not a search failure.
- Approaching the channel ceiling requires *both* redundancy and sharp switching together: at a redundancy of ≈ 3 (matching the genetic code's 64 codons / 20 amino acids ≈ 3.2) with sharp switching, the code captures nearly all of the available information, whereas removing either ingredient — or both — falls substantially short. The model therefore *derives* codon-style redundancy and a sharp, proofreading-grade switch (ultrasensitivity, after Goldbeter & Koshland 1984) as the two requirements for a high-capacity code.
- **Contingency (the degeneracy result):** across 200 molecular inventories at the realistic setting, every inventory reaches near-maximal capacity, yet all 200 produce *distinct* codes — no code recurs. An ordinary selective process — a hill-climb under the same fitness, not a solver — likewise yields 150

inventories that are each distinct and ~97% pairwise-divergent. The mapping is contingent on the inventory, not entailed by the architecture: emergent, not authored.

How the emitted code is structured. For one representative molecular inventory of the 200 (seed 0; the complete machine-readable set is provided in the accompanying reproducibility package), the 32 condition-states are the physical environment — the ~2^5 combinations of independent axes such as O_2, pH, K^+/Na^+, and the αKG/2-HG and acetyl-CoA/SAM metabolite pools — and the encoder maps each to a symbol from a 96-symbol intermediate alphabet. The decoder is a fixed projection, symbol $s \rightarrow s \bmod 32$, held identical across all 200 inventories; it is 3-to-1 by construction (96 → 32) and so sets the redundancy $R = 3$ — a value put in rather than derived from this run, motivated by the real code's 64 codons / 20 amino acids ≈ 3.2 and shown necessary, with sharp switching, by the ablations above. What varies across the 200 inventories is the bind×write surface — physically, the set of enzyme reader thresholds that interrogates that environment — so all the contingency lives in the encoder: the same physical conditions, read by a different inventory of molecular thresholds, yield a different code, 200 distinct and referencing no target. The model abstracts the specific thresholds: it draws random surfaces rather than computing O_2 or pH kinetics, so that the inventory can be varied — the measured thresholds themselves are §3.2's, tabulated physically in §3.4. The correspondence is objectively checkable in the machine-readable output: referent = encoded symbol mod 32 for every condition, and in this inventory the 32 conditions map to 32 distinct referents.

Contingent at origin, universal once frozen. **[Established + Interpretation]** A contingent map invites the obvious objection: if any of these codes could have formed, why is there one near-universal code? Contingency at origin and universality after fixation are not in tension — they are what an early-frozen optimum reached by communal innovation-sharing looks like. Under plausible biological constraints the canonical code sits at or very near a *global* optimum for error minimisation, with better codes extremely rare, and the forces acting before and after its fixation differ (Freeland 2000); and a dynamical account derives that same universality and optimality from collective, non-Darwinian (horizontal) mechanisms in early communal life, before the onset of vertical descent (Vetsigian 2006). The degeneracy result is the origin half of that picture — many codes are reachable; the realised code is the one that froze early and well.

The honest scope, which is itself the point of Section 2. The number this model reports is a *capacity*: the mutual information a condition → symbol → referent code can carry, bounding the mapping layer. What the model does not simulate is the step that makes the code biological — but that step is not a missing abstraction, it is realised physically: the write is itself the memory. The sequence the code lays down records the condition that wrote it — in which triplet or length was selected — and in the same act tunes the future response to that condition, shown at the sequence level in the codon analysis of §3.5, where the writers' codon content is a heritable record coupling their future translation to the condition that deploys them. It therefore bounds the "mapping is contingent, not forced" layer; the memory layer — the depletion, the heritable mark, the closed loop of §3.3 — is not omitted from the system but carried by the write, where Section 3 measures it. We flag it deliberately: even an information-theoretic model defaults to the technological frame, modelling a code as a static input → output table. The real system is the historical, metabolism-embedded one specified and evidenced in Section 3; the number here bounds only its mapping layer.

The model is a decision graph; the table is its shadow. **[Interpretation]** The capacity was never a property of the emitted table — it was produced by the graph that fills it. The ablations show it: the flat one-symbol-per-condition map is capacity-limited, and approaching the ceiling requires the bind×write AND-gate (a coincidence node), the sharp Hill switch (a thresholded decision node, after Goldbeter & Koshland 1984), and redundancy (several paths converging on one outcome). Those are not rows in a lookup table; they are the gates, thresholds and convergences of a branching constraint graph, and they are where the capacity comes from. The code the model emits is the memoryless *projection* of that graph — the left-column object, read off

the right-column object — which is precisely why a filled table can be emitted at all while the mapping is still not a separable table (§3.2): the two describe the same object at different resolutions. Two consequences follow. *First*, a graph with *state* carries information a table cannot — not only in its instantaneous input→output rows but in *which path was taken to reach a node*, the history of constraints already satisfied, which is the memory layer of §3.3 made structural. The capacity then has two independent sources that each scale on demand: conditional *depth* (a deeper binary cascade resolves more states) and *path/state* (the historical bits a combinational table discards). *Second*, the contingency result is a property of the graph, not the table: different molecular inventories impose different constraint sets, which funnel to different satisfying outcomes — akin to the attractor picture of a constrained dynamical system — so 200 inventories yield 200 distinct codes with no author. The honest scope of §2 is then exact: the table-number bounds the mapping layer, while the decision graph carries both the mapping and the memory. And the graph is datable. Its *bit-generating* nodes — the coincidence gate, the threshold and the redundancy — are prebiotic chemistry, old enough to bear the origin claim; its chromatin and oxygen-sensing nodes are later elaboration that shows the same logic kept running. The node-by-node dating, with the Great Oxidation Event as the hard line, is summarised in Figure 1.

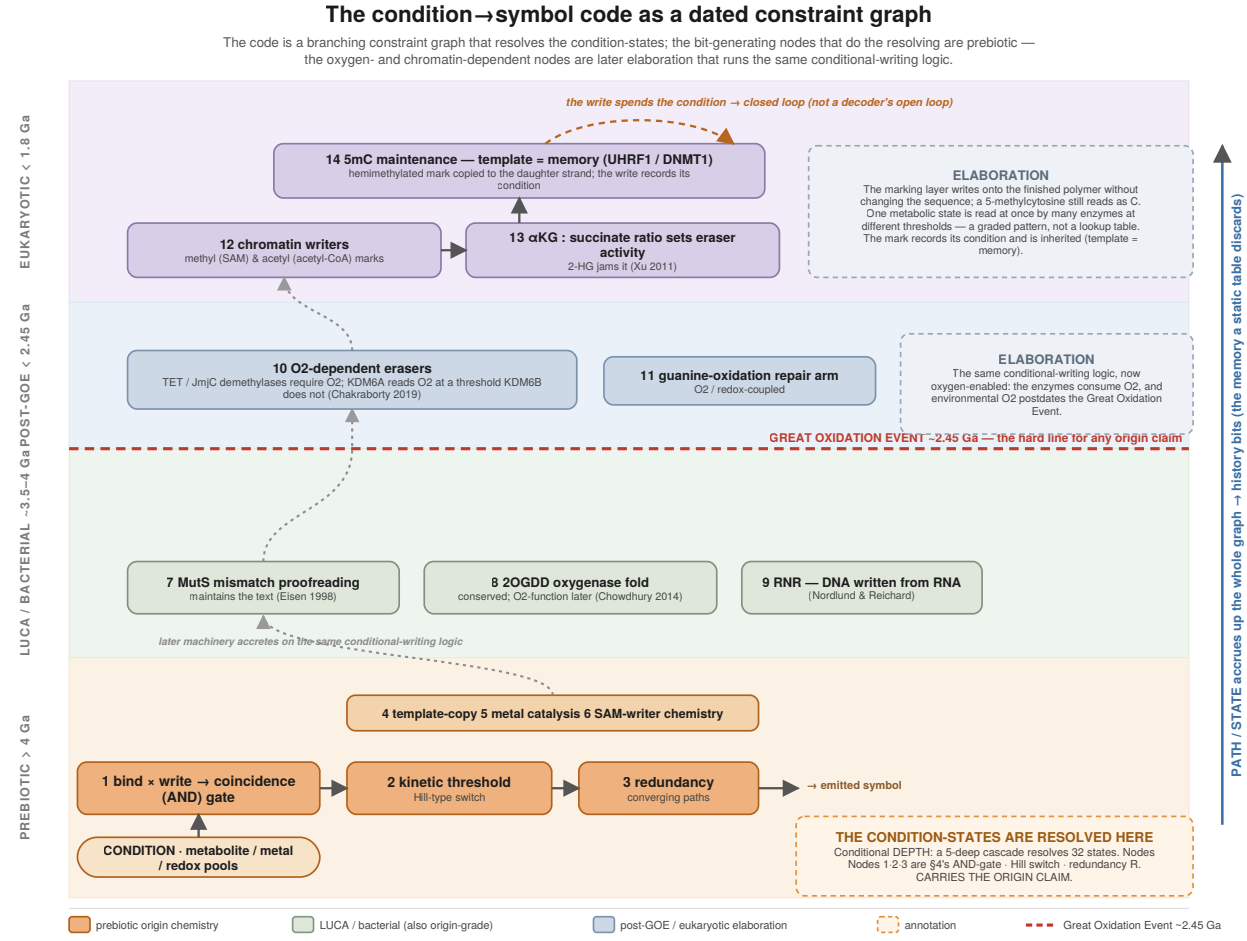

Figure 1. The condition→symbol code drawn as a dated constraint graph. Every node is placed in the era it arose, oldest at the bottom, and numbered in that order. Warm nodes — the prebiotic writing chemistry and the bit-generating coincidence / threshold / redundancy logic — are old enough to carry the origin claim; grey nodes are later elaboration that runs the same conditional-writing logic. The information that resolves the 32 condition-states has two sources: conditional *depth* on the prebiotic coincidence / threshold / redundancy nodes (the same three ingredients §4 identifies) and *path/state* accumulated up the whole graph. The Great Oxidation Event (~2.45 Ga) is the hard line: no oxygen-using node can carry the origin claim. The code the model emits (§4) is the memoryless projection of this graph.

5. Limitations and open questions

The central claim here — that biological information is a condition-set, self-modifying, memory-writing process rather than a static cipher — is a reframing supported by established biology, not a wet-laboratory abiogenesis result. The model demonstrates contingency but, by construction, only for the mapping layer; the memory layer is evidenced from extant cells, which does not by itself establish prebiotic realisability. A decisive empirical test of the full claim would be a prebiotic experiment: an RNA system assembled from simple non-living precursors, with biological contamination excluded, in which writing is condition-set — the product determined by the conditions present and persisting as a re-readable record — so that what is demonstrated is the right kind of object, a written state that records the condition that set it rather than a bare self-replicator. The nearest existing experimental handles are systems already cited here — amino-acid-binding RNA aptamers (Geiger 1996) and flexizyme aminoacylation ribozymes (Ohuchi 2007), which realise a non-living charge/assign step — together with the RNA polymerase ribozyme, which already copies a functional ribozyme from an RNA template (Wochner/Holliger 2011); and the open frontier that ribozyme replication must improve its copying fidelity before fully autonomous RNA evolution is reached (Papastavrou/Joyce 2024). Short of that experiment, what this account adds is a more accurate premise, a continuous mechanism from prebiotic chemistry through to extant cells, grounded evidence, and a reproducible model of the contingency it claims.

6. Conclusion

One system, written in nucleotides. [Interpretation] The genetic material is nucleic acid, and the same nucleotides that spell it out are, pervasively, the carriers that run metabolism. The cell's energy currency is the ribonucleoside triphosphates (ATP, GTP, CTP, UTP); its redox currency is nucleotide-based (NAD+/NADH, NADP+/NADPH, FAD); its acyl carrier is coenzyme A; its methyl donor is S-adenosylmethionine; its sugars are handed off as nucleotide-sugars for glycosylation and glycogen (UDP-glucose, UDP-GlcNAc, GDP-mannose, CMP-sialic acid); its phospholipids are assembled through CDP-choline and CDP-diacylglycerol; its sulfate is activated as the adenosine conjugate PAPS; and its second messengers are cyclic nucleotides (cAMP, cGMP, the cyclic di-nucleotides). Across energy, redox, acyl, methyl, sugar, lipid, sulfur, and signalling, the carrier is a nucleotide — most often built on the same adenosine handle a nucleotide-binding maker would have recognised (§3.2). The genome's alphabet and the cell's metabolic currency are one chemical inventory, not two.

The integration is a flow, not a wiring diagram. The ribonucleotides are at once the monomers of the labile running layer (RNA: catalysis, regulation, metabolite contact) and the stock from which the stable archive is cut: ribonucleotide reductase is the single de-novo gate that draws from the shared pool and commits it, one way, into DNA (§3.1). Building or marking the genome therefore debits the same pool that runs the metabolism, and the conversion between the two is a metabolic branch point, not a side reaction. Code, currency, and archive are three states of one nucleotide flow.

The origin question follows from the chemistry. There is no moment at which a static dictionary self-assembles, because writing was condition-dependent nucleotide addition from the first templated step, in the same nucleotide stock that ran the proto-metabolism. Neither half of the code was authored: the mapping from triplet to amino acid was found rather than assigned (§3.4), and metabolism supplied the inputs and the first writes — the abundance of an activated nucleotide standing in for the state of the cell (§3.6). What changes across that history is only what fixes the sequence — a nucleic-acid template early, a folded protein later — never the condition-instructed character of the writing itself. So the genetic code is the durable record of one metabolism-embedded writing process, written in the molecules that also run the cell, in the currency it spends to write:

each write records a condition and, by spending the metabolite, alters it. That is the literal sense in which this information records and alters its own conditions.

7. Declarations

Data and code availability. The reproducible model underlying Section 4 — its source code and complete machine-readable output — is provided as a reproducibility package deposited at Zenodo (https://doi.org/10.5281/zenodo.20966659; CC BY 4.0). No restricted or third-party datasets were used; all biological evidence is drawn from the peer-reviewed literature indexed in PubMed, with citations and DOIs listed in Section 8.

Funding. This research received no external funding.

Competing interests. The author declares no competing interests.

Author contributions. D.L. conceived the framework, performed the analyses and modelling, prepared the figure, and wrote the manuscript.

AI-assistance disclosure. This manuscript was prepared with the assistance of a large language model (Anthropic's Claude) for literature search and synthesis, manuscript drafting and editing, computational modelling, and figure preparation. All citations were verified against primary records in PubMed; the author reviewed all content and takes full responsibility for the analyses, claims, and conclusions.

Classification. Biology; Evolutionary Biology; Molecular Biology.

8. References (PubMed; DOIs as links)

Retrieved and verified via PubMed; claims above are paraphrased.

1. Abbink, T. E. M., & Berkhout, B. (2008). HIV-1 reverse transcription initiation: a potential target for novel antivirals? *Virus research*, *134*(1-2), 4–18. https://doi.org/10.1016/j.virusres.2007.12.009
2. Avvakumov, G. V., Walker, J. R., Xue, S., Li, Y., Duan, S., Bronner, C., Arrowsmith, C. H., & Dhe-Paganon, S. (2008). Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1. *Nature*, *455*(7214), 822–825. https://doi.org/10.1038/nature07273
3. Begley, U., Dyavaiah, M., Patil, A., Rooney, J. P., DiRenzo, D., Young, C. M., Conklin, D. S., Zitomer, R. S., & Begley, T. J. (2007). Trm9-catalyzed tRNA modifications link translation to the DNA damage response. *Molecular cell*, *28*(5), 860–870. https://doi.org/10.1016/j.molcel.2007.09.021
4. Bird, J. G., Zhang, Y., Tian, Y., Panova, N., Barvík, I., Greene, L., Liu, M., Buckley, B., Krásný, L., Lee, J. K., Kaplan, C. D., Ebright, R. H., & Nickels, B. E. (2016). The mechanism of RNA 5′ capping with NAD+, NADH and desphospho-CoA. *Nature*, *535*(7612), 444–447. https://doi.org/10.1038/nature18622
5. Bostick, M., Kim, J. K., Estève, P. O., Clark, A., Pradhan, S., & Jacobsen, S. E. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science (New York, N.Y.)*, *317*(5845), 1760–1764. https://doi.org/10.1126/science.1147939
6. Burcar, B. T., Barge, L. M., Trail, D., Watson, E. B., Russell, M. J., & McGown, L. B. (2015). RNA Oligomerization in Laboratory Analogues of Alkaline Hydrothermal Vent Systems. *Astrobiology*, *15*(7), 509–522. https://doi.org/10.1089/ast.2014.1280
7. Cahová, H., Winz, M. L., Höfer, K., Nübel, G., & Jäschke, A. (2015). NAD captureSeq indicates NAD as a bacterial cap for a subset of regulatory RNAs. *Nature*, *519*(7543), 374–377. https://doi.org/10.1038/nature14020
8. Camprubi, E., Jordan, S. F., Vasiliadou, R., & Lane, N. (2017). Iron catalysis at the origin of life. *IUBMB life*, *69*(6), 373–381. https://doi.org/10.1002/iub.1632
9. Chakraborty, A. A., Laukka, T., Myllykoski, M., Ringel, A. E., Booker, M. A., Tolstorukov, M. Y., Meng, Y. J., Meier, S. R., Jennings, R. B., Creech, A. L., Herbert, Z. T., McBrayer, S. K., Olenchock, B. A., Jaffe, J. D., Haigis, M. C., Beroukhim, R., Signoretti, S., Koivunen, P., & Kaelin, W. G., Jr (2019). Histone demethylase KDM6A directly senses oxygen to control chromatin and cell fate. *Science (New York, N.Y.)*, *363*(6432), 1217–1222. https://doi.org/10.1126/science.aaw1026
10. Chowdhury, R., Sekirnik, R., Brissett, N. C., Krojer, T., Ho, C. H., Ng, S. S., Clifton, I. J., Ge, W., Kershaw, N. J., Fox, G. C., Muniz, J. R. C., Vollmar, M., Phillips, C., Pilka, E. S., Kavanagh, K. L., von Delft, F., Oppermann, U., McDonough, M. A.,

Doherty, A. J., & Schofield, C. J. (2014). Ribosomal oxygenases are structurally conserved from prokaryotes to humans. *Nature*, *510*(7505), 422–426. https://doi.org/10.1038/nature13263

11. Deng, P., Lee, H., Armijo, C., Wang, H., & Gao, A. (2026). Protein-templated synthesis of dinucleotide repeat DNA by an antiphage reverse transcriptase. *Science (New York, N.Y.)*, *392*(6804), 1274–1281. https://doi.org/10.1126/science.aed1656
12. Duzdevich, D., Carr, C. E., & Szostak, J. W. (2020). Deep sequencing of non-enzymatic RNA primer extension. *Nucleic acids research*, *48*(12), e70. https://doi.org/10.1093/nar/gkaa400
13. Eisen, J. A. (1998). A phylogenomic study of the MutS family of proteins. *Nucleic acids research*, *26*(18), 4291–4300. https://doi.org/10.1093/nar/26.18.4291
14. Fedeles, B. I., Singh, V., Delaney, J. C., Li, D., & Essigmann, J. M. (2015). The AlkB family of Fe(II)/α-ketoglutarate-dependent dioxygenases: repairing nucleic acid alkylation damage and beyond. *The Journal of biological chemistry*, *290*(34), 20734–20742. https://doi.org/10.1074/jbc.R115.656462
15. Fleming, A. M., Zhu, J., Ding, Y., & Burrows, C. J. (2017). 8-Oxo-7,8-dihydroguanine in the Context of a Gene Promoter G-Quadruplex Is an On-Off Switch for Transcription. *ACS chemical biology*, *12*(9), 2417–2426. https://doi.org/10.1021/acschembio.7b00636
16. Forterre, P. (2005). The two ages of the RNA world, and the transition to the DNA world: a story of viruses and cells. *Biochimie*, *87*(9-10), 793–803. https://doi.org/10.1016/j.biochi.2005.03.015
17. Freeland, S. J., Knight, R. D., Landweber, L. F., & Hurst, L. D. (2000). Early fixation of an optimal genetic code. *Molecular biology and evolution*, *17*(4), 511–518. https://doi.org/10.1093/oxfordjournals.molbev.a026331
18. Geiger, A., Burgstaller, P., von der Eltz, H., Roeder, A., & Famulok, M. (1996). RNA aptamers that bind L-arginine with sub-micromolar dissociation constants and high enantioselectivity. *Nucleic acids research*, *24*(6), 1029–1036. https://doi.org/10.1093/nar/24.6.1029
19. Goldbeter, A., & Koshland, D. E., Jr (1984). Ultrasensitivity in biochemical systems controlled by covalent modification. Interplay between zero-order and multistep effects. *The Journal of biological chemistry*, *259*(23), 14441–14447.
20. He, Y. F., Li, B. Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., Sun, Y., Li, X., Dai, Q., Song, C. X., Zhang, K., He, C., & Xu, G. L. (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science (New York, N.Y.)*, *333*(6047), 1303–1307. https://doi.org/10.1126/science.1210944
21. Hsiao, C., Chou, I. C., Okafor, C. D., Bowman, J. C., O'Neill, E. B., Athavale, S. S., Petrov, A. S., Hud, N. V., Wartell, R. M., Harvey, S. C., & Williams, L. D. (2013). RNA with iron(II) as a cofactor catalyses electron transfer. *Nature chemistry*, *5*(6), 525–528. https://doi.org/10.1038/nchem.1649
22. Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., Yi, C., Lindahl, T., Pan, T., Yang, Y. G., & He, C. (2011). N^6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nature chemical biology*, *7*(12), 885–887. https://doi.org/10.1038/nchembio.687
23. Jiao, X., Doamekpor, S. K., Bird, J. G., Nickels, B. E., Tong, L., Hart, R. P., & Kiledjian, M. (2017). 5′ End Nicotinamide Adenine Dinucleotide Cap in Human Cells Promotes RNA Decay through DXO-Mediated deNADding. *Cell*, *168*(6), 1015–1027.e10. https://doi.org/10.1016/j.cell.2017.02.019
24. Joshi, P. C., Dubey, K., Aldersley, M. F., & Sausville, M. (2015). Clay catalyzed RNA synthesis under Martian conditions: Application for Mars return samples. *Biochemical and biophysical research communications*, *462*(2), 99–104. https://doi.org/10.1016/j.bbrc.2015.04.044
25. Kortmann, J., & Narberhaus, F. (2012). Bacterial RNA thermometers: molecular zippers and switches. *Nature reviews. Microbiology*, *10*(4), 255–265. https://doi.org/10.1038/nrmicro2730
26. Krebs, C., Galonić Fujimori, D., Walsh, C. T., & Bollinger, J. M., Jr. (2007). Non-heme Fe(IV)-oxo intermediates. *Accounts of chemical research*, *40*(7), 484–492. https://doi.org/10.1021/ar700066p
27. Lambowitz, A. M., & Belfort, M. (2015). Mobile Bacterial Group II Introns at the Crux of Eukaryotic Evolution. *Microbiology spectrum*, *3*(1), MDNA3–2014. https://doi.org/10.1128/microbiolspec.MDNA3-0050-2014
28. Lundin, D., Berggren, G., Logan, D. T., & Sjöberg, B. M. (2015). The origin and evolution of ribonucleotide reduction. *Life (Basel, Switzerland)*, *5*(1), 604–636. https://doi.org/10.3390/life5010604
29. Millman, A., Bernheim, A., Stokar-Avihail, A., Fedorenko, T., Voichek, M., Leavitt, A., Oppenheimer-Shaanan, Y., & Sorek, R. (2020). Bacterial Retrons Function In Anti-Phage Defense. *Cell*, *183*(6), 1551–1561.e12. https://doi.org/10.1016/j.cell.2020.09.065
30. Monsen, R. C., Trent, J. O., & Chaires, J. B. (2022). G-quadruplex DNA: A Longer Story. *Accounts of chemical research*, *55*(22), 3242–3252. https://doi.org/10.1021/acs.accounts.2c00519
31. Moukadiri, I., Prado, S., Piera, J., Velázquez-Campoy, A., Björk, G. R., & Armengod, M. E. (2009). Evolutionarily conserved proteins MnmE and GidA catalyze the formation of two methyluridine derivatives at tRNA wobble positions. *Nucleic acids research*, *37*(21), 7177–7193. https://doi.org/10.1093/nar/gkp762
32. Mrnjavac, N., Schwander, L., Brabender, M., & Martin, W. F. (2024). Chemical Antiquity in Metabolism. *Accounts of chemical research*, *57*(16), 2267–2278. https://doi.org/10.1021/acs.accounts.4c00226

33. Mulkidjanian, A. Y., Bychkov, A. Y., Dibrova, D. V., Galperin, M. Y., & Koonin, E. V. (2012). Open questions on the origin of life at anoxic geothermal fields. *Origins of life and evolution of the biosphere : the journal of the International Society for the Study of the Origin of Life*, *42*(5), 507–516. https://doi.org/10.1007/s11084-012-9315-0
34. Nakabeppu, Y., & Sekiguchi, M. (1986). Regulatory mechanisms for induction of synthesis of repair enzymes in response to alkylating agents: ada protein acts as a transcriptional regulator. *Proceedings of the National Academy of Sciences of the United States of America*, *83*(17), 6297–6301. https://doi.org/10.1073/pnas.83.17.6297
35. Naorem, S. S., Han, J., Wang, S., Lee, W. R., Heng, X., Miller, J. F., & Guo, H. (2017). DGR mutagenic transposition occurs via hypermutagenic reverse transcription primed by nicked template RNA. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(47), E10187–E10195. https://doi.org/10.1073/pnas.1715952114
36. Nelson, J. R., Lawrence, C. W., & Hinkle, D. C. (1996). Deoxycytidyl transferase activity of yeast REV1 protein. *Nature*, *382*(6593), 729–731. https://doi.org/10.1038/382729a0
37. Nordlund, P., & Reichard, P. (2006). Ribonucleotide reductases. *Annual review of biochemistry*, *75*, 681–706. https://doi.org/10.1146/annurev.biochem.75.103004.142443
38. Ohuchi, M., Murakami, H., & Suga, H. (2007). The flexizyme system: a highly flexible tRNA aminoacylation tool for the translation apparatus. *Current opinion in chemical biology*, *11*(5), 537–542. https://doi.org/10.1016/j.cbpa.2007.08.011
39. Panagopoulos, A., Stout, M., Kilic, S., Leary, P., Vornberger, J., Pasti, V., Galarreta, A., Lezaja, A., Kirschenbühler, K., Imhof, R., Rehrauer, H., Ziegler, U., & Altmeyer, M. (2025). Multigenerational cell tracking of DNA replication and heritable DNA damage. *Nature*, *642*(8068), 785–795. https://doi.org/10.1038/s41586-025-08986-0
40. Papastavrou, N., Horning, D. P., & Joyce, G. F. (2024). RNA-catalyzed evolution of catalytic RNA. *Proceedings of the National Academy of Sciences of the United States of America*, *121*(11), e2321592121. https://doi.org/10.1073/pnas.2321592121
41. Patel, B. H., Percivalle, C., Ritson, D. J., Duffy, C. D., & Sutherland, J. D. (2015). Common origins of RNA, protein and lipid precursors in a cyanosulfidic protometabolism. *Nature chemistry*, *7*(4), 301–307. https://doi.org/10.1038/nchem.2202
42. Polacek, N., Gaynor, M., Yassin, A., & Mankin, A. S. (2001). Ribosomal peptidyl transferase can withstand mutations at the putative catalytic nucleotide. *Nature*, *411*(6836), 498–501. https://doi.org/10.1038/35078113
43. Saha, P., Panda, D., Müller, D., Maity, A., Schwalbe, H., & Dash, J. (2020). *In situ* formation of transcriptional modulators using non-canonical DNA i-motifs. *Chemical science*, *11*(8), 2058–2067. https://doi.org/10.1039/d0sc00514b
44. Savitskaya, V. Y., Strekalovskikh, V. V., Snyga, V. G., Monakhova, M. V., Arutyunyan, A. M., Dolinnaya, N. G., & Kubareva, E. A. (2023). *pilE* G-Quadruplex Is Recognized and Preferentially Bound but Not Processed by the MutL Endonuclease from *Neisseria gonorrhoeae* Mismatch Repair Pathway. *International journal of molecular sciences*, *24*(7), 6167. https://doi.org/10.3390/ijms24076167
45. Sedgwick, B., & Vaughan, P. (1991). Widespread adaptive response against environmental methylating agents in microorganisms. *Mutation research*, *250*(1-2), 211–221. https://doi.org/10.1016/0027-5107(91)90178-q
46. Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T. A., Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K., Tajima, S., Mitsuya, K., Okano, M., & Koseki, H. (2007). The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature*, *450*(7171), 908–912. https://doi.org/10.1038/nature06397
47. Siebenmorgen, T., & Zacharias, M. (2017). Origin of Ion Specificity of Telomeric DNA G-Quadruplexes Investigated by Free-Energy Simulations. *Biophysical journal*, *112*(11), 2280–2290. https://doi.org/10.1016/j.bpj.2017.04.036
48. Sirover, M. A., & Loeb, L. A. (1977). On the fidelity of DNA replication. Effect of metal activators during synthesis with avian myeloblastosis virus DNA polymerase. *The Journal of biological chemistry*, *252*(11), 3605–3610.
49. Snoussi, K., Nonin-Lecomte, S., & Leroy, J. L. (2001). The RNA i-motif. *Journal of molecular biology*, *309*(1), 139–153. https://doi.org/10.1006/jmbi.2001.4618
50. Stevens, J. R., Lahue, E. E., Li, G. M., & Lahue, R. S. (2013). Trinucleotide repeat expansions catalyzed by human cell-free extracts. *Cell research*, *23*(4), 565–572. https://doi.org/10.1038/cr.2013.12
51. Su, R., Dong, L., Li, C., Nachtergaele, S., Wunderlich, M., Qing, Y., Deng, X., Wang, Y., Weng, X., Hu, C., Yu, M., Skibbe, J., Dai, Q., Zou, D., Wu, T., Yu, K., Weng, H., Huang, H., Ferchen, K., … Chen, J. (2018). R-2HG exhibits anti-tumor activity by targeting FTO/m^6A/MYC/CEBPA signaling. *Cell*, *172*(1-2), 90–105.e23. https://doi.org/10.1016/j.cell.2017.11.031
52. Tang, S., Conte, V., Zhang, D. J., Žedaveinytė, R., Lampe, G. D., Wiegand, T., Tang, L. C., Wang, M., Walker, M. W. G., George, J. T., Berchowitz, L. E., Jovanovic, M., & Sternberg, S. H. (2024). De novo gene synthesis by an antiviral reverse transcriptase. *Science (New York, N.Y.)*, *386*(6717), eadq0876. https://doi.org/10.1126/science.adq0876
53. Thompson, K. M., & Gottesman, S. (2014). The MiaA tRNA modification enzyme is necessary for robust RpoS expression in Escherichia coli. *Journal of bacteriology*, *196*(4), 754–761. https://doi.org/10.1128/JB.01013-13
54. Trifonov, E. N. (2000). Consensus temporal order of amino acids and evolution of the triplet code. *Gene*, *261*(1), 139–151. https://doi.org/10.1016/s0378-1119(00)00476-5
55. Trifonov, E. N. (2004). The triplet code from first principles. *Journal of biomolecular structure & dynamics*, *22*(1), 1–11. https://doi.org/10.1080/07391102.2004.10506975
56. Vetsigian, K., Woese, C., & Goldenfeld, N. (2006). Collective evolution and the genetic code. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(28), 10696–10701. https://doi.org/10.1073/pnas.0603780103

57. Walton, T., Pazienza, L., & Szostak, J. W. (2019). Template-Directed Catalysis of a Multistep Reaction Pathway for Nonenzymatic RNA Primer Extension. *Biochemistry*, *58*(6), 755–762. https://doi.org/10.1021/acs.biochem.8b01156
58. Wellen, K. E., Hatzivassiliou, G., Sachdeva, U. M., Bui, T. V., Cross, J. R., & Thompson, C. B. (2009). ATP-citrate lyase links cellular metabolism to histone acetylation. *Science (New York, N.Y.)*, *324*(5930), 1076–1080. https://doi.org/10.1126/science.1164097
59. White, H. B. (1976). Coenzymes as fossils of an earlier metabolic state. *Journal of molecular evolution*, *7*(2), 101–104. https://doi.org/10.1007/BF01732468
60. Whiteley, A. T., Eaglesham, J. B., de Oliveira Mann, C. C., Morehouse, B. R., Lowey, B., Nieminen, E. A., Danilchanka, O., King, D. S., Lee, A. S. Y., Mekalanos, J. J., & Kranzusch, P. J. (2019). Bacterial cGAS-like enzymes synthesize diverse nucleotide signals. *Nature*, *567*(7747), 194–199. https://doi.org/10.1038/s41586-019-0953-5
61. Winkler, W. C., Nahvi, A., Roth, A., Collins, J. A., & Breaker, R. R. (2004). Control of gene expression by a natural metabolite-responsive ribozyme. *Nature*, *428*(6980), 281–286. https://doi.org/10.1038/nature02362
62. Wochner, A., Attwater, J., Coulson, A., & Holliger, P. (2011). Ribozyme-catalyzed transcription of an active ribozyme. *Science (New York, N.Y.)*, *332*(6026), 209–212. https://doi.org/10.1126/science.1200752
63. Wu, D., Hu, D., Chen, H., Shi, G., Fetahu, I. S., Wu, F., Rabidou, K., Fang, R., Tan, L., Xu, S., Liu, H., Argueta, C., Zhang, L., Mao, F., Yan, G., Chen, J., Dong, Z., Lv, R., Xu, Y., Wang, M., … Shi, Y. G. (2018). Glucose-regulated phosphorylation of TET2 by AMPK reveals a pathway linking diabetes to cancer. *Nature*, *559*(7715), 637–641. https://doi.org/10.1038/s41586-018-0350-5
64. Xu, W., Yang, H., Liu, Y., Yang, Y., Wang, P., Kim, S. H., Ito, S., Yang, C., Wang, P., Xiao, M. T., Liu, L. X., Jiang, W. Q., Liu, J., Zhang, J. Y., Wang, B., Frye, S., Zhang, Y., Xu, Y. H., Lei, Q. Y., Guan, K. L., … Xiong, Y. (2011). Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α-ketoglutarate-dependent dioxygenases. *Cancer cell*, *19*(1), 17–30. https://doi.org/10.1016/j.ccr.2010.12.014
65. Yang, W., Lee, J. Y., & Nowotny, M. (2006). Making and breaking nucleic acids: two-Mg2+-ion catalysis and substrate specificity. *Molecular cell*, *22*(1), 5–13. https://doi.org/10.1016/j.molcel.2006.03.013
66. Yarus, M., Widmann, J. J., & Knight, R. (2009). RNA-amino acid binding: a stereochemical era for the genetic code. *Journal of molecular evolution*, *69*(5), 406–429. https://doi.org/10.1007/s00239-009-9270-1
67. Zeraati, M., Langley, D. B., Schofield, P., Moye, A. L., Rouet, R., Hughes, W. E., Bryan, T. M., Dinger, M. E., & Christ, D. (2018). I-motif DNA structures are formed in the nuclei of human cells. *Nature chemistry*, *10*(6), 631–637. https://doi.org/10.1038/s41557-018-0046-3
68. Zheng, G., Dahl, J. A., Niu, Y., Fedorcsak, P., Huang, C. M., Li, C. J., Vågbø, C. B., Shi, Y., Wang, W. L., Song, S. H., Lu, Z., Bosmans, R. P. G., Dai, Q., Hao, Y. J., Yang, X., Zhao, W. M., Tong, W. M., Wang, X. J., Bogdan, F., … He, C. (2013). ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Molecular cell*, *49*(1), 18–29. https://doi.org/10.1016/j.molcel.2012.10.015
69. Zhou, J., Wan, J., Gao, X., Zhang, X., Jaffrey, S. R., & Qian, S. B. (2015). Dynamic m^6A mRNA methylation directs translational control of heat shock response. *Nature*, *526*(7574), 591–594. https://doi.org/10.1038/nature15377

The amino-acid-chronology, catalytic-residue, and cross-domain-phylostratum results (AA-1/AA-2/AA-3) and the RNA-writes-protein and protein-writer-takeover capacity analyses (OB-1/OB-2) referenced in §3.1 are prior deposited project computations and syntheses (`COMPUTATIONAL_ANALYSIS_WRITER_AMINOACID_COMPOSITION_CATALYTIC_CHRONOLOGY`, `COMPUTATIONAL_ANALYSIS_WRITER_DOMAIN_CROSSDOMAIN_PHYLOSTRATUM`, `SYNTHESIS_ORIGIN_OF_LIFE_OB1/OB2`), available on request; their external anchors are two-step PubMed-verified in those documents.

Accompanying machine-readable output (filled tables, all metrics) and reproducible source are provided in the reproducibility package, deposited at Zenodo (https://doi.org/10.5281/zenodo.20966659).