APPLICATION NOTE · BIOLOGY

Information Looking for Its Owner in DNA Repair

Nucleotide excision repair as a search for an owner, a control that spends energy without making a record, and a first look at public Damage-seq data

Erkan Bayar

Version 1.0 · 28 October 2026
DOI: 10.5281/zenodo.23725581

Uses the method of Information Looking for Its Owner, doi:10.5281/zenodo.23724281. Does not test Time-Gel Theory.

Abstract

Ultraviolet light writes a mark into DNA: two neighbouring pyrimidines fuse into a cyclobutane pyrimidine dimer (CPD). The mark is information without an owner. Somewhere among billions of intact base pairs, a set of proteins has to find it, confirm it, cut it out and rewrite the gap. Aziz Sancar and colleagues worked out how that machine operates. This note borrows a way of asking questions, not a result, from a protocol written for a different system, the freezing of supercooled water, and applies it to excision repair.

Three things transfer. The first is a separation of five roles: source, carrier, processor, state builder and meaning-giver. The second is the idea that what matters is the irreversible act, not the state it leaves behind. The third is a control that spends the same energy with the same machinery but produces no record. In repair, that control is the full excision system working on undamaged DNA. One thing does not transfer: the clock effect that Time-Gel Theory attaches to irreversible events is, by its own numbers, about 10^{-11} s per cell, far below any clock that could sit inside a cell.

The note proposes a biochemical experiment that measures what it costs, in ATP, for information to find its correct owner, and whether paying less produces more wrong owners. Conventional kinetic proofreading, proposed for this very step by Reardon and Sancar, predicts the same outcomes; the note claims no discrimination. It then takes a first look at a public CPD Damage-seq time course in human fibroblasts. Thirty-six hours after UV, about 17% of the genome sits in 100 kb bins whose damage has not gone down at all. These bins form blocks with a mean length of about 560 kb and are concentrated on chromosomes 21, X, 13 and 18. Nothing here is new physics. Every number is either taken from public data, computed from it, or an order-of-magnitude estimate, and each claim is labelled.

01 // How to read this document

Every claim carries one label. The labels are those of the parent protocol, with one addition for the data analysis in Section 09.

[MEASURED]	Established by experiment in published biology or physics. Where the source was not reopened for this note, it says so.
[OBSERVED]	Computed by this note from public data. Not peer-reviewed and not independently repeated.
[MODEL]	Calculated from equations, definitions or stated assumptions.
[HYPOTHESIS]	A claim about repair that the experiment in this note can test. None of them is a test of Time-Gel.
[DECISION]	A choice made by the author, recorded before any data.
[OPEN]	A point not settled.
[INTERPRETATION]	A reading offered by this note. Not yet a hypothesis.
[ANALOGY]	An image. Not a physical claim.

A borrowed question is not borrowed evidence.

02 // What transfers, and what does not

The parent protocol designs a clock test around freezing water. Its value for biology lies in the structure of the thinking and in the logic of its controls, not in its clock prediction.

From the ice clock	Counterpart in excision repair	Transfers?
Five roles (source, carrier, processor, state builder, meaning-giver)	Lesion, DNA strand, recognition proteins, dual incision, cell response	Yes
Search time and the first irreversible step	Finding the lesion; the two incisions	Yes
Control B2: same entropy, no record	Full repair system burning ATP on undamaged DNA	Yes. It is the core of the experiment
Clock slowing during the event	Clock shift inside a repairing cell	No. About 10^{-11} s, unmeasurable (Appendix A)
“Clock”	Circadian clock, which changes repair capacity over the day	No. Same word, different thing: a chemical oscillator, not a local rate of time

03 // The chain of thought

1. **Spark.** A UV photon is absorbed by DNA and two adjacent pyrimidines become a CPD [MEASURED]. The mark carries information: *something happened here*. Nothing yet reads it [INTERPRETATION].
2. **Search.** In global genome repair, the recognition factor XPC (with RAD23B) has to find the lesion among intact DNA. Recognition factors bind and leave many times before a complex commits [MEASURED, FROM MEMORY]. The length of the search shows up as the time a lesion survives.
3. **The owner appears.** Once TFIIH has verified the lesion with its XPB and XPD subunits, and XPA and RPA are in place, the complex is committed [MEASURED, FROM MEMORY]. The owner here is not born from the information, as the ice nucleus is born from the water. It is a separate machine that arrives.
4. **Birth moment.** XPF–ERCC1 cuts on the 5′ side and XPG on the 3′ side; an oligonucleotide of about 24–32 nucleotides leaves the genome [MEASURED]. That double cut is the first step that cannot be undone. The gap is then filled and sealed.
5. **Two routes.** When an RNA polymerase stalls at the lesion, the owner is already standing on the information. Transcription-coupled repair needs no search. Global genome repair does [MEASURED]. One

lesion type, two arms, “search” and “no search”, provided by the cell itself.

6. **The clock.** Time-Gel ties a clock effect to irreversible events. In repair, that effect is far too small to see. What remains is the question the clock was meant to answer, asked of energy instead: what does it cost for information to find its correct owner?

04 // Who is the owner

In freezing water the source and the state builder are the same molecules. In repair, the five roles come apart cleanly: the source arrives from outside as light, and the owner is an independent machine [INTERPRETATION].

Role	In excision repair	Reversible?	Measurable counterpart
Source	UV photon and the CPD it makes	The lesion does not revert on its own; photolyase reverses it in organisms that have it	Damage maps (Damage-seq), lesion counts
Carrier	The DNA strand, which carries the lesion to the next reading	—	Position of the lesion in the genome
Processor	XPC binding and release; verification by TFIIH	Yes: a protein bound at the wrong place leaves	Single-molecule dwell times, ATPase rate
State builder	Dual incision and repair synthesis	No: the first step that does not come back	The excised oligonucleotide (excision assay, XR-seq)
Meaning-giver	Cell response: checkpoints, restart of transcription	—	Partly: response markers can be measured, “meaning” cannot [OPEN]

The ice protocol reads the coincidence of source and state builder as an example of “the recorder is the recorded”. Repair is the opposite case. Comparing the two is the reason to have both in one framework [INTERPRETATION].

05 // The decision this note rests on

The parent protocol records one decision [DECISION]: the clock looks at the moment the state changes, not at the state that results. Its definition is $\rho_I \equiv \sigma/k_B$, the local entropy production rate. Translated into repair, the decision says: what counts is the act of repairing, not the repaired DNA, and the cost of the act is dissipation. In a biochemical system that cost can be measured directly, as ATP hydrolysed and, if resolution allows, as heat.

$$\sigma \approx (r_{ATP} \cdot |\Delta G_{ATP}|) / T \tag{1}$$

Here r_{ATP} is the hydrolysis rate per volume and ΔG_{ATP} the free energy per hydrolysis under the reaction’s own ATP/ADP/P_i concentrations [MODEL]. Equation (1) is the quantity the B arm below must match.

06 // The experiment: the price of finding the owner

Question. How much energy does excision repair spend to find the correct owner for a lesion, and does spending less produce more wrong owners, that is, incisions in undamaged DNA?

System. Human excision nuclease reconstituted from purified factors (XPC–RAD23B, TFIIH, XPA, RPA, XPF–ERCC1, XPG) and the bacterial UvrABC system. Substrate: duplex DNA carrying a single site-specific CPD, and the same sequence without it. The excision assay itself is a long-standing method of the Sancar laboratory [MEASURED]. What is new is the second quantity recorded in the same tube: energy.

Measured. Rate of correct excision (oligonucleotides released from damaged DNA); rate of wrong excision (oligonucleotides released from undamaged DNA); ATP hydrolysis rate; heat by isothermal titration calorimetry where sensitivity allows.

Arm	What happens	What it controls
A · Correct owner	Full system, damaged substrate	Test arm: excision and ATP rates
B · Spending without a record (counterpart of B2)	Full system, undamaged substrate	How much the same proteins dissipate without forming a record, and how often they cut anyway
C · Energy dial	ATP/ADP ratio stepped at fixed total nucleotide, so $\lvert\Delta G_{ATP}\rvert$ decreases step by step	The curve of accuracy against dissipation
D · Verification off	ATPase-dead XPB or XPD variants	Which step spends the energy
E · Reversal instead of excision	Photolyase, which reverses the CPD with light and leaves no gap	Cost of repair that leaves no trace
F · Search in the cell	XR-seq time course after a UV pulse in normal cells, in XPC-deficient cells (transcription-coupled only) and CSB-deficient cells (global only)	Search time when the owner is already present versus when it must look

Birth moment. In the ice clock nature chooses the moment; in the cell arm the experimenter does, by delivering the UV pulse. Pulse times are therefore set by a random schedule, and circadian phase is recorded, because repair capacity changes over the day [MEASURED] [8]. That rhythm is a confounder here, not a signal.

Pre-registration. As in the parent protocol: analysis plan, thresholds and exclusions fixed and timestamped before data; the entropy estimate for arm B frozen before unblinding; results published whatever their sign.

07 // Predictions and the conventional rival

In this experiment Time-Gel and conventional biophysics predict the same things. The note claims no discrimination. Its contribution is the design of the controls.

Kinetic proofreading (Hopfield 1974, Ninio 1975) [MEASURED] as theory: an enzyme can make fewer errors than equilibrium binding allows by adding an irreversible, energy-consuming verification step. Each such step can lower the error rate by at most one equilibrium discrimination factor, and it costs energy. For CPD repair specifically, Reardon and Sancar found that none of the six factors alone discriminates a CPD efficiently from undamaged DNA, and proposed that RPA, XPA and XPC locate the lesion cooperatively while TFIIH supplies kinetic proofreading [MEASURED] [7]. Arm C asks for the energy–accuracy curve of exactly that proposal.

The Time-Gel reading [INTERPRETATION]: a clock that looks at the moment of change says the price is paid during search and verification, not in the repaired DNA. That is what kinetic proofreading already says.

Arm	Kinetic proofreading	Time-Gel (rate reading)	Discriminates?
C · Less energy	Wrong-incision rate rises	Same	No
B · Undamaged DNA	Futile ATP consumption	Same: dissipation without a record	No
D · ATPase dead	Verification and incision stop	Same	No
E · Photolyase	Different energy path, no excision	Same	No
Clock shift in the cell	None	About 10^{-11} s	In principle; not in practice

08 // What would count as what

These conditions test the repair hypothesis of this note [HYPOTHESIS], not Time-Gel.

- If the arm C curve is flat (accuracy does not change while $|\Delta G_{ATP}|$ changes), CPD selection by excision repair is not an energy-limited proofreading process in this reconstituted system.
- If arm B shows no measurable ATP consumption, the “spending without a record” arm cannot be built, and the B2 analogy fails for this system.
- If accuracy improves with energy beyond the bound kinetic proofreading allows for the number of verification steps, either the model is incomplete or the measurement is wrong. This does not count in favour of Time-Gel; the explanation is sought in biochemistry first.

Inconclusive: if the share of ATP going to verification cannot be separated from the share going to DNA unwinding, the energy–accuracy curve cannot be interpreted.

09 // A first look at public data

Before any new experiment, existing data can show where information waits longest for its owner. This section reanalyses a public CPD Damage-seq time course. It describes the data; it tests nothing.

09.1 Data

Normal human skin fibroblasts (NHF1), CPD damage mapped by HS-Damage-seq at 0 h, 1 h, 8 h, 24 h, 36 h and 48 h after UV, two replicates each, plus two unirradiated and two input libraries; GEO samples GSM2585685–GSM2585698, GSM2585719 and GSM2585720 [5, 6]. The author downloaded the reads (about 2 GB) and counted them in 100 kb bins of the human genome. Sample labels were taken from the GEO file names. Strand information is not kept in this summary.

09.2 Method [MODEL]

Each bin’s damage count is divided by its input count, which removes copy-number and mappability differences. For each time point, the bin’s share of all damage is divided by its share at 0 h. If the total damage remaining at time t is a fraction F of the starting damage, a bin that has not repaired at all has a ratio of exactly 1/F. All such bins pile up at the same ceiling, and partially repaired bins fall below it. The ceiling therefore gives an upper bound on F without needing absolute damage counts. Bins with too little input (below 30% of the autosomal median, 15% for the single-copy X) or fewer than 50 reads at 0 h were excluded, leaving 28,359 bins. ChrY was excluded.

09.3 Results [OBSERVED]

- **Replicates agree.** Spearman correlation between replicate pairs, per bin: 0.94 (0 h), 0.88 (1 h), 0.96 (8 h), 0.97 (24 h), 0.99 (36 h). The second 48 h replicate has only 1.3 million reads and correlates at 0.46; 48 h results are therefore read with caution.
- **A ceiling appears at 36 h.** The distribution of $\log_2$ ratios develops a sharp upper peak at +0.96 (Fig. 1). 4,751 bins (16.8%) lie within ±0.15 of it. The peak is present in each 36 h replicate separately (+0.96 and +0.84).
- **Total remaining damage.** The ceiling implies that at most about 51% of the initial CPDs remain at 36 h. It is an upper bound, because bins at the ceiling may have repaired a little. The 48 h estimate (56%) is higher than at 36 h, which points to the low-depth replicate rather than to new damage.
- **Repair is organised in blocks.** Neighbouring bins behave alike (lag-1 autocorrelation of the ratio 0.63–0.77 across time points; 0.01 after shuffling). The unrepaired bins form 845 runs with a mean length of 5.6 bins (about 560 kb); 78% of them lie in runs of at least 500 kb; the longest run is 5.6 Mb.
- **Chromosomes differ.** The unrepaired share is 5–7% on chr22, chr17, chr19 and chr1, and 23–29% on chr18, chr13, chrX and chr21 (Fig. 2).

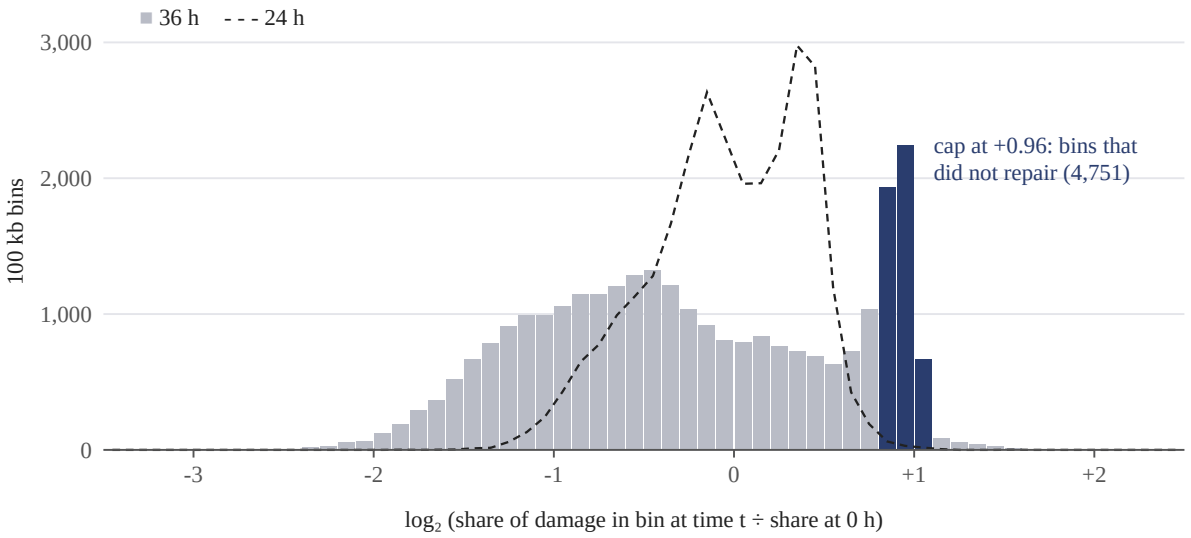


Figure 1. Distribution, over 28,359 bins of 100 kb, of the change in each bin's share of CPD damage relative to 0 h. Grey bars: 36 h; dashed line: 24 h. At 36 h a ceiling forms at $\log_2$ ratio +0.96 (navy): bins whose damage has not gone down. Its position bounds the remaining global damage at $\leq 2^{-0.96} \approx 0.51$. Data: GEO GSM2585685–2585720 [5, 6]; analysis by this note.

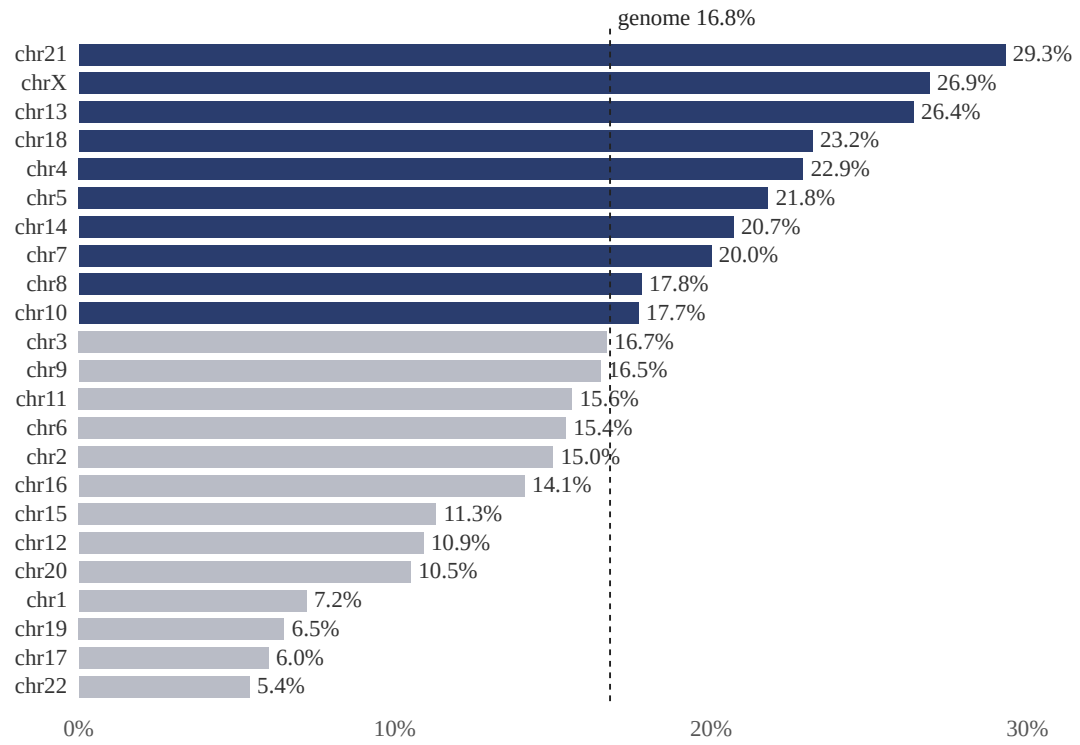


Figure 2. Share of each chromosome's 100 kb bins that sit at the 36 h ceiling, i.e. show no repair. Navy: above the genome-wide 16.8%. ChrX is counted with a halved input threshold because its input coverage is about half the autosomal level, as expected for a single X.

09.4 What this says, and what it does not

In the language of this note: part of the damage has not found its owner after 36 hours, and that part is not scattered at random but sits in blocks of hundreds of kilobases [INTERPRETATION]. Conventional biology says the same in its own terms: repair is slower in closed chromatin and in regions without transcription, and the chromosomes with the most unrepaired bins here are those generally known to be gene-poor [MEASURED, GENERAL KNOWLEDGE]. Gene density, chromatin state and replication timing were not measured in this analysis, because the annotation tracks could not be downloaded in the environment where it was done [OPEN]. The original authors report that repair depends on chromatin state, transcription and transcription-factor binding [5]. Without strand information, transcription-coupled and global repair cannot be separated here. This analysis does not test Time-Gel.

Where the owner is late, the information waits; the data show where, not why.

10 // Feasibility

Arms A–E can be run with present methods in a laboratory that already reconstitutes excision repair; arm F needs XR-seq. The hardest part is not excision but energy accounting: ATP hydrolysis in a mixture of six factors includes unwinding, translocation and futile cycling, and separating the verification share needs the ATPase-dead variants of arm D and, ideally, single-molecule data. Section 09 can be extended at once with public annotation tracks (gene density, chromatin states, lamina-associated domains, replication timing) to replace general knowledge with measurement.

11 // Interpretation: two kinds of owner

This section is not tested by anything above. It records why the note was written.

- **Born owner and arriving owner** [INTERPRETATION]. In ice the owner is born from the information itself; in repair it arrives from outside. Transcription-coupled repair lies between: the owner is already passing by. A framework that holds all three is more useful than one built for a single case.
- **The cost is paid in the act** [INTERPRETATION]. Under the rate reading, a finished repair leaves no ongoing cost, as standing ice produces no entropy. The trace remains in the chain of evidence: the excised oligonucleotide, the repair patch.
- **Reversal leaves no trace** [ANALOGY]. Photolyase undoes the mark rather than replacing it. In the language of records it is erasure, not rewriting.

12 // Open problems

1. How to apportion ATP between unwinding, translocation, futile cycling and verification.
2. Whether arm B can match arm A's dissipation closely enough to be a true "same entropy, no record" control.
3. Gene density, chromatin and replication-timing annotation of the unrepaired blocks of Section 09.
4. A strand-resolved repeat of Section 09 to separate the two repair routes.
5. Any measure for the meaning-giver role.

Appendix A. Calculations

- **Clock shift in a repairing cell** [MODEL], all inputs assumed: about 10^5 lesions per cell after UV, 10–100 ATP per lesion, about 20 k_BT per ATP. Entropy produced $\lesssim 2 \times 10^8$ k_B; cell volume about 2×10^{-15} m^3; $N/V \approx 10^{23}$ m^{-3}. Phase step $N/(2V\rho_{crit}) \approx 5 \times 10^{-11}$ s at $\rho_{crit} = 10^{33}$ s^{-1} m^{-3}.
- **Ceiling bound:** a bin with no repair has ratio 1/F; a bin with a fraction r of its damage left has r/F < 1/F. The upper mode m of $\log_2$ ratios gives $F \le 2^{-m}$: m = 0.96 → F ≤ 0.51 (36 h, pooled); 0.96 and 0.84 for the two 36 h replicates.
- **Block length:** 4,751 ceiling bins in 845 runs → 5.62 bins per run on average, about 562 kb.

Appendix B. Provenance

The question, the five-role frame and the rate-reading decision come from the author's protocol for the ice clock [1]. The author downloaded the Damage-seq reads, summarised them into 100 kb bins and supplied the file names. Claude (Anthropic) assigned sample labels from those names, wrote and ran the analysis (Python; scripts available from the author), drafted the experiment and this English text, and checked the bibliographic details of references [5] and [7] by web search. Statements marked "from memory" were not re-checked against primary sources for this version. Decisions are the author's. AI agreement is not independent validation, and this note does not represent the views of Aziz Sancar or his laboratory.

References

[1] E. Bayar, *Information Looking for Its Owner: A supercooled-water clock test that separates two readings of Time-Gel*, v1.0, Zenodo (2026). doi:10.5281/zenodo.23724281.

[2] E. Bayar, *Time-Gel Theory: Time as the Fundamental Medium — A Falsifiable Framework*, v12.1, Zenodo (2026). doi:10.5281/zenodo.23106516.

[3] J. J. Hopfield, Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity, *Proc. Natl. Acad. Sci. USA* 71, 4135 (1974).

[4] J. Ninio, Kinetic amplification of enzyme discrimination, *Biochimie* 57, 587 (1975).

[5] J. Hu, O. Adebali, S. Adar, A. Sancar, Dynamic maps of UV damage formation and repair for the human genome, *Proc. Natl. Acad. Sci. USA* 114, 6758 (2017). doi:10.1073/pnas.1706522114.

[6] NCBI Gene Expression Omnibus, series GSE98025 (series assignment inferred from the study; sample numbers taken from the downloaded files), samples GSM2585685–GSM2585698, GSM2585719, GSM2585720.

[7] J. T. Reardon, A. Sancar, Recognition and repair of the cyclobutane thymine dimer, a major cause of skin cancers, by the human excision nuclease, *Genes Dev.* 17, 2539 (2003).

[8] T.-H. Kang, J. T. Reardon, M. Kemp, A. Sancar, Circadian oscillation of nucleotide excision repair in mammalian brain, *Proc. Natl. Acad. Sci. USA* 106, 2864 (2009).

[9] J. Hu, S. Adar, C. P. Selby, J. D. Lieb, A. Sancar, Genome-wide analysis of human global and transcription-coupled excision repair of UV damage at single-nucleotide resolution, *Genes Dev.* 29, 948 (2015).

[10] A. Sancar, Mechanisms of DNA repair by photolyase and excision nuclease, Nobel Lecture (2015); *Angew. Chem. Int. Ed.* 55, 8502 (2016).