cGAS-STING V4 Whitespace & Trajectory Analysis

Companion to: cGAS-STING V4 KC Portfolio Analysis (native `kc/v1`, gpt-5.5, family-scoped) **Date:** 2026-07-22 **Substrate:** 126 native on-target families → **101 on-pathway** (25 excluded as off-pathway co-mentions) **Basis:** structured-IP reading layer (native KC + compiled Markush genus + scope-neighbors) projected onto a literature+Wikidata-grounded scope ontology **Framing:** This is **candidate generation and a whitespace/whitespace-trajectory signal — NOT a freedom-to-operate (FTO) determination.** Genus membership/overlap is 3-valued (IN/OUT/INDETERMINATE) and recall is a **lower bound**; native KC is **not calibrated** (Gold Set pending). "Whitespace" here means *absence of strong structural claim coverage*, which is necessary but not sufficient for opportunity — it must be read against the reachable biology (the ontology denominator).

---

0. Why this report exists

The V4 KC portfolio report grades *what exists*. It does not tell an inventor *where the defensible room is* or *where the field is going*. Whitespace is the structured **complement** of claim coverage; it only exists relative to a coordinate system. This companion adds the three layers the KC report lacked:

1. a **scope-coordinate system** (mechanism-node × direction × modality × indication) — the whitespace denominator, grounded externally, not derived from the patent set;
2. a **coverage-density read + thin-cell detector** integrating genus breadth, granted-active status, and native claim-strength; and
3. the **dynamic layer** — lapse-driven decay, combination adjacencies, cross-class transfer, and a trajectory vector.

Everything is exported as a **knowledge graph** (`cgas_sting_knowledge_graph.json`, 304 nodes / 657 edges) so the whitespace question is answerable as a graph query.

---

1. Headline findings

1. **Only 101 of 126 families genuinely claim a cGAS-STING mechanism.** An LLM claim-scope read reclassified **25 families as off-pathway** — they co-mention the pathway but claim something else (SIRPα/CD47 antibodies, LAG-3/PD-1 combos, ATM/Chk1/DNA-PK DDR inhibitors, CAR constructs, A2A modulators, radiotherapy scheduling). Treating raw hit counts as coverage overstates the estate by ~25%.

2. **The field has inverted direction** (on real filing dates). Pre-2018 filings were **8:1 agonist:inhibitor** (STING-agonist immuno-oncology). 2022-2026 filings are **39 inhibitor : 9 agonist** — a decisive pivot to cGAS/STING **inhibition for**

autoimmune/inflammatory disease. (All 101 on-pathway families carry a real filing or publication year; no proxy dates.)

3. **The first-wave STING-agonist foundation is being actively abandoned.** Five foundational CDN-pocket patents are **LAPSED or DISCONTINUED for non-payment** — 4 STING *agonist* (Aduro US-9724408; GSK US-10047115, US-10450341, US-9994607) plus 1 CDN-pocket *inhibitor* (Aduro US-9549944) — including GSK US-9994607 discontinued 2026-07-20, two days before this run. The CDN-pocket agonist cell is decaying to ~10 live grants; a maintained Aduro/InvivoGen/ImmuneSensor core survives to 2033-2038.

4. **Four reachable mechanism cells are genuinely OPEN** (zero on-pathway families): **STING degrader (PROTAC)**, **STING trafficking/palmitoylation inhibitor**, **cGAS DNA-binding-surface inhibitor**, and **cGAS degrader**. Caveat: the `degrader` *direction* is unpopulated across the **entire** 126-family corpus, so those two degrader cells are "open" partly by corpus/axis construction (§7); the trafficking-inhibitor and DNA-binding cells are the more defensible open calls.

5. **The cGAS-catalytic small-molecule inhibitor cell is the crowded, well-defended core** (18 families, 9 live grants, median genus recall 0.326 across n=12 genus families) — the *least* whitespace, dominated by ImmuneSensor, Boehringer Ingelheim, and 4 named scaffold lanes (quinoline, indazole, pyridine/pyrrolidine, other-heterocycle). New entrants here face a genus thicket, though recall ~0.33 means each lane still has analog headroom.

6. **ENPP1 inhibition is genus-broad but grant-immature** (5 families, median genus recall **0.582** — the *highest* of any cell — but only 1 live grant, 4 pending). The scaffolds are well-drafted; the *grants* have not consolidated. Contestable now, closing soon — and watch the TXinno (Korea) same-applicant continuation cluster.

---

2. The scope-coordinate system (ontology / whitespace denominator)

Four axes, human-gated and **externally grounded** (V4 §2.3 — the denominator cannot be derived from the patent set). Full definitions in `data/ontology.json`.

- **mechanism_node** (12 values): where in the axis the claim acts — `cGAS_catalytic`, `cGAS_DNA_binding`, `cGAMP_ligand`, `STING_CDN_pocket`, `STING_nonCDN_allosteric`, `STING_trafficking`, `STING_degradation`, `ENPP1_hydrolase`, `downstream_TBK1_IRF3`, `upstream_DNA_sensing`, `pathway_delivery`, `pathway_biomarker_dx`.
- **direction**: agonist / antagonist-inhibitor / degrader / modulator / diagnostic.
- **modality_chemotype** (12 values): quinoline / indazole / pyridine-pyrrolidine / other-heterocycle small molecule; CDN; non-CDN SM agonist; ADC-conjugate; nanoparticle-polymer; peptide; biologic-cell-bacteria; gene/nucleic-acid; dx-signature.

- **indication**: IO-oncology / autoinflammatory / neuroinflammation / fibrosis / cardiometabolic / infectious-vaccine / aging-other.

Grounding (Wikidata SPARQL, confirmed): proteins cGAS (Q22676633 / CGAS / UniProt Q8N884), STING (Q21112990 / STING1-TMEM173 / Q86WV6), TBK1 (Q21111280), IRF3 (Q5973468); **gene→disease genetic associations (P2293)** — STING1→**SAVI**, TBK1→**FTD/ALS**, ENPP1→**arterial calcification of infancy / OPLL**; pathway items Q45317067, Q99303932. These anchor the indication axis and the KG disease nodes. Mechanism/direction/modality axes are literature-grounded (cGAS-STING review corpus: Chen, Ablasser, Vance, Hopfner).

---

3. Coverage-density grid + thin-cell whitespace

3.1 Mechanism-node × direction (on-pathway families; `whitespace_grid_node_dir.json`)

`live` = granted **and currently maintained** (lapsed patents propagated out per audit M3); `laps` = granted-then-lapsed; `medRcl(N)` = median `family_best` genus recall over N genus-carrying families in the cell (recall is a **lower bound**).

mechanism_node	direction	n	live	laps	pend	genus fams	medRcl (N)	verdict
cGAS_catalytic	inhibitor	18	9	0	9	17	0.326 (12)	**COVERED**
STING_CDN_pocket	agonist	14	10	4	0	11	0.365 (11)	**COVERED / decaying**
STING_nonCDN_allosteric	inhibitor	11	3	0	8	9	0.528 (7)	PARTIAL→covered
STING_trafficking	agonist	7	3	0	4	0	–	PARTIAL (no genus)
STING_nonCDN_allosteric	agonist	6	2	0	4	5	0.536 (4)	PARTIAL→covered
upstream_DNA_sensing	inhibitor	6	0	0	6	0	–	PARTIAL (all pending)
pathway_delivery	modulator	6	2	0	4	1	0.5 (1)	PARTIAL
pathway_delivery	agonist	6	1	0	5	1	–	PARTIAL
ENPP1_hydrolase	inhibitor	5	1	0	4	5	**0.582 (5)**	genus-broad, grant-immature
pathway_biomarker_dx	inhibitor	5	0	0	5	1	0.648 (1)	PARTIAL
STING_nonCDN_allost	modulat	3	0	0	3	3	0.049	THIN

mechanism_node	direction	n	live	laps	pend	genus fams	medRcl (N)	verdict
eric	or						(3)	
cGAMP_ligand	modulator	2	1	0	1	1	0.744 (1)	THIN
ENPP1_hydrolase	agonist	2	0	0	2	0	–	THIN
STING_CDN_pocket	inhibitor	1	0	1	0	0	–	THIN — near-open **& lapsed**
cGAMP_ligand	agonist	1	1	0	0	1	0.947 (1)	THIN
cGAS_DNA_binding	modulator	1	1	0	0	0	–	THIN

(A `best KC3` column was dropped: native KC is uncalibrated — Gold Set 0/200-300 — and a high KC3 next to a thin cell invites over-reading. KC is in the KC report, not used to grade whitespace here.)

3.2 Thin-cell verdicts (`whitespace_thin_cells.json`)

Verdict rule: OPEN = 0 on-pathway families; THIN = ≤2 families and no broad live-granted genus; PARTIAL = coverage present but no broad live-granted genus (recall ≥0.30 over ≥1 genus family); COVERED = broad live-granted genus present.

cell	verdict	evidence
STING_trafficking / inhibitor	**OPEN**	0 families; palmitoylation/trafficking inhibitors for autoimmune — mechanistically distinct from pocket binders
cGAS_DNA_binding / inhibitor	**OPEN**	0 families; alternative to the crowded catalytic-site genus
STING_degradation / degrader	OPEN*	0 families; *but `degrader` is unpopulated corpus-wide (axis/corpus artifact — treat as hypothesis prompt)
cGAS_catalytic / degrader	OPEN*	0 families; same corpus-wide-`degrader` caveat
downstream_TBK1_IRF3 / inhibitor	OPEN	0 families (largely off this target's core)
STING_CDN_pocket / **inhibitor**	**THIN — near-open & lapsed**	1 family (Aduro US-9549944) and it is **itself lapsed** → CDN-pocket antagonist direction for autoimmune is effectively unclaimed *and* undefended
cGAMP_ligand / agonist	COVERED (n=1, recall	single family but a very broad genus

cell	verdict	evidence
	0.947)	
ENPP1_hydrolase / inhibitor	genus-broad, grant-immature	5 families, recall 0.582, only 1 live grant
STING_nonCDN_allosteric / inhibitor	COVERED	11 families, recall 0.528; 8 pending — actively consolidating
STING_CDN_pocket / agonist	COVERED / **decaying**	14 families, 10 live + 4 lapsed; recall 0.365
cGAS_catalytic / inhibitor	**COVERED**	18 families, 9 live grants, recall 0.326 — crowded, 4 scaffold lanes

> **Read.** The genuinely open/near-open cells cluster on the **STING trafficking-inhibitor, cGAS DNA-binding-inhibitor, and CDN-pocket-antagonist** axes — the autoimmune-directed *mechanistic alternatives* the field is pivoting toward but has not fenced. The cGAS-catalytic inhibitor and non-CDN cells are already genus-covered; the CDN-pocket agonist cell is covered-but-decaying. (The corrected recall data upgraded several cells from "thin" to "covered" — the earlier draft under-counted genus coverage due to a recall-field bug caught in audit.)

---

4. The actionable core — 58 genus-bearing families (56 on-pathway)

The user-flagged genus small-molecule cohort reconciles as: **58 has-genus families across the 126**; the structural layer profiles the **56 on-pathway** genus families (2 off-pathway genus families excluded); of those, **36 are strictly small-molecule *inhibitor/agonist* chemotypes** (`sm_*`) — the remainder are CDN, non-CDN small-molecule agonist, or delivery-polymer genera. Median `family_best` genus recall across the 36 SM = **0.41** (n=29 with a recall value; a lower bound), i.e. compiled Markush claims capture ~40% of the enumerable analog space they gesture at — leaving material *in-scaffold* room even inside "covered" cells. (Recall convention standardized to `family_best.recall` after audit; the earlier `genera[].recall` read was null for ~90% of families and produced spurious single-point medians.)

Structural sub-signals (from `genus_layer.json` + `scope_neighbors.json`):

- **cGAS-catalytic inhibitor scaffolds are diversifying, not converging.** Four distinct chemotype lanes (quinoline — ImmuneSensor; indazole — Carna/BI; pyridine-pyrrolidine — BI/Janssen; other-heterocycle — Ventus/Lilly/GSK/Cornell). No single genus dominates; median recall 0.326 (n=12) means each lane still has analog headroom.
- **The TXinno (Korea) ENPP1 families show a same-applicant scaffold cluster.** US-20240116882 ↔ US-20240140944 have high ECFP similarity (0.955) but the Markush overlap is **entirely INDETERMINATE** (`in_trips=0`; 48/48 candidate trips) and the families are *different named chemotypes* — consistent with same-

applicant continuation fencing, not confirmed identical scope. ENPP1 *grant* consolidation has not happened (1 of 5 live-granted) despite the high genus recall (0.582).

- **The Aduro non-CDN agonist genus (US-10738056) is broad** (recall 0.629, enumeration estimate ~1e9) and **maintained** — the strongest surviving structural claim on the agonist side. Caveat: its genus is `markush_status: parse_failed`, so the 0.629 is from a partial compilation.
- **CDN-pocket genera are moderate-to-narrow** (GSK recall 0.286-0.318; cell median 0.365) and **partly lapsing** — the decaying foundation is structurally moderate, not deep.

---

5. Dynamic layer

5.1 Lapse-driven decay timeline (`legal_decay.json`)

patent	assignee	node / direction	event	date	est. expiry
US-10047115	GSK	CDN-pocket **agonist**	LAPSED (non-payment)	2022-10-11	2036
US-10450341	GSK	CDN-pocket **agonist**	LAPSED (non-payment)	2023-12-19	2035
US-9549944	Aduro	CDN-pocket **inhibitor**	LAPSED (non-payment)	2025-03-25	2034
US-9724408	Aduro/UC	CDN-pocket **agonist**	LAPSED (non-payment)	2025-10-07	2034
US-9994607	GSK	CDN-pocket **agonist**	DISCONTINUATION	2026-07-20	2037

Maintained core (survives): Aduro US-9770467 (→2033), US-9695212 (→2033); InvivoGen US-10011630 (→2035), US-11730817 (→2038); ImmuneSensor US-10519188 (→2038); Aduro non-CDN US-10738056 (→2038).

> **Combinatorial whitespace-as-lapse.** Four of the five lapses are in the **CDN-pocket agonist** cell (GSK ×3, Aduro ×1) — first-generation *systemic* CDN chemistry being abandoned. The fifth (Aduro US-9549944) is the *only* CDN-pocket **inhibitor** family and it too has lapsed — so the CDN-pocket antagonist direction is now both unclaimed and undefended. The agonist cell is not fully opened (10 live grants remain) but the *original broad CDN genus* is no longer defended; freedom expands specifically for **non-systemic / delivery-formulated / non-CDN** agonist chemistry designing around the maintained core. The cGAS-**inhibitor** core, by contrast, is entirely fresh (2021-2026 filings) and accreting.

5.2 Combination adjacencies (from the 25 off-pathway co-claims)

cGAS-STING is being **co-claimed** with: checkpoint (anti-LAG-3 + PD-1 — BMS; generic checkpoint combos — Flagship), **innate-checkpoint** (SIRPα ×2, CD47 — Arch Oncology), **DDR inhibitors** (ATM — Duke; Chk1 — Sierra; DNA-PK — U Washington), **cytokine** (IL-18 — Yale; APRIL-TACI — Dana-Farber), **cell therapy** (CAR — USC/UPenn), and **A2A** (Ryvu). These are the empirically-revealed combination axes. Combination whitespace = biologically-rational pairs *not* yet co-claimed — notably **cGAS/STING-inhibitor × DDR** for autoimmune (the DDR co-claims are all agonist/oncology-directed) and **ENPP1-inhibitor × checkpoint**.

5.3 Cross-class transfer candidates (53 cross-class structural edges)

The knowledge graph surfaces **53 cross-class edges** — genus-overlap + sequence-identity edges to families *outside* the cGAS-STING substrate (of 242 raw scope-neighbor edges, capped per family in the KG) — scaffolds whose Markush genus reaches into cGAS-STING chemical space but were filed for other purposes (e.g. old heterocyclic-chemistry patents US-5360914, US-4504470, US-6252041; kinase/nucleotide families US-11427610). These are **transfer hypotheses**: chemistry proven elsewhere, structurally adjacent to a cGAS-STING node, not yet claimed *for* the pathway. (All genus overlap is INDETERMINATE-driven — candidate generation, lower-bound, not FTO.)

5.4 Trajectory vector — "what is the trajectory of this target?"

Composed from filing velocity × direction × modality × entrant type by era, on **real filing/publication dates** (no proxy; all 101 on-pathway families dated) — `dynamic_signals.json`:

era	dominant modality	direction (ago:inh)	entrants
pre-2018	CDN (9)	**8 : 1**	company (Aduro/GSK/Portnoy), 3 academic
2018-2021	CDN (7) + first SM (9) + non-CDN agonist (5) + biologic (5)	21 : 8	company-led; first China (2) + Korea (1) entrants
2022-2026	**small-molecule heterocycle (33 + 12 named-scaffold)**	**9 : 39**	37 company, 12 academic, 4 China-origin, 3 Korea-origin

Trajectory statement: *cGAS-STING is transitioning from a lapsing first-wave systemic-CDN **STING-agonist / IO** foundation toward (i) a crowded, well-drafted, multi-scaffold small-molecule **cGAS-catalytic-inhibitor** race for **autoimmune** indications (the current center of gravity, 2021→2024), and (ii) an emerging **ENPP1-inhibitor** and **non-CDN STING-inhibitor** frontier (non-CDN inhibitor filings peak 2025) — with genuine whitespace concentrating in the **STING-trafficking-inhibitor, cGAS-DNA-binding-inhibitor, and CDN-pocket-antagonist** cells. Entrants broaden from a US-pharma/academic core to include China- and Korea-origin filers, though the count is modest (China 2→4, Korea 1→3) and driven mostly by*

the ENPP1 and non-CDN-inhibitor cells. Node-emergence timing (real dates): cGAS-catalytic inhibitors appear 2021 and surge 2023-2024; non-CDN STING inhibitors build through 2025; biomarker/dx and delivery emerge 2022+.

---

6. Knowledge graph

`data/cgas_sting_knowledge_graph.json` — **304 nodes, 657 edges**, generic nodes/edges JSON (Cytoscape / networkx / Neo4j-loadable).

- **Node types:** target_pathway, mechanism_node, direction, modality, indication, **disease** (Wikidata-grounded), assignee, patent_family, year.
- **Edge types:** `acts_on`, `directed_as`, `uses_modality`, `targets_indication`, `assigned_to`, `part_of_pathway`, `co_claims` (off-pathway adjacency), `lapses_on` (family→year), `genus_overlaps` (structural, incl. 53 cross-class), `sequence_identity`, `node_associates_disease` (Wikidata P2293).

The 5 `node_associates_disease` edges each source from a real `node:` id (verified: `node:STING_CDN_pocket`→SAVI, `node:downstream_TBK1_IRF3`→FTD/ALS, `node:ENPP1_hydrolase`→GACI/OPLL/PPK), so the disease-grounding subgraph is queryable (an earlier kwarg-collision bug that dangled these edges was fixed in audit).

The whitespace query becomes a graph query: *mechanism_node reachable in the biology (has a `node_associates_disease` edge or ontology entry) with no incident `acts_on` edge from a live-granted, broad-genus family* → returns the OPEN cells of §3.2. **The trajectory query:** family nodes grouped by real filing-year, projected on `directed_as` / `uses_modality` → the §5.4 vector. **The lapse-whitespace query:** cells whose `acts_on` families carry `lapses_on` edges → the decaying CDN-pocket cells.

---

7. Limits & honest framing

- **Not FTO, not calibrated.** Whitespace = absence of *strong structural claim coverage*, a necessary-not-sufficient condition for opportunity. Genus recall is a lower bound; native KC is uncalibrated (Gold Set 0/200-300) and is deliberately **not** used to grade whitespace here. An OPEN cell is a *hypothesis-generation prompt*, not clearance.
- **`degrader` is unpopulated corpus-wide.** No family in any of the 126 (on- or off-pathway) claims a `degrader`. So the two "OPEN" degrader cells (STING-degrader, cGAS-degrader) are open partly because the axis value is empty everywhere — this is the case where absence may reflect ontology-axis choice + a 126-family corpus rather than a searched-and-empty gap. Treat them as the weakest of the open calls; the trafficking-inhibitor / DNA-binding / CDN-pocket-antagonist cells are better-grounded (they sit next to *populated* neighboring cells).
- **Denominator is operator-defined.** The reachable space is set by the ontology axes (literature + Wikidata), not the patent corpus. Different axes → different whitespace. The axes are auditable in `ontology.json`.

- **Classification is a model read.** The 101/25 on/off-pathway split and the per-family bins come from an LLM claim-scope read; each family carries `scope_evidence`. Systematic error is possible; the report emphasizes cell-level patterns over single-family calls.
- **Corpus-bounded.** Whitespace is relative to the 126 ingested families. Four 2026 PGPUBs failed to ingest (BQ snapshot lag, per the ingest ledger) and are not represented; genuinely un-filed art is invisible by construction.
- **Two misresolved foundational anchors** surfaced (US-10166291 = Caltech targeted-nanoparticle, not a CDN foundation — it remains in the on-pathway *delivery* set and one of the pre-2018 trajectory rows, but does not change any cell-level conclusion; consistent with the KC report's re-anchoring recommendation).

---

8. Actionable summary for an inventor

priority	where	why
1	**STING trafficking / palmitoylation inhibitor** — autoimmune	OPEN cell, 0 families; mechanistically distinct from pocket binders; sits next to a populated (non-CDN inhibitor) neighbor, so a real gap not an axis artifact
2	**CDN-pocket STING *antagonist*** — autoimmune	THIN & lapsed: the sole family (Aduro US-9549944) has lapsed → direction unclaimed *and* undefended
3	**cGAS DNA-binding-surface inhibitor**	OPEN; design-around the crowded catalytic-site genus (median recall 0.326)
4	**ENPP1 inhibitor** — IO	Genus-broad (recall 0.582) but only 1 of 5 live-granted; contestable *now*, before grants consolidate; watch the TXinno (Korea) same-applicant cluster
5	**Non-systemic / non-CDN STING agonist** (delivery-formulated)	Foundation partly lapsing (4 agonist patents); design around the maintained Aduro/InvivoGen/ImmuneSensor core (→2033-2038)
6 *(speculative)*	**STING or cGAS degrader (PROTAC)**	OPEN but `degrader` is unpopulated corpus-wide — treat as a hypothesis prompt, verify with a dedicated degrader-corpus search before relying on it
— avoid	cGAS-catalytic occupancy inhibitor (generic scaffold)	COVERED, crowded, 4 defended scaffold lanes

---

Artifacts: `data/ontology.json`, `families_scoped_full.json` (126, with on/off-pathway + 4-axis bins + scope_evidence), `genus_layer.json`, `scope_neighbors.json`,

legal_decay.json, *whitespace_grid_node_dir.json*, *whitespace_grid_node_ind.json*, *whitespace_thin_cells.json*, *dynamic_signals.json*, *cgas_sting_knowledge_graph.json*. *Scripts: bin_scope_grid.py*, *whitespace_analysis.py*, *dynamic_signals.py*, *build_kg.py*.