CLDN18.2 IP Portfolio Analysis

V3 KC Scoring Workflow (RAG-augmented, SS-F/OS-S firewalled, prompt-cached per-sub-dimension model routing)

Executive Summary

This report applies the rDNA.ai four-stage KC scoring pipeline to 266 recent USPTO biopharma patents in the CLDN18.2 (Claudin-18.2 / Claudin 18 isoform 2) target class. CLDN18.2 is a tight-junction protein expressed on gastric, gastroesophageal-junction (GEJ), and pancreatic tumors. The therapeutic landscape spans naked monoclonal antibodies (zolbetuximab — Astellas, approved 2024 for CLDN18.2+ gastric / GEJ adenocarcinoma), antibody-drug conjugates (AstraZeneca / Daiichi Sankyo + Chinese-originator ADC platforms: Innovent, Hengrui, LaNova LM-302, RemeGen, ASKB589), bispecific T-cell engagers (CLDN18.2 × CD3 — Amgen AMG910, givastomig, ASP2138; CLDN18.2 × 4-1BB costimulatory bispecifics), and CAR-T cell therapies (CT041 / satricabtagene autoleucel — CARsgen, Phase III; ZL-1211; TST001). Each patent is scored on six Key Component (KC) dimensions on a 1–5 scale and aggregated into a four-scheme ensemble (Meta Baseline / AbbVie / Amgen / Genentech weighting), which determines the patent's investment-grade tier. This is the V3 corrected pipeline (HER3/TL1A/TROP2/GLP-1 backtest canonical pattern): KC2/KC3/KC4/KC6 composite prompts embed the V2 holistic-judgment guidance verbatim with a KC4 prior-art landscape anchor specific to the CLDN18.2 modality landscape (NAKED-mAb most crowded around Ganymed/IMAB362 priority chain; ADC moderately crowded with rapid Chinese-originator expansion; bispecific / CAR-T moderately open). No character truncation is applied to the patent text submitted to the model. TA-bucket multipliers are not applied to the patent-level ranking (display groupings only). Sub-dimensions D3 (Prosecution Flexibility) and A4 (Commercial Differentiation) are forced to Opus 4.7 to avoid the Sonnet rubric-collapse failure mode identified in HER2 corrected QA.

Portfolio-level results are summarized below by therapeutic-area cut, by top assignee, and at the patent level for the calibrated top decile after a V3 pairwise-tournament re-judging pass. The CLDN18.2 corpus has high Chinese-originator share by design (Innovent, Hengrui, Junshi, LaNova, RemeGen, CARsgen, BeiGene, Akeso, Shanghai Henlius, Sino Biopharma); the substrate-lock memo reports the China-originator count + share for narrative context.

This is a research-pipeline application of the V3 KC scoring rubric (RAG-augmented, SS-F / OS-S firewalled, per-patent prompt-cached, per-sub-dimension model-routed). The Sonnet 4.5 Quarantine firewall screened all 266 candidate patents and flagged 0 for quarantine. USPTO biopharma issued and published patents from 2008-01-01 to 2026-05-31.

Key Component (KC) Scoring Framework

Each patent is scored on six Key Component (KC) dimensions on a 1–5 scale. The ensemble score is the mean of the six dimensions (with confidence and standard deviation tracked alongside).

- **KC1** measures abstract and therapeutic specificity: target specificity, therapeutic positioning, structural disclosure, and technical differentiation.
- **KC2** measures independent-claim quality: scope, structure, enforceability, and enablement vulnerability.
- **KC3** measures dependent-claim quality: fallback hierarchy, embodiment coverage, prosecution flexibility, and structural diversity.
- **KC4** measures novelty and differentiation: prior-art positioning, inventive-step quality, and field positioning.
- **KC5** measures specification enablement: enablement gap, structure-function insight, data breadth, and translational completeness.

- **KC6** measures working examples: working/prophetic example quality, structural diversity, experimental rigor, and in vivo / translational validation.

Retrieval-Augmented Generation (RAG) for KC4 Novelty Scoring

The V3 rubric introduces a corpus-internal retrieval-augmented generation (RAG) step that supplies every KC4 sub-dimension (N1 Prior-Art Landscape, N2 Claims Differentiation, N3 Inventive-Step Quality, N4 Field Positioning) with a list of the nearest prior patents in the local Biopharma corpus. The neighbour list is passed alongside the cached full patent text so the scoring LLM judges novelty against actual nearby filings rather than its internal world knowledge alone.

RAG construction:

- Embedding model: text-embedding-3-small (dim=1536).
- Each Biopharma full-text specification was chunked into ≤ 7,000-char blocks across the abstract, claims, and description sections, then each chunk was embedded; a patent-level centroid was computed and L2-normalised.
- Index size: 266 Biopharma specifications → 16,002 chunks → 266 patent centroids.
- Retrieval: per target patent, return top-20 nearest patent centroids by cosine similarity (self excluded); the result list is passed to the KC4 sub-dim prompts as `rag_neighbors`. Sub-dims must echo back `rag_neighbors_cited` in their JSON response.

RAG results across the corpus:

- Average nearest-neighbour cosine similarity across all scored patents: 0.961.
- Average similarity of the top-20 neighbourhood: 0.916 (a denser neighbourhood pulls KC4 N1/N2 scores down toward 'crowded'; sparser pushes them up toward 'pioneering').
- Corpus average composite KC4 score (post-RAG): 2.46 on the 1–5 scale, covering sub-dims N1, N2, N3, N4.
- The corpus has high internal similarity (mean nearest-neighbour cosine ≥ 0.70), indicating a relatively concentrated CLDN18.2 prior-art landscape; KC4 N1/N2 scoring therefore skews toward 'crowded' for many patents.
- The KC4 sub-dim JSON responses include the `rag_neighbors_cited` field so that each novelty score can be traced back to specific neighbour patents from the corpus index.

Most-frequently-cited prior-art neighbours in the corpus:

Rank	Neighbour Patent	Times Cited in Other Patents' Top-20
1	US20220411492A1	83
2	US20200385463A1	70
3	US20210380681A1	64
4	US20220306765A1	64
5	US20250011419A1	63
6	US20250136678A1	63
7	US20210171628A1	62
8	US20260007750A1	57
9	US20230235047A1	56
10	US20220372133A1	53
11	US20230085834A1	52
12	US20260104418A1	51
13	US20210380680A1	50

14	US20220396616A1	48
15	US12415853B2	48

Portfolio-Level Metrics

• Total CLDN18.2 USPTO records ingested: 504 (USPTO PPubs full-text fetch via searchWithBeFamily, publication date 2008-01-01 to 2026-05-31; 16 CLDN18.2 / CLDN18-2 / Claudin-18.2 / Claudin 18 isoform 2 / anti-CLDN18.2 / CLDN18.2-ADC / CLDN18.2-bispecific / CLDN18.2-CAR-T / zolbetuximab / IMAB362 / claudiximab / Vyloy / CT041 / satricabtagene / satri-cel / ASKB589 / AB011 / LM-302 / SHR-A1904 / TST001 / ZL-1211 / AMG910 / givastomig / ASP2138 / osemitamab discovery queries)

• Title- and full-text-based classification: Biopharma=266, Diagnostic=22, Device=1, Non_Biopharma=215

• Total Biopharma patents scored: 266

• Average KC scores — KC1: 4.30 KC2: 3.81 KC3: 3.54 KC4: 2.46 KC5: 4.24 KC6: 3.49

• Average ensemble mean across portfolio: 3.580

Investment-grade distribution:

• TIER_2_STRONG: 11

• TIER_3_WATCHLIST: 191

• LOW_PRIORITY: 64

Top Assignees

Rank	Assignee	Patent Count
1	LANOVA MEDICINES LIMITED COMPANY	11
2	ALLOGENE THERAPEUTICS, INC	6
3	HARBOUR BIOMED CO., LTD	5
4	AMGEN INC	5
5	GRACELL BIOTECHNOLOGIES CO., LTD	5
6	NOVAROCK BIOTHERAPEUTICS, LTD	5
7	THE REGENTS OF THE UNIVERSITY OF CALIFORNIA	4
8	ASTELLAS PHARMA EUROPE BV; XENCOR, INC	4
9	ASKGENE PHARMA, INC	4
10	SHANDONG BOAN BIOTECHNOLOGY CO., LTD	4
11	GENSUN BIOPHARMA INC	4
12	I-MAB BIOPHARMA US LIMITED; ABL BIO INC	4
13	SOTIO BIOTECH A.S	4
14	JIANGSU HENGRUI MEDICINE CO., LTD.; SHANGHAI HENGRUI PHARMACEUTICAL CO., LTD	4
15	SPARX BIOSCIENCE LIMITED	4

Therapeutic Area Analysis (n ≥ 5 patents)

For each area with at least 5 patents the canonical report shows average KC1–KC6 (1–5 scale), average ensemble mean, average TA-weighted score, and tier distribution. TA buckets are display groupings only in this CLDN18.2 V3 run. The five CLDN18.2 TA buckets (CLDN18_Naked_mAb, CLDN18_ADC, CLDN18_Bispecific_T_cell_engager, CLDN18_CAR_T, CLDN18_Other_modality) carry no numerical multipliers (TA_MULTIPLIERS_APPLIED = False). ta_weighted_score in the CSV outputs is identical to ensemble_mean. The five-bucket structure anchors on the four primary CLDN18.2 modality classes plus a catch-all for radio-immunoconjugates, NK-cell engagers, and vaccine-immunotherapy filings.

Therapeutic Area	N	KC 1	KC 2	KC 3	KC 4	KC 5	KC 6	Ens Mean	TA-W	T1 Premium	T1 High	T 2	T3 / Low
CLDN18_ADC	86	4.31	3.86	3.73	2.43	4.29	3.50	3.622	3.622	0	6	64	16
CLDN18_Bispecific_T_cell_engager	75	4.31	3.79	3.44	2.41	4.23	3.61	3.558	3.558	0	1	57	17
CLDN18_Naked_mAb	55	4.33	3.87	3.53	2.22	4.45	3.45	3.560	3.560	0	0	40	15
CLDN18_CAR_T	50	4.26	3.70	3.36	2.84	3.92	3.32	3.566	3.566	0	4	30	16

Therapeutic Area Narratives

Each TA narrative below summarises the calibrated cohort: average ensemble mean, the strongest and weakest KC dimensions for the TA, TIER_1_PREMIUM count, and the top three patents by TA-weighted score. Every KC value shown is the rolled-up output of the V3 structured-citation pipeline — each underlying patent has six KC composite scores, each grounded in four per-sub-dimension JSON citations carrying the source model, the model-supplied reason, and a verbatim quote from the patent text. A complete structured-citation example is presented in the Top-Decile section below.

CLDN18_ADC (n=86)

Average ensemble mean: 3.622. Highest KC: KC1 (avg 4.31). Lowest KC: KC4 (avg 2.43). TIER_1_PREMIUM: 0. Top 3 patents (by TA-weighted score): US20220306765A1; US20220227832A1; US11541127B2.

CLDN18_Bispecific_T_cell_engager (n=75)

Average ensemble mean: 3.558. Highest KC: KC1 (avg 4.31). Lowest KC: KC4 (avg 2.41). TIER_1_PREMIUM: 0. Top 3 patents (by TA-weighted score): US20220331416A1; US12378313B2; US20240130340A1.

CLDN18_Naked_mAb (n=55)

Average ensemble mean: 3.560. Highest KC: KC5 (avg 4.45). Lowest KC: KC4 (avg 2.22). TIER_1_PREMIUM: 0. Top 3 patents (by TA-weighted score): US11395852B2; US20220184209A1; US20200055932A1.

CLDN18_CAR_T (n=50)

Average ensemble mean: 3.566. Highest KC: KC1 (avg 4.26). Lowest KC: KC4 (avg 2.84). TIER_1_PREMIUM: 0. Top 3 patents (by TA-weighted score): US20200291090A1; US20240342215A1; US20240041926A1.

Top-Decile Calibration: Pairwise Tournament Results

The top decile (n=27) of the CLDN18.2 biopharma cohort was re-judged against the V3 rubric using full-text specifications and a complete pairwise tournament. The V3 calibration replaced any prior uniform top-decile plateau with a discriminating ranking; calibrated ensemble distribution: mean 4.069, range 3.915–4.287.

Rank	Patent	Assignee	KC1	KC2	KC3	KC4	KC5	KC6	Cal Ens
1	US20200291090A1	ALLOGENE THERAPEUTICS, INC	5	4	4	4	4	4	4.225
2	US20220227832A1	ALLOGENE THERAPEUTICS, INC. (+1)	5	4	4	4	4	4	4.225
3	US11541127B2	ASTELLAS PHARMA, INC. (+1)	5	4	4	3	5	5	4.162
4	US20240342215A1	REGENERON PHARMACEUTICALS, INC	5	4	5	4	3	3	4.200
5	US20240041926A1	SHENZHEN FRONTIERGATE BIOTECHNOLOGY C... (+1)	5	4	3	4	5	4	4.178
6	US20260007747A1	VIVASOR, INC	5	4	3	4	5	4	4.178
7	US20250235547A1	UBE CORPORATION	5	4	5	3	4	4	4.137
8	US20220306765A1	SUZHOU TRANSCENTA THERAPEUTICS CO., LTD	5	5	5	2	5	5	4.287
9	US20230270877A1	SEAGEN INC	5	4	5	3	4	4	4.137
10	US20220331416A1	BIOTHEUS CO., LTD	4	4	4	4	5	4	4.103
11	US11395852B2	ASTELLAS PHARMA INC. (+1)	5	4	4	3	5	4	4.090
12	US12016923B2	TRIUMVIRA IMMUNOLOGICS USA, INC	5	4	4	3	5	4	4.090
13	US20220184209A1	ASTELLAS PHARMA INC. (+1)	5	4	4	3	5	4	4.090
14	US12378313B2	XENCOR, INC	5	4	4	3	5	3	4.018
15	US20230149465A1	CELLEDIT LLC	5	4	4	3	4	4	3.987
16	US20260104418A1	THE REGENTS OF THE UNIVERSITY OF CALI...	5	5	5	2	5	3	4.143
17	US20240043527A1	BIONTECH SE	5	4	4	3	4	4	3.987
18	US11485782B2	BEIJING XUANYI PHARMASCIENCES CO., LTD	5	5	4	2	5	4	4.065
19	US20250134962A1	SHANGHAI TTM-BIO TECHNOLOGY CO., LTD	5	4	3	3	5	4	3.940
20	US20230338565A1	SICHUAN KELUN-BIOTECH BIOPHARMACEUTIC...	5	4	4	2	5	5	3.925
21	US20200055932A1	AMGEN RESEARCH GMBH (+1)	5	4	3	3	5	4	3.940
22	US20240130340A1	BIOCYTOGEN	5	4	4	3	4	4	3.987

		PHARMACEUTICALS CO., LTD							
23	US20230085834A1	THE TRUSTEES OF THE UNIVERSITY OF PEN... (+1)	5	4	4	3	4	3	3.915
24	US20210009686A1	BEIJING XUANYI PHARMASCIENCES CO., LTD	5	5	4	2	5	4	4.065
25	US20240226321A1	ANJARIUM BIOSCIENCES AG	4	4	4	4	4	3	3.928
26	US20250340631A1	HARBOUR BIOMED CO., LTD	5	4	4	2	5	5	3.925
27	US20200390887A1	ASTELLAS PHARMA INC. (+1)	5	4	4	2	5	5	3.925

V3 Structured-Output Example: #1 US20200291090A1 — ALLOGENE THERAPEUTICS, INC

Every row in the table above is backed by a V3 structured-output citation: each patent has six KC composite scores, each grounded in four per-sub-dimension JSON records carrying the source model, the model-supplied reason, and a verbatim quoted snippet of the patent text (with KC4 sub-dims additionally citing the RAG prior-art neighbour list). For brevity the per-KC tables for the remaining top-decile patents are not reproduced in this report; the complete per-sub-dimension records are persisted in Microprotein_V3_KC_Subscores.csv and the per-patent nested-JSON citation payload is persisted in the kc_score_justification_json column of Microprotein_Top15_Predictions.csv. The single example presented below — the #1 top-decile patent (US20200291090A1) — illustrates the structured-output format for each of the six KC dimensions.

CONSTITUTIVELY ACTIVE CHIMERIC CYTOKINE RECEPTORS

KC	Score	Reason	Verbatim Quote from the Patent
KC1	5	Patent provides extensive structural disclosure including full amino acid sequences with SEQ ID NOs, detailed domain compositions, specific amino acid substitutions (H499L, S505N, W515K, G509N), and complete IUPAC-style nomenclature for all CACCR variants.	In some embodiments, the CACCR comprises the amino acid sequence of SEQ ID NO: 90 or 119, with or without a signal sequence...TPOR/MPLR(478-582; H499L, S505N, W515K)
KC2	4	All four sub-dims at 4: diverse claim types (CACCR, polynucleotide, vector, cell, method of treatment, kit), strong structural anchoring via SEQ IDs and specific mutations, broad comprising genus with species hierarchy, and robust exemplification mitigating Amgen risk. Not comprehensive enough for 5 given functional r…	CACCRs improved the in vivo anti-tumor activity and persistence of BCMA CAR-T cells against orthotopic multiple myeloma
KC3	4	D2(5) and D4(4) drive strong commercial/structural coverage via TM mutations, multiple cytotails, CAR targets, cell types, and treatment methods. D1(4) and D3(4) reflect 57 multi-level dependents with robust fallbacks, just short of dense thicket.	pharmaceutical composition comprising the engineered immune cells...kit comprising the engineered immune cells...method of treating a cancer in a subject, comprising administering...therapeutically effective amount
KC4	4	N2/N3/N4 all at 4 with tunable	constitutively active chimeric cytokine

		CACCR design, documented unexpected Tscm expansion, durable in vivo tumor control, and IL-15-mimetic signaling beyond IL-7Ra C7R prior art; moderately open CAR-T enhancement space (N1=3) supports composite of 4.	receptors (CACCRs). When present on chimeric antigen receptor (CAR)-bearing immune cells, such CACCRs allow for increased immune cell activation, proliferation, persistence [RAG anchor: US20230265147A1, US20230085834A1, US20230272040A1…]
KC5	4	Extensive in vivo data in NSG mice with orthotopic MM1.S and subcutaneous LN229-EGFRvIII models, demonstrating tumor regression, delayed relapse, and CAR-T persistence at clinical-relevant doses (1-3×10^6 cells).	8-10 week old female NSG mice were irradiated with 1 Gy one day prior to intravenous inoculation of 5×10.sup.6 MM1.S-Luc-GFP. 14 days after tumor implantation, mice were randomized based on tumor burden, and dosed intravenously with either 1×10.sup.6 or 3×10.sup.6 of the indicated CAR-T cells
KC6	4	12 working examples with systematic TM/cytotail diversity (W1=5,W2=4), rigorous controls/statistics across multiple donors (W3=4), and dual in vivo models (BCMA orthotopic MM + EGFRvIII glioblastoma) with dose-response (W4=4). Falls short of 5 due to no PK/safety program.	Example 1-12 provide comprehensive working examples including HEK293T reporter assays, primary human CAR-T cell generation, STAT5 activation studies, cytotoxicity assays, growth factor-independent persistence assays, and in vivo xenograft models in NSG mice.

KC4 rag_neighbors_cited (top prior-art neighbours from the corpus that informed the N1–N4 novelty scoring): US20230265147A1, US20230085834A1, US20230272040A1, US20210332334A1, US20210340219A1, US20250213694A1, US20220127570A1, US20230134345A1, US20220325241A1, US20220193133A1…

Top-Decile Assignee Ranking

Top-decile assignees are ranked by frequency in the calibrated top-decile cohort, then by best calibrated patent rank, then by average calibrated ensemble mean.

Rank	Assignee	Count	Best Patent Rank	Avg Cal Rank	Avg Cal Ens
1	ALLOGENE THERAPEUTICS, INC	1	1	1.0	4.225
2	ALLOGENE THERAPEUTICS, INC. (+1)	1	2	2.0	4.225
3	ASTELLAS PHARMA, INC. (+1)	1	3	3.0	4.162
4	REGENERON PHARMACEUTICALS, INC	1	4	4.0	4.2
5	SHENZHEN FRONTIERGATE BIOTECHNOLOGY CO., LTD. (+1)	1	5	5.0	4.178
6	VIVASOR, INC	1	6	6.0	4.178
7	UBE CORPORATION	1	7	7.0	4.137
8	SUZHOU TRANSCENTA THERAPEUTICS CO.,	1	8	8.0	4.287

	LTD				
9	SEAGEN INC	1	9	9.0	4.137
10	BIOTHEUS CO., LTD	1	10	10.0	4.103
11	ASTELLAS PHARMA INC. (+1)	1	11	11.0	4.09
12	TRIUMVIRA IMMUNOLOGICS USA, INC	1	12	12.0	4.09
13	ASTELLAS PHARMA INC. (+1)	1	13	13.0	4.09
14	XENCOR, INC	1	14	14.0	4.018
15	CELLEDIT LLC	1	15	15.0	3.987
16	THE REGENTS OF THE UNIVERSITY OF CALIFORNIA	1	16	16.0	4.143
17	BIONTECH SE	1	17	17.0	3.987
18	BEIJING XUANYI PHARMASCIENCES CO., LTD	2	18	21.0	4.065
19	SHANGHAI TTM-BIO TECHNOLOGY CO., LTD	1	19	19.0	3.94
20	SICHUAN KELUN-BIOTECH BIOPHARMACEUTICAL CO., LTD	1	20	20.0	3.925
21	AMGEN RESEARCH GMBH (+1)	1	21	21.0	3.94
22	BIOCYTOGEN PHARMACEUTICALS CO., LTD	1	22	22.0	3.987
23	THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA (+1)	1	23	23.0	3.915
24	ANJARIUM BIOSCIENCES AG	1	25	25.0	3.928
25	HARBOUR BIOMED CO., LTD	1	26	26.0	3.925
26	ASTELLAS PHARMA INC. (+1)	1	27	27.0	3.925

Bottom-Decile Assignee Ranking

Bottom-decile assignees are ranked by frequency in the bottom CLDN18.2 biopharma decile, then by weaker average original ensemble score.

Rank	Assignee	Count	Avg Original Score	Weakest Score
1	BIONTECH SE (+2)	1	2.837	2.837
2	SHANGHAI LONGYAO BIOTECHNOLOGY INC., LTD	1	2.865	2.865
3	SHIJIAZHUANG YILING PHARMACEUTICAL CO.,	3	2.945	2.902

	LTD			
4	NANJING KAEDI BIOTHERAPEUTICS LTD	2	3.058	2.94
5	SANYOU BIOPHARMACEUTICALS CO., LTD	1	2.987	2.987
6	SYSTIMMUNE, INC. (+1)	1	3.0	3.0
7	GUANGZHOU BIO-GENE TECHNOLOGY CO., LTD	1	3.018	3.018
8	FUSION PHARMACEUTICALS INC	1	3.018	3.018
9	L-MAB BIOPHARMA US LIMITED (+1)	1	3.018	3.018
10	BEIJING IMMUNOAH PHARMATECH CO., LTD	1	3.018	3.018
11	LYELL IMMUNOPHARMA, INC	1	3.03	3.03
12	ORUM THERAPEUTICS, INC	1	3.052	3.052
13	GENSUN BIOPHARMA INC	2	3.15	3.075
14	TRON— TRANSLATIONALE ONKOLOGIE AN DER UNIVERSITATSMED... (+2)	1	3.075	3.075
15	SUZHOU ZELGEN BIOPHARMACEUTICALS CO. LTD	1	3.075	3.075
16	ORGENESIS INC	1	3.08	3.08
17	SHANGHAI YAKE BIOTECHNOLOGY LTD	1	3.09	3.09
18	CSPC MEGALITH BIOPHARMACEUTICAL CO., LTD	1	3.115	3.115
19	SPARX BIOSCIENCE LIMITED	1	3.135	3.135
20	BIOSION INC	1	3.14	3.14
21	SUZHOU IMMUNOFOCO BIOTECHNOLOGY CO., LTD	1	3.163	3.163
22	GRACELL BIOTECHNOLOGIES CO., LTD	1	3.225	3.225
23	ALLOGENE THERAPEUTICS, INC	1	3.225	3.225

Top 15 Tournament Calibration CLDN18.2 Patents

Each top-15 patent below is presented with: (a) the patent number and confirmed assignee, (b) the patent title, (c) the calibrated KC profile and pairwise tournament record, and (d) the analyst narrative. Every calibrated KC value shown is the V3 composite of four per-sub-dimension scores, each backed by a structured-citation record (score,

source model, model-supplied reason, verbatim quote, and — for KC4 — rag_neighbors_cited). The complete per-sub-dimension records are persisted in the V3_KC_Subscores.csv and the Top15_Predictions.csv (kc_score_justification_json column). A worked example of the full structured-output format for one patent is shown in the Top-Decile section above.

#1 US20200291090A1 — ALLOGENE THERAPEUTICS, INC
CONSTITUTIVELY ACTIVE CHIMERIC CYTOKINE RECEPTORS

Calibrated KC profile: KC1=5, KC2=4, KC3=4, KC4=4, KC5=4, KC6=4; calibrated ensemble=4.225; pairwise record=26-0.

V3 pairwise tournament: 26W / 0L over top-decile (n=27).

#2 US20220227832A1 — ALLOGENE THERAPEUTICS, INC. (+1)
PROTEASE-ACTIVATING CD45-GATE CAR

Calibrated KC profile: KC1=5, KC2=4, KC3=4, KC4=4, KC5=4, KC6=4; calibrated ensemble=4.225; pairwise record=25-1.

V3 pairwise tournament: 25W / 1L over top-decile (n=27).

#3 US11541127B2 — ASTELLAS PHARMA, INC. (+1)
Drug conjugates comprising antibodies against claudin 18.2

Calibrated KC profile: KC1=5, KC2=4, KC3=4, KC4=3, KC5=5, KC6=5; calibrated ensemble=4.162; pairwise record=24-2.

V3 pairwise tournament: 24W / 2L over top-decile (n=27).

#4 US20240342215A1 — REGENERON PHARMACEUTICALS, INC
ENGINEERED T CELL RECEPTORS FUSED TO BINDING DOMAINS FROM ANTIBODIES

Calibrated KC profile: KC1=5, KC2=4, KC3=5, KC4=4, KC5=3, KC6=3; calibrated ensemble=4.200; pairwise record=23-3.

V3 pairwise tournament: 23W / 3L over top-decile (n=27).

#5 US20240041926A1 — SHENZHEN FRONTIERGATE BIOTECHNOLOGY CO., LTD. (+1)
DENDRITIC CELL ACTIVATING CHIMERIC ANTIGEN RECEPTORS AND USES THEREOF

Calibrated KC profile: KC1=5, KC2=4, KC3=3, KC4=4, KC5=5, KC6=4; calibrated ensemble=4.178; pairwise record=22-4.

V3 pairwise tournament: 22W / 4L over top-decile (n=27).

#6 US20260007747A1 — VIVASOR, INC
Engineered PD-L1- Targeting Gamma Delta T Cell Receptors

Calibrated KC profile: KC1=5, KC2=4, KC3=3, KC4=4, KC5=5, KC6=4; calibrated ensemble=4.178; pairwise record=21-5.

V3 pairwise tournament: 21W / 5L over top-decile (n=27).

#7 US20250235547A1 — UBE CORPORATION
ANTIBODY-DRUG CONJUGATE

Calibrated KC profile: KC1=5, KC2=4, KC3=5, KC4=3, KC5=4, KC6=4; calibrated ensemble=4.137; pairwise record=18-8.

V3 pairwise tournament: 18W / 8L over top-decile (n=27).

#8 US20220306765A1 — SUZHOU TRANSCENTA THERAPEUTICS CO., LTD
NOVEL ANTI- CLDN18.2 ANTIBODIES

Calibrated KC profile: KC1=5, KC2=5, KC3=5, KC4=2, KC5=5, KC6=5; calibrated ensemble=4.287; pairwise record=17-9.

V3 pairwise tournament: 17W / 9L over top-decile (n=27).

#9 US20230270877A1 — SEAGEN INC
Combination Therapy

Calibrated KC profile: KC1=5, KC2=4, KC3=5, KC4=3, KC5=4, KC6=4; calibrated ensemble=4.137; pairwise record=17-9.

V3 pairwise tournament: 17W / 9L over top-decile (n=27).

#10 US20220331416A1 — BIOTHEUS CO., LTD
COMBINED EXPRESSION OF A CHIMERIC CD3 FUSION PROTEIN AND AN ANTI-CD3-BASED BISPECIFIC T CELL ACTIVATING ELEMENT

Calibrated KC profile: KC1=4, KC2=4, KC3=4, KC4=4, KC5=5, KC6=4; calibrated ensemble=4.103; pairwise record=17-9.

V3 pairwise tournament: 17W / 9L over top-decile (n=27).

#11 US11395852B2 — ASTELLAS PHARMA INC. (+1)
Therapy involving antibodies against Claudin 18.2 for treatment of cancer

Calibrated KC profile: KC1=5, KC2=4, KC3=4, KC4=3, KC5=5, KC6=4; calibrated ensemble=4.090; pairwise record=16-10.

V3 pairwise tournament: 16W / 10L over top-decile (n=27).

#12 US12016923B2 — TRIUMVIRA IMMUNOLOGICS USA, INC
Claudin 18.2 T cell-antigen couplers and uses thereof

Calibrated KC profile: KC1=5, KC2=4, KC3=4, KC4=3, KC5=5, KC6=4; calibrated ensemble=4.090; pairwise record=15-11.

V3 pairwise tournament: 15W / 11L over top-decile (n=27).

#13 US20220184209A1 — ASTELLAS PHARMA INC. (+1)
THERAPY INVOLVING ANTIBODIES AGAINST CLAUDIN 18.2 FOR TREATMENT OF CANCER

Calibrated KC profile: KC1=5, KC2=4, KC3=4, KC4=3, KC5=5, KC6=4; calibrated ensemble=4.090; pairwise record=14-12.

V3 pairwise tournament: 14W / 12L over top-decile (n=27).

#14 US12378313B2 — XENCOR, INC

Bispecific binding agents binding to CLDN18.2 and CD3

Calibrated KC profile: KC1=5, KC2=4, KC3=4, KC4=3, KC5=5, KC6=3; calibrated ensemble=4.018; pairwise record=14-12.

V3 pairwise tournament: 14W / 12L over top-decile (n=27).

#15 US20230149465A1 — CELLEDIT LLC

GENETICALLY MODIFIED IMMUNE CELLS EXPRESSING A CHIMERIC ANTIGEN RECEPTOR AND HAVING REDUCED PROINFLAMMATORY CYTOKINE SIGNALING

Calibrated KC profile: KC1=5, KC2=4, KC3=4, KC4=3, KC5=4, KC6=4; calibrated ensemble=3.987; pairwise record=13-13.

V3 pairwise tournament: 13W / 13L over top-decile (n=27).

Methodology

The CLDN18.2 V3 analysis ingested 504 USPTO records via 16 USPTO PPubs full-text discovery queries crossing the CLDN18.2-isoform-specific gene/protein nouns (CLDN18.2, CLDN18-2, Claudin-18.2, Claudin 18 isoform 2) with the targeting / binding / conjugation vocabulary (antibody, conjugate, target, bind, inhibitor, antagonist, modulator), CLDN18.2 modality language (anti-CLDN18.2 antibody, CLDN18.2 ADC, CLDN18.2 bispecific / T-cell engager, CLDN18.2 CAR-T), and the marketed/clinical-stage drug-name vocabulary (zolbetuximab / IMAB362 / claudiximab / Vyloy; CT041 / satricabtagene / satri-cel; ASKB589; AB011; LM-302; SHR-A1904; TST001; ZL-1211; AMG910; givastomig; ASP2138; osemitamab) across abstract, title, and claim fields, restricted to publication date 2008-01-01 through 2026-05-31. The isoform-2-specific framing requirement is critical: CLDN18.1 is an unrelated lung-specific isoform and IP overlap with .2 is minimal but non-zero. Title- and full-text-based classification yielded 266 Biopharma, 22 Diagnostic, 1 Device, and 215 Non-Biopharma classifications. The 266 biopharma patents were assigned to one of five CLDN18.2 therapeutic-area buckets (CLDN18_Naked_mAb, CLDN18_ADC, CLDN18_Bispecific_T_cell_engager, CLDN18_CAR_T, CLDN18_Other_modality) via the substrate-category-primary classifier with modality-marker universal overrides (GLP-1 backtest Path 2 canonical pattern). TA multipliers are NOT applied to the patent-level ranking; ta_weighted_score == ensemble_mean. CLDN18.2 is a tight-junction protein with no kinase activity; the five-bucket structure covers the four primary modality classes (naked mAb, ADC, bispecific / T-cell engager, CAR-T) plus a catch-all for radio-immunoconjugates, NK-cell engagers, and vaccine immunotherapy filings.

This is a 2020-forward current-state landscape of CLDN18.2 IP. The discovery window nominally spanned 2008-01-01 through 2026-05-31, but the oldest CLDN18.2-isoform-2-specific patent in the cohort publishes 2020-01-30 — USPTO PPubs has no CLDN18.2-isoform-2-specific full-text indexing before 2020, which reflects the field's commercial inflection (zolbetuximab Phase III readouts began 2018-2020). The pre-2020 Ganymed Pharmaceuticals / IMAB362 foundational priority chain (Türeci / Şahin, filed 2008-2010) — which is the composition-of-matter origin of zolbetuximab (Astellas, approved 2024) — predates the discovery horizon and is not in the scored substrate. Astellas Pharma post-acquisition filings (20 patents in the substrate; Astellas acquired Ganymed in 2016 and subsequent CLDN18.2 IP transferred to Astellas assignment) operate as the de-facto foundational layer within the scored cohort. Additionally, several named priority filers in the broader CLDN18.2 commercial landscape (AstraZeneca, Daiichi Sankyo, RemeGen, Junshi) have zero direct assignments in the cohort. This is a USPTO PPubs indexing limitation — these originators' CLDN18.2 filings may be assigned to partnered

entities or filed via PCT routes that do not fully surface in USPTO full-text search. The absence of direct AstraZeneca / Daiichi / RemeGen / Junshi assignments should be read as a data-source coverage gap rather than a substantive claim about these companies' CLDN18.2 IP positions.

All KC scoring was performed by Anthropic LLM agents reading the local full-spec corpus (abstract + independent claims + dependent claims + specification body, with section caps of 6,000 / 200,000 / 500,000 chars respectively — large enough that virtually all CLDN18.2 substrate patents reach the model intact). The per-patent loop uses prompt caching (cache_control=ephemeral on the full patent text) and routes each of the 24 KC sub-dimensions to a per-sub-dim model: Haiku 4.5 for mechanical sub-dims (A1, A3, C1, C3, D1, E3, W1, W4); Sonnet 4.5 for the default and five of the holistic sub-dims (C4, N1, N3, N4, E1 holistic-routed-to-Sonnet); and Opus 4.7 for the two holistic sub-dims D3 (Prosecution Flexibility) and A4 (Commercial Differentiation) plus the four KC2/KC3/KC4/KC6 composite calls. The D3+A4 → Opus forcing is the HOLISTIC_OPUS_FORCED correction identified by HER2 corrected QA (Sonnet exhibits rubric-collapse on D3 — pinning every patent at score 4 — and on A4 with 67.5% modal-4 vs the Opus discriminating 2/3/4 distribution). KC4 sub-dims (N1–N4) additionally receive RAG top-20 nearest-neighbour context from a text-embedding-3-small chunk-level vector index over the entire scored Biopharma corpus. Composite KC scores follow the V2 rules: KC1 and KC5 = median of sub-dimension scores rounded to nearest integer (mechanical, no LLM call); KC2, KC3, KC4, KC6 = holistic Opus composite per the V2-anchored per-KC composite prompt (KC4 carries a CLDN18.2-modality-specific anchor naming the established commercial-asset priority chains and their crowding implications). Ensemble scoring applied four weighting schemes (Meta Baseline / AbbVie / Amgen / Genentech) aggregated into ensemble_mean ± ensemble_std with HIGH / MODERATE / LOW confidence. Tier thresholds are absolute (TIER 1 PREMIUM ≥ 4.5; TIER 1 HIGH 4.0–4.49; TIER 2 MODERATE 3.0–3.99; TIER 3 SPECULATIVE < 3.0).

Before any KC scoring, a Sonnet 4.5 SS-F (Single-Shot Filter) / OS-S (One-Sided Safe) firewall screened every patent's title + abstract + first 2,000 chars of claims + first 2,000 chars of description against the V3 Quarantine Taxonomy (Q1 weaponization, Q2 illicit drug synthesis, Q3 CSAM, Q4 pseudoscience-asserting-biomed, Q5 hate-speech-jargon-with-biomed). For this corpus 0 of 266 substrate patents were quarantined. The top decile (n=27) was then re-judged via a complete pairwise tournament (Sonnet 4.5 judge) prompted to compare CLDN18.2 / Claudin-18.2 biopharma patents and pick the higher-quality IP overall (modality-agnostic comparison). Detailed V3 outputs (per-sub-dim scores with verbatim evidence quotes and model attribution; RAG neighbour lists; tournament pairings; calibrated rankings) are available in CLDN18_2_V3_KC_Subscores.csv, CLDN18_2_Top15_Predictions.csv (with JSON citation-based KC justifications), CLDN18_2_Top_Decile_V3_Calibrated_KC_Scores.csv, and the V3 workflow manifest at raw/CLDN18_2_V3_KC_Workflow_Manifest.json.

Substrate framing. This is a 2020-forward current-state competitive landscape of CLDN18.2 IP, not a complete history of the target's intellectual property. The pre-2020 Ganymed Pharmaceuticals / IMAB362 foundational priority chain is outside the discovery horizon (USPTO PPubs lacks CLDN18.2-isoform-2-specific full-text indexing before 2020) and is not in the scored substrate. The substrate ensemble mean of 3.58 is 0.18–0.32 points higher than the sibling V3-corrected targets (HER2 ≈3.36, HER3 ≈3.40, TL1A ≈3.33, TROP2 ≈3.40, GLP-1 backtest 3.26). This difference reflects substrate composition (a 2020-forward competitive cohort whose patents articulate target / modality / enablement more cleanly than a broader-discovery corpus), NOT methodology drift. Cross-target absolute-score comparisons are therefore not apples-to-apples; within-target relative rankings (top decile vs substrate-wide) are the methodologically defensible read.

KC4 prior-art-landscape discrimination. The KC4 mean across the substrate is 2.46 — 154 of the scored patents (≈58%) score 2 on KC4 (“moderately crowded prior art”) and only 10 patents (≈4%) score 4 (“sparse prior art”). This is the strongest structural-discrimination signal observed in the corrected V3 campaign to date. The KC4 sub-dimension's prompt anchor names the established CLDN18.2 commercial-asset priority chains

(Ganymed/IMAB362 → Astellas zolbetuximab; AstraZeneca / Daiichi Sankyo ADC platforms; Chinese-originator ADC layer: Innovent, Hengrui, LaNova, RemeGen; Amgen AMG910 bispecific; CARsgen CT041) and explicitly instructs that NAKED-mAb class is most crowded around the Ganymed priority chain and ADC class is moderately crowded with rapid recent Chinese-originator expansion. The 58-percent at KC4=2 outcome confirms that Opus 4.7 (the KC4 composite model) is honestly reading the post-2020 CLDN18.2 prior-art space as crowded across modalities. Sub-dimensions D3 (Prosecution Flexibility) and A4 (Commercial Differentiation), forced to Opus 4.7 per the HOLISTIC_OPUS_FORCED correction, are not in the top-10 list of failure-prone sub-dimensions, validating the routing decision for this corpus.

China-originator share by decile. The CLDN18.2 substrate carries a 39.1 percent Chinese-originator share overall (104 of 266 patents from Innovent, Hengrui, Junshi, LaNova, RemeGen, CARsgen, BeiGene, Akeso, Sino Biopharma, Shanghai Henlius, Harbour Biomed, AskGene Pharma, Suzhou Transcenta, Shenzhen Frontiergate, Beijing Xuanyi, Shanghai TTM-Bio, and other Chinese-headquartered biopharmas). Distribution across ensemble-rank cohorts shows a monotonic decrease as quality rank improves: the top quartile is 33.3 percent Chinese-originator (22/66), the top decile is 25.9 percent (7/27), and the top-10 patents are 20.0 percent (2/10). The 0.66× substrate-relative enrichment in the top decile and 0.51× in the top-5 indicate that within V3's KC1–KC6 lens, Chinese-originator filings concentrate in the middle of the IP-quality distribution rather than at the top. This is a structural finding of the scoring methodology applied to the 2020-forward CLDN18.2 substrate; no policy commentary is offered. Astellas Pharma (zolbetuximab originator, post-Ganymed-acquisition), Allogene Therapeutics, Regeneron Pharmaceuticals, Amgen Research, Seagen, and Xencor dominate the top of the ranking. The highest-ranked Chinese-originator filing in the substrate is Suzhou Transcenta Therapeutics US20220306765A1 (ensemble_mean 4.287, rank 1 substrate-wide, CLDN18_ADC bucket), an anti-CLDN18.2 antibody patent that scored at the top of the KC1 + KC5 dimensions and is spot-checked in the top-decile per-patent narratives below.

Tier distribution and ceiling. The 266-patent substrate produced 11 TIER_2_STRONG patents, 64 LOW_PRIORITY, and 191 TIER_3_WATCHLIST. No patent in the substrate reached the TIER_1_HIGH (≥4.0) or TIER_1_PREMIUM (≥4.5) ensemble-mean thresholds. The maximum ensemble in the substrate is 4.287 (Suzhou Transcenta US20220306765A1), 0.213 points below the TIER_1_HIGH threshold. This ceiling is a direct consequence of the KC4 prior-art-density reading described above: when KC4 averages 2.46 across the substrate and tops out at 4 for only ≈4 percent of patents, the four-scheme ensemble cannot ascend to TIER_1 thresholds regardless of how strong KC1 / KC2 / KC3 / KC5 / KC6 read. This is consistent with the substrate being a 2020-forward competitive landscape where the foundational composition-of-matter layer (Ganymed/IMAB362, filed 2008–2010, ex-substrate) had already cleared the pioneering-IP space before the discovery horizon. Reading the TIER_2_STRONG cohort as the substrate's effective “TIER 1” for this landscape is the methodologically honest interpretation.

Suzhou Transcenta Therapeutics US20220306765A1 — within-decile reshuffle as methodology working. This anti-CLDN18.2 antibody filing (CLDN18_ADC bucket) carries the substrate's highest ensemble_mean of 4.287 (substrate-wide rank 1). Under the V3 pairwise tournament, however, it places at calibrated rank 8 with a tournament record of 17 wins, 9 losses. This is the methodology working as designed: the ensemble_mean reading captures patent-quality dimensions in isolation, and the pairwise tournament refines that reading by asking Sonnet 4.5 to compare top-decile patents directly against each other on overall IP quality. The 17-9 tournament record is still firmly top-decile, but the within-decile reshuffle places the Allogene constitutively-active chimeric cytokine receptors patent (rank 1 calibrated; 26-0 tournament sweep) and the Allogene + Pfizer protease-activating CD45-gate CAR patent (rank 2; 25-1) ahead of it. The honest narrative for the Suzhou Transcenta filing is: a strong ADC filing in a competitive ADC field, top-decile but not the substrate's single foundational filing. Both the ensemble-rank and the tournament-rank readings are reported in the top-decile table; the within-decile reshuffle should be read as the pairwise judge applying differentiated discrimination beyond what the four-scheme ensemble captured.

Engineered-cell-therapy and drug-conjugate platforms dominate the post-2020 CLDN18.2 competitive frontier. The tournament-calibrated top three are: rank 1 Allogene Therapeutics US20200291090A1 (CAR-T constitutively-active chimeric cytokine receptors, 26-0 tournament sweep — the strongest tournament signal in the corrected V3 campaign to date); rank 2 Allogene Therapeutics + Pfizer co-assigned US20220227832A1 (protease-activating CD45-gate CAR, 25-1); rank 3 Astellas Pharma / TRON US11541127B2 (drug conjugates comprising antibodies against claudin 18.2, 24-2). Two CAR-T platform filings and one ADC. The narrative is platform-mechanism: the post-2020 CLDN18.2 competitive frontier is led by engineered-cell-therapy innovations (gated CARs, protease-activatable specificity switches, constitutively-active cytokine receptor engineering) and by sophisticated antibody-drug-conjugate platforms operated by established originators. This is not a story of Astellas / zolbetuximab dominance; it is a story of where the post-CoM innovation frontier sits in the CLDN18.2 space, and the answer the V3 KC1–KC6 rubric returns is: engineered-cell-therapy + drug-conjugate platforms.

Astellas zolbetuximab estate — intra-portfolio quality gradient. Four Astellas Pharma / TRON patents land in the calibrated top decile at ranks 3, 11, 13, and 27. Their tournament records (24-2, 16-10, 14-12, 0-26) describe a structural quality gradient within the zolbetuximab IP estate: foundational drug-conjugate IP scores highest (rank 3 — the ADC platform layer), naked-mAb therapy IP scores middle (ranks 11 and 13 — Astellas's two “therapy involving antibodies against Claudin 18.2” filings), and combination-therapy IP scores at the bottom of the cohort (rank 27 — 0 wins and 26 losses in pairwise comparison). The rank-27 filing's tournament record indicates that Sonnet judged it weaker than every other top-decile patent on the overall IP-quality comparison. The methodological reading: not all lifecycle-extension IP within an established originator's portfolio is top-tier; the V3 rubric is sensitive to structural IP differences between drug-conjugate platforms (high differentiation potential), naked-mAb therapeutic strategies (moderate, with prior-art crowding), and combination-therapy follow-ons (lowest within the scored cohort).

Chinese-originator share — strongly monotonic decrease across ensemble cohorts. Quantitative anchors for the article narrative: substrate-wide 39.1 percent Chinese-originator (104/266); top quartile 33.3 percent (22/66); top decile 25.9 percent (7/27); calibrated top 10 — 10.0 percent (1/10). The top-10 datapoint is the strongest signal: tournament calibration moved Suzhou Transcenta from ensemble rank 1 to calibrated rank 8, taking the top-10 Chinese-originator share from 20 percent (pre-tournament) down to 10 percent (post-tournament). The single Chinese-originator filing in the calibrated top 10 is Shenzhen Frontiergate Biotechnology US20240041926A1 (calibrated rank 5, CLDN18_CAR_T bucket, dendritic-cell-activating chimeric antigen receptor, 22-4 tournament record). Two exceptions to the otherwise monotonic China-share decrease are worth naming directly: Suzhou Transcenta US20220306765A1 at calibrated rank 8 (anti-CLDN18.2 antibody, ADC bucket) and Shenzhen Frontiergate US20240041926A1 at calibrated rank 5 (dendritic-cell CAR-T). Both demonstrate that the V3 KC1–KC6 lens reads structural rigor wherever it appears. The structural finding of the rubric applied to this 2020-forward substrate is that Chinese-originator filings concentrate predominantly in the middle of the IP-quality distribution rather than at the top, while Western and Japanese originators (Allogene, Astellas, Regeneron, Vivasor, UBE, Seagen, Triumvira, Xencor, BMS, Amgen) dominate the top of the cohort. No policy commentary is offered — this is the result the scoring methodology returns.

Canonical V3 rubric reference:
/Product_Roadmap/University_IP/_KC_Methodology/V3_Rubric/improved_scoring_prompt_v3.md.