Journal of High School Science

Target-family differences in stereochemical discrimination: a same-assay analysis of enantiomer pairs in ChEMBL

Ashraf D

Submitted: July 5, 2026, Revised: version 1, August 6, 2026, version 2, September 8, 2026
Accepted: September 10, 2026

Abstract
Chiral molecules can differ substantially in biological potency, yet the extent to which this stereochemical discrimination varies across protein families has not been systematically measured. Here, same-assay enantiomer pairs were extracted from ChEMBL 37 to construct an archive-scale, target-family-stratified distribution of the eudysmic ratio while minimizing variation arising from comparisons across different assays. The analysis identified 18,346 eligible structural pairs, including 12,344 assigned unambiguously to kinases, ion channels, G protein-coupled receptors, nuclear receptors, or proteases. Across all eligible pairs, the median enantiomeric potency difference was approximately threefold; 27.7% exceeded tenfold and 7.5% exceeded 100-fold. Stereochemical discrimination was generally greater among pairs assayed against proteases and nuclear receptors, intermediate for G protein-coupled receptors, and lower for kinases and ion channels. The ordering at the upper end depended on whether chemical series, targets, or individual pairs were weighted, but the broader separation between the higher- and lower-discrimination families was robust to controls for ligand chemistry, stereocenter count, assay type, congeneric-series clustering, selection bias, and recoverable censoring. Aggregation by protein produced a reproducible target-associated index across disjoint chemical series, although uncontrolled assay and ligand-design factors preclude interpreting it as an intrinsic protein property. Target identity explained substantially more variation than protein family, and the eudysmic ratio alone carried little information for predicting family membership. These findings provide the first large-scale, same-assay map of enantiomeric potency differences across major protein families and establish a reproducible framework for investigating the molecular determinants of stereochemical discrimination.

Keywords
Chirality, Enantiomers, Stereoselectivity, Eudysmic ratio, Eudysmic index, Protein–ligand interactions, Target families, Bioactivity data, ChEMBL, Cheminformatics

Dylan Ashraf, American High School, 36300 Fremont Blvd, Fremont, CA 94536, USA.
dylanashraf56014@gmail.com

https://doi.org/10.64336/001c.170283

1. Introduction

Chirality is central to how small-molecule drugs work. A protein binding pocket may itself be chiral, such that it can distinguish the two enantiomers of a chiral ligand, and those enantiomers can differ markedly in target potency (1). Roughly two-thirds of small-molecule drugs approved between 2013 and 2022 are chiral, and the great majority of those are marketed as single enantiomers (2), so the magnitude of the enantiomeric potency difference is a routine concern in drug discovery rather than a curiosity. The more potent enantiomer is called the eutomer, the less potent the distomer; their potency ratio is the eudysmic ratio, and its logarithm is the eudysmic index introduced by Lehmann, Rodrigues de Miranda and Ariëns (3) and developed quantitatively by Lehmann (4). The clinical importance of this asymmetry is textbook — thalidomide is the cautionary archetype — and a classical generalization, Pfeiffer's rule (5), holds that more potent drugs tend to show larger enantiomeric differences, although its universality has been questioned (6).

The eudysmic ratio is not a property of a ligand. It is produced when a chiral ligand meets a chiral pocket, so it is a property of the *interaction*. Holding the ligand pair fixed and changing the protein changes the number; that is precisely what makes a comparison across protein families meaningful. Interpreted in this context, the quantity of interest reflects the extent to which a given protein discriminates between two mirror-image molecules and the classical literature already contains target-level formulations of that idea: Lehmann's eudysmic affinity quotient, the slope of eudysmic index as a function of eutomer affinity across a ligand series for one target (3, 7), and, in enzymology, the enantiomeric ratio E of Chen and co-workers, the canonical kinetic measure of an enzyme's enantioselectivity (8). Eudysmic indices were shown to differ between two closely related targets — β1 versus β2 adrenoceptors (9) — three decades ago. What has not previously been done is to estimate such a quantity for hundreds of targets at once from a bioactivity archive, under a measurement control, and to determine whether its distribution varies systematically across binding-pocket classes.

Prior work addresses related questions but does not directly answer this one. The most closely related study, “Chiral Cliffs” (10), analysed approximately 3,800 chiral compound pairs from ChEMBL and trained a machine-learning model to predict whether chirality produces an activity cliff. Its aim was therefore to classify the presence or absence of a cliff, rather than to quantify the magnitude of the effect and compare its distribution across target families. Within a single target family, chirality has also been explored as a design principle across the human kinome (11). A separate study of kinase inhibitors found that whole-protein sequence properties were associated with activity-cliff propensity, whereas binding-site descriptors alone were not predictive (12). This finding informed our decision not to model the effect using pocket descriptors.

Other related approaches provide complementary evidence. Matched-molecular-pair analysis typically examines changes

produced by replacing one chemical substructure with another (13), making pure stereochemical inversion without a change in molecular connectivity poorly suited to the standard framework. More recently, a large-scale analysis using stereochemically aware bioactivity descriptors showed that many spatial-isomer pairs differ in bioactivity (14), establishing the importance of stereochemistry at scale without examining how the magnitude of the effect varies across pocket classes. Finally, an analysis of enantiomeric pairs in common in vitro ADME assays compared enantiomer differences with experimental variability for permeability, CYP, and hERG endpoints (15). The present study extends this benchmarking principle to on-target potency and asks whether the resulting effect differs systematically among target families.

A second obstacle is measurement noise. Bioactivity values aggregated across heterogeneous public assays carry a substantial dispersion: the standard deviation of pIC50 values pooled from different, unknown assays is ~ 0.68 log units (16), and combining IC50 or Ki values from different sources has been re-confirmed as a significant source of noise in ChEMBL-derived data sets (17). Any attempt to characterize genuine handedness effects — many of which are smaller than one log unit — must separate real chemistry from that dispersion. Here, this was done by a strict design choice: two enantiomers were compared only when both had been measured using the same assay. Within a single assay the enantiomers were run under identical conditions on the same protein preparation, so the between-source component of the dispersion was removed by construction. The same-assay design does not remove intra-assay (replicate, plate, day) error, which was not separately quantified; individual small differences therefore remain subject to it, and the family-level analysis rests on aggregation over thousands of pairs rather than on the resolution of any single measurement.

This paper makes four contributions. First, the archive-scale distribution of the log eudysmic ratio was built from strictly defined same-assay enantiomer pairs in ChEMBL 37. Second, that distribution was grouped by five target families and shown to have reproducible structure that survived controls for stereocenter count, congeneric-series clustering, assay type, selection bias and the broader medicinal chemistry of the ligand sets. Third, the ratio was aggregated to the protein to give a target-level eudysmic index, and that index was tested for reliability across disjoint chemical series, which is what distinguishes a property of the protein from a property of the compounds. Fourth, the mutual information between the ratio and target family was quantified to assess how informative the ratio is about target-family identity. Throughout, we distinguish what was measured—experimental potency differences between real enantiomeric pairs—from what was inferred—family-level associations and the assumptions underlying them. Accordingly, the family effect is interpreted as an association with the target class against whi each pair was assayed, rather than as evidence of an intrinsic causal property of the binding pocket.

2. Methods

2.1 Data source

All data were drawn from ChEMBL 37 (18), the manually curated open bioactivity database, obtained as the public SQLite release (dated 2026-05-29). ChEMBL 37 contains 24,527,044 activity records, of which 4,969,278 carry a pChEMBL value — the standardized −log10 molar potency that ChEMBL computes only for exact ("=" relation) measurements. All analysis used a local read-only copy.

2.2 Enantiomer-pair detection

Two molecules are true enantiomers if they share an identical constitution and relative configuration but are non-superimposable mirror images — every stereocenter inverted (as opposed to diastereomers or epimers, where only some differ). The standard InChI representation encodes this cleanly: enantiomers have identical formula, connectivity (`/c`), hydrogen (`/h`) and tetrahedral-parity (`/t`) layers, differing only in the mirror flag `/m` (`/m0` vs `/m1`), with the same absolute-stereo flag `/s1` (19). Pairs were therefore detected by parsing ChEMBL's precomputed standard InChI: for each molecule the `/m` and `/s` layers were stripped to form a key, and two molecules were taken to form an enantiomer pair when they shared that key but carried opposite, defined `/m` values (0 vs 1) and `/s1`. This guarantees a single defined configuration at the stereocenter(s) that define the enantiomer relationship. Racemates, entries with no defined `/m`, and isotopically labeled records were excluded. A residual 7.84% of pairs carried an additional unspecified stereo element elsewhere in the molecule (a `?` in the `/t` layer); excluding these left every family median within 0.08 log units and the ordering unchanged (results/revision_analyses.json), so they were retained in the primary set and reported as a sensitivity.

Table 1. Ground-truth validation of enantiomer detection. Both the InChI `/m`-flag rule and an independent RDKit mirror-inversion rule gave the expected answer in all eight cases.

Test case	Relationship	Detected as enantiomers	Correct
(R)- vs (S)-alanine	enantiomers	Yes	Yes
(R)- vs (S)-ibuprofen	enantiomers	Yes	Yes
(R,R)- vs (S,S)-tartaric acid	enantiomers (2 centers)	Yes	Yes
(R,R)- vs (S,S)-threonine	enantiomers (2 centers)	Yes	Yes
(R,R)- vs meso-tartaric acid	diastereomers	No	Yes
threonine vs allo-threonine	diastereomers	No	Yes
(S)-alanine vs (S)-alanine	identical	No	Yes
glycine	achiral (no stereocenter)	No	Yes

This detector was validated against eight ground-truth cases before any counting was performed (Table 1): four known enantiomer pairs, including two two-stereocenter cases (tartaric acid and threonine), were correctly identified; two diastereomer pairs (meso- vs (R,R)-tartaric acid, and threonine vs allo-threonine) were correctly rejected; and identical-molecule and achiral (glycine) controls produced no `/m` distinction. The InChI rule is exact for the non-isotopic case because InChI's canonical atom numbering is

achiral, so both enantiomers receive identical `/t` parities and are distinguished solely by `/m`. As real-drug checks, the known single-enantiomer drugs dexibuprofen, levodopa and esomeprazole were each correctly matched to their opposite-configuration counterparts in ChEMBL.

As an independent check, detection was re-implemented with RDKit (20): for each candidate a fully defined single stereoisomer was confirmed via `FindPotentialStereo` (a stricter, whole-molecule test than the `/m,/s1` rule), the true mirror image was generated by inverting every tetrahedral tag, and molecules sharing the {self, mirror} InChI set were paired. Axial and planar chirality were not separately treated; the analysis was limited to tetrahedral stereocenters as encoded by InChI `/t`. The two detection routines were written independently and were compared (Section 3.10).

2.3 The same-assay constraint and the log eudysmic ratio

For every ChEMBL assay, its member compounds were grouped by enantiomer key. An assay contributed a qualifying pair when it contained at least one `/m0` and one `/m1` member of the same key, both with a pChEMBL value. Replicate measurements of the same compound in the same assay were aggregated by their median. The log eudysmic ratio for a pair is given by,

$$\Delta = | \text{median}(p\_A) - \text{median}(p\_B) | \quad \text{(eq1)}$$

where p_A and p_B are the pChEMBL values of the two enantiomers in that assay. Because pChEMBL is −log10 of molar potency, Δ (eq1) is exactly log10 of the eudysmic ratio, that is, the eudysmic index,

$$\Delta = \log10(\text{potency(eutomer)} / \text{potency(distomer)}) \quad \text{(eq2)}$$

Hence Δ = 1 is a 10-fold and Δ = 2 a 100-fold potency difference. Requiring both enantiomers from the same assay identifier means that they were measured under identical conditions, which removed the between-source component of the dispersion reported for pooled public potency data (16, 17). It does not remove intra-assay error, which was not quantified here.

2.4 Target-family assignment

Each assay was mapped to its target via the ChEMBL target dictionary, and each target to protein families through `target_components → component_class → protein_classification`. ChEMBL 37 stores the classification as a parent–child tree, so a class was assigned to a family by walking its ancestor chain to one of five roots: GPCR (any G protein-coupled receptor subfamily under "Membrane receptor"), Kinase, Protease, Ion channel and Nuclear receptor. Targets whose single unambiguous family fell in this set were retained; targets outside these families, or mapping to more than one, were excluded from family comparisons. Family assignment is independent of the potency data, hence the

analysis is non-circular. Targets were ChEMBL target identifiers, so orthologues of the same protein in different organisms were counted separately.

2.5 Filters and unit of analysis

For the family analysis the data were restricted to binding (`B`) and functional (`F`) assays with target confidence score ≥ 8 (direct single-protein assignment). The primary unit of analysis is the unique structural pair: a given pair may be measured in many assays, so each pair's Δ was aggregated across its assays (median) before all family statistics, so that each pair contributes one value per family. Where a per-protein quantity was required (Section 2.7), the unit was instead the (pair × target) combination.

2.6 Ligand descriptors

To test whether family differences reflect the medicinal chemistry of the ligand sets rather than the proteins, fourteen 2D descriptors were computed once per structural pair with RDKit (20): molecular weight, Crippen cLogP, topological polar surface area, hydrogen-bond donors and acceptors, rotatable bonds, fraction of sp3 carbons, aromatic-ring count, total ring count, heavy-atom count, amide-bond count (a peptide-likeness proxy), QED, Bertz complexity, and a normalized flexibility index (rotatable bonds per heavy atom). The number of defined stereocenters was taken from the InChI `/t` layer. Computing these once per pair is exact rather than approximate, because the two members of an enantiomer pair are identical in constitution and every descriptor used here is a function of the 2D graph alone; the descriptors therefore characterize the ligand series, which is the confound at issue. Collinear descriptors were removed by iteratively dropping the largest variance inflation factor until all remained below 10. This removed heavy-atom count, Bertz complexity and rotatable bonds and retained eleven descriptors.

2.7 The target-level eudysmic index

For each target t with at least 20 enantiomer pairs, a target-level eudysmic index was defined as the median of Δ over the pairs assayed against that target. The minimum of 20 pairs trades per-target precision against coverage: it retains 224 targets and 83.1% of all pairs, at a median bootstrap interval width of 0.39 log units, whereas a minimum of 5 would retain 96.0% of pairs but widen the median interval to 0.49, and a minimum of 50 would narrow it to 0.32 while discarding 41% of the data. As with the equivalence margin, the threshold is a convention, so every conclusion drawn from the index is reported across minima of 20, 30 and 50 pairs (Section 3.12). This is the target-level aggregate of the classical per-compound eudysmic index (2, 3); it is not a new concept, and it is deliberately not given a new acronym. Candidate abbreviations such as PSD and SDI are already heavily used in this literature for postsynaptic density, post-source decay and the clinical-chemistry standard deviation index, so a new abbreviation would collide rather than clarify. The quantity is related in intent to Lehmann's eudysmic affinity quotient (3, 4) and is the bioactivity-side analogue of the kinetic enantiomeric ratio *E* used in enzymology (8); what is new here is the estimator — archive-scale, same-assay-controlled, and computed for hundreds of targets at once — not the definition.

Because an index of this kind is only useful if it is reproducible, three properties were tested. Split-half reliability was estimated by randomly partitioning each qualifying target's pairs into halves, computing the index in each half, and correlating the two halves across targets (Spearman), averaged over 300 partitions and corrected to full length by the Spearman–Brown formula. Scaffold-disjoint split-half reliability repeated this with the partition made over Bemis–Murcko scaffolds (21) rather than over individual pairs, so that no chemical series contributes to both halves; this is the test that distinguishes a property of the protein from a property of the series that happened to be assayed against it. Variance decomposition used a linear mixed model with a random intercept for target to estimate the intraclass correlation, that is, the share of variance in Δ lying between targets (22, 23), before and after including family as a fixed effect.

2.8 Statistics

Family distributions of Δ were compared with the Kruskal–Wallis omnibus test, followed by pairwise Mann–Whitney U tests with Holm correction (24) for the ten family pairs; effect sizes are the rank-biserial correlation,

$$r = 2U / (n_A \cdot n_B) - 1 \quad \text{(eq3)}$$

Because the overlap of two per-family confidence intervals is not a test of the difference between them, the primary object reported for each family pair is a 4,000-resample bootstrap confidence interval on the difference between the two medians, together with the rank-biserial correlation and its bootstrap interval. Two families were described as equivalent only when that difference interval lay entirely inside an equivalence margin of ±0.10 log units; a merely non-significant test was not treated as evidence of equivalence. The margin is anchored to the reproducibility of the underlying measurements, not to the observed family differences. A margin of 0.10 log units corresponds to a 1.26-fold potency difference. That is below the standard deviation of repeated determinations of the same compound in the same assay, reported as σ = 0.17–0.22 log units for in-house repeat measurements, and far below the σ ≈ 0.68 log units of potency values pooled across heterogeneous public sources, for which 68% of measurements agree only within a factor of 4.8 (16, 17). The ±0.10-log margin was selected as a conservative threshold for practical equivalence because it is smaller than typical repeat-measurement variability and lies within the range commonly treated as experimental variation in medicinal chemistry (16, 17). Because the numerical value of any such margin is a convention rather than a fact, equivalence verdicts are reported across a range of margins (Section 3.3) instead of at a single threshold. Confidence intervals elsewhere are 95% (2,000-resample bootstrap for medians; Wilson score interval for proportions (25)). All tests were two-sided. Analyses used SciPy 1.16 (26), statsmodels 0.14 (27) and scikit-learn 1.7 (28). Because reported p-values do not correct for residual non-independence among congeneric compounds, the effect sizes, confidence

intervals and robustness analyses — not the omnibus p-values — are treated as the load-bearing evidence.

Adjusted family effects were estimated by median (quantile) regression (29), which targets the same median that the descriptive analysis reports, with inference from a 400-resample bootstrap clustered on Bemis–Murcko scaffold; ordinary least squares with scaffold-cluster-robust standard errors (30) is reported alongside. Because R^2 is not defined for quantile regression and no pseudo-R^2 was substituted, every figure reported as "variance explained" or "R^2" in Section 3.11 is the ordinary-least-squares coefficient of determination from a separate least-squares fit of the same predictors to Δ, and is intended only to put the competing effects on a common scale. The tail statistic was modeled by logistic regression for $P(\Delta > 1)$ with the same clustering. As a non-parametric complement, each family was compared with a kinase set matched 1:1 on a propensity score built from ligand descriptors alone, within the conventional caliper of 0.2 standard deviations of the logit of the propensity score.

Information content was estimated as the mutual information between Δ and family, using a nearest-neighbour estimator for a continuous variable against a discrete label (31, 32), converted to bits and referenced to the family entropy, with a 200-permutation null. Classification used multinomial logistic regression and random forests evaluated under 5-fold cross-validation grouped by Bemis–Murcko scaffold, so that no chemical series appears in both training and test folds; ungrouped splits are known to overstate performance in cheminformatics benchmarks (33, 34). No claim is made that any model recovers chemical knowledge: the models are used only as estimators of how much information the measured quantity carries.

2.9 Robustness analyses

Six controls tested whether the family association is an artifact. (i) Stereocenter count: pairs were stratified by the number of defined stereocenters and the comparison re-run within strata. (ii) Congeneric-series non-independence: the Bemis–Murcko scaffold (21) of each pair was computed (stereochemistry stripped) and every scaffold within a family collapsed to a single median Δ. (iii) Assay type: the binding-versus-functional composition was recorded, the comparison re-run within binding-only assays, and the functional-only complement examined. (iv) Selection bias: assay size (distinct pChEMBL compounds per assay) was used as a proxy for incidental testing and the comparison re-run within assay-size strata, reporting effect sizes rather than omnibus p-values, together with the target and assay concentration inside each stratum. The stratum boundaries (2–4, 5–30, 31–200, >200 compounds) are conventional round numbers separating focused pairwise comparisons from medium series and from large screening panels; they were not selected against the outcome, and the argument rests on the monotone trend in effect size across strata rather than on any single boundary. (v) Ligand chemistry: the multivariable adjustment and matching described in Sections 2.6 and 2.8. (vi) Censoring: the effect of the requirement that both enantiomers carry an exact value,

described in Section 2.10.9 and quantified in Section 3.14. The number of distinct targets per family, a leave-one-target-out check, and the correlation of Δ with eutomer potency (Pfeiffer's rule) are also reported.

2.10 Explicit assumptions

The following assumptions underlie the design and the analyses that follow.

2.10.1 *The InChI mirror-flag rule identifies true enantiomers*

Justified by the achirality of InChI canonical numbering, validated on eight ground-truth cases (Table 1) and cross-checked by an independent RDKit implementation. It is exact for non-isotopic tetrahedral cases and silent on axial and planar chirality, which are therefore out of scope.

2.10.2 *Two measurements sharing an assay identifier were made under the same conditions*

This is the design's central assumption. It is what the ChEMBL assay identifier is intended to denote, and it is supported empirically by the observation that 99.9% of pairs share an activity type within their assay. It would fail if a single assay record aggregated heterogeneous experiments; the impact of such failures would be to add noise to Δ and so to bias effect sizes toward zero.

2.10.3 *pChEMBL values are comparable within an assay*

ChEMBL computes pChEMBL only for exact measurements of a defined set of potency types. Comparability across assays is not assumed anywhere in this work.

2.10.4 *Aggregating replicates and assays by the median is appropriate*

The median was chosen for robustness to outlying records. Because Δ is defined as an absolute difference (eq1), aggregation is applied to the two potencies before the difference is taken within an assay, and across assays to the resulting Δ values.

2.10.5 *Taking the absolute value discards eutomer identity*

Which enantiomer is the eutomer is not analysed here; the question addressed is the magnitude of discrimination, not its direction. This assumption would matter for any question about configuration preference (for example, whether R or S is favoured within a family), which is outside the present scope.

2.10.6 *ChEMBL protein classification is accurate, and independent of the potency data*

Independence is what makes the analysis non-circular. Accuracy is assumed, not verified; misclassification would attenuate family differences.

2.10.7 *2D descriptors characterize an enantiomer pair*

Exactly true for the descriptors used, since the two members share a constitution (Section 2.6). It also means that no descriptor used in the adjustment can capture the 3D difference between the enantiomers themselves, which is the quantity under study rather than a confounder.

2.10.8 *Pairs are exchangeable within a family after scaffold collapse*

Only approximately true; residual dependence through shared targets is why target-level analyses (Section 3.12) are reported alongside pair-level ones.

2.10.9 *Pairs excluded because one enantiomer lacks an exact value are missing at random with respect to Δ*

This assumption is false, in a knowable direction, and Section 3.14 quantifies it. The weaker enantiomer is reported as a bound precisely when it is too weak to quantify, which is when Δ is largest, so the observed distribution is truncated from above and the reported statistics are conservative.

2.10.10 *Absolute rates describe assayed pairs, not chiral drugs in general*

Both enantiomers are typically made and tested together only when a chemist already suspects that chirality matters, so no population rate of stereoselectivity is claimed anywhere in this work.

3. Results

3.1 Pair counts by family

After enantiomer detection and the same-assay constraint, ChEMBL 37 yielded 47,350 qualifying pair instances collapsing to 23,086 unique structural pairs. Restricting to binding/functional assays against high-confidence (score ≥ 8) targets left 18,346 unique pairs, of which 12,344 mapped unambiguously to one of the five families and entered every family comparison below. Every family contributed at least 921 unique structural pairs (Table 2), spread across 37–331 distinct targets, and the per-family bootstrap intervals on the median spanned 0.04–0.24 log units, which is narrow relative to the between-family gaps. All 201,799 enantiomer-pair activity records carried the exact ("=") relation — necessarily so, because ChEMBL computes pChEMBL only for exact measurements, so censored values are excluded by construction rather than by any filter applied here (Section 3.14 quantifies the consequence). Empirically, 99.9% of pairs shared the same activity type within their assay, so mixing IC50, Ki and EC50 values is not a concern.

3.2 The overall distribution

Across the 18,346 unique structural pairs, the log eudysmic ratio returned a median of 0.48 (interquartile range 0.18–1.10; Figure 1). Handedness effects were common among these assayed pairs — 27.7% exceeded one log unit (10-fold) and 7.5% exceeded two log units (100-fold) — with a tail beyond 4 log units. These are rates *among assayed enantiomer pairs*, not a population rate over all chiral drugs (assumption 2.10.10). Much of the distribution lay below 0.68 log units, the standard deviation reported for potency values pooled across heterogeneous sources (16); because the pairs are same-assay, that between-source component does not apply to them. This does not establish that any individual sub-threshold Δ is resolved, since intra-assay error was not quantified.

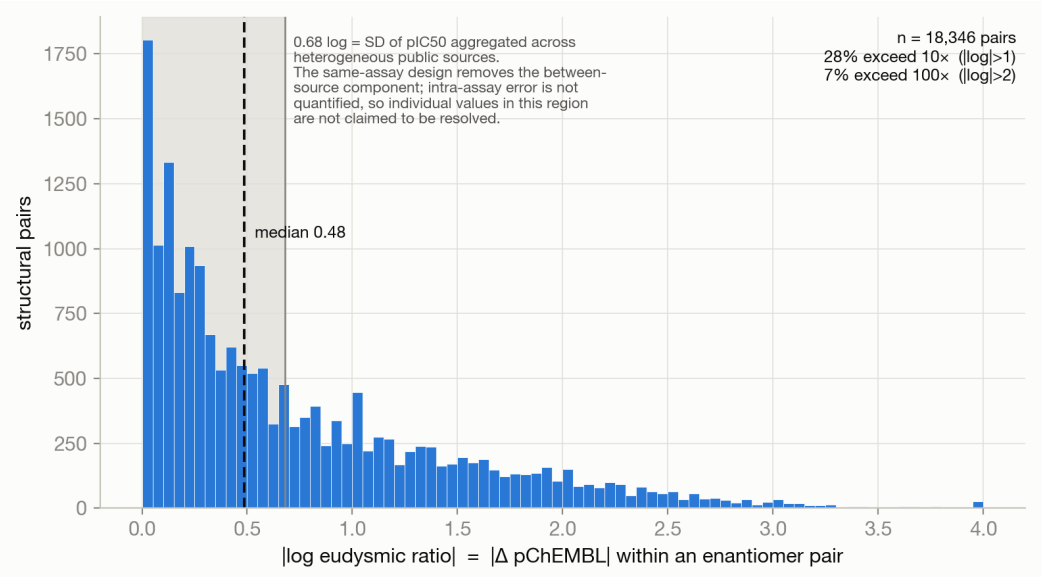

Figure 1. Archive-scale distribution of the log eudysmic ratio over 18,346 same-assay enantiomer structural pairs. Dashed line, median (0.48); shaded region, the 0.68-log standard deviation reported for potency data aggregated across heterogeneous public sources (16), whose between-source component the same-assay design removes. Intra-assay error is not quantified, so individual values inside the shaded region are not claimed to be resolved. Values are clipped at 4 log units for display.

3.3 Association with family

Grouping by target family gave a monotone ordering of medians (Table 2, Figure 2): kinase 0.39, ion channel 0.40, GPCR 0.59, nuclear receptor 0.77, protease 0.98. The families differed highly significantly (Kruskal–Wallis H = 451, p = 2.6×10^{-96}), but that statistic is inflated by the n = 12,344 pairs entering the comparison and by residual non-independence among congeneric analogues, hence the differences between medians and their confidence intervals are reported as the primary evidence.

Of the ten family pairs, kinase and ion channel were equivalent: their median difference was −0.010 with a 95% bootstrap interval of [−0.04, 0.03], lying entirely inside the ±0.10-log equivalence margin. Every one of the other nine differences had an interval excluding zero, including nuclear receptor versus protease (−0.21 [−0.36, −0.09]). After Bemis–Murcko scaffold collapse (Section 3.5) the picture was the same except at the top of the range: kinase versus ion channel remained equivalent (−0.010 [−0.065, 0.035]) and nuclear receptor versus protease became indeterminate (−0.125 [−0.255, 0.032]) — neither separated nor equivalent within the margin. The honest description is therefore a monotone gradient with one genuine tie at the bottom and an unresolved boundary at the top, rather than three clean tiers.

The kinase–ion channel equivalence conclusion is robust across margins from ±0.05 to ±0.50 log units because its confidence interval lies within ±0.04. For the other comparisons,

exclusion of zero establishes statistical separation but does not preclude practical equivalence at wider margins; their equivalence classifications were therefore recalculated at each margin and are reported in results/margin_sensitivity.json. After scaffold collapse, kinase versus ion channel requires a margin of approximately ±0.065 for equivalence, whereas nuclear receptor versus protease requires approximately ±0.255 (1.80-fold). At the prespecified ±0.10 margin, the former is equivalent and the latter is indeterminate.

Effect sizes were modest throughout (rank-biserial 0.03–0.31), so the family distributions overlapped substantially; this is a systematic shift in location, not a categorical divide between "chiral" and "flat" pockets.

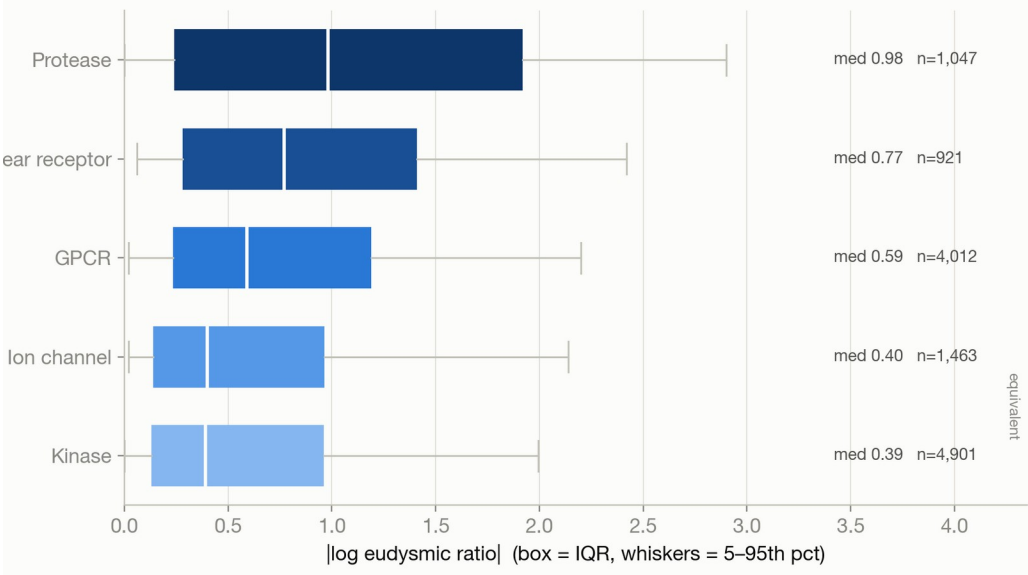

Figure 2. Per-family distributions of the log eudysmic ratio (box, interquartile range; whiskers, 5th–95th percentiles; white line, median), ordered by median. Kinase and ion channel are statistically equivalent within a ±0.10-log margin; all other adjacent contrasts are separated in the per-pair analysis.

The tail carries a sharper family signal than the median (Figure 3). The fraction exceeding 10-fold rose from 22% (kinase) and 24% (ion channel) through 31% (GPCR) to 39% (nuclear receptor) and 49% (protease); the fraction exceeding 100-fold rose from 4.8% [4.2–5.4] for kinase to 22.1% [19.7–24.7] for protease. Among protease-assayed pairs, a >100-fold enantiomeric difference is therefore about 4.6 times as frequent as among kinase-assayed pairs (~ 3.6–5.9 allowing for the interval on each proportion).

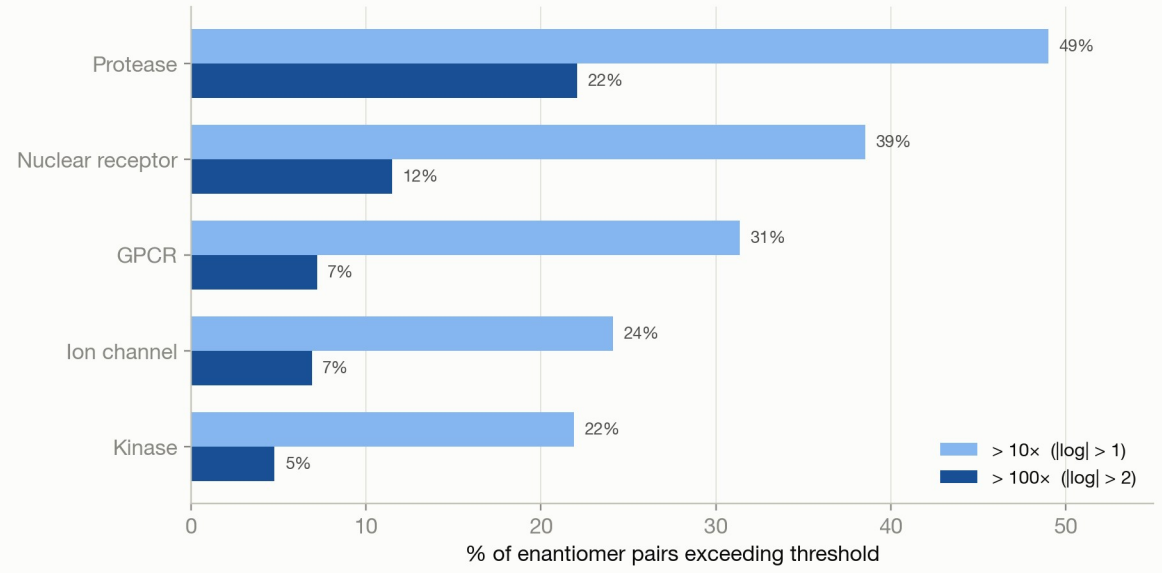

Figure 3. Fraction of enantiomer pairs per family exceeding 10-fold and 100-fold stereoselectivity, ordered by family median. The tail is the sharper family signal.

Table 2. Per-family distribution of the log eudysmic ratio (unique structural pairs; binding/functional assays, high-confidence single-protein targets), ordered by increasing median. Brackets are 95% confidence intervals (bootstrap for medians, Wilson for proportions). "ER" = eudysmic ratio.

Family	Pairs / targets	Median log ER [95% CI]	>10-fold [CI]	>100-fold [CI]
Kinase	4,901 / 217	0.39 [0.37–0.41]	21.9% [20.8–23.1]	4.8% [4.2–5.4]
Ion channel	1,463 / 102	0.40 [0.37–0.43]	24.1% [22.0–26.4]	6.9% [5.7–8.3]
GPCR	4,012 / 331	0.59 [0.56–0.62]	31.4% [30.0–32.8]	7.2% [6.4–8.0]
Nuclear receptor	921 / 37	0.77 [0.68–0.82]	38.5% [35.5–41.7]	11.6% [9.7–13.8]
Protease	1,047 / 125	0.98 [0.86–1.10]	49.0% [46.0–52.0]	22.1% [19.7–24.7]

3.4 Stereocenter count does not explain the observed effect

Families might differ not in pocket chemistry but in the molecules made for them — protease programs favouring rigid peptidomimetics with several stereocenters, kinase programs flatter binders with one. Stratifying by the number of defined stereocenters tested the simplest version of that concern. All five families had a median of one defined stereocenter, and 54–65% of pairs in every family were single-stereocenter. Within the single-stereocenter stratum the ordering was essentially unchanged (kinase 0.35, ion channel 0.45, GPCR 0.55, nuclear receptor 0.88, protease 0.91; Kruskal–Wallis $p = 4.3 \times 10^{-69}$). The two-stereocenter stratum ($n = 3{,}060$) is also highly significant ($p = 2.1 \times 10^{-27}$) but the ordering was *not*

preserved: ion channel fell to 0.34, below kinase at 0.55, and GPCR (0.78) exceeded nuclear receptor (0.675). In the three-or-more stratum (n = 1,603) the ordering was again different (ion channel 0.33, kinase 0.39, nuclear receptor 0.50, GPCR 0.51, protease 1.43). What was stable across all three strata was that protease ranked highest and that kinase and ion channel fell in the lower half; the relative placement of GPCR and nuclear receptor was not. The association is therefore not explained by the number of stereocenters, but stereocenter count is only one facet of molecular complexity, and the broader ligand-chemistry control is reported in Section 3.11.

Even unique structural pairs are not fully independent: one program may contribute dozens of analogues sharing a scaffold and a stereocenter. Collapsing every Bemis–Murcko chemical series within a family to a single median value tested this. The pseudo-replication was mild (1.6–1.9 pairs per scaffold; largest series 67), and after collapse the family ordering was preserved and remained highly significant (Kruskal–Wallis $p = 8.1 \times 10^{-55}$; Table 3, Figure 4). The difference intervals after collapse are given in Section 3.3: eight of ten contrasts remained separated, kinase versus ion channel remained equivalent, and only nuclear receptor versus protease became indeterminate.

3.5 Congeneric series do not explain the observed effect

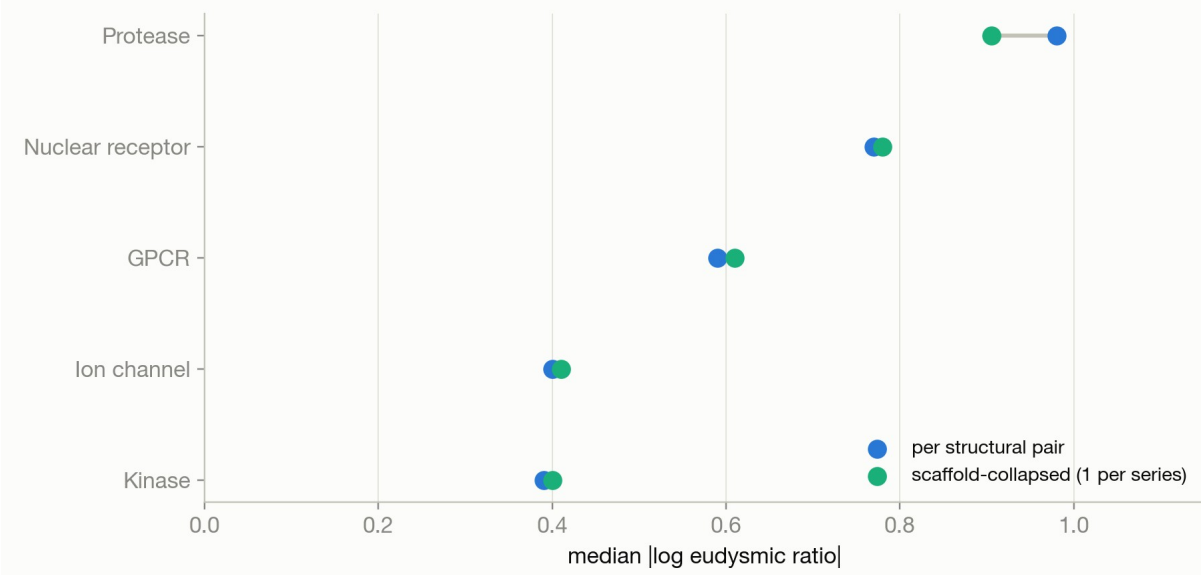

Figure 4. Median log eudysmic ratio per family computed per structural pair (blue) versus per Bemis–Murcko series (green; scaffold-collapsed, one value per chemical series). The medians shift negligibly and the ordering is preserved (Kruskal–Wallis on the collapsed data, $p = 8.1 \times 10^{-55}$).

3.6 Assay type does not explain the observed effect

Because functional (pEC50) potency can reflect efficacy and signal amplification rather than pocket binding alone, differing binding-versus-functional composition across families

could in principle drive the ordering. In every family, 83–99.8% of pairs were measured in at least one binding assay, and restricting to binding-only assays reproduced the ordering and medians almost exactly (kinase 0.39, n = 4,892; ion channel 0.40, n = 1,410; GPCR 0.59, n = 3,332; nuclear receptor 0.78, n = 857; protease 0.985, n = 1,030; Kruskal–Wallis $p = 2.3 \times 10^{-96}$; Table 3). Because this restriction removes only about 7% of the family-assigned pairs, it is a weak test on its own. The complementary set — pairs never measured in a binding assay — was thin in four families (kinase 9, protease 17, ion channel 53, nuclear receptor 64 pairs) and could not support a comparison there. For GPCR, where it was adequate (n = 680), the functional-only median was 0.61 against 0.59 for binding, essentially identical. The ordering was therefore not created by functional assays; no claim is made about how enantioselectivity differs between binding and functional readouts in general.

3.7 Target dominance and within-family heterogeneity

Because the thinner families sit at the top of the ordering, a natural question is whether a family's median is carried by a few dominant targets. For kinase, GPCR and ion channel it was not: they spanned 217, 331 and 102 distinct targets, the most-represented target contributed at most 12.4% of pairs (ion channel; kinase 7.5%, GPCR 6.3%), and removing any single target shifted the family median by at most 0.04 log units. The two high-ranked families were both concentrated, and to a similar degree. Nuclear receptor spanned only 37 targets, with ROR-γ contributing 46.7% of its pairs and a maximum leave-one-target-out shift of 0.157. Protease spanned 125 targets, but β-secretase 1 (BACE1) supplied 16.5% and its maximum leave-one-target-out shift was 0.16 — indistinguishable from that of nuclear receptor. Protease was in addition far more heterogeneous across its targets (interquartile range of per-target medians 0.33–1.42, width 1.10) than nuclear receptor (0.58–0.97, width 0.39). Both high-ranked families should therefore be read with caution: nuclear receptor because one target dominates its pair count, protease because its targets disagree with one another. Target count alone is therefore a poor measure of the robustness of a family's position in the ranking.

3.8 The family ordering is not explained by eutomer potency (Pfeiffer's rule)

Consistent with Pfeiffer's rule (5), the log eudysmic ratio correlated positively with eutomer potency, both overall (Spearman $\rho = 0.23$, $p < 10^{-100}$) and within every family ($\rho = 0.19$–0.38). Eutomer potency was in fact the strongest single predictor of Δ examined here ($R^2 = 0.073$ on its own, Section 3.11). This within-family potency gradient did not, however, explain the between-family ordering: median eutomer potency spanned only 0.7 log units across the five families (pChEMBL 7.22–7.92) and did not track the Δ ordering, the most potent family (kinase, 7.92) being the one with the smallest median Δ. Barlow's objection to Pfeiffer's rule (6) — that the rule admits marked exceptions and neglects differences in molecular flexibility — anticipates exactly this kind of discrepancy.

3.9 Selection bias

The most serious concern is publication and selection bias: both enantiomers tend to be made and assayed together mainly when a chemist had already suspected that chirality mattered, so the sampled distribution is conditioned on "chirality was interesting here". If that drove the family differences, they should weaken where pairs are tested incidentally, in a large screening panel rather than a focused study. Stratifying on assay size (Table 4, Figure 5) and measuring separation with the sample-size-independent rank-biserial correlation rather than an omnibus p-value, the protease-versus-kinase effect size was 0.24 [−0.03, 0.49] in the smallest stratum, 0.21 [0.14, 0.27] and 0.21 [0.15, 0.27] in the intermediate strata, and rose to 0.61 [0.53, 0.69] in large screening panels (>200 compounds) — the opposite of what selection bias predicts. Nuclear receptor versus kinase behaved similarly (0.34 [0.28, 0.39] in large panels).

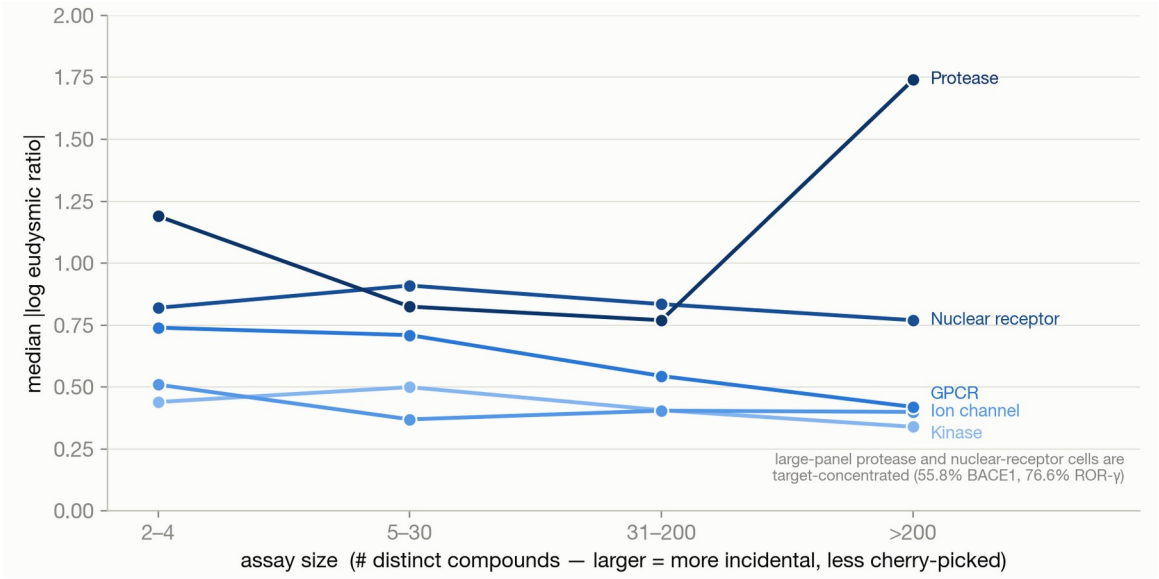

Figure 5. Median log eudysmic ratio per family across assay-size strata (larger panels = more incidentally tested). Protease and nuclear receptor remain elevated in every stratum and the protease–kinase effect size is largest in large screening panels; the GPCR/ion-channel/kinase distinction does not survive there. The large-panel protease and nuclear-receptor cells are dominated by BACE1 and ROR-γ respectively (Section 3.9).

Three qualifications are necessary. First, the elevation that survived in the least selection-prone stratum was that of protease and nuclear receptor only: in large panels GPCR (0.42) was level with ion channel (0.40) and kinase (0.34), and the GPCR-versus-ion-channel effect size was 0.00 [−0.08, 0.06]. What the large-panel data support is a two-level pattern, not the full five-way ordering. Second, GPCR and kinase medians declined as panels increased, which is a genuine, mild selection effect consistent with focused studies enriching for chirally active compounds. Third, and most important, the large-panel cells were themselves target-

concentrated: 55.8% of protease measurements in that stratum come from BACE1 and 54.5% from just three assays, and 76.6% of the nuclear-receptor measurements come from ROR-γ. The large-panel result therefore shows that the effect persists under incidental testing at those particular targets; it does not establish that the effect generalizes across the families as a whole.

Table 3. Robustness of the family ordering: median log eudysmic ratio per family under the baseline analysis and three controls. Every column preserves the position of protease at the top and of kinase near the bottom.

Family	Baseline	Single-stereocenter	Scaffold-collapsed	Binding-only
Kinase	0.39	0.35	0.40	0.39
Ion channel	0.40	0.45	0.41	0.40
GPCR	0.59	0.55	0.61	0.59
Nuclear receptor	0.77	0.88	0.78	0.78
Protease	0.98	0.91	0.91	0.99
Kruskal–Wallis p	2.6×10^{-96}	4.3×10^{-69}	8.1×10^{-55}	2.3×10^{-96}

Table 4. Median log eudysmic ratio per family across assay-size strata (number of unique pairs in parentheses), with the protease-versus-kinase rank-biserial effect size and its 95% bootstrap interval.

Family	tiny (2–4)	small (5–30)	medium (31–200)	large (>200)
Kinase	0.44 (41)	0.50 (952)	0.41 (2,134)	0.34 (2,104)
Ion channel	0.51 (60)	0.37 (419)	0.41 (634)	0.40 (480)
GPCR	0.74 (189)	0.71 (1,757)	0.55 (1,813)	0.42 (620)
Nuclear receptor	0.82 (37)	0.91 (270)	0.84 (374)	0.77 (395)
Protease	1.19 (33)	0.83 (442)	0.77 (471)	1.74 (167)
Protease vs kinase, *r*	0.24 [−0.03, 0.49]	0.21 [0.14, 0.27]	0.21 [0.15, 0.27]	0.61 [0.53, 0.69]

3.10 Cross-validation of the two detection routines

The RDKit mirror-inversion routine yields 44,675 pair instances over all targets (21,497 unique structural pairs, median Δ = 0.425); restricted to single-protein targets it yields 33,148 instances (17,038 unique pairs, median Δ = 0.450). Per-family unique structural pairs are GPCR 3,878, kinase 4,651, protease 861, ion channel 1,280 and nuclear receptor 835, with medians of 0.560, 0.350, 1.000, 0.550 and 0.800 respectively. The family rank order was identical to that of the primary analysis, and four of the five medians agreed to within 0.02–0.04 log units. The exception was the ion-channel family, whose median was 0.15 log units higher in the cross-check (0.55 versus 0.40), bringing it level with GPCR (0.56). The low end of the ordering is therefore implementation- and filter-dependent in a specific way: the separation of ion channel from GPCR appears under the target-

confidence and binding/functional filters of the primary analysis but not in the unfiltered cross-check, and Section 3.12 shows the same instability when targets rather than pairs are weighted equally. The count difference between routines is expected, since the RDKit routine restricts to single-protein targets and additionally drops approximately 5,300 of approximately 52,000 candidate molecules (≈10%) that carry an unspecified stereo element under the stricter whole-molecule test; this is a molecule-level rate and is not directly comparable with the 7.84% pair-level rate of Section 2.2. It should also be noted that only the stereo-detection step differs between the two routines — same-assay grouping, pChEMBL filtering, Δ computation and family mapping are shared — so this is a check on enantiomer detection, not an independent replication of the family result.

3.11 Ligand chemistry differs sharply between families, and the family effect survives adjustment for it

Different families have historically been pursued with different medicinal chemistry, so the association could reflect ligand rigidity, peptide-likeness or pharmacophore style rather than the protein. That confound was real and large.

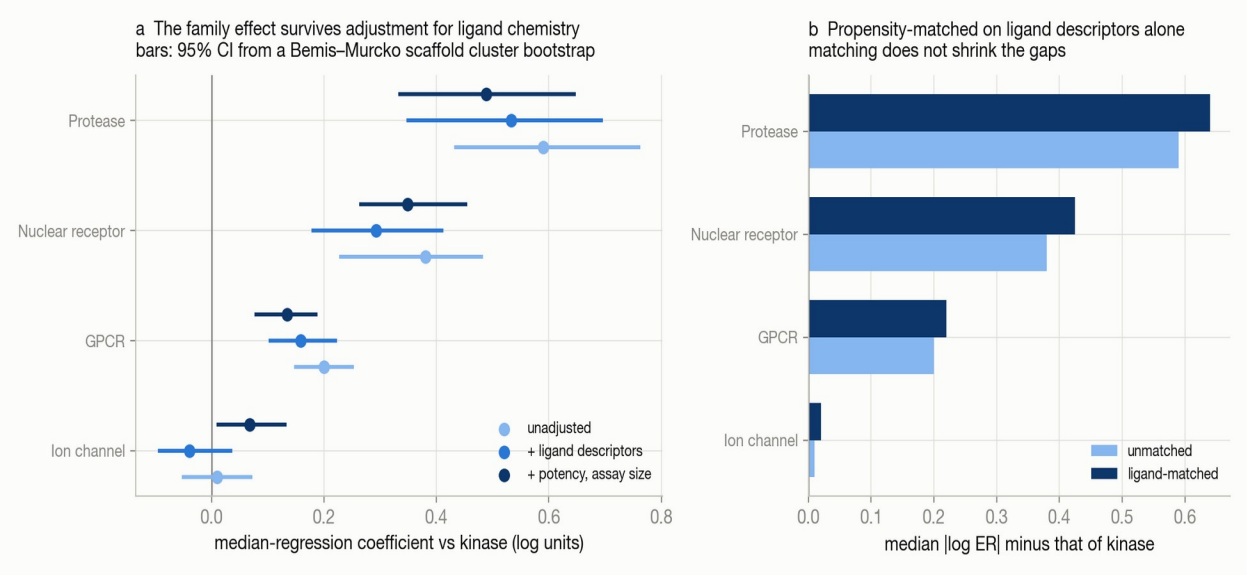

Figure 6. (a) Median-regression coefficients for each family relative to kinase, unadjusted, adjusted for ligand descriptors and stereocenter count, and additionally adjusted for eutomer potency and assay size; bars are 95% intervals from a Bemis–Murcko scaffold cluster bootstrap. (b) Difference in median log eudysmic ratio from kinase, unmatched and after 1:1 propensity matching on ligand descriptors alone. Adjustment and matching leave the gaps essentially intact.

Table 5. Median ligand descriptors per family, for all fifteen descriptors examined (one value per structural pair; identical for the two enantiomers by construction, since every descriptor is computed from the 2D graph). Every descriptor differs significantly across families. Three descriptors — heavy atoms, Bertz complexity and rotatable bonds — were dropped from the adjusted regressions for collinearity, at the conventional variance-inflation-factor threshold of 10, but are tested and reported here.

Descriptor	Kinase	Ion channel	GPCR	Nuclear receptor	Protease	Kruskal–Wallis p
Molecular weight	448.6	443.9	421.6	509.5	454.5	7.7×10^{-129}
Heavy atoms	33	31	30	35	32	2.5×10^{-135}
cLogP	3.44	3.64	3.83	4.89	3.42	3.8×10^{-191}
TPSA (Å²)	103.8	83.8	69.4	81.8	99.9	$< 10^{-300}$
H-bond donors	2	1	1	1	2	$< 10^{-300}$
H-bond acceptors	6	5	4	5	5	$< 10^{-300}$
Rotatable bonds	5	5	6	6	6	4.4×10^{-48}
Flexibility (rot. bonds / heavy atom)	0.167	0.167	0.192	0.176	0.192	2.5×10^{-102}
Fraction sp³ C	0.346	0.371	0.391	0.381	0.314	5.1×10^{-94}
Rings	5	4	4	4	4	2.2×10^{-206}
Aromatic rings	3	3	2	3	3	$< 10^{-300}$
Amide bonds	1	1	1	0	1	3.3×10^{-59}
Bertz complexity	1233.9	1103.8	984.1	1271.7	1150.3	1.2×10^{-263}
QED	0.501	0.580	0.602	0.438	0.492	5.2×10^{-120}
Defined stereocenters	1	1	1	1	1	2.5×10^{-17}

All fifteen descriptors examined differed significantly across families (Table 5): nuclear-receptor ligands are the heaviest and most lipophilic (median MW 509, cLogP 4.89) and carry no amide bond at the median, protease ligands have the lowest fraction of sp3 carbons of the five families (0.314) together with the second-highest polar surface area — they are the flattest by that measure, yet show the largest Δ, and GPCR ligands were the smallest and least polar. The consequence is quantified in Section 3.13: target family could be predicted from ligand descriptors alone at a macro one-vs-rest AUC of 0.86.

Nevertheless the family effect was almost undiminished by adjustment for that chemistry. In a median regression of Δ on family plus the eleven retained ligand descriptors and stereocenter count, with inference from a scaffold-clustered bootstrap, the protease coefficient fell only from +0.590 to +0.533 (9.7% attenuation, 95% CI [0.350, 0.694]), nuclear receptor from +0.380 to +0.293 (22.9%, [0.181, 0.410]) and GPCR from +0.200 to +0.158 (21.0%, [0.104, 0.221]); ion channel remained indistinguishable from kinase (−0.040, [−0.093, 0.034]). Adding eutomer potency and assay size left protease at +0.489 [0.335, 0.645] (Figure 6a). For the tail statistic, the ligand-adjusted odds of exceeding

10-fold relative to kinase are 3.00 [2.40, 3.75] for protease, 1.85 [1.42, 2.39] for nuclear receptor, 1.43 [1.20, 1.72] for GPCR and 0.97 [0.75, 1.27] for ion channel.

The non-parametric check agreed and was if anything stronger. Matching each protease pair 1:1 to a kinase pair on a propensity score built from ligand descriptors alone gave a median Δ of 0.98 against 0.34 for the matched kinase set — a gap of 0.64, slightly *larger* than the unmatched gap of 0.59 ($n = 1{,}047$, $p = 2.3 \times 10^{-38}$). The same held for nuclear receptor (0.425 matched versus 0.38 unmatched) and GPCR (0.22 versus 0.20), while ion channel remained null (0.02, $p = 0.33$) (Figure 6b). Ligand chemistry does not account for the family differences; matching on it does not close the gaps.

The variance partition puts all of this in proportion. These shares are ordinary-least-squares coefficients of determination (R^2) for Δ on the untransformed scale, computed separately from the median-regression coefficients above (Section 2.8). Family alone explained 4.6% of the variance in Δ, ligand descriptors alone 4.2%, eutomer potency alone 7.3%, and the full model 17.8%, leaving 82.2% unexplained. Δ was dominated by pair-specific variation that none of these variables captured.

3.12 A reproducible target-associated eudysmic index

Aggregating Δ to the protein gave a target-level eudysmic index for 224 targets with at least 20 pairs, covering 83.1% of all family-assigned pairs (median interval width 0.39 log units). The index spanned a wide range: at the top, mitogen-activated protein kinase 6 (2.05, n = 33), coagulation factor XI (1.85, n = 107), P2X purinoceptor 7 (1.84, n = 68), NMDA receptor subunit 2B (1.79, n = 37), BACE1 (1.75, n = 173) and BACE2 (1.61, n = 21); at the bottom, several kinases and a deubiquitinase with an index of 0.00 (protein kinase C ι, n = 46; USP28, n = 22; GCN2, n = 43; ATR, n = 33), where the median pair shows no measurable enantiomeric preference at all.

The index behaved as a reproducible quantity. Splitting each target's pairs at random into halves and correlating the two halves across targets gave a Spearman correlation of 0.755 [0.711, 0.801], which corrects to 0.86 at full length. The demanding version of the test partitions each target's pairs by chemical series, so that the two halves share no Bemis–Murcko scaffold; there the correlation was 0.633 [0.584, 0.688], correcting to 0.775 (Figure 7b). An index estimated from one set of chemical series therefore predicts the index estimated from a disjoint set of series against the same protein, which is the operational meaning of a target-associated rather than a series-associated quantity. Reliability across series does not by itself establish an intrinsic property of the protein: stereocenter position relative to the pharmacophore, assay design, and the research programme behind a given target's compound collection are shared across that target's series and remain uncontrolled (Section 4). Consistent with this, a random-intercept model places 26.5% of the total variance in Δ between targets.

The 20-pair minimum is not what produces this. Raising it to 30 pairs (159 targets) or 50

pairs (91 targets) leaves the reliability essentially unchanged — random split-half 0.755, 0.743 and 0.792 at minima of 20, 30 and 50, and scaffold-disjoint 0.633, 0.601 and 0.652 — so the index is no more reproducible when estimated only from the best-sampled targets, which is what would be expected if the 20-pair result were being carried by sampling noise. The target-weighted family comparison is significant at every minimum ($p = 1.2 \times 10^{-10}$, 1.4×10^{-8} and 4.0×10^{-7} at minima of 10, 20 and 30). One ordering is threshold-dependent and is reported rather than suppressed: at a minimum of 20 pairs nuclear receptor (0.78) exceeds protease (0.657), while at a minimum of 30 the two invert (0.70 versus 0.76). This is the same nuclear-receptor/protease boundary that Section 3.3 reports as indeterminate after scaffold collapse, and it should be read as one instability observed under three different analyses, not as three separate findings.

The index was not an artifact of data density: across the 224 targets it was uncorrelated with the number of pairs ($\rho = 0.087$, $p = 0.19$) and with the number of distinct scaffolds ($\rho = 0.095$, $p = 0.16$). It did correlate weakly with median eutomer potency ($\rho = 0.218$, $p = 0.001$), the target-level form of Pfeiffer's rule.

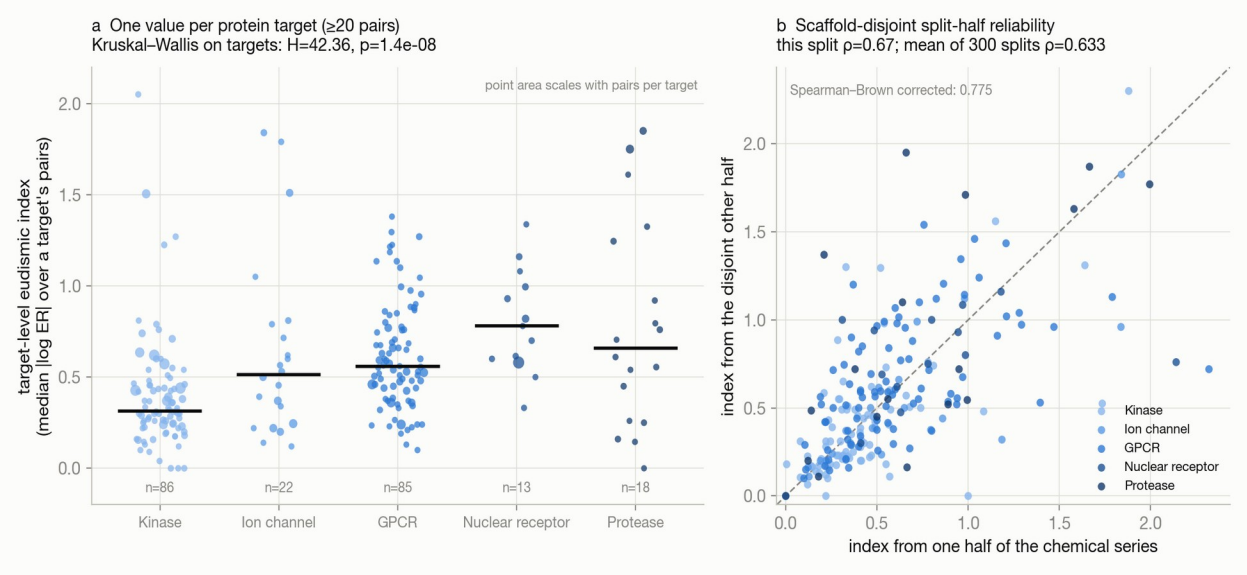

Figure 7. (a) The target-level eudysmic index for the 224 targets with at least 20 enantiomer pairs, one point per target, area scaled by pair count; horizontal bars are family medians. (b) Scaffold-disjoint split-half reliability: the index computed from one half of a target's chemical series against the index computed from the disjoint other half.

Two consequences follow for the family analysis. First, family is a coarse surrogate for the protein: of the between-target variance, family explains only 12.3%, so most of the reproducible protein-level signal is target-specific and not captured by family membership. Second, when each target counts once instead of each pair, the family comparison was still significant (Kruskal–Wallis $H = 42.4$, $p = 1.4 \times 10^{-8}$; Figure 7a) and

the ordering was broadly preserved, but with 13–86 targets per family only three pairwise contrasts survived Holm correction — kinase versus GPCR, versus nuclear receptor, and versus protease. Two differences from the pair-weighted analysis are worth stating plainly. Ion channel rose to 0.515, level with GPCR (0.56) rather than with kinase (0.315), reproducing the instability seen in the cross-check of Section 3.10; and protease fell to 0.657, below nuclear receptor (0.78), because its pair-level median was carried by a few intensively studied and highly discriminating targets such as BACE1 and factor XI. Protease's position at the top of the pair-weighted ordering is thus a statement about the protease pairs in ChEMBL, not about the typical protease.

3.13 Target-family information carried by the eudysmic ratio

The association is reproducible but weak in information terms, and it is worth stating that bound explicitly. The entropy of the family label was 2.002 bits. The mutual information between Δ and family was 0.057 bits — 2.9% of that entropy — against a permutation null of 0.003 (95th percentile 0.013, $p < 0.005$). Under scaffold-grouped cross-validation, a random forest using Δ alone reached a balanced accuracy of 0.226 and a macro one-vs-rest AUC of 0.571, against a chance balanced accuracy of 0.200 and a chance AUC of 0.500; it never predicted two of the five classes. Adding eutomer potency raised this to 0.261 and 0.639.

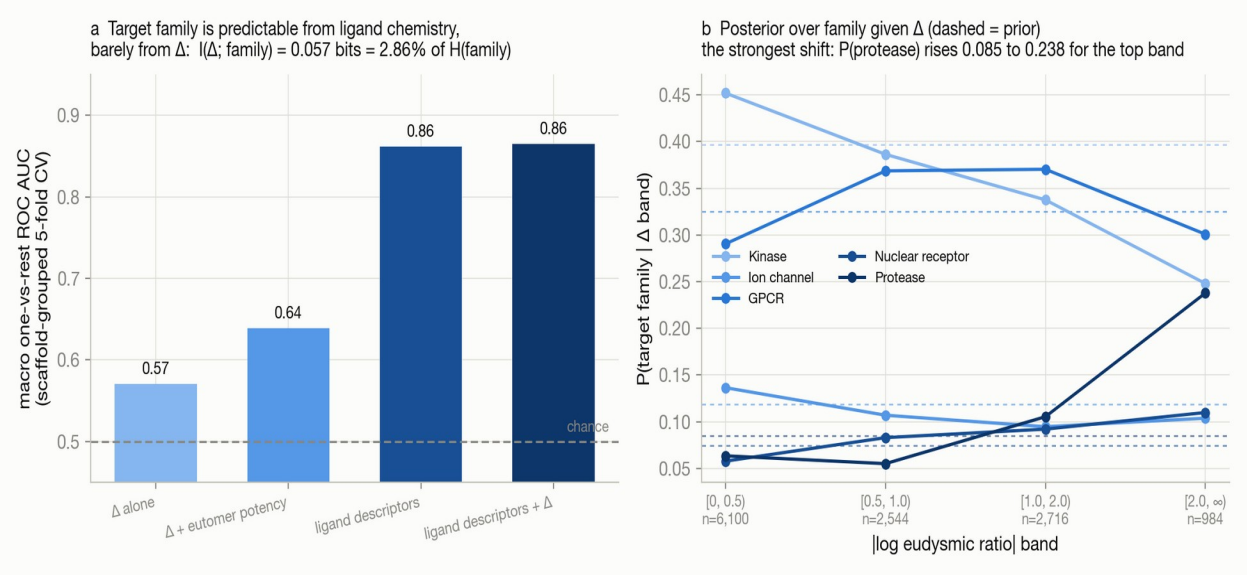

Figure 8. (a) Macro one-vs-rest AUC for predicting target family under scaffold-grouped 5-fold cross-validation, from Δ alone, from Δ with eutomer potency, from ligand descriptors alone, and from both. (b) Posterior probability of each family given the Δ band (dashed lines, prior).

The contrast with ligand chemistry is stark. The same model given ligand descriptors alone reached a balanced accuracy of 0.443 and a macro AUC of 0.862 — target family is

strongly encoded in the medicinal chemistry of the compounds, which is a restatement of Section 3.11's confound rather than a discovery about stereochemistry. Adding Δ to those descriptors improved the macro AUC by 0.004 and the balanced accuracy by 0.010 (Figure 8a). Practically, then, Δ carries almost no incremental information about which family a compound was tested against.

The Bayesian form of the same statement is the most interpretable one (Figure 8b). Among pairs with Δ > 2 — a greater-than-100-fold enantiomeric difference — the probability that the target was a protease rose from a prior of 0.085 to 0.238, while the probability that it was a kinase fell from 0.397 to 0.248. Two distinct quantities describe that shift and should not be conflated: the probability enrichment is 0.238/0.085 = 2.8-fold, whereas the odds ratio — equivalently the Bayes factor, (0.238/0.762)/(0.085/0.915) — is 3.4. A very large eudysmic ratio therefore does shift belief about the target class, multiplying the prior probability of a protease by 2.8 and its prior odds by 3.4, but it leaves 76% of such pairs belonging to other families. This is therefore, at best, a useful nudge and should not be considered a classifier.

3.14 Censoring: the distribution is truncated and the reported values are conservative

The requirement that both enantiomers carry a pChEMBL value means that pairs in which the weaker enantiomer was reported only as a bound — for example "IC50 > 10,000 nM" — are excluded, because ChEMBL assigns a pChEMBL value only to exact measurements. Assumption 2.10.9 above is that such pairs are missing at random with respect to Δ. They are not: the weaker enantiomer is reported as a bound precisely when it is too weak to quantify, which is when Δ is largest.

Searching the same assays for pairs in which one enantiomer has an exact value and the other only a "greater-than" bound yielded 780 such pairs against 12,335 fully exact pairs (5.95%). Their implied lower bounds on Δ had a median of 1.22, against a median of 0.50 for the exact pairs — the excluded pairs were indeed the highly stereoselective ones. Reinstating them at their lower bounds, which can only understate their true values, raised the overall median from 0.50 to 0.53, the fraction exceeding 10-fold from 28.8% to 30.5% and the fraction exceeding 100-fold from 7.8% to 8.6%.

Crucially, the censoring rate differed by family ($\chi^2 = 66.5$, $p = 1.3 \times 10^{-13}$), and it was highest for protease (10.2%) and lowest for kinase (4.5%) and nuclear receptor (4.7%). The family most affected by the truncation was thus the one with the largest observed Δ, so correcting the bias widened rather than narrowed the protease–kinase gap: family medians with censored pairs reinstated at their bounds were kinase 0.41, ion channel 0.42, GPCR 0.63, nuclear receptor 0.80 and protease 1.02. With respect to this recoverable form of censoring — one enantiomer exact, the other reported as a ">" bound in the same assay — the reported medians and tail fractions are therefore conservative, as is the protease–kinase gap. That direction is established for those pairs only. Pairs excluded for other reasons — both enantiomers censored, or one enantiomer never

assayed alongside the other — could not be recovered, and their effect on the distribution is unknown. The direction of the censoring bias on the effect sizes and on the information-theoretic and reliability statistics was not determined.

4. Discussion

At archive scale, the size of the enantiomeric potency difference among assayed pairs is associated with the target family the pair was tested against, in a reproducible monotone ordering: pairs assayed against proteases and nuclear receptors show the largest differences, GPCRs intermediate, kinases and ion channels the smallest, with kinase and ion channel statistically equivalent and the nuclear-receptor/protease boundary unresolved after scaffold collapse.

Because the eudysmic ratio is generated by a chiral ligand meeting a chiral pocket, differences of this kind are most naturally read as differences in how sharply the *protein* discriminates stereochemical information, with ligand chemistry contributing to the measured value. Three results support that reading without establishing it. The family effect survived adjustment for the ligand chemistry that differs so markedly between families, and propensity matching on ligand descriptors alone did not shrink it (Section 3.11); the target-level index reproduced across disjoint chemical series at $\rho = 0.63$ (Section 3.12), which is difficult to explain if the quantity were mainly a property of the compounds; and target identity accounted for 26.5% of the variance in Δ. It is tempting to go further and read the ordering as pocket architecture — deep, enclosed protease and nuclear-receptor sites making more directional, configuration-sensitive contacts than the conserved hinge region of the kinase ATP site — but no structural variable was measured here, and the position of the inverted stereocenter relative to the pharmacophore was not controlled. The mechanistic reading is stated as a hypothesis in Section 5, not as a result.

Protein family is a surrogate for the underlying pocket properties, and Section 3.12 quantifies how coarse a surrogate it is: family accounts for only 12.3% of the reproducible between-target variance in stereochemical discrimination. Most of the protein-level signal is target-specific. A pocket-descriptor regression is the natural next step and is not attempted here, for two reasons. It would require structures for a large fraction of the 224 indexed targets and a pocket-detection pipeline, which is a study in itself; and the one published attempt to relate binding-site descriptors to a closely analogous quantity — activity-cliff propensity among kinase inhibitors — reports that pocket-only models performed near chance and concludes that the whole protein matrix, not the site, is controlling (12). Attempting a shallow version of that analysis here would be more likely to mislead than to inform.

Several features distinguish the finding from a restatement of intuition. It is anchored by removing the between-source component of measurement dispersion (16, 17): the pairs are 100% exact-relation, same-activity-type, same-assay comparisons. Six alternative explanations were tested — stereocenter count, congeneric-series clustering, assay type, selection bias,

ligand chemistry and censoring. The elevation of protease and nuclear receptor above kinase and ion channel persisted under all six, with the selection-bias and ligand-chemistry tests the most informative and the censoring analysis showing that the reported values are conservative. The finer distinctions did not persist uniformly: the full five-way ordering is not preserved in the higher-stereocenter strata (Section 3.4), the nuclear-receptor/protease contrast becomes indeterminate after scaffold collapse (Section 3.3), and the GPCR/ion-channel/kinase separation does not survive in large screening panels (Section 3.9). Pfeiffer's rule (5) holds internally while failing to explain the between-family pattern. And the effect is bounded honestly in information terms: 0.057 bits, 2.9% of the family entropy.

The findings refine rather than echo prior work. Chiral Cliffs (10) established that chirality-driven cliffs exist and are learnable; the present work adds the continuous, family-stratified distribution with a measurement anchor. Work on kinase cliffs (12) related whole-protein properties to cliff propensity within kinases; kinases here sit at the low-discrimination end of a five-family ordering, using a strictly stereochemical transformation that standard matched-molecular-pair analyses (13) do not treat. The enantiomeric-pair ADME study (15) benchmarked enantiomer differences against measurement noise for off-target assays; the same is done here for on-target potency and across pocket families. Relative to classical eudysmic analysis (2–4, 9), which compared indices between two or a few targets, the contribution is the estimator and the scale rather than the concept.

Limitations bound the interpretation. (i) This is an association, not causation; the pocket reading is a hypothesis. (ii) No unbiased population rate of stereoselectivity can be claimed: absolute numbers are conditioned on the decision to make and test both enantiomers, so they describe assayed pairs. (iii) Both high-ranked families are target-concentrated — nuclear receptor because ROR-γ supplies 46.7% of its pairs, protease because its per-target medians are highly dispersed and its pair-level median is carried by BACE1 and factor XI; when targets are weighted equally, protease falls below nuclear receptor. (iv) The position of ion channel is unstable, sitting with kinase in the pair-weighted primary analysis but with GPCR in both the unfiltered cross-check and the target-weighted analysis. (v) Stereocenter count was controlled, and a broad panel of ligand descriptors, but not the position of the inverted stereocenter relative to the pharmacophore. (vi) Effect sizes are modest and distributions overlap, so the ordering is a shift in central tendency and tail weight, not a rule for individual compounds. (vii) The same-assay constraint removes the between-source component of measurement dispersion but not intra-assay error, which was not separately quantified. (viii) Only tetrahedral stereocenters (InChI `/t`) are treated; axial and planar chirality are excluded. (ix) The large-panel selection-bias result, though the strongest evidence against cherry-picking, rests substantially on BACE1 and ROR-γ screening data. None of these undermines the central measurement; they define its scope.

5. Perspectives

The results above are empirical measurements.

This section outlines the hypotheses they motivate, but do not establish, and identifies the evidence needed to test each one. None of these hypotheses is presented as a conclusion of this study.

5.1 Binding-pocket architecture

The most direct hypothesis is that stereochemical discrimination tracks pocket geometry: depth, enclosure, buriedness, hydrogen-bond density, contact density and flexibility. Deep, enclosed sites should constrain ligand orientation more tightly and so penalize an inverted configuration more heavily than shallow, solvent-exposed ones. Testing this means computing pocket descriptors — with tools such as fpocket (35) or druggability models of the kind used to characterize binding sites (36, 37) — for the targets with a structure and a well-estimated index, and regressing the index on them. The prior on success should be tempered: the analogous attempt for activity-cliff propensity in kinases found pocket-only descriptors near chance (12). A negative result would itself be informative, pointing towards whole-protein or dynamic determinants.

5.2 Multiple binding modes

A low index has at least three distinct explanations that the present data cannot separate: the pocket may be genuinely permissive; stereochemistry may sit away from the pharmacophore in the particular series assayed; or the two enantiomers may adopt different but energetically comparable poses. The targets identified here with an index of essentially zero across dozens of pairs are natural candidates for co-crystallography or molecular dynamics on both enantiomers, which would distinguish these possibilities directly.

5.3 Evolutionary determinants

Proteases recognize peptide substrates through stereochemically defined subsites (38), and the classical demonstration that chemically synthesized D- and L-HIV-1 protease show reciprocal chiral substrate specificity (39) is the sharpest statement that protease recognition is stereochemically strict. Kinases, by contrast, evolved around a conserved ATP site whose hinge-binding requirement is satisfied by largely planar heteroaromatic scaffolds (40). The hypothesis is that a family's typical discrimination reflects the stereochemical strictness of its natural ligand recognition. It could be tested by relating the target-level index to substrate stereospecificity measured independently, or to residue conservation in the site.

5.4 Druggability and optimization

If a protein discriminates configuration sharply, then configuration is a productive optimization axis for that target, and one would expect high-index targets to yield more selective ligands per unit of medicinal chemistry effort. This predicts measurable relationships between the index and hit rates, optimization trajectories or achievable selectivity, which could be tested against project-level data that this archive-scale analysis does not contain.

5.5 Ligand promiscuity

A pocket indifferent to configuration might be indifferent along other axes too, in which case the target-level index would correlate

negatively with the diversity of scaffolds a target binds. Promiscuity has been quantified from bioactivity archives before (41, 42), hence this hypothesis is testable with data of the kind used here, and it is the most immediately tractable of the six.

5.6 Resistance and clinical pharmacology

Resistance mutations that remodel a drug-binding site provide a direct perturbation with which to test the structural basis of stereochemical discrimination (43, 44). Two related proteins with similar wild-type eudysmic indices may respond very differently to such mutations: the eudysmic ratio may collapse in one but remain comparatively stable in another, indicating that similar baseline levels of stereochemical discrimination can arise from interactions with different susceptibility to structural perturbation. These changes should be interpreted together with the corresponding loss of eutomer potency, because a change in the ratio alone cannot distinguish selective disruption of stereochemical recognition from a general loss of binding. The particularly informative comparison is therefore whether resistance-associated mutations impose different potency penalties on eutomers whose wild-type targets initially show similar stereochemical discrimination, and whether those differences can be traced to specific stereochemical interactions within the binding site. Such comparisons could reveal binding geometries in which the eutomeric interaction is less vulnerable to resistance-associated structural change. If those features are generalizable, they could suggest a strategy for stereochemical optimization in which ligand geometry is designed to reduce the potency loss caused by likely resistance mutations, thereby raising the mutational barrier to resistance rather than preventing resistance altogether. Testing this hypothesis would require matched wild-type and resistance-mutant potency measurements for both enantiomers across related targets and chemical series; such data are not available in the present analysis, so this remains a prospective hypothesis rather than a conclusion of the study.

A general caution applies to all six. The index reported here has a median 95% interval width of 0.39 log units at 20 pairs per target, and family membership explains only 12.3% of between-target variance. Studies built on it will need targets with substantially more data than the minimum used here, and should treat the index as a noisy estimate rather than a fixed protein constant.

6. Conclusion

Using strictly defined same-assay enantiomer pairs from ChEMBL 37, the first target-family-stratified distribution of the eudysmic ratio built under a same-assay measurement constraint was constructed, and the size of the enantiomeric potency difference was shown to be associated with binding-pocket family in a monotone ordering: protease and nuclear receptor above GPCR, above kinase and ion channel, the last two statistically equivalent. The pattern is anchored by removing the between-source component of measurement dispersion, and robust to controls for stereocenter count, congeneric-series clustering, assay type, selection bias and — most importantly for interpretation — the

medicinal chemistry of the ligand sets, which differs sharply between families yet accounts for almost none of the effect. Aggregating to the protein gives a target-level eudysmic index that reproduces across disjoint chemical series, hence stereochemical discrimination behaves as a measurable, reproducible target-associated characteristic — though not, on this evidence, an established intrinsic property of the protein, since stereocenter position, assay design and target-specific research programmes remain uncontrolled; family, in turn, is only a coarse surrogate for it, capturing 12.3% of the between-target variance. The ratio carries little information about the family in any individual case (0.057 bits, 2.9% of the family entropy), and the reported values are conservative with respect to the censored pairs that could be recovered, because the requirement of an exact value for both enantiomers truncates the distribution from above; pairs excluded for other reasons were not recovered and their effect is unknown. The eudysmic ratios themselves are measured quantities; the family-level pattern is an association whose assumptions and limits have been stated. The extraction and analysis pipeline is openly available at https://github.com/gnarlylace51842/eudysmic-ratio-target-family so the distribution can be regenerated and extended as ChEMBL grows.

One implication of these findings is that the target-level eudysmic index may provide a useful framework for studying mutational drug resistance. Proteins with similar wild-type eudysmic indices need not respond similarly to resistance-associated mutations: stereochemical discrimination may collapse in one target but remain comparatively stable in another, revealing differences in how the interactions responsible for enantiomer recognition respond to structural perturbation. The informative quantity would not be the change in eudysmic ratio alone, but that change considered together with the corresponding loss of eutomer potency. Systematic comparison of matched wild-type and mutant proteins—especially among related targets with similar starting eudysmic indices—could therefore determine whether comparable wild-type stereochemical discrimination is associated with different susceptibility to resistance mutations and identify the structural interactions responsible for those differences. In particular, resistance mutations that produce a smaller potency penalty for one stereochemical binding mode than another could reveal ligand geometries that are less vulnerable to mutational disruption. If such features prove generalizable, stereochemical optimization could potentially be used to reduce the potency loss produced by resistance-associated mutations or increase the structural change required for the target to escape inhibition, thereby raising the mutational barrier to resistance. This possibility is not tested by the present data, but follows as a testable hypothesis from the reproducible target-associated differences observed here.

Acknowledgments

The author thanks the reviewers for comments that materially improved this work, in particular the recommendation to reframe the measurement as a property of the protein–ligand interaction and to adjust for ligand chemistry, which led to the analyses in Sections 3.11–3.13, and the request to state all

assumptions explicitly, which led to Section 2.10 and to the censoring analysis in Section 3.14.

The author declares the following assistance in the preparation of this manuscript. An AI assistant (Claude, Anthropic) was used for technical and analytical help. It helped with generating figures, organizing the data / scripts, and aiding with handling the computation. The author conceived the study, specified the analyses, verified the outputs against the source data, and is responsible for the content of the manuscript, including all interpretations and conclusions. No other editorial, technical, analytical or writing support was received. No funding was received for this work.

Code and data availability

All analysis used the public ChEMBL 37 SQLite release (18). The complete extraction and analysis pipeline (enantiomer detection, same-assay pairing, family mapping, ligand descriptors, the target-level index and its reliability tests, the multivariable adjustment, the information analysis, the censoring analysis, the equivalence-margin sensitivity analysis, statistics and figures) and the derived pair-level and target-level datasets are openly available at https://github.com/gnarlylace51842/eudysmic-ratio-target-family . All results can be regenerated from the public database with the released code.

7. References

1. Brooks, W. H., Guida, W. C., & Daniel, K. G. (2011). The significance of chirality in drug design and development. Current Topics in Medicinal Chemistry, 11(7), 760–770. https://doi.org/10.2174/156802611795165098

2. McVicker, R. U., & O'Boyle, N. M. (2024). Chirality of new drug approvals (2013–2022): Trends and perspectives. Journal of Medicinal Chemistry, 67(4), 2305–2320. https://doi.org/10.1021/acs.jmedchem.3c02239

3. Lehmann, F. P. A., Rodrigues de Miranda, J. F., & Ariëns, E. J. (1976). Stereoselectivity and affinity in molecular pharmacology. In E. Jucker (Ed.), Progress in Drug Research (Vol. 20, pp. 101–142). Birkhäuser. https://doi.org/10.1007/978-3-0348-7094-8_4

4. Lehmann, F. P. A. (1982). Quantifying stereoselectivity or how to choose a pair of shoes when you have two left feet. Trends in Pharmacological Sciences, 3, 103–106. https://doi.org/10.1016/0165-6147(82)91043-4

5. Pfeiffer, C. C. (1956). Optical isomerism and pharmacological action, a generalization. Science, 124(3210), 29–31. https://doi.org/10.1126/science.124.3210.29

6. Barlow, R. (1990). Enantiomers: How valid is Pfeiffer's rule? Trends in Pharmacological Sciences, 11(4), 148–150. https://doi.org/10.1016/0165-6147(90)90065-G

7. Lehmann, F. P. A. (1986). Stereoisomerism and drug action. Trends in Pharmacological Sciences, 7, 281–285. https://doi.org/10.1016/0165-6147(86)90353-6

8. Chen, C.-S., Fujimoto, Y., Girdaukas, G., & Sih, C. J. (1982). Quantitative analyses of biochemical kinetic resolutions of enantiomers. Journal of the American Chemical Society, 104(25), 7294–7299. https://doi.org/10.1021/ja00389a064

9. El Tayar, N., Testa, B., van de Waterbeemd, H., Carrupt, P. A., & Kaumann, A. J. (1988). Influence of lipophilicity and chirality on the selectivity of ligands for β1- and β2-adrenoceptors. Journal of Pharmacy and Pharmacology, 40(9), 609–612. https://doi.org/10.1111/j.2042-7158.1988.tb05319.x

10. Schneider, N., Lewis, R. A., Fechner, N., & Ertl, P. (2018). Chiral cliffs: Investigating the influence of chirality on binding affinity. ChemMedChem, 13(13), 1315–1324. https://doi.org/10.1002/cmdc.201700798

11. Saha, D., Kharbanda, A., Yan, W., Lakkaniga, N. R., Frett, B., & Li, H. Y. (2020). The exploration of chirality for improved druggability within the human kinome. Journal of Medicinal Chemistry, 63(2), 441–469. https://doi.org/10.1021/acs.jmedchem.9b00640

12. Daoud, S., & Taha, M. (2024). Protein characteristics substantially influence the propensity of activity cliffs among kinase inhibitors. Scientific Reports, 14, 9058. https://doi.org/10.1038/s41598-024-59501-w

13. Dalke, A., Hert, J., & Kramer, C. (2018). mmpdb: An open-source matched molecular pair platform for large multiproperty data sets. Journal of Chemical Information and Modeling, 58(5), 902–910. https://doi.org/10.1021/acs.jcim.8b00173

14. Comajuncosa-Creus, A., Lenes, A., Sánchez-Palomino, M., Dalton, D., & Aloy, P. (2024). Stereochemically-aware bioactivity descriptors for uncharacterized chemical compounds. Journal of Cheminformatics, 16, 70. https://doi.org/10.1186/s13321-024-00867-4

15. Bagdanoff, J. T., Xu, Y., Dollinger, G., & Martin, E. (2015). Effect of chirality on common in vitro experiments: An enantiomeric pair analysis. Journal of Medicinal Chemistry, 58(15), 5781–5788. https://doi.org/10.1021/acs.jmedchem.5b00552

16. Kalliokoski, T., Kramer, C., Vulpetti, A., & Gedeck, P. (2013). Comparability of mixed IC50 data – A statistical analysis. PLoS ONE, 8(4), e61007. https://doi.org/10.1371/journal.pone.0061007

17. Landrum, G. A., & Riniker, S. (2024). Combining IC50 or Ki values from different sources is a source of significant noise. Journal of Chemical Information and Modeling, 64(5), 1560–1567. https://doi.org/10.1021/acs.jcim.4c00049

18. Zdrazil, B., Felix, E., Hunter, F., Manners, E. J., Blackshaw, J., Corbett, S., de Veij, M., Ioannidis, H., Mendez Lopez, D., Mosquera, J. F., Magarinos, M. P., Bosc, N., Arcila, R., Kizilören, T., Gaulton, A., Bento, A. P., Adasme, M. F., Monecke, P., Landrum, G. A., & Leach, A. R. (2024). The ChEMBL database in 2023: A drug discovery platform spanning multiple bioactivity data types and time periods. Nucleic Acids Research, 52(D1), D1180–D1192. https://doi.org/10.1093/nar/gkad1004

19. Heller, S. R., McNaught, A., Pletnev, I., Stein, S., & Tchekhovskoi, D. (2015). InChI, the IUPAC International Chemical Identifier. Journal of Cheminformatics, 7, 23. https://doi.org/10.1186/s13321-015-0068-4

20. RDKit. (n.d.). RDKit: Open-source cheminformatics (Version 2026.03) [Computer software]. https://www.rdkit.org/

21. Bemis, G. W., & Murcko, M. A. (1996). The properties of known drugs. 1. Molecular frameworks. Journal of Medicinal Chemistry, 39(15), 2887–2893. https://doi.org/10.1021/jm9602928

22. Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. Biometrics, 38(4), 963–974. https://doi.org/10.2307/2529876

23. Nakagawa, S., Johnson, P. C. D., & Schielzeth, H. (2017). The coefficient of determination R^2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. Journal of the Royal Society Interface, 14(134), 20170213. https://doi.org/10.1098/rsif.2017.0213

24. Holm, S. (1979). A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics, 6(2), 65–70. https://www.jstor.org/stable/4615733

25. Wilson, E. B. (1927). Probable inference, the law of succession, and statistical inference. Journal of the American Statistical Association, 22(158), 209–212. https://doi.org/10.1080/01621459.1927.10502953

26. Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. Nature Methods, 17(3), 261–272. https://doi.org/10.1038/s41592-019-0686-2

27. Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and statistical modeling with Python. In Proceedings of the 9th Python in Science Conference (pp. 92–96). https://doi.org/10.25080/Majora-92bf1922-011

28. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825–2830. https://www.jmlr.org/papers/v12/pedregosa11a.html

29. Koenker, R., & Bassett, G. (1978). Regression quantiles. Econometrica, 46(1), 33–50. https://doi.org/10.2307/1913643

30. Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. Journal of Human Resources, 50(2), 317–372. https://doi.org/10.3368/jhr.50.2.317

31. Ross, B. C. (2014). Mutual information between discrete and continuous data sets. PLoS ONE, 9(2), e87357. https://doi.org/10.1371/journal.pone.0087357

32. Kraskov, A., Stögbauer, H., & Grassberger, P. (2004). Estimating mutual information. Physical Review E, 69(6), 066138. https://doi.org/10.1103/PhysRevE.69.066138

33. Wallach, I., & Heifets, A. (2018). Most ligand-based classification benchmarks reward memorization rather than generalization. Journal of Chemical Information and Modeling, 58(5), 916–932. https://doi.org/10.1021/acs.jcim.7b00403

34. van Tilborg, D., Alenicheva, A., & Grisoni, F. (2022). Exposing the limitations of molecular machine learning with activity cliffs. Journal of Chemical Information and Modeling, 62(23), 5938–5951. https://doi.org/10.1021/acs.jcim.2c01073

35. Le Guilloux, V., Schmidtke, P., & Tuffery, P. (2009). Fpocket: An open source platform for ligand pocket detection. BMC Bioinformatics, 10, 168. https://doi.org/10.1186/1471-2105-10-168

36. Halgren, T. A. (2009). Identifying and characterizing binding sites and assessing druggability. Journal of Chemical Information and Modeling, 49(2), 377–389. https://doi.org/10.1021/ci800324m

37. Schmidtke, P., & Barril, X. (2010). Understanding and predicting druggability. A high-throughput method for detection of drug binding sites. Journal of Medicinal Chemistry, 53(15), 5858–5867. https://doi.org/10.1021/jm100574m

38. Schechter, I., & Berger, A. (1967). On the size of the active site in proteases. I. Papain. Biochemical and Biophysical Research Communications, 27(2), 157–162. https://doi.org/10.1016/S0006-291X(67)80055-X

39. Milton, R. C. de L., Milton, S. C. F., & Kent, S. B. H. (1992). Total chemical synthesis of a D-enzyme: The enantiomers of HIV-1 protease show reciprocal chiral substrate specificity. Science, 256(5062), 1445–1448. https://doi.org/10.1126/science.1604320

40. Zhang, J., Yang, P. L., & Gray, N. S. (2009). Targeting cancer with small molecule kinase inhibitors. Nature Reviews Cancer, 9(1), 28–39. https://doi.org/10.1038/nrc2559

41. Paolini, G. V., Shapland, R. H. B., van Hoorn, W. P., Mason, J. S., & Hopkins, A. L. (2006). Global mapping of pharmacological space. Nature Biotechnology, 24(7), 805–815. https://doi.org/10.1038/nbt1228

42. Hu, Y., & Bajorath, J. (2012). Growth of ligand-target interaction data in ChEMBL is associated with increasing and activity measurement-dependent compound promiscuity. Journal of Chemical Information and Modeling, 52(10), 2550–2558. https://doi.org/10.1021/ci3003304

43. Gorre, M. E., Mohammed, M., Ellwood, K., Hsu, N., Paquette, R., Rao, P. N., & Sawyers, C. L. (2001). Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification. Science, 293(5531), 876–880. https://doi.org/10.1126/science.1062538

44. Ali, A., Bandaranayake, R. M., Cai, Y., King, N. M., Kolli, M., Mittal, S., Murzycki, J. F., Nalam, M. N. L., Nalivaika, E. A., Özen, A., Prabu-Jeyabalan, M. M., Thayer, K., & Schiffer, C. A. (2010). Molecular basis for drug resistance in HIV-1 protease. Viruses, 2(11), 2509–2535. https://doi.org/10.3390/v2112509