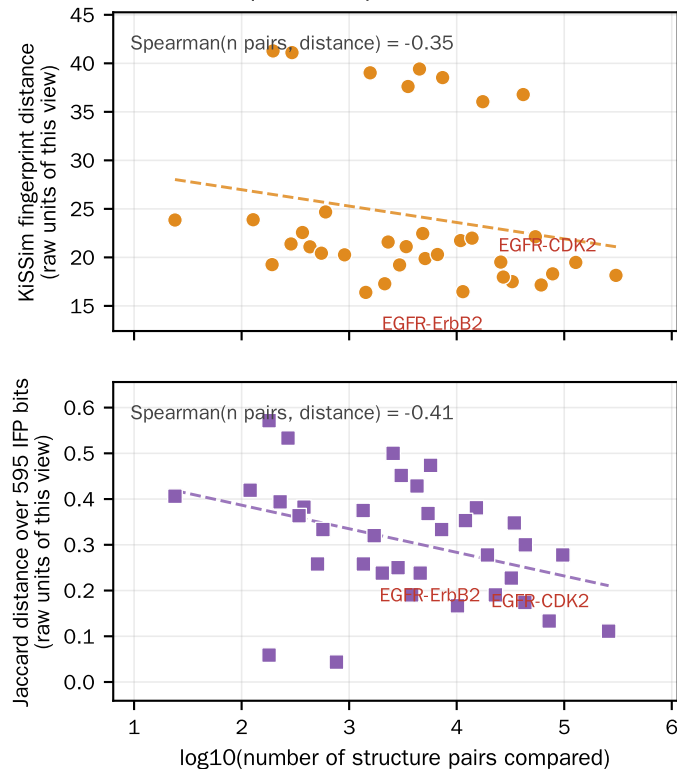
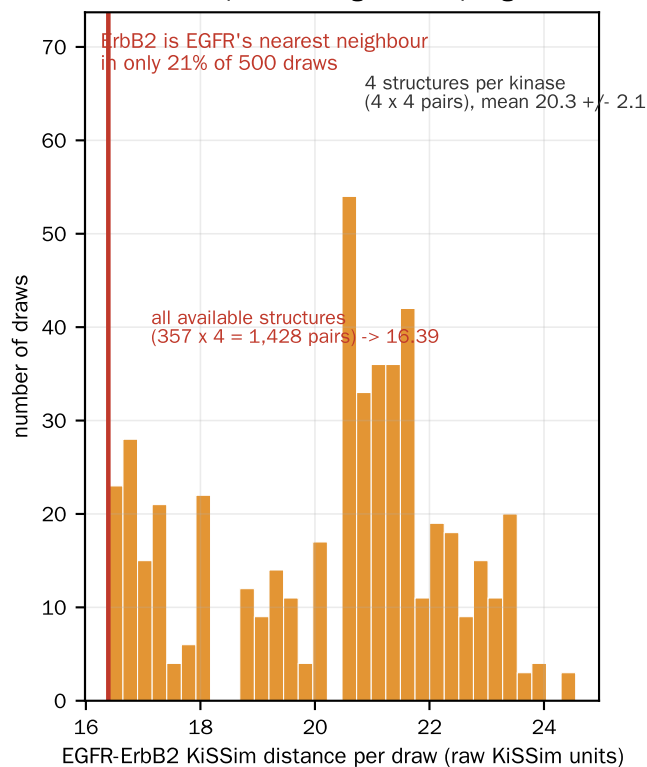


Why the structure-derived views can disagree with the rest

A More pairs compared, smaller distance



B Equal-coverage resampling



C Rank relative to EGFR (1 = closest)

	T024	T025	T026	T027	combined
ErbB2	1	1	6	1	1
LCK	3	2	2	5	2
VEGFR2	2	3	5	3	3
MET	4	4	4	6	4
BRAF	5	6	7	2	5
p38a	7	7	3	4	6
CDK2	6	5	1	7	7
PI3K	8	8	8	8	8

Panels A1/A2 keep each structure-derived view in its OWN units (KiSSim distance ~16-41; Jaccard distance 0-1) on separate y-axes. Both reduce a kinase pair to the MINIMUM over available structure pairs, and both show the same behaviour: more pairs compared, smaller reported distance. The two views do NOT share structure counts: EGFR x CDK2 vs EGFR x ErbB2 is 357 x 848 = 302,736 vs 357 x 4 = 1,428 pairs for KiSSim, and 339 x 763 = 258,657 vs 339 x 4 = 1,356 pairs for IFP. Panel B is a coverage-sensitivity test, not a real selectivity probability: with 4 structures per kinase, ErbB2 remains EGFR's nearest neighbour in only 21% of 500 draws. That is a KiSSim-only result: it shows the KiSSim ranking depends on which structures are available, and it does NOT quantify how much coverage contributes to the IFP ranking. Panel C ranks the other kinases relative to EGFR per view; in the combined column the order is ErbB2 1, LCK 2, VEGFR2 3, MET 4, BRAF 5, p38a 6, CDK2 7, PI3K 8.