Supplementary Data

Statistical comparison performance details

For each figure wherein statistical comparisons were made, below we describe the statistical testing (all performed in Graphpad Prism v10.2.3) and significance found. Supplementary Data includes full data, descriptive statistics (percentiles, mean, error, N), and statistical assessment outputs for each experiment.

Figure 1C: Flow cytometry assay. Original dataset was lost in a lab move. Average and confidence interval were in a processed data file, but statistics testing cannot be performed on this sample.

Figure 1D: Flow cytometry assay. All four TfR-stained samples were compared to one another by Kruskal-Wallis test with Dunn's correction; Med affn hTfR vs High affn hTfR was not significant ($P > 0.9999$), while all other comparisons were $P < 0.0001$.

Figure 1F: Flow cytometry assay. All non-Parental samples within a given pH (7.4 or 5.5) were compared to the respective Parental sample by Kruskal-Wallis test with Dunn's correction. At pH 7.4, His sub 1 ($P > 0.9999$), His sub 3 ($P = 0.1161$), and His sub 1+3 ($P > 0.9999$) were not significant, while all others were $P < 0.0001$. At pH 5.5, His sub 1 ($P = 0.1666$) and His sub 3 ($P > 0.9999$) were not significant, while all others were $P < 0.0001$.

Figure 2F: Flow cytometry assay. All three samples were compared to one another by Kruskal-Wallis test with Dunn's correction; CT-4212-1 vs CT-4212-3 was not significant ($P = 0.1647$), while the other two comparisons were $P < 0.0001$.

Figure 2G: Flow cytometry assay. All three samples were compared to one another by Kruskal-Wallis test with Dunn's correction; CT-4212-1 vs CT-4212-3 was $P = 0.0027$, while the other two comparisons were $P < 0.0001$.

Figure 2H: Flow cytometry assay. All three samples were compared to one another by Kruskal-Wallis test with Dunn's correction; CT-4212-1 vs CT-4212-3 was not significant ($P = 0.0908$), while the other two comparisons were $P < 0.0001$.

Figure 2I: Flow cytometry assay. All three samples were compared to one another by Kruskal-Wallis test with Dunn's correction; Untreated vs CT-4212-3 was $P = 0.0002$, while the other two comparisons were $P < 0.0001$.

Figure 2J: Flow cytometry assay. All three samples were compared to one another by Kruskal-Wallis test with Dunn's correction; all comparisons were $P < 0.0001$.

Figure 2K: Flow cytometry assay. All three samples were compared to one another by Kruskal-Wallis test with Dunn's correction; Untreated vs CT-4212-1 was not significant ($P = 0.3139$), while the other two comparisons were $P < 0.0001$.

Figure 3G: Flow cytometry assay. PBS vs CT-1212-1 pairwise comparisons for each quintile were all $P < 0.0001$ via two-tailed Kolmogorov-Smirnov test. Quintile 1 vs Quintile 3 was $P = 0.0062$, while all other comparisons were $P < 0.0001$.

Figure 3H: Flow cytometry assay. PBS vs CT-1212-1 pairwise comparisons at each time point were all $P < 0.0001$ via two-tailed Kolmogorov-Smirnov test. All four CT-1212-1 samples were compared to one another by Kruskal-Wallis test with Dunn's correction; 30 min vs Withdrawal was not significant ($P > 0.9999$), while all other comparisons were $P < 0.0001$.

Figure 3I: Flow cytometry assay. PBS vs CT-1212-1 pairwise comparisons at each time point were all $P < 0.0001$ via two-tailed Kolmogorov-Smirnov test. All four CT-1212-1 samples were compared to one another by Kruskal-Wallis test with Dunn's correction; 24 hr vs Withdrawal was $P = 0.0008$, while all other comparisons were $P < 0.0001$.

Figure 4B: Flow cytometry assay. Within each of the four cell lines (A549, H1975, H1650, H358), all five samples were compared to one another by Kruskal-Wallis test with Dunn's correction; all comparisons for all cell lines were $P < 0.0001$.

Figure 4C: Flow cytometry assay. All four samples were compared to one another by Kruskal-Wallis test with Dunn's correction; CT-1212-1 2 hr vs Cetuximab 2 hr was $P = 0.0001$, while all other comparisons were $P < 0.0001$. Furthermore, all four samples were compared via one sample two-tailed T test to a value of 0; Cetuximab Withdrawal was not significant ($P = 0.372$), while all others were $P < 0.0001$. Wilcoxon signed-rank test vs 0 was also performed for all four samples; all were $P < 0.0001$, though it should be noted that the Cetuximab Withdrawal sample has a negative median (-18.54), so this significant comparison should not be interpreted as the sample having >0 signal.

Figure 4E: Flow cytometry assay. All six treated samples were compared to Untreated by Kruskal-Wallis test with Dunn's correction; all comparisons were $P < 0.0001$. All three molecules had their 24 hr and Withdrawal samples compared to each other by Kolmogorov-Smirnov test; all three comparisons were $P < 0.0001$.

Figure 4F: Flow cytometry assay. All six treated samples were compared to Untreated by Kruskal-Wallis test with Dunn's correction; all comparisons were $P < 0.0001$. All three molecules had their 24 hr and Withdrawal samples compared to each other by Kolmogorov-Smirnov test; all three comparisons were $P < 0.0001$.

Figure 4G: Flow cytometry assay. Within all four cell lines, all samples were compared one another by Kruskal-Wallis test with Dunn's correction. In A549, 1 hr vs 1 day was $P = 0.0019$, and 1 hr vs 2 day was not significant ($P = 0.9911$); all other comparisons were $P < 0.0001$. In H1975, 1 day vs 2 day was not significant ($P = 0.8709$), 1 day vs 3 day was $P = 0.0004$, and 2 day vs 3 day was not significant ($P = 0.3246$); all other comparisons were $P < 0.0001$. In H1650, 1 hr vs 3 day was $P = 0.0188$ and 1 day vs 2 day was $P = 0.0069$; all other comparisons were $P < 0.0001$. In H358, 1 day vs 2 day was not significant ($P > 0.9999$); all other comparisons were $P < 0.0001$.

Figure 4H: Flow cytometry assay. Each treatment dose was compared to Untreated and to the other doses for that same molecule by Kruskal-Wallis test with Dunn's correction. For CT-1212-1, 2 nM and 50 nM were not significant ($P > 0.9999$); all other comparisons were $P < 0.0001$. For CT-5212-3, 10 nM vs 50 nM was not significant ($P = 0.4598$), 10 nM vs 200 nM was $P = 0.0202$, and 50 nM vs 200 nM was not significant; all other comparisons were $P < 0.0001$. For CT-6212-3, 2 nM vs 10 nM was $P = 0.0001$, 10 nM vs 50 nM was not significant, 10 nM vs 200 nM was not significant ($P = 0.6048$), and 50 nM vs 200 nM was not significant; all other comparisons were $P < 0.0001$.

Figure 5C: Flow cytometry assay. Each CYpHER treatment (CT-1212-1, CT-1112-1, CT-1211-1, CT-1111-1) underwent all-vs-all comparison within the treatment (including Untreated) by Kruskal-Wallis test. Within CT-1212-1, 2 nM vs 10 nM was not significant ($P = 0.1494$), while all other comparisons were $P < 0.0001$. Within CT-1112-1, 2 vs 10 nM was $P = 0.0031$, while all other comparisons were $P < 0.0001$. Within CT-1211-1, 2 vs 10 nM was $P = 0.0002$, while all other comparisons were $P < 0.0001$. Within CT-1111-1, all comparisons were $P < 0.0001$.

Figure 5G: Flow cytometry assay. Each CYpHER treatment group was compared to the No CYpHER treatment (data point represented by the 100% line) by Kruskal-Wallis test with Dunn's correction. All comparisons were $P < 0.0001$. Each CYpHER + holoTF treatment was separately compared to the CYpHER + no holoTF group (data point represented by the "No holoTF" line at ~21%) by Kruskal-Wallis test with Dunn's correction. The 100 nM holoTF treatment was not significant ($P = 0.9831$); all three other comparisons were $P < 0.0001$.

Figure 6C: Flow cytometry assay. For each of the four cell lines (H358, H1650, H1975, A549), the pre-treatment sample was compared to its respective Untreated sample via two-tailed Kolmogorov-Smirnov test. All four comparisons were $P < 0.0001$. The four pre-treatments were also compared all-vs-all to one another by Kruskal-Wallis test with Dunn's correction. All comparisons were $P < 0.0001$.

Figure 6D: Flow cytometry assay. Each sample was compared all-vs-all with the others by Kruskal-Wallis test with Dunn's correction. CT-6212-1 Pre-treat vs CT-5212-3 Pre-treat was not significant ($P > 0.9999$), while all other comparisons were $P < 0.0001$.

Figure 6E: Flow cytometry assay. Each MDA-MB-231 sample was compared all-vs-all with the others by Kruskal-Wallis test with Dunn's correction. All comparisons were $P < 0.0001$. Each H1650 sample was compared all-vs-all with the others by Kruskal-Wallis test with Dunn's correction. All comparisons were $P < 0.0001$.

Figure 7, D-H: CellTiter-Glo 2.0 viability assay. Each CT-1212-1 curve fit (asymmetric sigmoidal, 5PL) was compared pairwise to the comparator compounds with extra sum-of-squares F test ("Do the best-fit EC50 values differ between the selected comparator compound and CT-1212-1?"). In A431 (Panel D), CT-1212-1 vs cetuximab was $P = 0.0194$; CT-1212-1 vs gefitinib was $P = 0.0031$; and CT-1212-1 vs osimertinib was $P = 0.0011$. In H1975 (Panel E), the software could not compute an EC50 comparison for CT-1212-1 vs cetuximab or for cetuximab vs a fixed EC50 of 2.00 nM (the EC50 of CT-1212-1 in H1975 cells), but an alternative comparison, "Does one curve fit both data sets?", resulted in $P < 0.0001$; CT-1212-1 vs gefitinib was $P < 0.0001$; and CT-1212-1 vs osimertinib was not significant ($P = 0.2213$). In H1650 (Panel F), cetuximab was not effective at suppressing growth and was not compared to CT-1212-1; the other two compounds were $P < 0.0001$ vs CT-1212-1. In A549 (Panel G), cetuximab was not effective at suppressing growth and was not compared to CT-1212-1; the other two compounds were $P < 0.0001$ vs CT-1212-1. In H358 (Panel H), cetuximab was not effective at suppressing growth and was not compared to CT-1212-1; the software could not compute curve-vs-curve EC50 comparison, but an alternative comparison, "Does the EC50 differ from 0.22?" (the EC50 of CT-1212-1 in H358 cells), yielded $P < 0.0001$ for gefitinib; and CT-1212-1 vs osimertinib was $P < 0.0001$.

Figure 8B: ELISA. At each time point (10 time points), 3 mice were sampled and ELISA run in technical triplicate at each of two dilutions (1:200 or 1:1000). The reading that fit the linear range of the 10-point standard curve was used; if both fit, the lower dilution was used. The native SoftMax Pro software was used to interpolate concentrations for each sample based on the technical triplicate with default settings; if it could not reach a quantitative value, that mouse sample was not used. This resulted in an N of 2 for CT-1212-1 at 0.5 hours, 8 hrs, and 7 days; an N of 2 at 4 hours for CT-1211-1; and an N of 2 at 2 hours for CT-1232-1; all other time points for all compounds were N of 3 for PK parameter determination as described in the Methods.

Figure 8, D and E: Western blot densitometry. Within each sample (three tumors per compound/dose), Western blot EGFR:actin densitometry ratios for CYpHER-dosed animals were tested with pairwise two-tailed Welch's T test vs ratios of vehicle mice. In the top blot, CT-1212-1 vs vehicle was not significant ($P = 0.2293$), CT-1222-1 vs vehicle was $P = 0.04$, and CT-5212-3 trended toward significance at $P = 0.07$, while in the bottom blot, CT-1212-1 (450 µg) vs vehicle trended towards significance at $P = 0.107$, CT-1212-1 (150 µg) vs vehicle was not significant ($P = 0.5887$), and CT-1212-1 (50 µg) vs vehicle was not significant ($P = 0.3919$)

Figure 8G: Automated nuclei counting from Ki67 immunohistochemistry. The five treatment groups were compared to vehicle by Kruskal-Wallis test with Dunn's correction. CT-1212-1 450 µg vs vehicle was $P = 0.0192$, CT-1212-1 150 µg vs vehicle was $P = 0.0033$, and CT-1222-1 150 µg was $P = 0.0013$; CT-1212-1 50 µg and CT-5212-3 150 µg were not significant vs vehicle ($P > 0.9999$).

Supplementary Fig. 1B: Flow cytometry assay. All samples were compared all-vs-all by Kruskal-Wallis test with Dunn's correction. All comparisons were $P < 0.0001$.

Supplementary Fig. 2C: Flow cytometry assay. Samples were compared pairwise via two-tailed Kolmogorov-Smirnov test. The comparison was significant at $P < 0.0001$.

Supplementary Fig. 2D: Flow cytometry assay. Samples were compared pairwise via two-tailed Kolmogorov-Smirnov test. The comparison was not significant ($P = 0.4902$).

Supplementary Fig. 2G: Flow cytometry assay. Variants of Interest were defined as those that retained >10% of Parental EGFR binding at pH 7.4 as well as having lost >25% of EGFR binding at pH 5.5 compared to pH 7.4. For each Variant of Interest (His subs 6, 10, 13, and 15), significance was compared between pH 7.4 and pH 5.5 samples by pairwise two-tailed Kolmogorov-Smirnov test. His sub 6 was $P = 0.0236$ between pH 7.4 and pH 5.5; His sub 10 was $P < 0.0001$ between pH 7.4 and pH 5.5; His sub 13 trended towards significant with $P = 0.0923$ between pH 7.4 and pH 5.5; and His sub 15 was $P < 0.0001$ between pH 7.4 and pH 5.5.

Supplementary Fig. 3C: Flow cytometry assay. All variants' full length EGFR staining were compared to that of wild type EGF (EGF WT) by Kruskal-Wallis test with Dunn's correction. For full length EGFR staining, EGFd1 was $P < 0.0001$ vs EGF WT; EGFd1.1 was $P < 0.0001$ vs EGF WT; EGFd1.2 was $P = 0.0063$ vs EGF WT; EGFd1.3 was $P = 0.0467$ vs EGF WT; EGFd1.4 was not significant vs EGF WT ($P > 0.9999$); and EGFd1.5 was $P = 0.0001$ vs EGF WT. All variants' EGFRvIII staining were separately compared to that of EGF WT by Kruskal-Wallis test with Dunn's correction. For EGFRvIII staining, all variants were $P < 0.0001$ vs EGF WT. Finally, all variants' full length EGFR vs EGFRvIII staining were also compared pairwise by Kolmogorov-Smirnov test. All variants' full length EGFR vs EGFRvIII comparisons were $P < 0.0001$.

Supplementary Fig. 3D: Flow cytometry assay. pH 7.4 and pH 5.5 samples were compared to one another by two-tailed Kolmogorov-Smirnov test. This comparison was $P < 0.0001$.

Supplementary Fig. 3E: Flow cytometry assay. Samples from all 36 variants of EGFd1.5 were compared to EGFd1.5 via Kruskal-Wallis test with Dunn's correction. All but variants 8, 9, 12, 28, 29, and 31 were $P < 0.0001$ vs EGFd1.5. Variant 8 was $P = 0.0002$ vs EGFd1.5; variant 9 was $P = 0.0021$ vs EGFd1.5; variant 12 was $P = 0.0277$ vs EGFd1.5; variant 28 was not significant vs EGFd1.5 ($P = 0.1540$); variant 29 was not significant vs EGFd1.5 ($P > 0.9999$); and variant 31 was $P = 0.0010$ vs EGFd1.5.

Supplementary Fig. 5A: Flow cytometry assay. All A549 treated samples were compared vs Untreated by Kruskal-Wallis test with Dunn's correction. All comparisons were $P < 0.0001$. Additionally, all Withdrawal samples were compared vs CT-1212-1 24 hr by Kruskal-Wallis test with Dunn's correction. All comparisons were $P < 0.0001$.

Supplementary Fig. 5B: Flow cytometry assay. All H1975 treated samples were compared vs Untreated by Kruskal-Wallis test with Dunn's correction. All comparisons were $P < 0.0001$. Additionally, all Withdrawal samples were compared vs CT-1212-1 24 hr by Kruskal-Wallis test with Dunn's correction. 24 hr vs 1 day withdrawal was not significant ($P > 0.9999$), all other comparisons were $P < 0.0001$.

Supplementary Fig. 5C: Flow cytometry assay. All A549 treated samples were compared vs Untreated by Kruskal-Wallis test with Dunn's correction. All comparisons were $P < 0.0001$. Additionally, all Withdrawal samples were compared vs CT-1212-1 24 hr by Kruskal-Wallis test with Dunn's correction. All comparisons were $P < 0.0001$.

Supplementary Fig. 5D: Flow cytometry assay. All H1975 treated samples were compared vs Untreated by Kruskal-Wallis test with Dunn's correction. All comparisons were $P < 0.0001$. Additionally, all Withdrawal samples were compared vs CT-1212-1 24 hr by Kruskal-Wallis test with Dunn's correction. All comparisons were $P < 0.0001$.

Supplementary Fig. 6: Flow cytometry assay. Within all four cell lines, all samples were compared to one another by Kruskal-Wallis test with Dunn's correction. In A549, Untreated vs 3 day was not significant ($P = 0.4123$); 1 hr vs withdrawal was $P = 0.0373$; 1 day vs 2 day was $P = 0.0072$; and all other comparisons were $P < 0.0001$. In H1975, Untreated vs 1 hr was not significant ($P > 0.9999$); Untreated vs 3 day was $P = 0.0009$; 1 hr vs 3 day was not significant ($P = 0.0874$); and all other comparisons were $P < 0.0001$. In H358, Untreated vs withdrawal was not significant ($P > 0.9999$); 1 day vs 2 day was not significant ($P > 0.9999$); and all other comparisons were $P < 0.0001$. In H1650, all comparisons were $P < 0.0001$.

Supplementary Fig. 7B. Flow cytometry assay. Each treatment was compared to PBS by Kruskal-Wallis test with Dunn's correction. CT-1212-1 vs PBS was $P < 0.0001$; CT-1232-1 vs PBS was $P < 0.0001$; and PBS vs CT-3212-1 was $P = 0.0150$.

Supplementary Fig. 7C. Flow cytometry assay. Each treatment was compared to PBS by Kruskal-Wallis test with Dunn's correction. CT-1212-1 vs PBS was $P < 0.0001$; CT-1232-1 vs PBS was not significant ($P = 0.5149$); and PBS vs CT-3212-1 was $P < 0.0001$.

Supplementary Fig. 8: All samples were compared to one another by Kruskal-Wallis test with Dunn's correction. All comparisons were $P < 0.0001$.

Supplementary Fig. 9B: CellTiter-Glo 2.0 viability assay. The CT-1212-1 curve fit (asymmetric sigmoidal, 5PL) was compared pairwise to those of CT-1211-1 and CT-1222-1 with extra sum-of-squares F test ("Do the best-fit EC50 values differ between the selected comparator compound and CT-1212-1?"). The comparison was $P < 0.0001$ for CT-1211-1 and not significant ($P = 0.134$) for CT-1222-1.

Supplementary Fig. 9C: CellTiter-Glo 2.0 viability assay. At each FAC dose, No CYpHER and 20 nM CT-1212-1 were compared pairwise by two-tailed Welch's T test. (K-S test did not perform well due to low N per sample.) No FAC was $P = 0.0007$; 4 nM FAC was $P = 0.0002$; 12 nM FAC was $P = 0.0119$; 40 nM FAC was $P = 0.0088$; 120 nM FAC was $P < 0.0001$; 400 nM FAC was $P = 0.0055$; 1.2 µM FAC was $P = 0.0466$; 4 µM, 12 µM, and 40 µM FAC were not significant.

Supplementary Fig. 9E: CellTiter-Glo 2.0 viability assay. Samples not treated with EGF were compared all vs all by Kruskal-Wallis test with Dunn's correction. No comparisons were significant. Separately, samples treated with 1 ng/mL EGF were compared all vs all by Kruskal-Wallis test with Dunn's correction. No CYpHER vs CT-1212-1 was $P = 0.0168$; no CYpHER vs CT-3212-1 was not significant; and CT-1212-1 vs CT-3212-1 was $P = 0.0014$. Finally, within the three CYpHER treatments (no CYpHER, CT-1212-1, and CT-3212-1), no EGF and 1 ng/mL EGF were compared pairwise by two-tailed Kolmogorov-Smirnov test. No CYpHER was $P = 0.0005$; CT-1212-1 was $P = 0.0023$; and CT-3212-1 was $P < 0.0001$.

Supplementary Fig. 9F: CellTiter-Glo 2.0 viability assay. Cetuximab was not effective at suppressing SW48 cell growth, so no comparison for statistical significance was made. N = 3 wells per condition.

Supplementary Fig. 9G: CellTiter-Glo 2.0 viability assay. Samples were compared all vs all by Kruskal-Wallis test with Dunn's correction. No CYpHER vs CT-1212-1 was $P < 0.0001$; no CYpHER vs CT-3212-1 was not significant; and CT-1212-1 vs CT-3212-1 was $P = 0.0058$.

Supplementary Fig. 10B (left): Flow cytometry assay. All surface EGFR samples were compared to one another by Kruskal-Wallis test with Dunn's correction. 1 day vs withdrawal was not significant; all other comparisons were $P < 0.0001$.

Supplementary Fig. 10B (right): Flow cytometry assay. All surface TfR samples were compared to one another by Kruskal-Wallis test with Dunn's correction. 1 hr vs withdrawal was not significant; 1 day vs 2 day was not significant; 1 day vs 3 day was not significant; 2 day vs 3 day was not significant; all other comparisons were $P < 0.0001$.

Supplementary Fig. 10C: CellTiter-Glo 2.0 viability assay. The CT-1212-1 curve fit (asymmetric sigmoidal, 5PL) was compared pairwise to the comparator compounds with extra sum-of-squares F test ("Do the best-fit EC50 values differ between the selected comparator compound and CT-1212-1?"). CT-1212-1 vs cetuximab was $P = 0.0446$; CT-1212-1 vs gefitinib was $P < 0.0001$; and CT-1212-1 vs osimertinib was $P < 0.0001$.

Supplementary Fig. 10D. Ratio of EC50s from CellTiter-Glo 2.0 viability assays. Due to the complexity of error propagation and subsequent significance calculations for ratios of best-fit EC50 values, significance was not calculated for drug-vs-drug comparisons. This data is presented as-is, without claims of performance differences.

Supplementary Fig. 11A: Complete blood count statistics from CT-1222-1-treated, tumor-bearing mice. Within each blood cell type, counts were compared between all samples (all vs all) via Kruskal-Wallis test with Dunn's correction. No comparisons were found to be significant in any cell type.

Supplementary Fig. 11B: Mouse masses from CT-1222-treated, tumor bearing mice. Grouped samples were tested all-vs-all by a two-way mixed-effects model with Tukey correction. Within each time point, there were no significant pairwise comparisons. The smallest P value was at Day 5, 50 μg vs 150 μg (P = 0.2286).